Machine Vision: Approaches and Limitations

Moisés Rivas López, Oleg Sergiyenko and Vera Tyrsa
Engineering Institute of Autonomous University of Baja California, Polytechnic University of Baja California
Mexicali, BC, México

1. Introduction

Machine vision in its common definition is a possibility of a machine (by sensing means and computer mathematic processing consecutively) to obtain an information about surrounding environment for further analytical treatment.
According to this common definition we can unite in a general classification various, sometimes quite different by its principle, technical systems.
These classification tables can be represented on the base of two different approaches: 1) practical causes (Soini, 2001) for necessity to "see surrounding environment", and 2) technical principle and means using for this task solution.
According to the common definition any complete Machine vision system combines two components: technical means (or hardware) and information processing mathematics and algorithm (or software). However, the various software analyses is not expedient in view of variety of mathematical methods and their object focused applications in each case (Mordohai & Medioni, 2006); and finally can't give clearer problem understanding.
We are now observing a rapid growth of 3D software and hardware capabilities for mainstream PCs, and 3D graphics accelerator boards with processing capabilities of roughly millions polygons per second are becoming commonplace (Petrov et al., 1998). At the same time, dynamic level-of-detail algorithms—built into standard 3D software packages— offer considerable acceleration of model viewing and progressive loading and transmission of 3D models. Despite the fast growth of computer 3D visualization capabilities, until recently data input technology has remained unchanged.
So, in our research for better understanding what is Machine vision, what is its modern state, which practical and technical tasks it decide, and which objective limitations and open problems recently it have, we'll based on the two mentioned above approaches.
In a part of practical reasons, which caused for necessity to develop Machine (or computer) vision concept, can be mentioned:

- security problems in static/dynamic image analysis in perimeter/volume protection (motion/dangerous object detection); (Chellappa et al., 2005), (Itti & Baldi, 2005)
- analysis of short/long term deformation of important engineering structures (more commonly known as 'structural health monitoring' or SHM); (Athavale et al., 1990), (Allen et al., 2005), (Mallet et al., 2004), (Tyrsa et al., 2004), (Ohno et al., 2002), (Benedetti et al., 2004), (Slob & Hack, 2004), (Stewart & Tsakiri, 2002), (Liwen Dai et al., 2002)

Source: Computer Vision, Book edited by: Xiong Zhihui,
ISBN 978-953-7619-21-3, pp. 538, November 2008, I-Tech, Vienna, Austria

- surface digital mapping and micro-surface analysis; (Winkelbach et al., 2006), (Slob & Hack, 2004), (Tunstel, 1995), (Vang et al., 2000), (Peng-Hsiang Weng & Fu-Jen Kao, 2004)
- automatic navigation of robot in unknown scene (Tunstel, 1995), (Diosi & Kleeman, 2005), (França et al, 2005), (Ibeanusi et al, 1999), (Chaumette & Hutchinson, 2006), (Surmann et al, 2003), (Kee et al, 2008), (Hada & Takase, 2001)

In a part of technical principle and means Machine vision system can be classified to:

- camera (or "stereo camera") principle with further image analysis algorithms; (Chellappa et al., 2005), (Itti & Baldi, 2005), (Selectes papers on CCD and CMOS imadges, 2003), (Lavelle et al., 2004)
- 2D-3D image reconstruction techniques (most frequently - 3D laser scanning on triangulation principle); (Tyrsa et al., 2004), (Peng-Hsiang Weng & Fu-Jen Kao, 2004), (França et al, 2005), (Surmann et al, 2003), (Lavelle et al., 2004), (*Handbook of Optical and Laser Scanning,* (2004), (Forman & Parry, 2001), (Pierce et al., 1992), (Lichti et al., 2000), (Petrov et al., 1998)
- terrestrial laser total stations, or airborne laser scanners; (Slob & Hack, 2004), (Baltsavias, b, 1999), (Baltsavias, a, 1999), (Wehr & Lohr U, 1999)
- obstacle detection y description techniques, based on signal "time-of-flight"(radar, lidar, sonar, rangefinders or UWB technologies); (Vang et al, 2000), (Pierce et al., 1992), (Yu et al., 2008)
- GPS-based systems for objects surface reconstruction; (Kee et al., 2008), (Stewart, & Tsakiri, 2002), (Hada & Takase, 2001), (Liwen Dai et al., 2002)
- combined systems, which use a certain combination of mentioned above basic means for increase total system resolution and noise robustness (Ohno et al., 2002), (Benedetti et al., 2004), (Lavelle et al., 2004), (Retscher, 2007), (Liu et al., 2006).

The goal of our work is to compare several of mentioned approaches to Machine vision design, compare their advantages over each other, and the most principal limitations for practical application.

In general, variety of practical applications which strongly requires the technical vision device it is just a confirmation of an actuality and high importance of a new technical vision methods development.

However, sometimes very different practical tasks use very similar techniques for practical task solution. And it has a different level of success. Because of various practical limitation and specific requirements which appears in each unique case. So, the key parameter for to analyze different technical principle for machine vision system design is a basic technical device for task solution.

2. Approaches to machine vision design

More attentive analysis of the mentioned above technical principle and means list permit us to make a simplified conclusion. There are four completely distinct technical approaches to technical vision device design. More truly, if to be rigorous in definitions, three relatively independent methods, and the fourth group which cannot be an independent basis for creation of the device, but possesses such important advantages, that at use of other methods it is not to forget about them, and it is desirable to use actively them as auxiliary means.

These four groups are:

- camera methods;
- laser scanning systems;
- GPS-based methods;
- numerous rangefinder devices (radar-, sonar- , laser-principle, etc.)

As evident the last one cannot be an independent basis for creation of the complete Machine vision system. Because of its physical nature this system it is only capable to estimate the distance to “averaged object”, but not reconstruct its surface point-by-point.
Let us carefully review all mentioned principles of Machine vision design for establish their strong and weak points.

2.1 Camera based machine vision

The machine vision system includes (Selectes papers on CCD and CMOS images, 2003) a stereo TV camera assembly and a processing unit. It detects obstacles in real time within its field of view in a range from 5 m to 50 m ahead of the vehicle with a viewing angle of 40 degrees. The cameras are arranged vertically at the front part of the vehicle. The system locates obstacles in the trapezoidal field of view. The scanning of each camera is synchronized and the processing unit uses hard-wired logic instead of a programmable device in order to realize high speed processing of video signals from the cameras. The principle of the obstacle detection is parallax. When two images from both of the cameras are compared, the two images of an obstacle are identical except the positions in the frames. On the other hand each image of figures on the ground differs due to the positions of the cameras. Fig. 1 illustrates the principle of the obstacle detection.

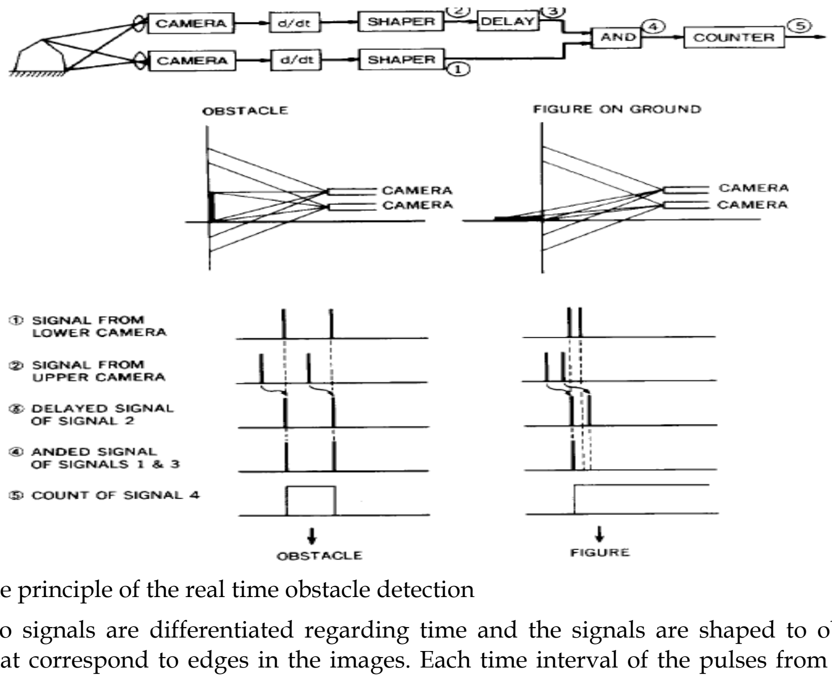


Fig. 1. The principle of the real time obstacle detection

The video signals are differentiated regarding time and the signals are shaped to obtain pulses that correspond to edges in the images. Each time interval of the pulses from each cameras, ***(signal*** 1 and ***signal*** 2 in Fig. l), discriminates an obstacle from a figure on a road.

An obstacle generates same time intervals, but a figure on a road generates different time intervals. The cameras have to be, thus, synchronized with each other, and have to employ vertical and progressive scanning techniques. The position of a scanning line corresponds to the direction to the obstacle, and the point where the optical axes of the cameras are crossing indicates the distance to the obstacle. Delaying of one of the signals from the TV cameras is equivalent IO rotation of the optical axis of the camera. Thus, varying the delay time enables us to detect obstacles at other locations. For enlargement of the field of view and detection of obstacle:; in the two-dimensional field of view during one scanning period, parallel processing with 16 kinds of delay time is employed, which yields the field of view of 16 zones arranged longitudinally at intervals of **1** m. Time required to detect obstacles is 35.6 ms, which consists of 33.3 ms of scanning OF one frame and 2.3 ms of processing to detect and locate obstacles. Fig. 2 shows an example of the obstacle detection. The guardrail is identified as a series of obstacles that are indicated by black elements in the figure at the bottom. Since the system had no measures against brightness, shadows, and shades, the operating condition was restricted.

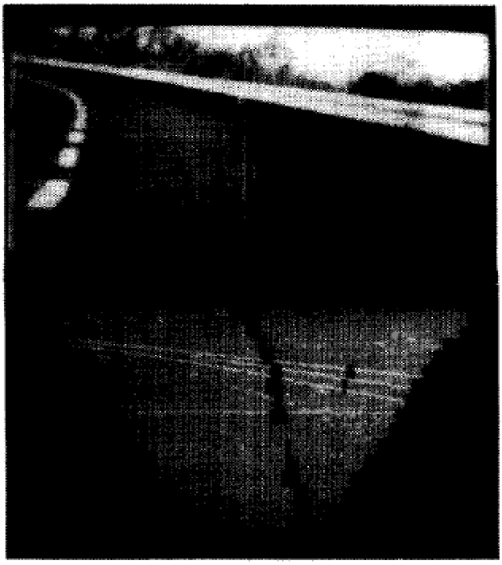


Fig. 2. The obstacle detection: a road scene (top) and obstacles in the scene (bottom) (Sadayuki Tsugawa, 1994)

In a basis of any camera method it is the principle of the human volumetric vision, capable to reconstruct a 3-dimensional picture and approximately estimate distances up to the objects within scene. That is in other words, stereovision.

Any stereovision technical system is approach to a multicamera system. In the elementary case under consideration it is two cameras system. If a stereovision system (Chaumette & Hutchinson, 2006) is used, and a 3-D point is visible in both left and right images (see Figure 3), it is possible to use as visual features s (vector ***s*** contains the desired values of the scene/features):

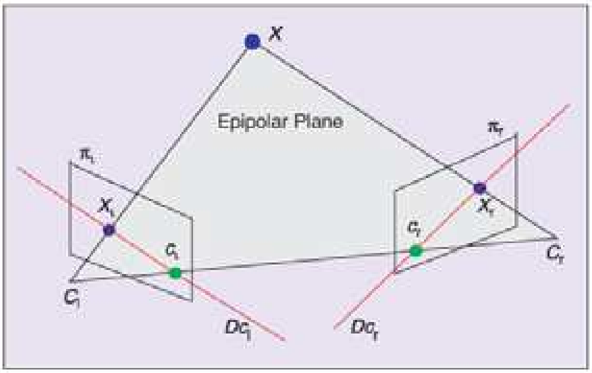


Fig. 3. A stereovision system

$$\mathbf{s} = \mathbf{x}s = (\mathbf{x}l, \mathbf{x}r) = (xl, y\ l, xr, yr),$$

i.e., to represent the point by just stacking in **s** the *x* and *y* coordinates of the observed point in the left and right images.
For a 3-D point with coordinates $X = (X,Y,Z)$ in the camera frame, which projects in two images as a 2-D point with coordinates $\mathbf{x} = (x, y)$, we have (Chaumette & Hutchinson, 2006):

$$\begin{cases} x = X/Z = (u - c_u)/f\alpha \\ y = Y/Z = (v - c_v)/f, \end{cases} \tag{1}$$

where image measurements matrix $\mathbf{m} = (u, v)$ gives the coordinates of the image point expressed in pixel units, and $\mathbf{a} = (c\ u, c\ v, f, \alpha)$ is the set of camera intrinsic parameters: $c\ u$ and $c\ v$ are the coordinates of the principal point, f is the focal length, and α is the ratio of the pixel dimensions. In this case, we take $\mathbf{s} = \mathbf{x} = (x, y)$, the image plane coordinates of the point.
Taking the time derivative of the projection equations (1), we obtain the result which can be written in general form

$$\mathbf{X} = \mathbf{L}_x \mathbf{V}_c, \tag{2}$$

where $\mathbf{V}_c$ is a spatial velocity of the camera be denoted by $\mathbf{v}_c = (v_c, \omega_c)$, (with v_c the instantaneous linear velocity of the origin of the camera frame and ω_c the instantaneous angular velocity of the camera frame) and the interaction matrix $\mathbf{L}_x$ (we consider here the case of controlling the motion of a camera with six degrees of freedom) related to **x** is

$$L_x = \begin{bmatrix} \frac{-1}{Z} & 0 & \frac{x}{Z} & xy & -\left(1+x^2\right) & y \\ 0 & \frac{-1}{Z} & \frac{y}{Z} & 1+y^2 & -xy & -x \end{bmatrix} \tag{3}$$

L_x is an interaction matrix related to *s, or* feature Jacobian. In the matrix $\mathbf{L}_x$, the value *Z* is the depth of the point relative to the camera frame. Therefore, any control scheme that uses this form of the interaction matrix must estimate or approximate the value of *Z*. Similarly

(Chaumette & Hutchinson, 2006), the camera intrinsic parameters are involved in the computation of x and y.
Let us analyze this basic "camera method" parameters regard to equations (1)-(3).

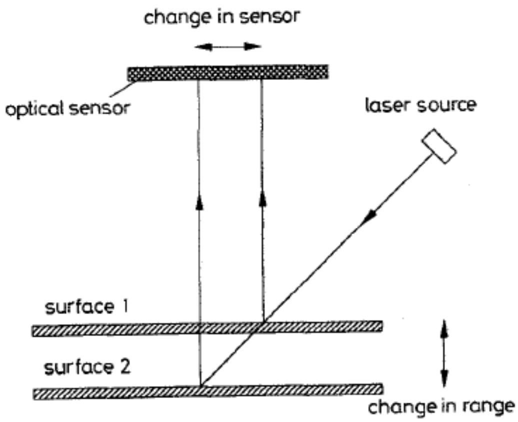
1. Focal distance f in (1), as well as the set of camera intrinsic parameters, shows us that camera uncertainty is strongly related to fabrication imperfections for each unique camera. And, hence the absolute value of this camera uncertainty is rising sufficiently with scene depth increase (proportionally to $1f$, $2f$ … nf).
2. $\mathbf{V}c$ in (2), which is camera velocity, shows us that any camera method error strongly related to any own camera motions. And, taking to account an arbitrary character of camera self-motions and consequently instant vector $\mathbf{V}c$ direction, it is very difficult to estimate real camera uncertainty.
3. Depth Z in (3) shows for components (1;1), (2;2) (2;3) for example, that method resolution and reasonable operating rate are strongly limited by own theory.

These reasons permits us to understand exactly clear a principal limitations for any practical application of camera based technical vision method. They are very critical for practical using in any application with self-moving camera positioning, are extremely sensible to any vibration and dynamic shocks. Moreover, in spite of significant achievements of camera technologies in last ten years, it still exist a possibility of "camera error", i.e. when one typical object image is assigned with typical properties of another. And this possibility is proportionally increased with distance growth. Finally, camera methods application is naturally limited with distances up to 50 m.

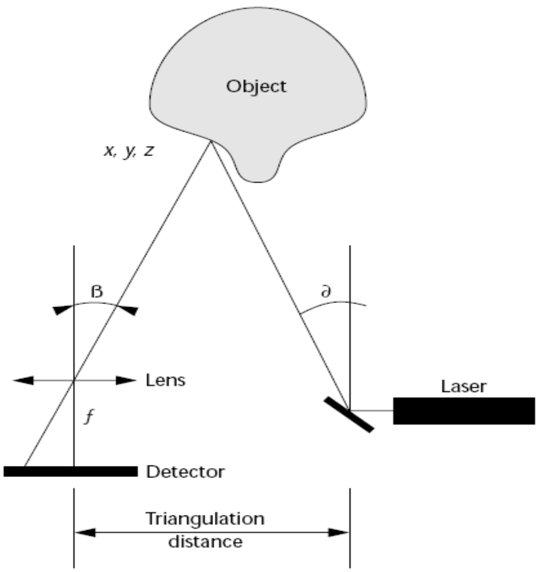
2.2 Laser principle based machine vision systems

Modern laser sensor technologies have made fast and accurate 3-D data acquisition possible as evidence by several commercialized 3-D laser sensors (Tyrsa et al., 2004), (Peng-Hsiang Weng & Fu-Jen Kao, 2004), (Diosi & Kleeman, 2005), (França et al., 2005), (Surmann et al., 2003), (Lavelle et al., 2004), (Handbook of Optical and Laser Scanning, 2004), (Forman & Parry, 2001), (Pierce et al., 1992), (Lichti et al., 2000), (Petrov et al., 1998). Other than some existing 3-D technologies based on CCD 2-D image reconstruction, these 3-D sensors are based on laser scanning and geometric methodologies such as triangulation (Petrov et al., 1998), (Tyrsa et al., 2006, a), (Tyrsa et al., 2006, b), (Rivas et al., 2008) and have achieved more and more attentions in machine vision application due to its robustness and simplicity. However, to expand functional application of these laser sensors, powerful algorithm add-on is still needed to effectively process measured data from them, normally depending on individual application case.
Laser triangulation principle (Handbook of Optical and Laser Scanning, 2004) in general can be based on two schemes represented in Fig 4 (a and b). The first one uses a fixed angle of emission and variable distance; the second one, on the contrary, fixed triangulation base and variable scanning angle.
The first one works as follows.
A laser beam is projected onto the measurement surface (Fig. 4, a), where it is scattered from the surface and its image is detected by an optical detector, usually a CCD camera. By using a suitable angular arrangement between the laser and sensor positions, the detected location of the laser spot on the image plane produces an accurate measurement of the distance between the sensor and the surface. Therefore, the profile of a surface can be measured by using laser triangulation. The laser beam is made to scan across the surface of the object. The

range data at each location is calculated according to its position within the image plane so that the whole 3-dimensional profile of the surface can be obtained. The positioning of the laser beam is normally controlled by an adjustable mirror system, which is able to change the angular direction of the laser beam over a 2-dimensional plane.



a)



b)

Fig. 4. Two principles of laser triangulation (a – "with fixed angle of emission"; b - "with fixed triangulation distance", consists of a laser dot or line generator and a 1D or 2D sensor located at a triangulation distance from the light source (Petrov et al., 1998). Optionally, the laser light can be steered by a motor)

A second one basic triangulation scheme (Fig. 4, b) is an active optical triangulation 3D digitizing systems, which visualize real-life objects. These active optical systems provide photorealistic representations of shapes and textures with reasonable speed. Figure 4,b shows a simple diagram of a triangulation scanner. The laser beam—reflected from a mirror—is projected on the object. The diffusely reflected light is collected by the sensor, which is a linear array if a laser dot is projected or a 2D matrix (typically a charge coupled device camera) if laser stripes are projected.

The laser positioning circuitry controls the angle α and is known. The angle β is determined in the sensor measurements if the focal length f and the detector pixel size are given. The triangulation distance—the distance between the sensor and the mirror—is also known.

As Figure 4, b shows, since all geometric parameters are known, the x, y, z coordinates of the point on the object can be computed in a trigonometric fashion. If a single laser dot is projected, the system measures the coordinates of just one point of the object. When a laser stripe is projected, all points along the stripe are digitized. The basic triangulation scheme can be enhanced to improve the optical quality and depth of field (see www.vit.iit.nrc.ca for several good examples). The modifications, however, require custom-manufactured components and are relevant mostly for scanners used in high-accuracy reverse engineering tasks. In general, any other kind of structured light can replace the laser dot or stripe. For example, several dots or laser stripes can be projected. However, if the system projects multiple patterns, it's difficult to identify individual elements. If, say, two stripes are projected, then the image processing software must separate the first and second lines. Solutions to the identification problem include using a sequence of different colored stripes. Such schemes typically become sensitive to ambient light, as we'll discuss later.

The most well-known optical triangulation 3D scanner is the one developed by Cyberware of Monterey, California. This legendary scanner was used by a generation of computer scientists. The company has maintained essentially the same product line for more than 10 years. Cyberware products can capture photorealistic images of a range of objects—from apple-size models to fullsize human bodies. The scanner head contains a laserline generator, a system of mirrors, and black-and-white and color video cameras. Scanning occurs by moving the object on a rotation and translation platform, or by moving the sensor around the object in a circular motion. In the basic Cyberware scanner model, a system of mirrors collects laser light from left and right triangulation paths relative to the laser. This scheme helps avoid shadows in the scans caused by the triangulation angle. However, it imposes strict requirements on the optical assembly's quality and the system's calibration, and increases the scanner's size. The scanners are complex, not portable, and prohibitively expensive for many applications. One version, the full-body scanner, digitizes a complete human body as a combination of four scans in about 17 seconds. Each scan has 250 × 1000 points of resolution. The four scans can be glued using commercial software packages (Petrov et al., 1998).

Another typical method for laser scanners is a Circular laser scanning vision sensor for detecting position of singular points (Tao Dai et al., 2005). Usually it called data acquisition device-circular scanning sensor (CSS). By singular point in this case we mean small convex or concave points deviated from smooth surface manifold, caused by either faulty or normal production/machining. The platform for this devise features a three degree of freedom (hence, accommodate 3-D) motion control through a processor, three servo motors and a circular scanning sensor (CSS), as well as a computing system. An emulated head is attached rigidly with CSS to indicate the tracking and focusing actions. The circular scanning sensor (CSS) (Figure 5), with embedded data acquisition device, is a laser scanning multi-axis machine vision device which produces 3-D data sequences of a target work piece for reconstruction of geometry characteristics.

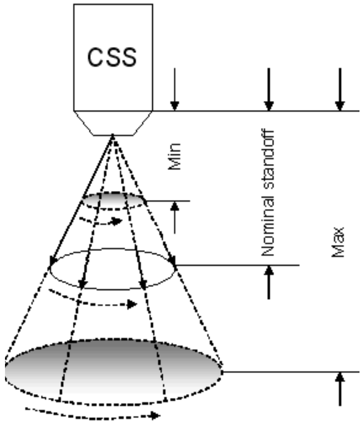


Fig.5. A Laser Circular Scanning Sensor and illustration of Circular Scanning Principle

The CSS and the emulated head are driven by three servo motors to generate 3-D (*x*, *y*, *z*-axis) motion: arbitrary planar scanning route, up and down motion. The main emulated head could track any 3-d singular point on the surface of the work piece while conducting scanning and concentrate on it at specified focusing distance.
The motors are controlled by a central processor which receives location information of singular points from a PC system. The loop is then closed by feedbacking 3-D geometric data of target points acquired by the CSS to the PC system which processes the data sequence continuously using DMA based algorithm (Fig. 6.) to return the location information of singular points, if any, on the surface of work pieces.

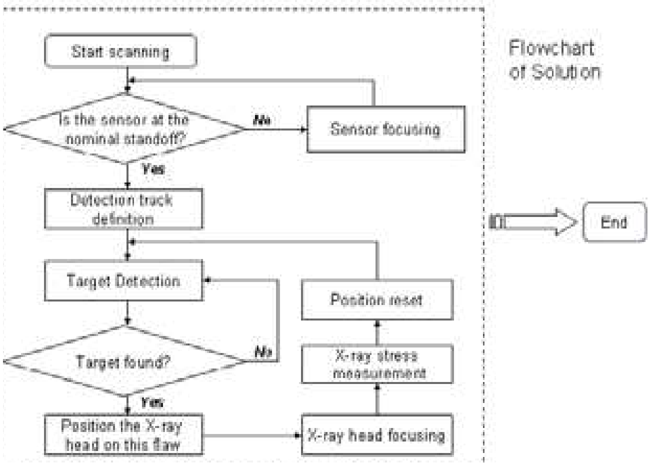


Fig. 6. Flow-Chart of DMA Based Operating Algorithm

For central laser scanning sensor, scanned data is obtained in sequence:

$$Xi = [x_0, x_2, \cdot\cdot\cdot, x_{N-1}],\ i = 1, 2, \cdot\cdot\cdot, \tag{4}$$

where Xi indicates the $i - th$ data sequence captured and N is the number of scanning times called the depth of the data sequence. This data sequence could represent data piece acquired by one sensor in different pre-set time period or by multiple sensors at same time. For example, a line or circular scanning laser sensor generates a geometric data sequence of objects in one scanning cycle. For circular scanned data is considered, hence the depth N is also the scanning cycle period. An example of a 1-D real circular scanning data cycle is illustrated in Figure 7, which is conducted, for example, on a horizontally flat iron plate placed on a table.

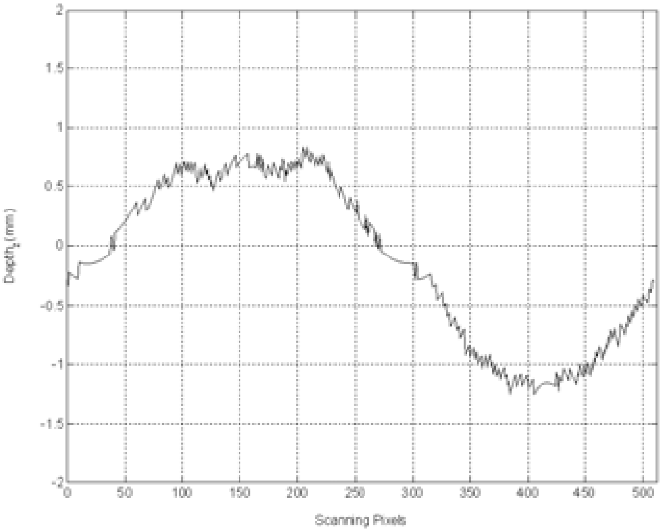


Fig. 7. A 1-D Circular Scanning Data Cycle

In ideal case, one would expect scanned data yields a horizontal line. However, this hardly happens in reality. Instead, as Figure 7 shows, a pseudo-sinusoid shape is presented reflecting either unevenness of plate surface or table leaning, or both. In the case that there is a hole-like singular point, caused by either normal or faulty machining, on the surface, the scanned data would normally reflect that point with a peak significantly deviated from the profile, unlike small peaks around the profile induced by noise. Such an example with a hole of depth 0.4 mm at the 100th pixel is shown in Figure 8 (1-D circular scanning data).
From these two examples, it is clear that two major tasks have to be performed in order for us to obtain accurate information of targets from laser scanned data: reduction of effect of noise on the measurement and detection of singular peak.
This method even it is a laser scanner practically is similar to 3-D technologies based on 2-D image reconstruction and posses all weak points, which are mentioned for these group of methods.
For static object shape monitoring also possible to use a multi-beam laser triangulation (Noll & Krauhausen, 2003) up to more than 120 laser beams are used to simultaneously measure geometric features of moving products such as e.g. the thickness, profile or surface topology. The laser beams form a series of measuring points or lines on the object surface.

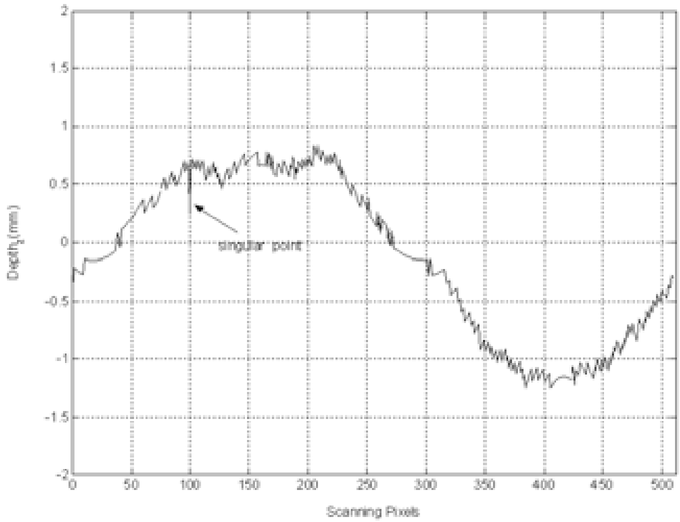


Fig. 8. A 1-D Circular Scanning Data Cycle with a Singular Point

By temporal modulation of the pulse width of the laser diodes and control of the exposure time of the CCD or CMOS detectors **a** dynamic range of more than $1 \div 10^4$ (without considering the dynamic range of the detector elements itself is realized enabling the measurement of object surfaces ranging from black to metallic bright. The use of different laser wavelengths allows suppressing cross-talking originating from overlapping laser projections on the surface of the specimen. However, this method is a good tool for dynamic surface control, but it is unsuitable for high-grade image reconstruction. It cannot be related to pure "camera methods" because CCD or CMOS devices are used only like "light 1/0 detectors", also for incident ray preliminary positioning. This example highlights those sometimes optical sensors matrixes are the better detectors because of their good sensitivity/resolution.

Another laser tools for static perimeter control is Laser radar (LADAR) (Sahba et al., 2006), (Stutz, 2005). LADAR scanning is far advanced over conventional active infrared (AIR) sensing. The transmitter and receiver units are usually coaxially located, making the installation less prone to misalignment or tremors than traditional opposed or retro-reflective sensor posts. Laser scanning also has much finer resolution that RADAR and microwave-RADAR, allowing the definition of the intruding target size. Time of flight pulse detection or continuous wave phase shift methods are used to derive range measurements. According to tests conducted to determine the feasibility of using laser scanning for perimeter protection (Sahba et al., 2006), it has been confirmed the capability of detecting humans and vehicles, with a maximum range of 25m and 80m respectively. The testing demonstrated large area coverage, the ability to determine intruder size, the definition of multiple detection zones, a good detection rate and low false alarm rate (Hosmer, 2004). It has also been claimed that camouflage and other methods for masking an object can deceive near infrared devices, but it is impossible to mask a physical volume moving through space, thereby implying that by scanning a plane at high speed and resolution, any object moving through this plane will be detected (Hancock et al., 1998). Riza, et. al. propose generic criteria for an ideal optical scanner including a wide scanning range, no moving parts, high resolution, high speed and low power consumption. Additionally, they place significance on reconfiguration (Riza & Muzammil, 2003). By projecting many spots, a laser curtain is produced and multiple range measurements from a dense array of laser spots provide a 3D contour of the perimeter. In current LADAR scanners, a laser beam is deflected by rotating mirrors as show in Fig. 9 where D is the incident beam diameter, a is the beam-feed angle, θ is the active scan angle and L is the scan length.

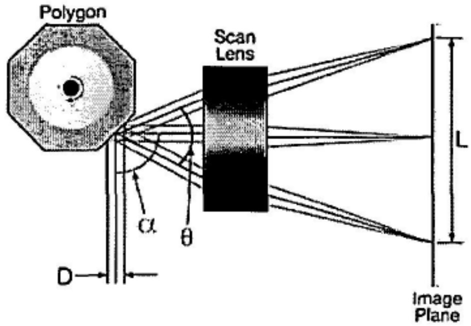


Fig. 9. Principal of laser beam deflection using a rotating mirror (Stutz, 2005)

As we can observe, commonly static object monitoring methods strongly related to basic optical property of focal distance and its practical consequence of a sharp image in a focal plane. From the other hand, it makes these technical solutions initially (apriori) unsuitable for such dynamic task like a mobile robot vision.

In applications where information from more than one physical point of an object needs to be derived, a scanning mechanism must be employed to deflect the light beam from the laser source, towards the desired point of interest on the object surface. This deflection is usually driven in two or three dimensions electro-mechanically. Laser scanning mechanisms have been used in a large variety of applications ranging from photo and document scanners to airborne topographic mapping, surveying of buildings and plants, counting and collision prevention. Approximately over the past ten years laser scanning has found new applications in photoelectric security sensing. Some common uses include providing light curtains in front of assets on a wall display or over the facade of a building. More commonly, laser scanners are employed to provide high resolution range measurements to the outer perimeter (e.g. fencing, walls) of a building and detect a change in range values when beams are intersected by an intruder.

To date, driving the deflection mirror in the desired direction has been accomplished using different devices such as electro-mechanics, rotating polygon shaped mirrors and galvanometers. Figs. 10. (c) and (d) show the layout of a typical commercially available laser scanner and its mirror driving mechanics. Typically, the whole optical head is rotated by a servo motor head around the azimuth in the order of a 1000 rpm, and the mirror rotates in elevation at the same time. This dual movement deflects the beam in the desired directions. However, the intrinsic motion of the mirror causes scanning problems which are the hardest to quantify and correct. Acceleration time taken to reach the constant scanning speed from a stationary position can result in range measurements being stored at the wrong corresponding points in the scanned image. Additionally, all mirrors and measurement components must be synchronized exactly. In general, laser scanning is much more sensitive to vibration than a multi-beam stationary optic approach. In particular, mirror device scanners are slow, bulky and expensive and being inherently mechanical they wear out as a result of acceleration, cause deflection errors and require regular calibration. Stutz in (Stutz, 2005) explains that the performance of polygonal scanners, especially with respect to maintaining an accurate deflection beam path, is prone to manufacturing tolerances.

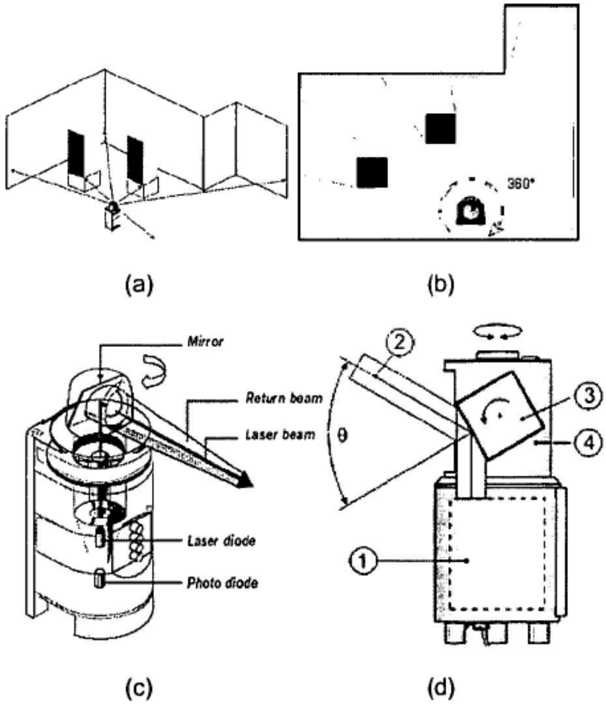


Fig. 10. (a) and (b) indoor room surveillance scanning, (c) scanner with mirror mounted on motor, (d) scanner with rotating mirror where 1 contains the range finding electronics, 2 is the deflected beam, 3 is the rotating mirror and 4 is the complete optical head which also rotates (Sahba et al., 2006)

Dynamic track and jitter errors are caused by tolerances for:

- Polygon machining errors
- Mounting errors
- Mounting-hub errors
- Random wobble caused by ball bearings
- Motor cogging
- Torque variations
- Noisy start of scan circuit

Machining and mounting errors directly cause the deviation of the deflection beam from the desired path of propagation towards the correct point at the image plane. Errors can range from 1 to 60 arc-seconds in angular deviation. Cogging, torque and facet flatness variations cause errors in the actual scan line.

Other problems listed with rotational scanners are as follows (Stutz, 2005):

- Synchronization with other time-dependent elements in the system is rather difficult.
- Motor stability and durability at higher rotation speeds also present problems.
- There is an upper limit to the rotation speed due to the tensile strength of the mirror material. The mirror must not disintegrate at the maximum rotation speed.

Another strong tool for 3D spatial data acquisition is terrestrial laser scanning (Slob & Hack, 2004), (Baltsavias, 1999, a), (Baltsavias, 1999, b). These are active scanners based on laser signals for measurement of slant distances to acquire information about the surface of the Earth and objects on it. As mentioned above, there are several laser scanning systems being operational. Concerning the measurement principle two different methods had been realized: *runtime* measurement using pulsed laser signals and *phase difference* measurement using continuous-wave lasers. Because most systems are based on the first mode, it will be described in more detail in the next section. A good overview on both topics can be found in (Wehr & Lohr, 1999). The pulsed mode laser scanner emits pulsed laser light in exactly determined time intervals. The system measures the runtime of these laser pulses, i.e. the elapsed time between emitting a signal and receiving it after reflection on the surface of the Earth or objects on it. Therefore, slant distances can be derived from these time differences by the wellknown formula v = Ds / Dt or Ds = v / Dt. By means of the exterior orientation of the sensor (recorded by differential GPS (dGPS) and INS systems) 3D co-ordinates of the illuminated surface points can be determined.

Laser systems need to be designed mainly regarding two components: the *emitting and receiving unit* and the *positioning unit*. Both will be described as the example by means of the operational system TopoSys (Wehr & Lohr, 1999) which was used to acquire the data sets about terrestrial landscape.

In this system, the emitting and receiving unit is realized by means of a glass fiber optic. The laser light is emitted on a nutating mirror, i.e. a rotating mirror which deflects it on a glass fiber bunch. The ends of the glass fibers are connected to a row-shaped optic, so the resulting measuring pattern on the surface of the Earth is a single scanning line. In addition to the movement of the airplane this results in a strip-wise data acquisition as shown in Figure 11.

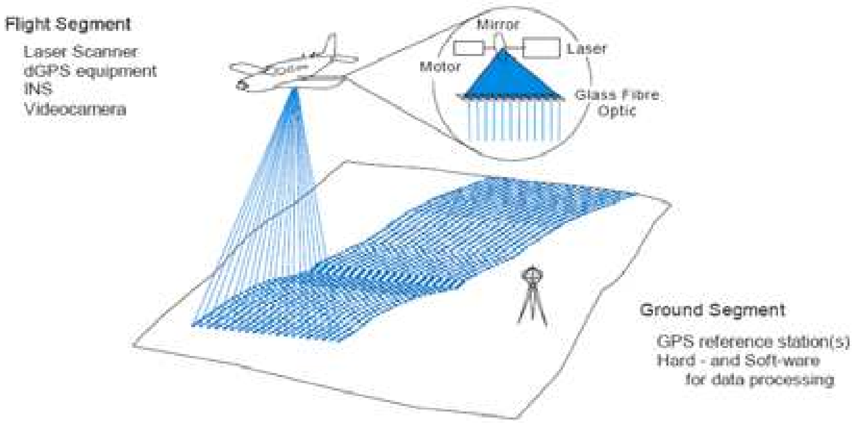


Fig. 11. Laser scanner system TopoSys (Wehr & Lohr, 1999)

The positioning component of the system consists of several elements. As navigation principle differential Global Positioning System (dGPS) is chosen. Therefore, GPS antennas

are mounted on the airplane, as well as on reference stations on the ground. Although this positioning strategy yields to good results concerning the accuracy of the obtained co-ordinates (in the range of some few centimeters), the measurement rate is lower than the one of the laser scanner. Therefore, additionally Inertial Navigation Systems (INS) are used, i.e. navigation units register the rotations of the airplane based on gyros. These are capable to determine a position with a higher temporal resolution.
In Table 1 the performance parameters of this system are listed. It should be mentioned that the system is capable of acquiring up to 5 points per m^2. This results in data sets of high point density and suitable accuracy in position as well as in elevation.

sensor type	pulse modulated laser Radar	range	< 1000 m
scanning principle	fibre optic line scanner	transmitter	solid state at 1.5 µm
measurement principle	run-time measurement	scan frequency	300 Hz (adjustable)
field of view	+/- 7°	number of pixels per scan	127
swath width (1000m flight height)	250 m	accuracy of a single distance measurement	< 0.3 m
accuracies of point coordinates x,y,z	˜ 0.3, 0.3, 0.1 m	resolution of a distance measurement	< 0.1 m

Table 1. Performance parameters of TopoSys laser system (Steinle & Vogtle, 2000)

It is relatively more expensive method over another laser application. Hence the principle limitation of it is the practical importance of the task. However, terrestrial laser scanning have a significant flexibility of provided practical use acts. For example, terrestrial scanning lidar (light detection and ranging) is applied to outcrop stratigraphic mapping enables researchers to capture laser range data at a rate of thousands of individual X, Y, Z and laser-intensity points per second (Bellian et al., 2005).

Fig. 12. A typical field setup in Patagonia in early 2002. All equipment was transported in a backpack by a single person (picture courtesy of Jeremy Willson, Shell Oil). Lithium Ion batteries have helped reduce battery weight to 1,540 grams (3.4 pounds) per approximately 2 hours of continuous scan time (Bellian et al., 2005)

Research group of (Bellian et al., 2005) store the 3D coordinates data set with an Optech Laser Imaging ILRIS 3D terrestrial (ground-based) scanning lidar (Fig.12). This instrument was chosen for its 1 km range, palm-type interface, and light weight. All data from the ground-based lidar surveys were processed on a standard laptop computer with at least a gigahertz processor speed, 1 gigabyte of RAM and Innovmetric Incorporated's Polyworks CAD (Computer Aided Design) software. The principal result of (Bellian et al., 2005) was: high-resolution lidar (approx. 1cm point spacing) is far more accurate in three dimensions than nearly all GPS currently available to the public.

2.3 GPS-based and combined machine vision principles

GPS-based solutions sometimes also are acceptable for machine vision, especially for navigation tasks. The good example of such system is presented in (Vang et al., 2000).
GPS based navigation for autonomous land vehicles has the capabilities of GPS to determine the locations of vehicles, as far as static objects, basing on satellites (Kee et al., 2008), (Stewart & Tsakiri, 2002).
GPS measurements consist of biased and noisy estimates of ranges to the orbiting satellites. The principal source of bias is the unknown receiver clock offset, whereas the remaining errors arise from:

- ′ modelling of the satellite clock and ephemeris.
- ′ modelling of the ionospheric and tropospheric delay.
- ′ measurement of the code and carrier phase influenced by both receiver noise and multipath.

DGPS is a technique that improves the user's position accuracy by measuring the infinitesimal changes in variables in order to provide satellite positioning corrections. It should contain a reference station, a data-link, and user applications. The reference station generates corrections by means of a measured pseudo-range or carrier-range, a calculated distance from the reference station to each satellite, and a satellite clock bias as well as an estimated receiver clock bias, and broadcasts them to the user application. The correction messages contain a Pseudo Range Correction (PRC) for DGPS, a Carrier Phase Correction (CPC) for CDGPS, and their time rates, Range Rate Correction (RRC) (Kee et al., 2008):

$$PRC = -(-b + I + T + \delta R) = d - \rho + \widehat{B} \tag{5}$$

$$CPC = -(-b - I + T + \delta R + N\lambda) = d - \varphi + \widehat{B} \tag{6}$$

Where:

ρ: pseudo range measurement
φ: carrier phase measurement
λ: wavelength of the carrier phase
N: integer ambiguity
d: distance from the reference station to the satellite
b: satellite clock bias error
$\widehat{B}$: estimated clock bias of the receiver
I : ionospheric delay
T: tropospheric delay
δR: orbit error

Unfortunately, the available accuracy for civilian applications is very poor because of all basic components in (5, 6). First of all, because of objective atmospheric delays (ionospheric and tropospheric). Although commercial DGPS (Differential GPS) service increases the accuracy up to several meters, it is still not sufficient for autonomous driving. Hence it is possible use an image processing to compensate this error. In spite of the fact that it is difficult to use image processing for recognition of a complex environment, it is possible to use it to determine some specified objects that haw obviously simple features. For example, lane tracking in highway, forward car tracking have been reported in (Vang et al., 2000). In this study, information about some specified objects is provided by a 3D map which is synthesized based on a real environment. The autonomous land vehicle uses these features to identify the indicated objects from the complex environments so as to position the vehicle's location. Because image processing consumes long time and positioning accuracy cannot satisfy the autonomous driving in many cases, environmental sensors, such as a 3D scanning laser rangefinder: ultrasonic environmental sensors, are also used.
Fig. 13 depicts the autonomous land vehicle used in the experiments (Vang et al., 2000). This vehicle was equipped with a color CCD camera, a GPS unit, a 3D scanning laser rangefinder, ultrasonic sensors, etc.

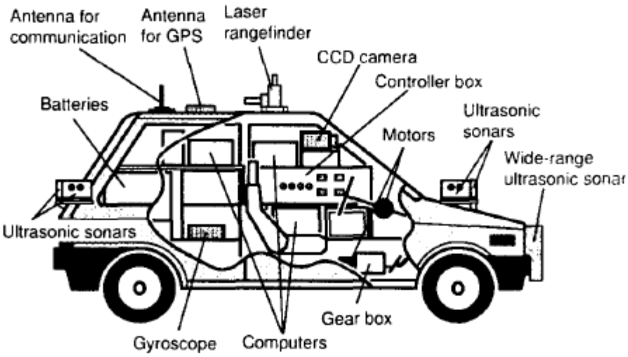


Fig. 13. The autonomous land vehicle used in the experiments (Vang et al., 2000)

Pedals and steering wheel are controlled with servomotors. The environmental sensing and control system is multitools. The position information from DGPS is used to retrieve environmental data from the 3D map. The 3D map provides the information about landmarks and features of some targets. Based the target features, the image processing system identify the landmarks by which the position of the vehicle is determined. An encoder attached to the wheel of a tire and a gyroscope are used for determination of the vehicle position and attitude. Because draft error may become unallowably large, these sensors are only for short distance measurement. A 3D scanning laser rangefinder is employed for compensation of the image processing system. The computers control the steering based on the information from the environmental recognition.
Positioning of GPS uses satellites. The resolution of the GPS used in the experiments is about 0.18m. However, the accuracy available for civilian applications is very poor. Even though DGPS makes the accuracy improved greatly, the positioning error is still insufficient for automatic driving. The positioning accuracy is not so good to navigate a vehicle, but enough to indicate the vehicle t o search some landmarks around the location. The synthesized 3D map provides geographic information that is referred by DGPS for positioning the vehicle's location. Because of the error of DGPS, some other information has