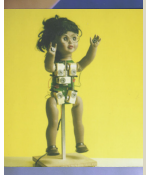
PDA Interface for Humanoid Robots

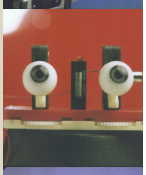
Sylvain Calinon, Aude Billard
{sylvain.calinon,aude.billard}@epfl.ch
ASL3-I2S-STI, Swiss Federal Institute of Technology Lausanne, EPFL
http://asl.epfl.ch
Autonomous Systems Lab


A growing trend in Robotics investigates the integration of user-friendly means of interacting and reprogramming humanoid robots. Unfortunately, efficient methods for speech and vision processing remain computationally expensive, and cannot be easily exploited on cost- and size-limited platforms.

The recent introduction on the market of affordable humanoids and toy robots has created a need for hardware and software solutions to provide simple speech and vision-based communication. This work shows that **Personal Digital Assistants (PDAs)** are ideal low-cost platforms to interface humanoid toy robots.

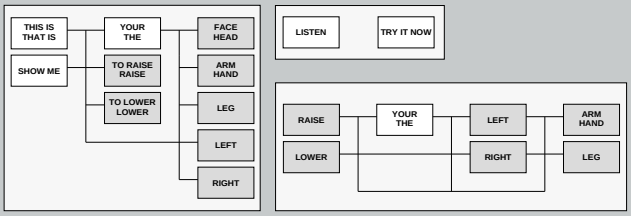
We present **PDA applications using vision processing, speech recognition, and speech synthesis**. In these applications, the robot can track and imitate the movement of the user's arms and head, and can learn a simple vocabulary to label objects and actions by associating the user's vocal utterances with the user's gestures.


ENVIRONMENT

SPEECH MODULE

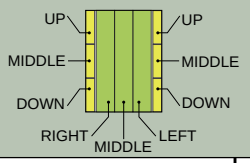
Parses the vocal utterances and extracts only the relevant keywords (shown in grey below) using syntactic rules. Speech is translated into sets of strings by the *Conversay* speech engine for PDAs.

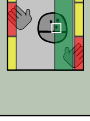


Preprocessing of sensory data before learning and retrieval

VISION MODULE

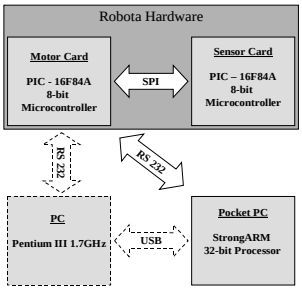
Builds a simplified representation of the user's gestures to track the vertical motion of the arms and the rotation of the head. Visual processing is performed by the PDA, using a *Flycam-CF* camera connected to the *CompactFlash* cards slot. Tracking of the arms is based on luminosity and optical flow detection. Tracking of the head is based on a local template matching of the tip of the nose. This feature has symmetry and convexity properties, that ensure robust tracking during the rotation of the head.




Tracking of gestures

HARDWARE SPECIFICATIONS

Robota is a mini-humanoid robot of 45cm, 1.5kg, with 5 DC motors and potentiometers. An ***iPAQ-3850 PocketPC*** (*StrongARM* Processor, 206MHz, 64Mb RAM) is fixed on the front of the robot, with a *FlyCam-CF* camera.

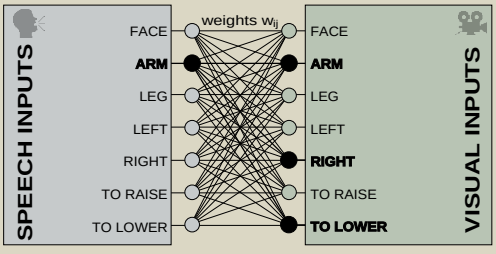


Robota is used in robotics classes at the undergraduate level, and in pilot studies with normal children and children with autism.

A. Billard. Robota: Clever Toy and Educational Tool. Robotics & Autonomous Systems, 2003

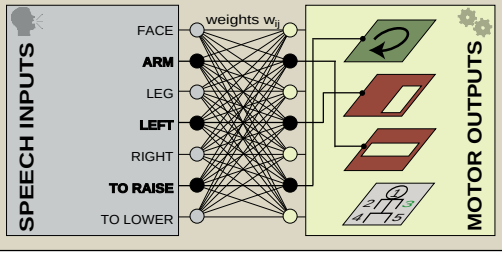
LEARNING MODULE

Key features extracted from the vision and speech modules are associated in a 1-layer Artificial Neural Network (ANN) through Hebbian learning. In this example, the user says *"This is your arm"* while lowering his left arm. Robota imitates the user by lowering its right arm (mirror imitation). The connection weights between co-activated inputs are increased.



RETRIEVAL MODULE

In the retrieval mode, speech inputs reactivate motor outputs. Based on the activity of the speech nodes, the ANN computes the activity of the motor nodes, following a winners-take-all mechanism.



FEEDBACK MODULE

- **Vocal feedback** is provided by the *Elan* speech synthesis engine. The robot repeats the keywords extracted by the speech module to acknowledge correct parsing.
- **Visual feedback** is given through the display of the user's image on the PDA screen, and by the robot imitating the user's gestures.

MOTORS COMMAND MODULE

Interprets the outputs of the ANN, and combines motor masks and basic motor commands, to produce the final motion.

"Raise left arm" produces one command to define the sense of rotation of the motor, one mask that activates only the motors controlling the left part of the body, and one mask that activates only the arms. The combination of the 3 outputs leads to a motor command that raises the left arm of the robot.