

# 色情報を用いた生活支援ロボットに対する意思伝達方法の高精度化

## Accuracy improvement of a communication method for life support robots using color information

知能機械工学コース

知能ロボティクス研究室 森 大地

### 1. 緒言

少子高齢化に伴い、要介護（要支援）認定者数は平成 19 年度の約 453 万人から令和 5 年度の約 708 万人へ約 1.6 倍に増加<sup>(1)</sup>しており、在宅で支援を必要とする高齢者の増加に対し介護人材は不足している。介護サービスの中心は居宅であり、自宅環境での介護負担軽減が重要である。このような背景のもと、介護者の負担軽減と要介護者の自立生活支援を目的とした介護・生活支援ロボットの研究開発が進められている。

一方で、ロボットが十分な機械的機能を有していても、要介護者が要求や意図を適切に伝達できなければ実用的な支援は実現できない。音声認識やタブレット操作等は手順理解や細かい操作を要し、発話・認知機能が低下した対象者には負担が大きく、認識の不安定さも課題である。また、在宅では照明条件や背景物等の影響が大きく、ロボラスト性とリアルタイム性の両立が求められる。

そこで本研究では、人型単体多機能な生活支援ロボットの在宅支援を見据え、4つの色を組み合わせたカード提示によりタスクを指示する意思伝達方法に着目する。

本研究では、先行研究<sup>(2)</sup>で構築した色カード認識アルゴリズムを、自宅環境での背景写り込み等を考慮して一部処理フローを改良した改良基準手法をベースラインとして定義する。この枠組みを固定した上で色空間変換およびノイズ除去方法を差し替え、照明・距離・角度・背景条件下での認識精度と処理時間を比較評価する。

### 2. 生活支援ロボットシステムの概用

本研究室では、自宅環境における要介護者の生活支援および介護者負担の軽減を目的として、単体多機能型生活支援ロボット KUT-PCR を開発している。本ロボットは、人型上半身と全方向移動可能な下半身機構を備え、物品操作や室内移動など多様な支援タスクを行うことができる。

一方、ロボットが実用的に生活支援を行うためには、要介護者の要求や意図を正確に認識するための入力インターフェースおよび認識処理系が不可欠である。

KUT-PCR には頭部および胸部に RGB-D カメラが搭載されており、本研究では比較的遠距離からの意思伝達を想定し、頭部に搭載された Intel RealSense D435i を用いて画像を認識する。認識処理の再現性を確保するため、自動露光および自動ホワイトバランスを無効化し、解像度 1280×720 pixel、フレームレート 15 fps の条件で撮影した。

### 3. 意思伝達方法の概用

本研究では、寝たきり状態にある要介護者が在宅環境において生活支援ロボットに対して自身の要求や意図を簡便に伝達する手段として、色カードを提示する方式に着目する。要介護度が高く発話や複雑な操作が困難な場合、音声入力やタブレット端末等の従来インターフェースは入力として精度が低下し、安定させることが難しいと指摘されている<sup>(3)(4)</sup>。また、在宅環境では介護者が常時操作を代行できないため、

要介護者自身が直感的に利用可能で、かつロボット側で安定して認識できる入力方式が求められる。そこで本研究では、上肢の微細な操作や複雑な手順を必要とせず、視覚的に明確な物体をカメラに向けてだけで意思伝達が可能であるカード提示方式を採用した。

ロボット側の識別情報としては、文字認識や幾何学図形等も候補となるが、これらは照明変動や距離・角度変化、透視歪み、一部遮蔽の影響を受けやすい。一方、色領域は面積をもつ領域特徴として扱うことができ、多少のぼけや解像度低下が生じても領域全体の色分布が保持されやすい。また、色に基づく領域抽出は計算量が比較的少なく、リアルタイム処理への適用が期待される。以上より、本研究ではロボットが認識する情報として「色」を用いることが妥当であると判断した。

本研究で用いる色カードは、1枚のカード表面を4つの色領域に分割して構成し、カードの4つの回転状態および色の組合せ・配置順を情報として利用する。これにより、1枚のカードで4種類の指示を表現可能となるため、カード枚数を削減でき、要介護者の操作負担の低減が期待される。ロボットは搭載カメラから取得した画像に対して色領域抽出および形状解析を行い、カードの有無、向き、使用色の種類および並び順を認識することで、提示されたカードに対応する意思内容を推定する。実験で使用した色カードを図 1 に示す。裏面には利用者向けにイラストを配置し、その表示が正立となる向きに提示した場合その意志が伝達される。



Fig 1 Color Card

### 4. 色カード認識アルゴリズムの構成と基準手法の定義

本研究では、学習データ収集を必要とせず、再現性の高い実装が可能である点を重視し、深層学習等の学習を伴う手法ではなく、OpenCVに基づく典型的画像処理により認識系を構成する。

#### 4.1 先行研究における基準手法

本基準手法では、RGB カメラ画像を入力とし、色相 (Hue)、彩度 (Saturation)、明度 (Value) で表現される HSV 色空間への変換後、各色に対して閾値処理により二値化を行い、5枚の色別二値画像を生成する。各二値画像に対してモルフォロジー処理によるノイズ除去を施し、白画素領域から重心および主軸角度を算出する。主軸角度に基づきアフィン変換による傾き補正を行った後、外接矩形のアスペクト比を実カー

ドの値と比較することでカードの有無を判定する。さらに、アスペクト比により縦向き・横向きを識別し、各色の白画素数の大小関係から使用色を、上位4色として決定する。色順は、各色領域の重心座標を縦向き時はy座標、横向き時はx座標で比較することで推定する。本手法は処理構成が単純で計算量も小さい一方、色種判定および色順判定を画像全体で行うため、背景にカードと同色の物体が存在する場合、白画素数や重心の算出が影響を受けやすく、精度が低下する課題がある。

#### 4.2 改良基準手法（本研究のベースライン）

まず、HSV変換および色別二値化、ノイズ除去までは基準手法と同様に行う。その後、5枚の色別二値画像を論理和により合成し、カード候補領域を表す合成二値画像を生成する。合成画像に対して再度ノイズ除去を行う。合成画像に対し回転外接矩形を求め、そのアスペクト比や白ピクセル数、矩形充填率を設定値と比較することでカードの有無および向きを判定する。

カードが存在すると判定された場合、この回転外接矩形を用いてROIマスクを生成し、元の色別二値画像と論理積を取ることで、カード領域内のみを抽出する。色種判定はROI内の白画素数の大小関係から上位4色を選択し、色順判定はROI内の各色領域の重心位置を縦向き時はy座標、横向き時はx座標で比較することで行う。このように、カード領域を先に確定し、その内部に限定して色判定を行うことで、背景中の同色物体の影響を抑制できる構成とした。

#### 5. 比較評価方法

本研究では、第4章で定義した改良基準手法の枠組みを固定し、色空間およびノイズ除去方法を変更したときの認識性能への影響を比較評価する。評価は、在宅環境での運用を想定し、撮影条件（距離・照度・角度・背景）の変化に対して、

(1) カードを正しく認識できるか、(2) 処理時間、の2点から行う。比較対象として、色空間はHSVと、明度(L\*)と色成分(a\*, b\*)で表現されるLabの2種類とした。HSVは基準手法で用いており、色相・彩度・明度を分離して扱えるため閾値に基づく色抽出を設計しやすい。一方で照明変動の影響を受けやすい可能性があるため、在宅環境で差が現れやすい条件でLabと比較する。ノイズ除去は、モルフォロジー処理単体（基準）に加え、Medianフィルタ→モルフォロジー、モルフォロジー→連結成分解析(CC)の3方式を比較する。以上より、色空間2種×ノイズ除去3種の計6構成を評価対象とする。

実験環境は自宅の一室に設置したベッド上とし、在宅環境で寝たきり要介護者がベッド上でカードを提示する利用シーンを模擬した。居室内の壁・家具・生活用品等が視野に入り得る状況を再現することで、実運用に近い外乱（背景物体や照明変動）の影響を評価する。

撮影条件は、距離（1 m, 2 m, 3 m）、照度（160 lx, 600 lx, 1500 lx）、角度（正対、ピッチ±30°、ヨー30°）、背景（他物体なし、あり）の4因子とした。評価カードは縦向き1種類・横向き1種類の計2種類を使用した。

認識成否は「カードの有無」「向き（縦／横）」「使用色（4色）」「色順」がすべて正しく判定できた場合を成功と定義し、処理時間は各条件で複数回計測し、平均値で評価する。

以上の条件より、本実験では合計864通りの評価を行った。

#### 6. 結果

表1に6構成の総合認識率と処理時間を示す。比較の結果、HSV+Medianフィルタ→モルフォロジーにおいて75.0%で最も高い成功率を示した。一方、処理時間はHSV+モルフォロジーのみが50 ms/フレームで最短であった。Median→モルフォロジーはHSVで71 ms/フレーム、Labで222 ms/フレームとなり、他の方法より処理負荷が大きい傾向が見られた。また、HSV系はいずれのノイズ除去法でもLab系（63.9%）より高い成功率を示し、処理時間はHSV系（50～71 ms/フレーム）に対してLab系（189～222 ms/フレーム）は大きい傾向を示した。

背景干渉においては、最良構成（HSV+Median→モルフォロジー）では、背景干渉なし84.7%に対し背景干渉あり65.3%となり、背景混入が主要な性能低下要因となった。

先行研究では背景中に同色物体が存在する場合、認識が不可能であったが、本研究ではカード領域を推定してROI内で判定する手法に改良し、背景に同色物体がある場合でも外乱（距離・照度・角度）の条件によっては認識可能となった。

Table 1 Recognition rate and processing time for six configurations

Color space	Noise removal method	Recognition rate (%)	Processing time (ms/frame)
HSV	Morph only	73.6	50
HSV	Median → Morph	75.0	71
HSV	Morph → CC	73.6	62
Lab	Morph only	63.9	189
Lab	Median → Morph	63.9	222
Lab	Morph → CC	63.9	200

#### 7. 結言

本研究では、寝たきり要介護者が色カードを提示して生活支援ロボットに意思を伝達する状況を想定し、色カード認識アルゴリズムの基準手法（先行研究）と、合成二値画像に基づくROI確定を含む改良基準手法（本研究）を定義した。さらに、色空間（HSV, Lab）とノイズ除去（モルフォロジーのみ、Median→モルフォロジー、モルフォロジー→連結成分）を組み合わせた6構成について、在宅環境（距離・照度・角度・背景干渉）下で総合認識成功率と処理時間を比較評価した。その結果、HSV+Medianフィルタ→モルフォロジーの構成が6構成中で最も高い総合認識成功率を示し、本研究条件下における有効な処理構成であることを確認した。

#### 文献

- (1) 厚生労働省, “令和5年度介護保険事業状況報告(年報)”
- (2) 森大地, “色情報を用いた生活支援ロボットに対する意志伝達方法の開発”, 高知工科大学学士論文, (2023)
- (3) 谷岡稔真, 江頭広幸, 高田真由美, 岡崎泰久, 渡辺健次, 近藤弘樹, “発話障害のある肢体不自由者のための音声による文字入力とPC操作を可能にするシステムの開発,”人工知能学会論文誌, 2008
- (4) 畠中 順子, 高橋 美和子, 下村 義弘, 原田 悦子, “タッチ端末操作における操作特性の加齢変化 —2024年度計測結果の報告—,” 人間工学, 2008