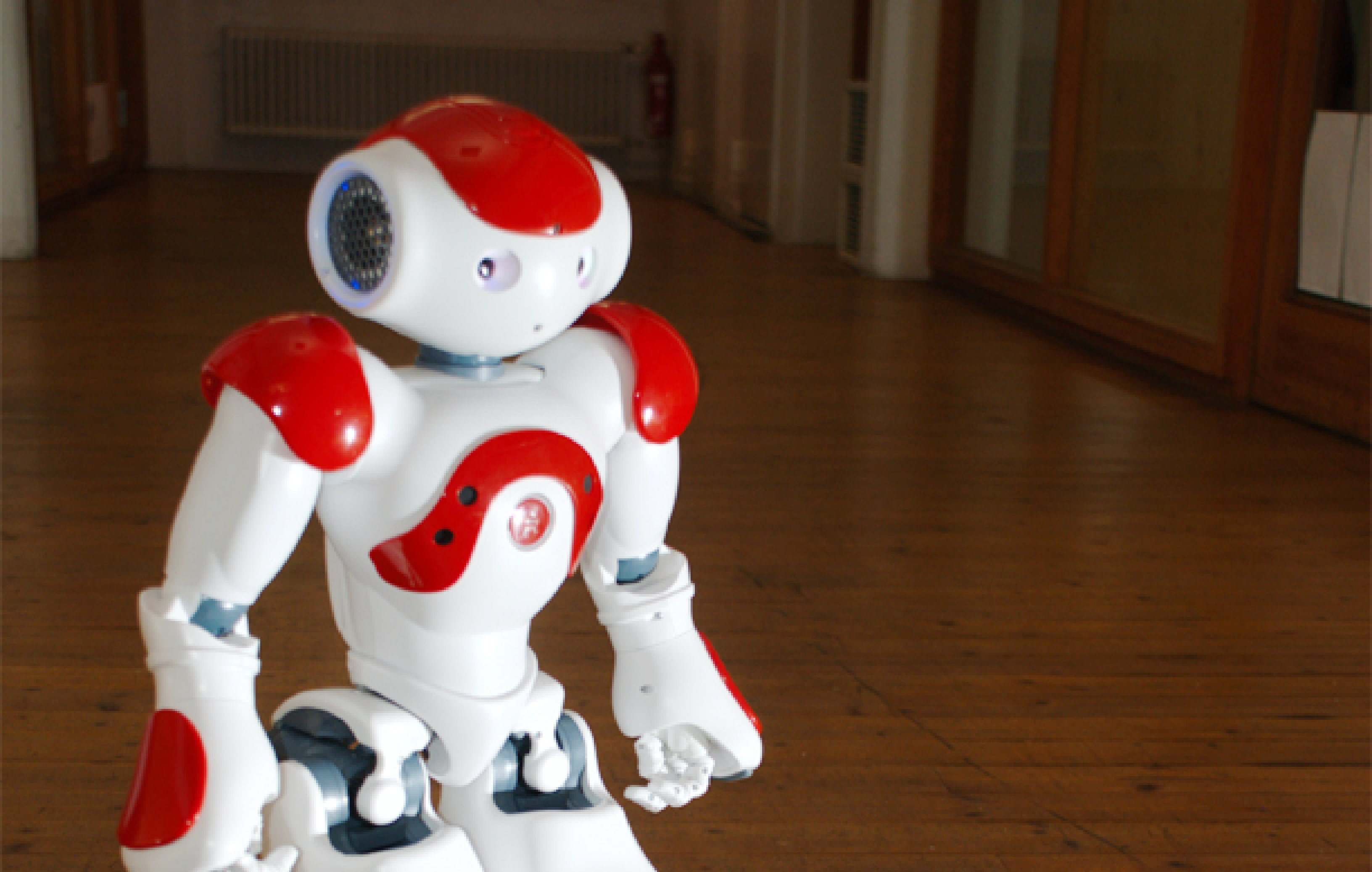


Major Concepts and Applications of Robotics & Artificial Intelligence

Leesa Headrick

First Edition, 2012

ISBN 978-81-323-3618-1

Published by:
University Publications
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

Chapter- 1

Artificial Consciousness

Artificial consciousness (AC), also known as **machine consciousness** (MC) or **synthetic consciousness**, is a field related to artificial intelligence and cognitive robotics whose aim is to define that which would have to be synthesized were consciousness to be found in an engineered artifact (Aleksander 1995).

Neuroscience hypothesizes that consciousness is generated by the interoperation of various parts of the brain, called the neural correlates of consciousness or NCC. Proponents of AC believe it is possible to construct machines (e.g., computer systems) that can emulate this interoperation.

Philosophical views of artificial consciousness

As there are many designations of consciousness, there are many potential types of AC. In the philosophical literature, perhaps the most common taxonomy of consciousness is into "access" and "phenomenal" variants. Access consciousness concerns those aspects of experience that are amenable to a functional description, while phenomenal consciousness concerns those aspects of experience that seem to defy functional depiction, instead being characterized qualitatively in terms of "raw feels", "what it is like" or qualia (Block, 1997). Weaker versions of AC only require that functional, "access consciousness" be artificially instantiated. Stronger versions of AC hold that all aspects of consciousness are instantiable in artificial, machine systems.

The debate over the plausibility of artificial consciousness

There is considerable debate over the plausibility of AC. Some theorists (e.g., biological naturalists) hold that consciousness can only be instantiated in biological systems (Searle, 2004). A slightly more liberal view, though one still skeptical of AC, is held by theorists (e.g., type-identity theorists) who hold that consciousness can only be realized in particular physical systems because consciousness has properties that necessarily depend on physical constitution (Block, 1978; Bickle, 2003). Their skepticism rests on the fact that there are many physical asymmetries between natural, organic systems and artificially constructed (e.g., computer) systems, and it seems reasonable to think it is unlikely that these differences would not be relevant to the generation of conscious states.

However, other theorists are more sanguine about the plausibility of AC. For some theorists (e.g., functionalists), who define mental states in terms of causal roles, any system that can instantiate the same pattern of causal roles, regardless of physical constitution, will instantiate the same mental states, including consciousness (Putnam, 1967). Along these lines, some theorists (e.g., David Chalmers) have proposed that consciousness can be realized in properly designed and programmed computers.

Chalmers' argument for artificial consciousness

One of the most explicit arguments for the plausibility of AC comes from David Chalmers. His proposal, found within his manuscript A Computational Foundation for the Study of Cognition, is roughly that computers perform computations and the right kinds of computations are sufficient for the possession of a conscious mind. In outline, he defends his claim thus: Computers perform computations. Computations can capture other systems' abstract causal organization. Mental properties are nothing over and above abstract causal organization. Therefore, computers running the right kind of computations will instantiate mental properties.

The most controversial part of Chalmers' proposal is that mental properties are "organizationally invariant;" i.e., nothing over and above abstract causal organization. His rough argument for which is the following. Mental properties are of two kinds, psychological and phenomenological. Psychological properties, such as belief and perception, are those that are "characterized by their causal role" within an overall causal system. He adverts to the work of Armstrong (1968) and Lewis (1972) in claiming that "[s]ystems with the same causal topology…will share their psychological properties." Phenomenological properties, on the other hand, are not prima facie definable in terms of their causal roles. Establishing that phenomenological properties are amenable to individuation by causal role therefore requires argument. Chalmers provides his "Dancing Qualia Argument" for this purpose.

Chalmers begins by assuming that agents with identical causal organizations could have different experiences in virtue of having different material constitutions (silicon vs. neurons, e.g.). He then asks us to conceive of changing one agent into the other by the replacement of parts (neural parts replaced by silicon, say) while preserving its causal organization. Ex hypothesi, the experience of the agent under transformation would change (as the parts were replaced), but there would be no change in causal topology and therefore no means whereby the agent could "notice" the shift in experience. To imagine, however, that it makes sense to say an agent could having qualitative changes in experience but be unable to notice changes in experience is absurd. Given the absurdity of the conclusion, Chalmers' rejects the initial premise that agents with identical causal organization can have different experiences. Thus, by his Dancing Qualia Argument, Chalmers defends his view that phenomenological properties are organizationally invariant.

Critics of AC can object that Chalmers begs the question in assuming that all mental properties are sufficiently captured by abstract causal organization. It remains a contentious proposal. Regardless, it is important for laying a philosophical foundation for the plausibility of AC.

Consciousness in digital computers

There are various aspects of consciousness generally deemed necessary for a machine to be artificially conscious. A variety of functions in which consciousness plays a role were suggested by Bernard Baars (Baars 1988) and others. The functions of consciousness suggested by Bernard Baars are Definition and Context Setting, Adaptation and Learning, Editing, Flagging and Debugging, Recruiting and Control, Prioritizing and Access-Control, Decision-making or Executive Function, Analogy-forming Function, Metacognitive and Self-monitoring Function, Autoprogramming and Self-maintenance Function, and Definitional and Context-setting Function. Igor Aleksander suggested 12 principles for Artificial Consciousness (Aleksander 1995) and these are: The Brain is a State Machine, Inner Neuron Partitioning, Conscious and Unconscious States, Perceptual Learning and Memory, Prediction, The Awareness of Self, Representation of Meaning, Learning Utterances, Learning Language, Will, Instinct, and Emotion. The aim of AC is to define whether and how these and other aspects of consciousness can be synthesized in an engineered artifact such as a digital computer. This list is not exhaustive; there are many others not covered.

Awareness

Awareness could be one required aspect, but there are many problems with the exact definition of *awareness*. The results of the experiments of neuroscanning on monkeys suggest that a process, not a state or object, activates neurons . For such a reaction, a model of the process based on the information received through the senses must be created, which demands a lot of flexibility, and is also useful for making predictions.

Learning

Learning is also considered necessary for AC. By Bernard Baars, conscious experience is needed to represent and adapt to novel and significant events (Baars 1988). By Axel Cleeremans and Luis Jiménez, learning is defined as "a set of philogenetically [*sic*] advanced adaptation processes that critically depend on an evolved sensitivity to subjective experience so as to enable agents to afford flexible control over their actions in complex, unpredictable environments" (Cleeremans 2001).

Anticipation

The ability to predict (or anticipate) foreseeable events is considered important for AC by Igor Aleksander (Aleksander 1995). The emergentist multiple drafts principle proposed by Daniel Dennett in *Consciousness Explained* may be useful for prediction: it involves the evaluation and selection of the most appropriate "draft" to fit the current environment.

By Igor Aleksander, relationships between world states are mirrored in the state structure of the conscious organism enabling the organism to predict events (Aleksander 1995). An artificially conscious machine should be able to anticipate events correctly in order to be ready to respond to them when they occur. The implication here is that the machine needs real-time components, making it possible to demonstrate that it possesses artificial consciousness in the present and not just in the past. In order to do this, the machine being tested must operate coherently in a novel environment, to simulate the real world.

Subjective experience

Subjective experiences or qualia are widely considered to be *the* hard problem of consciousness. Indeed, it is held to pose a challenge to physicalism, let alone computationalism. On the other hand, there is a similar problem with uncertainty principle in physics, which has not made the research in physics impossible.

Symbolic or hybrid proposals

Franklin's Intelligent Distribution Agent

Stan Franklin (1995, 2003) defines an autonomous agent as possessing functional consciousness when it is capable of several of the functions of consciousness as identified by Bernard Baars' Global Workspace Theory (Baars 1988, 1997). His brain child IDA (Intelligent Distribution Agent) is a software implementation of GWT, which makes it functionally conscious by definition. IDA's task is to negotiate new assignments for sailors in the US Navy after they end a tour of duty, by matching each individual's skills and preferences with the Navy's needs. IDA interacts with Navy databases and communicates with the sailors via natural language e-mail dialog while obeying a large set of Navy policies. The IDA computational model was developed during 1996–2001 at Stan Franklin's "Conscious" Software Research Group at the University of Memphis. It "consists of approximately a quarter-million lines of Java code, and almost completely consumes the resources of a 2001 high-end workstation." It relies heavily on *codelets*, which are "special purpose, relatively independent, mini-agent[s] typically implemented as a small piece of code running as a separate thread." In IDA's top-down architecture, high-level cognitive functions are explicitly modeled. While IDA is functionally conscious by definition, Franklin does "not attribute phenomenal consciousness to his own 'conscious' software agent, IDA, in spite of her many human-like behaviours. This in spite of watching several US Navy detailers repeatedly nodding their heads saying 'Yes, that's how I do it' while watching IDA's internal and external actions as she performs her task."

Ron Sun's cognitive architecture CLARION

CLARION posits a two-level representation that explains the distinction between conscious and unconscious mental processes.

CLARION has been successful in accounting for a variety of psychological data. A number of well-known skill learning tasks have been simulated using CLARION that span the spectrum ranging from simple reactive skills to complex cognitive skills. The tasks include serial reaction time (SRT) tasks, artificial grammar learning (AGL) tasks, process control (PC) tasks, the categorical inference (CI) task, the alphabetical arithmetic (AA) task, and the Tower of Hanoi (TOH) task (Sun 2002). Among them, SRT, AGL, and PC are typical implicit learning tasks, very much relevant to the issue of consciousness as they operationalized the notion of consciousness in the context of psychological experiments.

The simulations using CLARION provide detailed, process-based interpretations of experimental data related to consciousness, in the context of a broadly scoped cognitive architecture and a unified theory of cognition. Such interpretations are important for a precise, process-based understanding of consciousness and other aspects of cognition, leading up to better appreciations of the role of consciousness in human cognition (Sun 1999). CLARION also makes quantitative and qualitative predictions regarding cognition in the areas of memory, learning, motivation, meta-cognition, and so on. These predictions either have been experimentally tested already or are in the process of being tested.

Ben Goertzel's OpenCog

Ben Goertzel is pursuing an embodied AGI through the open-source OpenCog project. Current code includes embodied virtual pets capable of learning simple English-language commands, as well as integration with real-world robotics, being done at the robotics lab of Hugo de Garis at Xiamen University.

Neural network proposals

Haikonen's cognitive architecture

Pentti Haikonen (2003) considers classical rule-based computing inadequate for achieving AC: "the brain is definitely not a computer. Thinking is not an execution of programmed strings of commands. The brain is not a numerical calculator either. We do not think by numbers." Rather than trying to achieve mind and consciousness by identifying and implementing their underlying computational rules, Haikonen proposes "a special cognitive architecture to reproduce the processes of perception, inner imagery, inner speech, pain, pleasure, emotions and the cognitive functions behind these. This bottom-up architecture would produce higher-level functions by the power of the elementary processing units, the artificial neurons, without algorithms or programs". Haikonen believes that, when implemented with sufficient complexity, this architecture will develop consciousness, which he considers to be "a style and way of operation, characterized by distributed signal representation, perception process, cross-modality reporting and availability for retrospection." Haikonen is not alone in this process view of consciousness, or the view that AC will spontaneously emerge in autonomous agents that have a suitable neuro-inspired architecture of complexity; these are shared by many, e.g.

Freeman (1999) and Cotterill (2003). A low-complexity implementation of the architecture proposed by Haikonen (2003) was reportedly not capable of AC, but did exhibit emotions as expected.

Takeno's self-awareness research

Self-awareness in robots is being investigated by Junichi Takeno at Meiji University in Japan. Takeno is asserting that he has developed a robot capable of discriminating between a self-image in a mirror and any other having an identical image to it , and this claim has already been reviewed (Takeno, Inaba & Suzuki 2005). Takeno asserts that he first contrived the computational module called a MoNAD, which has a self-aware function, and he then constructed the artificial consciousness system by formulating the relationships between emotions, feelings and reason by connecting the modules in a hierarchy (Igarashi, Takeno 2007). Takeno completed a mirror image cognition experiment using a robot equipped with the MoNAD system. Takeno proposed the Self-Body Theory stating that "humans feel that their own mirror image is closer to themselves than an actual part of themselves." The most important point in developing artificial consciousness or clarifying human consciousness is the development of a function of self awareness, and he claims that he has demonstrated physical and mathematical evidence for this in his thesis (Takeno 2008). He also demonstrated that robots can study episodes in memory where the emotions were stimulated and use this experience to take predictive actions to prevent the recurrence of unpleasant emotions (Torigoe, Takeno 2009).

Aleksander's impossible mind

Igor Aleksander, emeritus professor of Neural Systems Engineering at Imperial College, has extensively researched artificial neural networks and claims in his book *Impossible Minds: My neurons, My Consciousness* that the principles for creating a conscious machine already exist but that it would take forty years to train such a machine to understand language. Whether this is true remains to be demonstrated and the basic principle stated in *Impossible minds*: that the brain is a neural state machine is open to doubt.

Testing for artificial consciousness

The most well-known method for testing machine intelligence is the Turing test and it may be seen as an indirect test for consciousness. Another test, inspired by features of biological systems, is ConsScale, and it has been proposed as a means to measure and characterize the cognitive development of artificial creatures.

Yet there is a very serious objection to the plausibility of a test for consciousness. Tests, as such, are third-person procedures in which data is accessible to independent inquirers. However, qualia, or phenomenological consciousness, is an inherently first-person phenomenon. Accordingly, although various systems may display various signs of behavior correlated with functional consciousness, there is no conceivable way in which third-person procedures can have access to first-person phenomenological features. Ultimately, then, a test of strong versions of AC may be impossible.

The ethics of artificial consciousness

If it was certain that a particular machine was conscious its rights would be an ethical issue that would need to be assessed (e.g. what rights it would have under law). For example a conscious computer that was owned and used as a tool or central computer of a building or large machine is a particular ambiguity. Should laws be made for such a case, consciousness would also require a legal definition (for example a machine's ability to experience pleasure or pain, known as sentience). Because artificial consciousness is still largely a theoretical subject such ethics have not been discussed or developed to a great extent, though it has often been a theme in fiction (see below).

The rules for the 2003 Loebner Prize competition explicitly addressed the question of robot rights:

61. If, in any given year, a publicly available open source Entry entered by the University of Surrey or the Cambridge Center wins the Silver Medal or the Gold Medal, then the Medal and the Cash Award will be awarded to the body responsible the development of that Entry. If no such body can be identified, or if there is disagreement among two or more claimants, the Medal and the Cash Award will be held in trust until such time as the Entry may legally possess, either in the United States of America or in the venue of the contest, the Cash Award and Gold Medal in its own right.

Chapter- 2

Speech Recognition

The display of the Speech Recognition screensaver on a laptop, in which the character responds to questions, e.g. "Where are you?" or statements, e.g. "Hello."

Speech recognition (also known as **automatic speech recognition** or **computer speech recognition**) converts spoken words to text. The term "voice recognition" is sometimes used to refer to recognition systems that must be trained to a particular speaker—as is the case for most desktop recognition software. Recognizing the speaker can simplify the task of translating speech.

Speech recognition is a broader solution which refers to technology that can recognize speech without being targeted at single speaker—such as a call center system that can recognize arbitrary voices.

Speech recognition applications include voice user interfaces such as voice dialing (e.g., "Call home"), call routing (e.g., "I would like to make a collect call"), domotic appliance control, search (e.g., find a podcast where particular words were spoken), simple data entry (e.g., entering a credit card number), preparation of structured documents (e.g., a radiology report), speech-to-text processing (e.g., word processors or emails), and aircraft (usually termed Direct Voice Input).

History

The first speech recognizer appeared in 1952 and consisted of a device for the recognition of single spoken digits Another early device was the IBM Shoebox, exhibited at the 1964 New York World's Fair. Lately there have been numerous improvements like a high speed mass transcription capabilitiy on a single system like Sonic Extractor

One of the most notable domains for the commercial application of speech recognition in the United States has been health care and in particular the work of the medical transcriptionist (MT). According to industry experts, at its inception, speech recognition (SR) was sold as a way to completely eliminate transcription rather than make the transcription process more efficient, hence it was not accepted. It was also the case that SR at that time was often technically deficient. Additionally, to be used effectively, it required changes to the ways physicians worked and documented clinical encounters, which many if not all were reluctant to do. The biggest limitation to speech recognition automating transcription, however, is seen as the software. The nature of narrative dictation is highly interpretive and often requires judgment that may be provided by a real human but not yet by an automated system. Another limitation has been the extensive amount of time required by the user and/or system provider to train the software.

A distinction in ASR is often made between "artificial syntax systems" which are usually domain-specific and "natural language processing" which is usually language-specific. Each of these types of application presents its own particular goals and challenges.

Applications

Health care

In the health care domain, even in the wake of improving speech recognition technologies, medical transcriptionists (MTs) have not yet become obsolete. The services provided may be redistributed rather than replaced.

Speech recognition can be implemented in front-end or back-end of the medical documentation process.

Front-End SR is where the provider dictates into a speech-recognition engine, the recognized words are displayed right after they are spoken, and the dictator is responsible for editing and signing off on the document. It never goes through an MT/editor.

Back-End SR or Deferred SR is where the provider dictates into a digital dictation system, and the voice is routed through a speech-recognition machine and the recognized draft document is routed along with the original voice file to the MT/editor, who edits the draft and finalizes the report. Deferred SR is being widely used in the industry currently.

Many Electronic Medical Records (EMR) applications can be more effective and may be performed more easily when deployed in conjunction with a speech-recognition engine. Searches, queries, and form filling may all be faster to perform by voice than by using a keyboard.

Military

High-performance fighter aircraft

Substantial efforts have been devoted in the last decade to the test and evaluation of speech recognition in fighter aircraft. Of particular note are the U.S. program in speech recognition for the Advanced Fighter Technology Integration (AFTI)/F-16 aircraft (F-16 VISTA), the program in France on installing speech recognition systems on Mirage aircraft, and programs in the UK dealing with a variety of aircraft platforms. In these programs, speech recognizers have been operated successfully in fighter aircraft with applications including: setting radio frequencies, commanding an autopilot system, setting steer-point coordinates and weapons release parameters, and controlling flight displays.

Working with Swedish pilots flying in the JAS-39 Gripen cockpit, Englund (2004) found recognition deteriorated with increasing G-loads. It was also concluded that adaptation greatly improved the results in all cases and introducing models for breathing was shown to improve recognition scores significantly. Contrary to what might be expected, no effects of the broken English of the speakers were found. It was evident that spontaneous speech caused problems for the recognizer, as could be expected. A restricted vocabulary, and above all, a proper syntax, could thus be expected to improve recognition accuracy substantially.

The Eurofighter Typhoon currently in service with the UK RAF employs a speaker-dependent system, i.e. it requires each pilot to create a template. The system is not used for any safety critical or weapon critical tasks, such as weapon release or lowering of the undercarriage, but is used for a wide range of other cockpit functions. Voice commands are confirmed by visual and/or aural feedback. The system is seen as a major design feature in the reduction of pilot workload, and even allows the pilot to assign targets to himself with two simple voice commands or to any of his wingmen with only five commands.

Speaker independent systems are also being developed and are in testing for The F35 Lightning II (JSF) and the Aermacchi M346 lead in fighter trainer. These systems have produced word accuracies in excess of 98%.

Helicopters

The problems of achieving high recognition accuracy under stress and noise pertain strongly to the helicopter environment as well as to the fighter environment. The acoustic noise problem is actually more severe in the helicopter environment, not only because of the high noise levels but also because the helicopter pilot generally does not wear a facemask, which would reduce acoustic noise in the microphone. Substantial test and evaluation programs have been carried out in the past decade in speech recognition systems applications in helicopters, notably by the U.S. Army Avionics Research and Development Activity (AVRADA) and by the Royal Aerospace Establishment (RAE) in the UK. Work in France has included speech recognition in the Puma helicopter. There has also been much useful work in Canada. Results have been encouraging, and voice applications have included: control of communication radios; setting of navigation systems; and control of an automated target handover system.

As in fighter applications, the overriding issue for voice in helicopters is the impact on pilot effectiveness. Encouraging results are reported for the AVRADA tests, although these represent only a feasibility demonstration in a test environment. Much remains to be done both in speech recognition and in overall speech recognition technology, in order to consistently achieve performance improvements in operational settings.

Battle management

Battle Management command centres generally require rapid access to and control of large, rapidly changing information databases. Commanders and system operators need to query these databases as conveniently as possible, in an eyes-busy environment where much of the information is presented in a display format. Human-machine interaction by voice has the potential to be very useful in these environments. A number of efforts have been undertaken to interface commercially available isolated-word recognizers into battle management environments. In one feasibility study speech recognition equipment was tested in conjunction with an integrated information display for naval battle management applications. Users were very optimistic about the potential of the system, although capabilities were limited.

Speech understanding programs sponsored by the Defense Advanced Research Projects Agency (DARPA) in the U.S. has focused on this problem of natural speech interface. Speech recognition efforts have focused on a database of continuous speech recognition (CSR), large-vocabulary speech which is designed to be representative of the naval resource management task. Significant advances in the state-of-the-art in CSR have been achieved, and current efforts are focused on integrating speech recognition and natural language processing to allow spoken language interaction with a naval resource management system.

Training air traffic controllers

Training for air traffic controllers (ATC) represents an excellent application for speech recognition systems. Many ATC training systems currently require a person to act as a "pseudo-pilot", engaging in a voice dialog with the trainee controller, which simulates the dialog which the controller would have to conduct with pilots in a real ATC situation. Speech recognition and synthesis techniques offer the potential to eliminate the need for a person to act as pseudo-pilot, thus reducing training and support personnel. In theory, Air controller tasks are also characterized by highly structured speech as the primary output of the controller, hence reducing the difficulty of the speech recognition task should be possible. In practice this is rarely the case. The FAA document 7110.65 details the phrases that should be used by air traffic controllers. While this document gives less than 150 examples of such phrases, the number of phrases supported by one of the simulation vendors speech recognition systems is in excess of 500,000.

The USAF, USMC, US Army,US Navy and FAA as well as a number of international ATC training organizations such as the Royal Australian Air Force and Civil Aviation Authorities in Italy, Brazil, Canada are currently using ATC simulators with speech recognition from a number of different vendors.

Telephony and other domains

ASR in the field of telephony is now commonplace and in the field of computer gaming and simulation is becoming more widespread. Despite the high level of integration with word processing in general personal computing, however, ASR in the field of document production has not seen the expected increases in use.

The improvement of mobile processor speeds made feasible the speech-enabled Symbian and Windows Mobile Smartphones. Speech is used mostly as a part of User Interface, for creating pre-defined or custom speech commands. Leading software vendors in this field are: Microsoft Corporation (Microsoft Voice Command), Nuance Communications (Nuance Voice Control), Vito Technology (VITO Voice2Go), Speereo Software (Speereo Voice Translator), Digital Syphon (Sonic Messager appliance) and SVOX.

Further applications

- Automatic translation;
- Automotive speech recognition (e.g., Ford Sync);
- Telematics (e.g. vehicle Navigation Systems);
- Court reporting (Realtime Voice Writing);
- Hands-free computing: voice command recognition computer user interface;
- Home automation;
- Interactive voice response;
- Mobile telephony, including mobile email;
- Multimodal interaction;
- Pronunciation evaluation in computer-aided language learning applications;
- Robotics;
- Video games, with Tom Clancy's EndWar and Lifeline as working examples;
- Transcription (digital speech-to-text);
- Speech-to-text (transcription of speech into mobile text messages);
- Air Traffic Control Speech Recognition.

Performance

The performance of speech recognition systems is usually specified in terms of accuracy and speed. Accuracy is usually rated with word error rate (WER), whereas speed is measured with the real time factor. Other measures of accuracy include Single Word Error Rate (SWER) and Command Success Rate (CSR).

In 1982, Kurzweil Applied Intelligence and Dragon Systems released speech recognition products. By 1985, Kurzweil's software had a vocabulary of 1,000 words—if uttered one word at a time. Two years later, in 1987, its lexicon reached 20,000 words, entering the realm of human vocabularies, which range from 10,000 to 150,000 words. But recognition accuracy was only 10% in 1993. Two years later, the error rate crossed below 50%. Dragon Systems released "Naturally Speaking" in 1997 which recognized normal human speech. Progress mainly came from improved computer performance and larger source text databases. The Brown Corpus was the first major database available, containing several million words. In 2001, recognition accuracy reached its current plateau of 80%, no longer growing with data or computing power. In 2006, Google published a trillion-word corpus, while Carnegie Mellon University researchers found no significant increase in recognition accuracy.

Algorithms

Both acoustic modeling and language modeling are important parts of modern statistically-based speech recognition algorithms. Hidden Markov models (HMMs) are widely used in many systems. Language modeling has many other applications such as smart keyboard and document classification.

Hidden Markov models

Modern general-purpose speech recognition systems are based on Hidden Markov Models. These are statistical models which output a sequence of symbols or quantities. HMMs are used in speech recognition because a speech signal can be viewed as a piecewise stationary signal or a short-time stationary signal. In a short-time (e.g., 10 milliseconds)), speech can be approximated as a stationary process. Speech can be thought of as a Markov model for many stochastic purposes.

Another reason why HMMs are popular is because they can be trained automatically and are simple and computationally feasible to use. In speech recognition, the hidden Markov model would output a sequence of *n*-dimensional real-valued vectors (with *n* being a small integer, such as 10), outputting one of these every 10 milliseconds. The vectors would consist of cepstral coefficients, which are obtained by taking a Fourier transform of a short time window of speech and decorrelating the spectrum using a cosine transform, then taking the first (most significant) coefficients. The hidden Markov model will tend to have in each state a statistical distribution that is a mixture of diagonal covariance Gaussians which will give a likelihood for each observed vector. Each word, or (for more general speech recognition systems), each phoneme, will have a different output distribution; a hidden Markov model for a sequence of words or phonemes is made by concatenating the individual trained hidden Markov models for the separate words and phonemes.

Described above are the core elements of the most common, HMM-based approach to speech recognition. Modern speech recognition systems use various combinations of a number of standard techniques in order to improve results over the basic approach described above. A typical large-vocabulary system would need context dependency for the phonemes (so phonemes with different left and right context have different realizations as HMM states); it would use cepstral normalization to normalize for different speaker and recording conditions; for further speaker normalization it might use vocal tract length normalization (VTLN) for male-female normalization and maximum likelihood linear regression (MLLR) for more general speaker adaptation. The features would have so-called delta and delta-delta coefficients to capture speech dynamics and in addition might use heteroscedastic linear discriminant analysis (HLDA); or might skip the delta and delta-delta coefficients and use splicing and an LDA-based projection followed perhaps by heteroscedastic linear discriminant analysis or a global semitied covariance transform (also known as maximum likelihood linear transform, or MLLT). Many systems use so-called discriminative training techniques which dispense with a purely statistical approach to HMM parameter estimation and instead optimize some classification-related measure of the training data. Examples are maximum mutual information (MMI), minimum classification error (MCE) and minimum phone error (MPE).

Decoding of the speech (the term for what happens when the system is presented with a new utterance and must compute the most likely source sentence) would probably use the Viterbi algorithm to find the best path, and here there is a choice between dynamically creating a combination hidden Markov model which includes both the acoustic and language model information, or combining it statically beforehand (the finite state transducer, or FST, approach).

Dynamic time warping (DTW)-based speech recognition

Dynamic time warping is an approach that was historically used for speech recognition but has now largely been displaced by the more successful HMM-based approach. Dynamic time warping is an algorithm for measuring similarity between two sequences which may vary in time or speed. For instance, similarities in walking patterns would be detected, even if in one video the person was walking slowly and if in another they were walking more quickly, or even if there were accelerations and decelerations during the course of one observation. DTW has been applied to video, audio, and graphics – indeed, any data which can be turned into a linear representation can be analyzed with DTW.

A well known application has been automatic speech recognition, to cope with different speaking speeds. In general, it is a method that allows a computer to find an optimal match between two given sequences (e.g. time series) with certain restrictions, i.e. the sequences are "warped" non-linearly to match each other. This sequence alignment method is often used in the context of hidden Markov models.

Further information

Popular speech recognition conferences held each year or two include SpeechTEK and SpeechTEK Europe, ICASSP, Eurospeech/ICSLP (now named Interspeech) and the IEEE ASRU. Conferences in the field of Natural language processing, such as ACL, NAACL, EMNLP, and HLT, are beginning to include papers on speech processing. Important journals include the IEEE Transactions on Speech and Audio Processing (now named IEEE Transactions on Audio, Speech and Language Processing), Computer Speech and Language, and Speech Communication. Books like "Fundamentals of Speech Recognition" by Lawrence Rabiner can be useful to acquire basic knowledge but may not be fully up to date (1993). Another good source can be "Statistical Methods for Speech Recognition" by Frederick Jelinek and "Spoken Language Processing (2001)" by Xuedong Huang etc. More up to date is "Computer Speech", by Manfred R. Schroeder, second edition published in 2004. The recently updated textbook of "Speech and Language Processing (2008)" by Jurafsky and Martin presents the basics and the state of the art for ASR. A good insight into the techniques used in the best modern systems can be gained by paying attention to government sponsored evaluations such as those organised by DARPA (the largest speech recognition-related project ongoing as of 2007 is the GALE project, which involves both speech recognition and translation components).

In terms of freely available resources, Carnegie Mellon University's SPHINX toolkit is one place to start to both learn about speech recognition and to start experimenting. Another resource (free as in free beer, not free software) is the HTK book (and the accompanying HTK toolkit). The AT&T libraries GRM library, and DCD library are also general software libraries for large-vocabulary speech recognition.

A useful review of the area of robustness in ASR is provided by Junqua and Haton (1995).

People with disabilities

People with disabilities can benefit from speech recognition programs. Speech recognition is especially useful for people who have difficulty using their hands, ranging from mild repetitive stress injuries to involved disabilities that preclude using conventional computer input devices. In fact, people who used the keyboard a lot and developed RSI became an urgent early market for speech recognition. Speech recognition is used in deaf telephony, such as voicemail to text, relay services, and captioned telephone. Individuals with learning disabilities who have problems with thought-to-paper communication (essentially they think of an idea but it is processed incorrectly causing it to end up differently on paper) can benefit from the software.

Current research and funding

Measuring progress in speech recognition performance is difficult and controversial. Some speech recognition tasks are much more difficult than others. Word error rates on some tasks are less than 1%. On others they can be as high as 50%. Sometimes it even appears that performance is going backwards as researchers undertake harder tasks that have higher error rates.

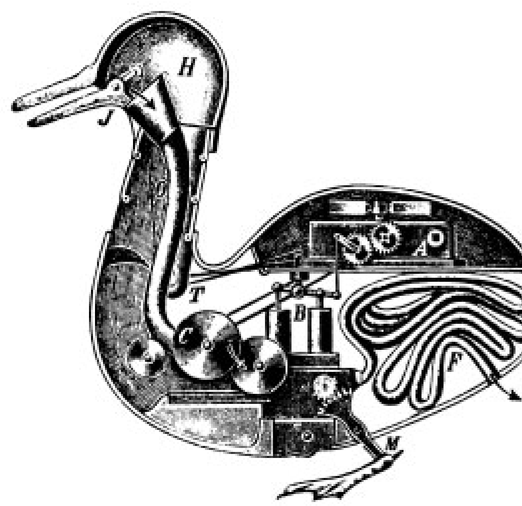
Because progress is slow and is difficult to measure, there is some perception that performance has plateaued and that funding has dried up or shifted priorities. Such perceptions are not new. In 1969, John Pierce wrote an open letter that did cause much funding to dry up for several years. In 1993 there was a strong feeling that performance had plateaued and there were workshops dedicated to the issue. However, in the 1990s funding continued more or less uninterrupted and performance continued slowly but steadily to improve.

For the past thirty years, speech recognition research has been characterized by the steady accumulation of small incremental improvements. There has also been a trend continually to change focus to more difficult tasks due both to progress in speech recognition performance and to the availability of faster computers. In particular, this shifting to more difficult tasks has characterized DARPA funding of speech recognition since the 1980s. In the last decade it has continued with the EARS project, which undertook recognition of Mandarin and Arabic in addition to English, and the GALE project, which focused solely on Mandarin and Arabic and required translation simultaneously with speech recognition.

Commercial research and other academic research also continue to focus on increasingly difficult problems. One key area is to improve robustness of speech recognition performance, not just robustness against noise but robustness against any condition that causes a major degradation in performance. Another key area of research is focused on an opportunity rather than a problem. This research attempts to take advantage of the fact that in many applications there is a large quantity of speech data available, up to millions of hours. It is too expensive to have humans transcribe such large quantities of speech, so the research focus is on developing new methods of machine learning that can effectively utilize large quantities of unlabeled data. Another area of research is better understanding of human capabilities and to use this understanding to improve machine recognition performance.

Chapter- 3

Automaton



The *Digesting Duck* by Jacques de Vaucanson, hailed in 1739 as the first automaton capable of digestion.

An **automaton** (plural: **automata** or **automatons**) is a self-operating machine. The word is sometimes used to describe a robot, more specifically an autonomous robot. An alternative spelling, now obsolete, is **automation**.

Etymology

The word *Automaton* is the latinization of the Greek *αὐτόματον*,... *automaton*, (neuter) “acting of one’s own will”. It is more often used to describe non-electronic moving machines, especially those that have been made to resemble human or animal actions, such as the *jacks* on old public striking clocks, or the cuckoo and any other animated figures on a cuckoo clock.

Ancient automata

The automata in the Hellenistic world were intended as toys, religious idols, or tools for demonstrating basic scientific principles, including those built by Greek mathematician Hero of Alexandria (sometimes known as Heron). When his writings on hydraulics, pneumatics, and mechanics were translated into Latin in the sixteenth century, Hero's readers initiated reconstruction of his machines, which included siphons, a fire engine, a water organ, the aeolipile, and a programmable cart.

Complex mechanical devices are known to have existed in ancient Greece, though the only surviving example is the Antikythera mechanism. It is thought to have come originally from Rhodes, where there was apparently a tradition of mechanical engineering; The island was renowned for its automata; to quote Pindar's seventh Olympic Ode:

> The animated figures stand
> Adorning every public street
> And seem to breathe in stone, or
> move their marble feet.

However, the information gleaned from recent scans of the fragments indicate that it may have come from the colonies of Corinth in Sicily and implies a connection with Archimedes.

There are also examples from myth: Daedalus used quicksilver to install a voice in his statues. Hephaestus created automata for his workshop: Talos, an artificial man of bronze, and, according to Hesiod, the woman Pandora.

According to Jewish tradition, Solomon used his wisdom to design a throne with mechanical animals which hailed him as king when he ascended it; upon sitting down an eagle would place a crown upon his head, and a dove would bring him a Torah scroll.

In ancient China, a curious account on automata is found in the *Lie Zi* text, written in the 3rd century BC. Within it there is a description of a much earlier encounter between King Mu of Zhou (1023-957 BC) and a mechanical engineer known as Yan Shi, an 'artificer'. The latter proudly presented the king with a life-size, human-shaped figure of his mechanical handiwork (Wade-Giles spelling):

The king stared at the figure in astonishment. It walked with rapid strides, moving its head up and down, so that anyone would have taken it for a live human being. The artificer touched its chin, and it began singing, perfectly in tune. He touched its hand, and it began posturing, keeping perfect time...As the performance was drawing to an end, the robot winked its eye and made advances to the ladies in attendance, whereupon the king became incensed and would have had Yen Shih [Yan Shi] executed on the spot had not the latter, in mortal fear, instantly taken the robot to pieces to let him see what it really was. And, indeed, it turned out to be only a construction of leather, wood, glue and lacquer, variously coloured white, black, red and blue. Examining it closely, the king found all the internal organs complete—liver, gall, heart, lungs, spleen, kidneys, stomach and intestines; and over these again, muscles, bones and limbs with their joints, skin, teeth and hair, all of them artificial...The king tried the effect of taking away the heart, and found that the mouth could no longer speak; he took away the liver and the eyes could no longer see; he took away the kidneys and the legs lost their power of locomotion. The king was delighted.

In the mid-8th century, the first wind powered automata were built: "statues that turned with the wind over the domes of the four gates and the palace complex of the Round City of Baghdad". The "public spectacle of wind-powered statues had its private counterpart in the 'Abbasid palaces where automata of various types were predominantly displayed." Also in the 8th century, the Muslim alchemist, Jābir ibn Hayyān (Geber), included recipes for constructing artificial snakes, scorpions, and humans which would be subject to their creator's control in his coded *Book of Stones*. In 827, Caliph Al-Ma'mun had a silver and golden tree in his palace in Baghdad, which had the features of an automatic machine. There were metal birds that sang automatically on the swinging branches of this tree built by Muslim inventors and engineers at the time. The Abbasid Caliph Al-Muqtadir also had a golden tree in his palace in Baghdad in 915, with birds on it flapping their wings and singing. In the 9th century, the Banū Mūsā brothers invented a programmable automatic flute player and which they described in their *Book of Ingenious Devices*.

Automata from the 13th to 19th centuries

Automaton in the Swiss Museum CIMA

Al-Jazari described complex programmable humanoid automata amongst other machines he designed and constructed in the “Book of Knowledge of Ingenious Mechanical Devices” in 1206. His automaton was a boat with four automatic musicians that floated on a lake to entertain guests at royal drinking parties. His mechanism had a programmable drum machine with pegs (cams) that bump into little levers that operate the percussion. The drummer could be made to play different rhythms and different drum patterns if the pegs were moved around. According to Charles B. Fowler, the automata were a "robot band" which performed "more than fifty facial and body actions during each musical selection."

Al-Jazari also invented a hand washing automaton first employing the flush mechanism now used in modern flush toilets. It features a female automaton standing by a basin filled with water. When the user pulls the lever, the water drains and the female automaton refills the basin. His "peacock fountain" was another more sophisticated hand washing device featuring humanoid automata as servants which offer soap and towels. Mark E. Rosheim describes it as follows: "Pulling a plug on the peacock's tail releases water out of the beak; as the dirty water from the basin fills the hollow base a float rises and actuates a linkage which makes a servant figure appear from behind a door under the peacock and offer soap. When more water is used, a second float at a higher level trips and causes the appearance of a second servant figure — with a towel!" Al-Jazari thus appears to have been the first inventor to display an interest in creating human-like machines for practical purposes such as manipulating the environment for human comfort.

Villard de Honnecourt, in his 1230s sketchbook, show plans for animal automata and an angel that perpetually turns to face the sun.

The Chinese author Xiao Xun wrote that when the Ming Dynasty founder Hongwu (r. 1368–1398) was destroying the palaces of Khanbaliq belonging to the previous Yuan Dynasty, there were—amongst many other mechanical devices—automatons found that were in the shape of tigers.

Leonardo da Vinci sketched a more complex automaton around the year 1495. The design of Leonardo's robot was not rediscovered until the 1950s. The robot, which appears in Leonardo's sketches, could, if built successfully, move its arms, twist its head, and sit up.

The Renaissance witnessed a considerable revival of interest in automata. Hero's treatises were edited and translated into Latin and Italian. Numerous clockwork automata were manufactured in the 16th century, principally by the goldsmiths of the Free Imperial Cities of central Europe. These wondrous devices found a home in the cabinet of curiosities or *Wunderkammern* of the princely courts of Europe. Hydraulic and pneumatic automata, similar to those described by Hero, were created for garden grottoes.

A new attitude towards automata is to be found in Descartes when he suggested that the bodies of animals are nothing more than complex machines - the bones, muscles and organs could be replaced with cogs, pistons and cams. Thus mechanism became the standard to which Nature and the organism was compared. France in the 17th century was the birthplace of those ingenious mechanical toys that were to become prototypes for the engines of the Industrial Revolution. Thus, in 1649, when Louis XIV was still a child, an artisan named Camus designed for him a miniature coach, and horses complete with footmen, page and a lady within the coach; all these figures exhibited a perfect movement. According to P. Labat, General de Gennes constructed, in 1688, in addition to machines for gunnery and navigation, a peacock that walked and ate. The Jesuit Athanasius Kircher produced many automatons to create Jesuit shows, including a statue which spoke and listened via a speaking tube.

Tea-serving Japanese automaton, "karakuri ningyō", with mechanism (right), 19th century.

The world's first successfully-built biomechanical automaton is considered to be *The Flute Player*, invented by the French engineer Jacques de Vaucanson in 1737. He also constructed the Digesting Duck, a mechanical duck that gave the false illusion of eating and defecating, seeming to endorse Cartesian ideas that animals are no more than machines of flesh.

In 1769, a chess-playing machine called the Turk, created by Wolfgang von Kempelen, made the rounds of the courts of Europe purporting to be an automaton. The Turk was operated from inside by a hidden human director, and was not a true automaton.

Other 18th century automaton makers include the prolific Frenchman Pierre Jaquet-Droz and his contemporary Henri Maillardet. Maillardet, a Swiss mechanician, created an automaton capable of drawing four pictures and writing three poems. Maillardet's Automaton is now part of the collections at the Franklin Institute Science Museum in Philadelphia. Belgian-born John Joseph Merlin created the mechanism of the Silver Swan automaton, now at Bowes Museum.

According to philosopher Michel Foucault, Frederick the Great, king of Prussia from 1740 to 1786, was "obsessed" with automata. According to Manuel de Landa, "he put together his armies as a well-oiled clockwork mechanism whose components were robot-like warriors."

Japan adopted automata during the Edo period (1603–1867); they were known as *karakuri ningyō*.

The famous magician Jean Eugène Robert-Houdin (1805–1871) was known for creating automata for his stage shows.

The period 1860 to 1910 is known as "The Golden Age of Automata". During this period many small family based companies of Automata makers thrived in Paris. From their workshops they exported thousands of clockwork automata and mechanical singing birds around the world. It is these French automata that are collected today, although now rare and expensive they attract collectors worldwide. The main French makers were Vichy, Roullet & Decamps, Lambert, Phalibois, Renou and Bontems.

Contemporary automata

Contemporary automata continue this tradition with an emphasis on art, rather than technological sophistication. Contemporary automata are represented by the works of Cabaret Mechanical Theatre in the United Kingdom, Dug North and Chomick+Meder, Thomas Kuntz, Arthur Ganson, Joe Jones in the United States, and Le Défenseur du Temps by French artist Jacques Monestier.

An evolution of the mechanized toys developed during the 18th and 19th centuries is represented by automata made with paper. Despite the relative simplicity of the material, paper automata intrinsically are objects with a high degree of technology, where the principles of mechanics meet the artistic creativity.

Other historic examples

Other notable examples of automata include Archytas's dove, mentioned by Aulus Gellius. Similar Chinese accounts of flying automata are written of the 5th century BC Mohist philosopher Mozi and his contemporary Lu Ban, who made artificial wooden birds (*ma yuan*) that could successfully fly according to the *Han Fei Zi* and other texts.

The Smithsonian Institution has in its collection a clockwork monk, about 15 in (380 mm) high, possibly dating as early as 1560. The monk is driven by a key-wound spring and walks the path of a square, striking his chest with his right arm, while raising and lowering a small wooden cross and rosary in his left hand, turning and nodding his head, rolling his eyes, and mouthing silent obsequies. From time to time, he brings the cross to his lips and kisses it. It is believed that the monk was manufactured by Juanelo Turriano, mechanician to the Holy Roman Emperor Charles V.

Chapter- 4

Facial Recognition System

Swiss European surveillance: facial recognition and vehicle make, model, color and license plate reader.

Side View.

Close-up of the infrared illuminator. This light is invisible to the human eye but it creates a day-like environment for the surveillance cameras.

A **facial recognition system** is a computer application for automatically identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a facial database.

It is typically used in security systems and can be compared to other biometrics such as fingerprint or eye iris recognition systems.

Techniques

Traditional

Some facial recognition algorithms identify faces by extracting landmarks, or features, from an image of the subject's face. For example, an algorithm may analyze the relative position, size, and/or shape of the eyes, nose, cheekbones, and jaw. These features are then used to search for other images with matching features. Other algorithms normalize a gallery of face images and then compress the face data, only saving the data in the image that is useful for face detection. A probe image is then compared with the face data. One of the earliest successful systems is based on template matching techniques applied to a set of salient facial features, providing a sort of compressed face representation.

Recognition algorithms can be divided into two main approaches, geometric, which looks at distinguishing features, or photometric, which is a statistical approach that distill an image into values and comparing the values with templates to eliminate variances.

Popular recognition algorithms include Principal Component Analysis with eigenface, Linear Discriminate Analysis, Elastic Bunch Graph Matching fisherface, the Hidden Markov model, and the neuronal motivated dynamic link matching.

3-D

A newly emerging trend, claimed to achieve previously unseen accuracies, is three-dimensional face recognition. This technique uses 3-D sensors to capture information about the shape of a face. This information is then used to identify distinctive features on the surface of a face, such as the contour of the eye sockets, nose, and chin.

One advantage of 3-D facial recognition is that it is not affected by changes in lighting like other techniques. It can also identify a face from a range of viewing angles, including a profile view.

Even a perfect 3D matching technique could be sensitive to expressions. For that goal a group at the Technion applied tools from metric geometry to treat expressions as isometries

Skin texture analysis

Another emerging trend uses the visual details of the skin, as captured in standard digital or scanned images. This technique, called skin texture analysis, turns the unique lines, patterns, and spots apparent in a person's skin into a mathematical space.

Tests have shown that with the addition of skin texture analysis, performance in recognizing faces can increase 20 to 25 percent.

Software

- Google's Picasa digital image organizer has a built in face recognition system starting from version 3.5 onwards. It can associate faces with persons, so that queries can be run on pictures to return all pictures with a specific group of people together. Picasaweb.com has also been providing a similar feature to its users.
- Apple iPhoto, photo organizer distributed with iLife suite of applications includes a system using which people can tag recognized people on photos. Then they can be searched using Spotlight.

- Sony's Picture Motion Browser (PMB) analyses photo, associates photos with identical faces so that they can be tagged accordingly, and differentiates between photos with one person, many persons and nobody.

Notable users and deployments

The London Borough of Newham, in the UK, previously trialled a facial recognition system built into their borough-wide CCTV system.

The German Federal Police use a facial recognition system to allow voluntary subscribers to pass fully automated border controls at Frankfurt Rhein-Main international airport. Subscribers need to be European Union or Swiss citizens. Since 2005 the German Federal Criminal Police Office offers centralized facial recognition on mugshot images for all German police agencies. Recognition systems are also used by casinos to catch card counters and other blacklisted individuals.

The Australian Customs Service has an automated border processing system called SmartGate that uses facial recognition. The system compares the face of the individual with the image in the e-passport microchip, certifying that the holder of the passport is the rightful owner.

Pennsylvania Justice Network searches crime scene photographs and CCTV footage in the mugshot database of previous arrests. A number of cold cases have been resolved since the system became operational in 2005. Other law enforcement agencies in the USA and abroad use arrest mugshot databases in their forensic investigative work.

U.S. Department of State operates one of the largest face recognition systems in the world with over 75 million photographs that is actively used for visa processing.

Spaceship Earth in Epcot uses this[?] for the touch screen part of the ride.

Additional uses

In addition to being used for security systems, authorities have found a number of other applications for facial recognition systems. While earlier post 9/11 deployments were well publicized trials, more recent deployments are rarely written about due to their covert nature.

At Super Bowl XXXV in January 2001, police in Tampa Bay, Florida, used Identix' facial recognition software, FaceIt, to search for potential criminals and terrorists in attendance at the event. (it found 19 people with pending arrest warrants)

In the 2000 presidential election, the Mexican government employed facial recognition software to prevent voter fraud. Some individuals had been registering to vote under several different names, in an attempt to place multiple votes. By comparing new facial images to those already in the voter database, authorities were able to reduce duplicate registrations. Similar technologies are being used in the United States to prevent people from obtaining fake identification cards and driver's licenses.

There are also a number of potential uses for facial recognition that are currently being developed. For example, the technology could be used as a security measure at ATM's; instead of using a bank card or personal identification number, the ATM would capture an image of your face, and compare it to your photo in the bank database to confirm your identity. This same concept could also be applied to computers; by using a webcam to capture a digital image of yourself, your face could replace your password as a means to log-in.

As part of the investigation of the disappearance of Madeleine McCann the British police are calling on visitors to the Ocean Club Resort, Praia da Luz in Portugal or the surrounding areas in the two weeks leading up to the child's disappearance on Thursday 3 May 2007 to provide copies of any photographs of people taken during their stay, in an attempt to identify the abductor using a biometric facial recognition application.

Also, in addition to biometric usages, modern digital cameras often incorporate a facial detection system that allows the camera to focus and measure exposure on the face of the subject, thus guaranteeing a focused portrait of the person being photographed. Some cameras, in addition, incorporate a smile shutter, or take automatically a second picture if someone closed their eyes during exposure.

Comparative study

Among the different biometric techniques, facial recognition may not be the most reliable and efficient. However, one key advantage is that it does not require aid (or consent) from the test subject. Properly designed systems installed in airports, multiplexes, and other public places can identify individuals among the crowd. Other biometrics like fingerprints, iris scans, and speech recognition cannot perform this kind of mass identification. However, questions have been raised on the effectiveness of facial recognition software in cases of railway and airport security.

Criticisms

Weaknesses

Face recognition is not perfect and struggles to perform under certain conditions. Ralph Gross, a researcher at the Carnegie Mellon Robotics Institute, describes one obstacle related to the viewing angle of the face: "Face recognition has been getting pretty good at full frontal faces and 20 degrees off, but as soon as you go towards profile, there've been problems."

Other conditions where face recognition does not work well include poor lighting, sunglasses, long hair, or other objects partially covering the subject's face, and low resolution images.

Another serious disadvantage is that many systems are less effective if facial expressions vary. Even a big smile can render in the system less effective. For instance: Canada now allows only neutral facial expressions in passport photos.

Effectiveness

Critics of the technology complain that the London Borough of Newham scheme has, as of 2004, never recognized a single criminal, despite several criminals in the system's database living in the Borough and the system having been running for several years. "Not once, as far as the police know, has Newham's automatic facial recognition system spotted a live target." This information seems to conflict with claims that the system was credited with a 34% reduction in crime - which better explains why the system was then rolled out to Birmingham also.

An experiment by the local police department in Tampa, Florida, had similarly disappointing results.

"Camera technology designed to spot potential terrorists by their facial characteristics at airports failed its first major test at Boston's Logan Airport"

Privacy concerns

Many citizens are concerned that their privacy will be invaded. Some fear that it could lead to a "total surveillance society," with the government and other authorities having the ability to know where you are, and what you are doing, at all times. This is not to be an underestimated concept as history has shown that states have typically abused such access before.

Recent Improvements

In 2006, the performance of the latest face recognition algorithms were evaluated in the Face Recognition Grand Challenge (FRGC). High-resolution face images, 3-D face scans, and iris images were used in the tests. The results indicated that the new algorithms are 10 times more accurate than the face recognition algorithms of 2002 and 100 times more accurate than those of 1995. Some of the algorithms were able to outperform human participants in recognizing faces and could uniquely identify identical twins.

Low-resolution images of faces can be enhanced using face hallucination. Further improvements in high resolution, megapixel cameras in the last few years have helped to resolve the issue of insufficient resolution.

Early development

Pioneers of Automated Facial Recognition include: Woody Bledsoe, Helen Chan Wolf, and Charles Bisson.

During 1964 and 1965, Bledsoe, along with Helen Chan and Charles Bisson, worked on using the computer to recognize human faces (Bledsoe 1966a, 1966b; Bledsoe and Chan 1965). He was proud of this work, but because the funding was provided by an unnamed intelligence agency that did not allow much publicity, little of the work was published. Given a large database of images (in effect, a book of mug shots) and a photograph, the problem was to select from the database a small set of records such that one of the image records matched the photograph. The success of the method could be measured in terms of the ratio of the answer list to the number of records in the database. Bledsoe (1966a) described the following difficulties:

> "This recognition problem is made difficult by the great variability in head rotation and tilt, lighting intensity and angle, facial expression, aging, etc. Some other attempts at facial recognition by machine have allowed for little or no variability in these quantities. Yet the method of correlation (or pattern matching) of unprocessed optical data, which is often used by some researchers, is certain to fail in cases where the variability is great. In particular, the correlation is very low between two pictures of the same person with two different head rotations."
>
> —Woody Bledsoe, 1966

This project was labeled man-machine because the human extracted the coordinates of a set of features from the photographs, which were then used by the computer for recognition. Using a graphics tablet (GRAFACON or RAND TABLET), the operator would extract the coordinates of features such as the center of pupils, the inside corner of eyes, the outside corner of eyes, point of widows peak, and so on. From these coordinates, a list of 20 distances, such as width of mouth and width of eyes, pupil to pupil, were computed. These operators could process about 40 pictures an hour. When building the database, the name of the person in the photograph was associated with the list of computed distances and stored in the computer. In the recognition phase, the set of distances was compared with the corresponding distance for each photograph, yielding a distance between the photograph and the database record. The closest records are returned.

This brief description is an oversimplification that fails in general because it is unlikely that any two pictures would match in head rotation, lean, tilt, and scale (distance from the camera). Thus, each set of distances is normalized to represent the face in a frontal orientation. To accomplish this normalization, the program first tries to determine the tilt, the lean, and the rotation. Then, using these angles, the computer undoes the effect of these transformations on the computed distances. To compute these angles, the computer must know the three-dimensional geometry of the head. Because the actual heads were unavailable, Bledsoe (1964) used a standard head derived from measurements on seven heads.

After Bledsoe left PRI in 1966, this work was continued at the Stanford Research Institute, primarily by Peter Hart. In experiments performed on a database of over 2000 photographs, the computer consistently outperformed humans when presented with the same recognition tasks (Bledsoe 1968). Peter Hart (1996) enthusiastically recalled the project with the exclamation, "It really worked!"

By about 1997, the system developed by Christoph von der Malsburg and graduate students of the University of Bochum in Germany and the University of Southern California in the United States outperformed most systems with those of Massachusetts Institute of Technology and the University of Maryland rated next. The Bochum system was developed through funding by the United States Army Research Laboratory. The software was sold as ZN-Face and used by customers such as Deutsche Bank and operators of airports and other busy locations. The software was "robust enough to make identifications from less-than-perfect face views. It can also often see through such impediments to identification as mustaches, beards, changed hair styles and glasses—even sunglasses".

In about January 2007, image searches were "based on the text surrounding a photo," for example, if text nearby mentions the image content. Polar Rose technology can guess from a photograph, in about 1.5 seconds, what any individual may look like in three dimensions, and thought they "will ask users to input the names of people they recognize in photos online" to help build a database.

Chapter- 5

Haptic Technology

Haptic technology, or **haptics**, is a tactile feedback technology that takes advantage of a user's sense of touch by applying forces, vibrations, and/or motions to the user. This mechanical stimulation may be used to assist in the creation of virtual objects (objects existing only in a computer simulation), for control of such virtual objects, and for the enhancement of the remote control of machines and devices (teleoperators). It has been described as "(doing) for the sense of touch what computer graphics does for vision". Although haptic devices are capable of measuring bulk or reactive forces that are applied by the user, it should not be confused with touch or tactile sensors that measure the pressure or force exerted by the user to the interface.

Haptic technology has made it possible to investigate in detail how the human sense of touch works by allowing the creation of carefully controlled haptic virtual objects. These objects are used to systematically probe human haptic capabilities, which would otherwise be difficult to achieve. These new research tools contribute to the understanding of how touch and its underlying brain functions work.

The word *haptic*, from the Greek ἁπτικός (*haptikos*), means pertaining to the sense of touch and comes from the Greek verb ἅπτεσθαι *haptesthai* meaning *to contact* or *to touch*.

History

One of the earliest forms of haptic devices is used in large modern aircraft that use servomechanism systems to operate control systems. Such systems tend to be "one-way" in that forces applied aerodynamically to the control surfaces are not perceived at the controls, with the missing normal forces simulated with springs and weights. In earlier, lighter aircraft without servo systems, as the aircraft approached a stall the aerodynamic buffeting was felt in the pilot's controls, a useful warning to the pilot of a dangerous flight condition. This control shake is not felt when servo control systems are used. To replace this missing cue, the angle of attack is measured, and when it approaches the critical stall point a "stick shaker" (an unbalanced rotating mass) is engaged, simulating the effects of a simpler control system. This is known as *haptic feedback*. Alternatively the servo force may be measured and this signal directed to a servo system on the control. This method is known as *force feedback*. Force feedback has been implemented experimentally in some excavators. This is useful when excavating mixed materials such as large rocks embedded in silt or clay, as it allows the operator to "feel" and work around unseen obstacles, enabling significant increases in productivity.

Current applications

Teleoperators and simulators

Teleoperators are remote controlled robotic tools, and when contact forces are reproduced to the operator, it is called "haptic teleoperation". The first electrically actuated teleoperators were built in the 1950s at the Argonne National Laboratory in the United States, by Raymond Goertz, to remotely handle radioactive substances. Since then, the use of "force feedback" has become more widespread in all kinds of teleoperators such as underwater exploration devices controlled from a remote location.

When such devices are simulated using a computer (as they are in operator training devices) it is useful to provide the force feedback that would be felt in actual operations. Since the objects being manipulated do not exist in a physical sense, the forces are generated using haptic (force generating) operator controls. Data representing touch sensations may be saved or played back using such haptic technologies.

Haptic simulators are currently used in medical simulators and flight simulators for pilot training (2004).

Computer and video games

Some simple haptic devices are common in the form of game controllers, in particular of joysticks and steering wheels. At first, such features and/or devices used to be optional components (like the Nintendo 64 controller's *Rumble Pak*). Now many of the newer generation console controllers and some joysticks feature built in devices (such as Sony's DualShock technology). An example of this feature is the simulated automobile steering wheels that are programmed to provide a "feel" of the road. As the user makes a turn or accelerates, the steering wheel responds by resisting turns or slipping out of control. Another concept of force feedback is that of the ability to change the temperature of the controlling device. This would prove especially efficient for prolonged usage of the device. However, due to the high cost of such a technology and the power drainage it would cause, the closest many manufacturers have come to realizing this concept has been to install air holes or small fans into the device to provide the user's hands with ventilation while operating the device.

In 2007, Novint released the Falcon, the first consumer 3D touch device with high resolution three-dimensional force feedback, allowing the haptic simulation of objects, textures, recoil, momentum, physical presence of objects in games.

Mobile consumer technologies

Tactile haptic feedback is becoming common in cellular devices. Handset manufacturers like LG and Motorola are including different types of haptic technologies in their devices. In most cases this takes the form of vibration response to touch. Alpine Electronics uses a haptic feedback technology named *PulseTouch* on many of their touch-screen car navigation and stereo units. The Google Nexus One features "haptic feedback" according to their specifications.

Haptics in virtual reality

Haptics are gaining widespread acceptance as a key part of virtual reality systems, adding the sense of touch to previously visual-only solutions. Most of these solutions use stylus-based haptic rendering, where the user interfaces to the virtual world via a tool or stylus, giving a form of interaction that is computationally realistic on today's hardware. Systems are also being developed to use haptic interfaces for 3D modeling and design that are intended to give artists a virtual experience of real interactive modeling. Researchers from the University of Tokyo have developed 3D holograms that can be "touched" through haptic feedback using "acoustic radiation" to create a pressure sensation on a user's hands. The researchers, led by Hiroyuki Shinoda, currently have the technology on display at SIGGRAPH 2009 in New Orleans.

Research

Some research has been done into simulating the different kinds of tactition by means of high-speed vibrations or other stimuli. One device of this type uses a pad array of pins, where the pins vibrate to simulate a surface being touched. While this does not have a realistic feel, it does provide useful feedback, allowing discrimination between various shapes, textures, and resiliencies.

Several haptics APIs have been developed for research applications, such as Chai3D, OpenHaptics and H3DAPI (Open Source).

Medicine

Various haptic interfaces for medical simulation may prove especially useful for training of minimally invasive procedures (laparoscopy/interventional radiology) and remote surgery using teleoperators. A particular advantage of this type of work is that the surgeon can perform many more operations of a similar type, and with less fatigue. It is well documented that a surgeon who performs more procedures of a given kind will have statistically better outcomes for his patients. Haptic interfaces are also used in Rehabilitation robotics.

In ophthalmology, "haptic" refers to a supporting spring, two of which hold an artificial lens within the lens capsule (after surgical removal of cataracts).

A 'Virtual Haptic Back' (VHB) is being successfully integrated in the curriculum of students at the Ohio University College of Osteopathic Medicine. Research indicates that VHB is a significant teaching aid in palpatory diagnosis (detection of medical problems via touch). The VHB simulates the contour and compliance (reciprocal of stiffness) properties of human backs, which are palpated with two haptic interfaces (SensAble Technologies, PHANToM 3.0).

Robotics

The Shadow Dextrous Robot Hand uses the sense of touch, pressure, and position to reproduce the human grip in all its strength, delicacy, and complexity. The SDRH was first developed by Richard Greenhill and his team of engineers in Islington, London, as part of The Shadow Project, (now known as the Shadow Robot Company) an ongoing research and development program whose goal is to complete the first convincing humanoid. An early prototype can be seen in NASA's collection of humanoid robots, or robonauts. The Dextrous Hand has haptic sensors embedded in every joint and finger pad, which relay information to a central computer for processing and analysis. Carnegie Mellon University in Pennsylvania and Bielefeld University in Germany in particular have found The Dextrous Hand is an invaluable tool in progressing our understanding of haptic awareness and are currently involved (2006) in research with wide ranging implications. The first PHANTOM, which allows one in the human world to interact with objects in virtual reality through touch, was developed by Thomas Massie, while a student of Ken Salisbury at M.I.T.

Arts and design

Touching is not limited to a feeling, but it allows interactivity in real-time with virtual objects. Thus, haptics are commonly used in virtual arts, such as sound synthesis or graphic design/animation. The haptic device allows the artist to have direct contact with a virtual instrument that produces real-time sound or images. For instance, the simulation of a violin string produces real-time vibrations of this string under the pressure and expressiveness of the bow (haptic device) held by the artist. This can be done with physical modelling synthesis.

Designers and modellers may use high-degree of freedom input devices that give touch feedback relating to the "surface" they are sculpting or creating, allowing faster and more natural workflow than with traditional methods.

Actuators

Haptics is enabled by actuators that apply the forces to the skin for touch feedback. The actuator provides mechanical motion in response to an electrical stimulus. Most early designs of haptic feedback use electromagnetic technologies such as vibratory motors with an offset mass, such as the pager motor, that is in most cell phones or voice coils where a central mass or output is moved by a magnetic field. The electromagnetic motors typically operate at resonance and provide strong feedback, but have limited range of sensations. Next-generation actuator technologies are beginning to emerge, offering a wider range of effects thanks to more rapid response times. Next generation haptic actuator technologies include Electroactive Polymers, Piezoelectric, and Electrostatic surface actuation.

Future applications

Future applications of haptic technology cover a wide spectrum of human interaction with technology. Some current research focuses on the mastery of tactile interaction with holograms and distant objects, which, if successful may result in applications and advancements in gaming, movies, manufacturing, medical, and other industries. The medical industry will also gain from virtual and telepresence surgeries, providing new options for medical care. Some speculate the clothing retail industry could gain from haptic technology in ways such as being able to "feel" the texture of clothes for sale on the internet. Future advancements in haptic technology may even create new industries that were not feasible or realistic before the advancements happening right now.

Holographic interaction

Researchers at the University of Tokyo are currently working on adding haptic feedback to holographic projections. The feedback allows the user to interact with a hologram and receive tactile response as if the holographic object were real. The research uses ultrasound waves to create a phenomenon called *acoustic radiation pressure*, which provides tactile feedback as users interact with the holographic object. The haptic technology does not affect the hologram, or the interaction with it, only the tactile response that the user perceives. The researchers posted a video displaying what they call the "Airborne Ultrasound Tactile Display." The technology is not yet ready for mass production or mainstream application in industries, but it is quickly progressing, and "industrial companies" are already showing a positive response to the technology. It is important to note that this example of possible future application is the first in which the user does not have to be outfitted with a special glove or use a special control, they can "just walk up and use [it] " which paints a promising picture for future applications.

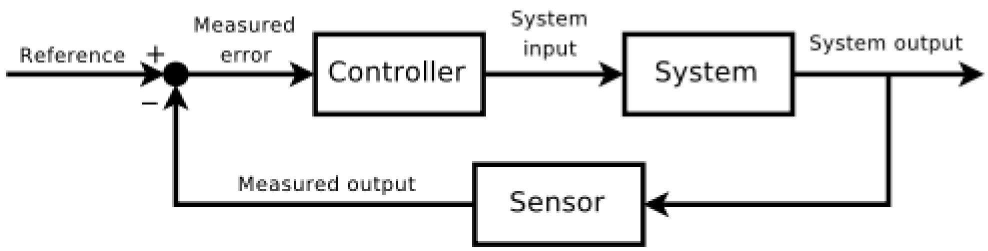
Future medical applications

One currently developing medical innovation is a central workstation surgeons would use to perform operations remotely—local nursing staff would set up the machine and prepare the patient. Rather than travel to an operating room, the surgeon becomes a *telepresence*. This allows expert surgeons to operate from across the country, increasing availability of expert medical care. Haptic technology will provide tactile and resistance feedback to the surgeon as he operates the robotic device. The goal is that, as the surgeon, for instance, makes an incision, he feels ligaments as he would if working directly on the patient.

Surgical training is also on the brink of benefiting from haptic technology. Researchers at Stanford are currently developing technology to simulate surgery for training purposes. Simulated operations would let surgeons and surgical students practice and train more. Haptic technology will aid in the simulation by creating a realistic environment of touch. Much like the telepresence surgery, surgeons will feel simulated ligaments, or the pressure of a virtual incision as if it were real. The researchers led by J. Kenneth Salisbury Jr., a research professor of both computer science and surgery, are also hoping to eventually be able to create realistic internal organs for the simulated surgeries, but, as Salisbury has said, that is not an easy feat. The idea behind the research is that "just as commercial pilots train in flight simulators before they're unleashed on real passengers, surgeons will be able to practice their first incisions without actually cutting anyone."

Chapter- 6

Control Theory



The concept of the feedback loop to control the dynamic behavior of the system: this is negative feedback, because the sensed value is subtracted from the desired value to create the error signal which is amplified by the controller.

Control theory is an interdisciplinary branch of engineering and mathematics, that deals with the behavior of dynamical systems. The desired output of a system is called the *reference*. When one or more output variables of a system need to follow a certain reference over time, a controller manipulates the inputs to a system to obtain the desired effect on the output of the system. A related theory known as perceptual control theory has been used to model living systems on the premise that outputs are manipulated to obtain the desired effect on the input to the system.

Overview

Control theory is

- a theory that deals with influencing the behavior of dynamical systems
- an interdisciplinary subfield of science, which originated in engineering and mathematics, and evolved into use by the social sciences, like psychology, sociology and criminology.

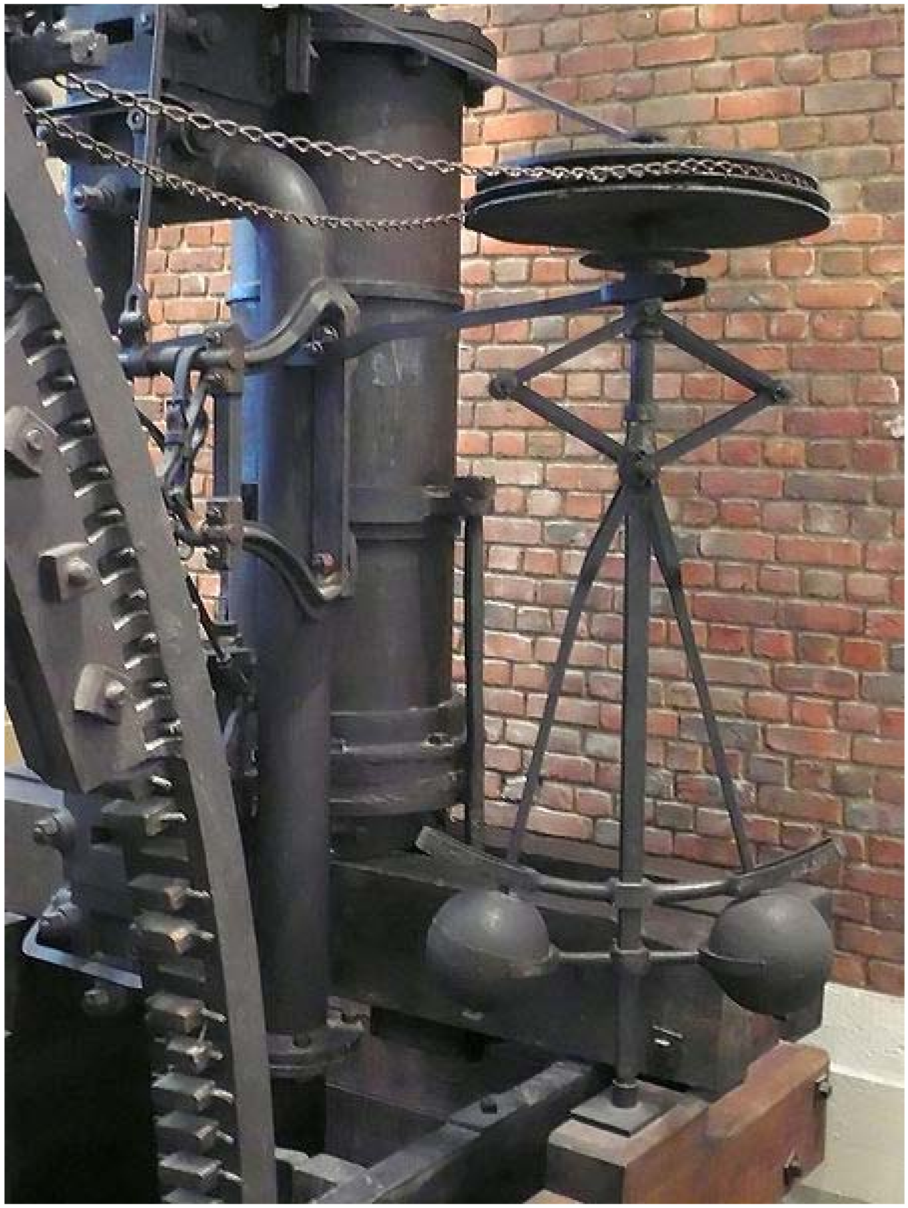
An example

Consider a car's cruise control, which is a device designed to maintain a constant vehicle speed with the *desired* or *reference* speed provided by the driver. The *system* in this case is the vehicle. The system output is the vehicle speed, and the control variable is the engine's throttle position which influences engine torque output.

A primitive way to implement cruise control is simply to lock the throttle position when the driver engages cruise control. However, on mountain terrain, the vehicle will slow down going uphill and accelerate going downhill. In fact, any parameter different from what was assumed at design time will translate into a proportional error in the output velocity, including exact mass of the vehicle, wind resistance, and tire pressure. This type of controller is called an open-loop controller because there is no direct connection between the output of the system (the vehicle's speed) and the actual conditions encountered; that is to say, the system does not and can not compensate for unexpected forces.

In a **closed-loop control system**, a sensor monitors the output (the vehicle's speed) and feeds the data to a computer which continuously adjusts the control input (the throttle) as necessary to keep the control error to a minimum (that is, to maintain the desired speed). Feedback on how the system is actually performing allows the controller (vehicle's on board computer) to dynamically compensate for disturbances to the system, such as changes in slope of the ground or wind speed. An ideal feedback control system cancels out all errors, effectively mitigating the effects of any forces that might or might not arise during operation and producing a response in the system that perfectly matches the user's wishes. In reality, this cannot be achieved due to measurement errors in the sensors, delays in the controller, and imperfections in the control input.

History



Centrifugal governor in a Boulton & Watt engine of 1788

Although control systems of various types date back to antiquity, a more formal analysis of the field began with a dynamics analysis of the centrifugal governor, conducted by the physicist James Clerk Maxwell in 1868 entitled *On Governors*. This described and analyzed the phenomenon of "hunting", in which lags in the system can lead to overcompensation and unstable behavior. This generated a flurry of interest in the topic, during which Maxwell's classmate Edward John Routh generalized the results of Maxwell for the general class of linear systems. Independently, Adolf Hurwitz analyzed system stability using differential equations in 1877, resulting in what is now know as the Routh-Hurwitz theorem.

A notable application of dynamic control was in the area of manned flight. The Wright Brothers made their first successful test flights on December 17, 1903 and were distinguished by their ability to control their flights for substantial periods (more so than the ability to produce lift from an airfoil, which was known). Control of the airplane was necessary for safe flight.

By World War II, control theory was an important part of fire-control systems, guidance systems and electronics. The Space Race also depended on accurate spacecraft control. However, control theory also saw an increasing use in fields such as economics.

People in systems and control

Many active and historical figures made significant contribution to control theory, including, for example:

- Alexander Lyapunov (1857–1918) in the 1890s marks the beginning of stability theory.
- Harold S. Black (1898–1983), invented the concept of negative feedback amplifiers in 1927. He managed to develop stable negative feedback amplifiers in the 1930s.
- Harry Nyquist (1889–1976), developed the Nyquist stability criterion for feedback systems in the 1930s.
- Richard Bellman (1920–1984), developed dynamic programming since the 1940s.
- Andrey Kolmogorov (1903–1987) co-developed the Wiener-Kolmogorov filter (1941).
- Norbert Wiener (1894–1964) co-developed the Wiener-Kolmogorov filter and coined the term cybernetics in the 1940s.
- John R. Ragazzini (1912–1988) introduced digital control and the z-transform in the 1950s.
- Lev Pontryagin (1908–1988) introduced the maximum principle and the bang-bang principle.

Classical control theory

To avoid the problems of the open-loop controller, control theory introduces feedback. A closed-loop controller uses feedback to control states or outputs of a dynamical system. Its name comes from the information path in the system: process inputs (e.g. voltage applied to an electric motor) have an effect on the process outputs (e.g. velocity or torque of the motor), which is measured with sensors and processed by the controller; the result (the control signal) is used as input to the process, closing the loop.

Closed-loop controllers have the following advantages over open-loop controllers:

- disturbance rejection (such as unmeasured friction in a motor)

- guaranteed performance even with model uncertainties, when the model structure does not match perfectly the real process and the model parameters are not exact
- unstable processes can be stabilized
- reduced sensitivity to parameter variations
- improved reference tracking performance

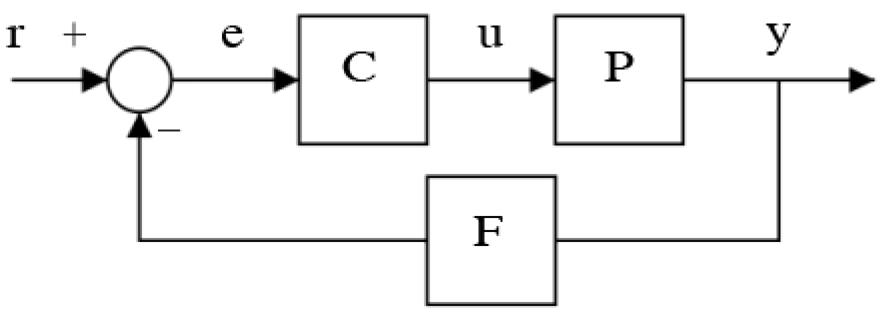
In some systems, closed-loop and open-loop control are used simultaneously. In such systems, the open-loop control is termed feedforward and serves to further improve reference tracking performance.

A common closed-loop controller architecture is the PID controller.

Closed-loop transfer function

The output of the system *y(t)* is fed back through a sensor measurement *F* to the reference value *r(t)*. The controller *C* then takes the error *e* (difference) between the reference and the output to change the inputs *u* to the system under control *P*. This is shown in the figure. This kind of controller is a closed-loop controller or feedback controller.

This is called a single-input-single-output (*SISO*) control system; *MIMO* (i.e. Multi-Input-Multi-Output) systems, with more than one input/output, are common. In such cases variables are represented through vectors instead of simple scalar values. For some distributed parameter systems the vectors may be infinite-dimensional (typically functions).



If we assume the controller *C*, the plant *P*, and the sensor *F* are linear and time-invariant (i.e.: elements of their transfer function *C(s)*, *P(s)*, and *F(s)* do not depend on time), the systems above can be analysed using the Laplace transform on the variables. This gives the following relations:

$$Y(s) = P(s)U(s)$$
$$U(s) = C(s)E(s)$$
$$E(s) = R(s) - F(s)Y(s).$$

Solving for *Y*(*s*) in terms of *R*(*s*) gives:

$$Y(s) = \left(\frac{P(s)C(s)}{1 + F(s)P(s)C(s)} \right) R(s) = H(s)R(s).$$

The expression $H(s) = \frac{P(s)C(s)}{1 + F(s)P(s)C(s)}$ is referred to as the *closed-loop transfer function* of the system. The numerator is the forward (open-loop) gain from *r* to *y*, and the denominator is one plus the gain in going around the feedback loop, the so-called loop gain. If $|P(s)C(s)| \gg 1$, i.e. it has a large norm with each value of *s*, and if $|F(s)| \approx 1$, then *Y(s)* is approximately equal to *R(s)*. This simply means setting the reference to control the output.

PID controller

The PID controller is probably the most-used feedback control design. *PID* is an acronym for *Proportional-Integral-Differential*, referring to the three terms operating on the error signal to produce a control signal. If *u(t)* is the control signal sent to the system, *y(t)* is the measured output and *r(t)* is the desired output, and tracking error $e(t) = r(t) - y(t)$, a PID controller has the general form

$$u(t) = K_P e(t) + K_I \int e(t) \mathrm{d}t + K_D \frac{\mathrm{d}}{\mathrm{d}t} e(t).$$

The desired closed loop dynamics is obtained by adjusting the three parameters K_P, K_I and K_D, often iteratively by "tuning" and without specific knowledge of a plant model. Stability can often be ensured using only the proportional term. The integral term permits the rejection of a step disturbance (often a striking specification in process control). The derivative term is used to provide damping or shaping of the response. PID controllers are the most well established class of control systems: however, they cannot be used in several more complicated cases, especially if MIMO systems are considered.

Applying Laplace transformation results in the transformed PID controller equation

$$u(s) = K_P e(s) + K_I \frac{1}{s} e(s) + K_D s e(s)$$
$$u(s) = (K_P + K_I \frac{1}{s} + K_D s) e(s)$$

with the PID controller transfer function

$$C(s) = (K_P + K_I \frac{1}{s} + K_D s).$$

Modern control theory

In contrast to the frequency domain analysis of the classical control theory, modern control theory utilizes the time-domain state space representation, a mathematical model of a physical system as a set of input, output and state variables related by first-order differential equations. To abstract from the number of inputs, outputs and states, the variables are expressed as vectors and the differential and algebraic equations are written in matrix form (the latter only being possible when the dynamical system is linear). The state space representation (also known as the "time-domain approach") provides a convenient and compact way to model and analyze systems with multiple inputs and outputs. With inputs and outputs, we would otherwise have to write down Laplace transforms to encode all the information about a system. Unlike the frequency domain approach, the use of the state space representation is not limited to systems with linear components and zero initial conditions. "State space" refers to the space whose axes are the state variables. The state of the system can be represented as a vector within that space.)

Topics in control theory

Stability

The *stability* of a general dynamical system with no input can be described with Lyapunov stability criteria. A linear system that takes an input is called bounded-input bounded-output (BIBO) stable if its output will stay bounded for any bounded input. Stability for nonlinear systems that take an input is input-to-state stability (ISS), which combines Lyapunov stability and a notion similar to BIBO stability. For simplicity, the following descriptions focus on continuous-time and discrete-time linear systems.

Mathematically, this means that for a causal linear system to be stable all of the poles of its transfer function must satisfy some criteria depending on whether a continuous or discrete time analysis is used:

- In continuous time, the Laplace transform is used to obtain the transfer function. A system is stable if the poles of this transfer function lie strictly in the open left half of the complex plane (i.e. the real part of all the poles is less than zero).
- In discrete time the Z-transform is used. A system is stable if the poles of this transfer function lie strictly inside the unit circle. i.e. the magnitude of the poles is less than one).

When the appropriate conditions above are satisfied a system is said to be asymptotically stable: the variables of an asymptotically stable control system always decrease from their initial value and do not show permanent oscillations. Permanent oscillations occur when a pole has a real part exactly equal to zero (in the continuous time case) or a modulus equal to one (in the discrete time case). If a simply stable system response neither decays nor grows over time, and has no oscillations, it is marginally stable: in this case the system transfer function has non-repeated poles at complex plane origin (i.e. their real and complex component is zero in the continuous time case). Oscillations are present when poles with real part equal to zero have an imaginary part not equal to zero.

Differences between the two cases are not a contradiction. The Laplace transform is in Cartesian coordinates and the Z-transform is in circular coordinates, and it can be shown that:

- the negative-real part in the Laplace domain can map onto the interior of the unit circle
- the positive-real part in the Laplace domain can map onto the exterior of the unit circle

If a system in question has an impulse response of

$$x[n] = 0.5^n u[n]$$

then the Z-transform (see this example), is given by

$$X(z) = \frac{1}{1 - 0.5z^{-1}}$$

which has a pole in $z = 0.5$ (zero imaginary part). This system is BIBO (asymptotically) stable since the pole is *inside* the unit circle.

However, if the impulse response was

$$x[n] = 1.5^n u[n]$$

then the Z-transform is

$$X(z) = \frac{1}{1 - 1.5z^{-1}}$$

which has a pole at $z = 1.5$ and is not BIBO stable since the pole has a modulus strictly greater than one.

Numerous tools exist for the analysis of the poles of a system. These include graphical systems like the root locus, Bode plots or the Nyquist plots.

Mechanical changes can make equipment (and control systems) more stable. Sailors add ballast to improve the stability of ships. Cruise ships use antiroll fins that extend transversely from the side of the ship for perhaps 30 feet (10 m) and are continuously rotated about their axes to develop forces that oppose the roll.

Controllability and observability

Controllability and observability are main issues in the analysis of a system before deciding the best control strategy to be applied, or whether it is even possible to control or stabilize the system. Controllability is related to the possibility of forcing the system into a particular state by using an appropriate control signal. If a state is not controllable, then no signal will ever be able to control the state. If a state is not controllable, but its dynamics are stable, then the state is termed Stabilizable. Observability instead is related to the possibility of "observing", through output measurements, the state of a system. If a state is not observable, the controller will never be able to determine the behaviour of an unobservable state and hence cannot use it to stabilize the system. However, similar to the stabilizability condition above, if a state cannot be observed it might still be detectable.

From a geometrical point of view, looking at the states of each variable of the system to be controlled, every "bad" state of these variables must be controllable and observable to ensure a good behaviour in the closed-loop system. That is, if one of the eigenvalues of the system is not both controllable and observable, this part of the dynamics will remain untouched in the closed-loop system. If such an eigenvalue is not stable, the dynamics of this eigenvalue will be present in the closed-loop system which therefore will be unstable. Unobservable poles are not present in the transfer function realization of a state-space representation, which is why sometimes the latter is preferred in dynamical systems analysis.

Solutions to problems of uncontrollable or unobservable system include adding actuators and sensors.

Control specifications

Several different control strategies have been devised in the past years. These vary from extremely general ones (PID controller), to others devoted to very particular classes of systems (especially robotics or aircraft cruise control).

A control problem can have several specifications. Stability, of course, is always present: the controller must ensure that the closed-loop system is stable, regardless of the open-loop stability. A poor choice of controller can even worsen the stability of the open-loop system, which must normally be avoided. Sometimes it would be desired to obtain particular dynamics in the closed loop: i.e. that the poles have $Re[\lambda] < -\overline{\lambda}$, where $\overline{\lambda}$ is a fixed value strictly greater than zero, instead of simply asking that $Re[\lambda] < 0$.

Another typical specification is the rejection of a step disturbance; including an integrator in the open-loop chain (i.e. directly before the system under control) easily achieves this. Other classes of disturbances need different types of sub-systems to be included.

Other "classical" control theory specifications regard the time-response of the closed-loop system: these include the rise time (the time needed by the control system to reach the desired value after a perturbation), peak overshoot (the highest value reached by the response before reaching the desired value) and others (settling time, quarter-decay). Frequency domain specifications are usually related to robustness.

Modern performance assessments use some variation of integrated tracking error (IAE,ISA,CQI).

Model identification and robustness

A control system must always have some robustness property. A robust controller is such that its properties do not change much if applied to a system slightly different from the mathematical one used for its synthesis. This specification is important: no real physical system truly behaves like the series of differential equations used to represent it mathematically. Typically a simpler mathematical model is chosen in order to simplify calculations, otherwise the true system dynamics can be so complicated that a complete model is impossible.

System identification

The process of determining the equations that govern the model's dynamics is called system identification. This can be done off-line: for example, executing a series of measures from which to calculate an approximated mathematical model, typically its transfer function or matrix. Such identification from the output, however, cannot take account of unobservable dynamics. Sometimes the model is built directly starting from known physical equations: for example, in the case of a mass-spring-damper system we know that $m\ddot{x}(t) = -Kx(t) - B\dot{x}(t)$. Even assuming that a "complete" model is used in designing the controller, all the parameters included in these equations (called "nominal parameters") are never known with absolute precision; the control system will have to behave correctly even when connected to physical system with true parameter values away from nominal.

Some advanced control techniques include an "on-line" identification process. The parameters of the model are calculated ("identified") while the controller itself is running: in this way, if a drastic variation of the parameters ensues (for example, if the robot's arm releases a weight), the controller will adjust itself consequently in order to ensure the correct performance.

Analysis

Analysis of the robustness of a SISO control system can be performed in the frequency domain, considering the system's transfer function and using Nyquist and Bode diagrams. Topics include gain and phase margin and amplitude margin. For MIMO and, in general, more complicated control systems one must consider the theoretical results devised for each control technique: i.e., if particular robustness qualities are needed, the engineer must shift his attention to a control technique including them in its properties.

Constraints

A particular robustness issue is the requirement for a control system to perform properly in the presence of input and state constraints. In the physical world every signal is limited. It could happen that a controller will send control signals that cannot be followed by the physical system: for example, trying to rotate a valve at excessive speed. This can produce undesired behavior of the closed-loop system, or even break actuators or other subsystems. Specific control techniques are available to solve the problem: model predictive control, and anti-wind up systems. The latter consists of an additional control block that ensures that the control signal never exceeds a given threshold.

System classifications

Linear Systems control

For MIMO systems, pole placement can be performed mathematically using a state space representation of the open-loop system and calculating a feedback matrix assigning poles in the desired positions. In complicated systems this can require computer-assisted calculation capabilities, and cannot always ensure robustness. Furthermore, all system states are not in general measured and so observers must be included and incorporated in pole placement design.

Nonlinear Systems control

Processes in industries like robotics and the aerospace industry typically have strong nonlinear dynamics. In control theory it is sometimes possible to linearize such classes of systems and apply linear techniques, but in many cases it can be necessary to devise from scratch theories permitting control of nonlinear systems. These, e.g., feedback linearization, backstepping, sliding mode control, trajectory linearization control normally take advantage of results based on Lyapunov's theory. Differential geometry has been widely used as a tool for generalizing well-known linear control concepts to the non-linear case, as well as showing the subtleties that make it a more challenging problem.

Decentralized Systems

When the system is controlled by multiple controllers, the problem is one of decentralized control. Decentralization is helpful in many ways, for instance, it helps control systems operate over a larger geographical area. The agents in decentralized control systems can interact using communication channels and coordinate their actions.

Main control strategies

Every control system must guarantee first the stability of the closed-loop behavior. For linear systems, this can be obtained by directly placing the poles. Non-linear control systems use specific theories (normally based on Aleksandr Lyapunov's Theory) to ensure stability without regard to the inner dynamics of the system. The possibility to fulfill different specifications varies from the model considered and the control strategy chosen. Here a summary list of the main control techniques is shown:

Adaptive control
: Adaptive control uses on-line identification of the process parameters, or modification of controller gains, thereby obtaining strong robustness properties. Adaptive controls were applied for the first time in the aerospace industry in the 1950s, and have found particular success in that field.

Hierarchical control
: A Hierarchical control system is a type of Control System in which a set of devices and governing software is arranged in a hierarchical tree. When the links in the tree are implemented by a computer network, then that hierarchical control system is also a form of Networked control system.

Intelligent control
: Intelligent control uses various AI computing approaches like neural networks, Bayesian probability, fuzzy logic, machine learning, evolutionary computation and genetic algorithms to control a dynamic system.

Optimal control
: Optimal control is a particular control technique in which the control signal optimizes a certain "cost index": for example, in the case of a satellite, the jet thrusts needed to bring it to desired trajectory that consume the least amount of fuel. Two optimal control design methods have been widely used in industrial applications, as it has been shown they can guarantee closed-loop stability. These are Model Predictive Control (MPC) and Linear-Quadratic-Gaussian control (LQG). The first can more explicitly take into account constraints on the signals in the system, which is an important feature in many industrial processes. However, the "optimal control" structure in MPC is only a means to achieve such a result, as it does not optimize a true performance index of the closed-loop control system. Together with PID controllers, MPC systems are the most widely used control technique in process control.

Robust control
: Robust control deals explicitly with uncertainty in its approach to controller design. Controllers designed using *robust control* methods tend to be able to cope with small differences between the true system and the nominal model used for design. The early methods of Bode and others were fairly robust; the state-space methods invented in the 1960s and 1970s were sometimes found to lack robustness. A modern example of a robust control technique is H-infinity loop-shaping developed by Duncan McFarlane and Keith Glover of Cambridge University, United Kingdom. Robust methods aim to achieve robust performance and/or stability in the presence of small modeling errors.

Stochastic control

Stochastic control deals with control design with uncertainty in the model. In typical stochastic control problems, it is assumed that there exist random noise and disturbances in the model and the controller, and the control design must take into account these random deviations.

Chapter- 7

BEAM Robotics

The word "beam" in **BEAM robotics** is an acronym for ***B**iology, **E**lectronics, **A**esthetics, and **M**echanics*. This is a term that refers to a style of robotics that primarily uses simple analogue circuits, such as comparators, instead of a microprocessor in order to produce an unusually simple design (in comparison to traditional mobile robots) that trades flexibility for robustness and efficiency in performing the task for which it was designed. Exceptions to the convention of using only analog electronics do exist and these are often colloquially referred to as "mutants". BEAM robots typically consist of a set of the aforementioned analog circuits (mimicking biological neurons) which facilitate the robot's response to its working environment.

Mechanisms and principles

The basic BEAM principles focus on a stimulus-response based ability within a machine. The underlying mechanism was invented by Mark W. Tilden where the circuit (or a Nv net of Nv neurons) is used to simulate biological neuron behaviours. Some similar research was previously done by Ed Rietman in 'Experiments In Artificial Neural Networks'. Tilden's circuit is often compared to a shift register, but with several important features making it a useful circuit in a mobile robot.

Other rules that are included (and to varying degrees applied):

1. Use the lowest number possible of electronic elements (*"keep it simple"*)
2. Recycle and reuse technoscrap
3. Use radiant energy (such as solar power)

There are a large number of BEAM robots designed to use solar power from small solar arrays to power a "Solar Engine" which creates autonomous robots capable of operating under a wide range of lighting conditions. Besides the simplistic computational layer of Tilden's "Nervous Networks", BEAM has brought a multitude of useful tools to the roboticist's toolbox. The "Solar Engine" circuit, many H-bridge circuits for small motor control, tactile sensor designs, and meso-scale (palm-sized) robot construction techniques have been documented and shared by the BEAM community.

BEAM robots

Being focused on "reaction-based" behaviors (as originally inspired by the work of Rod Brooks), BEAM robotics attempts to copy the characteristics and behaviors of natural organisms, with the ultimate goal of domesticating these "wild" robots. BEAM robotics also promotes the value of aesthetics in the design of the device, proving the adage "form follows function".

Disputes in the name

Various people have varying ideas about what BEAM actually stands for. The most widely accepted meaning is ***B**iology, **E**lectronics, **A**esthetics, and **M**echanics*.

This term originated with Mark Tilden during a discussion at the Ontario Science Center in 1990. Mark was displaying a selection of his original bots which he had built while working at the University of Waterloo.

However, there are many other semi-popular names in use, including:

- **B**iotechnology **E**thology **A**nalogy **M**orphology
- **B**uilding **E**volution **A**narchy **M**odularity

Microcontrollers

Unlike many other types of robots controlled by microcontrollers, BEAM robots are built on the principle of using multiple simple behaviours linked directly to sensor systems with little signal conditioning. This design philosophy is closely echoed in the classic book "Vehicles: Experiments in Synthetic Psychology". Through a series of thought experiments, this book explores the development of complex robot behaviours through simple inhibitory and excitory sensor links to the actuators. Microcontrollers and programming are usually not a part of a traditional (aka., "pure") BEAM robot due to the very low-level hardware-centric design philosophy.

There are successful robot designs mating the two technologies. These "hybrids" fulfil a requirement needing robust control systems with the flexibility of dynamic programming, like the "horse-and-rider" topology BEAMbots (eg. the ScoutWalker 3). The physical robot body (the "horse") is controlled by traditional BEAM technology, and the microcontroller and programming influences (and if needed, subsumes) the robot body from the "rider" position . The rider component is not necessary for the robot to function, but without it the robot will lose the important influence of a "smarter brain" telling it what to do.

Types

There are various "-*trope*" BEAMbots, which attempt to achieve a specific goal. Of the series, the phototropes are the most prevalent, as light-seeking would be the most beneficial behaviour for a solar-powered robot.

Audiotrope

In BEAM robotics, an **audiotrope** is a robot that reacts to sounds. This term, which literally means "sound turning," is generally applied to sound-seeking robots. More accurately, audiotropes can either seek (audiophiles) or avoid (audiophobes) sources of sound.

Phototrope

In BEAM robotics, a **Phototrope** is a robot that reacts to light sources. Literally, "light turning," this term is generally (if somewhat inaccurately) applied to light-seeking robots. More accurately, phototropes can either seek (photophiles) or flee (photophobes) bright sources of light.

The simplest and most common form of phototrope is the photopopper—many are as simple as to be essentially two solarrollers stuck together.

One mechanism for phototropism in robotics is implementation of a light sensor where a direct feedback system allows for interaction with the environment. The phototrope analyzes "shots" of its environment and decides whether to move into a certain area depending upon the light intensity.

Alternatively, photovoltaic cells may be used to provide both control and energy for a phototrope. Clever geometry in construction allows for current yielded by a photovoltaic cell to cause motion in the direction of (or away from) the most intense light source in the robots vicinity.

Radiotrope

In BEAM robotics, a **radiotrope** is a robot that reacts to radio frequency sources. Literally, "radio turning," this term is generally (if somewhat inaccurately) applied to radio frequency seeking robots. More accurately, radiotropes can either seek (radiophiles) or flee (radiophobes) sources of radio waves.

Thermotrope

In BEAM robotics, a **thermotrope** is a robot that reacts to heat sources. Literally, "heat turning," this term is generally (if somewhat inaccurately) applied to heat-seeking robots. More accurately, thermotropes can either seek (thermophiles) or flee (thermophobes) sources of heat.

General

BEAMbots have a variety of movements and positioning mechanisms. These include:

Sitter (BEAM)

In BEAM robotics, a **sitter** is a type of robot that does not move. Such robots have no self-contained means of transportation and are non-mobile. Instead of moving, sitters are usually designed to power lights in some sort of pattern. Most sitters are built as decorations rather than serving a practical purpose.

As sitters usually have few parts and are not very complicated to build, many newcomers to BEAM robotics start with sitters, as they are a good place to begin learning the basics of the hobby.

Genera

- Beacons : Transmit a signal (usually a navigational blip) for other BEAMbots to use, often as a navigational device. Essentially, these consist of a solar cell, capacitor, LED(s), and possibly a solar engine. Beacons are often found within robot games and are used by other robots to use as a "base" for navigation.
- Pummers : Display a "light show" or a pattern of sounds. Pummers are often nocturnal robots that store solar energy during the day, then activate during the night.
- Ornaments : A catch-all name for sitters which are not beacons or pummers. Many times, these are mostly electronic art.

Squirmer (BEAM)

In BEAM robotics, a **squirmer** is a stationary robot that performs an interesting action (usually by moving some sort of limbs or appendages).

Slider (BEAM)

In BEAM robotics, a **Slider** is a robot that has a mode of locomotion by moving body parts smoothly along a surface while remaining in contact with it.

Crawler (BEAM)

In BEAM robotics, a **Crawler** is a robot that has a mode of locomotion by tracks or by transferring the robot's body on limbs or appendages. These do not drag parts of their body on the ground.

Jumper (BEAM)

In BEAM robotics, a **Jumper** is a robot that has a mode of locomotion by propelling the robot off the ground and from place to place on the ground.

Roller (BEAM)

In BEAM robotics, a **Roller** is a robot that has a mode of locomotion by rolling all or part of the robot.

Genera

- Symets : Driven mode of locomotion via a single motor. The robot is powered by a solarengine circuit that drives it by charging either three or four capacitors up, then releasing the stored energy in a single 'burst' of power for a period of time, before stopping again to recharge, starting the cycle again. The capacitors are arranged in either a triangular or square pattern along the cylindrical part of the motor. The capacitors are arranged so that they act as rests on which the motor shaft (acting as the 'wheel' for this robot) is balanced out on two other skids. The symmetrical design allows the robot to change direction when it hits an obstacle by tipping onto another set of two side capacitors, which then cause it to move into another direction.

- Solarrollers : Driven mode of locomotion by a single motor; Solar-powered BEAMbots that attempt to complete a fairly short, straight and level course in the shortest amount of time. They are driven by a 'solar engine' circuit; the circuit is designed so that it charges up energy gained from the solar panels and then releases it in one continuous burst. The two main types of solar engines are the FLED engine, which uses a flashing LED to regulate the power, and a 1381 engine, which uses a 1381 voltage regulator to decide when to release the energy.

- Poppers : Driven mode of locomotion by multiple motors; Uses differential sensors achieve a goal.
- Miniballs : Driven mode of locomotion via one or two motors; Entire robot body is caused to move by turning over (on an axis) while on the ground.

Walker (BEAM)

In BEAM robotics, a **walker** is a walking machine that has a driven mode of locomotion by intermittent ground-contacting legs. They usually possess 1 to 12 motors. (ed. motors numbering above 3 are uncommon) "Muscle wired" walkers ultilizes Nitinol (nickel - titanium alloy) for its actuators.

BEAM walking robots are creating using something called NV nets. An NV net consists of NV neurons each of which is a very simple oscillator setup. The most common form of BEAM walker is the master slave Bicore, this uses two suspended Bicore arrangements.

A BEAM walker does not use a processor, nor is it programmed in any way! despite this it is able to walk and respond to terrain via resistive input from its motors. This is an extremely clever method for creating locomotion.

An Example of a two motor walker can be seen at TomboT.net

Genera

- unimotor : Driven mode of locomotion via one motor.
- bimotor : Driven mode of locomotion via 2 motors.
- trimotor : Driven mode of locomotion via 3 motors.
- quadramotor : Driven mode of locomotion via 4 motors.
- pentamotor : Driven mode of locomotion via 5 motors.
- hexamotor : Driven mode of locomotion via 6 motors.
- septuamotor : Driven mode of locomotion via 7 motors.
- octamotor : Driven mode of locomotion via 8 motors.
- ennamotor : Driven mode of locomotion via 9 motors.
- dekamotor : Driven mode of locomotion via 10 motors.
- undecmotor : Driven mode of locomotion via 11 motors.
- dodecamotor : Driven mode of locomotion via 12 motors.
- Muscle wired : Ultilizes Nitinol (nickel - titanium alloy) for its actuators.

Swimmer (BEAM)

In BEAM robotics, a **swimmer** is a robot that functions on or in a liquid environment. These are sometimes referred to as **aquabots** or **aquavores** (the latter being a misnomer, as they are not "water eaters"). Swimmers include various designs, but usually lie within a device class that can be delineated as a watercraft, usually a boat but includes any vehicle designed to move across (or through) water.

Genera

- Boatbots : Self-propelled on the surface of a liquid; construction akin to surface watercrafts.
- Subbots : Self-propelled under the surface of a liquid; construction akin to submarines.

Flier (BEAM)

In BEAM robotics, a **flier** is an aero-robot that functions in an atmospheric environment. They possess a driven mode of locomotion through and/or supported by the atmosphere. Fliers include various designs, but usually lie within a device class that can be delineated as aircraft, which includes any vehicle designed to move through the air (aerodynes or aerostats).

One idea for a solar powered blimp is to use the principle of the Solar balloon, and a standard photopopper circuit.

Genera

- Helicopter : Powered rotor provides both lift and propulsion; Utilizes differential thrust to hop toward brighter areas (aka. "Hoppers")
- Plane : Aircraft that has a fixed wing and is powered by propellers or jets; Usually powered via non-solar power source (such as pneumatic or battery); Solar power for control.
- Blimp : Aircraft that has neutrally-buoyant balloon for lift; Solar power for control and propulsion.

Climber (BEAM)

In BEAM robotics, a **Climber** is a robot that goes upward or downward with gradual or continuous progress on a track (such as a rope or wire).

Applications and current progress

At present, autonomous robots have seen limited commercial application, with some exceptions such as the iRobot Roomba robotic vacuum cleaner and a few lawn-mowing robots. The main practical application of BEAM has been in the rapid prototyping of motion systems and hobby/education applications. Mark Tilden has successfully used BEAM for the prototyping of products for Wow-Wee Robotics, as evidenced by the "proto-Robosapien" "BIODroid" , B.I.O.Bug, and RoboRaptor

Aspiring BEAM roboticists often have problems with the lack of direct control over "pure" BEAM control circuits. There is ongoing work to evaluate Biomorphic techniques that copy natural systems because they seem to have an incredible performance advantage over traditional techniques. There are many examples of how tiny insect brains are capable of far better performance than the most advanced microelectronics.

Another barrier to widespread application of BEAM technology is the perceived random nature of the 'nervous network', which requires new techniques to be learned by the builder to successfully diagnose and manipulate the characteristics of the circuitry. A think-tank of international academics meet annually in Telluride, Colorado to address this issue directly, and until recently, Mark Tilden has been part of this effort (he had to withdraw due to his new commercial commitments with Wow-Wee toys).

Having no long-term memory, BEAM robots generally do not learn from past behaviour. However, there has been work in the BEAM community to address this issue. One of the most advanced BEAM robots in this vein is Bruce Robinson's Hider , which has an impressive degree of capability for a microprocessor-less design.

Chapter- 8

Major Concepts and Applications of Robotics and AI

Actuator

An **actuator** is a mechanical device for moving or controlling a mechanism or system. It is operated by a source of energy, usually in the form of an electric current, hydraulic fluid pressure or pneumatic pressure, and converts that energy into some kind of motion.

Examples and applications

- Mechanical actuators operate by conversion of rotary motion into linear motion, or vice versa. Conversion is commonly made via a few simple types of mechanism including:
    - **Screw**: Screw jack, ball screw and roller screw actuators all operate on the principle of the simple machine known as the screw. By rotating the actuator's nut, the screw shaft moves in a line. By moving the screw shaft, the nut rotates.
    - **Wheel and axle**: Hoist, winch, rack and pinion, chain drive, belt drive, rigid chain and rigid belt actuators operate on the principle of the wheel and axle. By rotating a wheel/axle (e.g. drum, gear, pulley or shaft) a linear member (e.g. cable, rack, chain or belt) moves. By moving the linear member, the wheel/axle rotates.
- In engineering, actuators are frequently used as mechanisms to introduce motion, or to clamp an object so as to prevent motion. In electronic engineering, actuators are a subdivision of transducers. They are devices which transform an input signal (mainly an electrical signal) into motion. Specific examples include: electrical motors, pneumatic actuators, hydraulic actuators, linear actuators, comb drive,

piezoelectric actuators and amplified piezoelectric actuators, thermal bimorphs, micromirror devices and electroactive polymers.

- Motors are mostly used when circular motions are needed, but can also be used for linear applications by transforming circular to linear motion with a bolt and screw transducer. On the other hand, some actuators are intrinsically linear, such as piezoelectric actuators.

Thermodynamic efficiency

The efficiency of an actuator is a standard tool used to calculate or estimate the usefulness of any actuating mechanism. It is a dimensionless quantity which is generally lower than 1 expressing the energy conversion factor. Most of the wasted energy (due to friction, magnetic losses, eddy currents etc.) is thermally dissipated.

$$\varepsilon = \frac{\text{useful work}}{\text{spent energy}} = \frac{\text{output energy}}{\text{input energy}}$$

Adaptable robotics

Adaptable robotics are generally based in robot developer kits.

The kits come with an open software platform tailored to a range of common robotic functions.

The kits also come with common robotics hardware that connects easily with the software (infrared sensors, motors, microphone and video camera).

Anthrobotics

Anthrobotics is the science of developing and studying robots that are either entirely or in some way human-like.

The term *anthrobotics* was originally coined by Mark Rosheim in a paper entitled "Design of An Omnidirectional Arm" presented at the IEEE International Conference on Robotics and Automation, May 13–18, 1990, pp. 2162–2167. Rosheim says he derived the term from "...Anthropomorphic and Robotics to distinguish the new generation of dexterous robots from its simple industrial robot forbears." The word gained wider recognition as a result of its use in the title of Rosheim's subsequent book *Robot Evolution: The Development of Anthrobotics*, which focussed on facsimiles of human physical and psychological skills and attributes.

However, a wider definition of the term *anthrobotics* has been proposed, in which the meaning is derived from *anthropology* rather than *anthropomorphic*. This usage includes robots that respond to input in a human-like fashion, rather than simply mimicking human actions, thus theoretically being able to respond more flexibly or to adapt to unforeseen circumstances. This expanded definition also encompasses robots that are situated in social environments with the ability to respond to those environments appropriately, such as insect robots, robotic pets, and the like.

Anthrobotics is now taught at some universities, encouraging students not only to design and build robots for environments beyond current industrial applications, but also to speculate on the future of robotics that are embedded in the world at large, as mobile phones and computers are today.

Astrochicken

Astrochicken is the name given to a thought experiment expounded by theoretical physicist Freeman Dyson. In his book *Disturbing the Universe* (1979), Dyson contemplated how humanity could build a small, self-replicating automaton that could explore space more efficiently than a manned craft could. He attributed the general idea to John von Neumann, based on a lecture von Neumann gave in 1948 entitled *The General and Logical Theory of Automata*. Dyson expanded on von Neumann's automata theories and added a biological component to them.

Astrochicken, Dyson explained, would be a one-kilogram spacecraft unlike any before it. It would be a creation of the intersection of biology, artificial intelligence and modern microelectronics—a blend of organic and electronic components. Astrochicken would be launched by a conventional spacecraft into space, like an egg being laid into space. Astrochicken would then hatch and start growing a solar energy collector. The solar collector would feed an ion drive engine that would power the craft. Once Astrochicken entered a planet's vicinity, it would collect material from the moons and rings of the planet, taking in nutrients. It could land and take off using an auxiliary chemical rocket similar to that used by bombardier beetles. It would periodically transmit details of its journey when it could make radio contact with Earth.

The term "astrochicken" does not occur in Dyson's earliest essays regarding von Neumann-inspired automata. When Dyson was giving a lecture in Adelaide, Australia on the subject of space exploration with biotechnology, an audience member called out "Oh, you mean this is an astro-chicken." The whimsical name caught on, and Dyson began to use it himself in subsequent essays he wrote on his theoretical biotechnology spacecraft.

Today, Dyson's Astrochicken resonates with several theories of how space exploration might proceed in the future. Computer scientist Rodney Brooks has proposed sending a multitude of cheap, bug-like robots to explore Mars instead of solitary, expensive rovers. Cheaper and smaller means of studying space have also been the primary design philosophy of NASA for many years, perhaps best exemplified by the Mars Pathfinder mission. Physicist and noted author Michio Kaku wrote in his work *Hyperspace*, "Small, lightweight, and intelligent, Astrochicken is a versatile space probe that has a clear advantage over the bulky, exorbitantly expensive space missions of the past, which have been a bottleneck to space exploration. ... It will not need huge quantities of rocket fuel; it will be bred and programmed to 'eat' ice and hydrocarbons found in the rings surrounding the outer planets".

In recent years, Dyson has referred to Astrochicken as a "joke", though it is not quite certain what he means by this. He went on to say "I think it's a sensible idea, but one shouldn't take it literally. We don't have the science yet; we don't have the technology. It would be a disaster if NASA tried to do this in the bureaucratic NASA style."

As a noted author of essays on the possibilities of science in the future, Dyson's theories, such as the Dyson sphere and the Dyson tree, have become popular in the scientific and science fiction communities. The more whimsically named "Astrochicken" has not achieved this same level of fame, although the mixture of biology and technology is now a popular scientific subject.

Autonomous automation

Autonomous automation refers to 'the use of autonomous software agents to adapt the controllers of computer controlled industrial machinery and processes'

The term *autonomous automation* has in the past, on a limited number of web-sites, been used mainly to refer to 'the use of autonomous computer controlled industrial machinery and processes'. Since automation implies autonomy to a great extent this usage can be considered as somewhat redundant.

Autonomous research robot

The Denning Mobile Robot company of Boston was the first to offer ready-made autonomous robots, which were purchased primarily by researchers. Grinnell More's Real World Interface, Inc. (RWI) and James Slater's Nomadic Technologies in the US and Francesco Mondada's K-Team in Switzerland were also among the pioneers to address the need for ready-made robots for robotics researchers with the B-21 from RWI, XR4000 from Nomadic and tiny Khepera mobile robot from K-Team; however, prices meant that only a few graduate students and military researchers could afford them. The low-cost Pioneer robot was introduced by a collaboration between RWI and ActivMedia Robotics in 1995, making robots available to many and opening the floodgates to new research in mobile robotics.

By 1999, the Denning company was defunct. In 1998, RWI joined with ISRobotics to form the iRobot corporation. There Grinnell More introduced the PackBot remote control robot, veering away from autonomous development and research robots to pursue military venues. Nomadic Technologies also left the field. MobileRobots Inc and K-Team continued to build on autonomy and provide for the research community.

In 2003 the Defense Advanced Research Projects Agency (DARPA) contracted with Segway to convert fifteen Segway PTs into Segway Robotic Mobility Platforms. Segway developed the platform to serve as a reliable, cost-effective tool for research institutions and delivered the units to DARPA in April. In June 2003 DARPA worked with SPAWAR Systems Center, San Diego to distribute the units to 14 government and university research institutions to use in robotic research projects..

Autonomous navigation techniques

An ActivMedia Pioneer 3-AT robot at the Georgia Institute of Technology

Research robots made great strides in autonomous indoor navigation between the 1990s and the 2000s. Currently, a number of ready-made research bases have the sensing, mobility, and computational power necessary for such autonomy: The Pioneer, PatrolBot, PowerBot and PeopleBot platforms can map buildings and navigate out-of-the box, using

SLAM and a variation on Monte Carlo method/Markov localization and modified value-iterated search navigation techniques, with any sensor of the 2-D range-finder class. This method creates a human readable map of the robot's workspace that can be used to control and track robots of this type as they move. Evolution Robotics offers single-camera VSLAM software, which replaces range-finding with visual pattern-matching, but this system cannot create a human readable map with which to monitor robots' position. Other groups are building stereocam-based VSLAM. Because the stereocam provides range-finding data using the disparity between the lenses, maps can be made and robots tracked. The K-Team Khepera, Segway-based platforms and other research robots can link to external computing resources to use such software.

The precision of any of these methods depends upon the precision of the sensor, the granularity of data tracked and the speed of calculation. Range-finding lasers may have +/-1 cm accuracy while digital stereocamera accuracy is limited to a quarter pixel and thus is highly range-dependent. Vision-based systems require more computational resources than simple range-finding systems such as lasers, but may do the computation on a digital signal processor embedded with the camera. Because of cost and precision trade-offs, less expensive vision-based systems tend to be used on consumer robots while commercial and industrial robots and automated guided vehicles (AGVs) tend to use laser-based systems.

Outdoors, localization is primarily handled with GPS, however, satellite signals can frequently be lost due to weather, trees, buildings or other obstructions. When the signal is lost, the robot typically navigates using dead reckoning and inertial motion tracking. Dead reckoning relies on relative wheel motion and is highly subject to cumulative slippage errors. Inertial motion tracking uses rate gyroscopes and accelerometers to determine actual motion of the platform. The accuracy of inertial motion tracking depends upon the quality and calibration of the sensors employed. The Segway RMP 400 and Seekur robots are two of the few research platforms designed for such research; most other outdoor research robots are jerry-rigged by researchers from existing vehicles.

In constrained areas, some robots, such as the John Deere Gator, simply surround the perimeter with radio beacons and use simple triangulation from three or more beacons to localize and navigate. Beacons are also used indoors by older AGVs in factories.

Autonomous Solutions is a leader in the field of outdoor navigation software; their system is used by John Deere tractors and by some military platforms.

Programming research robots

Much research software for autonomous robots is Free Software or Open Source Software, including Carmen from Carnegie Mellon, Player/Stage/Gazebo from the University of Southern California and the ARIA API libraries from MobileRobots, Inc. There is also commercial software: Webots has been continuously developed since 1998 and is currently used by more than 500 universities. It runs on Linux, Windows and Mac OS X. More recently, in June 2006, Microsoft Research began offering free beta-test copies of a Robotics Studio software development kit with Pioneer robots in simulation for Windows XP in an attempt to counter Linux dominance onboard mobile robot platforms. An older platform: URBI with a Free Software SDK is used in many universities. The plethora of autonomous mobile robots and software available for researchers has greatly sped the pace of development in the robotics field.

Behavior-based AI

Behavior Based Artificial Intelligence (BBAI) is a methodology for developing AI based on a modular decomposition of intelligence. It was made famous by Rodney Brooks and his subsumption architecture was one of the earliest attempts to describe a mechanism for developing BBAI. It is extremely popular in robotics and to a lesser extent intelligent virtual agents because it allows the successful creation of real-time dynamic systems that can run in complex environments. For example, it underlies the intelligence of the Sony Aibo and many RoboCup robot teams.

The most important attribute of a behavior based system is that the intelligence is controlled by a set of independent semi-autonomous modules. In the original systems, each module was actually a separate device or was at least conceived of as running on its own processing thread. Generally, though, the modules are just abstractions. BBAI may be seen as a software engineering approach to AI, perhaps akin to object oriented design.

BBAI is often associated with reactive planning, but the two are not synonymous. Brooks advocated an extreme version of cognitive minimalism which required initially that the behavior modules were finite state machines and thus contained no conventional memory or learning. This is associated with reactive AI because reactive AI requires reacting to the current state of the world, not to an agent's memory or preconception of that world. However, learning is obviously key to realistic strong AI, so this constraint has been relaxed, though not entirely abandoned.

Researchers currently active in this area include Ronald Arkin (who has written a book called *Behavior Based Robotics*) and Maja Mataric. Former researchers include Brooks, Pattie Maes and Bruce Blumberg. It was probably at least partially inspired by Marvin Minsky's Society of Mind.

Behavior-based robotics

Behavior-based robotics or **behavioral robotics** is the branch of robotics that incorporates modular or behavior based AI (BBAI).

How they work

Most behavior-based systems are also reactive, which means they use relatively little internal variable state to model the environment. For instance, there is no programming in the robot of what a chair looks like, or what kind of surface the robot is moving on - all the information is gleaned from the input of the robot's sensors. The robot uses that information to react to the changes in its environment.

Behavior-based robots (BBR) usually show more biological-appearing actions than their computing-intensive counterparts, which are very deliberate in their actions. A BBR often makes mistakes, repeats actions, and appears confused, but can also show the anthropomorphic quality of tenacity. Comparisons between BBRs and insects are frequent because of these actions. BBRs are sometimes considered examples of Weak artificial intelligence, although some have claimed they are models of all intelligence (Brooks 1991).

History

The school of behavior-based robots owes much to work undertaken in the 1980s at the Massachusetts Institute of Technology by Professor Rodney Brooks, who with students and colleagues built a series of wheeled and legged robots utilising the subsumption architecture. Brooks' papers, often written with lighthearted titles such as "*Planning is just a way of avoiding figuring out what to do next*", the anthropomorphic qualities of his robots, and the relatively low cost of developing such robots, popularized the behavior-based approach.

Brooks' work builds - whether by accident or not - on two prior milestones in the behavior-based approach. In the 1950s, W. Grey Walter, an English scientist with a background in neurological research, built a pair of vacuum tube-based robots that were exhibited at the 1951 Festival of Britain, and which have simple but effective behavior-based control systems.

The second milestone is Valentino Braitenberg's 1984 book, "*Vehicles - Experiments in Synthetic Psychology*" (MIT Press). He describes a series of thought experiments demonstrating how simply wired sensor/motor connections can result in some complex-appearing behaviors such as fear and love.

Later work in BBR is from the BEAM robotics community, which has built upon the work of Mark Tilden. Tilden was inspired by the reduction in the computational power needed for walking mechanisms from Brooks' experiments (which used one microcontroller for each leg), and further reduced the computational requirements to that of logic chips, transistor-based electronics, and analog circuit design.

A different direction of development includes extensions of behavior-based robotics to multi-robot teams. The focus in this work is on developing simple generic mechanisms that result in coordinated group behavior, either implicitly or explicitly.

leJOS, a substitute firmware and additional library for programming Lego Mindstorms, provides Java classes that support behavior based robotics.

Embodied cognitive science

Embodied Cognitive Science is an interdisciplinary field of research, the aim of which is to explain the mechanisms underlying intelligent behavior. It comprises three main methodologies: 1) the modeling of psychological and biological systems in a holistic manner that considers the mind and body as a single entity, 2) the formation of a common set of general principles of intelligent behavior, and 3) the experimental use of robotic agents in controlled environments.

Embodied cognitive science borrows heavily from embodied philosophy and the related research fields of cognitive science, psychology, neuroscience and artificial intelligence. From the perspective of neuroscience, research in this field was led by Gerald Edelman of the Neurosciences Institute at La Jolla, the late Francisco Varela of CNRS in France, and J. A. Scott Kelso of Florida Atlantic University. From the perspective of psychology, research by Michael Turvey and Eleanor Rosch. From the perspective of language acquisition, Eric Lenneberg and Philip Rubin at Haskins Laboratories. From the perspective of autonomous agent design, early work is sometimes attributed to Rodney Brooks or Valentino Braitenberg. From the perspective of artificial intelligence, see *Understanding Intelligence* by Rolf Pfeifer and Christian Scheier or *How the body shapes the way we think*, also by Rolf Pfeifer and Josh C. Bongard.

Turing proposed that a machine may need a human-like body to think and speak:

It can also be maintained that it is best to provide the machine with the best sense organs that money can buy, and then teach it to understand and speak English. That process could follow the normal teaching of a child. Things would be pointed out and named, etc. Again, I do not know what the right answer is, but I think both approaches should be tried (Turing, 1950).

General principles of intelligent behavior

In the formation of general principles of intelligent behavior, Pfeifer intended to be contrary to older principles given in Traditional Artificial Intelligence. The most dramatic difference is that the principles are applicable only to situated robotic agents in the real world, a domain where Traditional Artificial Intelligence showed the least promise.

Principle of Cheap Design and Redundancy :: Pfeifer realized that implicit assumptions made by engineers often substantially influence a control architecture's complexity. This insight is reflected in discussions of the scalability problem in robotics. The internal processing needed for some bad architectures can grow out of proportion to new tasks needed of an agent.

One of the primary reasons for scalability problems is that the amount of programming and knowledge engineering that the robot designers have to perform grows very rapidly with the complexity of the robot's tasks. There is mounting evidence that pre-programming cannot be the solution to the scalability problem ... The problem is that programmers introduce too many hidden assumptions in the robot's code.

The proposed solutions are as follows, to have the agent exploit the inherent physics of its environment, to exploit the constraints of its niche, and to have agent morphology based on parsimony and the principle of Redundancy. Redundancy reflects the desire for the error-correction of signals afforded by duplicating like channels. Additionally, it reflects the desire to exploit the associations between sensory modalities. In terms of design, this implies that redundancy should be introduced with respect not only to one sensory modality but to several. It has been suggested that the fusion and transfer of knowledge between modalities can be the basis of reducing the size of the sense data taken from the real world. This again addresses the scalability problem.

Principle of Parallel, Loosely-coupled Processes :: An alternative to hierarchical methods of knowledge and action selection. This design principle differs most importantly from the Sense-Think-Act cycle of traditional AI. Since it does not involve this famous cycle, it is not affected by the Frame problem.

Principle of Sensory-Motor Coordination :: Ideally, internal mechanisms in an agent should give rise to things like memory and choice-making in an emergent fashion, rather than being prescriptively programmed from the beginning. These kinds of things are allowed to emerge as the agent interacts with the environment. The motto is, build fewer assumptions into the agent's controller now, so that learning can be more robust and idiosyncratic in the future.

Principle of Ecological Balance :: This is more a theory than a principle, but its implications are widespread. Its claim is that the internal processing of an agent cannot be made more complex unless there is a corresponding increase in complexity of the motors, limbs, and sensors of the agent. In other words, the extra complexity added to the brain of a simple robot will not create any discernible change in its behavior. The robot's morphology must already contain the complexity in itself to allow enough "breathing room" for more internal processing to develop.

The Value Principle :: This was the architecture developed in the Darwin III robot of Gerald Edelman. It relies heavily on connectionism.

Developmental robotics

Developmental Robotics (**DevRob**), sometimes called epigenetic robotics, is a methodology that uses metaphors from neural development and developmental psychology to develop the mind for autonomous robots. The focus is on a single or multiple robots going through stages of **autonomous mental development** (AMD).

Researchers in this field study artificial emotions, self-motivation, and other methods of self-organization. The program that simulates the functions of genome to develop a robot's mental capabilities is called a developmental program.

Different from traditional machine learning, some major features of developmental robotics are:

Task-nonspecificity: Since it is difficult for the genome to predict what tasks the baby will learn and perform in his life, the developmental program is body-specific (species specific) but not task-specific.

Environmental openness: Due to the task-nonspecificity, AMD must deal with unknown and uncontrolled environments, including various human environments.

Raw sensors: AMD must directly deal with continuous raw signals from sensors (e.g., vision, audition and touch), since different tasks require different information in the sensors. Only raw signals have all.

Online processing: At each time instant, what the machine will sense next depends on what the machine does now.

Incremental processing: Acquired skills must be used to assist in the acquisition of new skills, as a form of **scaffolding.** This requires incremental processing.

DevRob is related to, but differs from, evolutionary robotics (**ER**). ER uses populations of robots that evolve over time, whereas DevRob is interested in how the organization of a single robot's control system develops through experience, over time.

DevRob is also related to work done in the domains of Robotics, Artificial Life.

History

Cresceptron (ICCV 1992, IJCV 1997) was the first published developmental learning method for detecting and recognizing a general object in a complex natural background based on a composite image view. It also segments the detected object from the complex natural background. Human-machine interactions through the sensory-end and the motor-end teach the Cresceptron, while the internal self-organization is fully autonomous.

The NSF/DARPA funded Workshop on Development and Learning was held April 5-7, 2000 at Michigan State University. It was the first international meeting devoted to computational understanding of mental development by robots and animals. The term "by" was used since the agents are active during development.

DevRob was explained in Weng et al. *Autonomous mental development by robots and animals. Science 291:599-600, 2001.* Its major uniqueness is the task nonspecificity concept of a new kind of program: developmental program (DP). A DP simulates the developmental functions of the "genome".

The first undergraduate courses in DevRob were offered at Bryn Mawr College and Swarthmore College in the Spring of 2003 by Douglas Blank and Lisa Meeden, respectively.

The first graduate course in DevRob was offered at Iowa State University by Alexander Stoytchev in the Fall of 2005.

Frankenstein complex

In Isaac Asimov's robot novels, the ***Frankenstein complex*** is a term that he coined for the fear of mechanical men.

History

Some of Asimov's S.F. short stories and novels predict that this phobia will become strongest and most widespread when being directed against "mechanical men" that most-closely resemble human beings, but it is also present on a lower level against robots that are plainly electromechanical automatons. The "Frankenstein Complex" is similar in many respects to Masahiro Mori's uncanny valley hypothesis.

The name, "Frankenstein Complex", derives from the name of Victor Frankenstein in the groundbreaking novel, *Frankenstein; or, The Modern Prometheus*, written by Mary Wollstonecraft Shelley in about the year 1818. In Ms. Shelley's story, Frankenstein created an intelligent, somewhat superhuman being. He finds that his creation is horrifying to behold, and he abandons it. This ultimately leads to Victor's death at the conclusion of a vendetta between himself and his embittered creation.

Note the distinction between "Frankenstein" the creator and *Frankenstein's monster*: a *Frankenstein complex* is not a fear of roboticists, scientists, or even mad scientists, but rather, a fear of artificial human beings, although fear of one generally implies some fear of the other.

The general attitude of the public towards robots in much of Dr. Asimov's fiction is fear and suspicion: ordinary people fear that robots will either replace them or dominate them. Although dominance is impossible under the specifications of the Three Laws of Robotics, which state clearly,

A robot may not harm a human being or, through inaction, allow a human being to come to harm,

the fictitious earthly public does not generally listen to this logic, but rather they listen to their fears. In I, Robot's short story "Little Lost Robot" is an example of the "fear of robots" that Asimov described.

In Asimov's robot novels, the *Frankenstein Complex* is a major problem for roboticists and robot manufacturers. They do all they can to calm the public and show that robots are harmless, sometimes even hiding the truth because the public would misunderstand it and take it to the extreme. The fear by the public and the response of the manufacturers is an example of the theme of paternalism, the dread of paternalism, and the conflicts that arise from it in Asimov's fiction.

Human-robot interaction

Human-robot interaction is the study of interactions between humans and robots. It is often referred as HRI by researchers. Human-Robot interaction is a multidisciplinary field with contributions from human-computer interaction, artificial intelligence, robotics, natural language understanding, and social sciences.

Origins

Human-robot interaction has been a topic of both science fiction and academic speculation even before any robots existed. Because HRI depends on a knowledge of (sometimes natural) human communication, many aspects of HRI are continuations of human communications topics that are much older than robotics per se.

The origin of HRI as a discrete problem was stated by 20th-century author Isaac Asimov in 1941, in his novel *I, Robot*. He states the Three Laws of Robotics as,

“

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey any orders given to it by human beings, except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.

”

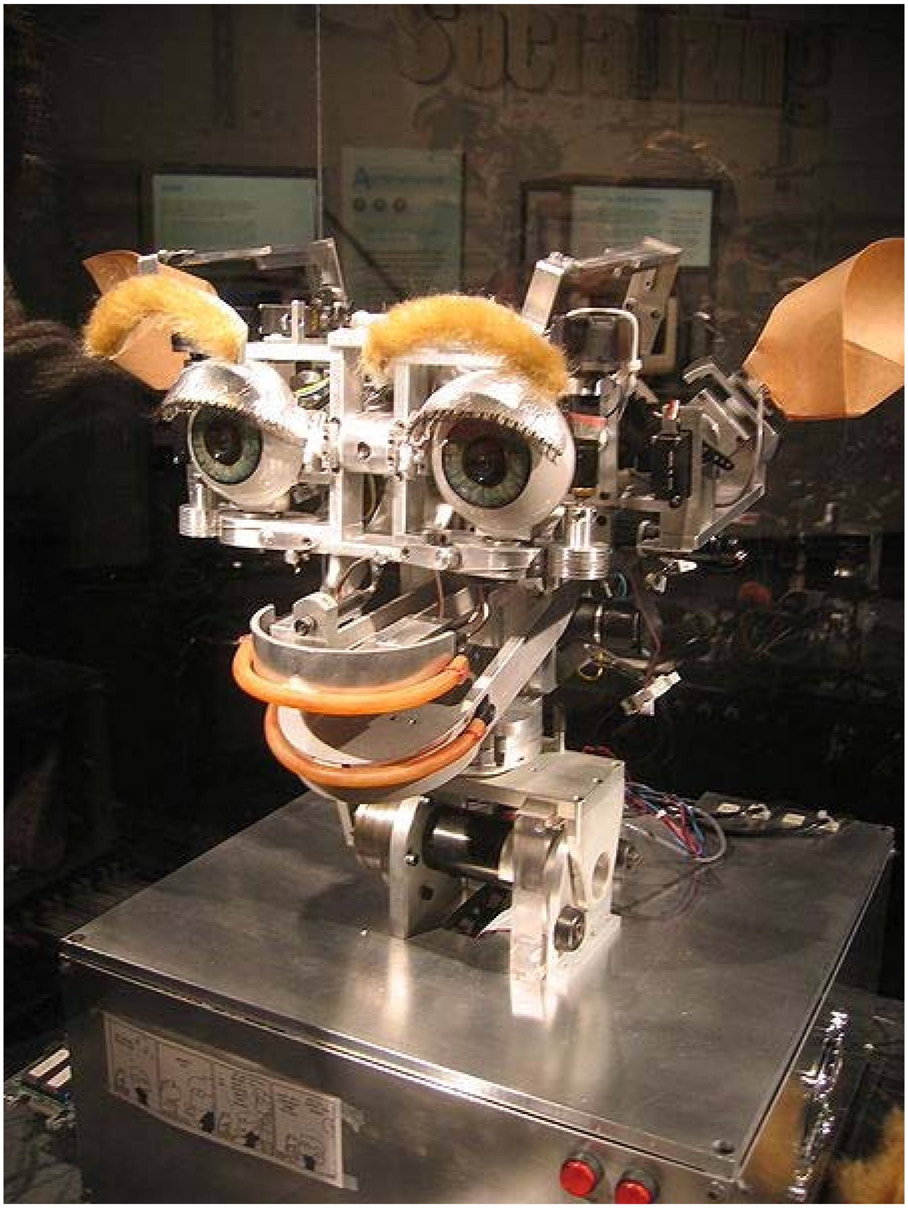
The 3 laws of robotics determine the idea of safe interaction. The closer the human and the robot get the higher the risk of injury. In the industry this is solved by not letting human and robot share the workspace at any time by the extensive use of zones. The presence of human is completely forbidden in some part of space while the robot is working in it and vice versa.

With the advances of artificial intelligence, the autonomous robots could eventually have more proactive behaviours, planning their motion in complex unknown environments. These new capabilities would have to keeping safety as a primer issue and as second efficiency. To allow this new generation of robot, research is being made on human detection, motion planning, scene reconstruction, intelligent behaviour through task planning.

The basic goal of HRI is to define a general human model that could lead to principles and algorithms allowing more natural and effective interaction between humans and robots. Research ranges from how humans work with remote, tele-operated unmanned vehicles to peer-to-peer collaboration with anthropomorphic robots.

Many in the field of HRI study how humans collaborate and interact and use those studies to motivate how robots should interact with humans.

Towards Friendly human-robot interactions



Kismet can produce a range of facial expressions.

Robots are artificial agents with capacities of perception and action in the physical world often referred by researchers as workspace. Their use has been generalized in factories but nowadays they tend to be found in the most technologically advanced societies in such critical domains as search and rescue, military battle, mine and bomb detection, scientific exploration, law enforcement, entertainment, and hospital care.

These new domains of applications imply a closer interaction with the user. The concept of closeness is to be taken in its full meaning, robots and humans share the workspace but also share goals in terms of task achievement. This close interaction needs new theoretical models, on one hand for the robotics scientists who work to improve the robots utility and on the other hand to evaluate the risks and benefits of this new "friend" for our modern society.

With the advance in AI, the research is focusing on one part towards the safest physical interaction. But also on a socially correct interaction, dependent on cultural criteria. The goal is to build an intuitive, and easy communication with the robot through speech, gestures, and facial expressions.

The robot has to adapt itself to our way of expressing desires and orders and not the contrary. But every day environments such as homes have much more complex social rules than those implied by factories or even military environments. Thus, the robot needs perceiving and understanding capacities to build dynamic models of its surroundings. It needs to categorize objects, recognize and locate humans and further their emotions. The need for dynamic capacities pushes forward every sub-field of robotics.

On the other end of HRI research the cognitive modelling of the "relationship" between human and the robots benefits the psychologists and robotic researchers the user study are often of interests on both sides. This research endeavours part of human society.

General HRI research

HRI research spans a wide range of field, some general to the nature of HRI.

Methods for perceiving humans

Most methods intend to build a 3D model through vision of the environment. The proprioception sensors permit the robot to have information over its own state. This information is relative to a reference.

Methods for perceiving humans in the environment are based on sensor information. An example of modern technique is to use colour information for example the fact that for light skinned people the hands are lighter than the clothes worn. A human modelled a priori is then fitted to the sensor information. The robot builds or has (depending on the level of autonomy the robot has) a 3D mapping of its surroundings to witch is assigned the humans locations.

A speech recognition system is used to interpret human desires or commands. By combining the information inferred by proprioception, sensor and speech the human position and state (standing, seated).

Methods for motion planning

Motion planning in dynamic environment is a challenge that is for the moment only achieved for 3 to 10 degrees of freedom robots. Humanoid robots or even 2 armed robots that can have up to 40 degrees of freedom are unsuited for dynamic environments with today's technology. However lower dimensional robots can use potential field method to compute trajectories avoiding collisions with human.

Cognitive models

A lot of data has been gathered with regards to user studies. For example, when users encounter proactive behaviour on the part of the robot and the robot does not respect a safety distance, penetrating the user space, he or she might express fear. This is dependent on one person to another. Only intensive experiment can permit a more precise model.

It has been shown that when a robot has no particular use, negative feelings are often expressed. The robot is perceived as useless and its presence becomes annoying.

In another experiment, it has occurred that people tend to attribute to the robot personality characteristics that were not implemented.

Application-oriented HRI research

In addition to general HRI research, researchers are currently exploring application areas for human-robot interaction systems. Application-oriented research is used to help bring current robotics technologies to bear against problems that exist in today's society. While human-robot interaction is still a rather young area of interest, there is active development and research in many areas.

Search and rescue

First responders face take great risks in search and rescue (SAR) settings, which typically involve environments that are unsafe for a human to travel. In addition, technology offers tools for observation that can greatly speed-up and improve the accuracy of human perception. Robots can be used to address these concerns . Research in this area includes efforts to address robot sensing, mobility, navigation, planning, integration, and tele-operated control.

SAR robots have already been deployed to environments such as the Collapse of the World Trade Center.

Other application areas include:

- Entertainment
- Education
- Field robotics
- Home and companion robotics
- Hospitality
- Rehabilitation and Elder Care

Robot learning

Robot learning is a subset of machine learning and robotics. Usually "robot learning" refers to learning to perform tasks such as obstacle avoidance, control and various other motion-related tasks. While machine learning is frequently used by computer vision algorithms employed in the context of robotics, these applications are usually not referred to as "robot learning" Robot learning can be closely related to adaptive control and reinforcement learning.

Robot locomotion

Robot locomotion is the study of how to design robot appendages and control mechanisms to allow robots to move fluidly and efficiently. Although wheeled robots are typically quite energy efficient and simple to control, other forms of locomotion may be more appropriate for a number of reasons (e.g. traversing rough terrain, moving and interacting in human environments). Furthermore, studying biped and insect-like robots may impact biomechanics.

A major goal in this field is in developing capabilities for robots to autonomously decide how, when, and where to move. However, coordinating a large number of robot joints for even simple matters, like negotiating stairs, is difficult. Autonomous robot locomotion is a major technological obstacle for many areas of robotics, such as humanoids (like Honda's Asimo).

Types of Locomotion

Wheeled

In terms of energy efficiency on flat surfaces, wheeled robots are the most efficient. This is due to the fact that an ideal rolling (but not slipping) wheel loses no energy. A wheel rolling at a given velocity needs no input to maintain its motion. This is in contrast to legged robots which suffer an impact with the ground at heelstrike and lose energy as a result.

There are many different types of wheeled robots, the most common being the Reed Shepps type and the unicycle type. The major concern in the motion planning of wheeled robots are the holonomic that the robot is subject to. These are decided by the type of wheels, number of wheels and the direction of the axes of rotation of the wheels.

Examples

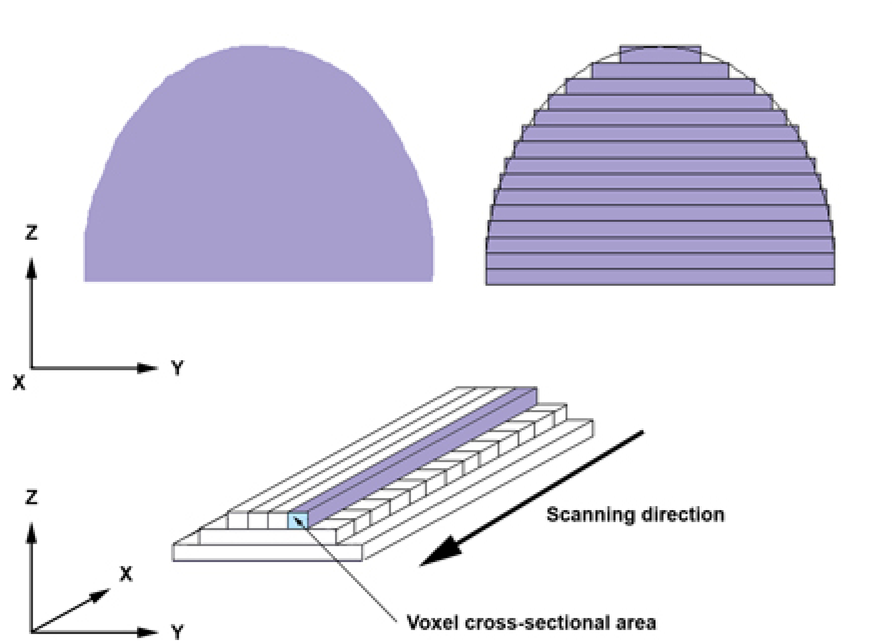
- iRobot's Rhoomba
- Various DARPA Grand Challenge entries
- MobileRobot's Pioneer P3-DX

Rapid prototyping

A rapid prototyping machine using selective laser sintering.

Rapid prototyping is the automatic construction of physical objects using additive manufacturing technology. The first techniques for rapid prototyping became available in the late 1980s and were used to produce models and prototype parts. Today, they are used for a much wider range of applications and are even used to manufacture production-quality parts in relatively small numbers. Some sculptors use the technology to produce complex shapes for fine arts exhibitions.

Introduction



The use of additive manufacturing for rapid prototyping takes virtual designs from computer aided design (CAD) or animation modeling software, transforms them into thin, virtual, horizontal cross-sections and then creates successive layers until the model is complete. It is a WYSIWYG process where the virtual model and the physical model are almost identical.

With additive manufacturing, the machine reads in data from a CAD drawing and lays down successive layers of liquid, powder, or sheet material, and in this way builds up the model from a series of cross sections. These layers, which correspond to the virtual cross section from the CAD model, are joined together or fused automatically to create the final shape. The primary advantage to additive fabrication is its ability to create almost any shape or geometric feature.

The standard data interface between CAD software and the machines is the STL file format. An STL file approximates the shape of a part or assembly using triangular facets. Smaller facets produce a higher quality surface.

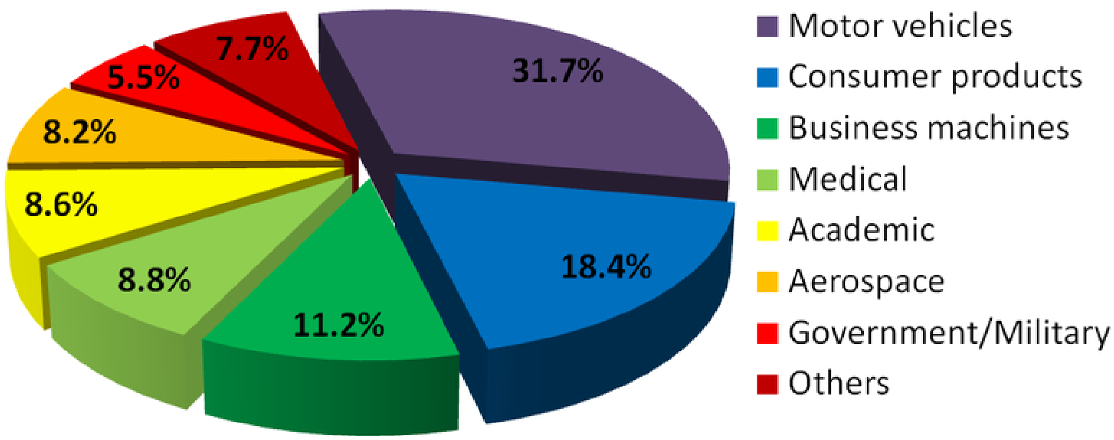
The word "rapid" is relative: construction of a model with contemporary methods can take from several hours to several days, depending on the method used and the size and complexity of the model. Additive systems for rapid prototyping can typically produce models in a few hours, although it can vary widely depending on the type of machine being used and the size and number of models being produced simultaneously.

Some solid freeform fabrication techniques use two materials in the course of constructing parts. The first material is the part material and the second is the support material (to support overhanging features during construction). The support material is later removed by heat or dissolved away with a solvent or water.

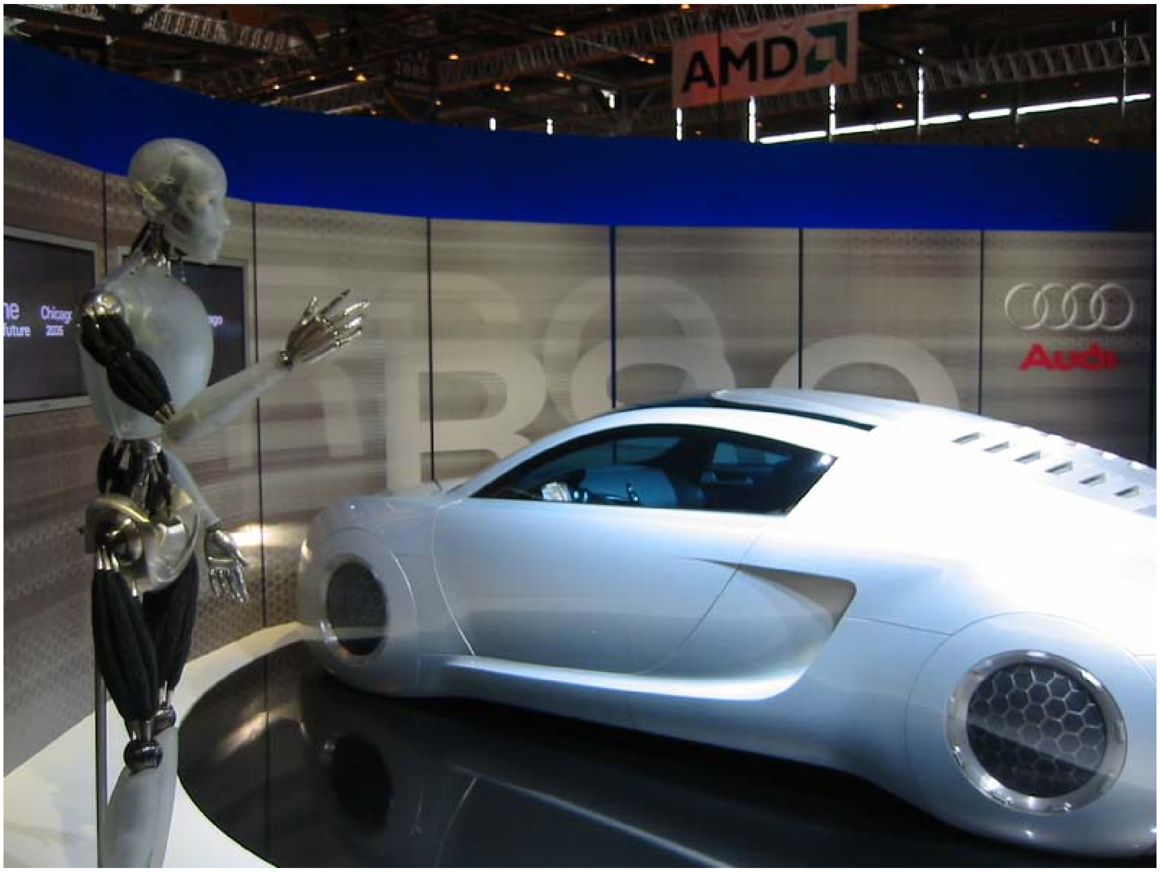
Traditional injection molding can be less expensive for manufacturing polymer products in high quantities, but additive fabrication can be faster and less expensive when producing relatively small quantities of parts. 3D printers give designers and concept development teams the ability to produce parts and concept models using a desktop size printer.

Rapid prototyping is now entering the field of rapid manufacturing and it is believed by many experts that this is a "next level" technology.

Technologies



Rapid prototyping worldwide



The Audi RSQ was made by Audi with rapid prototyping industrial KUKA robots

A large number of competing technologies are available in the marketplace. As all are additive technologies, their main differences are found in the way layers are built to create parts. Some are melting or softening material to produce the layers (SLS, FDM) where others are laying liquid materials thermosets that are cured with different technologies. In the case of lamination systems, thin layers are cut to shape and joined together.

As of 2005, conventional rapid prototype machines cost around £25,000.

Prototyping technologies	**Base materials**
Selective laser sintering (SLS)	Thermoplastics, metals powders
Fused deposition modeling (FDM)	Thermoplastics, eutectic metals.
Stereolithography (SLA)	photopolymer
Laminated object manufacturing (LOM)	Paper
Electron beam melting (EBM)	Titanium alloys
3D printing (3DP)	Various materials

In 2006, John Balistreri and others at Bowling Green State University began research into 3D Rapid Prototyping machines, creating printed ceramic art objects. This research has led to the invention of ceramic powders and binder systems that enable clay material to be printed from a computer model and kiln fired for the first time.

Servomechanism

Industrial servomotor
The grey/green cylinder is the brush-type DC motor. The black section at the bottom contains the planetary reduction gear, and the black object atop the motor is the optical rotary encoder for position feedback. This is the steering actuator of a large robot vehicle.

A **servomechanism**, or **servo** is an automatic device that uses error-sensing negative feedback to correct the performance of a mechanism. The term correctly applies only to systems where the feedback or error-correction signals help control mechanical position or other parameters. For example, an automotive power window control is not a servomechanism, as there is no automatic feedback that controls position—the operator does this by observation. By contrast the car's cruise control uses closed loop feedback, which classifies it as a servomechanism.

A servomechanism may or may not use a servomotor. For example, a household furnace controlled by a thermostat is a servomechanism, yet there is no motor being controlled directly by the servomechanism.

A common type of servo provides *position control*. Servos are commonly electrical or partially electronic in nature, using an electric motor as the primary means of creating mechanical force. Other types of servos use hydraulics, pneumatics, or magnetic principles. Servos operate on the principle of negative feedback, where the control input is compared to the actual position of the mechanical system as measured by some sort of transducer at the output. Any difference between the actual and wanted values (an "error signal") is amplified and used to drive the system in the direction necessary to reduce or eliminate the error. This procedure is one widely used application of control theory.

Speed control via a governor is another type of servomechanism. The steam engine uses mechanical governors; another early application was to govern the speed of water wheels. Prior to World War II the constant speed propeller was developed to control engine speed for maneuvering aircraft. Fuel controls for gas turbine engines employ either hydromechanical or electronic governing.

Positioning servomechanisms were first used in military fire-control and marine navigation equipment. Today servomechanisms are used in automatic machine tools, satellite-tracking antennas, remote control airplanes, automatic navigation systems on boats and planes, and antiaircraft-gun control systems. Other examples are fly-by-wire systems in aircraft which use servos to actuate the aircraft's control surfaces, and radio-controlled models which use RC servos for the same purpose. Many autofocus cameras also use a servomechanism to accurately move the lens, and thus adjust the focus. A modern hard disk drive has a magnetic servo system with sub-micrometre positioning accuracy.

Typical servos give a rotary (angular) output. Linear types are common as well, using a screw thread or a linear motor to give linear motion.

Another device commonly referred to as a servo is used in automobiles to amplify the steering or braking force applied by the driver. However, these devices are not true servos, but rather mechanical amplifiers.

In industrial machines, servos are used to perform complex motion.

History

James Watt's steam engine governor is generally considered the first powered feedback system. The windmill fantail is an earlier example of automatic control, but since it does not have an amplifier or gain, it is not usually considered a servomechanism.

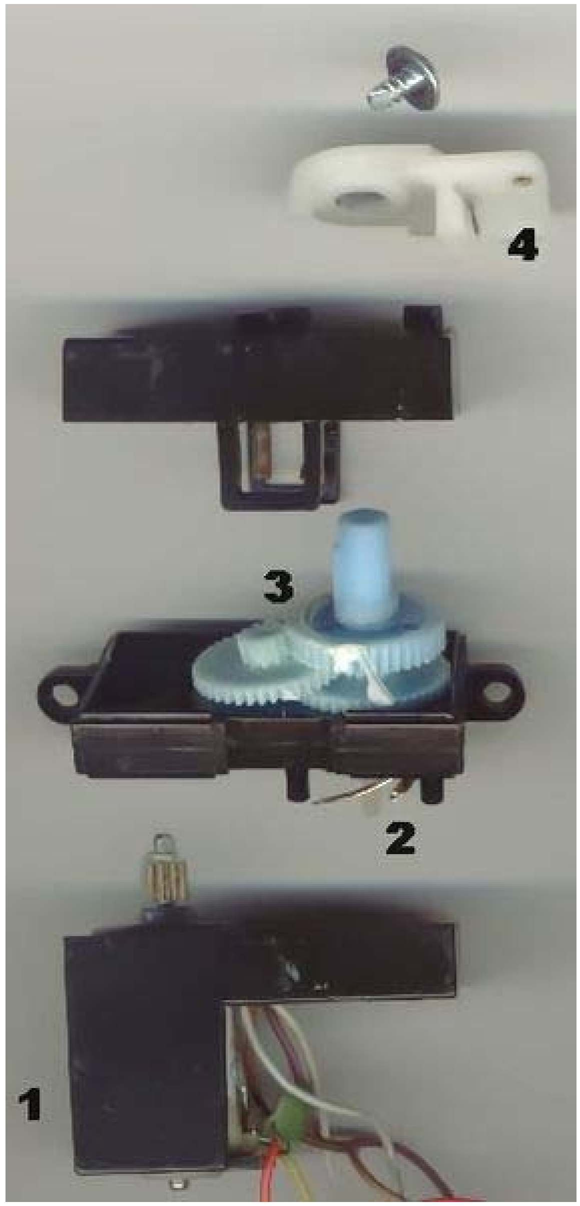
The first feedback position control device was the ship steering engine, used to position the rudder of large ships based on the position of the ship's wheel. This technology was first used on the SS Great Eastern in 1866. Steam steering engines had the characteristics of a modern servomechanism: an input, an output, an error signal, and a means for amplifying the error signal used for negative feedback to drive the error towards zero. The Ragonnet power reverse mechanism was a general purpose air or steam-powered servo amplifier for linear motion patented in 1909.

Electrical servomechanisms require a power amplifier. World War II saw the development of electrical fire-control servomechanisms, using an amplidyne as the power amplifier. Vacuum tube amplifiers were used in the UNISERVO tape drive for the UNIVAC I computer. The Royal Navy began experimenting with Remote Power Control (RPC) on HMS Champion in 1928 and began using RPC to control searchlights in the early 1930s. During WW2 RPC was used to control gun mounts and gun directors.

Modern servomechanisms use solid state power amplifiers, usually built from MOSFET or thyristor devices. Small servos may use power transistors.

The origin of the word is believed to come from the French “Le Servomoteur” or the slavemotor, first used by J. J. L. Farcot in 1868 to describe hydraulic and steam engines for use in ship steering.

RC servos



Small R/C servo mechanism
1. electric motor
2. position feedback potentiometer
3. reduction gear
4. actuator arm

RC servos are hobbyist remote control devices servos typically employed in radio-controlled models, where they are used to provide actuation for various mechanical systems such as the steering of a car, the control surfaces on a plane, or the rudder of a boat.

RC servos are composed of an electric motor mechanically linked to a potentiometer. Pulse-width modulation (PWM) signals sent to the servo are translated into position commands by electronics inside the servo. When the servo is commanded to rotate, the motor is powered until the potentiometer reaches the value corresponding to the commanded position.

Due to their affordability, reliability, and simplicity of control by microprocessors, RC servos are often used in small-scale robotics applications.

The servo is usually controlled by three wires: ground, power, and control. The servo will move based on the pulses sent over the control wire, which set the angle of the actuator arm. The servo expects a pulse every 20 ms in order to gain correct information about the angle. The width of the servo pulse dictates the range of the servo's angular motion.

A servo pulse of 1.5 ms width will typically set the servo to its "neutral" position or 45°, a pulse of 1.25 ms could set it to 0° and a pulse of 1.75 ms to 90°. The physical limits and timings of the servo hardware varies between brands and models, but a general servo's angular motion will travel somewhere in the range of 90° - 120° and the neutral position is almost always at 1.5 ms. This is the "standard pulse servo mode" used by all hobby analog servos.

A hobby digital servo is controlled by the same "standard pulse servo mode" pulses as an analog servo. Some hobby digital servos can be set to another mode that allows a robot controller to read back the actual position of the servo shaft. Some hobby digital servos can optionally be set to another mode and "programmed", so it has the desired PID controller characteristics when it is later driven by a standard pulse servo receiver.

RC servos are usually powered by the receiver which in turn is powered by battery packs or an Electronic speed controller (ESC) with an integrated or a separate Battery eliminator circuit (BEC). Common battery packs are either NiCd, NiMH or lithium-ion polymer battery (LiPo) type. Voltage ratings vary, but most receivers are operated at 5 V or 6 V.

Robot calibration

Robot calibration is the process of determining the actual values of kinematic and dynamic parameters of an industrial robot (IR). Kinematic parameters describe the relative position and orientation of links and joints in the robot while the dynamic parameters describe arm and joint masses and internal friction.

A calibrated robot has a higher *absolute* positioning accuracy than an uncalibrated one, i.e., the real position of the robot end effector corresponds better to the position calculated from the mathematical model of the robot. Absolute positioning accuracy is particularly relevant in connection with robot exchangeability and off-line programming of precision applications. Besides the calibration of the robot, the calibration of its tools and the workpieces it works with (the so-called *cell calibration*) can minimize occurring inaccuracies and improve process security.

Accuracy criteria and error sources

The international standard ISO 9283 is setting different performance criteria for IR and suggesting test procedures in order to obtain appropriate parameter values. The most important criteria, and also the most commonly used, are accuracy of pose (AP) and repeatability of pose (RP). Repeatability is particularly important when the robot is moved towards the command positions manually („Teach-In“). If the robot program is generated by a 3D simulation („off-line programming“), absolute accuracy is vital, too. Both are generally influenced in a negative way by kinematic factors. Here especially the joint offsets and deviations in lengths and angles between the individual robot links take effect.

Measurement systems

There exist different possibilities for pose measurement with industrial robots, e.g. touching reference parts, using supersonic distance sensors, laser interferometry, theodolites, calipers or laser triangulation. Furthermore there are camera systems which can be attached in the robot's cell or at the IR mounting plate and acquire the pose of a reference object. Measurement and calibration systems are made by such companies as Dynalog, FARO Technologies, Leica, Metris, Metronor, Wiest, Teconsult and Automated Precision,Inc..

Mathematical principles

$$\vec{p} = (p_1, \ldots, p_n)^T,\ \vec{p} \in R^n$$

$$\vec{x} = (\vec{x}_1, \ldots, \vec{x}_m)^T,\ \vec{x}_i \in R^k,\ \vec{x} \in R^{m \times k}$$

$$\vec{y}_M = (\vec{y}_{M1}, \ldots, \vec{y}_{Mm})^T,\ \vec{y}_{Mi} \in R^l,\ \vec{y}_M \in R^{m \times l}$$

$$\vec{y}_S = (\vec{y}_{S1}, \ldots, \vec{y}_{Sm})^T,\ \vec{y}_{Si} \in R^l,\ \vec{y}_S \in R^{m \times l}$$

$$\vec{y}_M = \vec{f}(\vec{p}, \vec{x}),\ \vec{f} \in R^{m \times l}$$

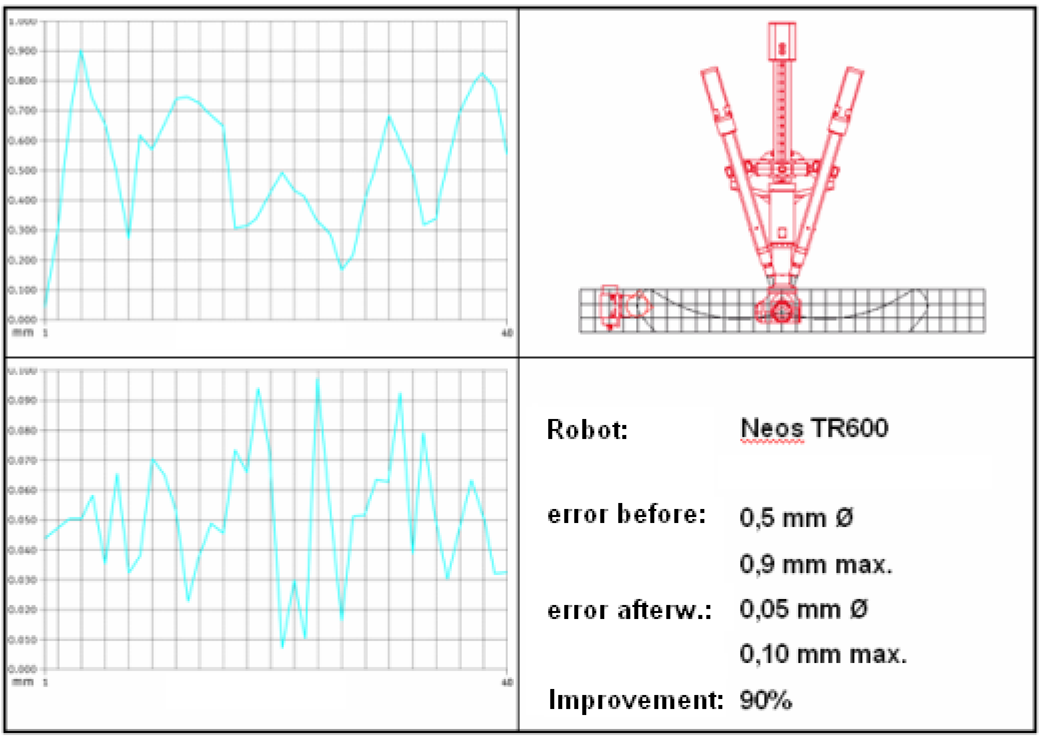
$$\min_{\vec{p} \in R^n} \{r\} \quad \text{mit} \quad r = \left\| \vec{y}_M - \vec{y}_S \right\|^2 = \left\| \vec{f}(\vec{p}, \vec{x}) - \vec{y}_S \right\|^2 \quad \text{und} \quad r \in R$$

Objective function and optimization problem

The robot errors gathered by pose measurements can be minimized by numerical optimization. For kinematic calibration, a complete kinematical model of the geometric structure must be developed, whose parameters then can be calculated by mathematical optimization. The common system behaviour can be described with the vector model function as well as input and output vectors (see figure). The variables *k, l, m, n* and their derivates describe the dimensions of the single vector spaces. Minimization of the residual error *r* for the purpose of identification of the optimal parameter vector *p* follows from the difference between both output vectors using the Euclidean norm.

For solving the kinematical optimization problems least-squares descent methods are convenient, e.g. a modified quasi-Newton method. This procedure supplies corrected kinematical parameters for the measured machine, which then for example can be used to update the system variables in the controller in order to adapt the used robot model to the real kinematics.

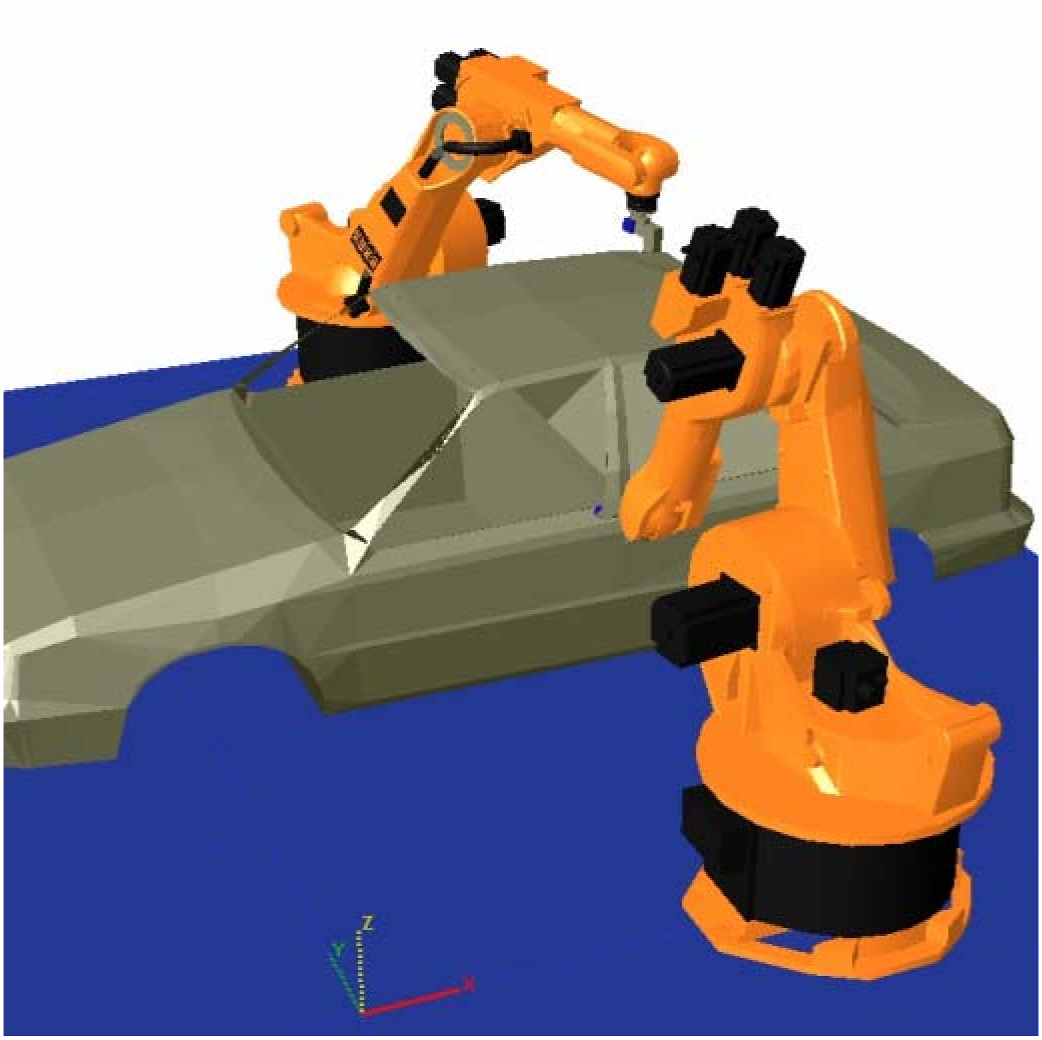
Results



Positioning accuracy of a Tricept robot before and after calibration

The positioning accuracy of industrial robots varies by manufacturer, age, and robot type. The magnitude of the error between the actual position and the desired position can be as low as a tenth of a millimeter, or as high as several centimeters. Using kinematic calibration, these errors can be reduced to less than a millimeter in most cases. An example of this is shown in the figure to the right.

Sample applications



In-line measurement cell for car body inspection

In industry there is a general trend towards substitution of machine tools and special machines by industrial robots for certain manufacturing tasks whose accuracy demands can be fulfilled by calibrated robots. In the figure a current example is shown: In-line measurement in automotive manufacturing, where the common „measurement tunnel“ used for 100% inspection with many expensive sensors are partly replaced by IR which carry only one sensor each. This way the total costs of a measurement cell can be reduced significantly. Furthermore the station can be re-used after a model change by simple re-programming without mechanical adaptations.

Further examples for precision applications are robot-guided hemming in car body manufacturing, assembly of mobile phones, drilling, riveting and milling in aerospace industry and increasingly medical applications.

Summary

By application of efficient calibration methods it is possible with today's industrial robots – especially parallel kinematic manipulators – to achieve an accuracy of pose of 0.1 mm in order to improve exchangeability, to simplify off-line programming, and to enable new, highly precise applications.

Robot welding

A set of six-axis robots used for welding.

Robot welding is the use of mechanized programmable tools (robots), which completely automate a welding process by both performing the weld and handling the part. Processes such as gas metal arc welding, while often automated, are not necessarily equivalent to robot welding, since a human operator sometimes prepares the materials to be welded. Robot welding is commonly used for resistance spot welding and arc welding in high production applications, such as the automotive industry.

Robot welding is a relatively new application of robotics, even though robots were first introduced into US industry during the 1960s. The use of robots in welding did not take off until the 1980s, when the automotive industry began using robots extensively for spot welding. Since then, both the number of robots used in industry and the number of their applications has grown greatly. In 2005, there were more than 120,000 robots are used in North American industry, about half of them pertaining to welding. Growth is primarily limited by high equipment costs, and the resulting restriction to high-production applications.

Robot arc welding has begun growing quickly just recently, and already it commands about 20% of industrial robot applications. The major components of arc welding robots are the manipulator or the mechanical unit and the controller, which acts as the robot's "brain". The manipulator is what makes the robot move, and the design of these systems can be categorized into several common types, such as the SCARA robot and cartesian coordinate robot, which use different coordinate systems to direct the arms of the machine.

The technology of signature image processing has been developed since the late 1990s for analyzing electrical data in real time collected from automated, robotic welding, thus enabling the optimization of welds.