ジェスチャ認識を用いたインテリジェントルームの構築
- 手のジェスチャによる家電製品の操作 -

入江耕太(日立製作所)　○若村直弘(中央大学)　梅田和昇(中央大学)

Construction of an Intelligent Room Using Gesture Recognition
-Operation of Electric Appliances with Hand Gestures-

Kota IRIE,Hitachi Ltd., *Naohiro WAKAMURA, Kazunori UMEDA,Chuo University

Abstract － This paper proposes an intelligent room that is free of operator's position based on gesture recognition technologies. Intention and position of an operator is recognized by detecting the hand waving, and pan-tilt cameras are zoomed and focused on the operator. The hand region is extracted using color infomation, and direction or number of fingers and motion of the hand region are detected. Electric applinances such as a television set are controlled by using the direction or number of fingers and motion of the hands.

***Key Words*:** Intelligent Room, Man-Machine Interface, Gesture Recognition, Image Processing

1. 序論

現在，我々の生活環境の情報化，インテリジェント化が進んでいる．それに応じて，直感的でかつ非接触での操作を可能にするマン・マシン・インタフェース技術が必要になると考えられる．中でも，人間と機械が自然なコミュニケーションを実現するためには，ジェスチャを用いることが有効である．これまでに，動画像からジェスチャを認識する研究が数多く報告されている[1)2)]．また，それらのジェスチャ認識技術を用いて部屋全体を知能ロボット化したインテリジェントルームの研究も行われている[3)4)]．しかし，これらの多くは，ジェスチャを認識できる場所が限定されているため，生活空間での実用性に欠ける．

本研究では，指の本数や方向などのジェスチャ認識によるマン・マシン・インタフェースの実利用を目指し，操作位置の拘束がないインテリジェントルームの構築を行う．

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は，室内にカメラを取り付け，部屋全体を知能ロボット化した部屋であり，一般的なオフィスや家庭などへの適用を想定している．

本研究で提案するインテリジェントルームのシステムは，視覚となるCCDカメラを設置し，自律的に操作者の特定ならびにジェスチャの認識を行う．まず，2台のCCDカメラの画像から「手振りの検出」を行い，操作者の発見とともに3次元位置情報の取得を行う．その3次元位置情報によってカメラからの距離を算出し，カメラをズームさせ，ジェスチャ認識処理を行うための候補領域を絞る．続いて，手振り位置の色情報から手領域の抽出を行う．抽出された手領域の面積や形状，領域の個数などの特徴量から「指さし方向の認識」や「指の本数認識」を行い，登録されたジェスチャの中からジェスチャを識別する．操作対象となる家電製品は，赤外線リモコンで操作可能な機器を対象とし，本システムではテレビ，照明機器，扇風機とする．各機器への制御信号は，PCに接続された赤外線リモコンを用いて送信する．Fig.1にインテリジェントルームの概念図を示す．

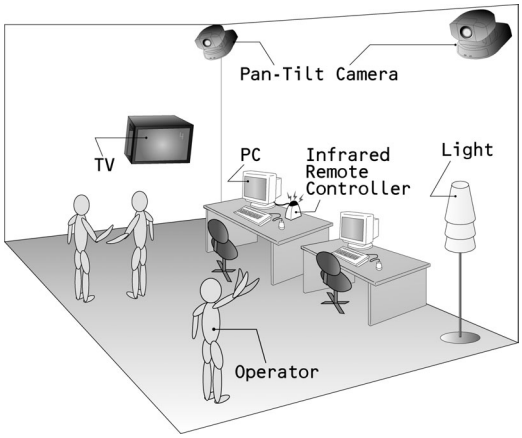

Fig.1　The conceptual figure of an intelligent room

3. 操作者特定のための手振り検出[5)]

インテリジェントルームにおいて，操作者を特定する手法として，手振り動作の検出を行う．低解像度化した濃淡画像の各画素に対して時間軸方向のFFTを行い，その中で手振り領域の検出を行う．本手法は，カラー情報を用いていないため，照明条件や肌色の個人差に対してロバストである．さらに，あらかじめ手領域を抽出するといった画像処理は不必要であり，極めて簡潔な処理である．

4. 肌色登録

肌色の個人差や環境による変化に対するロバスト性の向上を図るために，前処理として「肌色の登録」を行う．肌色登録における色空間は，照明条件などの明るさによる影響をうけにくい H（色相），S（彩度），I（明度）空間を用いる．

HSI 空間において，肌色の特徴を表現するのに適した HS を特徴空間とし，肌色のクラスタを形成する．認識時には，各画素で得られた H と S との値からマハラノビス距離を求め，設定した閾値以内の画素を手領域として抽出する．肌色のクラスタの平均ベクトルを $\mathbf{M}=[M_H\ M_S]^T$ とし，共分散行列を $\mathbf{V}$ とする．計測された特徴量ベクトル $\mathbf{X}$ とのマハラノビス距離 d_M は式(1)で与えられる．Fig.2 に特徴空間におけるマハラノビス距離の模式図を示す．

$$d_M{}^2 = (\mathbf{X}-\mathbf{M})^T \mathbf{V}^{-1}(\mathbf{X}-\mathbf{M}) \tag{1}$$

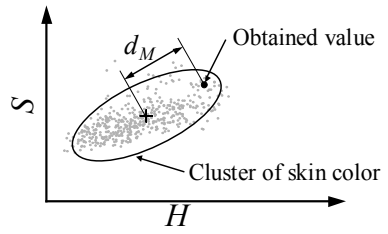

Fig.2　Distribution of HS value in feature space

5. ジェスチャ認識

4 章で登録された肌色情報に基づき，手の領域の抽出を行い，指の本数認識，指さし方向の認識，手の動作の認識を行う．

5.1 指の本数認識

指の本数認識手法を示す．認識は以下の手順で行う（Fig.3 参照）．

(1) 肌色情報に基づき，手の領域を抽出する．
(2) 手の領域に収縮処理を行う．
(3) (2)の領域に膨張処理を行う．
(4) (1)の領域から(2)の領域の差をとり，領域を分離し，残った領域から形状や面積の特徴から指領域のみを抽出する．

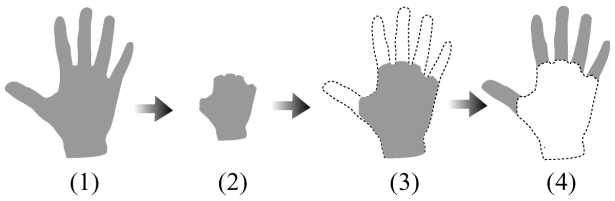

Fig.3　Recognition of number of fingers

5.2 指さし方向の認識

2 台のカメラからの画像に対し，5.1 節で述べた指の本数認識手法を用いて指領域の抽出を行い，抽出された領域の主軸ベクトルを求める．それぞれの主軸ベクトルを 3 次元空間に投影し，投影された 2 平面の交線を指さし方向とする．

5.3 手の動作特徴の抽出

手領域の移動量を計測するために重心移動ベクトルを利用する．手の動作に重心移動ベクトルを用いる利点として，以下があげられる．

(1) 動作の特徴抽出や識別手法の構築が容易
(2) 計算コストが低く，高速な処理が可能

時系列で算出される重心座標の差をとることにより，重心移動ベクトルを算出する．Fig.4 に重心移動ベクトルの模式図を示す．重心移動ベクトル (dx_i, dy_i) は，

$$dx_i = x_{i+1} - x_i \tag{2}$$

$$dy_i = y_{i+1} - y_i \tag{3}$$

により求められる．

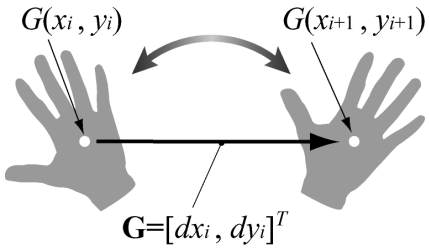

Fig.4　Movement vector of center of gravity

5.4 操作ジェスチャの認識

機器を操作するためのジェスチャを「操作ジェスチャ」とし，前節で述べた指の本数情報と動作特徴量とを組み合わせることで実現する．操作ジェスチャは，指の本数認識に加え，キャンセル動作認識，アップ・ダウン動作認識とする．

6. ジェスチャによる家電製品の操作

機器を操作するためには，まず，操作対象を決定する必要がある．そこで，前章で述べた「指さし方向認識」を利用し，操作する機器を指さすことで，操作対象を特定する．操作対象を選択した後，ジェスチャ認識により各機器の操作を行う．

7. 実験

7.1 実験システム

画像処理には HALCON(MVTec)を用い，各種演算処理は PC(DELL, Pentium 4 2.2GHz)で行った．画像は Pan-Tilt-Zoom 機能を搭載した 2 台の CCD カメラ

(SONY EVI-D100)で撮像し，画像分割ユニット(Panasonic WJ-MS488)で合成された映像をキャプチャボード(Leutron PicPort Color)を用いて PC に入力した．また，家電製品を操作するために，PC で制御可能な学習型赤外線リモコン(SUGIYAMA ELECTRON クロッサム 2+USB)を使用した．

7.2 手振りの認識

3 章で述べた手法に基づき，手振りの検出実験を行った．被験者 5 人を対象に，計測距離を変化させて認識率を算出した．手振り方法に関して次のことを条件とした．(1)画像中の任意の場所で手振りを約 2[s]間行う．(2)手振りを止めてから約 2[s]おいて次の手振りを再開する．以上の動作を 20 回試行し，そのうち約 2[s]以内で手振りを検出した場合を成功とした．距離による認識率を Table 1 に示す．広い計測範囲で高い認識率が得られた．認識失敗例として，手振りの幅が小さい場合が挙げられる．

Table 1　Recognition rate

4m	5m	6m	7m	8m
96%	96%	97%	92%	83%

7.3 指の本数認識

5 章で述べた手法に基づき，被験者 5 人に対して指の本数認識を行った．まず，被験者ごとに 4 章で述べた肌色登録を行い，続いて指の本数を 0～5 本に変化させ，それぞれの認識率を算出した．指の本数による認識率と検出結果の例をそれぞれ Table 2 と Fig.5 に示す．全ての指の本数において良好な認識率が得られた．認識の失敗例として，掌が極端に下方または上方へ傾いている場合が挙げられる．

Table 2　Recognition rate

0	1	2	3	4	5
90%	98%	96%	82%	88%	88%

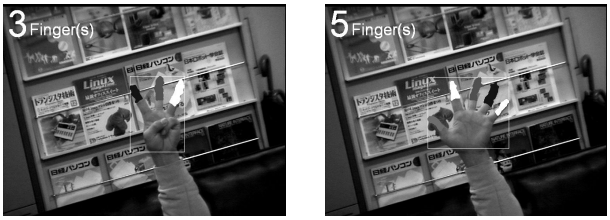

Fig.5　Experiment result

7.4 家電製品の操作実験

構築したジェスチャ認識を用いて家電製品の操作実験を行った．実験の様子を Fig.6 に示し，Fig.7,8 に操作ジェスチャの認識例を示す．誤動作に関しては，肌色登録の失敗による指の本数の誤認識が主な原因であった．その他はほぼ正しく動作した．

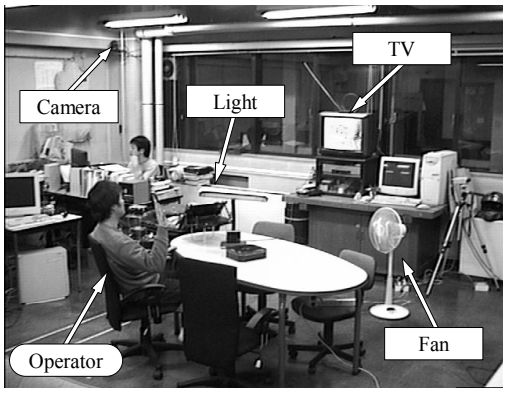

Fig.6　Situation of an experiment

Fig.7　“DOWN”

Fig.8　“CANCEL”

8. 結論

手を使ったジェスチャ認識手法を提案し，それをインタフェースとしたインテリジェントルームの構築を行った．また，実験により本手法の有効性を確認した．今後は，新たなジェスチャを追加し，機器の操作性向上を目指す．

参考文献

1) P. Hong, M.Turk, T. S. Huang：Gesture Modeling and Recognition Using Finite State Machines, IEEE Int. Conf. on Automatic Face and Gesture Recognition (2000).
2) H. Wu, T. Shioyama, and H. Kobayashi : Spotting Recognition of Head Gestures from Color Image Series, Proc. of the International Conference on Pattern Recognition, pp.83-85(1998).
3) Taketoshi MORI and Tomomasa SATO : Robotic Room “Its concept and Realization”, Robotics and Autonomous Systems, Vol.28, No.2, pp.141-144 (1999).
4) 山岸崇志，中西裕一，梅田和昇：インテリジェントルーム実現のためのジェスチャ認識，画像の認識・理解シンポジウム (MIRU'2000)，I，pp.289-294(2000).
5) 入江耕太，梅田和昇：濃淡値の時系列変化を利用した画像からの手振り検出，画像の認識・理解シンポジウム(MIRU’2002)，I，pp.285-290(2002).