

Kinect 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

A simple operation system of home appliances by hand waving using Kinect

精密工学専攻 18号 鹿野巧
Takumi Kano

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が、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 [1] が多くなされている。

入江ら [2] は、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ている。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なシステムであるが、手振りから始まるジェスチャを数種類行う必要がある。また、浅野ら [3] は、家電操作のコマンドを関連づけた空間（以後、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出来るシステムを提案している。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 [4] は、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することで、室内のどこでも操作が可能としている。

これらのシステムでは、複数台のカメラを用いているため、カメラの設置及びキャリブレーションに手間がかかるという問題がある。また、操作を行う際には、操作者は止まって、手振りを行う必要がある。

本論文では、今村らと同様の家電操作システムを Microsoft 社の Kinect v2 を用いて構築する。Kinect は、カラー画像と距離画像、骨格情報を取得できるため、1 台でシステムを構築可能であり、設置が容易である。また、既存の手振り検出手法を応用して移動人物の手振りを検出することで、移動しながら家電を操作することを可能とする。

2. インテリジェントルームの概要

本研究で扱うインテリジェントルームとは、室内の家電機器をジェスチャのみで操作する智能化された部屋のことであり、リビングやオフィスなどでの使用を想定している。

本研究では、Kinect から得られるカラー画像と距離画像を用いて、手振りの検出、及びその三次元座標の取得を行う。取得した三次元座標に応じて、あらかじめ設定した家電操作を行う。インテリジェントルームの概念図を Fig. 1 に示す。

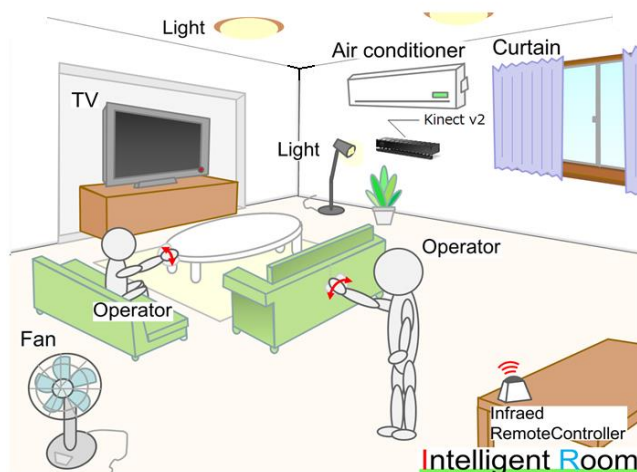


Fig.1 The conceptual diagram of the Intelligent Room

3. 相対座標系

3.1 相対座標系の設定

Kinect から得られる骨格情報を用いて、操作者を基準とした相対座標系を設定する。Kinect から得られる骨格情報の内、両肩の座標を結んだ軸を第一軸、首、背骨中央を結んだ軸を第二軸とし、第一軸と第二軸の外積を第三軸とする。Fig. 2 に概略図を示す。

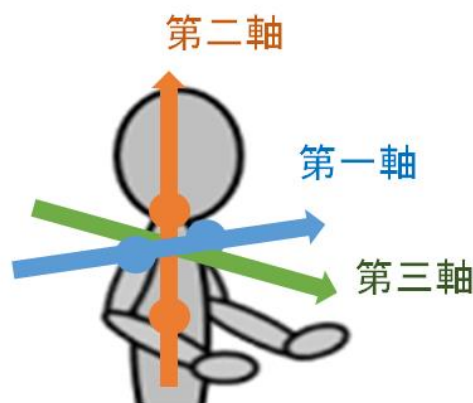


Fig.2 Relative coordinate based on the operator

3.2 相対座標系の補正

Kinect から得られる骨格情報は、人物が移動しているときや、取得する骨格点にもものが重なった場合などに、誤差が大きくなる。そこで、各骨格点の前フレームからの変位量が閾値以上だった場合に、前々フレームから、前フレームでの変位量をもとに補正を行うことで、誤差を小さくする。

4. 手振り検出手法

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FFTを行うことで、周波期的な動作である手振りの検出を行う。

4.1 前処理

Kinect から取得したカラー画像をグレースケール化し、低解像度化する。画像を低解像度化することで、計算量の低減やノイズの抑制といった効果がある。また、手振りによる濃淡値に変化が滑らかになり、より正弦波に近い濃淡変化のパターンが得られるといった効果がある。



(a) Input image

(b) Low resolution image

Fig.3 Preprocessing

4.2 従来の周波数特徴量の抽出

低解像度化した画像の各画素に対して時間軸方向のFFTを行う。手振り領域に対応する画素では、Fig. 4 に示すように濃淡値が手の場合と背景の場合で変化する。この濃淡値の変化は、周期的な変化であるため、FFTを行うことで特徴量を抽出する。特徴量にはスペクトルの最大値と平均値を用いる。

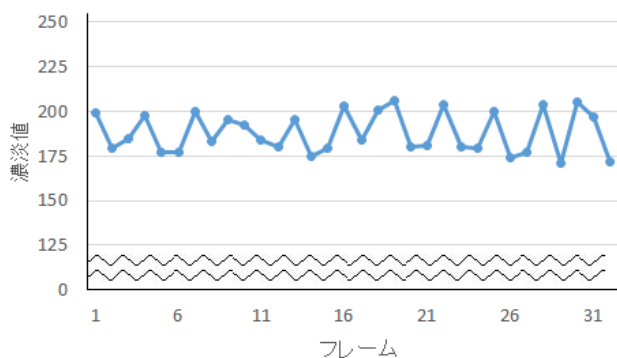
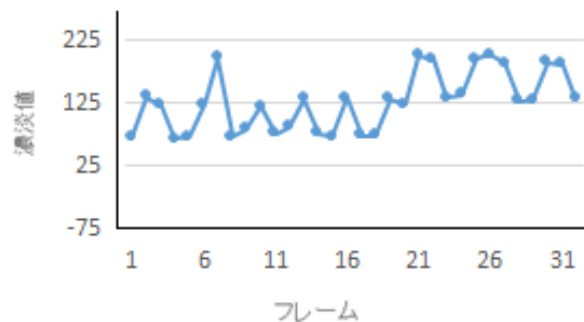


Fig.4 Changes in pixel values of pixels corresponding to hand waving

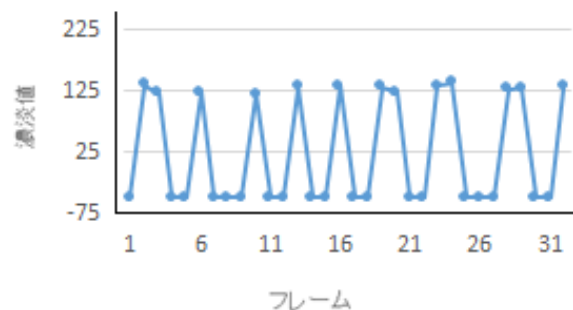
4.3 移動人物の手振り認識のための周波数特徴量の抽出

従来手法では、入力画像の1画素に注目し、その時系列変化に対してFFTを行うことで周波数特徴を抽出するため、操作者は止まって操作する必要があった。そこで、相対座標系の1点に注目し、その濃淡値の時系列変化に対しFFTを行うことで、移動人物の手振り動作の周波数特徴量を抽出する。また、移動人物の手振り認識に際しては、背景の濃淡値が一定とならないことがあるため、Fig. 5(a)に示すように、濃淡値の時系列変化が、別々の正弦波をつなぎ合わせたような変

化になることがある。そこで、Kinect から得られる手のひらの距離値と閾値以上離れた距離値を持つ画素の画素を一定値に補正することで、Fig. 5(b)の様に、より単一の正弦波に近い濃淡変化のパターンを得る。



(a) Before correction



(b) After correction

Fig.5 Changes in pixel values in moving human hand waving

4.4 手振りの検出

FFTにより求めた特徴量から、手振りのクラスと手振り以外のクラスを識別する。識別には、サポートベクターマシン (Support Vector Machine, SVM) [5]を用いる。

5. 手振りによる機器操作手法

家電操作を空間に関連付けることで、手振りのみを用いた家電操作を実現する。

5.1 コマンド空間による機器操作

空間に情報を埋め込む新妻らの空間メモリ [6]の考え方をを用いて、家電操作のコマンドを空間に関連付ける。この家電操作を関連付けられた空間 (コマンド空間) で手振り動作を行うことで、関連付けられた家電操作を行うことができる。本研究におけるコマンド空間のイメージ図を Fig. 6 に示す。Fig. 6 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に関連付けられており、そのコマンド空間で手振り動作を行うことで、関連付けられていたTVのON/OFFの操作を行っている。

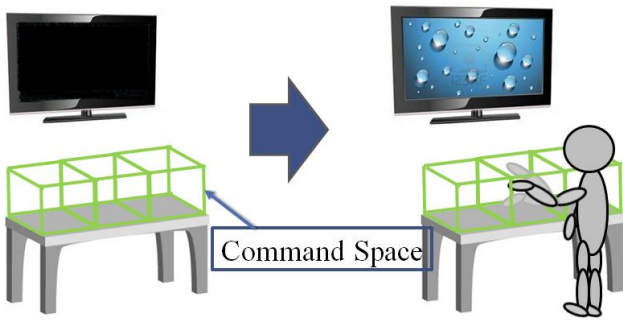


Fig.6 Image diagram of command space

5.2 コマンド空間内の手振り検出

設定したコマンド空間内に、複数点を設定し、その各点に対応する画素の画素値を各フレームにおいて取得し、濃淡値の時系列データとして、前述の手振り検出手法を用いて、手振り検出を行い、対応した操作を行う。

6. 実験

構築したシステムの評価を行うために、実験を行った。

6.1 実験条件

実験はFig.7に示す環境で行った。Kinectを高さ1400[mm]地点に仰角0[deg]で設置し、立位(立った姿勢)、座位(座った姿勢)、仰臥位(仰向けに寝た姿勢)の3つの姿勢で実験を行った(Fig.8参照)。



Fig.7 Experiment environment



(a) Standing position

(b) Sitting position



(c) Supine position

Fig.8 Experimental posture

6.2 コマンド空間の配置

コマンド空間は、Fig.9に示すように、操作者の前方に、横に3個並べたものを上下2段、計6個配置した。6個のコマンド空間は、高さ方向は、上下段の境目が第一軸と同じ高さになるように、横方向はコマンド空間群の中心が、操作者の正面に来るように配置した。また、コマンド空間を操作者から50[mm]離れた位置に配置した

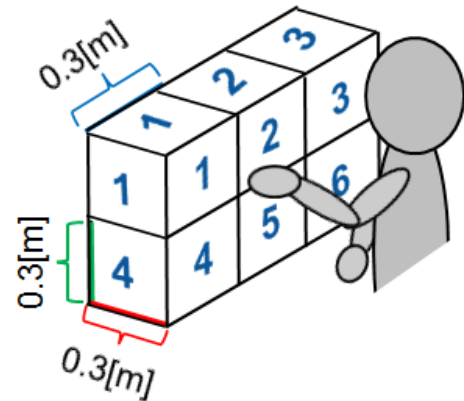


Fig.9 arranging the command space

6.3 座標系の設定精度検証のための機器実験

提案手法の検証実験を行った。各コマンド空間で20回ずつ手振りを行うことで認識率を調べた。被験者は本手法の操作経験のある人物1名とした。

結果をTable 1に提案手法と参考として今村らの手法(以下従来手法)でのコマンド空間の認識率を示す。認識率は6つのコマンド空間の平均認識率である。提案手法ではコマンド空間の平均認識率は立位で92[%]、座位で93[%]、仰臥位では32[%]となり、立位と座位では90[%]を上回ったが、仰臥位では50[%]を下回る結果となった。

仰臥位では、Kinectから取得した画像に写る人物の領域が小さくなり、骨格情報が正しく推定できないことがあり、認識率の低下につながった。

Table 1 Recognition rate for each posture [%]

	立位	座位	仰臥位
提案手法	92	93	32
従来手法	100	85	92

6.4 ジェスチャの認識率検証実験

次に、被験者を増やし、提案手法の検証実験を行った。Fig. 8に示すように操作者の前方に6個のコマンド空間を配置し、各コマンド空間で5回ずつの手振りを行い、認識率を調べた。被験者は20代男性5名で、実験に当たって大まかなコマンド空間の配置のみ説明を行った。

Table 2に被験者毎の認識率を、Table 3にコマンド空間毎の認識率を示す。平均認識率は、立位で77[%]、座位で75[%]と、ともに70[%]を上回る結果となった。コマンド空間毎の認識率は、下段のコマンド空間全てで認識率80[%]を上回っていたのに対し、上段のコマンド空間では、コマンド空間3を除く2つの空間で、認識率70[%]を下回っていた。これは、手振りをした際に腕の周期運動が検出されたためであると考えられる。とくに、コマンド空間2では立位、座位ともに、60[%]を下回る結果となった。これは、コマンド空間2で手振りをした際に、Kinectに腕と体が重なって写り、座標系の設定に悪影響を及ぼしたためと考えられる。

Table 2 Recognition rate for each subject [%]

	立位	座位
被験者 1	77	87
被験者 2	70	67
被験者 3	73	83
被験者 4	67	63
被験者 5	93	73
平均値	77	75

Table 3 Recognition rate for each command space [%]

	立位	座位
コマンド空間 1	68	68
コマンド空間 2	52	32
コマンド空間 3	88	92
コマンド空間 4	84	84
コマンド空間 5	84	96
コマンド空間 6	84	80
平均値	77	75

7. 結言

本研究では、今村らと同様の家電操作システムを、Kinectを用いることで、より構築が容易なシステムとした。また、既存の手振り検出手法を応用した移動人物の手振りを認識する手法を提案した。

今回の実験では、立位、座位では7割以上の認識率を得られたが、仰臥位や、腕と体が重なって移った際には、座標系設定制度が問題となり、認識率が低下した。

今後は、座標系設定の手法を見直し、システムの改善に取り組む。また、移動人物の手振り認識手法の評価実験を行う。

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, **91-5** (2008) pp. 402-410.
- (2) 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, **73-C-725** (2007) pp. 258-265.
- (3) 浅野秀胤, 織茂達也, 永易武, 寺林賢司, 太田陸, 梅田和昇, 小さな手振り検出を用いた家電操作システムの構築, 映像メディア学会年次大会講演予稿集, **2011-9-9** (2011).
- (4) 今村勇也, 永易武, 浅野秀胤, 増山岳人, 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稿集, **2D2-03** (2013).
- (5) Cortes, C. and Vapnik, V., Support-vector-networks, Machine Learning, **20-3** (1995) pp. 273-297.
- (6) Mihoko Niitsuma, Hiromu Kobayashi, Ayumu Shiraishi, Enhancement of Spatial Memory for Applying to Sequential Activity, Journal of Advanced Sciences, **9-1** (2012) pp. 121-137.