Artificial Eye, Artificial Vision: How does my robot see?
Transcript

Date: Wednesday, 9 April 2014 - 6:00PM
Location: Museum of London
9 April 2014

Artificial Eye, Artificial Vision: How does my robot see?

Professor William Ayliffe

Mechanical devices that move without human force have fascinated people for millennia. Until recently these have been blind pre-programmed mechanical devices. Advances in machine vision have enabled autonomous machines to problem solve, recognize objects, faces and to navigate around obstacles in complex environments.

Sophisticated mechanical devices are known to have existed in ancient Greece, though the only surviving example is the Antikythera mechanism.

Manufactured in 87 BC, it was recovered in 1900 from the ancient Antikythera shipwreck. The system of wheels and teeth functioned as an ancient analog computer designed to calculate astronomical positions. Earlier in 270BC, Ctesibius, an engineer in Alexandria, had made water clocks with moving figures. The tradition was continued and Heron, in Roman Alexandria, 10-70AD, invented amongst other things, the first steam engine (aeolipile) which rotated when its contained water was boiled, and a number of mechanical devices operated by cogwheels driven by the metal balls falling sequentially into a drum. The Islamic and European Renaissance was a fertile period for programmable automatic moving devices from clockwork astronomical calculators to human-like machines that could perform a limited number of pre-programmed tasks. Amazingly some of these devices still exist, including the oldest working clock in the World at Salisbury Cathedral and its near contemporary astronomical clock from Wells Cathedral. Giovanni Fontana (c1395 – c1455) a Venetian physician and engineer at the University of Padua, describes many such devices in his Bellicorum instrumentorum liber including siege engines a magic lantern to project images onto walls; rocket-propelled bird, fish, and rabbit.

In 1495, Leonardo designed a humanoid automaton as a knight clad in armour. It was possibly displayed at a celebration hosted by Ludovico Sforza in Milan. The sketchbook plans were discovered in 1950. The automaton could stand, sit, raise its visor and manoeuvre its arms with two interior independent systems:

The legs were powered by an external crank driving cable, connected to the ankle, knee, and hip, allowing three-degrees-of-freedom.

A mechanical, analog-programmable controller in chest provided power and control for the upper limbs, enabling four-degree-of-freedom for the arms with articulated shoulders, elbows, wrists, and hands.
In the 18th and 19th century many beautiful and sophisticated mechanical devices were designed including the Canard Digérateur of Vaucanson (1737) which ate grain and excreted material; prompting the sarcastic remark of Voltaire “without...the duck of Vaucanson, you will have nothing to remind you of the glory of France.”

Between 1768 and 1774, Pierre Jaquet-Droz made several programmable Machine Dolls with up to 6,000 moving pieces. Used as advertising to sell watches, they are still in working order. In 1800, Henri Maillardet, a Swiss clockmaker working in London built the Draughtsman-Writer, with the largest "memory" of any automaton, contained in brass discs, allowing four drawings and three poems (two in French and one in English). As the cams are turned by the clockwork motor, three steel brushes follow their irregular edges translating the movements into side to side, front and back, and up and down movements of the doll's writing hand through a system of levers and rods that produce the markings on paper.

The term robot, replaced automaton, and was first used by Karel Capek, a Czech dramatist in his 1920 Play "R.U.R" (Rossum's Universal Robots)

These intelligent biological machines were designed to serve humans and were named from the Czech, "robota” meaning forced and slavish work derived from rab, meaning "slave." Rossum is also an allusion to a Czech word rozum, meaning "reason, wisdom".

However the robots take over the world and destroy humanity. He distinguished the robot from man by the absence of emotion.

In 1927, Thea Gabriele von Harbou (1888-1954) actress, author and film director, with her husband Fritz Lang produced the hugely expensive and influential film, Metropolis, set in a futuristic urban dystopia, Brigitte Helm played both the robot form and human incarnation, and was the first robot ever depicted in cinema. An early attempt at making human-like robot, Alpha was later shown to be elaborate fake.

However, Elektro built by the Westinghouse Corporation and displayed at the 1939 New York World’s Fair had 26 routines and responded to speech by sensing the vibrations and used a preset vocabulary of 700 words recorded on 78rpm records.

He was powered by a plugged in cable, although Nicola Tesla had already discovered remote control using radio waves in 1898, the technology was not advanced enough to control a complex android robot.

A series of programable model machine arms led to the first industrial robots in 1961 with George Devol's Ultimate being used at General Motors.

Programmed with a set of step-by-step instructions, stored on a magnetic drum, the arm moved hot pieces of die-cast metal. Welding these parts on automobile bodies, was a dangerous task for humans.
These types of robots are commonly found in manufacturing performing roles such as welding, painting and handling materials.

Their six independent joints, allow six degrees of freedom. Placing a solid body in space requires six parameters; three to specify the location \((x, y, z)\) three to specify the orientation \((\text{roll}, \text{yaw}, \text{pitch})\).

When combined with a vision system it can move product from conveyor belt to package at a very high rate of speed

**ROBOT SENSING**

Robot vision need not be limited to light as is the case for humans. Vision in the infra-red, ultraviolet and even X-ray for seeing into objects can all be incorporated into the machine.

However interpreting these various images remains problematic and a current area of intense research. An early attempt at having automatic direction of movement was the robot dog built by John Hammond, Jr. and Ben Miessner in 1912. This ancestor of phototropic robots, automatically moved to lights using Selenium cell 'eyes'.

If one 'eye' received light: its Relay activated the power supply to the Motor, and to the Solenoid on the opposite side: The Robot Dog moved forward steering towards the light.

If both 'eyes' received sufficient light then power would be supplied to the Motor, but the Solenoids would be OFF so the Dog would steer straight ahead.

Attempting to understand how a few circuits could lead to complex behavior in animals, in 1948, Grey Walter, a Neurophysiologist in Bristol built electric Tortoises. These slow, autonomous machines on three wheels used phototaxis, to locate their recharging station when low on battery power. They also avoided bumping into objects, a leap in robotics because they operated without preset directions.

These machines confirmed that a few simple circuits enabled conditioned behavior.

In 1951 three new tortoises were displayed at the Festival of Britain and were then auctioned off. They inspired Edmund C. Berkeley & Jack Koff to build Squee, a robot squirrel, that hunted tennis balls and brought them to a collection point. Squee had four sense organs (two phototubes, two contact switches), and three acting organs (a drive motor, a steering motor, and a motor which opens and closes the scoop or "hands"). These were controlled by a small brain of only half a dozen relays.

When a torch was shone at the robot, it approached, scooped the ball, then ignored the light, and took the ball to its nest marked by a flashing light. Squee was the first robot to be able to carry out a defined task, as opposed to just steer towards light.

It was also the first robot to have a manipulator under automatic control.

Analogue Robots can be programmed to achieve a simple goal such as finding more light or responding to sound.

They use just a few analogue circuits, not a microprocessor and include Phototropes, audiotropes, thermotropes that can sit, squirm, crawl, jump, swim, or fly in response to their stimulating sense. Commercially
the iRobot Corporation manufactures a variety of these robots, Roombas for cleaning rooms, Looj for gutters and Mirra for pools.

Modern industrial robots are marvels of engineering. They have replaced human beings in a wide variety of industries. Robots out-perform humans in jobs that require precision, speed, endurance and reliability, safely performing dirty and dangerous jobs.

A robot the size of a person can quickly carry heavy loads with a repeatability of +/- 0.006".

A basic problem with industrial assembly is a process: **parts feeding**.

Objects/parts required for product assembly stored in a bin.

The assembly process requires a single part to be isolated from the bin of parts.

In structured environments, with fixed part presentation a blind robot can repeatedly perform these tasks. However, if there is variability in the way the parts are presented, blind robots need some assistance.

**Robot’s guide - machine vision**

A typical task would be for a robot to pick up a tray from a flat surface.

1: Acquire a suitable image.

2: “Find” the object of interest (the Tray, or the piece that must be grabbed.)

3: Determine the object’s position $X$ vs $Y$ and orientation: $R_z$

4: Translate this location to the robot’s co-ordinate system.

5: Send the information to the robot.

6: Using that information, the robot can then move to the proper position and orientation to grasp the object in a prescribed way.

The machine vision portion (steps #1 through #5) are executed within a few hundredths of a second.
Cameras do not create the image, they transform optical signals (light) into electrical charge (voltage) and process (digitize) them into electronic signal (raw digital image).

At each location the brightness of the image is stored as a number.

The robot now has a series of numbers, related to the intensity of light, for each location on the x,y plane. In order to make sense of these numbers it has to extract meaningful information, feature analysis. For black and white images, this is relatively straightforward, and a number of algorithms will allow the robot to recognise simple shapes.

It gets more complicated if there are several shapes, overlapping shapes or variations in contrast. In order to create meaningful sense of a chaotic array of pixels of different values, the scene must be reconstructed.

**Edge detection**

A useful first step to understand the various components of a scene is to determine where the borders of objects lie in the image. Once the computer finds the edges it can assign shape and start to process what objects might have that particular shape.

Similarly, colours or textures could be extracted to give a sophisticated analysis of the visual scene. Indeed the robot is not limited to light and might use infra-red, ultra-violet, X-rays or radio waves to build a map of the external world.

However, edge detection is a critical first step, whatever the imaging modality might be. An **edge** is the **boundary** between an object and the background.

One relatively simple method to determine where edges lie is to use a 3x3 window of the image. The local threshold is calculated as the mean of the 9 intensity values of the pixels in the window.

If a pixel has an intensity value greater than this threshold, it is set to a 1.

If a pixel has an intensity value less than this threshold, it is set to a 0.

This gives a binary pattern of the 3x3 window.
More commonly, methods to locate edges calculate the gradients in the image.

Vision is of course much more complicated than the analysis of edges. Advanced analysis of the pixels and using multiple sensors, allows robots to build a more accurate model of their environs. They can now move around a complicated three-dimensional world with accuracy.

Once limited to flat surfaces and with limited vocal skills, Robots now have excellent speech recognition allowing realistic conversations. When they are made to look like humans this is unnerving. Furthermore, robots can play football, perform complex tasks on Mars, and use algorithms for facial and personal recognition that enable surveillance or target location.

The ethical considerations are enormous and must be taken on board as engineers make sophisticated machines with locomotive and visual capabilities that exceed those of humans.

© Professor William Ayliffe, 2014