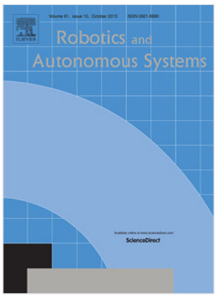
Accepted Manuscript



Visual recognition for humanoid robots

Sean Ryan Fanello, Carlo Ciliberto, Nicoletta Noceti, Giorgio Metta, Francesca Odone

PII: S0921-8890(15)30181-0
DOI: http://dx.doi.org/10.1016/j.robot.2016.10.001
Reference: ROBOT 2714

To appear in: *Robotics and Autonomous Systems*

Received date : 1 November 2015
Accepted date : 4 October 2016

Please cite this article as: S.R. Fanello, et al., Visual recognition for humanoid robots, *Robotics and Autonomous Systems* (2016), http://dx.doi.org/10.1016/j.robot.2016.10.001

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

ACCEPTED MANUSCRIPT

Visual Recognition for Humanoid Robots

Journal name
000(00):1–xx
©The Author(s) 2010
Reprints and permission:
sagepub.co.uk/journalsPermissions.nav
DOI:doi number
http://mms.sagepub.com

Sean Ryan Fanello[3] *, Carlo Ciliberto[1,2], Nicoletta Noceti[2], Giorgio Metta[1] and Francesca Odone[2]

[1] *iCub Facility, Istituto Italiano di Tecnologia, Genova, Italy,*
[2] *DIBRIS, Università degli Studi di Genova, Genova, Italy*
[3] *Microsoft Research Labs, Redmond, WA 98051, USA.*

Abstract

Visual perception is a fundamental component for most robotics systems operating in human environments. Specifically, visual recognition is a prerequisite to a large variety of tasks such as tracking, manipulation, human-robot interaction. As a consequence, the lack of successful recognition often becomes a bottleneck for the application of robotics system to real-world situations. In this paper we aim at improving the robot visual perception capabilities in a natural, human-like fashion, with a very limited amount of constraints to the acquisition scenario. In particular our goal is to build and analyze a learning system that can rapidly be re-trained in order to incorporate new evidence if available. To this purpose, we review the state-of-the-art coding-pooling pipelines for visual recognition and propose two modifications which allow us to improve the quality of the representation, while maintaining real-time performances: a coding scheme, Best Code Entries (BCE), and a new pooling operator, Mid-Level Classification Weights (MLCW). The former focuses entirely on sparsity and improves the stability and computational efficiency of the coding phase, the latter increases the discriminability of the visual representation, and therefore the overall recognition accuracy of the system, by exploiting data supervision. The proposed pipeline is assessed from a qualitative perspective on a Human-Robot Interaction (HRI) application on the iCub platform. Quantitative evaluation of the proposed system is performed both on in-house robotics data-sets (iCubWorld) and on established computer vision benchmarks (Caltech-256, PASCAL VOC 2007). As a byproduct of this work, we provide for the robotics community an implementation of the proposed visual recognition pipeline which can be used as perceptual layer for more complex robotics applications.

Keywords
Human-Robot Interaction, Learning and Interaction, Visual Recognition, Sparse Representations, iCub

* Corresponding author; e-mail: seanfa@microsoft.com

Download English Version:

https://daneshyari.com/en/article/4948789

Download Persian Version:

https://daneshyari.com/article/4948789

Daneshyari.com