3U-9

距離センサーと画像処理による移動体認識

広瀬 大樹† 三好 力†

龍谷大学理工学部†

1. はじめに

人間は，ある対象物を見たとき，その 3 次元形状，サイズ，テクスチャ等の情報を瞬時に分析し，複数のカテゴリーの認識を同時平行的に行うことが出来る．しかし機械は，画像を多少幾何学的変換が行われただけでも物体抽出・認識を行うことが非常に難しくなる．これは，人間が物体を 3 次元的に捉え，理解するのに対し，機械は平面的に画像を捉えることに起因すると考えれば，機械で人間と類似の高度な認識を行うには処理に人間と類似のアプローチを行う必要がある．本研究は赤外線距離センサと RGB カメラを用いて 3 次元形状，サイズ，テクスチャ等の特徴から，複数のカテゴリの認識が可能なシステムの構築を目指し，その基礎となる物体認識処理について Kinect による距離画像を用いたフィルタを作成し，RGB 画像とあわせた画像領域分割システムを作成することを目的とする．

2. 物体認識の概要

物体認識とは画像をシステムに入力し，入力に対し適切なラベルを付与する過程のことをいう．一般に実世界画像に対する物体認識は大きく分けて同定と分類の 2 種類の認識がある．同定とは画像内の物体を個々に区別する認識であり，入力画像とデータベース内のモデルと照合を行い，画像内の物体がデータベース内に存在するか，またどのモデルと一致するかを出力結果とする一方分類は画像内の物体の種類を区別する認識であり，人間が定めたクラスと画像内の物体を対応付け，対応したクラス名を出力結果とする．物体認識で行われている研究は特定物体認識と一般物体認識の 2 種類に分類することができる．特定物体認識は同定，一般物体認識は分類の分野を扱っている．物体認識処理の流れは学習時と識別時のどちらでもデータ入力，特徴抽出，識別記の順に処理されていくのが一般的である．[1]

2.1. 領域分割

画像の解析，理解を行う場合に画像中に含まれる対象を分離，抽出することが必要となる．この処理を領域分割を呼ぶ．エッジや線の抽出，追跡して領域を分割する方法も考えられるが，同一領域は類似の特徴を持つという性質を利用して，領域を分割する手法が一般に用いられている．後者の手法では対象の境界線は必ず閉じたものとして得られ，ノイズの影響を受けにくいという利点がある．ここでいう領域とは画像中で共通の性質を持つ部分のことをいう．通常，画像中でひとかたまりになっているが，お互いに離れた複数の部分から構成される場合もある．[2]

3. 3D 特徴を用いた画像領域分割

3.1. 概要

従来の画像領域分割では，RGB 画像の画素情報を使った領域分割が行われているが画面の全領域を RGB データのみで分割すると物体の配置や光の反射具合によって非常に難しくなる場合がある．本研究ではゲームコントローラ Kinect に搭載されている近赤外線距離画像センサを用いてフィルタを作成し画像の切り出しを行うことによってより複雑に物体が配置された画像から中心に配置された物体の領域を容易に抽出できるシステムを提案する．

3.2. フィルタの作成

画像領域分割を行うために Kinect から得た情報を用いて，領域分割に用いる特徴量を決定する必要がある．特徴量の決定には Kinect の近赤外線距離画像センサから得られる深度画像を利用する．深度画像から距離データを用いて有効な情報を決定し深度画像を 2 値化する．この 2 値化画像をフィルタとして画像領域分割を行う．特定の範囲内に物体が存在する座標を 1，存在しない座標を 0 とし，2 値化画像の作成を行う．

3.3. フィルタを用いた画像領域分割

画像領域分割には 3.2 で作成した 2 値化画像をフィルタとし 1 となった部分とカメラ画像を重ねる．重なった部分を切り出す．距離画像を元に切り出しを行っているため，RGB 画像では領域分割が難しい画像でも容易に領域分割が行える．

4. 実験

提案手法の有効性を検証するため，RGB 画像を基に作成したフィルタを用いて，目標物体の抽出を行う従来手法と，提案手法の出力画像を比較し，どちらが目標物体の抽出が正確に行えているかを評価した．また，画像取得から切り出された画像の出力までの処理時間の比較し，どちらの処理が早く，リアルタイム処理に適しているかの評価も行った．

Migrants recognition by image processing using distance sensor

†HIROSE Daiki , MIYOSHI Tsutomu , Faculty of Science and Engineering, Ryukoku University

Copyright ©2013 Information Processing Society of Japan.
All Rights Reserved.

4.1. 実験方法

従来手法，提案手法の比較実験を行うために以下の手順で実験を行った．

・Kinectから約1.5m離れた位置でカメラ画像の中心となるよう目標物体を配置
・KinectからRGB画像，距離画像を取得
・フィルタの作成
・フィルタを用いてカメラ画像から目標物体の切り出し
・切り出された画像と処理時間の出力
・出力された画像，処理時間の比較

上記手順の内フィルタの作成については，従来手法はRGB画像を基にフィルタを作成し，提案手法はKinectの深度画像を基にフィルタを作成した．

4.2. 実験環境

本実験では目標物体の認識距離を1.0m～1.5mとして認識を行った．Kinectは地面より0.7mの高さに配置，全物体の配置距離はKinectの認識可能距離である0.5m~5.0mの位置に配置した．開発環境はVisual C++ 2010，RGBカメラからのカメラ画像の取得，近赤外線距離画像センサからの深度画像の取得にOpenNI 1.5.2.23，取得した画像の変換，表示についてはOpenCV 2.2.0を用いて行った．

4.3. 実験結果

4.3.1. 実験1

目標物体のみが画面内に配置された場合

実験に用いた元画像を図1，従来手法の出力結果を図2，提案手法の出力結果を図3に示す．従来手法では目標物体が光を反射している部分の切り出しがうまくいっていないが，提案手法では目標物体全体の切り出しが正確に行われている．また，従来手法の出力結果は壁の一部が切り出されてしまっているが，提案手法では目標物体のみの切り出しができている．

図1. 実験1:元イメージ

図2. 実験1:従来手法

図3. 実験1:提案手法

4.3.2. 実験2

画面内に物体が複数配置された場合

実験に用いた元画像を図4，従来手法の出力結果を図5，提案手法の出力結果を図6に示す．従来手法では4.3.1.の結果と同様に目標物体の全体がうまく切り出されず，さらに目標物体以外の物体も切り出してしまっている．提案手法では，目標物体の全体が正確に切り出され，他の物体の切り出しも行われていない．

図4. 実験2:元イメージ

図5.実験2:従来手法

図6.実験2:提案手法

4.3.3. 処理時間の比較

従来手法と提案手法の処理時間の平均を表1に示す．処理時間のサンプル数は各200個とした．フィルタを作成する際に従来手法ではRGBカメラから得た画像をグレースケールに変換する処理をしたが，提案手法では近赤外線距離画像センサから得た距離画像をそのままフィルタ作成に用いたことによって処理時間の差が出たと思われる．

表1. 従来手法と提案手法の平均処理時間

	目標物体のみ配置	複数物体配置
従来手法	16.02[ms]	15.45[ms]
提案手法	14.80[ms]	15.01[ms]

5. おわりに

本論文では3Dセンサが搭載されているKinectによる距離画像を用いたフィルタを作成し，RGB画像とあわせた画像領域分割システムを作成した．実験の結果から目標物体のみの単純な画面，物体が複数存在する複雑な画面共に中心物体の切り出し精度が向上し，処理時間も削減できる利点を確認した．

今後は，距離情報とRGB情報で領域分割された画像を用いて，中心物体の形状，サイズ，テクスチャなどの特徴を用いた物体認識システムの特徴抽出部分の研究を進める．

参考文献

[1]佐藤元昭，一般物体認識における画像の学習効率および認識率向上に関する検討，早稲田大学(2010.2)
[2]“画像内の領域分割と統合”
http://cryst.tv/experiment/ImageProcessing/DomainDivisionAndIntegration/

Copyright ©2013 Information Processing Society of Japan.
All Rights Reserved.