OUKA
Osaka University Knowledge Archive

Title	Machine Visual Perception in Cluttered Environments
Author(s)	Almaddah, Amr Reda
Citation	大阪大学, 2013, 博士論文
Version Type	
URL	https://hdl.handle.net/11094/26263
rights	
Note	やむを得ない事由があると学位審査研究科が承認したため、全文に代えてその内容の要約を公開しています。全文のご利用をご希望の場合は、大阪大学の博士論文についてをご参照ください。

Osaka University Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

Osaka University

Form 3

Synopsis of Thesis

Title: Machine Visual Perception in Cluttered Environments
（複雑環境下の視覚認識）

Name of Applicant: AMR REDA R. ALMADDAH

Machine vision is the science concerned with developing theoretical and algorithmic approaches by which useful information is obtained and analyzed from an image or a set of images of a scene of interest. Cluttered environment, due to harsh illumination or complex multi colored scene components, comes out as one of the biggest challenges to the field of machine vision. In this dissertation, we propose the use of the relationship between light and objects albedo to achieve robust visual perception in cluttered environments. As part of the face recognition task in a robust security system, we propose a novel approach for the illumination recovery of faces with cast shadows and specularities. Given a single 2D face image, we relight the face object by extracting the nine spherical harmonic basis and the face spherical illumination coefficients by using the face spherical spaces properties. First, an illumination training database is generated by computing the properties of the spherical spaces out of face albedo and normal values estimated from 2D training images. The training database is then discriminately divided into two directions in terms of the illumination quality and light direction of each image. Based on the generated multi-level illumination discriminative training space, we analyze the target face pixels and compare them with the appropriate training subspace by using pre-generated tiles. When designing the framework, practical real time processing speed and small image size were considered. In contrast to other approaches, our technique requires neither 3D face models, nor restricted illumination conditions for the training process. Furthermore, the proposed approach uses one single face image to estimate the face albedo and face spherical spaces. For autonomous robots, the ability to manipulate unknown objects is pivotal to successfully operate in a complex environment. To tackle the issue in hand, we propose a novel multi lighting two-step method for unknown objects segmentation using albedo characteristics of object shape. We address the task of detecting unknown multi colored objects of interest, specifically ones that are being seen for the first time and located in the proximity of unidentified obstacles. Illumination at different wavelengths and angles are projected by a robot, thus acquiring additional information about the scene and exploiting it for successful unknown objects segmentation. By analyzing shades and reflections of a scene's objects, we were able to form and identify true edges and visually separate items of interest. Our proposed method does not require predefined models of target objects and assumes no previously assigned targets pose. We validated our theoretical methodology and technical solutions by presenting promising experimental results showing significant improvements in machine visual perception after actively and passively making use of the light-object relationship.

様式7

論文審査の結果の要旨及び担当者

氏　名　（AMR REDA R. ALMADDAH）		
	（職）	氏　名
論文審査担当者	主査　教授　新井健生 副査　教授　佐藤宏介 副査　教授　石黒浩	

論文審査の結果の要旨

実世界の様々な照明条件下において有効に機能するマシンビジョンの実現は極めて困難である．しかし，様々な照明条件において撮像された画像から推定される物体表面の反射特性を利用することにより，ロバストな視覚認識が実現できる可能性がある．本論文では始めに顔認識のための任意照明下における一枚の入力画像からのリライティング手法を提案する．顔画像データベースとして，各人物について少なくとも一枚の顔画像が与えられているものとする．提案手法では，顔画像データベースから区分された2次元の訓練空間を構築することによって頑健な顔画像のリライティング手法を実現している．訓練空間は，訓練画像から推定される顔のアルベドと法線ベクトルからなる球面調和空間の特性を計算することにより得られる．訓練空間は顔画像の照明の明るさと方向の2次元空間であり，各顔画像は，区分された訓練空間の対応する区分へ自動的に配置され，区分訓練空間を得る．入力画像の画素の明るさを解析し，この区分された訓練空間から入力画像が対応する区分を選択する．画像は，あらかじめ決めた3×4画素の大きさの複数のタイルから構成されているとし， 区分に属する訓練画像の中から最も類似したタイルを見つけ，そのタイル部分を出力画像の一部とする．この処理を全てのタイルについて行うとライティングされた出力画像を得る．入力は，一つの2次元画像であり，出力はリライティングされた顔画像である．球面調和基底を用いた画像表現と，画像からアルベドと法線方向を推定することによって，訓練データとして顔の3次元データを用いずに，１枚の入力となる顔画像からリライティングされた画像の生成を可能としている．

また，顔認識だけではなく，サービスロボットによる物体認識への応用も図った．サービスロボットには，人の指示に従って目標とする物を見つけ，物を人に渡すといった作業が求められる．このような高度な作業を実現するためには，三次元空間での目標物体の検出，物体の位置姿勢の推定を行う視覚機能が必要となる．これらの機能が様々な対象物体や環境の変化にも頑健であることが求められる．特に，物体モデルを用いずに未知環境において物体ごとに領域分割をすることは困難な問題である．本論文では，人の屋内生活環境を想定した未知環境において未知物体の領域を抽出することを目的とする．このような作業に求められる視覚機能を実現するために，本論文では，能動的に照明条件を変化させ，物体の空間配置や物体の反射特性の違いから生じるシーンの各部の変化を解析し，シーンを物体領域に分割する視覚情報処理技術を提案する．ロボットは複数の異なる波長と角度でシーンを照明し，複数の照明条件下においてシーンを撮像する．このように能動的に照明されたシーンを解析することにより，物体境界に生じるエッジや物体表面の反射特性の違いから未知物体の境界を決定し，物体領域を分離して抽出することが可能となった．

以上の通り，本論文は視覚認識に関する新たな方法論やシステムの構築を行い，従来困難とされた複雑な環境下における対象物や様々な照明条件における顔認識の実用的な手法を提案しており，学術的かつ技術的な貢献が大きく，博士（工学）の学位論文として価値のあるものと認める．