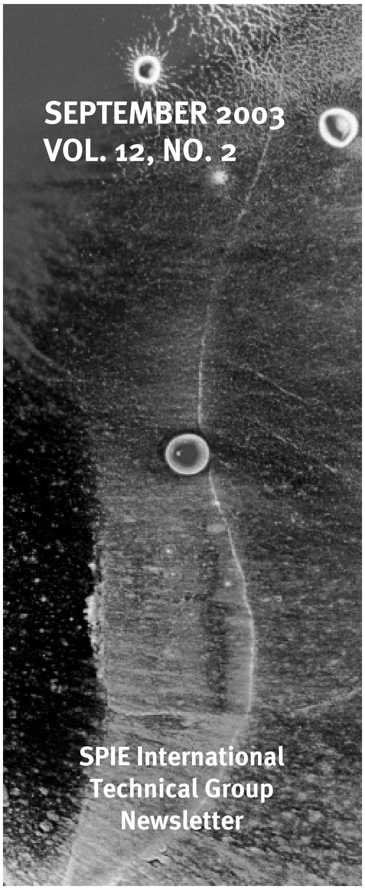


Robotics and Machine Perception

A robot that flies with a neuromorphic eye

Small unmanned air vehicles (UAV) and micro air vehicles (MAV) are being tested for urban surveillance. Future missions will require sensors and flight control systems (FCS) dedicated to obstacle avoidance and guidance so that the remote operators can concentrate on navigation and observation. Flying insects use the wide field of view (FOV) of their compound eyes to avoid obstacles and follow terrain. Insects use the retinal motion of contrasts also known as optic flow (OF): their nervous system fuses visual, inertial, and aerodynamic senses to control flight.

Our robotic aircraft demonstrates how insect vision can be applied to UAV flight control systems.[1] It also shows how biologically-inspired sensing can enable a flying test bed (Figures 1 and 2) to follow terrain and avoid obstacles in flight conditions that a operator using remote-control would find daunting.

Vision system

Although insects possess compound eyes, it is possible to design a camera eye that is equivalent for the analysis of OF. Our aircraft's camera eye contains a 20-pixel linear photoreceptor array and an aspheric lens (focal length 24mm) set at only

Continues on page 2.

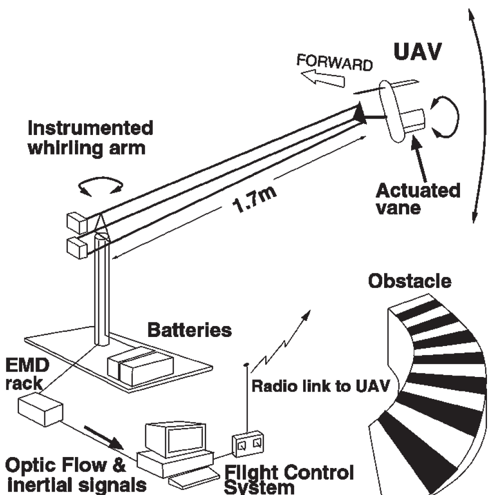


Figure 1. The aircraft flies at 2-3m/s and is tethered to an instrumented whirling arm. The flight control system uses two feedback loops: (1) flight speed is maintained with an inclinometer to sense aircraft pitch and by actuating the aerodynamic vane, (2) height above ground is maintained with an airborne photoreceptive array to sense optic flow and then commanding rotor thrust.

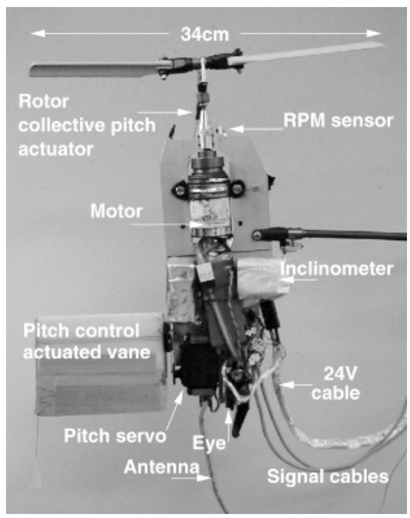


Figure 2. Front view of visually-guided rotorcraft (weight = 0.84 kg).

Calendar*—See page 11*

Technical Group Registration Form*—See page 10*

Newsletter now available on-line

Technical Group members are being offered the option of receiving the Robotics and Machine Perception Newsletter electronically. An e-mail is being sent to all group members with advice of the web location for this issue, and asking members to choose between the electronic and printed version for future issues. If you are a member and have not yet received this message, then SPIE does not have your correct e-mail address.

To receive future issues electronically, please send your e-mail address to:
`spie-membership@spie.org`
with the word `Robotics` in the subject line of the message and the words `electronic version` in the body of the message.

If you prefer to receive the newsletter in the printed format, but want to send your correct e-mail address for our database, include the words `print version preferred` in the body of your message.

Editorial

Welcome to this issue of the R&MP newsletter. The set of articles in this issue reflect many interrelated aspects of robotics and machine perception. There are four topics represented, namely visual skills, mapping and control, humanoid robots, and applications of robotics technology including automated highway systems, online museum robots, and smart meeting rooms.

The articles by Netter & Franceschini and by Bianco reflect low-level and high-level approaches to visual navigation. The former explores a very specific mechanism using neural network techniques, while the latter seeks a unified model, based on potential fields, for a range of visual skills including topological navigation, landmark learning and obstacle avoidance.

The articles by Liu and by Bernabeu also study navigation, and also contrast low and high-level aspects of robotic intelligence. Liu focuses on the construction of spatial representations based on temporal sequence processing networks (TSPNs), while Bernabeu discusses obstacle avoidance techniques for automated highway systems. The former emphasizes map-building, while the latter emphasizes more localized trajectory planning.

The issue contains three articles on humanoid robots. The first two, by Taylor & Kleeman and by Fitzpatrick, focus on perception of the robot's environment. The former describes a robust technique based on stereoscopic light-stripe scanning for the measurement of object features while the latter describes an active perception technique using manipulation to segment objects from their background. Both articles reflect the importance placed on developing techniques for representing and learning about novel objects in home and office environments. The article by Duffy, in contrast, takes a more general look at the issue of socially capable robots, not restricting the discussion necessarily to humanoid robots. A number of projects are described that illustrate novel perspectives on human-robot interactions.

The final two articles put humans squarely in the picture, but in very different ways. The article by Nait-Charif & McKenna describes the use of visual tracking techniques, incorporating Bayesian uncertainty models and particle filtering techniques, to track and annotate the behaviors of humans for application in smart meeting rooms. The article by Maeyama et al., in contrast, describes a museum robot that can interact with humans locally (visiting the museum) and remotely (via the Internet). These two articles, along with the article by Bernabeu on automated highway systems, represent the three application settings highlighted above.

There are a number of threads that can be drawn through the set of articles in this issue, illustrating important links and overlapping themes across many aspect of robotics and machine perception. We encourage you to read the articles and pursue the references for further details.

Gerard McKee
Technical Group Chair
The University of Reading, UK
E-mail: Gerard.McKee@Reading.ac.uk

A robot that flies with a neuromorphic eye

Continued from cover.

13mm from the array. Defocusing the retinal image reduces aliasing errors, improves OF measurement, and increases the FOV. The eye is tilted (-50˚) so that its FOV (75˚) covers the forward and downward region.

During flight, the velocities of contrasts in the retinal image are measured by an array of analog electronic Elementary Motion Detectors (EMDs). These are neuromorphic in that their circuits mimic the computation of biological neurons. Each EMD detects motion in a particular direction within the small part of the visual field seen by a pair of adjacent photoreceptors, and outputs a pulse whose voltage is nearly inversely proportional to the time delay between both photoreceptor excitations: i.e. it is quasi-proportional to speed. The EMD pulses are digitized and aggregated by a flight control computer.

Flight control using vision

The FCS regulates two things: flight at constant pitch, by commanding the aerodynamic vane to approximate flight at constant speed; and flight at constant OF, by commanding thrust to approximate flight at a constant height above the ground

The pitch is measured by an inclinometer and regulated using forward and backward flight PID (proportional integral derivative) regulators with bumpless transfer and anti-windup. The PID's parameters were adjusted by identifying the aircraft's pitch response to manually-commanded step inputs to the aerodynamic vane.

The aggregated OF value used to command thrust corresponds to a weighted average of the elementary retinal velocities computed by the EMDs. The weighing rule rectifies and fuses a reference OF distribution when the aircraft is flying at a preferred speed and altitude over flat terrain. The rule gives more weight to the forward FOV than to the downward FOV. This paradigm is inspired by the response fields and dendritic structures of frontal neurons VS1 and VS2 of the blowfly's vertical vision system.[2] During flight, the current OF is compared to the reference OF and the FCS modulates thrust to vary the height.

A Scilab simulation of the flight control paradigm showed that it can also be used to control aircraft landings by voluntarily decreasing flight speed while retaining the reference OF. Such a strategy can be compared to that of insects.[3]

Flight tests

The aircraft is tethered to a light whirling arm that carries visual, inertial, and tachymetric signal lines from the aircraft to the ground via slip-rings. The aircraft is powered at 24V, 8A, by two car batteries. The FCS runs on a ground-based PC with the Real-Time Linux operating system. The PC digitizes signals and generates the commands which are broadcast via remote-control.

The aircraft was piloted manually for system identification. After that, over 50 vision-guided terrain-following flights were demonstrated with the aircraft whirling four times over a contrasted 30° ramp whose peak is at 1.5m.

Thomas Netter[*] and Nicolas Franceschini[+]
[*]The Institute of Neuroinformatics
Zurich, Switzerland
E-mail: tnetter@ini.unizh.ch
http://www.ini.unizh.ch/~tnetter
[+]CNRS Biorobotics
31 Ch. J. Aiguier
13204 Marseille Cedex 20, France
E-mail: franceschini@laps.univ-mrs.fr

References

1. T. Netter and N. Franceschini, *A Robotic Aircraft that Follows Terrain Using a Neuromorphic Eye,* **Proc. IEEE/RSJ Int'l. Conf. on Intelligent Robots and Systems IROS'02,** paper #593, 2002.
2. H. G. Krapp, B. Hengstenberg, and R. Hengstenberg, *Dendritic structure and receptive-field organization of optic flow processing interneurons in the fly,* **J. Neurophysiology 79,** pp 1902-1917, 1998.
3. M. V. Srinivasan, S. W. Zhang, J. S. Chahl, E. Barth, and S. Venkatesh, *How honeybees make grazing landings on flat surfaces,* **Biol. Cyber. 83,** pp 171-183, 2000.

Temporal coding of spatial knowledge for mobile robot navigation

In navigation systems, it is straightforward to construct a spatial model of physical environments as a metric or topological map, then cyclically find the robot's location in the map and execute a planned path leading to the goal.[1] Although the bird's eye mapping strategy is prevalent in a number of applications—shipping, aviation, expedition and so on—it appears not to be the only mechanism underlying human beings' way-finding ability. Usually we have no GPS, compasses or range finders to localize ourselves. But we seldom ask, "What are my current coordinates?" or, "What is the next step in the planned path?" We select the right route, not reactively, but intuitively according to our intention and perception. How do humans quickly learn spatial knowledge, memorize it in neural cells, and recall it when necessary with neither a precise map nor Cartesian coordinates? Here we consider the problem of constructing and using internal spatial representations for mobile robot navigation in a connectionist way, and explores the mechanism behind biological route-learning behavior.

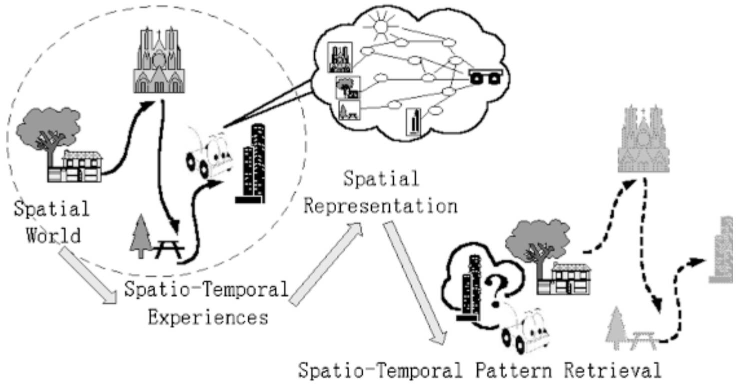


Figure 1. Construction and retrieval of an internal world representation.

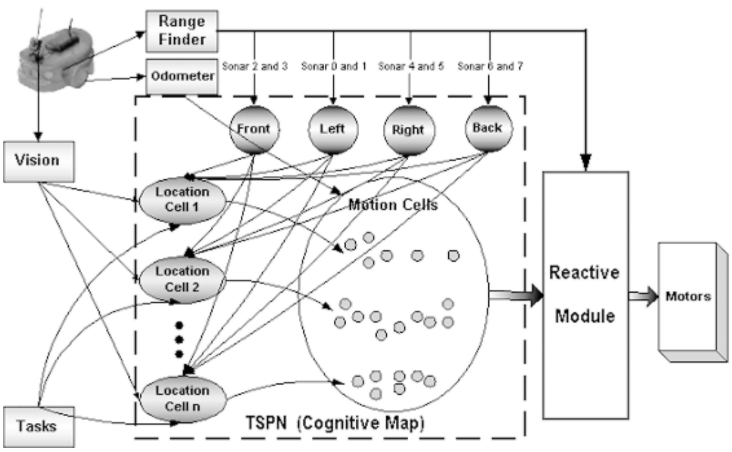


Figure 2. Sketch map of the navigation system.

Spatio-temporal transformation in navigation

Grounded on the fact that learning, recognizing and recalling temporal patterns contribute greatly to human intelligence, we conceive that robots may also learn spatial knowledge from the regularity of temporal sequences of sensory and action flows. When people walk through a given territory, the spatial structure of the world is transformed into spatio-temporal patterns that are perceived sequentially by our sensors and stored in short-term memory. For example, the duration of an action can reflect the distance or the change of headings.

In order to maintain these patterns in long-term memory, another transformation is involved to encode the spatio-temporal pattern into a spatial one, i.e. a neural network with cells and connections, like those in the grooves of a vinyl record. Most often they are dormant, but can be activated by internal desire and perception of external stimuli. The cells and strengths of various synapses predispose the system to produce various spatio-temporal patterns. A sequence of cell-firing patterns represents a specific route and its environmental contexts. Therefore, the construction and retrieval of an internal world representation can be interpreted as a *spatial world* → *spatio-temporal experiences* → *spatial representations* → *spatio-temporal patterns* procedure, as shown in Figure 1.

Temporal sequence processing network

What can influence our way-finding action? The proposed list contains environmental cues (sensor inputs), intention, instinctive behavior (innate obstacle avoidance and other safeguard actions), spatial knowledge (learned from self experience or other information, maintained in long-term memory as a cognitive map[2]) and short-term memory of recent action and perception.

An adequate connectionist model for "intuitive" navigation should have the ability to deal with spatio-temporal patterns and continuously integrate robot's sensation and action into an interrelated whole.

Temporal Sequence Processing Network (TSPN)[3] fulfills them via activity leakage, cell differentiation, and the postsynaptic-potential activation mechanism. It memorizes and correlates the robot's own spatial-temporal experiences, including its past and current sensory inputs and behaviors, and effectively retrieves them when exposed to similar stimuli in later runs. Unlike the models using place cells, the network itself is not a topological graph of the environment. In TSPN-based systems (see Figure 2), spatial information is implicitly coded in temporal characteristics of cells and connections that are incrementally constructed at run-time when the robot is exploring the environment. Neuron activations are assumed to be action decisions. Their execution is influenced by innate safeguard modules so that dangerous actions will be inhibited.

The navigation strategy is similar to that underlying our intuitive way-finding behaviors, which does not depend on maps and coordinates. The robot learns a goal-directed cognitive map from its own viewpoint. Different from most existing navigation systems, the system is irrational: its behaviors are not grounded on reason. However many can afterwards be broken down into component elements and their origins are brought into harmony with the laws of reason.[4] Action selection is an immediate decision based on sensation and memory. Similar phenomena happen when our intuition plays a more important role than reasoning in the decision-making process. Our idea of doing something presents itself whole and complete.

Conclusion

TSPN is a connectionist model for autonomous robots to learn spatial information from its own temporal experience. The routes between lo-

Continues on page 11.

Visual skills development in robotics: a unified view

Algorithms for developing vision-based behaviors for autonomous robots are well-documented in the technical literature. Apart from a few examples, the principles followed to introduce vision-based skills often make use of finely-tailored techniques with a very narrow working scope. Here we show how different vision-based skills—such as learning of distinctive features, visual guidance, topological navigation, obstacle avoidance, and visual navigation enhancement—share common principles based on the navigation vector field.

The computation of the vectorial output of a motion strategy over the whole environment defines a vector field. A key concept related to this vector field is that it is usually the gradient of a scalar potential function that states the convergence properties or robustness of a given path. This holds when the field is conservative. Unfortunately, vector fields produced by a real planner are barely conservative. Luckily, this represents one of the most important principles behind explaining many visual behaviors. Figure 1 shows examples of a vector field, a versor field (not considered in this paper), the degree of conservativeness of the vector field, and its potential function.

Visual landmark learning

Landmarks are widely used in mobile robots since diverse tasks can take advantage of distinctive features in the sensorial space. In Figure 2, landmarks are shown as box-shaped regions of the image. This image has been grabbed from a color camera mounted on the top of a Nomad 200 robot.

In the robotics literature, extraction of visual landmarks from the environment—apart from a few examples—is performed while considering a still picture of the goal, and not while the agent is moving. However, motion plays a hidden but important role in feature extraction: it influences the conservativeness of the navigation vector field, thus providing strong bases for further speculation about strategy (mainly convergence and robustness). To extract good landmarks an agent performs the well-known learning scheme of the bee's: *turn back and look*. Tests show how motion influences the conservativeness of the field.

Visual guidance

After reliable landmarks have been chosen, navigation information can be extracted from them. The underlying principle is that real movement is represented by an attraction force: the agent tries to restore the original position and size of every landmark. The data can then be fused together by weighted addition.

The physical principle that drives the agent is computed by integrating the navigation vector field, thus obtaining the potential function. In Figure 3, the actual trajectory followed by the agent is highlighted. The agent follows the gradient of the potential to get to the goal location (minimum potential). The actual path and its potential function profile followed by the agent are reported at the left of Figure 2 (at the bottom of the rectangle, and at its top, respectively).

Visual obstacle avoidance

A fundamental and recently-discovered principle is that moving obstacles generate a non-conservative wavefront that can be exploited to implement real-time practical obstacle-avoidance mechanisms. At run-time, conservativeness instabilities are discovered by calculating the variance of the navigation vector: when the variance is above the given threshold, the field is considered to be unstable and the robot assumes this variation as being generated by moving obstacles crossing the field of view of the camera. Thus the robot can act accordingly either stopping or moving elsewhere. The calculation of the navigation-vector variance is shown above the circle at the bottom-center of Figure 2. In this case, the variance reports that a moving obstacle is crossing the environment.

Visual topological navigation

A vector field provides information about the topology of an environment. Usually, when an agent exploits visual navigation, its working principle consist of comparing the actual view and the goal image to compute a navigation step and so reduce discrepancies between the images. Practically, two sets of vector fields are considered: the global-vector field produced by global landmarks (i.e. globally-relevant features viewed from every place within the environment), and the set of regional local-vector fields produced by the respective local features. Both sets allow for the computation of potential functions: the global potential function encodes region adjacency properties and reciprocal directions. The local potential functions allow for precise positioning of sub goals within specific regions.

Conclusion

Apparently different visual behaviors share common working principles. Vi-

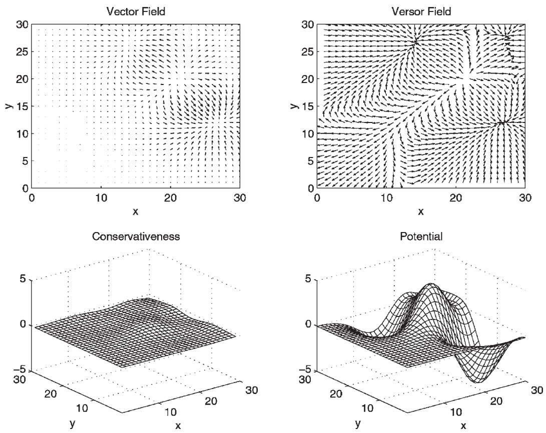


Figure 1. Features related to a vector field (top left): the versor field (top right), the conservativeness of the vector field (bottom left) and the scalar-potential function generating the field (bottom right).

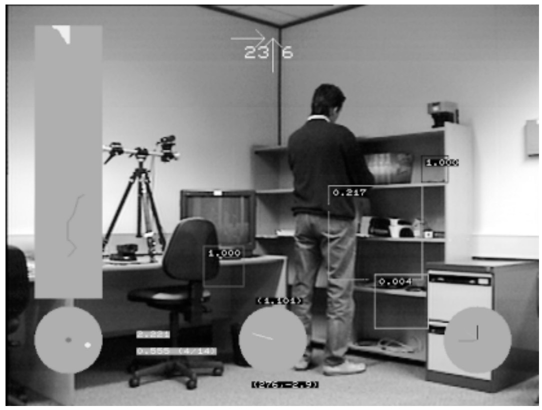


Figure 2. A frame captured along the agent's path.

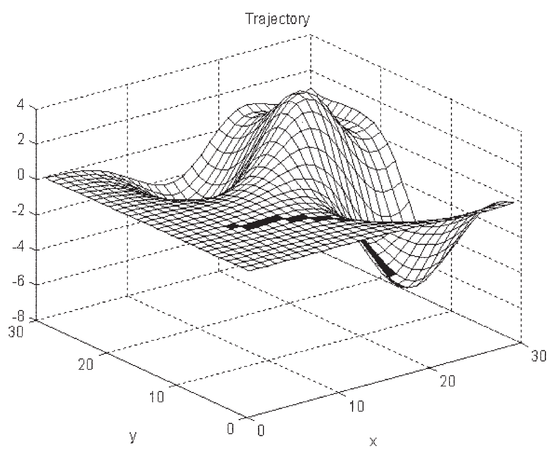


Figure 3. The potential function drives the agent to the goal. The actual trajectory followed by the agent's states is highlighted.

Continues on page 11.

Towards automatic driving of vehicles in automated highway systems

The generation of safe maneuvers for unmanned vehicles surrounded by other vehicles (obstacles) has been extensively studied.[4] The problem is considered to be a particular case of trajectory-planning algorithms for mobile robots because, apart from emergency situations, obstacle and unmanned vehicles are moving along the same direction and in a bounded space defined by a set of lanes. Kinematics constraints can be also avoided in the automatic-driving process, this despite the fact that automobiles have nonholonomic constraints, and taking speed into account. In fact, both lane-changing and keeping-the-same-lane maneuvers can be generated on this basis.

An important requirement is that maneuvers be computed as quickly as possible. For this reason, geometric-based trajectory-planning algorithms (which consider a geometric model of the involved vehicles) are the most appropriate.[3]

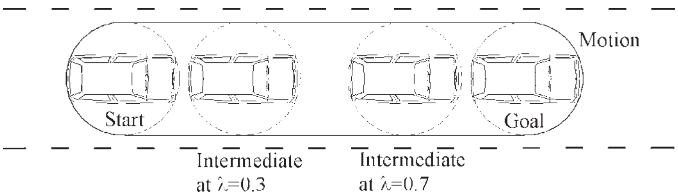


Figure 1. Motion representation. Start, goal, and intermediate configurations with λ=0.3 and λ=0.7 are shown.

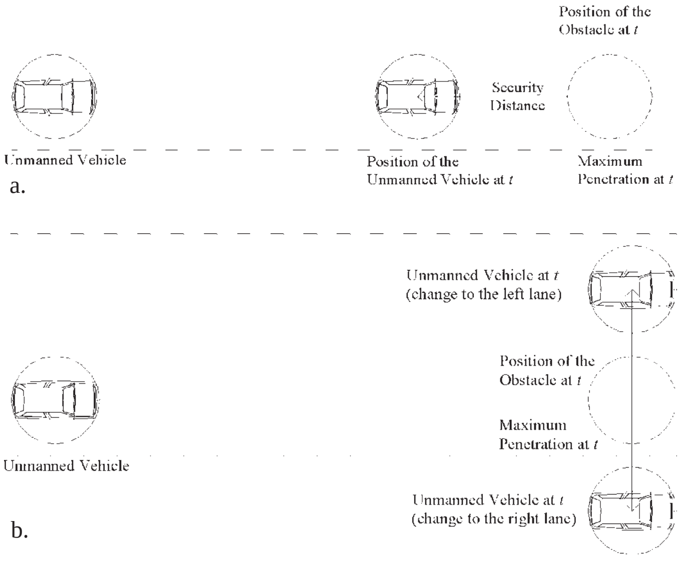


Figure 2. Maneuvers generated for avoiding a predicted collision. (a) braking. (b) lane changing.

Vehicle representation

Both vehicles and their motions are modelled by basic 2D spherically-extended polytopes (s-tope).[1] Essentially, all the infinite intermediate time positions, from the first to the final configuration of a given vehicle (enveloped by a circle, see Figure 1), are modelled by cylinders. Each one of these intermediate configurations obeys a linear function in $\lambda \in (0,1)$.

Collision avoidance

The unmanned vehicle has to be equipped with an onboard sensor system able to provide the current positions and speeds of neighboring vehicles. Each time this information is received, the motion of each obstacle vehicle is considered. Start time and position correspond to the current situation. Goal time and position are estimated by computing the position of the vehicle in a prediction horizon time Δt. This parameter is greater than the sampling rate T of the sensor system and it is chosen with reference to the safe distance between two vehicles in the same lane.

Next, the distance between the unmanned-vehicle motion and the obstacle motions is computed. The distance-computation algorithm returns a λ parameter, which characterizes a predicted collision (if one is likely), and the position and time of the maximum penetration.[1] In addition, it defines configurations that avoid this collision. This distance is computed without dividing the motions into different time intervals:[2] the algorithm is fast enough to be run at the same rate as the sensors. Consequently, it can be used to aid the maneuver planner.

Maneuver planner

When a collision is predicted between the motions of the unmanned vehicle and the leading vehicle in same lane, braking and double-lane-changing maneuvers are generated. In a general sense, these maneuvers imply a deceleration and acceleration respectively, so if a maneuver does not verify the dynamic constraints of the unmanned vehicle it will be rejected.[3]

When such a collision is predicted, the position and time of the future maximum penetration between both vehicles is characterized by parameter λ. If λ is greater than one, it means the collision will take place later than Δt, and consequently no changes are applied to the unmanned-vehicle actuators yet. Otherwise, a braking maneuver is generated by translating the position of the maximum penetration backwards until a safe distance is confirmed. In other words, the unmanned vehicle has to be located at this new position at the time determined by λ. A graphical example is shown in Figure 2a.

Additionally, two-lane-changing (right and left) maneuvers can be generated. These are defined by translating the position of the maximum penetration from the current lane to the center lines of its right and left lanes (see Figure 2b). Consequently, the unmanned vehicle is forced to be at one of these positions at the time determined by λ. Before starting a lane-changing maneuver, a collision test (between the motion associated with the maneuver at issue and the motions of vehicles in the target lane) has to be applied.

This maneuver planner has been implemented in C and run on Pentium-III 700Mhz. Under situations with multiple obstacles, the computational time has always been lower than 1ms.

Enrique J. Bernabeu
Department of System Engineering and Control
Universidad Politecnica de Valencia, Spain
Tel: +34 96 387 9572
Fax: +34 96 387 9579
E-mail: ebernabe@isa.upv.es

References

1. E. J. Bernabeu and J. Tornero, *Hough transform for distance computation and collision avoidance*, **IEEE Trans. on Robotics and Automation 18** (3), pp. 393-398, June 2002.
2. E. J. Bernabeu, J. Tornero, and M. Tomizuka, *Collision Prediction and Avoidance Amidst Moving Objects for Trajectory Planning Applications,* **Proc. of the 2001 IEEE Int'l Conf. on Robotics and Automation,** pp. 3801-3806, 2001.
3. I. Papadimitriou and M. Tomizuka, *Fast lane changing computations using polynomials,* **Proc. of American Control Conf.,** pp. 48-53, 2003.
4. M. Tomizuka, *Automated highway systems. an intelligent transportation system for the next century,* **IEEE/ASME Int'l Conf. on Advanced Mechatronics,** pp. 16-20, 1997.

Visual sensing for perception and control in robotic applications

Humanoid robots are natural vehicles for human-machine cooperation in office/domestic environments. Our research aims to explore the building blocks required for a humanoid robot to perform interactive manipulation tasks (grasping and placing objects) in this type of unstructured setting. We have focused our attention on visual perception because it makes important contributions to the areas of task specification, planning and actuation, gesture recognition, object modelling/localization, and visual-feedback control of robotic limbs.

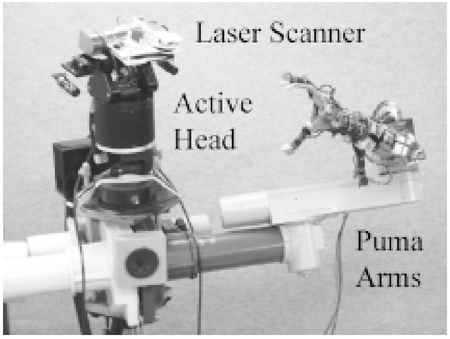
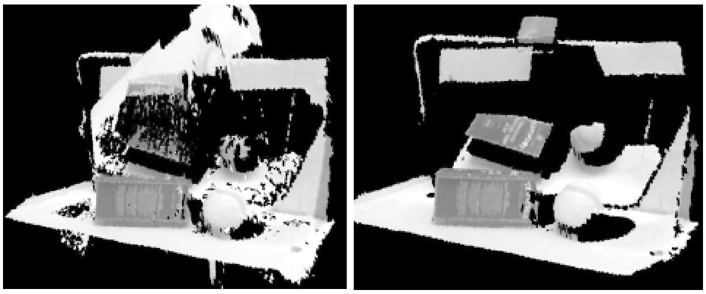


Figure 1. An experimental, upper-torso humanoid robot.

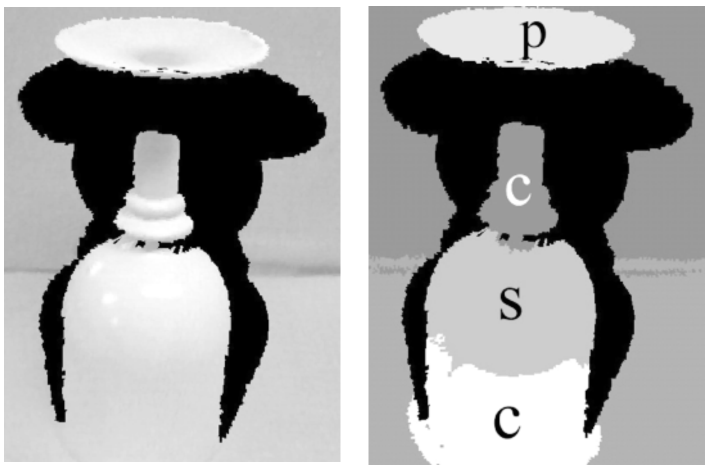
Figure 1 shows our experimental upper-torso humanoid platform. The arms consist of two 6-DOF (degrees of freedom) Puma 260 robots with 1-DOF prosthetic hands. Vision is provided by stereo cameras on a Biclops robotic head. A vertical laser stripe generator is mounted above the cameras and consists of a 5mW laser diode with cylindrical lens, and a DC motor/encoder to scan the stripe across the workspace.



a. b.

Figure 2. 3D scan of objects and mirror: (a) conventional approach (no noise rejection); (b) with robust noise rejection.

Early work concentrated on developing a position-based visual servoing scheme that provides on-line hand-eye calibration to allow the hands to be accurately positioned relative to a target.[1] Red LEDs (light-emitting diodes) are attached to the robot to provide robust features for visual sensing, and the pose of the hand is reconstructed and tracked using a Kalman-filter framework. Recent work has focused on robust measurement of the target objects.



a. b.

Figure 3. (a) Raw 3D scan of an inverted wine goblet. (b) Segmentation into (s)pherical, (c)ylindrical and (p)lanar components.

Stereoscopic light-stripe scanning

For maximum flexibility in performing *ad hoc* tasks, the robot should be capable of autonomously locating and classifying *a priori* unknown objects. Object modelling requires the acquisition of dense and reliable color/range measurements that passive stereopsis cannot offer. Light-stripe ranging is a suitable alternative, but conventional scanners do not distinguish the stripe from secondary reflections and crosstalk, making them unsuitable for robotic applications. Furthermore, robust stripe scanners proposed in previous work suffer from issues including assumed scene structure and lack of error recovery.

The scanner we have developed uses two cameras to measure the stripe, and exploits redundancy to disambiguate it from noisy measurements.[2] Each stereo image is processed using edge filters to determine candidate stripe locations. On each scan line, the pair of candidates corresponding to the actual stripe are identified by minimizing an error function derived from the expected relationship between valid stereo measurements given the camera parameters and light plane position. After each complete scan, a color image is captured and implicitly registered with the range data. Our validation and reconstruction algorithms also provide a framework for simple self-calibration using measurements of an arbitrary non-planar target.

Figure 2 demonstrates the robustness of our stereoscopic stripe scanner. A mirror is placed behind the objects to create a reflection of the laser stripe and simulate the effect of crosstalk and specularities. Figure 2(a) shows the range data measured using a conventional single-camera scanner, while our robust scan is shown in Figure 2(b). The inability of the conventional scanner to distinguish the laser from its reflection results in phantom measurements, while our method provides dense, reliable range data.

Range data segmentation and classification

Once the measurements are acquired, the robot must localize objects of interest. Identifying *a priori* unknown objects is particularly challenging when instances of the same type vary in size and shape. We overcome this problem by representing objects using data-driven geometric primitives, which are sufficient to describe many common domestic objects (cups, bowls, boxes etc).

Range-data segmentation is based on the notion that geometric primitives fit more robustly to large segments rather than small patches. Thus, we discard segmentation by aggregation in preference to a split-and-merge approach that attempts to maintain large segments. Depth data is first split at discontinuities and creases, and further splitting at changes in local surface type[3] occurs only if the initial segments cannot be accurately modelled. Surface-type classification is typically based on local curvatures calculated by fitting analytic curves to the range data, but the result usually depends on the arbitrary selection of approximating

Continues on page 11.

Computer vision for smart meeting rooms

There are many situations in which it is useful to automatically detect, track, and interpret people's activity using computer vision. Applications include surveillance, monitoring, smart rooms, low-cost motion capture, human-computer interaction, and gesture-control interfaces. The Computer Vision and Imaging Group at the Division of Applied Computing, University of Dundee, has been conducting research into vision systems for these and related application areas. Here we describe a system we developed for automatically annotating the activities of participants in a meeting. The system detects and simultaneously tracks multiple people, handles person-person occlusion, and combines data from two wall-mounted cameras on opposite sides of the meeting room to annotate the activity of all the participants throughout a meeting. The video data sets (see Figure 1) were provided by project FGnet (IST-2000-26434) for one of a series of IEEE workshops on Performance Evaluation of Tracking and Surveillance (ICVS-PETS).[1] We originally developed the core tracking algorithms used here for a different application: that of monitoring older people living alone in order to detect falls, with the aim of helping them to maintain their independence for longer.[2]

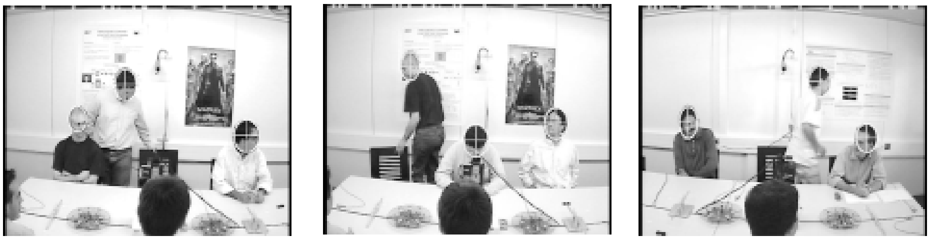


Figure 1. Three images from a long meeting sequence in which six participants enter, sit down and get up several times to use the white-board, before finally exiting the room.

Head tracking

Visual tracking is often formulated from a Bayesian perspective as a problem of estimating some degree of belief in the state of an object at each time step given a sequence of observations. Here we adopt such an approach using a likelihood model based on region (color) and boundary (edge) information. A person's head shape in an image is reasonably well approximated as elliptical, irrespective of 3D pose. The likelihood model combines intensity gradient information along an ellipse boundary with a color model of the ellipse's interior region. The color-based measurement is obtained by computing the divergence of a color histogram of the ellipse's interior from a stored model histogram. The gradient-based measurement involves searching for maximum-gradient magnitude points along short radial search-line segments centered on the ellipse boundary.

Particle (non-parametric) filtering has become popular in the computer vision community for tracking since it can cope well with visual clutter by propagating multi-modal probability densities over time. We modified the frequently used *condensation* algorithm[3] to more effectively explore the search space induced by our likelihood model. This modification, called *iterated likelihood weighting*, outperformed condensation using our model, particularly when the dynamic model was poor: often the case with human motion.

Tracker initialization and occlusion handling made use of scene-specific context. The room layout and the maximum and minimum height of a person meant that the heads of people on the far side of the table always appeared between upper and lower limits. Furthermore, when people pass in front of the camera on the opposite wall, they occlude the view of the people on the near side of the table from that opposing camera. When such an occlusion event is detected, any tracks in the corresponding regions of the opposing camera's field of view are suspended until the occlusion is over. Provided that the occluded people do not move too much while occluded, their trackers will recover and continue to track them. A background-subtraction algorithm was applied in each frame within the scene-entrance regions. Whenever significant change was detected, an initial particle set of head ellipses was instantiated—centered within the region—and a tracker was initialized. When a tracker's estimated head ellipse left the field of view in the direction of the white-board, that tracker waited for the background subtraction routine to signal re-entry. When a tracker's estimated head ellipse left the view in the direction of the exit, the tracker was terminated.

Figure 2. The automatically-extracted head trajectories for three of the participants overlaid on the empty meeting room.

Tracking results and activity recognition

All meeting participants were successfully tracked through long image sequences with automatic initialization and termination of tracking. They were tracked through occlusion using views from two different cameras. Figure 2 shows the extracted trajectories of the centers of three of the participants' heads for an entire meeting sequence.

Given the reliable head tracking just described, along with some simple scene-specific constraints, recognition of several actions in the meeting became straightforward. These actions were entering, exiting, going to the white-board, getting up and sitting down. All such actions were detected without false detections. The first three can be recognised by detecting where trackers initialize and terminate. Standing up and sitting down were detected using rules based on the head crossing the horizontal lines shown in Figure 2.

Hammadi Nait-Charif and Stephen J. McKenna
Division of Applied Computing
University of Dundee
Dundee DD1 4HN, Scotland, UK
E-mail: {hammadi, stephen}@computing.dundee.ac.uk
http://www.computing.dundee.ac.uk/projects/vision/

References

1. H. Nait-Charif and S.J. McKenna, *Head Tracking and Action Recognition in a Smart Meeting Room*, **4th IEEE Int'l Workshop on Performance Evaluation of Tracking and Surveillance (ICVS-PETS),** pp. 24-31, 2003.
2. F. Marquis-Faulkes and S.J. McKenna, P. Gregor, A. F. Newell, *Scenario-based Drama as a Tool for Investigating User Requirements with Application to Home Monitoring for Elderly People*, **Proc. HCI Int'l 3,** pp. 512-516, 2003.
3. M. Isard and A. Blake, *CONDENSATION: Conditional Density Propagation for Visual Tracking*, **Int'l J. of Computer Vision 29** (1), pp. 5-28, 1998.

Mobile robots in a remote art museum

At the University of Tsukuba, we implemented a five-year research plan (April 1997 to March 2002) towards modeling the evaluation structure of Kansei (a Japanese concept that has to do with the psychological, emotional, and aesthetic reaction that we have to objects).[1] To do this, we analyzed the how people appreciated works of art. What was unique about this project was that a mobile robot was used to act as an eye for those far from the museum. These observers could choose what to look at by using remote control to move the TV camera on the robot. Our goal is shown in Figure 1: a remote viewing system that enables ordinary people at home or in the office to view works of art remotely in a museum by manipulating the vision of the robot using an ordinary personal computer connected to the Internet. In this project, we initially used the system as an experimental tool, interpreting the positions and postures of the robot as an avatar of the viewer. Through the computer images from the robot we were able to gain insight into the behavior of remote visitors.

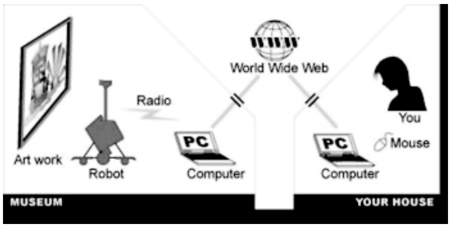


Figure 1. Basic concept of remote art appreciation via the Internet.

Figure 2. Graphical user interface for remote art appreciation shown on a web browser.


Figure 3. The mobile robotic avatar Kapros in the Tsukuba art museum.

Mobile robotic avatar

Our avatar does not have to be perfectly humanoid in shape: instead, it is equipped with limited functions that substitute for human eyes and feet. Though humans have stereo color vision a single color TV camera is sufficient since the images will only be displayed on a conventional computer monitor. As for movement, though human beings are bipedal, we chose a wheel drive mechanism for better moving efficiency on the flat museum floors. However, since the avatar must co-exist with actual humans in the museum, it was important that the speed of the avatar be the same as normal walking speed. Likewise, the avatar must be able to pass through corridors in the same way as ordinary visitors. We assumed the robot would be viewing ordinary-sized pictures exhibited at 140 cm from the floor, and set up a TV camera on the avatar at that height. We call our robotic avatar Kapros.

Tele-driving

The most difficult problem associated with the tele-operation of a mobile robot via the Internet is the randomness of time delays that occur while issuing operational instructions. For this reason, the authors have designed the remote-control commands about movement of the robot by sending the specification of a sub goal as a global coordinate inside the museum site. Once determining the unique absolute position of this sub goal, this information does not change with time.

All commands are sent by CGI (common gateway interface) scripts inside a Java program for firewall-free transmission of control commands. The GUI (graphical user interface), as shown in Figure 2, is roughly divided into a section for movement and another for TV-camera operation. The GUI must be intuitive for beginners. To define a movement, the user clicks the required destination on the floor of the live image. This allows easy and intuitive operation by non-specialists as well as interactive operation. As live images for movement, we use panoramic images with a -90˚ to +90˚ view angle generated from an omni-view sensor: normal images do not have enough angle of view for safe driving. Though the panoramic image has distortion, it does not cause serious problem, since only the destination point is clicked for movement.

To cope with the random time delay, some motions must be performed through autonomous operation of the robot. In our system, we use the position-based tele-operation scheme: the motion control related to dynamics, such as acceleration and velocity, is autonomously controlled by the robot itself. The trajectory from its current position to the destination, as requested by the remote visitor, is automatically calculated. The robot's wheels are controlled to follow the designed trajectory using feedback of the estimated position every short time interval. The robot also performs autonomous obstacle avoidance.

Trials at the Tsukuba art museum

We have carried out many experiments on seven exhibitions of art and design at the Tsukuba art museum since February of 2000. Figure 3 was taken at one of our first experiments: at the graduation exhibition for students of the University of Tsukuba Art School. In November 2000, we also demonstrated the remote art appreciation from the IROS'00 conference site during our presentation.[2] We have looked at 3D art works as well as pictures: at the exhibition of statues and sketches by Kunihiko Isski in 2001, for example. Consequently, we can confirm the technical feasibility of our system and acquire data about the appreciation attitude of remote visitors.

However, some problems remain: the slow update cycle of live images, the fact that only one viewer can be in control at a time, and the avatar disappearing from the viewer's GUI. These issues cause inconvenience for remote visitors. In the future we intend to extend our system to allow the remote viewing of 3D art works and, more importantly, to have multiple mobile robots operating on the same floor. At this point we will more-fully realize this new application of robotics and IT.

Shoichi Maeyama,* Shin'ichi Yuta†, and Akira Harada†
*Osaka Electro-Communication University
Neyagawa Osaka 572-8530, Japan
E-mail: maeyama@isc.osakac.ac.jp
†University of Tsukuba
Tsukuba Ibaraki 305-8573, Japan
E-mail: yuta@roboken.esys.tsukuba.ac.jp

Reference
1. A. Harada, *Modeling the Evaluation Structure of KANSEI Using Networked Robot*, **Proc.Int'l Conf. on Advanced Intelligent Mechatronics**, pp. 48-55, 1997.
2. S. Maeyama, S. Yuta, and A. Harada, *Experiments on a Remote Appreciation Robot in an Art Museum*, **Proc. 2000 IEEE/RSJ Int'l Conf. on Intelligent Robot and Systems**, pp.1008-1013, 2000.

Open object recognition for humanoid robots

Robots must be able to adapt gracefully to frequent and dramatic changes in their workspace if they are to operate successfully in human-centered environments, as opposed to controlled industrial settings. At the MIT Humanoid Robotics Group, we are developing methods that permit our robots to deduce the structure of novel activities, adopt the vocabulary appropriate for communication about the task at hand, and learn about the appearance and behavior of unfamiliar objects. This latter ability is discussed here. The humanoid robot Cog[1] uses active exploration to resolve visual ambiguity in its workspace.[2] As Cog accumulates experience, it clusters episodes of object interaction to learn the appearance and properties of novel, unfamiliar objects. This process is called open object recognition.[3] An operator can then introduce names for objects to facilitate further task-related communication.

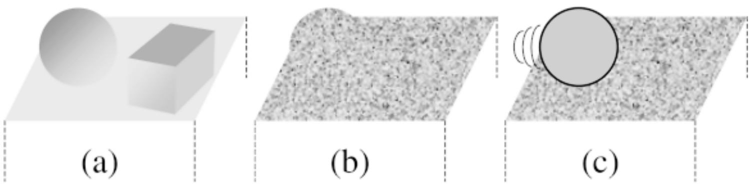


Figure 1. Cartoon motivation for active segmentation. Human vision is excellent at figure/ground separation (top left), but machine vision is not (center). Coherent motion is a powerful cue (right) and the robot can invoke it by simply reaching out and poking around.

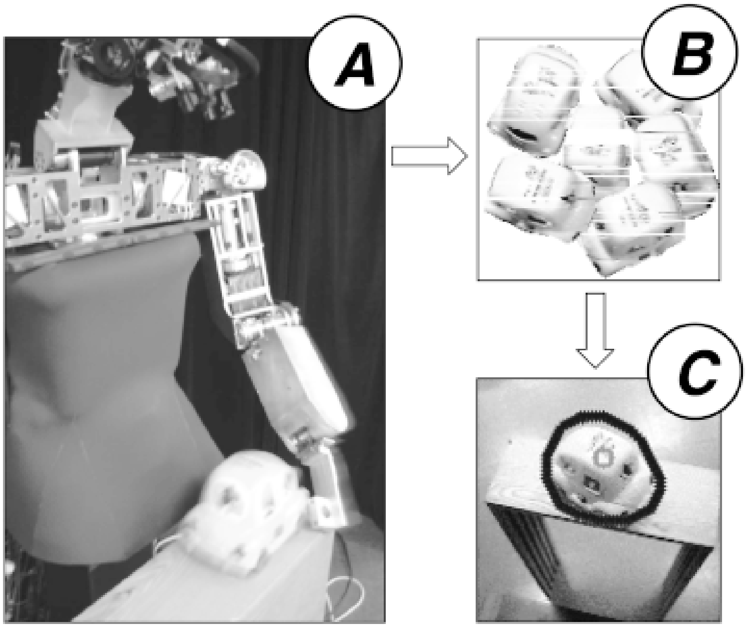


Figure 2. Object boundaries are not always easy to detect visually. The robot Cog (A) solves this by sweeping its arm through areas of ambiguity. If object motion results, the motion helps distinguish the object from its background (B). As the robot gains experience and becomes familiar with the appearance of an object, it learns to recognize and segment that object without further contact (C).

Figure/ground separation is a long-standing problem in computer vision, due to the fundamental ambiguities involved in interpreting the 2D projection of a 3D world. Cog can bypass this philosophical and practical dilemma by physical experimentation (see Figure 1). Cog has a 'poking' behavior that prompts it to select locations in its environment that may contain an object of interest, and sweep through them with its arm.[2] If an object is within the area swept, then the motion generated by the impact of the arm can be used to segment the object from its background, and obtain a reasonable estimate of its boundary. This is called active segmentation, and is a form of active perception.[4] Once Cog can reliably segment objects, then it learns about their appearance and how they move. Of course, active segmentation does not work for all objects—if an object is very small or very large, the procedure is likely to fail. But manipulable objects are, almost by definition, on the right scale for the method to work, and this is a particularly important class of object for robots.

Open object recognition is the ability to recognize a flexible set of objects, where new objects can be introduced at any time.[3] Cog can learn autonomously to recognize new objects by interacting with them (see Figure 2). Conventional object recognition systems do not need to be open: for example, the set of objects an industrial robot needs to interact with is likely to be fixed. But a humanoid robot in an unconstrained environment could be presented with just about anything, and trying to collect and train for all the possible objects the robot might encounter is simply not practical. Active segmentation gives Cog the ability to collect its own training data for machine learning. A variant of geometric hashing is used for object localization, with clustering of object models occurring both on- and off-line. The online clustering procedure is fast and responsive (on the order of seconds), but relatively coarse. The off-line clustering procedure is slower (on the order of tens of minutes), but can make subtler distinctions between objects. Both clustering methods are integrated so that the robot can distinguish visually distinctive objects quickly and more difficult cases over time.

The methods touched upon here allow our humanoid robot Cog to build up and maintain a perceptual system for object localization, segmentation, and recognition, starting from very little. Beyond this, Cog can track known objects to learn about activities they occur in, such as a sorting task or object search.[3] The overall goal of this effort is to develop a perceptual system for a humanoid robot that is as general-purpose and adaptable as the robot's physical form.

Paul Fitzpatrick
Humanoid Robotics Group
Computer Science and Artificial Intelligence Laboratory
Room 936, 200 Technology Square, Cambridge MA 02139
E-mail: paulfitz@ai.mit.edu
Tel: +1 617 253-6849
Fax: +1 617 253-0039
http://www.ai.mit.edu/people/paulfitz/

References

1. R. A. Brooks, C. Breazeal, M. Marjanovic, and B. Scassellati, *The Cog project: building a humanoid robot,* **Lecture Notes in Artificial Intelligence 1562,** C. Nehaniv (ed)., pp. 52-87, Springer, 1999.
2. P. Fitzpatrick and G. Metta, *Grounding vision through experimental manipulation,* **Philosophical Trans. of the Royal Society: Mathematical, Physical, and Engineering Sciences,** *in press.*
3. P. Fitzpatrick, *From first contact to close encounters: a developmentally deep perceptual system for a humanoid robot,* **Ph.D. Thesis and Technical Report AITR-2003-008,** MIT Department of Electrical and Computer Engineering, 2003.
4. R. Bajcsy, *Active perception,* **Proc. IEEE 76** (8), pp. 966-1005, August 1988.

Join the Technical Group

...and receive this newsletter

Membership Application

Please Print ☐ Prof. ☐ Dr. ☐ Mr. ☐ Miss ☐ Mrs. ☐ Ms.

First Name, Middle Initial, Last Name ______________________________

Position ______________________ SPIE Member Number ______________________

Business Affiliation ______________________________

Dept./Bldg./Mail Stop/etc. ______________________________

Street Address or P.O. Box ______________________________

City/State ______________________ Zip/Postal Code ____________ Country ____________

Telephone ______________________ Telefax ______________________

E-mail Address ______________________________

Technical Group Membership fee is $30/year, or $15/year for full SPIE members.

❑ Robotics & Machine Perception
Total amount enclosed for Technical Group membership $ ____________

❑ **Check enclosed.** Payment in U.S. dollars (by draft on a U.S. bank, or international money order) is required. Do not send currency. Transfers from banks must include a copy of the transfer order.

❑ **Charge to my:** ❑ VISA ❑ MasterCard ❑ American Express ❑ Diners Club ❑ Discover

Account # ______________________ Expiration date ______________________

Signature ______________________________
(required for credit card orders)

This newsletter is printed as a benefit of the **Robotics & Machine Perception Technical Group.** Membership allows you to communicate and network with colleagues worldwide.

As well as a semi-annual copy of the *Robotics & Machine Perception* newsletter, benefits include SPIE's monthly publication, **oe***magazine*, and a membership directory.

SPIE members are invited to join for the reduced fee of $15. If you are not a member of SPIE, the annual membership fee of $30 will cover all technical group membership services. For complete information and an application form, contact SPIE.

Send this form (or photocopy) to:
SPIE • P.O. Box 10
Bellingham, WA 98227-0010 USA
Tel: +1 360 676 3290
Fax: +1 360 647 1445
E-mail: spie@spie.org

http://www.spie.org/info/robotics

Please send me
☐ Information about full SPIE membership
☐ Information about other SPIE technical groups
☐ FREE technical publications catalog

Reference Code: 3537

ROBOTICSONLINE

Robotics Web Discussion Forum

You are invited to participate in SPIE's online discussion forum on Robotics & Machine Perception. To post a message, log in to create a user account. For options see "subscribe to this forum."

You'll find our forums well-designed and easy to use, with many helpful features such as automated email notifications, easy-to-follow 'threads,' and searchability. There is a full FAQ for more details on how to use the forums.

Main link to the new Robotics & Machine Perception forum: http://spie.org/app/forums/tech/

Related questions or suggestions can be sent to forums@spie.org.

The International Society for Optical Engineering

Robotics & Machine Perception

This newsletter is published semi-annually by SPIE—The International Society for Optical Engineering, for its International Technical Group on Robotics & Machine Perception.

Editor and Technical Group Chair Gerard McKee
Technical Editor Sunny Bains
Managing Editor/Graphics Linda DeLano

Articles in this newsletter do not necessarily constitute endorsement or the opinions of the editors or SPIE. Advertising and copy are subject to acceptance by the editors.

SPIE is an international technical society dedicated to advancing engineering, scientific, and commercial applications of optical, photonic, imaging, electronic, and optoelectronic technologies. Its members are engineers, scientists, and users interested in the development and reduction to practice of these technologies. SPIE provides the means for communicating new developments and applications information to the engineering, scientific, and user communities through its publications, symposia, education programs, and online electronic information services.

SPIE—The International Society for Optical Engineering, P.O. Box 10, Bellingham, WA 98227-0010 USA. Tel: +1 360 676 3290. Fax: +1 360 647 1445.

European Office: Karin Burger, Manager, karin@spieeurope.org, Tel: +44 7974 214542. Fax: +44 29 2040 4873.

In Russia/FSU: 12, Mokhovaja str., 119019, Moscow, Russia • Tel/Fax: +7 095 202 1079 E-mail: edmund.spierus@relcom.ru

Temporal coding of spatial knowledge

Continued from page 3.

cations are stored as the temporal characteristics of cell firing. The TSPN-based method is not proposed as a substitute to a rational navigation strategy, but as complementary to it. From the perspective of navigation, it cannot surpass classical systems using coordinates and maps with their precise localization, path planning, and tracking. However, this explanation of biological way-finding mechanisms may be treated as a primitive form of procedural/episodic memory: navigation is not its only application. With the expansion of sensory inputs and motor capacities, the same mechanism can be applied to other tasks such as language grounding: constructing internal representations of words and sentences.

Juan Liu
College of Information Science & Engineering
Central South University
Changsha, Hunan, 410083
China
E-mail: ljcic@263.net

References

1. S. Thrun, *Robotic mapping: a survey,* **Exploring Artificial Intelligence in the New Millennium,** Gerhard Lakemeyer, Bernhard Nebel (eds.), pp. 1-35, Morgan Kaufmann Publishers, 2002.
2. E. C. Tolman, *Cognitive maps in rats and men,* **The Psychological Review 55,** pp. 189-208, 1948.
3. J. Liu et al, *A Connectionist Model for Localization and Route Learning Based on Remembrance of Perception and Action,* **Proc. of IEEE/RSJ IROS-02,** pp. 655-660, Lausanne, Switzerland, Sept.30–Oct. 4, 2002.
4. C. G. Jung, **Psychological types**, pp.453-454, Princeton University Press, N.J. Princeton, 1971.

Calendar

2003 SPIE International Symposium Photonics East
27-30 October
Providence, Rhode Island USA

- **Photonics Technologies for Automation and Manufacturing**
- ***Optical Technologies for Industrial, Environmental, and Biological Sensing***

Advance Registration Ends 17 October 2003
http://spie.org/conferences/programs/03/pe/

2004 SPIE Defense and Security Symposium *formerly* AeroSense
12-16 April
Orlando, Florida USA
Abstracts Due 15 September 2003
http://spie.org/conferences/calls/04/or/

For more information on these events go to **www.spie.org** and go to the Meetings Calendar.

Visual skills development in robotics: a unified view

Continued from page 4.

sual skills, such as learning, obstacle-avoidance or sub-goal placement—and many others—can take advantage of the described mechanism. For example, one can argue about the efficacy of a landmark learning schema by evaluating the degree of conservativeness of the vector field the landmarks produce. All the behaviors described have been implemented in real robots.

Giovanni Bianco
Information Science Service
University of Verona, Italy
E-mail: giovanni.bianco@univr.it

References

1. M. V. Srinivasan and S. Venkatesh, **From Living Eyes to Seeing Machine,** Oxford University Press, Oxford, New York, Tokyo, 1997.
2. M. Lehrer and G. Bianco, *The Turn-back-and-look Behaviour: Bee versus Robot,* **Biological Cybernetics 83,** pp. 211-229, 2000.
3. G. Bianco and A. Zelinsky, *The convergence property of goal-based visual navigation,* **IEEE/RSJ Int' Conf.on Intelligent Robots and Systems (IROS 2002),** Switzerland, September 2002.

The human and the machine

Continued from page 12.

through simply engaging people in social interaction with an artificial entity? While arguments have prevailed for many years over the nature of intelligence and whether it can be realized in a machine, this work aims to demonstrate the power of perceived intelligence and people's willingness to interpret a social robot's interactions according to human-like social references. The key issue becomes a balance between function and form.

Researchers at the Anthropos Project at Media Lab Europe are: Brian Duffy, John Bradley, John Bourke, Eva Jacobus, Alan Martin, and Bianca Schoen.

Dr. Brian R. Duffy
Media Lab Europe
Sugar House Lane, Bellevue
Dublin 8, Ireland
Tel: +353 1 474 2823
E-mail: brd@media.mit.edu

Visual sensing for perception and control in robotic applications

Continued from page 6.

functions. We avoid this problem by developing a novel non-parametric surface type classifier based on analysis of the Gaussian image and surface convexity. A final merging step compensates for over-segmentation. The extracted primitives can then be used for classification, tracking, and task planning.

Figure 3 shows the result of our segmentation algorithm applied to a compound surface (an inverted wine goblet). Additional stereoscopic light stripe/segmentation experiments have been performed on a variety of scenes with objects such as bowls, bottles, and funnels, and the results can be viewed online.[4]

Geoffrey Taylor and Lindsay Kleeman
Intelligent Robotics Research Centre
Department of Electrical and Computer Systems Engineering
Monash University, Clayton 3800
Victoria, Australia
E-mail: {Geoffrey.Taylor, Lindsay.Kleeman}@eng.monash.edu.au
http://www.ecse.monash.edu.au/centres/IRRC/

References

1. G. Taylor and L. Kleeman, *Flexible self-calibrated visual servoing for a humanoid robot,* **Proc. Australian Conf. on Robotics and Automation**, pp. 79-84, November 2001.
2. G. Taylor, L. Kleeman, and A. Wernersson, *Robust colour and range sensing for robotic applications using a stereoscopic light stripe scanner,* **Proc. 2002 IEEE/RSJ Int'l Conf. on Intelligent Robotics and Systems**, pp. 178-183, 2002.
3. R. M. Haralick, L. T. Watson, and T. J. Laffey, *The topographic primal sketch,* **The Int'l Journal of Robotics Research 2** (1), pp. 50-72, Spring 1983.
4. http://www.irrc.monash.edu.au/laserscans

Contributing to the RM&P Newsletter

News, Events, and Articles

Contact our technical editor Sunny Bains (sunny@spie.org) with ideas for articles you would like to write or read.
For full information on article submission go to: http://www.sunnybains.com/newslet.html

Special Issues

Proposals for Special Issues are welcomed and should include:

- A brief summary of the proposed topical area.
- The reasons why it is of current interest.
- A brief biographical sketch of the proposed Special Issue editor.

Special Issue proposals should be submitted jointly to Gerard McKee and Sunny Bains.

Upcoming Deadlines

7 November 2003: Ideas for articles you'd like to write (or read).

16 January 2004: Calendar items for the twelve months starting September 2004.

The human and the machine

The inimitable chasm between human intelligence and machine intelligence has provided motivating challenges for researchers for many decades. Alternate perspectives and their justifications, such as found in Classical AI and New AI, thrive on differing interpretations of what are the key features in achieving a notion of intelligence through some artificial mechanisms. To date, the emphasis on bringing the human and the machine closer together through, for example, humanoid robotics, has often ignored the inherent features of the machine and consequently its advantages rather than limitations. It is currently estimated that there are nearly 900,000 multi-purpose robots in use worldwide (*World Robotics 2000* survey). One of the fundamental aspects leading to this considerable population of robots is their use for tasks for which they are inherently very proficient. Machines are good at mechanistic tasks: this is not a flaw, but rather an advantage.

The Anthropos Project at Media Lab Europe investigates the use of robots in our physical and social space from the perspective of exploiting its mechanistic capabilities and achieving a balance in its form and function. From a human-machine-interaction perspective, a socially-capable robot facilitates our access to the digital world. It does this through intuitive social mechanisms to improve or provide alternative approaches in education, information dissemination, and our future daily interactions with machines. Once domestic robots move beyond the washing machine, our social interaction with such machines becomes inevitable.

In order to begin addressing these issues, the Anthropos Project seeks to decompose the interaction issues between man and machine. Current research areas are described in the following paragraphs.

Balancing function and form for social robots

Anthropos and JoeRobot (see Figure 1) are prototypes built to explore the development of socially-capable robots. Key to the notion of expandability and rapid prototyping, a modular nervous-system strategy uses standardized interface protocols (Firewire and USB) for actuator and perceptor components. Research on integrating a socially-capable robot into performance spaces has demonstrated the power of the form as an interface to the digital information domain (Figure 1). People's willingness to engage with a machine that judiciously employs anthropomorphic features we are familiar with in social contexts facilitates man-machine interaction.

Strength and degree of minimal expression and communication

The Emotion Robots work is a series of experiments to investigate how minimal the set of humanlike features can be for a social robot. Data illustrates people's propensity to attribute such concepts as emotions and intelligence to machines performing computationally simple behaviours.

Seamless integration of physical worlds and information space

The Agent Chameleons project strives to develop digital *minds* that can seamlessly migrate, mutate, and evolve on their journey between and within physical and digital information spaces. This challenges the traditional boundaries between the physical and virtual through the empowerment of mobile agents. Three key attributes—mutation, migration, and evolution—underpin this concept. Here, digital personal assistants are developed that opportunistically migrate and choose a *body* (whether a robot, an avatar in virtual reality, an animated character on a PDA, or a web agent) to facilitate its intentions. A PDA is no longer a device, but a digital friend capable of using many platforms.

Humans and their machines

Biometric, force-feedback, and video data is retrieved from Team Media Lab Europe's national motorcycle racing team in the Vicarious Adrenaline project. This work investigates rider and motorcycle performance in conjunction with creating a third party experience rather than passive observation of the racing event.

Machines have intrinsic properties that are often seen as hindrances when the reference is either humans or other biological entities. The objective is to embrace those aspects that are constructive and integrate these with a machine's inherent advantages, i.e. being a machine. The primary research goals are the challenges of understanding and establishing a bond between man and machine. Can the illusion of life and intelligence emerge

Continues on page 11.

Figure 1. Media Lab Europe's JoeRobot at the Flutterfugue performance in London 2002. (Photo courtesy of Brent Jones)

SPIE Society of Photo-Optical Instrumentation Engineers

P.O. Box 10 • Bellingham, WA 98227-0010 USA

Change Service Requested

DATED MATERIAL

Non-Profit Org.
U.S. Postage Paid
Society of
Photo-Optical
Instrumentation
Engineers