MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

A New Retina-Like Visual Sensor Performing the Polar Transform

Cheon Woo SHIN and Seiji INOKUCHI

Department of Systems Engineering , Osaka University
1-1, Machikaneyama, Toyonaka, 560 JAPAN
shin@sys.es.osaka-u.ac.jp

Abstract

This paper describes the development of an anthropomorphic visual sensor with retina-like structure to perform the polar mapping. The sensor consists of a dove prism for image rotation and a linear CCD sensor with 512 pixel/line resolution, and holds approximately 45 Kbytes of image data. The retina-like sensor has variable resolution with increasing density towards the center of the visual field, and performs the polar transforming directly. The motion equation that relates the egomotion and/or the motion of the object in the scene to the optical flow is considerably simplified if the velocity is represented in a polar coordinate system, as opposed to a cartesian representation. Development of this sensor holds promise in application to high speed tracking systems, such as the eyes of navigation robots, because it has a data reduction characteristics and mapping function onto the polar plane.

1. Introduction

In the human visual system, object image is clearly perceived the center of the visual field but not perceived at the fovea field. This occurs because retina receptor increases the density towards the center of the visual field, and decreases the density from the fovea towards the visual field periphery. The advantage of the anthropomorphic visual system is that a considerable reduction in data requirement is obtained with non-uniform sampling in conjunction with high resolution in the field of view corresponding to the focal point. Furthermore the radial component of the optical flow is represented on the polar plane, velocity of moving object is directly and simply computed on the polar plane.

This type of visual system allows efficient data processing with a minimum of image input data, thus reducing the processing time required for obtaining precision image data because the center position has a higher density of sensors. The sensor may prove useful in computer vision as well. Indeed a number of researchs has shown that it can be used in object recognition, computer graphics and motion stereo. It therefore holds promise for application to real time image processing, for example in the field of robot navigation.

The research team of Pennsylvania University, IMEC Corp., and Massimo Tistarelli and Guilio Sandini of Genoa University in Italy employed a retina-like visual sensor constructed using a semiconductor CCD sensor[16]. They studied the application of this system in mobile robots and high-speed navigation robot visual systems [2,3,7]. But the semiconductor CCD anthropomorphic visual sensor has some limitations. One is the CCD cell dimension problem, since the center position of the CCD cell has a small dimension but the fovea position CCD cell has a large dimension. Therefore, the output response of the sensor at the same light source is different, depending on the distance from fovea to center. Another problem is the resolution required to find a target or distinguish an object for the robot, a limitation due to semiconductor fabrication techniques.

Another solution to conduct the retina-like sensor performs the polar mapping is simulate the conventional image data to polar mapping[10, 14, 17]. This is convenient to perfoem the polar mapping using conventional camera and image processing system. But this methods has a some problems like as a low resolution and error is extremely increased in the center of the visual field because of low image pixel depends on CCD area which is important visual area to conceive a object[17].

To overcome the defects memtioned above, resolution problems, different light intensity problems depends on cell dimension and increased error in the center of visual field. The constructed retina-like visual sensor consists of a linear CCD sensor and dove prism to rotate the input image. Therefore the input image via a linear CCD sensor has a high resolution equivilant up to 10000 pixels and rotate the input image using the dove prism to acquire the 2-dimension image data. Consenquencly the acquired input image data has a equivilant light intensity depends on whole area of visual fields and high resolution image data at the center of visual field without error increased defects. Futhermore this sensor can acquire the polar mapping image simply and directly which needs no any transformation to mapping. Developed this sensor holds promise in application to high speed image processing systems, such as the eyes of navigation robots, measure the 3-dimension depth information, real time optical flow measurements and so on.

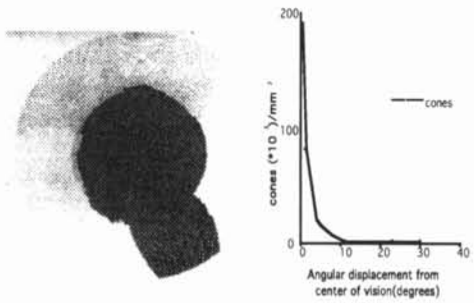


Figure 1 In the human visual system, the density of retina receptors is greater towards the center of the visual field.

2. The Polar Mapping

Regarding the image plane as a complex plane simplifies the discussion of the properties of the proposed grid beacuse geometric transformations are expressible as complex arithmetic. The complex plane is associated with the real cartesian plane by identifying the x- and y- axes with the real and imaginary axes as shown in Fig. 2. Any point in the plane is specified by a single complex number (z) whose real and imaginary parts, Re(z) and Im(z), correspond respectively to the point's x and y coordinates. That is

$$z = x+iy, \qquad \text{where } i = \sqrt{-1}$$

The vector from the orign to the point is represented by a complex number z; its length is called the modulus, written $|z|$ The angle which is called the argument is written Arg (z).

Thus, a complex number can be specifed in polar from as

$$z = r(\cos\theta + i\sin\theta)$$
$$\text{where } r = |z| \text{ and } \theta = \text{Arg}(z)$$

Multiplying two complex numbers yields a results whose modules is the product of moduli, meanwhile whose argument is the sum of arguments.

The notion of function also extends to complex variables; geometric considerations induced by the two-dimensionality of complex numbers are so important that such functions are called mapping. Referring to a region of the z-plane as a domain and a region of the w-plane(points u + iv) as a range, a mapping is a function such as

$$w = f(z)$$
$$w = u(z) + iv(z)$$
$$u(r, \theta) = r$$
$$v(r, \theta) = \theta$$

Mappings are useful in graphics and image processing because a picture can be considered a positive real function over a region of the complex plane. The value of the function at each point of the region is the brightness or gray level of the picture at that point. Geometric transformation of the picture can then be expressed by mapping the points of the original picture and their corresponding gray level range. For example, translation of a picture in the image plane in the direction($\cos\theta_1$, $\sin\theta_1$)through a real distance d_1 can be expressed as

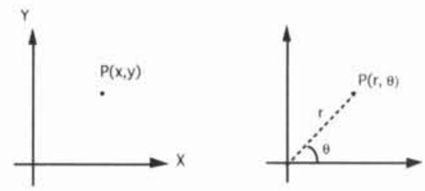
$$w = z + d_1 z_1, \quad \text{where } z_1 = \cos\theta_1 + \sin\theta_1.$$

Rotation through angle θ_1 can be expressed as

$$w = z_1 * z$$

and magnification by a real factor k as

$$w = k * z$$



P(x,y) on the cartesian plane P(r,θ) on the polar plane

Figure 2 The pointe P(x,y) on the cartesian plane and polar plane.

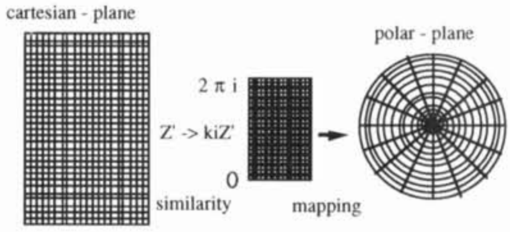


Figure 3 Mapping for the polar plane from the cartesian plane.

There are many attractive features of this mapping. From a mathematical viewpoint, it is the only analytical function which maps a circular region, such as an image on the retina, into a rectangular region. This is a desirable feature for the study and modeling of the human visual systems. The mappings of two regular patterns are shown in Figure 4 to result in a similarly regular pattern. Figure 4(a)shows that concentric circles in an image or in the z-plane become vertical lines in the mapped w-plane. This is because the constant radius, r, at all angles, θ, of the circle gives a constant u coordinate for all v coordinates in the mapped space. Similary in Figure 4(b), an image of radical lines, which have constant angles but variable radii, result in a map of horizontal lines.

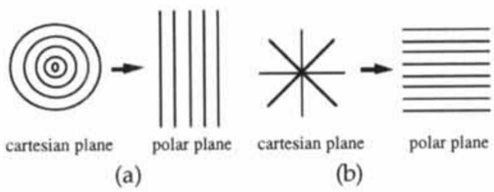


Figure 4 The polar mapping results in the transformation of certain regular patterns in the z-plane into other regular patterns in the w-plane.

Through such mapping, we can find some interesting properties of polar mapping. The first property is rotation invariant, we can see that for a circle in figure 4(a), all possible angular orientation of a point at the given radius will map to the same vertical line. In figure 4(b), as a radial line rotates about the origin, the horizontal line in polar mapping moves only vertically.

Another property is size invariance which can be seen in figure 4. The concentric circles of figure 4 remain vertical lines and moves horizontally as the circles change in size. However, the most important property is projection invariance, where the object translates in 3-Dimension space, the perceived image does not change on the striate cortex only moves vetically.

From the mentaioned invariance properties of polar mapping, estimation of the motion of moving object to the optical flow is considerably simplified since velocity is represented in a polar coordinate system, instead of a cartesian representation.

3. The Structure for Conducting the Polar Mapping system

There are many possible solutions to the construction of a retina-like visual sensor in which the sensitivity of the sensors increases towards the center of the visual field and decreases from the fovea towards the periphery. One is to increase the sensor density toward the center of the visual field using a circular CCD array and semiconductor construction techniques. This method does not allow an increase in density toward the center of the visual field beyond the limit of the semiconductor array density, however.

However, our solution was to employ a linear CCD sensor in conjunction with a prism which rotates the input image. In this way we constructed a retina-like visual sensor in which the density of receptors increased towards the center of the visual field, shown in Figure 5.

In order to create an anthropomorphic visual system using a linear CCD sensor and rotation image, a possible solution is to rotate the CCD sensor and accept linear CCD sensor image data. Another solution is to fix the linear CCD sensor and rotate the input image. Since there are many problems involved in rotating a linear CCD sensor, and constructing the necessary gears, we decided to produce a rotated input image using a dove prism. The proposed structed shown in Fig 5, the variations of N affect both the number of elements necessary to cover a given field and the spatical resolution at a given eccentricity (Fig.6).

The linear relationship between radius image pixel n and eccentricity E can be expressed, in terms of image line number N of equal size elements at a given eccentricity, by the following relation:

$$E = \frac{360 \times a^2 \times N}{\pi a^2 \left(n^2 - (n-1)^2\right)}$$

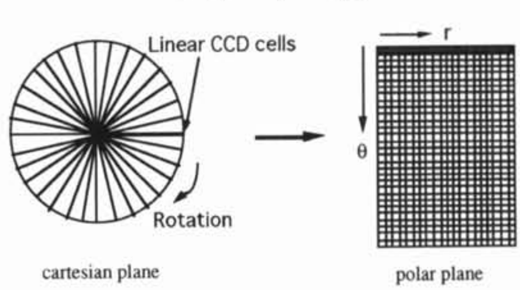


Figure 5 Method of increasing sensor density towards the center of the visual field

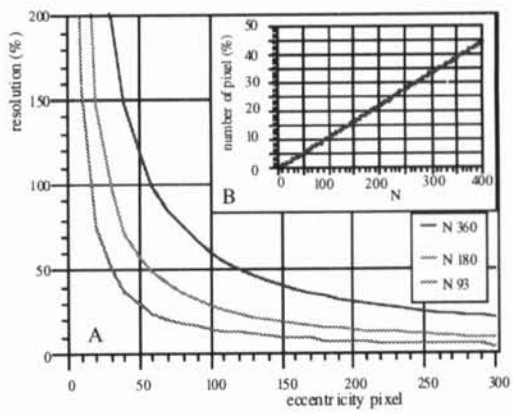


Fig6. (A). Percent variation of resolution as a function of eccentricity for different values of image line N. (B) Percentage of the number of elementer a given field as a function of N.

5. Experiment and results

To create our image rotation system, we employed two 512 pixel linear CCD sensors and a rotation prism. We positioned the linear sensors horizontally to the left and right sides of the camera, and then rotated the prism 90° , as shown in Figure 7. One of the linear CCD sensors therefore received the 180° rotated image.

In figure 10 we can see that the center of the image, which is high in sensor density as compared to the lower density periphery, is clearly displayed. Figure 10 displays the inputed image data directly from CCD the sensor which is transformed on the polar plane and the right image represented on a cartesian plane which is re-mapped. In the image of Figure 10(b), we can see the characters of "*TEC*" that the center part is seen with maximum resolution but decrease in the detection of details occurs with eccentricity. The image pixels obtained from the linear CCD sensor are overlaid in the center of the acquired image. The input image is rotated by the dove prism and detected by two 512 pixel linear CCD sensors located horizontally from the lens center. With 90 lines of about 512 pixels each, one image comprises approximately 45 Kbytes.

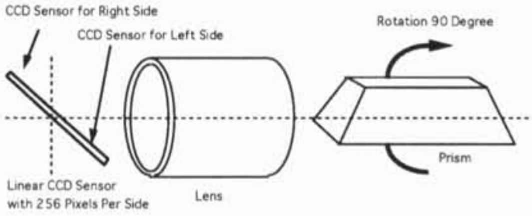


Fig.7 Construction of the image rotation system

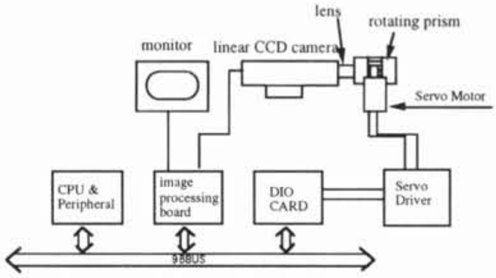


Fig. 8 System configulation to acquire the rotation image for retina-like sensor

6. Conclusion

We developed an anthropomorphic visual sensor in which sensor density increases towards the center of the visual field. The system consists of a dove prism to rotate the input image and a linear CCD sensor. The constructed sensor has a high resolution and eccentricity (the distance from the center of the sensor) independent intensity response specification with increasing density toward the center of the visual field response. It has data reduction characteristics and consisted of a polar mapping data array.

The time required for the linear CCD sensor to acquire a single image was approximately 2.5 msec using a 20 MHz scanning video. The maximum speed to access an image will be about 400 images per second if the mechanical problem in rotating the prism is overcome. We are attemptng to conjugate this sensor for active vision in the eye of a mobile robot and capture passing objects with high speed.

It is believed that this anthropomorphic visual system will prove useful in robot navigation and the tracking of high-speed moving targets.

References

[1] Eric R. Kandel, James H. Schwartz, Thomas M. Jessell, "Principles of Neural Science," Elsevier.

[2] Giulio Sandini, Vincenzo Tagliasco, "An Anthromoporphic Retina-like Structure for Scene Analysis," Computer graphics and image processing 14 PP365-372, 1980.

[3] Massimo Tistarelli and Giulio Sandini, "Direct Estimation of Time-to-imapct from Optical Flow,"DIST, LIRA-Lab Genoa, Italy.

[4] Massimo Tistarelli and Giulio Sandini, "Robot Navigation Using an Anthropomorphic Visual Sensor," IEEE Transactions on Pattern Analysis and Machine Intelligence PP374-381,1990.

[5] Donald L. Sullival , "Alignment of Rotational Prisms,"Applied Optics Vol.11, No.9 P2028-2032, September 1972.

[6] E. DeMicheli, G.Sandini, M.Tistarelli, V. Torre, "Estimation of visual motion and 3D motion parameters from singular points," In Proc. of intl. Workshop on Intelligent Robots and Systems Tokyo Japan, Oct 31 Nov.,1988.

[7] G.Sandini, F.Bosero, F.Bottino, A.Ceccherini, "The use of anthropomorphic visual sensor for motion estimation and object tracking,"

[8] L.Massone, G.Sandini, V.Tagliasco, " "Form-Invariant" Topological mapping strategy for 2D shape recognition", Computer vision, Graphics, and Image Procesing 30, 169-188, 1985.

[9] Carl F. R. Weiman, George Chaikin, "Logarithmic Spiral Grids for Image Processing andDisplay, " Computer vision, Graphics, and Image Procesing 11, 197-226, 1979.

[10] Rameh C. Jain, "Segmentation of Frame Sequences Obtained by a Moving Observer, " IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-6, NO.5, September 1984.

[11] Rameh C. Jain, "Direct Computation of the Focus of Expansion," IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. PAMI-5, NO.1, January 1983.

[12] Tomaso Poggio," Early Vision: From Computational Structure to Algorithm and Parallel Hardware," Computer vision, Graphics, and Image Procesing 31, 139-155, 1985.

[13] Sandra L. Bartlett, Ramesh Jain," Depth determation using Complex Logarithmic Mapping,"

SPIE Vol. 1382 Intelligent Robots and Computer Vision 14: Neural, Biological, and 3-D Methods, 1990.

[14] Ramesh Jain, Sandra L. Bartlett, Nancy O'Brien," Motion Stereo Using Ego-Motion Complex Logarithmic Mapping," IEEE Transactions on Pattern Analysis and Machine Intelligence, VOL. PAMI-9, NO.3, MAY 1987.

[15] Massimo Tistarelli and Giulio Sandini," Dynamic Aspects in Active Vision,"CVGIP:Image Understanding, Vol.56, No. 1, July, pp. 108-129, 1992.

[16] I Debusschere, E.Bronckaers and C.Claeys, G.Kreider and J.Van Der Spiegel, G.Sandini, P.Dario and F.Fantini, P.Bellutti aand G.Soncini, "A Retinal CCD Sensor for Fast 2D Shape Recognition and Tracking," In Proc. 5th Int. Solide-State Sensor and Transducers, Montreux, June 25-30, 1989.

[17] Sandra L. Bartlett and Ramesh Jain,"Depth determation using Complex Logarithm Mapping, "SPIE Vol.1382 Intelligent Robots and Computer Vision 14: Neural, Biologicaal, and 3-D Methods, 1990

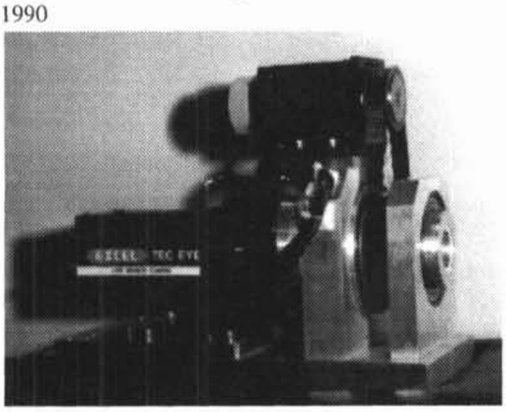


Fig.9 Picture of the anthropomorphic visual sensor

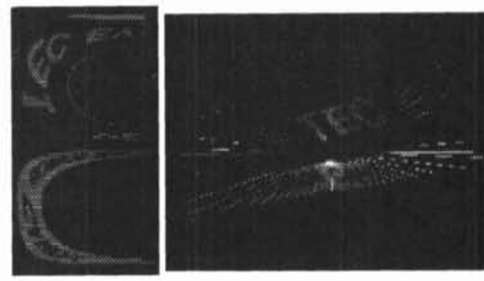


Fig.10 Polar mapping applied to the central of image, represented in Cartesian(left) and polar(r, θ) (right) planes.

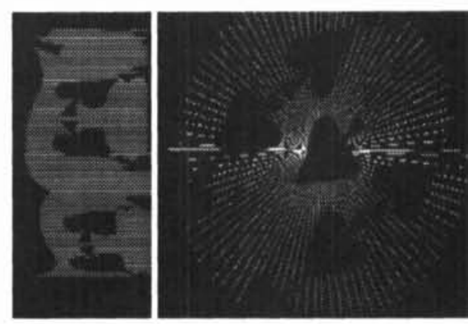


Fig.11 Spade image represented in Cartesian(left) and polar(r, θ) (right) planes.