ISR – Institute of Systems and Robotics
University of Coimbra - Portugal

http://paloma.isr.uc.pt
Human-Robot Interaction

Horopter Based Dynamic Background Segmentation applied to an Interactive Mobile Robots

José Prado, Luis Santos and Jorge Dias

14th International Conference on Advanced Robotics (ICAR 2009)
Human-Robot Interaction

Summary

1. The conceptual idea of Learning by Observation

2. The paper contribution (Horopter Based DBS applied to an Interactive Mobile Robot)

3. The Bayesian Model to be implemented in near future
Human-Robot Interaction

What is an Interactive Mobile Robot?

In HRI context, an Interactive Robot is a robot able to interact with humans, it is a candidate to become a social robot, but not necessarily have social skills.

An Interactive Guide Robot is a robot pre-programmed to some guidance among a place, but also can decide to stop its presentation to interact with the humans in front of it.
Human-Robot Interaction

HRI is a wide area of robotics

ConnectR shares sound and images, and the robot can be remote controlled by the internet

Haptic Master
FCS Robotics, Inc.
Larger workspace

SensAble Technologies, Inc.
Virtual sculpting

BCI (Brain Computer Interface) devices

RoboKnee
Yobotics, Inc.
Pratt, et al.
Body force input
1 DOF Knee
Enhances strength and endurance

University of Coimbra

Intuitive Surgical, Inc.

Larger workspace

FCS Robotics, Inc.

Larger workspace

FCS Robotics, Inc.

Larger workspace

SensAble Technologies, Inc.

Virtual sculpting

BCI (Brain Computer Interface) devices
1. All these approaches are passive, or in other words, the robot is somehow controlled by a human. He is not capable to make complex decisions by itself.

2. For machines to take on bigger roles in society they need to learn how to adapt to people’s behaviors. “Lola Canamero (University of Hertfordshire, UK), the coordinator on the Feelix Groing Project”

3. How to teach a robot to interact with humans?
Human-Robot Interaction

Possibility of recognizing human natural behaviours:
- Gesture Recognition
- Speech Recognition
- Expression Recognition

Belief
The belief, that we shall give the machines a way to learn by themselves how to interact with us.

Possible strategies:
1. Non-stochastic-models (deterministic logic),
2. Discriminative Models (deterministic classifier),
3. Descriptive Stochastic Model (Probabilistic Without Prediction),
4. Generative Stochastic Model (Bayesian Network – Probabilistic With Learning and Prediction)

Bayesian Approach for Robotic Behaviour Learning by Observation
Human-Robot Interaction

What we are not doing…

Our main research focus is the Bayesian Learning,
Our main focus for this paper (preliminary results) was the Horopter based Dynamic Background Segmentation:

1. We are not concerned in doing the best natural way to detect the human commands (because this would be just improving the robustness of detectors or adding some already existant classifiers).

2. We are not going to analyse expressions neither speech recognition, nor sound source detection (however we understand it would be a more natural way of commands)
Human-Robot Interaction

Study case – Guide Robot

1. We already know that giving to a robot a capability to learn can be useful in several aspects of our society.

2. As a study case we choose to do a Guide Robot, that is capable to learn a gesture Interaction between two humans.

3. After learning, he should be able to know what command to take when he want to create an expected reaction on the human.
Human-Robot Interaction

Current Scenario:

Problem: We need to segment the human to interact, from the background and possible noise.
Human-Robot Interaction

Approaches to Static Background Segmentation:

Static Background Segmentation usually applied for camera surveillance, is not a solution to mobile robots

Problem: In a mobile robot, the background is always changing
Human-Robot Interaction

The Horopter:

Our approach starts by segmenting the background by an algorithm based on the mathematical concept of Geometrical Horopter.
**Proposed Solution:**
Based on the mathematical concept of horopter, we calculate the disparity map between the images from a stereo system and we determine where is the horopter on the 3D space.
Human-Robot Interaction

Problem: Detecting the user to interact

Implementation was done by using SVS SRI libraries for stereo cameras. A modification was done to accept 2 monocular cameras instead of one VIDERE stereo camera.

...next step, detect face and hands
Human-Robot Interaction

Problem: Detecting interactive parts of user (face and hands)

The Mean Shift algorithm is a robust, non-parametric technique that climbs the gradient of a probability distribution to find the mode (peak) of the distribution (Fukunaga, 1990).

The CAMShift is the Continuously Adaptive Mean-Shift algorithm.
Human-Robot Interaction

Problem: Detecting interactive parts of user (face and hands)

CamShift results

Without skin color at the background

With skin color at the background
Human-Robot Interaction

Problem: Detecting interactive parts of user

CamShift results

Without background segmentation

With background segmentation
Human-Robot Interaction
Human-Robot Interaction

Implementation: High-level schema

- Dynamic Background Segmentation
  + Hand&Face Detection
  Contribution

- Laban Movement Analysis Gesture Recognition Algorithm

- Bayesian Learning by Observation
  Contribution

University of Coimbra

INSTITUTO DE SISTEMAS E ROBÓTICA

IMR Lab
Human-Robot Interaction

Laban Movement Analysis

Non kinematic components: Body (Body parts) and Space (Trajectories)
Kinematic components: Effort (dynamic qualities of the movement) and Shape (Body shape and pose)
Proposed Bayesian Network

Bayesian Inference

\[ P(C) = \text{Uniform} \]
\[ P(R|C) = \text{acquired from learning by observation} \]
\[ P(R) = \text{Uniform} \]
\[ P(C|R) = \frac{P(C) \cdot P(R|C)}{P(R)} \]

Based on Independence theorem

\[ P(L|C) = P(L) \& P(C|L) = P(C) \]
Where:

\[ G = \text{gesture} \]
\[ L = \text{laser, collision avoidance} \]
\[ C = \text{Command} \]
\[ R = \text{Response} \]

\[ P(R|L,C)P(C|G)P(L)P(G) \]
Proposed Bayesian Program

Program

Description

Variables:
\[ C_{0:t} = \text{: Gesture} \]
\[ C_{0:0} = \text{: Command} \]
\[ R_{0:t} = \text{: Response} \]
\[ L_{0:t} = \text{: Laser Collision Avoidance} \]

Decomposition:
\[ P(C^t \land R^t \land L^t) = \]
\[ P(R^t|L^t \land C^t) \]
\[ .P(C^t|G^t) \]
\[ .P(L^t).P(G^t) \]

Parametric forms:
\[ P(R^t|C^t) \equiv \text{: Histogram, given from the learning by observation} \]
\[ P(G^t) \equiv \text{Histogram} \quad \text{: PDF resultant from LMA Gesture BM} \]
\[ P(C^t) \equiv \text{Uniform} \quad 1 \quad \text{na} = \text{number of possible actions} \]
\[ P(R^t) \equiv \text{Uniform} \quad 1 \quad \text{nr} = \text{number of possible reactions} \]
\[ P(C^t) \equiv \text{Uniform} \quad 1 \quad \text{nc} = \text{number of possible collision in the scope of the sensor maximum range} \]
\[ P(L^t|C^t) \equiv P(L^t) \]
\[ P(C^t|L^t) \equiv P(C^t) \]

Identification:
Bayesian Unsupervised Learning for Human Behaviour Analysis

Question:
\[ P(C^t|R^t) = \frac{P(C^t)P(R^t|C^t)}{P(R^t)} \]
Future Work: Possible Expansion

Bayesian Inference

\[
P(C) = \text{Uniform}  \\
P(R|C) = \text{acquired from learning by observation}  \\
P(R) = \text{Uniform}  \\
P(C|R) = \frac{P(C) \cdot P(R|C)}{P(R)} \\
\]

Based on Independence theorem

\[
P(L|C) = P(L) \& P(C|L) = P(C) \\
\]

Where:

G = gesture  \\
S = sound (from human source)  \\
E = emotion  \\
L = laser, collision avoidance  \\
C = Command  \\
R = Response  \\
P(R|L, C)P(C|G, E, S)P(L)P(G)P(E)P(S)
Future Work: Interactive phase – active mode

As a robot guide, robot will do the pre-programmed guidance (for a small group), and during this, it observes the people in the group.

By using the knowledge acquired on the learning phase, robot will be able to actively interrupt the presentation if some “bad” behaviour is found.

Bad behaviours
If it detect that a person is “sad” (expression analysis)
Is looking at the wrong direction for a long time (head pose estimation)
Bad behaviour example arm gesture (we can force an example gesture to be considered a bad behaviour for our experimental setup)
Questions ?