

コマンド空間内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

眞嶋 宗一郎[†] 顔 世筈[†] Pathak Sarthak[‡] 梅田 和昇[‡]

[†] 中央大学大学院理工学研究科精密工学専攻 〒112-8551 東京都文京区春日 1-13-27

[‡] 中央大学理工学部精密機械工学科 〒112-8551 東京都文京区春日 1-13-27

E-mail: [†] {majima, karasu}@sensor.mech.chuo-u.ac.jp, [‡] {pathak, umeda}@mech.chuo-u.ac.jp

あらまし 本研究では、ジェスチャによる家電の簡単な操作を目的とし、複数ジェスチャ認識とコマンド空間を組み合わせることで、多数の家電操作コマンドを簡単に操作できる新規システムを提案している。操作者が部屋内任意の位置で2回手振りを行うことで擬似相対座標系を設定し、その座標系に設置して家電操作を関連付けたコマンド空間内でジェスチャを行うことで家電の操作を行う。実験により提案手法の検証を行っている。

キーワード ジェスチャ認識, ヒューマンインタフェース, 家電操作, コマンド空間

1. 序論

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。そこで近年、手振りなどの直感的で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。

入江らは、複数種類のジェスチャそれぞれに家電操作のコマンドを設定し、ジェスチャによる家電の操作を可能とする、室内に複数のカメラを設置したインテリジェントルームを構築した[1]。操作可能なコマンド数はジェスチャの種類により制限されるため、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った。これに対し、顔らは家電操作のコマンドを特定の空間に設定し（以降、コマンド空間と呼ぶ）、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[2]。しかしながら、各コマンド空間に設定されるコマンドは1つのため、1つの階層での操作可能なコマンド数は空間の数により制限されてしまう。

そこで本研究では、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を簡単に行えるシステムの構築を目的とする。

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。図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、操作者が任意の位置で2回手振りを行い、FFTを用いた手振り検出手法[1]を用いて手振りの三次元位置を求めて擬似相対座標系を定義する。次に、擬似相対座標系においてコマンド空間を定義する。操作者はコマンド空間内でジェスチャを行い、コマンド



図1 提案システムの概念図

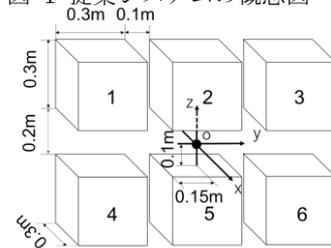


図2 コマンド空間の設定

空間に設定されているコマンドを実行して家電を操作する。

3. 擬似相対座標系によるコマンド空間の構築

擬似相対座標系とは、ユーザが任意な場所で定義する座標系である[2]。定義後に操作者が移動しても、定義された位置に維持される。擬似相対座標系においてコマンド空間を構築することにより、操作者は任意の位置で家電操作を行うことが可能となる。図2に設定されるコマンド空間を示す。なお、本家電操作システムでは、コマンド空間は一部屋に一つだけとしている。他の場所でコマンド空間を構築したいときは、任意の位置で再び手振りを行うことで擬似相対座標系ごと再構築する。

4. ジェスチャの検出とコマンド操作

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図3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、OpenPose[3]により抽出される人の骨格点情報を利用する。インテリジェントルーム内の4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、候補の中でエポポー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。ステレオ計測の原理により、手首の3次元座標(x_1, y_1, z_1)とひじの3次元座標(x_2, y_2, z_2)を求める。手首の3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。

5. 認識率の評価実験

2種類のジェスチャの6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被験者1人を対象とし、図2のコマンド空間を図4に示すように机の上に固定した。操作者は同じ位置から各コマンド空間内で2種類のジェスチャを50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。

2つのジェスチャの認識率を表1, 表2に示す。位置の認識率は、ジェスチャ1が96%, ジェスチャ2が98%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90%, ジェスチャ2が89%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。2種類のジェスチャだけを使い12個のコマンド操作をほぼ正確に行えることが確認できた。また、ジェスチャ1でのコマンド3とジェスチャ2でのコマンド2では他と比較して認識率が低く、位置認識率と認識率の差も大きくなった。コマンド2からコマンド3の順番でジェスチャ1を行ったため、コマンド3では肩が開き、腕が他と比べて手前に位置した。それに加え、ジェスチャ1はジェスチャ2に比べて手首が手前に位置するので、手首がコマンド空間の外に出てしまうことがある。また、手首を奥に位置させようとすると、手首とひじの高

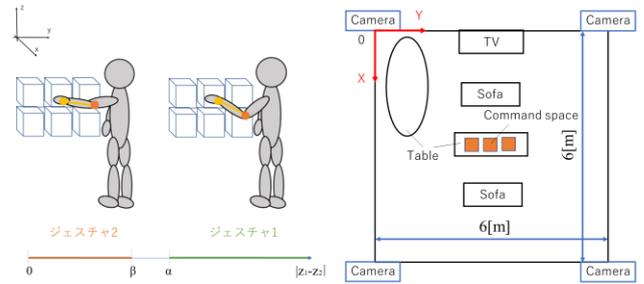


図3 2種類のジェスチャ 図4 実験環境

表1 ジェスチャ1の認識率 [%]

	1	2	3	4	5	6	平均
位置認識率	98	100	88	92	98	100	96
認識率	94	94	70	90	96	96	90

表2 ジェスチャ2の認識率 [%]

	1	2	3	4	5	6	平均
位置認識率	98	92	96	100	100	100	98
認識率	86	76	94	94	92	90	89

さの差が小さくなる。そのため、ジェスチャ1でのコマンド3では位置の認識率と認識率が下がったと考えられる。ジェスチャ2でのコマンド2は、ジェスチャ時に肩が上がる不自然な姿勢になるためジェスチャを正しく行いづらく、位置の認識率に比べて認識率が大きく下がったと考えられる。

6. 結論と今後の展望

本研究では、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。実験では平均約90%の認識率が得られた。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は正しいジェスチャを行いつらく、認識率が低くなった。今後は認識率の向上のため、どのコマンド空間内でも行いやすいジェスチャや、より使いやすいコマンド空間の構築を目指す。また、本システムでは1部屋で操作に利用できるコマンド空間は1つで、同時に操作できるのは1人だけという制限がある。そこで、室内にいる複数の人がそれぞれ1つずつ専用のコマンド空間を使用できるシステムの構築も目指す。

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] 顔世荀, 池勇勳, 梅田和昇: “簡単なジェスチャでコマンド空間を設定し家電操作可能なシステム”, 日本機械学会ロボティクス・メカトロニクス講演会講演論文集 2020, 2A2-F01, 2020.
- [3] Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh: “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” CVPR, 2017.