

Enhancing the vision segment architecture for AI-Robots using an orthogonal algorithm

M.Mohamed Sirajudeen
Periyar University,
Salem, India

Abstract: Object capture mechanism for robots is closely related to its vision architecture. Clarity and quality of the incoming object or object come closer to contact help to enhance the identification process. High quality and resolution cameras are used to improve the effectiveness of the acquired objects by the robots. In this paper, we focus on the angular specification of vision segment of the robots by using orthogonal algorithm. Configure the angular movement of eye ball rotation serve to cover all the directions and capture the object in an effective manner.

Key words: Vision, object and Robot

1. INTRODUCTION

Vision sensor in the robots architecture can be used to identify the object closer to its visible region. The inner components of the vision segment comprise of a high resolution cameras. For example, Contact Image Sensors (CIS) are a relatively recent technological innovation in the field of optical flatbed scanners that are rapidly replacing CCDs in low power and portable applications. As the name implies, CISs place the image sensor in near direct contact with the object to be scanned in contrast to using mirrors to bounce light to a stationary sensor, as is the case in conventional CCD scanners.

A CIS typically consists of a linear array of detectors, covered by a focusing lens and flanked by red, green, and blue LEDs for illumination. The use of LEDs allows the CIS to be highly power efficient, allowing scanners to be powered through the minimal line voltage supplied via a USB connection.

CIS devices typically produce lower image quality compared to CCD devices; in particular, the depth of field is greatly limited, which poses a problem for material that is not perfectly flat. However, a CIS contact sensor is smaller and lighter than a CCD line sensor, and allows all the necessary optical elements to be included in a compact module, thus helping to simplify the inner structure of the scanner. With a CIS contact sensor, the scanner can be portable, with a height of only around 30 mm. CIS is a both a key component of, and widely used in, scanners (especially portable scanners), electrographs, bar code readers and optical identification technology.

The proposed architecture insists the angular specification of the inbuilt digital camera in the place of vision for robots.

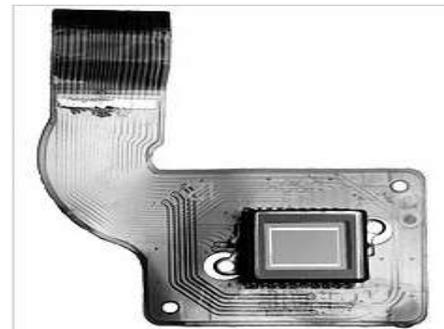


Fig 1 .0 A CCD image sensor on a flexible circuit board

[source:http://en.wikipedia.org/wiki/Contact_image_sensor].

2. RELATED WORK

2.1 Methods of Image capture – an overview

Since the first digital backs were introduced, there have been three main methods of capturing the image, each based on the hardware configuration of the sensor and color filters. The first method is often called *single-shot*, in reference to the number of times the camera's sensor is exposed to the light passing through the camera lens. Single-shot capture systems use either one CCD with a Bayer filter mosaic, or three separate image sensors (one each for the primary additive colors red, green, and blue) which are exposed to the same image via a beam splitter.

The second method is referred to as *multi-shot* because the sensor is exposed to the image in a sequence of three or more openings of the

lens aperture. There are several methods of application of the multi-shot technique. The most common originally was to use a single image sensor with three filters (once again red, green and blue) passed in front of the sensor in sequence to obtain the additive color information. Another multiple shot method is called Microscanning. This technique utilizes a single CCD with a Bayer filter but actually moved the physical location of the sensor chip on the focus plane of the lens to "stitch" together a higher resolution image than the CCD would allow otherwise.

A third version combined the two methods without a Bayer filter on the chip. The third method is called *scanning* because the sensor moves across the focal plane much like the sensor of a desktop scanner. Their *linear* or *tri-linear* sensors utilize only a single line of photo sensors, or three lines for the three colors. In some cases, scanning is accomplished by moving the sensor e.g. when using Color co-site sampling or rotate the whole camera; a digital rotating line camera offers images of very high total resolution.

The choice of method for a given capture is determined largely by the subject matter. It is usually inappropriate to attempt to capture a subject that moves with anything but a single-shot system.

However, the higher color fidelity and larger file sizes and resolutions available with multi-shot and scanning backs make them attractive for commercial photographers working with stationary subjects and large-format photographs. Dramatic improvements in single-shot cameras and raw image file processing at the beginning of the 21st century made single shot, CCD-based cameras almost completely dominant, even in high-end commercial photography. CMOS-based single shot cameras remained somewhat common. [Source: http://en.wikipedia.org/wiki/Digital_camera#Methods_of_image_capture].

2.2 Planar Position Recognition and Identification – an overview

The 2D Robot Vision System is a standard system for identification and/or position recognition without direct contact by the system. It delivers reliable information on position, angle, and variation of any part. The system's key benefit is its extremely high recognition speed at sub-pixel precise contour extraction.

Further development of the 2D technology resulted in an intelligent recognition method that enables determination of a part's height using just a single camera and a single camera position. This technology, called 2½ D Robot Visions, results in a system that is affordable, and yet, at the same time complete. 3D palletizing is a typical application using of 2½ D Robot Vision. In the past, the available processes and equipment for 3D measurement were too complex, too expensive and thus not economically efficient for many applications. Until now, at east two - and generally more - cameras, including the necessary peripheral equipment, were required to define the positions of objects in a room.

This is why they were only used, primarily, for highly sophisticated tasks that justified the expenditure, such as the position definition of entire auto bodies in the automotive industry. MONO3D is the solution to this problem.

The new 3D robot guidance process operates with only one camera, which opens up entirely new perspectives for cost-effective image processing applications in all the important industries.

It is now possible, from one single captured image, to precisely define a three-dimensional object based on the measurement of only three criteria in all six degrees of freedom (position and orientation).

- Increased accuracy
- Greater flexibility
- Cost-effective automation

Many modern production processes in manufacturing automation are highly complex. A high level of dimensional and fitting accuracy as well as flexibility is demanded as an extensive range of product variations must be produced on a single line in a short period of time.

The patented, revolutionary system combines information from several camera systems to determine the position of large objects in any given space at extremely high precision and speed. Its function includes the recognition of entire car bodies. 3D Robot Vision identifies variations of a given position in all six degrees of freedom.

The optical 3D stereo sensor can determine positions and measure objects in all six degrees of freedom by using edges, holes, curves or parts of the object that can be described by characteristic elements, edges, lines or polynome-like path contours.

The sensor can be mounted stationary as well as on the robot arm or handling system. To achieve even higher precision, multiple sensors can be used simultaneously.

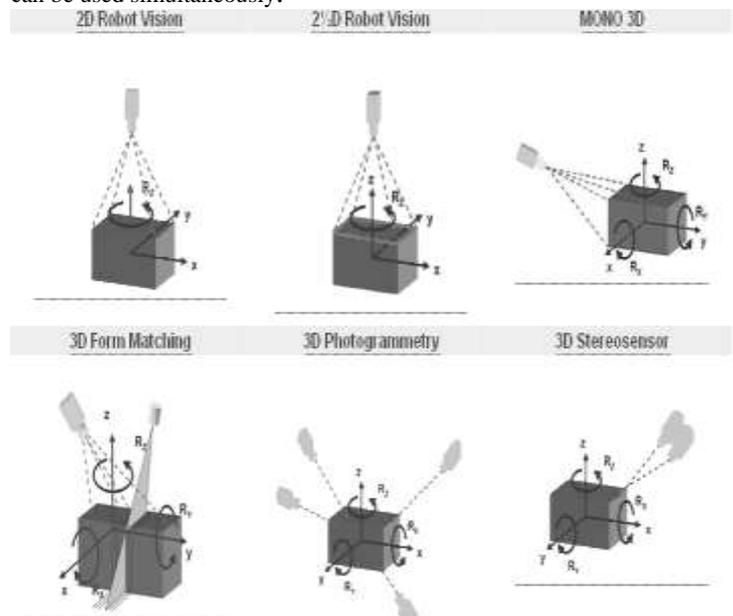


Fig 1.1 Existing vision segment structure [Source: <http://isravision.com>]

3. PROPOSED WORK

3.1 Architectural layout for the proposed vision segment :

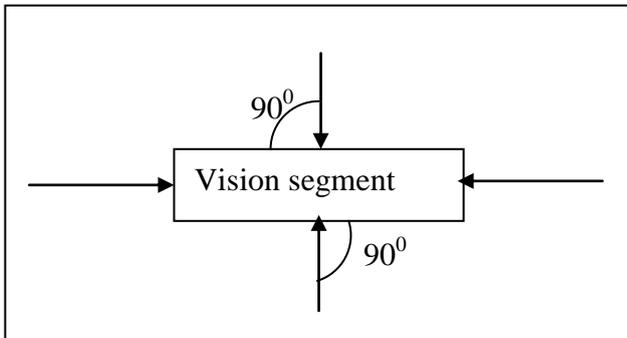


Fig 1.2 angular specification for the vision segment

The enhancement for the object/image capturing depends on the incidental and reflectional angle position for the incoming object/image. Identify the place to get the effective object/image structure help to improve the quality of the captured object.

The storage method for the incoming object helps us to reduce the complexity of the identification process. The boundary level points are used to perform the segmentation regarding to identify the layout of image.

The similarity between the physical segment of the stored object and the logical segment of the incoming object yields the result for identification process.

From the classification for the object storage make simplify the processes for identification. For example, ground based objects are occupied in one part of the memory; fly – based objects are resided in another part of the memory.

Initially we determine the nature of object enter into the visible region such as stable or moving sate, two-dimensional or three-dimensional objects either in different direction will be taken in to the account for further processing.

In the first stage any one of the segment will be taken into account for the comparison and get the probability to reach the reality of the image.

The identification process will start from a guess. Before, we start the comparison; the method of object storage will be specified by the algorithm without any ambiguity.

For this reason, the 100 percentage of success will depends on the storage classification process and if the object is new , then immediately allocate a new storage for the incoming object and display the message “ it is new in this zone.. “ along with the similarity for the image which one is already occupied into the memory.

A number of large-scale experiments [2], involving separate test images that evaluate performance with increasing number of items in the database, in the presence of clutter, background change, and occlusion, and also the results of some generic classification experiments where the system is tested on objects never previously seen or modelled.

Therefore, to avoid the process of store the entire image in the permanent storage part. The visual part of the device is deigned to cover all the directions like a rotating eye ball.

3.2 Specification for the Orthogonal Image Segment Comparison algorithm:

In order to use the system with an object, it’s any one of the view to be stored in the memory. Currently, this is done by obtaining a number of segments of the object from different directions. About a set of views are needed to cover the entire viewing sphere for the curve-based keys we have used.

Here, the searching technique for text and image will be performed by using a rotation of the image segment or text instead of using whole part of it. The way to segment and to store it will determine the efficiency of the algorithm.

The searching process will enter into the appropriate region is reduce our work for more than 80%, unless the progress will become a complicated and not possible to produce the expected result at end.

The categorization of the object will also be performed in a vey careful manner based on the standard characteristics of the nature of the objects.

An orthogonal matrix [9] is the real specialization of a unitary matrix, and thus always a normal matrix. Although we consider only real matrices here, the definition can be used for matrices with entries from any field. However, orthogonal matrices arise naturally from inner products, and for matrices of complex numbers that leads instead to the unitary requirement. Orthogonal matrices preserve inner product. So, for vectors “u, v” in an *n*-dimensional real inner product space,

$$\langle u, v \rangle = \langle Q_u, Q_v \rangle [10]$$

To see the inner product connection, consider a vector *v* in an *n*-dimensional real inner product space. Written with respect to an orthonormal basis, the squared length of “b” is “b^T b”. If a linear transformation, in matrix form *Q_b*, preserves vector lengths, then

$$v^T v = (Q_v)^T (Q_v) = b^T Q^T Q_v$$

The simplest orthogonal matrices are the 1×1 matrices [1] and [-1] which we can interpret as the identity and a reflection of the real line across the origin.[11]

The 2×2 matrices have the form

$$\begin{pmatrix} p & t \\ q & u \end{pmatrix}$$

Which orthogonality demands satisfy the three equations, [12]

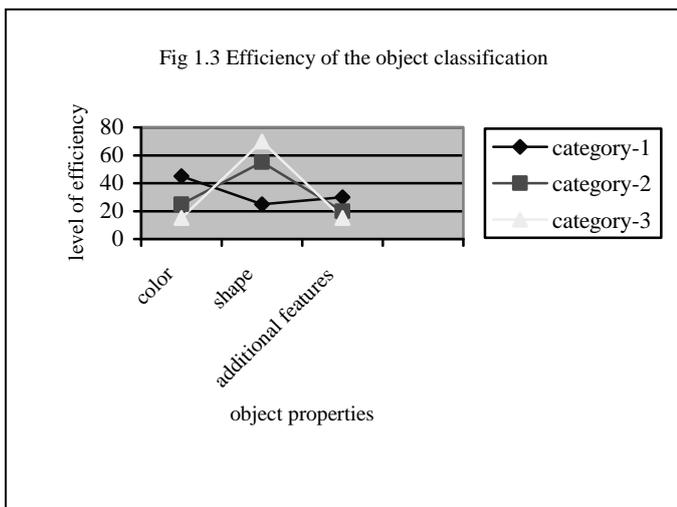
$$\begin{aligned} 1 &= p^2 + q^2 \\ 1 &= t^2 + u^2 \\ 0 &= pt + qu \end{aligned}$$

In consideration of the first equation, without loss of generality let $p = \cos \theta$, $q = \sin \theta$; then either $t = -q$, $u = p$ or $t = q$, $u = -p$. We can interpret the first case as a rotation by θ (where $\theta = 0$ is the identity), and the second as a reflection across a line at an angle of $\theta/2$. [13]

The reflection at 45° exchanges x and y ; it is a permutation matrix, with a single 1 in each column and row (and otherwise 0): [13]

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

The identity is also a permutation matrix [15]. A reflection is its own inverse, which implies that a reflection matrix is symmetric (equal to its transpose) as well as orthogonal. The product of two rotation matrices is a rotation matrix, and the product of two reflection matrices is also a rotation matrix.



4. Conclusion

The main objective of the work is to reduce the storage space for the incoming object and also avoid the complexity of multiple view storage representations in the AI-Robot's memory architecture. In general, the required storage space for the entire object take place huge amount of memory space and increase the complexity for the searching process. Instead of this, we consider only the segment of the incoming object for the identification process by using Orthogonal Image segment comparison algorithm.

References

[1]. Andrea Selinger and Randal C. Nelson, "A Perceptual Grouping Hierarchy for Appearance-Based 3D Object Recognition", Computer

Vision and Image Understanding, vol. 76, no. 1, October 1999, pp.83-92. Abstract, gzipped postscript (preprint) .

[2]. Randal C. Nelson and Andrea Selinger "Large-Scale Tests of a Keyed, Appearance-Based 3-D Object Recognition System", Vision Research **Special** issue on computational vision, Vol. 38, 15-16, Aug. 1998.

[3]. Abstract, gzipped postscript (preprint) Randal C. Nelson and Andrea Selinger "A Cubist Approach to Object Recognition", International Conference on Computer Vision (ICCV98), Bombay, India, January 1998, 614-621.

[4]. Abstract, gzipped postscript, also in an extended version with more complete description of the al

[5]. Randal C. Nelson, Visual Learning and the Development of Intelligence, In Early Visual Learning, Shree K. Nayar and Tomaso Poggio, Editors, Oxford University Press, 1996, 215-236.

[6]. Abstract, Randal C. Nelson, "From Visual Homing to Object Recognition", in Visual Navigation, Yiannis Aloimonos, Editor, Lawrence Earlbaum Inc, 1996, 218-250.

[7]. Abstract, Randal C. Nelson, "Memory-Based Recognition for 3-D Objects", Proc. ARPA algorithms, and additional experiments. Image Understanding Workshop, Palm Springs CA.

[8]. The Architecture of Brain and Mind Integrating Low-Level Neuronal Brain Processes with High-Level Cognitive Behaviours, in a Functioning Robot2.

[9]. Diaconis, Persi; Shahshahani, Mehrdad (1987), "The subgroup algorithm for generating uniform random variables", Prob. in Eng. and Info. Sci. 1: 15–32, ISSN 0269-9648.

[10]. Dubrulle, Augustin A. (1999), "An Optimum Iteration for the Matrix Polar Decomposition", lect. Trans. Num. Anal. 8: 21–25, <http://etna.mcs.kent.edu/>

[11]. Golub, Gene H.; Van Loan, Charles F. (1996), Matrix Computations (3/e ed.), Baltimore: Johns Hopkins University Press, ISBN 978-0-8018-5414-9

[12]. Higham, Nicholas (1986), "Computing the Polar Decomposition—with Applications", SIAM J. Sci. Stat. Comput. 7 (4): 1160–1174, doi:10.1137/0907079, ISSN 0196-5204, http://locus.siam.org/SISC/volume-07/art_0907079.html

[13]. Higham, Nicholas; Schreiber, Robert (July 1990), "Fast polar decomposition of an arbitrary matrix", SIAM J. Sci. Stat. Comput. 11 (4): 648–655, doi: 10.1137/0911038, ISSN 0196-5204, http://locus.siam.org/SISC/volume-11/art_0911038.html