UNIVERSITY OF TECHNOLOGY, SYDNEY

Affordance-Map : Learning Hidden Human Context in 3D Scenes Through Virtual Human Models

by

Jayaweera Mudiyanselage Lasitha Chandana Piyathilaka

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Engineering and IT
Centre of Autonomous Systems

February 2016

Declaration of Authorship

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis.

Signature of Student:

Date:

UNIVERSITY OF TECHNOLOGY, SYDNEY

Abstract

Faculty of Engineering and IT
Centre of Autonomous Systems

Doctor of Philosophy

by Jayaweera Mudiyanselage Lasitha Chandana Piyathilaka

Ability to learn human context in an environment could be one of the most desired fundamental abilities that a robot should possess when sharing workspaces with human coworkers. Arguably, a robot with appropriate human context awareness could lead to a better human robot interaction. This thesis addresses the problem of learning human context in indoor environments by only looking at geometrics features of the environment. The novelty of this concept is, it does not require to observe real humans to learn human context. Instead, it uses virtual human models and their relationships with the environment to map hidden human affordances in 3D scenes.

The problem of affordance mapping is formulated as a multi label classification problem with a binary classifier for each affordance type. The initial experiments proved that the SVM classifier is ideally suited for affordance mapping. However, SVM classifier recorded sub-optimum results when trained with imbalanced datasets. This imbalance occurs because in all 3D scenes in the dataset, the number of negative examples outnumbered positive examples by a great margin. As a solution to this, a number of SVM learners that are designed to tolerate class imbalance problem are tested for learning the affordance-map. These algorithms showed some tolerance to moderate class imbalances, but failed to perform well in some affordance types.

To mitigate these drawbacks, this thesis proposes the use of Structured SVM (S-SVM) optimized for F1-score. This approach defines the affordance-map building problems as a structured learning problem and outputs the most optimum affordance-map for a given set of features (3D-Images). In addition, S-SVM can be learned efficiently even on a large extremely imbalanced dataset. Further, experimental results of the S-SVM method outperformed previously used classifiers for mapping affordances.

Finally, this thesis presents two applications of the affordance-map. In the first application, affordance-map is used by a mobile robot to actively search for computer monitors in an office environment. The orientation and location information of humans models inferred by the affordance-map is used in this application to predict probable locations of computer monitors. The experimental results in a large office environment proved that the affordance-map concept simplifies the search strategy of the robot. In the second application, affordance-map is used for context aware path planning. In this application, human

context information of the affordance-map is used by a service robot to plan paths with minimal distractions to office workers.

Acknowledgements

I am sincerely in debt to my supervisor, Associate Professor Sarath Kodagoda for believing in me and for giving me enough flexibility to work on my lifelong passion. His continuous support and guidance made this difficult endeavor achievable.

I would also like to thank Professor Massimo Piccardi for the fruitful discussions we had together. He always had enough time for me in his busy schedule.

To all the academic and support staff at CAS, thanks for taking me in and making me part of the family. Never have I felt more at home than in these past four years.

I would also like to thank Australian Government and University of Technology Sydney for awarding me with International Postgraduate Research Scholarship (IPRS) and Australian Postgraduate Award (APA). Without these scholarships I certainly would have not studied for a PhD.

Finally, I am ever in debt to my family. Specially my parents who made great sacrifices to make me who am I today. A very special thanks to my wife Uthsu for continuously believing me in and encouraging me in this difficult journey. Thanks my lovely daughter Sethmi for making my life joyful and meaningful.

Contents

List of Figures

List of Tables

Abbreviations

3D	Three Dimensional
2D	Two Dimensional
DCM	Different Cost Model
FNB	Flexible Naive Bayes
GPMD	Global Path with Minimum Distractions
HMM	Hidden Markov Model
IFTM	Infinite Factored Topic Model
KDE	Kernel Density Estimation
MRF	Markov Random Fields
PRM	Probabilistic Road Map
RGB-D	Red, Green, Blue and Depth
SVM	Support Vector Machine
S-SVM	Structured Support Vector Machine
SLAM	Simultaneous Localization and Mapping
SMOTE	Synthetic Minority Over Sampling Technique

Nomenclature

$f(\cdots)$	A scalar valued function
$\boldsymbol{f}(\cdots)$	A vector valued function
$P(x)$	Probability of a random variable x
$P(x\|z)$	Conditional probability of x, given evidence z
$P(x, z)$	Joint probability of x and z
$DT[\cdot]$	Distance transform
$[\cdot]^T$	Transpose of a vector or a matrix
$\\|\cdot\\|$	The magnitude of a vector
$\Phi(\cdot)$	Feature function
$E[\cdot]$	The expectation of a random variable
$K((\cdot),(\cdot))$	Kernel function
	Specific Symbol Usage
A	Affordance-map
i_k	i^{th} Point cloud image
y_i	i^{th} Label
$\mathbf{x}_n$	Feature vector at i^{th} grid location
$\mathbf{g}_i$	i^{th} Grid location of the 3D map
C_k	k^{th} Class label
$\mathbf{w}$	The normal vector to the separating hyperplane of the SVM classifier
b	The offset of the hyperplane from the origin along the normal vector of the SVM classifier
ξ_i	i^{th} Slack variable

α_i	i^{th} Lagrange multiplier
$R_z(\theta_k)$	The rotational matrix about the z axis
H_l	Human skeleton model
$\bar{y}_k$	k^{th} Ground truth label
$\triangle(\bar{y}', \bar{y}_k)$	Loss function
μ_s	The mean of position and orientation of the robot in human skeleton co-ordinate system
Σ_s	The covariances of position and orientation of the robot in human skeleton co-ordinate system

Glossary of Terms

Affordance	All action possibilities latent in the environment, objectively measurable and independent of the individual's ability to recognize them.
Environment	A complex 3D unstructured place . Assumed to have some structural characteristics such as planar surfaces.
Features	Distinctly identifiable points in an image of a 3D environment.
Freespace	Areas in the environmental model or map that are known to be free of objects, obstacles and surfaces.
Iteration	A single step which is determined by optimisation.
Obstacle	An object within the environment which a robot can collide with.
Occlusion	Not visible from a viewpoint due to an obstruction.
Pose	The combination of position and orientation of an object in 3D space with respect to a reference coordinate frame.
Planning	The act of generating a path (and motion) course which the robot can then follow to get between two poses.
Surface	This is the face of an object in the environment.
Surface Normal	A 3D vector perpendicular to a surface.
Unstructured	Real-world environment that cannot be set up to facilitate experiment.
Viewpoint	A position in space and an orientation of a sensor that a corresponding robot pose can achieve.
Voxel	Volumetric Pixel which represents a 3D cube-like volume in Euclidean space.