

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

A System for Home Appliance Operation in User-definable Command Space by Simple Gestures

精密工学専攻 14号 顔世荀
Shixun Yan

1. 序論

日常生活において欠かすことのできない家電製品は、多機能化、高性能化している。しかし、その一方で操作の複雑化という問題も生じている。近年、家電製品の様に人に身近な製品を人のジェスチャを用いて直感的に操作することを目的とした研究が多くなされている[1][2]。

入江らは、高齢者がリモコンなどの小さいボタンを操作するのが苦手であるという問題を解決するため、室内に複数台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[3]。このシステムは、特殊な機器を用いず、室内のどこからでもジェスチャで機器の操作が可能である。しかし、このシステムを操作するために、複数種類の複雑なジェスチャをすべて記憶せざるを得ない。浅野らは、家電操作のコマンドに関連づけた仮想の空間（以降、コマンド空間と呼ぶ）を室内に配置し、その空間で手振りを行うことで任意の家電操作を行うことができるシステムを提案した[4]。しかし、このシステムではコマンド空間が固定されているため、操作位置が限定されてしまう。そこで、今村らは、ユーザを基準とした相対座標系を設定し、ユーザの周辺にコマンド空間を設置することで、室内のどこでも操作を可能にした[5]。さらに、鹿野らは、人の骨格情報を取得可能な RGB-D センサ 1 台を用いて相対座標系を設定する家電操作システムを構築した[6]。しかし、鹿野らのシステムでは、ユーザがカメラの正面を向いて手振りを遂行しない限り、骨格情報抽出の精度が落ちるため自由な家電操作が困難である。

これらの問題を踏まえ、擬似相対座標系という概念を導入する。複数のカメラ画像から人の手振りを検出し、その位置で擬似相対座標系を展開することで、部屋中任意の場所での

家電操作が可能である[7]。しかし、画像処理を使用し手振りを検出するため、時間がかかるという問題がある。さらに、家電製品を操作する際に何度も手を振る必要があるため、ユーザが疲労する可能性がある。これに対し、視線情報を使用し家電製品を操作する方法がある。しかし、部屋の位置に関わらず安定に視線を検出することは難しく、各操作に必要な時間も評価されていない[8]。この問題を解決するため、本研究では、ユーザに負担をかけずに、簡単なジェスチャで家電操作できる新しいシステムを提案する。簡単なジェスチャとは手かざしを用いる。手かざしとは、手を1箇所に置いて短時間で静止するジェスチャである。

2. 提案システムの概要

本研究で扱うシステムは、室内の家電機器を人のジェスチャのみで操作可能にしたシステムであり、リビングやオフィスなどでの使用を想定している。提案システムの概念図を Fig. 1 に示す。提案するシステムにおける操作手順を次に示す。まず、ユーザが異なる位置で 2 回手振りジェスチャを実行し、擬似相対座標系を定義する。コマンド空間は、擬似相対座標系に基づいて構築される。次に、ユーザはコマンド空間で手かざしにより、家電を操作する。

3. 手振り検出手法

インテリジェントルーム内のカメラから取得した画像を低解像度化し、画素値の時系列変化に対して高速フーリエ変換 (FFT: fast Fourier transform) を行うことで、周期的な動作である手振りの検出を行う。手振り領域に対応する画素では、Fig. 2 に示すように濃淡値が変化する。この濃淡値の変化は、周期的な変化であるため、FFT を行うことで周波数と振幅の算出が可能である。ある周波数の振幅が一定の閾値を超えた時手振りとして検出する。以上の手法を複数のカメラで行い、エピポーラ拘束を満たす画素を手振り画素として特定する。その後、ステレオ計測の原理により、手振りの三次元座標を求める。

4. 擬似相対座標の設置

4.1. 擬似相対座標系

擬似相対座標系とは、ユーザが任意の場所で定義する座標系である。定義後、ユーザが移動しても、定義された位置に

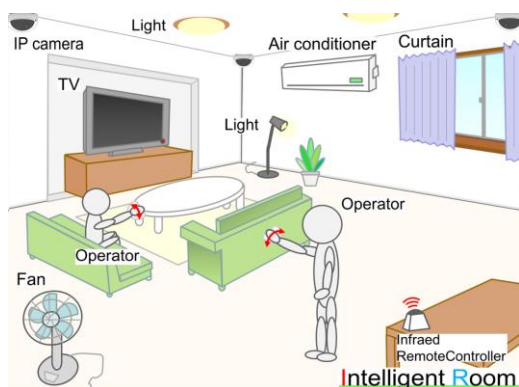


Fig. 1 Conceptual diagram of an intelligent room.

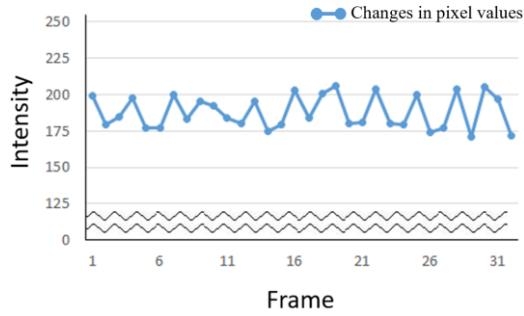


Fig. 2 Changes in pixel values corresponding to hand waving.

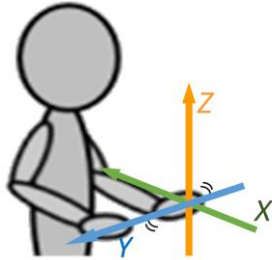


Fig. 3 Pseudo-relative coordinate system based on the position of hand waving.

固定したままとする．人の手振りを検出した三次元座標から擬似相対座標系を展開し，その擬似相対座標を基準にしてコマンド空間を設置することで，部屋中任意の場所での家電操作が可能になる．ユーザは家電操作をする前に，複数回手振りを遂行して，そこから得られた三次元座標で擬似相対座標系を設定する．また，手振りは左右どちらの手でも実行できる．本論文では，一回目の手振りの三次元座標を擬似相対座標系の原点とコマンド空間の中心とする．また，一回目と二回目の手振りの三次元座標同士を結んだ軸を 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．Fig. 3 に概略図を示す．

4.2. 擬似相対座標の計算

一回目の手振りの三次元座標を $O(x_1, y_1, z_1)$ ，二回目の座標を $Q(x_2, y_2, z_2)$ とする．ここで，コマンド空間を水平に展開するため， z_2 の値を z_1 と等しくする．絶対座標 (x, y, z) と擬似相対座標 (x', y', z') の関係を Fig. 4 と次式に示す．

$$\begin{pmatrix} x' \\ y' \\ z' \end{pmatrix} = \mathbf{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} + \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} \quad (1)$$

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

\mathbf{R} は絶対座標系と擬似相対座標系間の回転行列である．

5. コマンド空間の設置

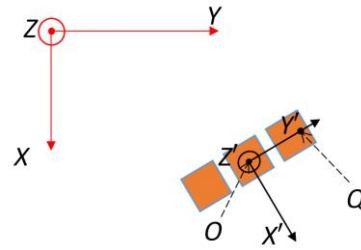


Fig. 4 Relationship between absolute coordinate system and pseudo-relative coordinate system. Orange rectangles represent the command spaces.

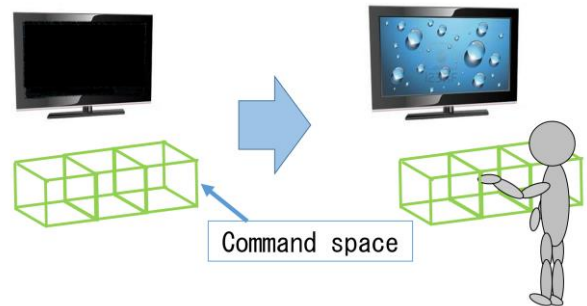


Fig. 5 Conceptual image of the command space.

5.1. コマンド空間

コマンド空間とは，家電操作と関連付けられた仮想空間である．空間に情報を埋め込んだ空間メモリ[9]の考え方をうけて，家電操作における特定のコマンドをコマンド空間に関連付ける[4][5]．このコマンド空間で手かざしを行うことで，関連付けられた家電の特定のコマンドを実行可能である．本研究におけるコマンド空間のイメージ図を Fig. 5 に示す．Fig. 5 のイメージ図では，真ん中にあるコマンド空間が TV の ON/OFF コマンドに関連付けられており，そのコマンド空間で手かざしを行うことで，TV の ON/OFF の操作が可能である．

5.2. コマンド空間の配置

コマンド空間を，Fig. 6 に示すように，横に 3 個上下に 2 段，計 6 個配置する．一回目の手振りによるコマンド空間の原点 O は Fig. 7 (正面図) に示すように全空間の中心位置である．また，Fig. 7 の点 Q は，二回目の手振りによる位置を示す．

展開されたコマンド空間の相対位置をユーザが直観的に分かりやすいように提示する方法として，Fig. 8 に示すような画面をディスプレイやプロジェクタで表示する．ここで，黄色の長方形は操作対象の家電を示す．

6. 手かざし検出

コマンド空間を構築した後，家電製品の操作は，手振り

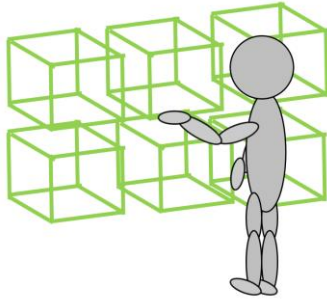


Fig. 6 Arrangement of the command spaces.

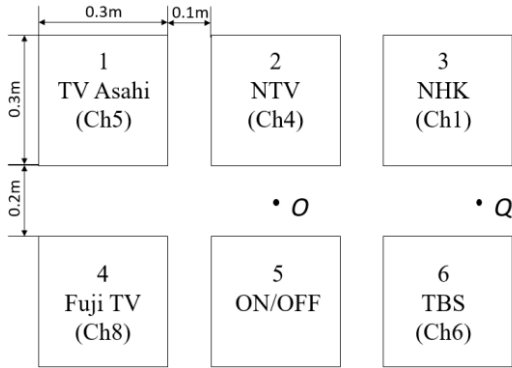


Fig. 7 Front view of the command spaces.

はなく、手かざしの位置に基づいて行う。低解像度化された画像に対し、背景差分とフレーム間差分を用いて、手かざしの検出を行う。手かざし検出処理の流れを Fig. 9 に示す。まず、フレーム $k-1$ の画像 (Fig. 9(a)) とフレーム k の画像 (Fig. 9(b)) を差分する、Fig. 9(e) に示した白いピクセルはフレーム間差分の結果である。次に、フレーム k の画像とフレーム $k+1$ の画像 (Fig. 9(c)) で差分を取る。手かざしが行われた時には、Fig. 9(f) に示したように差分領域が生じない。その後、フレーム $k+1$ の画像に対し背景差分を実行する (Fig. 9(g))。以上の手順を用いて、Fig. 9(e)、Fig. 9(f)、Fig. 9(g) の結果が白、黒、白の順で変化した領域を、手かざしが発生した領域として検出する。最後に、ステレオ計測の原理により、手かざしの三次元座標を求める。家電製品は手かざしの位置に配置されているコマンド空間に応じて実行される。

7. 実験

7.1. 実験環境

提案する家電操作システム、Fig. 10 に示すような部屋に構築した。四隅に設置したカメラによりユーザの手振りを認識し、コマンド空間を展開した。使用したカメラの画角は 55.8 [deg] \times 43.3 [deg]、画像の解像度は 640×480 、低解像度した画像の解像度は 80×60 であった。

7.2. 操作精度と操作時間の検証

Fig. 10 に示すように、コマンド空間を机の上に展開し、精

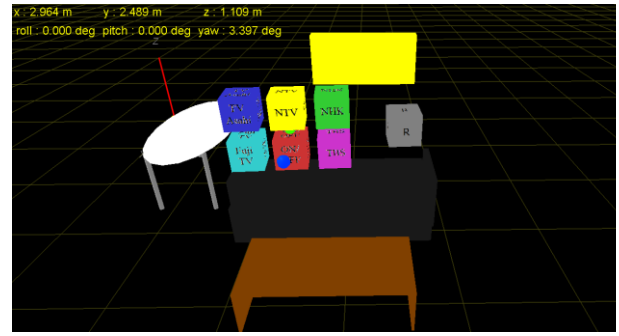


Fig. 8 Command space display.

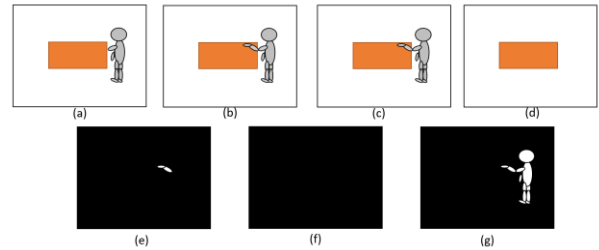


Fig. 9 Illustration of background subtraction and frame subtraction: (a) frame $k-1$, (b) frame k , (c) frame $k+1$, (d) background image, (e) frame subtraction between (a) and (b), (f) frame subtraction between (b) and (c), and (g) background subtraction result on (c).

度検証実験を行った。コマンド空間の展開と手かざしの精度を検証するため、手振りでコマンド空間を構築し手かざしによる家電操作を行う一連の動作を 20 回行い、その認識率を評価した。また、手かざしと手振りでの家電操作するための時間の比較実験を行った。展開されたコマンド空間でそれぞれのジェスチャで 6 つのコマンド操作を 10 回行い、操作時間を評価した。被験者は本手法の操作経験のある人物 1 名とした。実験の様子を Fig. 11 に示す。

提案手法のコマンド空間の認識率を Table 1 に示す。平均認識率は 88 % であり、全てのコマンド空間で認識率 85 % 以上の結果を得られた。しかし、手かざしの仕方によって認識率が低下することが実験中にあったので安定させる必要がある。また、操作時間の比較結果を Table 2 に示す。手かざしの平均操作時間は 21.21 [s]、手振りの平均操作時間は 23.92 [s] である。手かざし操作を手振り操作より 11 % 以上速くすることができた。従来の手振り操作に代わり、さらに簡単なジェスチャで速く検出することができたが、一回の手かざしにはまだ 3~4 [s] が必要なため、更なる高速化を考える必要がある。

8. 結論

本研究では、擬似相対座標系と手かざしを導入し、コマンド空間の展開位置に対する制限が少ない、簡単なジェスチャで操作できる家電操作システムを構築した。操作精度と時間の評価実験を行った。

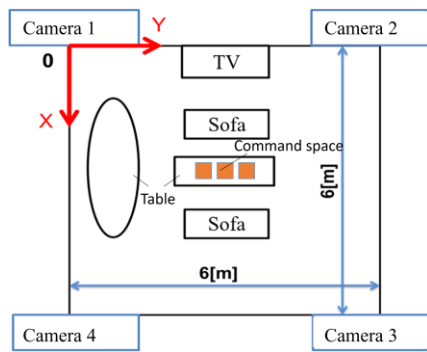


Fig. 10 Experimental environments of the home appliance operation system.



Fig. 11 Experimental situation.

今回の実験では、8割以上の認識率を得られた。
 今後は、認識率の向上と高速化を目標とする。

参考文献

- [1] 森武俊, 佐藤知正, “ロボティックルーム1・2・3の開発,” 電子情報通信学会誌, Vol. 91, No. 5, pp. 402-410, 2008.
- [2] C. Hsieh, D. Liou and D. Lee, “A Real Time Hand Gesture Recognition System using Motion History Image,” 2010 2nd International Conference on Signal Processing Systems, pp. 394-398, July, 2010.
- [3] 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.
- [4] 浅野秀胤, 永易武, 織茂達也, 寺林賢司, 太田陸, 梅田和昇, “フーリエ変換を用いた指振り検出と機器操作への応用,” 精密工学会誌, Vol.79-6, pp.565-570, 2013.
- [5] 今村勇也, 永易武, 浅野秀胤, 寺林賢司, 梅田和昇, “コマンド空間を用いた家電操作システムの改良と実験による認識率の検証,” 第31回日本ロボット学会学術講演会予

Table 1 Recognition rate for each position

	Recognition rate [%]
TV Asahi	90
NTV	85
NHK	95
Fuji TV	85
ON/OFF	90
TBS	85
Average	88

Table 2 Time required for operation

Times	Positioning [s]	Waving [s]
1	20.53	24.20
2	20.61	25.33
3	20.76	21.4
4	21.38	21.03
5	23.47	25.73
6	22.61	24.51
7	19.53	23.53
8	21.07	23.15
9	18.80	26.13
10	23.31	24.15
Average	21.21	23.92

稿集, 2D2-03, 2013.

- [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'18講演論文集, 2A2-E08, 2018.
- [7] S. Yan, Y. Ji and K. Umeda, “A System for Home Appliance Operation by Hand Waving in a User-definable Command Space,” Fourteenth International Conference on Quality Control by Artificial Vision, vol. 11172, no. 14, July 16, 2019.
- [8] H. Nakayama, N. Yabuki, H. Inoue, Y. Sumi and T. Tsukutani, “A control system for electrical appliances using eye-gaze input,” 2012 International Symposium on Intelligent Signal Processing and Communications Systems, pp. 410-413, Nov., 2012.
- [9] M. Niitsuma, H. Kobayashi, and A. Shiraishi, “Enhancement of Spatial Memory for Applying to Sequential Activity,” Journal of Advanced Sciences, Vol.9, No. 1, pp. 121-137, 2012.