AN EFFECTIVE LOW-COST VISION SYSTEM FOR THE ROBOT PINGPONG PLAYER

Willy C. K. Choi, Dennis N. K. Leung,
Kelvin S. Y. Yuen and Robin T. Bradbeer
Department of Electronic Engineering,
Citypolytechnic of Hong Kong.
83 Tat Chee Avenue, Hong Kong.
Email:EEKELVIN@CPHKVX.BITNET

Abstract

This paper introduces an effective means of sensing a rapidly moving white pingpong ball against a dark background. The method to determine the centre of gravity of the ball by hardware circuitry is fully explained. A comparison of our method with other conventional methods of sensing a pingpong ball is made. The means to integrate the information from our vision system with the fuzzy set theory so as to predict an optimal location and angle of the bat to hit the ball is also indicated.

1. INTRODUCTION

Nowadays, the thrust of work in industrial robotics is mainly towards enhancement of speed, acceleration and performance. The robot pingpong game was first proposed by Professor John Billingsley of U.K. in 1983 to challenge people in fulfilling special objectives of robotics -- speed and precision. During a robot pingpong contest, speed is everything, where deadlines just have to be met. The vision system must track the ball in real time as it moves across and bounces from the table. At the same time, the position system has to arrive just in time to hit the ball with the right strength so that the ball could fly back toward the opponent's side and land at the correct position.

The standardized size of the pingpong table is 2 metres in length and half a metre in width. Since the ball travels at a fairly high speed of 4-8 ms^{-1} along the table, the whole system must react within 0.3-0.5 second of time. The mechanical part of the pingpong player is comparatively slow. It is necessary for the robot pingpong designer to purchase or create a vision system that can respond fast enough to produce sufficient data for the expert controller to predict the subsequent position and the velocity of the ball far before it reaches the edge of the table.

Since the aerodynamics of the ping-pong ball flying in air is extremely complicated, finding the flight trajectory by solving physical equations would be very time-consuming. Parallel processing using several processors each dedicated for a special job is required in order to reduce the work load and to achieve the necessary speed. Some of the prize-winning designs in the world involve multi-million-dollar vision systems and high-speed computers. Our research at City Polytechnic of Hong Kong has shown that it is possible to achieve the goals of the robot ping-pong by following a different approach which requires less monetary investment and could achieve a result comparable, if not superior, to the conventional methods adopted by other parties -- that is, the use of fuzzy set theory incorporated with a self-designed low cost vision system.

2. POPULAR MEANS OF TRACKING THE PING-PONG

There are different means of sensing the ball in flight. The rotating mirror method is very efficient *in telling* the left-right position of the ball. The plane mirror stripe, which is half a metre in width, lies horizontally on top of the table edge. During rapid rotation, the mirror reflects the light from the ball onto an array of photodiodes lying in parallel with the mirror. An advantage of this method is that it is simple in hardware construction. Image processing is practically unnecessary. However, this method fails to produce reliable information of the height of the ball if only one row of sensors is used, especially when the ball comes close to the edge of the table. For reliable results, two sets of mirrors, one lying horizontally and one vertically, must be used. However, it may not be easy to synchronize the operation of these two sets of mirrors.

Since it is not permissible according to the competition rules to use active means to sense the ball, projecting infra-red light beam to scan the ball *is not applicable.* Ultrasound could be used. However ultrasonic transmission is only allowed when the ball is approaching the bat and it must cease on contact. It is practically not easy to construct an ultrasonic system that can automatically switch on and switch off at the right time. Furthermore, all vision systems, no matter active or passive, must stand behind the table and in no way can they be placed at other positions between the opposite edges of the table.

The use of video cameras, *in particular* CCD cameras, to grasp the image of the ball is the most popular method. In our project, we have used two low-resolution CCD cameras (OSCAR OS-15) to track the ball's trajectory. These miniature pin-hole cameras possess a 4.4mm x 3.3mm image area with 408 (Hori.) x 296 (Vert.) picture elements. An advantageous feature of these cameras is that they employ ***non-interlaced*** scanning technique with a horizontal frequency of 15.625 KHz and a vertical frequency of 50 Hz. As a result, there is no need to combine the odd and even frames. Another advantage of using CCD cameras over other alternatives (e.g., vidicon, MOS, CPD and CID) is that CCD cameras operate as pipelined devices, integrating one image while reading out the previous one. Thus there is no coupling from one image to the next, and therefore no time-varying characteristics. Each pixel is integrated over the same time interval, so there is no image distortion.

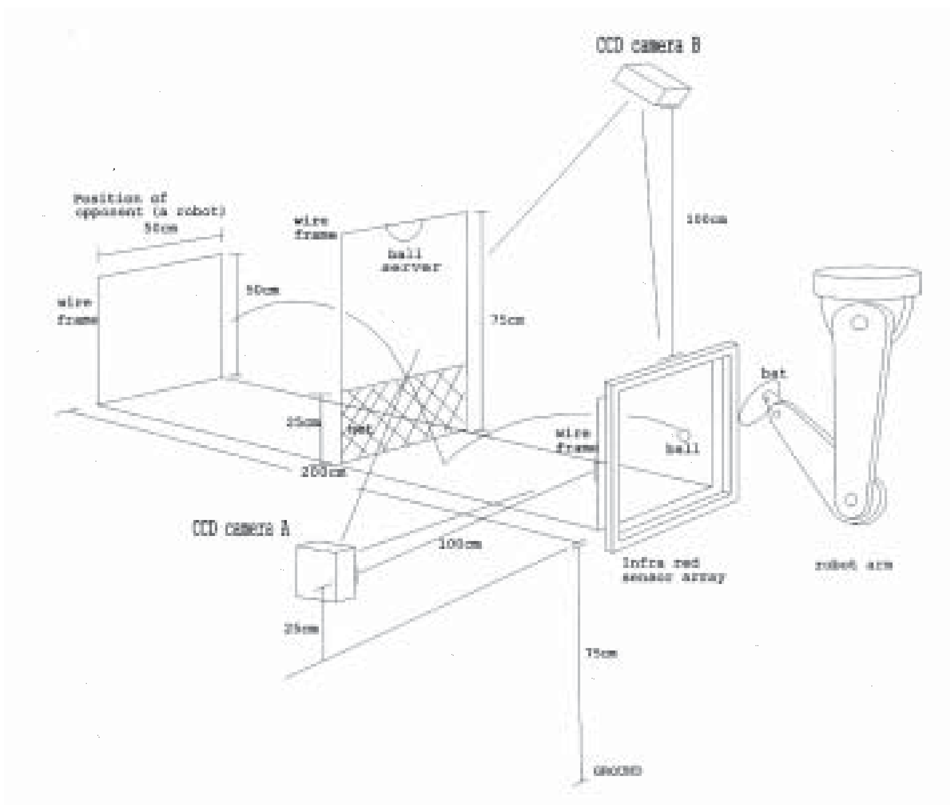

Figure 1 Setup of Fuzzy-Controlled Robot Ping-Pong Player *CITYMAN*.

3. PHYSICAL SETUP OF OUR CAMERAS

One of our cameras is set up at 1.5 metres above the table's edge while the other is placed on the left hand side at a distance of 1.0 metres from the table as shown in Figure 1. These two cameras lie on the same plane of the 50cmx50cm wire frame mounted at the near edge of the table. The two cameras are inclined at an angle of approximately 30° to the wire frame so that the near half of the table is clearly seen in the field of view of the cameras. Black curtains are set up on three adjacent sides around the table, the cameras and the mechanical arm. Since the ball appears as a white circular spot of finite radius in the picture, the c.g. of this spot must be found in order to locate the coordinates of the ball accurately. Camera A on the left observes the height of the ball and also its distance from the table's edge whereas camera B tells the left-right position of the ball and its distance from the table's edge. The combined information from these cameras is sufficient to uniquely define the position of the ball in the 3-dimensional space.

4. FINDING THE CENTRE OF THE PING-PONG IN REAL TIME

The cameras work in accordance with the CCIR Standard. The horizontal scanning frequency of the camera is 15.625 KHz which implies that each horizontal scan line will last for 1/15.625KHz = 64ms. For typical black and white cameras, about 16.5 to 18 percent of the scan time will be devoted to horizontal flyback. Thus the effective scan time is only 64ms x (1-16.5%) = 53.44ms. If we take 256 samples of the image along each scan line, the time allowed for each sample will only be 53.44ms/256 = 209 ns that is too short for any real time processing of the sampled pixel to be undertaken with ordinary microprocessors, even if we employ damn-fast digital signal processing chip like TI's TMS320C25 to process the image. We have therefore designed a *centre-of-gravity finder* (c.g. finder) with the use of binary logic to locate the centre-of-gravity of the ball in real time without any involvement of microprocessors.

Figure 2 is a simplified block diagram showing the main parts of this c.g. finder. This c.g. finder consists of two sets of logic circuits, the operation of which was controlled by clock pulses from a timing circuit. The on-off sequence of the clock

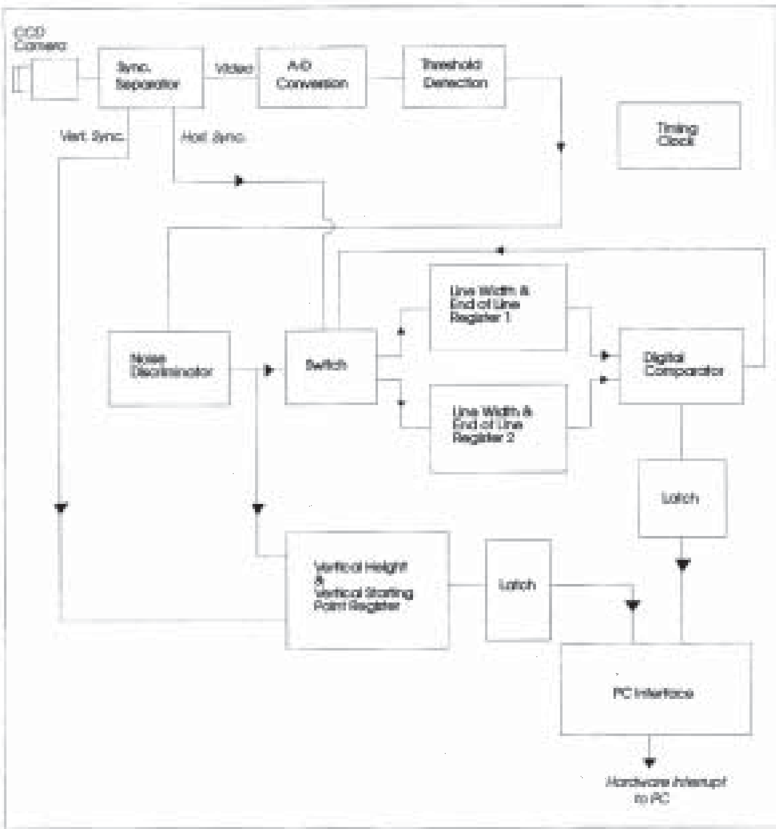

Figure 2. A Functional Block Diagram Showing the Essential Parts of the Centre-of-Gravity Finder

pulses is triggered by the vertical and horizontal synchronization of the video signals from the cameras. Since the ball is white and the background is dark (a matter of ping-pong competition rule), the flying ball appears as a single white circular disc in the picture. Two hundred and fifty-six samples are taken along each scan line of the picture. A total of 256x256=65,536 samples (pixels) would be taken for one frame since one frame is made up of 256 lines. After A-D conversion (by Datel's ADC-207 Flash AD-Converter), the intensity of the image is first hardware threshold-detected. Pixels with intensity less than certain preset value would be ignored. When the cameras scan line-by-line across the field of view, one set of logic will monitor the *ending position* and the *maximum width* of the ball's image in the horizontal direction. At the same time, another set of logic will continuously check the *maximum height* and the *starting position* of the ball's image. Once the cameras have finished scanning one complete frame, the width and height of the ball together with its far top and far right positions are latched in registers. These data are sent to the 80386 microcomputer via hardware interrupts. While the computer is reading the data, the c.g. finder could continue to monitor the next frame of the picture concurrently. As a result, this c.g. finder is able to produce a maximum of 50 sets of data about the c.g. of the ball in one second. This is the maximum rate of sampling available from the given video cameras. The horizontal position of the centre of gravity of the ball is found by first dividing the horizontal width by two and subtracting the resulting quotient from the ending position. Similarly, to find the vertical position of the ball's c.g., we simply divide the vertical height of the ball by two and add the result to the starting position.

5. CAMERA CALIBRATION & NOISE HANDLING

To prevent any unwanted non-black object enter into the field of view of the cameras and affect the function of the c.g. finder, we have to calibrate this vision system before use. Calibration is done by setting up the cameras in designated positions and angles. A ping-pong is then placed at the centre of the bench. When the calibration software program is run on the PC, the ball would show up on the screen monitor. At the same time, a small circle would overlap the image of the ball on the screen. The centre of this circle indicates the coordinates of the ball's centre of gravity found by the c.g. finder. If other unwanted obstacles appear on the screen or the ball's c.g. is incorrectly found, the orientation of the cameras and the brightness of the image may be adjusted until satisfactory results are obtained.

Special circuitry has been designed to reject the random noise spikes that may exist in the input signal. The method is to keep a historic record of the number of 'white' pixels continuously sampled by the A-D converter in the vertical and horizontal directions. Since noise spikes usually appear as scattered white dots on the screen whereas the ball's image consists of tenths of white pixels lying continuously in the horizontal and vertical directions, a few white pixels grouped together in one line is very likely due to the noise effect and these data must be ignored. The hardware circuitry is capable of monitoring the validity of the data, rejecting any faulty noise signals and resetting the logic circuit for new incoming data before the camera completes its scanning of the whole frame of the picture.

6. PREDICTING THE TRAJECTORY OF THE BALL

The two cameras are each observing the ball at an angle. The coordinates of the ball generated by the c.g. finders have to undergo coordinate transformation before the actual (x, y, z) coordinates of the ball with reference to a convenient Cartesian frame of the table can be obtained. At least three successive sets of the ball's coordinates with reference to a rectilinear frame must be obtained before we may apply the physical laws of dynamics to predict the subsequent motion of the ball. Both the work of coordinate transformation and solving the flight equation of the ball are time-consuming. In order to design a ping-pong player that is simple in structure, yet effective in operation, we have decided to make use of the information from the vision system as fuzzy input to our inference engine. Fuzzy variables with appropriate membership functions are defined in terms of the vertical and horizontal directions across the field of view of the two cameras. The fuzzy decision system is then 'taught' (by setting up decision rules according to human experience in playing ping-pong) to make proper decision on the position to place the bat, the angle to hit the ball and the strength required. The output from this decision system is finally used to control the motion of the mechanical arm to an optimal position to hit the ball when it comes to the edge of the table.

7. TEST RESULTS

The hardware and the software of the vision system have been completed. Thorough testing has been carried out. The trajectory of the ball plotted on the display of the personal computer is found to correspond closely to the actual flight path of the ball. In most cases, the error is less than 10%. This discrepancy is mainly due to the non-uniform lighting of the ball. The lamps are mounted on top of the table and the table itself is black, so there is a marked difference in the intensity of the light reflected respectively from the top and the bottom of the ball. The ball usually appears as a semicircle rather than a circular spot on the screen. At this moment, the c.g. finder is still unable to take care of this problem of non-uniform intensity. The present method to tackle this problem is to shift the c.g. slightly downward by means of software after the data have been read into the computer.

Figure 3 illustrates a calibration screen in which the image of the ping-pong has just been grasped. The centre of the little circle marked on the surface of the image is the location of the ball's c.g. determined by the c.g. finder. Fairly good agreement between the real c.g. position and the value of the c.g. given by the c.g. finder is clearly seen in this graphic capture from the computer. Figure 4 is another picture showing the trajectory of the ball following a curve path in air under the action of gravity. The dots indicate the centre-of-gravity of the ball at successive instants separated by 0.02s.

8. CONCLUSIONS

We have found this c.g. finder to be an effective and economical means of acquiring the c.g. of the ball in real time for playing the robot pingpong game. The whole vision system including the two cameras costs approximately HK$6000. Further improvement work that can be made on this vision system in the future includes refinement of the hardware/software to take care of non-uniform illumination, addition of hardware to monitor the speed of the ball and generalization of the above technique to finding the c.g.s of non-spherical objects.

Reference

[1] R. L. Andersson, A Robot Ping-Pong Player: Experiment in Real-Time Intelligent Control, Cambridge, MA: MIT Press, 1988.

[2] R. L. Andersson, "Dynamic Sensing in a Ping-Pong Playing Robot, " IEEE Trans. on Robotics and Automation, Vol. 5, No. 6, Dec., 1989.

[3] K. Hirota, Y. Arai & S. Hachisu, "Fuzzy Controlled Robot Arm Playing Two-dimensional Ping-Pong Game", Fuzzy Sets and Systems, Vol. 32, p.149-159, North-Holland, 1989.

[4] J. Knight and D. Lowery, "Pingpong-Playing Robot Controlled by a Microcomputer," Microprocessors and Microsystems, Vol. 10 No. 6, July/August, Butterworth & Co. (Publishers) Ltd., 1986.