Arnold: An Anthropomorphic Autonomous Robot for Human Environments

Thomas Bergener, Carsten Bruckhoff, Percy Dahm,
Herbert Janßen, Frank Joublin, Rainer Menzner
http://www.neuroinformatik.ruhr-uni-bochum.de/ini/VS/vS_group.html
Institut für Neuroinformatik, Lehrstuhl für theoretische Biologie
Ruhr–Universität Bochum

Abstract

We describe the general concept, system architecture, hardware and the first behavioral abilities of Arnold, a mobile autonomous service robot with a robot arm with seven degrees of freedom and a dual stereo camera head. Arnold is anthropomorphic for better adaptation to typical human environments and to allow for human-like behavioral strategies in solving complex tasks. Arnold was designed for service tasks like locating, recognizing and grasping objects in a rich initially unknown environment and to freely move in relatively small rooms with static and dynamic obstacles and conventional doors. Currently Arnold is able to fixate, move to and grasp for a human hand and avoid obstacles based on optical flow analysis.

Introduction

Why anthropomorphic autonomous service robots?

Robots - in the sense of human-like general purpose machines that are able to perform well in a human environment and can handle the various badly defined everyday tasks we humans spend a lot of time with - are still a matter of the future.

Besides the fascination to create our own technical analogon, the design of such robots poses a number of problems that are crucial for a lot of applications. The methods that need to be developed as a presupposition for successful robot design are important for many fields that at first glance have nothing to do with robotics. This includes automatic data acquisition, process control, user assistance for handling complex tasks like e.g. driving, and certainly entertainment. Robot design calls for a very high integration of mechanical, electrical, electronical and computational technology.

Contrary to how most mobile robots are constructed today, a look at the biological prototype reveals that not the extensive use of highly precise sensors is required, but that a few less perfect but universally applicable input channels - namely vision, audio and touch - can be used to great effect. As useful as a collection of special purpose devices for distance measurement, illumination etc. may be in the short run, they imply additional system load and unnecessary complexity for special cases and thus are a dead end for the development of more generic systems.

There is one more reason besides learning from an existing example for adopting a human-like *architecture* for an autonomous robot: General-purpose autonomous service robots for home or office environments have to deal with an environment that was intensively adapted to human anatomy, sensory and motor skills. Table height, position of door-handles and light switches, width of doors and other space to be crossed all match our specific abilities. A machine meant to perform well in a significant subset of typical human tasks in such an environment will better be designed after this anthropomorphic standard - especially since a lot of presently available technology does not even come close to the performance of existing biological solutions.

Figure 1: Arnold

Another common misconception for intelligent systems is the call for complete user or operator control. Total and detailed control over every aspect of the system is something that a typical classical computer system offers. This is also the reason why those systems are hard to use, require a lot of knowledge and training and are nevertheless easy to make erros on, i.e. to generate command with a different outcome than that indended by the user. Our interaction with more intelligent “systems“ like e.g. a trained dog is very different from that: Although the communication channel - speech and gestures - is highly redundant, the syntax and semantics are mostly ill-defined and the animal is never guaranteed to even obey a given command, the interaction is quite intuitive, robust even under difficult conditions and generally pretty powerful. The dog is easy to deal with because it will generally use its own intelligence and only do what his master requests if the commands are not in contradiction to his own "behavioral goals". A dog that is commanded "come here straight" and runs into a wall because it obeys the command would be called particularly stupid - with good reason. The same holds true for a robot. Combining real autonomy of behavior with a user interface may be quite straightforward by defining the command input on the same level as the sensory input and the general behavioral goals; such as e.g. obstacle avoidance.

Besides general purpose service robotics there is also the field of robotics for the disabled[9, 10]. Mobility assistances for the blind and wheelchair-mounted robot arms are only two examples posing a technical challenge for highly robust robot control which allows the user to easily specify tasks. One possible and maybe ideal approach for robotic tools for the disabled would be a library of semi-autonomous behaviors that perform tedious subtasks - like e.g. locating an object for grasping or determining trajectories. Thus robotics for the disabled share the concept of behaviours or skills as coarsely parametrized atoms by which more complex tasks can be successfully performed.

Anthropomorphic shape and the fluent, predictable movements of effectors controlled by dynamical systems rather than stiff trajectories also have a strong effect on users which makes working with such a robot aesthetically and emotionally more pleasing.

“Anatomy“ - Body Shape, Sensors and Effectors

Hardware Design determines Behavioral Abilities

As discussed before, one of the main goals followed in the construction of Arnold was to design the system as anthropomorphic and adapted to a human environment as possible with currently available and preferably well-established technology.

Arnold’s main features are the camera head, the arm and the vehicle. The camera head should perform well for Arnold’s typical range of tasks, which includes orientation and locomotion in an indoor environment, recognizing, locating and grasping of objects and perhaps manipulation of objects like doorhandles and switches.

Arnold is developed in the BMBF project NEUROS (Neural Robot Skills) to demonstrate how service tasks in a human environment can be performed by an autonomous anthropomorphic robot combining loosely coupled dynamical systems realizing basic behaviors to achieve a complex overall system behavior.

Visual constraints To acquire optimal sensory information for navigation, Arnold needs to see objects very close to its base. A logical consequence is to make the general shape of the body pyramidal. For orientation and obstacle avoidance a large field of view that allowing to track objects while passing them is ideal. On the other hand the recognition of small objects to be grasped or manipulated requires a high sensor resolution for at least a small field of view. Therefore Arnold is equipped with a double stereo camera system with a wide-angle monochrome “periphery“ camera and a color “fovea“ camera per side.

Visual Abilities Movements of a visual sensor are very useful for acquiring specific information about the world at relatively high speed. They also allow to avoid ambiguities and configurations error-prone for specific image processing algorithms. Vergence - i.e. the ability to move the point of intersection of the optical axes to any distance - specifically helps to minimize and simplify depth computation from stereo images. An ideal stereo camera head allows pure rotation of all cameras around their main point to avoid parallax errors in determining directions and additionally implements a rotating "neck" to enable symmetrical vergence configurations for any direction. In the real world, such a solution would be bulky and expensive, thus we use only three degrees of freedom for pan, tilt and vergence, where the fovea camera mainpoints lie at the intersection of the pan and tilt axes. The stereo base is at about 300 mm significantly wider than with human anatomy to compensate for the lower resolution of the fovea cameras with respect to the human fovea.

Our robot is equipped with no other sensors than cameras because we are strongly convinced that one versatile sensor proved to be the fittest for an extremely wide variety of environments and tasks is - in the long run - better suited for behavioral control than an assortment of sensors with special and supposedly ideal abilities for special tasks like e.g. distance measurements.[1]

Head Configuration The head position is important for both navigation and grasping. For navigation it is ideal to have identical axes for the camera pan and platform rotation.

For grasping the relative configuration of body, arm and head is cruicial: Humans typically use visually controlled arm and hand movements with the position of the head above and behind the grasping

[1]We are currently evaluating the use of acoustic sensors, though.

position. This and the fact that most visually controlled grasping is done from the side with the elbow pointing outwards gives an optimal visual control of the situation and especially the most critical part, the grip itself.

Grasping Space Humans typically grasp objects that are in a certain spatial area before the chest, while the manipulation of objects under visual control far away from this preferred space is often simplified by using additional degrees of freedom of the body - like bending the knees to grasp objects on the floor. For Arnold those additional degrees of freedom are technically not feasible so we decided to choose an anatomy best suited for the most probable handling position which we assume to be in about table height. With 1.35 m, Arnold has about the size of an 8 year old child and its grasping configuration with respect to typical tables is comparable to a sitting adult.

Arm Configuration Arnold's arm configuration is equivalent to a broadly simplified model of the human arm with a 3 degree of freedom (DoF) shoulder, an 1 DoF elbow and a 3 DoF wrist. The maximum grasping radius is about 1 m with a maximum load of about 1.5 kg.

The redundancy of the seventh degree of freedom allows the handling of situations in which additional movement constraints have to be met. Our main interest in utilizing this redundancy is in obstacle avoidance during grasping, which is important for handling of objects on tables, cupboards, etc.

This human-like arm configuration also has the advantage to be controlled by a distributed control scheme: the wrist position (3 DoF), the wrist orientation (3 DoF), and the elbow angle with respect to the shoulder-wrist axis (1 DoF) can be controlled by 3 independent control loops, allowing e.g. the elbow angle control to use a very specific and efficient obstacle avoidance criterion.

Mobile Platform The footprint is another important feature for any robot that has to move in the cluttered environments humans prefer to live in. Arnold's mobile base is approximately square with about 750 mm edge length. This is barely acceptable with respect to the typical width of doors. Of similar importance is the ability to rotate the body in place. Arnold's body rotation axis is centered with respect to the base area and equal to its head pan axis so ideal camera control is even possible during body rotation.

Limitations A discussion of Arnold's design would be incomplete without mentioning some major shortcomings with respect to basic human abilities. Probably the most serious limitation is the wheel based platform which - in comparison to the more elegant design of human legs - does not only increase Arnold's base area but does not even allow to bend knees for reaching objects outside the normal reaching volume, climb stairs or sit down at a table. So Arnold cannot grasp objects on the floor, and even if its arm length would allow it to do so, there would be almost no sensory control possible since there are no degrees of freedom of the body to accomodate for the head–hand configuration.

“Physiology“ - Hardware and System Software

Hardware Requirements

The technical realization of Arnold's design goals was mainly determined by the principle of using reliable “standard“ hardware in order to put the emphasis on behavioral control rather than tool development.

Arnold's hardware architecture meets the following main requirements:

- The whole platform is completely autonomous, including power supply, data connections and the necessary computing power.
- Realtime processing for images (e.g. optical flow field) and arm control is possible. We determined the maximum hard realtime requirement to be about 1 ms for the arm control module. Apart from fast processors, this requires a high-bandwidth internal bus and communication system.
- The software development allows convenient access from external hosts for multiple users simultaneously. This requires a capable multiuser/multitasking OS.

The requirements were met by using a small Arnold-wide network of Intel PCI PCs running QNX, a distributed realtime operating system. The network is currently realized as 10 Mbit Ethernet but will soon be changed to 100 Mbit.

Arnold's general hardware concept is modular. After power-on, only a minimum of modules are active: the "Master-PC", whose function will be explained later, a special interface board which manages the switching of the components, and the cameras and video electronics.

Arnold's other subsystems can be separately activated under software control. These are the power supply of the arm, the power supply of the camera head, the Labmate platform and the Slave-PCs.

Mobile Platform The mobile base platform is a TRC Labmate running on four passive casters in each corner and driven by two powered wheels. Thus Arnold is able to rotate around its center axis and requires a minimum operational area. The Labmate is controlled by a serial line interface. Its power supply uses two standard Pb-accumulators with 12V/66Ah each, resulting in a 24V main supply voltage. We modified the Labmate platform so that its power and controller board are controlled by the master PC via relays. Additionally the Labmates accumulators serve as one of two possible power sources for all of Arnold's systems, the other alternative being an external power supply as discussed later.

Robot Arm The arm is a modular system manufactured by Amtec. It has seven degrees of freedom, giving one degree of redundancy, which is useful for e.g. obstacle avoidance during grasping.

Each module carries a micro-controller with a CAN-bus interface. For communication with the arm an ISA-card implementing the CAN-bus protocol is used.

Camera Head The camera head is a TRC Zebra pan/tilt/vergence unit. Pan range is about $\pm$ 180° and tilt movement in the range of $\pm$ 90°. The vergence DOF was modified for better defined rotation axes and a larger camera base distance. The distance between the vergence axes is now 300 mm with a vergence angle between 0° and 40°, resulting in a minimum fixation distance of about 400 mm. The head's controller board is connected via a serial line to the master PC.

Two color "fovea" cameras with 35° field of view are positioned with their main point at the intersection of tilt and vergence axes and two monochrome "periphery" cameras with 90° field of view are mounted o the outside with 380 mm base distance.

Video Subsystem The video subsystem consists of the two color fovea cameras, the two monochrome periphery cameras and 12 video amplifiers. The video amplifiers have access to all 4 video signals, namely for input to the framegrabbers, camera synchronisation and external video output. Synchronization between camera pairs is crucial for stereo image processing of dynamical scenes.

The video signals are digitized using two PCI-bus framegrabbers manufactured by Imaging Technology Inc. with 2 MB buffer memory, color space conversion and about 80 MB/s framebuffer to host transfer rate. The driver software for QNX unfortunately had to be written in-house, since synchronous acquisition of stereo images with basic preprocessing options still seems to be an exotic requirement.

Power Supply System The power supply system allows for two operation modes with seamless switching via a Schottky diode:

- In external mode power is supplied via an external mains line for software development and automatic accumulator charging. The external supply voltage is 400V three-phase alternating current fed to a Sitec 24V/40A power supply
- In autonomous mode the whole platform is be powered by the Labmate's accumulators.

Besides the 24V basic voltage for the arm, platform and camera head, a few other supply voltages are generated via DC/DC-convertors: 12V for the PCs, the cameras and the video-electronics, -12V for the PCs and 5V for the PCs and the camera head.

To decouple the most critical components from momentary voltage drops generated by the high power consumers separate DC/DC-converters are used for the master and slave PCs.

Under normal development conditions Arnold's accumulators are sufficient for about 1 hour of continuous operation.

Computer Network The computational power is provided by a network of PCs - currently two 166MHz Pentiums - running QNX.

The master PC is central for Arnold's operation: It is the first piece of hardware activated when Arnold is switched on, controls all other components via an I/O card connected to various relays and acts as a gateway to the institute's LAN for external access via Ethernet. It also provides a simple status display via a couple of LEDs displaying the power and logic status of the hardware components.

The slave PCs are diskless and booted via TCP/IP from the master. Currently only one slave PC is installed which interfaces to the arm via a CAN-bus card. Space for three slave PCs is available, though, and we plan to extend the system as soon as the necessary computing power meets our budget. Since Arnold's internal LAN uses Fast Ethernet with its own 4-port hub, adding PC boards can be a matter of minutes.

User Interface and Security In addition to access via the external network connector used for general software development, Arnold provides several on-board interfaces for status information and security control. All video signals, keyboard and VGA connectors and reset buttons are directly available on a back panel. LEDs display the status of all supply voltages, the camera head's motor power, the bootup state of all PCs and harddisk accesses of the master. Two connectors implement emergency switches separately for the arm and the complete system. The former is convenient since our experience is that most emergencies are caused by the arm operation.

System Software

While the operating system provides basic communication and control interfaces, we developed the high level software system “Planet“ which provides a simple way to concurrently run behavior modules connected by transparent communication channels. Planet is controlled via a configuration file defining the process interconnection scheme and a system monitor allowing runtime inspection and control of the multiple processes. Planet is especially suited to implement and test behavioral architectures based on asynchronous data driven modules[1].

Arnold’s sensors and effectors are represented in Planet by three modules for communication with the platform, the arm and the camera head. Basically, the sensor and effector modules are servers that provide information and can be controlled concurrently or exclusively by any behavior module that is properly configured via Planet.

Internally Planet uses both shared memory and low level network communication in order to provide maximum speed for both CPU local and distributed interprocess communication.

“Mind“ - Behavioral Control

The main idea of Arnold’s behavior control stems from behavior based robotics, an approach first formulated by Braitenberg 1984 [3]. In 1985 Brooks proposed the subsumption architecture [4, 5] as an approach to define a hierarchical structure for behavior based autonomous systems. The complex behavior necessary to accomplish a given task results from coupling several simple behaviors. Each of the simple behaviors encompasses the perception, planning and task execution capabilities necessary to achieve a specific aspect of robot control. The different simple behaviors run in parallel, each performing a meaningful action. The clever combination of these behaviors realizes complex behavior of the robot.

To demonstrate our ideas, we have specified two example tasks Arnold should be able to perform at the end of the project. The first one is finding and grasping a cup. Arnold has no knowledge about its surroundings and thus needs an exploration and navigation system working properly in indoor environments. It has to search for the cup and establish a world model so that it can bring back the cup to the starting point. This system is implemented as a simulation and is currently being ported to the robot. The robot has to avoid static and dynamic obstacles in a “rich“ indoor environment. The robot has to approach the object, e.g. a table with a cup, very closely, so that the cup is in a good position to grasp. For this subtask a goal movement based on the 3-D object shape is currently being developed as a dynamical system. In the context of grasping we concentrate our efforts on realtime visually controlled arm movement in an environment with lots of obstacles. Hence, we also develop a method of estimating the 3-D position of the object to grasp and of the obstacles to avoid.

Figure 2: Grasping for a hand

The second task is opening a door. Arnold has to find the door and the handle and has to open the door without touching it with the robot’s body. In this scenario our main interest is in the coordination of concurrent arm and platform movement. While it would certainly be possible to control this behavior by one specialized complex algorithm, the challenge is to perform it as a combination of generic grasping and obstacle avoidance behaviors.

Another important aspect of Arnold is designing an appropriate architecture of the behavioral system. Due to the good results achieved in robot control by dynamical systems (i.e. Neven et al. 1996 [8]) we are planning to design each behavior as a dynamical system. The systems should not be coupled directly but indirectly by the sensor data. The time constants of the dynamical equations determine the hierarchy of the behaviors. In changing these constants during operation we can achieve different complex behaviors. This approach also allows us to formulate behavioral sequences (Steinhage, Schöner 1996 [11]).

The behavioral and preprocessing modules implemented so far have been kept simple on purpose because we want to show the possibilities coupling simple behaviors with dynamic systems to attain complex behavior. Also realtime performance is an important feature for any applications in a dynamical environment.

Hand-Following

For demonstrating our ideas of controlling Arnold by coupling dynamical systems achieving complex behavior we have implemented an example task: Arnold should look for the next human in its field of view, identified by skin color. The robot should keep a distance of about one meter to the human and point at the hand with its gripper.

To achieve this task, Arnold has to find the appropriate skin-colored areas, move the platform to an appropriate position and follow the hand with his head and with the gripper.

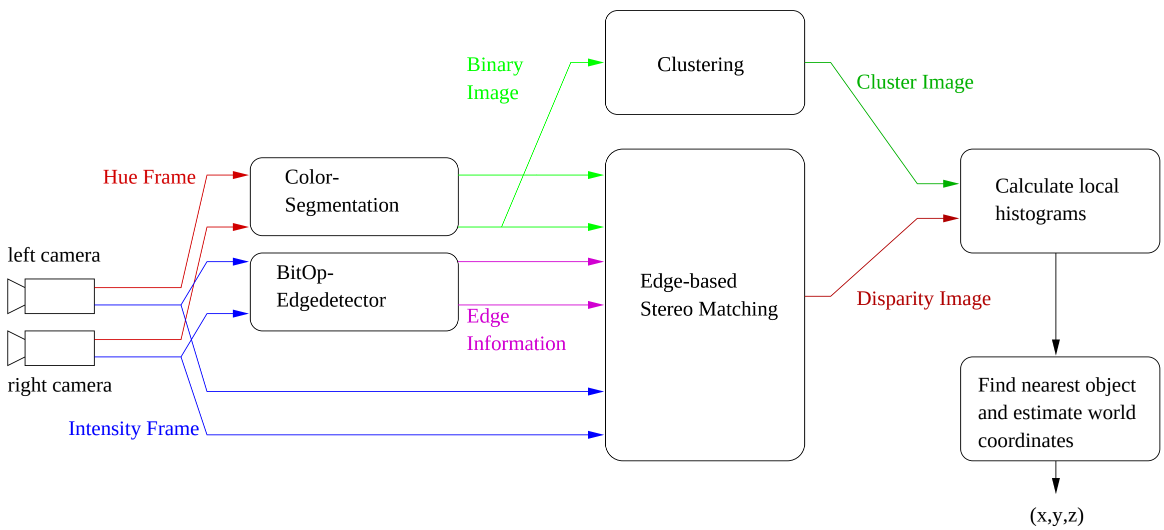

Figure 3: Skin Extraction Algorithm

The diagram in figure 3 shows the algorithm computing the 3-D coordinates of the skin area next to Arnold. The extraction of the image pixels of skin-color is done in the hue frame. In this frame, skin lies in a very characteristic range. The following clustering algorithm determines possible target areas in the image pair. In parallel to the skin detection an algorithm based on the Bit-Op (Goerick 1994 [6]) extracts the vertical edges. A correspondance algorithm computes the disparity of the edges. By calculating local histograms, the 3-D coordinates of each area are determined. The area nearest to Arnold is chosen as the target position for the given task.

For the hand following task Arnold is controlled by three simple behaviors. The first behavior moves the platform to a distance of about one meter to the target location. The second one moves the gripper to the 3-D coordinates of the hand and the third one moves the camera system in order to fixate the target. Each of these behaviors is working independently and is implemented as a dynamical system. They are only coupled by the 3-D sensor information of the skin-colored area next to Arnold. So the complex behavior of hand-following emerges from the coupling of three relatively simple control loops. A detailed description can be found in Dahm, Bergener 1997 [2].

Figure 4: Stereo image pair with corresponding skin colored image points and target location marked.

Obstacle Avoidance

As an example for the approach of dynamical systems on the sensor side, we have implemented a dynamical optical flow estimation. This algorithm takes advantage of the favorable stability properties of dynamical systems as robust estimators particularly suited for multimodal and noisy visual information. A detailed description can be found in Neven et al 1996 [8]. The result of the flow estimation is used in a simple behavior module that balances the overall flow in the left and the right image resulting in an obstacle avoidance that is particularly sensitive with respect to moving objects.

Some results for both hand following and obstacle avoidance are available as live movies in electronical form [7].

Conclusion and Outlook

This paper describes the design of Arnold, an autonomous mobile robot for service tasks. Since service robots have to perform well in an environment highly adapted to humans we propose that these robots have to be designed as anthropomorphic as possible. The best way to achieve complex behavior consists in coupling of simple behaviors. This is done best by designing each behavior as a dynamical system and coupling these system by sensor data as shown in the example task we have implemented so far.

In the future we are planning to implement the whole range of behaviors neccessary to cope with the discussed scenarios.

The discussion of Arnold's behavioral architecture, i.e. our methods to combine relatively simple behavioral modules into complex systems, is merely sketched in this paper and will be addressed in future publications.

We also plan to extend our work on user interfaces for autonomous robots driven by dynamical systems that gives the user predictable control yet does not compromise autonomy and clean architectural design.

References

[1] Rudi Alberts et al. Projektgruppe: Entwicklung der Netzwerk- und Steuer-Software eines autonomen mobilen Roboters. Technical report, Fachbereich Informatik der Universität Dortmund, 1997 (in press).

[2] Thomas Bergener and Percy Dahm. A framework for dynamic man-machine interaction implemented on an autonomous mobile robot. In *ISIE'97, IEEE International Symposium on Industrial Electronics*, 1997 (to appear).

[3] Valentino Braitenberg. *Vehicles: Experiments in Synthetic Psychology*. The MIT Press, 1984.

[4] Rodney A. Brooks. A robust layered control system for a mobile robot. Technical Report 864, MIT, Artificial Intelligence Laboratory, September 1985.

[5] Rodney A Brooks. From earwigs to humans. *Robotics and Autonomous Systems*, (in press).

[6] Christian Goerick. Local orientation coding and adaptive thresholding for real time early vision. Technical Report 94-05, Institut für Neuroinformatik, Ruhr-Universität Bochum, D-44780 Bochum, Germany, June 1994.

[7] NEUROS Project - von Seelen Group. Ruhr-Universität Bochum, Institut für Neuroinformatik. http://www.neuroinformatik.ruhr-uni-bochum.de/ini/PROJECTS/NEUROS/NEUROS.html.

[8] H. Neven, A. Steinhage, M.A. Giese, and C. Bruckhoff. Dynamical systems for vision-based autonomous mobile robot. In *From Animals to Animats 4: Proceedings of the Fourth International Conference on Simulation of Adaptive Behavior*, pages 113–123, September 1996.

[9] Stephen D. Prior. An electric wheelchair mounted robotic arm - survey of potential users. *Journal of Medical Engineering and Technology*, 14(4):143–154, July/August 1990.

[10] Carol A. Stanger, Carolyn Anglin, William S. Harwin, and Douglas P. Romilly. Devices for assisting manipulation: A summary of user task priorities. *IEEE Transactions on Rehabilitation Engineering*, 1994.

[11] A. Steinhage and G. Schöner. Self-calibration based on invariant view recognition: Dynamic approach to navigation. *Robotics and Autonomous Systems*, (in press).