Islamic Azad University

Science and Research Branch

Faculty of Engineering

Final Bachelor of Science

Sports Engeering

Title

3D Ball Detection and Trajectory Tracking, Case Study Volleyball

Supervisor

Dr. Reza Sarang

Student

Mohammad Almasi

February 2016

Abstract

In this thesis, we first introduce and then implement the new Striver vision model in volleyball to help the refereeing system. The stereo vision model is used to achieve the three-dimensional position of objects, this model consists of an object and two cameras with parallel or crossed optical axis or other arrangements. Camera calibration is one of the basic processes required in 3D machine vision in order to extract data related to the dimensions and 3D position of objects from 2D images. The method used in this research for calibration does not fall completely into the two categories of classical and self-calibration and is more flexible than them, and the main advantage of this method is its easy setup so that any person can create a calibration page. can use it. With stereo vision, we will reach a reasonably accurate distance, whose high computational volume makes it difficult to make it real-time, but with algorithm optimization methods based on reducing the search space and improving hardware, we can overcome this problem to some extent.

Keywords: stereo vision, calibration, volleyball, stereo matching, video check, refereeing.

Table of Contents

1-1 History of Volleyball .... **Error! Bookmark not defined.**

1-2 Referee system 6

1-3 Video Check system 7

1-4 Hawkeye System 8

Stereo vision system........ 11

Part one: Introduction to Stereo vision........ 11

2-1 Reasons and Importance of Stereo System........ 11

2-2 Different depth detction methods 11

2-3 Introduction to Stereo vision........ 13

2-4 Stereo vision algorithms........ 14

2-5 Terminolgy in Stereo vision........ 14

2-6 Introduction to Stereo vision........ 15

2-7 Review on Stereo vision........ 15

2-8 Area-based and feature based stereo vision 16

2-9 Dynamic stereo vision........ 17

2-10 Active stereo vision........ 18

2-11 Dense depth map 18

2-12 Camera-based and scene-based restrictions........ 19

2-12-1 Image geometry limitations 19

2-13 Camera calibration 20

2-14 Method for CaliBration........ 23

2-15 Streo Calibration 26

2-16 Image alignment........ 27

2-17 Calculation of distortion........ 27

2-18 Calculation the three-dimensional coordinates of a feature of the object........ 30

2-19 Calculation the Rotation and transfer of a three-dimensional object relative to a reference position .... **Error! Bookmark not defined.**

Camera 32

3-1 Callibration and 3D reconstruction 34

3-2 Object Detection 36

3-2-1 Appearance-Based method in object detection 37

3-3 Hough transfer..38
3-3-1 Parametric View..38
3-3-2 Warehouse..39
3-3-3 Hough Transfer Algorithm..39
3-3-4 Implementation 40
3-4 Correspondence of pixels in planes parallel to the image plane of cameras................**Error! Bookmark not defined.**
3-5 Blocking pixels in stereo correspondence..42
Impelementation.. **Error! Bookmark not defined.**
4-1 Calibration...43
4-2 Object Detection 45
Results.. **Error! Bookmark not defined.**
5-1Introduction...47
5-2 The equation of intersection of point and line in three-dimensional space 47
5-3 Routing ball coordinates in MATLAB 48
4-5 View the result using MATLAB...49
4-5 Forther Suggestions...51
5-5-Conclusion 51
References:... **Error! Bookmark not defined.**
Appendix:... **Error! Bookmark not defined.**

History of Volleyball

volleyball refereeing in 1895, by William. J. Morgan, director of the YMC It was invented in Holyoke, Massachusetts, USA. At first, this sport was called Mintonet. Under the influence of the great popularity of basketball among the public, Morgan decided to create a sport for his students that is played on the net and is enjoyable. Morgan started the game by using a tennis net on the grass, which was tied to both sides of the wall of the gym, and by using a basketball ball, which does not hurt the hands due to its light weight. Although this sport started slowly and slowly from YMC A, but it didn't take long for it to become popular in all the cities of Massachusetts and New England. In Springfield, after observing the game, Dr. TA Halstead changed the name of Mintonet to volleyball, because the main intention of playing is to send and return the ball from the net, which defines the meaning of the word volleyball. In the beginning, there were no specific rules for the game of volleyball, just as every innovation or invention is accompanied by defects in the beginning and is completed and corrected over time, volleyball was not an exception to this rule and rules were established for this game little by little. And technical methods and movements replaced the previous movements. The first foreign country to accept volleyball was Canada in 1900. The Soviet country established the National Volleyball Association in 1923 and did a lot of work for its development and transformation. The first international competition on the European continent between France and Czechoslovakia was held in Paris. In 1947, the International Volleyball Federation was established with the presence of fourteen countries from all over the world. In 1949, the first world volleyball tournament for men was held in Prague and in 1952, the first world volleyball tournament for women was held in Moscow. The volleyball in Iran was brought to Iran around the year 1299 by Mr. Mir Mahdi Varzandeh, a great master of sports, and it started to be taught at Darul-Mulemmin Sports. At the beginning of the establishment of sports federations in 1324, volleyball and basketball had joint federations, but at the end of 1336, volleyball became an independent federation and Mr. Amir Abbas Amin was appointed as the first president of the volleyball federation.

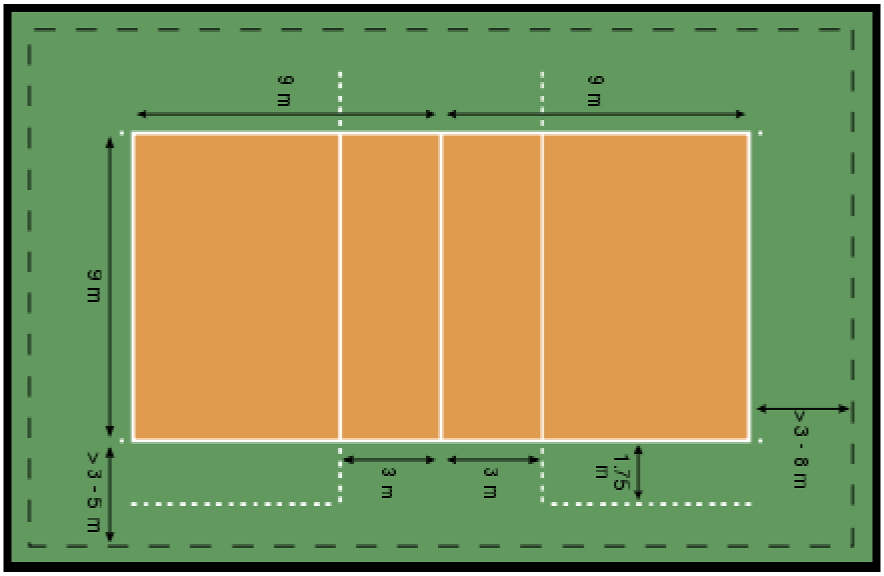


Fig 1. Characteristics of the volleyball court

1-2 Referee in Volleyball

Every volleyball match has two main referees, one is the first referee who is on top of the net and the other is the second referee who stands in front of the first referee and is allowed to whistle in certain cases. Four line referees stand around the volleyball court and help the first referee with the flag. Two referees are behind the secretary's desk, one is the secretary referee and the other is the assistant secretary, who is in charge of recording points, substitutions and warnings. The referee must be present in all meetings as an observer. This nine-person team is responsible for monitoring all areas of the land according to the following principles:

Volleyball Court Areas

Zone 2: Under the net on the right side of the field.

Zone 1: behind zone 2, the back of the field.

Zone 3: The part under the net in the middle of the field.

Area 6: Behind area 3 of the ground.

Zone 4: The part under the net on the left side.

Zone 5: behind zone 4 in the back of the field.

1-3 Digital Volleyball Referees (video checks)

In volleyball matches, the speed of the ball sometimes reaches more than 100 km/h, and for this reason, the possibility of human error of referees increases. Errors in the refereeing of volleyball games, video checks opened to volleyball. Video checkers are installed around the volleyball court and take videos of the court at certain angles. The captain of each team can request the referee to review the scene only 5 seconds after the score is announced. Also, during the review of these images by the second referee, and before announcing the final decision, the images will be played live on the stadium screen for the audience.

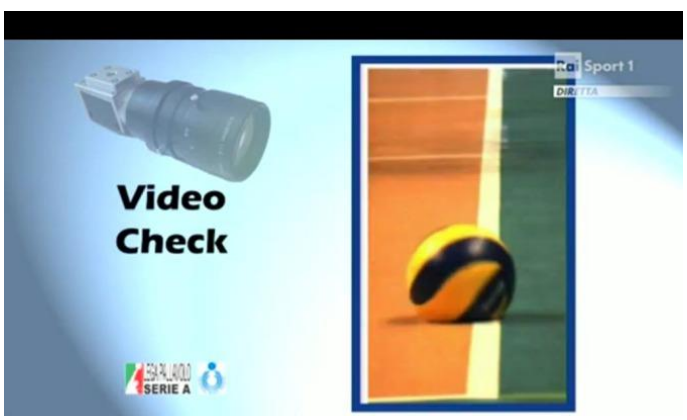


Fig 2. Videocheck one sample shot

The video check system used in the 12,000-seat Azadi hall has 26 cameras, six of which do not work due to inappropriate lenses, but this system is approved by the World Volleyball Federation. This system is one of the most up-to-date systems in the world and is used in many countries. Of course, there are other systems in the world that are used with more cameras, but they reduce the speed of the race in order to check many issues. The rules for using this technology are as follows:

1- The first referee of the match can ask for a video check as many times as he wants, if he doubts about giving points.

2- Each team can request a review of the scene only twice in each set, and if the result of the review is against the opinion of the referee and in favor of the requesting team, this number remains intact.

1-4 Hawkeye System

The hawkeye refereeing system was used for the first time in 2000, in a cricket match between England and Pakistan to track the ball in 3D. This system, in which 9 to 10 cameras are used to track the ball, nowadays also helps refereeing in sports such as tennis, volleyball, rugby and football. And it is also used in the production of sports advertisements [24]. The basis of this system is based on triangulation and the use of this method in ball detection. Triangulation has been described by astronomers as an ancient method of obtaining the coordinates of celestial bodies and estimating their distance from Earth.

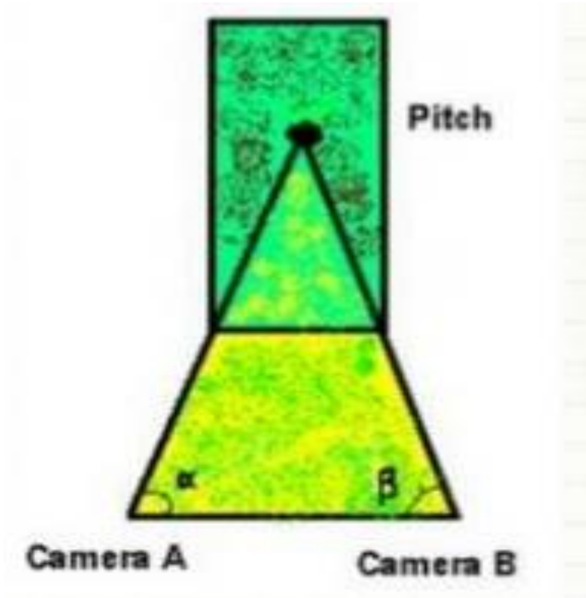


Fig 3. Triangulation

Triangulation Equations

$$\frac{d}{\tan\alpha} + \frac{d}{\tan\beta} = l$$

$$l \Big/ \frac{1}{\tan\alpha} + \frac{1}{\tan\beta} = d$$

where l is the distance between two cameras and d is the distance of two dorians to the pitch point.

Kalman Filter

Also referred to as a second-order linear filter, it is an algorithm that estimates the state of a dynamic system using a set of error-containing measurements over time. This filter usually provides a more accurate estimate than the estimate based on a single measurement based on Bayesian inference and estimating the joint probability distribution of a random variable at a point in time. This filter is named after Rudolph E. Kalman, one of the founders of this theory. Kalman filter has many applications in science and technology such as tracking and tracking vehicles, especially airplanes and spacecrafts. Hawk eye technology consists of two basic parts, which are:

A) imaging system:

In this part, with the help of speed gun, the speed of the ball and then the rate of rotation of the ball are obtained, so that, for example, in the sport of cricket, the reaction time of the opposing player can be measured. . The amount of swing and rising again is also one of the actions of this stage for different movements in different balls.

b) Movie playback system

In the video replay section, players and coaches are trying to learn about their performance with the help of replays.

Disadvantages of Cheshm Shahin technology system, this system uses 10 cameras to follow the ball in different directions in different parts of the field, for one monitor and then display the routing on the other monitor, it costs approximately 400,000 dollars (USD). And his income from each game is 10 thousand dollars [25].

Stereo Vision System

Introduction to Stereo Vision

2-1 Reasons and Importance of Stereo System

There are many ways to measure distance. In this century, due to the great scientific advances that have occurred, especially in the field of electronics and computers and new computing solutions, the method of measuring distances and distance finding has moved to the other side in such a way that having an accurate distance from the target and of course extracting this distance or 3D coordinates in real time is very important. For this purpose, 3D trackers are used to simulate visual information with high quality. Here we examine the new model of stereo vision to extract this information from the image [2,1].

The stereo model is a newly emerging model whose theory was proposed about a decade ago, and its use has been in progress for a short time. This system is inspired by the vision system and the dominant eye phenomenon in humans [3]. It is clear that in visual tasks, we take advantage of retinal difference and eye convergence to detect distance and depth, but the vision threshold is lower than when we follow with two eyes. One of the advantages of using two eyes is that looking with two eyes doubles the chance of making a difference. Binocular use appears to be an advantage for low-contrast stimuli and is used only in simple tasks. But the results show that people can benefit from the ability to see complex stimuli by using two eyes [11].

2-2 Different Depth Detction Methods

Humanity is currently entering a new world: the virtual world. In the virtual world, there should be a tool for locating body parts and objects. For example, in a military system, having an accurate distance from the target and extracting this distance in real time is very important. Measuring distance and finding three-dimensional coordinates is one of the big problems that many methods have been proposed to solve. Accurate measurement of the distance from an object or the distance of two objects from each other is of special importance in various industries such as: precision instruments, parts making, aviation, seafaring, robotics, and especially in military industries. Each of these solutions has certain advantages and disadvantages [2,1].

Here, we will examine the advantages and disadvantages and limitations of various methods for finding distance and compare them with each other. The use of infrared is one of the most conventional methods, the price of which is low, but its applications are limited to certain applications due to its small range, and the environmental effects are also the cause and add to the limitations of its use. Another method that has very wide applications is the use of ultrasonic sensors, whose low angular resolution is one of the problems of its use.

As you know, in 3D images, distance is also defined in 3D and is different from 2D distance. For example, if two-dimensional distances in two parts of an object are equal to the origin, these distances will not be equal in three-dimensional space. The ultrasonic system has a low price, but it cannot extract the three-dimensional distance and position for each image component, but an average distance is considered for the surfaces, and the low angular resolution limits the use of this method to applications that require less accuracy. limits.

Using a laser can be a very suitable option considering the progress made in this field and considering the divergence of normal light. Of course, the speed and accuracy of this method is very high, but the high cost and high power consumption limit its use to applications that require high accuracy. Also, not being suitable for transparent objects (such as glass) or shiny objects (such as shiny metals and surfaces that strongly deflect light) is the main disadvantage of this method [2].

In military systems, radar waves and millimeter wave range finders are also used, which have very high accuracy, but the environmental effects have adverse effects on it, which requires special systems to remove these effects and strengthen the signal compared to noise. It increases the use of them. Of course, it should be noted that all the systems introduced above are active and consume power.

Therefore, this research introduces and examines a new model for extracting distance from the image. One of the methods of 3D distance finding is to have a stereo system. With stereo vision or dual vision, which we will examine the structure of, we will reach a reasonably accurate distance, the high computational volume of which makes it difficult to make it in real time. But with hardware and software optimization methods, this problem can be overcome to some extent. In any case, the most appropriate system for each application should be selected by considering all the conditions of use [3]. The cost of this method can be low or high depending on the type of hardware used in it.

2-3 Introduction to Stereo Vision

Stereo vision model is a model consisting of one object and two cameras with parallel or crossed optical axis or other arrangements. A camera can also be used, in which case the camera or the object will move perpendicular to the optical axis of the camera so that the object is placed in the field of view of the camera. In stereo vision, different geometries of the camera-object are presented, cameras with crossed axes, moving the camera along the optical axis, rotating the camera around the optical axis, etc. The commonly used mode is the geometric model of an object and two cameras with parallel optical axes [11]. Due to the fact that at any time we need a separate calibration for the rotation of the cameras relative to each other, and this will increase the calculations and as a result decrease the speed, so we consider the cameras to be fixed relative to each other so that there is no movement over time. do not relate to each other. The cameras should be placed on the structure in such a way that they record the aligned images as much as possible (this work is necessary to reduce the calculations required for alignment and increase the speed of calculations) [1].

Camera calibration is one of the basic processes required in 3D machine vision in order to extract data related to the dimensions and 3D position of objects from 2D images. The geometry seen by the camera is related to a mathematical model depicted from it. In camera calibration, the goal is to solve this model. Camera calibration means calculating the internal and external parameters of the camera. The methods that have been introduced for calibration are selected and designed according to the type of system needs, application, imaging environment, required accuracy, speed and available facilities.

Based on factors such as the required accuracy, the amount of operator intervention and the metricness of the 3D reconstruction, there are various methods for camera calibration. Classical calibration methods mainly use a calibration object with specific physical dimensions (reference cube) which is placed in front of the camera. These methods enable us to fully calibrate the camera and perform a 3D reconstruction that matches the physical dimensions. Some of the new calibration methods use the features in the physical scene in front of the camera to calibrate it. The main advantage of these methods is that they do not require a calibration object and operator intervention, but their accuracy depends on the accuracy of extracting the features in the observed scene, which is not always possible to do with high accuracy. Therefore, the existing methods for camera calibration can be divided into two general categories, one group of methods that require a calibration object and the other are called classical methods. Another group of methods that do not need a calibration object and are called self-calibration.

The method used in this research for calibration does not fall completely into the above-mentioned two categories and is more flexible than them. The flexible calibration method is taken from reference [12]. Its main advantage is the easy setup method so that anyone can use it by creating a calibration page. In addition, it has high accuracy.

2-4 Stereo Vision Algorithms

The use of stereo vision is a new method based on images taken from a scene that can be used to find depth. Stereo vision helps to identify and locate the object in three dimensions.

Stereo vision algorithms, including pixel-based, feature-based and surface-based, all calculate the spatial difference. Pixel-based algorithms calculate the spatial difference for each image pixel and therefore greatly reduce the execution speed. In surface-based algorithms, special points are extracted in one image, for example the left image, and their matches are searched in another image based on a similarity criterion. In feature-based algorithms, feature points are extracted in both images and the similarity between them is measured according to the procedure. Because the number of matched candidates in this method is less than the previous two methods, this algorithm is the fastest [9,1].

In the feature-based algorithm, stereo matching includes two consecutive parts of detecting special points and establishing matching. In the special points detection stage, special points with specific properties, such as edges, corners, centers of gravity and texture areas, are extracted from the left and right images. Considering that high frequency points are more important and valuable, edges are usually used. Most of the existing stereo algorithms work on gray level images, but in this research algorithm, color images are used in order to reduce the error [10].

2-5 Terminolgy in Stereo Vision

Camera Calibration

Camera calibration consists of determining the internal optical and geometric characteristics of the camera (internal parameters of the camera) as well as the position and orientation of the camera relative to a global coordinate system (external parameters of the camera). Camera calibration is one of the basic processes required in 3D machine vision in order to extract data related to the dimensions and 3D position of objects from 2D images. The geometry seen by the camera is related to a mathematical model imaged from it. In camera calibration, the goal is to solve this model. The usual geometric model that is considered for the camera is the pinhole model, which is imaged by converting a point of three-dimensional space to the image plane of the camera [5].

Correspondence Vision

Visual matching means finding the corresponding pixels in two stereo images that target a common scene, in this matching we find pixels that do not have any corresponding pixels in the other image, that is, they have a position in space that is not in the field of view of the other camera, or The fact that there is an obstacle in the way of another camera's vision, these pixels are called blocked pixels [8].

Stereo Vision System

2-6 Introduction to Stereo System

Stereo is a well-known method for obtaining depth information from digital images. Currently, almost all robotic and intelligent systems are equipped with two cameras side by side to perceive depth. While humans generally have the ability to perceive depth by default, understanding depth is difficult for computers, and research is ongoing on this issue. Compared to stereo vision, there are different technologies for measuring distance and obtaining three-dimensional information, for example, laser scanners can extract accurate and high-density three-dimensional maps from large parts, but in most of these high-quality scanners, in At one moment, only one point or one line can be measured and they are enclosed in static and motionless environments. The usual method in photogrammetry is to use stereo techniques for three-dimensional measurement. In fact, the effort in the field of stereo vision system is simulating the human vision model, but with higher accuracy.

2-7 Review on Stereo Vision

The stereo vision model is a model consisting of an object and two cameras with parallel or crossed optical axis or other arrangements. A camera can also be used, in which case the camera or the object will move perpendicular to the optical axis of the camera so that the object is placed in the field of view of the camera. In stereo vision, different geometries of the camera-object are presented, cameras with crossed axes, moving the camera along the optical axis, rotating the camera around the optical axis, etc. The commonly used mode is the geometric model of an object and two cameras with parallel optical axes, two images are taken from the same scene, and the corresponding pixels are identified in the two images, which is done by image processing algorithms.

The main problem in stereo is how to find the corresponding points in the left image and the right image, which is known as the matching problem. Whenever a pair of corresponding points is found, the depth can be calculated by the triangulation method. Worldwide, many research activities dealing with stereo vision are known. In their solution, various methods have been used to solve the problem of matching and choosing the restrictions applied to the visibility of objects in the scene. In addition, various methods have been used in different tasks (for example, mobile robots, photogrammetry, stereo microscopy, etc.) and a large number of recognizable features have been used in the solutions. This allows direct comparison of methods.

In early 1989, Donde and Agarwal presented a journal on the extended stereo vision method, which was under development until late 1989 [13]. Their paper covered a class of algorithms that identify images based on differences in image geometry, matching geometric shapes, and using computational structure. In addition, the performance of these stereo methods in different classes was checked on test images. This comprehensive journal is highly recommended to all researchers dealing with stereo. However, many new stereo methods have been released over the past few years.

Stereo methods can be distinguished by several characteristics, for example, if they use region-based or feature-based methods, if they use static or dynamic scenes, if they use passive methods and or actively use and if they produce a sparse or dense depth map. All stereo methods use constraints to reduce errors in the matching process.

2-8 Area-based and Feature-Based Stereo Vision

Area-based stereo methods find the corresponding points based on the correlation (similarity) between the corresponding areas in the left and right images. First, a desired point in an image is selected. Correlation is measured and then used to search for a corresponding point in a corresponding region in another image. One of the disadvantages of area-based methods is being sensitive to photometric fluctuations during the image acquisition process and being sensitive to distortion that is the result of changing the observation position. This sensitivity arises from the direct ratio of the intensity values in the images.

Feature-based stereo methods, on the other hand, match features in the left images with features in the right images. Features are selected as the most salient parts in the image, for example, edge points or edge segments. The advantages of feature-based methods are less sensitivity to photometric fluctuations and being faster than region-based stereo, because there are fewer candidates to match corresponding points. In addition, features may be extracted from sub-pixel precision to increase the precision and accuracy of matching results.

Whenever a feature-based method is used, features must be extracted. Most methods still use edges as features, but there are exceptions that use regions or topological structures. The methods used for edge detection have not changed since 1989. Although stereo feature-based techniques solve the matching problem quickly and accurately, the number of matching structures is small due to the small number of features. Correlation methods seemed to fall out of fashion in the late eighties due to their disadvantages mentioned above, but in recent years, they have undergone a revival due to the aforementioned disadvantages of feature-based methods [1, 16].

2-9 Dynamic Stereo Vision

At first, images are captured by fixed cameras in a fixed environment. The main goal of all proposed stereo methods was and still is to automatically search for corresponding points in two images. Recently, a new trend has emerged in stereo research that uses motion to obtain more reliable depth estimates. This line of research, called dynamic stereo, is mainly pursued by members of the robot vision community using a moving mobile robot. Different methods have been published assuming a moving object, several moving objects or a fixed scene and a moving camera. Many stereo researchers exclude motion information from their solutions because motion does not significantly improve the accuracy of the results and is not always available (for example, in photogrammetry or stereo microscopy) [14].

2-10 Active Stereo Vision

In terms of computer vision, the localization range is usually classified into active or passive methods. Active methods transmit energy to the scene to measure the range (for example, ultrasound, radar, modulated light, etc.) while passive methods do not. They observe a point on the surface using a laser beam with the highest brightness. The range is easily obtained using the triangulation method using the bright points in both stereo images from the matching axiomatic analysis solutions.

2-11 Dense Depth Map

Still, feature search is an important criterion to reduce the computation time of matching stereo images. Unfortunately, feature-based stereo or edge-based stereo, respectively, only produce sparse disparity maps, ie, only sparse control points can calculate a successful surface interpolation process for a successful reconstruction of complex surfaces. However, it is necessary to calculate the dense disparity map defined for each pixel in the whole image [15,4].

Mass stereo matching is significantly affected by the assumption that the pixels of the corresponding images have similar intensities (photometric limitation). Several methods have been proposed to calculate the dense depth map. The defined energy function represents the constraints used for matching. A function to measure photometric similarity is defined, for example, as:

(2-1) $\mathrm{P_I}[\mathrm{D}(\mathrm{x}, y)] = |I_L(x, y) - I_R(x + D(x, y), y)|$

where $I_L(x, y)$And$I_R(x, y)$The intensity functions are in the left and right images. To ensure a smooth mismatch, the function may be defined as follows:

(2-2) $DS[D(x, y)] = \frac{1}{4}\sum_{(u,v)P\ N_4}(D(x, y) - D(u, v))^2$

where N_4 In contact with four neighbors(x, y)Is.The energy function representing both constraints with parameters$\propto$is given below:

(2-3) $E = \sum_x \sum_y \{P_I[D(x, y)] + \propto DS[D(x, y)]\}$

Finding a disparity map D(x, y) that minimizes equation (2-3) is a solution to the stereo matching problem. Several methods are presented to minimize the energy function. For example, let's apply another type of simulated annealing method to calculate the dense depth map, which uses the photometric limit and the smooth mismatch limit in the energy function [18]. Anneal simulation algorithm is a new method to reach absolute optimal or near optimal solution in multivariable functions with discrete design parameters.

2-12- Camera-Based and Scene-Based Restrictions

In general, there is no solution to the matching problem given the existence of fuzzy matching candidates. As a result, for each stereo method, several assumptions about image geometry and/or objects in the scene are used to reduce the number of ambiguities. These constraints can be specified as camera-based and scene-based. Camera-based constraints are related to image geometry, for example camera position, orientation, and the use of physical images. Scene-based constraints deal with assumptions about the nature of the allowed surfaces and their spatial relationship in the scene, many constraints are home-made and used only by the creator [17].

Although more defined than pre-1989 limitations, they are still fundamental to current research activities. However, most stereo articles only mention them, but do not explicitly define them. Therefore, examining the limitations of stereo can help to improve the readability of stereo articles.

2-12-1 Image Geometry Limitations

Epipolar limitation: One of the most powerful limitations in stereo vision is the knowledge of epipolar geometry. Figure 1-2 shows the imaging geometry of a pair of stereo images, OC_2 and OC_1 optical centers in both systems. They are cameras. The distance between these optical centers is called the baseline. In the scene, the point P(x, y, z) is obtained from the intersection of two points P_1 and P_2, respectively, in the corresponding image planes, from the base line and the radius of the spotlight that connects OC_1 to P, there is a plane that is called the epipolar plane, this epipolar plane It intersects both image planes in lines called epipolar correspondence lines. In this epipolar geometry is as follows, according to the point P_1 in the right image P_2, the matching point corresponding to it in the left image should be located in the corresponding epipolar line. Therefore, we reduce the search space for a corresponding matching point in the second image to a one-dimensional search space represented by the epipolar line. In addition, all epipolar planes bisect the image planes along horizontal lines if the camera is positioned parallel to the optical axis and the image planes are moved by a horizontal displacement. This arrangement of cameras is called standard geometry and is still used by many stereo methods due to its simplicity.

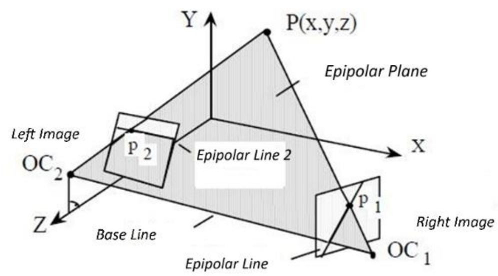


Fig 4. Epi geometry Thepolarand epi lines The polar [19]

The main problem with epipolar constraints is that the epipolar contours must be calculated using high-precision calibration techniques. Meanwhile, Hartley and Gupta have presented a technique to calculate the epipolar limit from relative orientations without precise camera calibration [20,19].

2-13 Camera Calibration

When two or more images are taken from a three-dimensional object from different perspectives, the three-dimensional position of the points on the object can be determined from the images taken. For each physical point of the surface of the object, if we can determine the pixel locations corresponding to it in all the images, the three-dimensional position of that point can be calculated. To perform such calculations, the calibration of each of the cameras and the finding of pixel locations corresponding to a 3D physical point in all images are required.

Many calibration methods have been introduced, each of which is selected according to the type of system needs, imaging environment, accuracy, speed and facilities. The classification of methods is usually based on the dimensions of the model used in the method. In this part, we will review a summary of the performance of the methods.

In the first category, the primary methods of camera calibration, which are the traditional or classic methods, use a three-dimensional object that has planes perpendicular to each other (such as a box) and has small squares like a checkerboard on its faces. Algorithms presented for this method are of two types (1): direct calculation of parameters, (2): use of projection matrix. One of the factors that determine the accuracy of calibration in this method is the dimensions of checkered windows located on the pattern pages. In addition to this, the method requires that the spatial position of the model is correctly determined. Therefore, the accuracy of metric measurement from outer space has a direct effect on the accuracy of calibration. In fact, during calibration, it is necessary to inform the system of at least 8 points in the image and the corresponding spatial coordinates of those 8 points in each stage with the help of human resources, which indicates the creation of a break. In the program and entering the role of the user during the execution of the algorithm. On the other hand, the computational cost and the complexity of its mathematical operations are very high.

The second category is the use of flat objects or two-dimensional objects as a calibration model. In this method, it is necessary to take pictures of the flat calibration pattern from different directions (at least three different directions) and then the calibration is done with the help of the pictures taken. The accuracy of this method is very good considering the optimization problem. In this method, there is no need to enter the information related to the spatial coordinates of the image points, but it is necessary to know the image coordinates of the feature points for each image, on the other hand, these feature points are different in different images and are never the same. Another problem with this method is the number of images that must be taken from a pattern in different directions, which itself causes the amount of information to be occupied and high processing.

The third category is the use of one-dimensional objects, which actually include thin and long objects such as television antennas, rulers, and thin sticks. In this method, it is necessary to determine three points on the pattern (beginning, middle and end). Calibration is done by rotating the model around a fixed point and taking at least 6 calibration images. Among the advantages of this method is easily finding the calibration pattern. The accuracy of this method is average, and to achieve more accuracy, the maximum likelihood algorithm should be used. But the computational volume is relatively high compared to the first category models.

The fourth category of calibration methods is self-calibration or calibration with the help of dimensionless objects, calibration is done only by matching the images taken from the surroundings. It is necessary to rotate the camera in different directions and since there is no model for calibration, many parameters must be extracted. Only taking three pictures is enough for calibration. Calibration with the help of matching turns into solving a complex mathematical problem with many parameters. In general, calibration itself is a difficult method to implement and its computational cost is high, and its only advantage is the calibration model that does not exist.

Among other calibration methods, the calibration is based on elimination points, this method requires a pattern that has three pairs of lines two by two parallel to each other. This method is more suitable for traffic cameras. It is necessary to know the information about the movement of the camera in this method [8].

Although the classical calibration method can achieve high-precision parameters, this method is only used for the laboratory or in situations where the calibration is very complicated, which usually results in more workload and problems in practical applications. Therefore, researchers are currently involved in modifying computer vision to focus on the calibration of self-calibrating cameras [12]. Due to the precision of classical methods, we examine them. Two categories of classical methods are briefly described below. Non-linear optimization: The transformation between global three-dimensional coordinates and two-dimensional image coordinates is a nonlinear transformation of calibration parameters. These methods, with a non-linear optimization, obtain an estimate of the calibration parameters, using the minimization of the residual error caused by re-imaging a set of known 3D coordinates with non-linear equations. The main benefit of this method is that no approximation is required and the camera model can be quite complex. The problems of this method are its high calculation volume and a good initial guess.

Calculating the perspective conversion matrix: Although the equations specifying the three-dimensional to two-dimensional conversion are non-linear equations of the camera parameters, these equations can be linearized by removing the deviation caused by the lens, and the coefficients of the perspective conversion matrix 4x3 as The unknowns of the equation were considered. These coefficients are a function of camera parameters. Having a set of global three-dimensional coordinates (obtained from the calibration object) and their corresponding two-dimensional image coordinates, a set of linear equations can be solved by the least squares method. Solving these equations enables us to obtain the perspective transformation matrix and thus the camera parameters. The advantage of this method is that no initial estimation is required. But it brings some problems. Among other things, the removal of lens distortion is unacceptable for 3D measurements, the resulting moment matrix is not orthogonal, and the calibration points cannot be located in the same scene, while if the calibration points can be located in the same scene, it becomes easy to make the calibration object and recognize the feature on it [5].

2-14 Method of Calibration

In the method chosen for calibration, the internal parameters that are calculated are:

Focal length: The values of the focal distance in terms of pixels in the x and y directions, which are shown in the equations of this section with the vector$fc_{2\times 1}$,The focal length f is the distance between the center of the image and the image plane.

Coordinates of the main point: In this section, it is represented by the vector$cc_{2\times 1}$ This point is where the image plane meets the line perpendicular to the image plane from the center of the image.

Error coefficients: These coefficients indicate several deviations that occur due to the non-linear effects of lenses in the image. These deviations are mainly radial and tangential. In this section, these coefficients are stored in a 1x4 vector called k_c. Radial deviations are visible at the outer edges of the image where straight lines appear curved. These effects are caused by the non-ideality of the lenses and as a result of different light refractions. Tangential deviations are also caused by non-ideal lenses.

We assume that it is a point in 3D space with a coordinate vector in the camera coordinate system.$Pxx_c = [x_c, y_c, z_c]$ The following relations show the perspective transformation of the point on the image plane using the internal parameters of the camera.P

x_n We define the normalized coordinates as (2-5).

$$(2\text{-}5) x_n = \begin{bmatrix} \frac{x_c}{z_c} \\ \frac{y_c}{z_c} \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}$$

If, after calculating the error coefficients,$r^2 = x^2 + y^2$ The normalized coordinates are shown as (2-6).

$$(2\text{-}6) x_d = \begin{bmatrix} x_d(1) \\ x_d(2) \end{bmatrix} = (1 + k_c(1) \times r^2 + k_c(2) \times r^4) \times x_n + dx$$

In above is the tangent error formula which is defined asdx

$$(2\text{-}7) dx = \begin{bmatrix} 2xyk_c(3) + k_c(4) \times (r^2 + 2x^2) \\ k_c(3) \times (r^2 + 2y^2) + 2xyk_c(4) \end{bmatrix}$$

Finally, the relationship between the pixel coordinates of the perspective image of point p on the image plane and the normalized coordinates ofx_d

$$(2\text{-}8) \begin{bmatrix} x_p \\ y_p \end{bmatrix} = KK \begin{bmatrix} x_d(1) \\ x_d(2) \\ 1 \end{bmatrix}$$

In the recent formula, KK is called the Doreen matrix and contains the following elements:

$$(2\text{-}9) KK = \begin{bmatrix} f_c(1) & 0 & cc(1) \\ 0 & f_c(2) & cc(2) \\ 0 & 0 & 1 \end{bmatrix}$$

After camera calibration, the KK matrix, which contains the camera's internal parameters, is calculated.

The external parameters of the camera calculated in the calibration are: the rotation matrix and the transfer matrix. These two matrices actually determine the relative position in the camera coordinate system and the global coordinate system. The following relationship converts the coordinates of point P in the global coordinate system to its coordinates in the camera coordinate system.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} R_{11} & R_{12} & R_{13} \\ R_{21} & R_{22} & R_{23} \\ R_{31} & R_{32} & R_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} T_x \\ T_y \\ T_z \end{bmatrix} \tag{2-10}$$

To calculate the internal and external parameters of the camera, we take an image from a checker board from different angles, in each image we calculate the pixel coordinates of the corners and define a global coordinate system similar to the figure below for each screen and the global coordinates of the corners in terms of square dimensions We calculate the checkerboard. The pixel coordinates of the corners of the checkerboard and their corresponding global coordinates for a number of images, for example 5 images, are given as input to the calibration algorithm. The calibration output is the internal parameters of the camera and the transfer matrix and the rotation of each position of the calibration plane relative to the camera. The minimum number of grid images to solve the calibration equations is three. Using the MATLAB environment: In the MATLAB software, there is a toolbox for detecting the internal and external matrices of the cameras in the calibration process. With the help of this toolbox, you can get all the mentioned values. The following points are the most important points of using this toolbox.

1) About 10 to 20 pieces of photos should be taken of the checkerboard moving frame. This number is because they can be used to obtain rotations and displacements on the page.

2) The size of each side of the small squares inside the checkerboard must be known, which is usually considered to be twenty-five centimeters, but it may be different.

3) It should be noted that the size of this screen is the same as the estimated size of the object. The large or small size of the checkerboard compared to the object may be associated with incorrect identification of the object.

4) It should be noted that calibration should be done from any distance where the object needs to be tracked. For example, if the object is located at a distance of 3 meters from the cameras, calibration should be done from a distance of less than 3 meters, so that the angularization of the Shatarji plane is simple. If it happens, this

problem can be considered one of the merits of this method because it does not require on-site calibration.

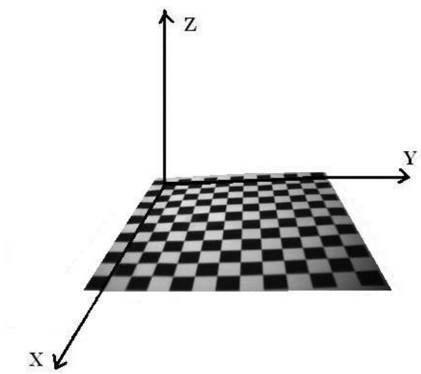


Fig 5. The global coordinate system defined in the corner of the page [5].

2-15 Stereo Calibration

Two cameras are used in stereo vision calculations. First, the cameras are calibrated with the calibration method mentioned in the previous section, then their relative position is determined with the following formulas.

$$(2\text{-}11)\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = R_{stereo} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} + T_{stereo}$$

The vector on the left shows the equality of the coordinates of a point in the right camera coordinate system and the vector on the right shows the coordinates of the same point in the left camera coordinate system. The purpose of stereo calibration is to determine the matrices and By calibrating the cameras and having the relative position of each camera in relation to a reference coordinate device, the two mentioned matrices can be determined.$R_{stereo}T_{stereo}$

$$(2\text{-}12)\begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = R_r \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T_r$$

$$(2\text{-}13)\begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = R_l \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + T_l$$

using the recent equations and are equal to: $R_{stereo} T_{stereo}$

(2-14) $R_{stereo} = R_r \times R_l^{-1}$

$T_{stereo} = T_r - R_r \times R_l^{-1} \times T_l$

Because after the calibration of each camera to the number of images of the calibration screen, the matrices R_r, R_l, T_r And T_l We have, by averaging, we determine and $R_{stereo} T_{stereo}$

2-16 Image Alignment

The images given by two stereo cameras are affected by the distortion of the lens and the probability that the coordinate axes of the cameras are parallel in the mechanical arrangement of the cameras is very low and it is not possible to find the coordinate axes due to this non-parallelism. Corresponding pixels in two images become difficult and the probability of error increases. In order to make it easier to find the corresponding pixels and for the corresponding pixels to be placed in the same horizontal direction (v coordinate of the image) in the two images, it is necessary that, firstly, lens distortion does not affect the captured images, secondly, the coordinate axes of the cameras are parallel to each other and the xy planes in two cameras should coincide and thirdly, the cameras should have the same intrinsic parameters, focal length and image center.

To do this, two right and left images must be corrected, in fact, two virtual images are produced, that is, the coordinate axes of the cameras are rotated virtually and by software in such a way that their axes are parallel to each other and the xy planes in the two cameras are on coincide, then the images should be lensed without any distortion and the same focal distance and center of the image are determined for these virtual images, in this case, finding the corresponding pixels will be easier and with less error [6].

2-17 Calculation of Distortion

Figure below Shows an example of a camera setup, with the camera tilted slightly inward. Suppose that the coordinates of two corresponding points and $(X_{left}, Y_{left})(X_{right}, Y_{right})$ is. ToTwo cameras that are properly aligned $Y_{right} = Y_{left}$.Inconsistency as $X_{left} - X_{right}$ is defined This value can be positive or negative depending on the camera angle and the distance to the object. The biggest problem in machine vision is searching for corresponding points and compressing images and videos. Calculating stereo misalignment is very similar to finding optimal motion vectors, and the solution methods are similar for both problems.

The most common method in both is stereo misalignment calculations and motion compensation where a block of one image slides over a captured second image. This approach is known as block matching algorithm. In each deviation, it is possible to calculate the square direction error. Finding the location of the most similar sub-image (where the least error occurs) is equivalent to calculating the disparity. Disparities usually have a small dynamic range (often smaller than 10 pixels) compared to the actual distances of objects. Therefore, measuring misalignment for an integral pixel value results in very low depth resolution.

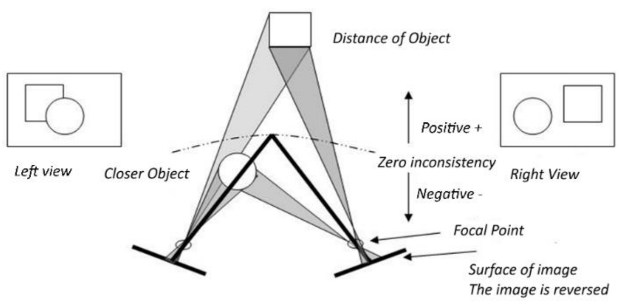


Fig 6. geometry Stereo view for camera settings [20]

Finding the corresponding points for each pixel in an image is computationally expensive. Consider a simple implementation: for each pixel in the left image, a block around that pixel (often 16x16 or 32x32) slides across a row of the right image (with the same height as the block in the left image, but with an equal width to the width of the entire image). At each position, the squared direction error (or other metric error) is calculated, which involves a large number of additions and coefficients.

Various optimizations have been proposed to reduce the amount of calculations. Instead of searching an entire row, a subset of it is selected, usually based on an estimate of the highest probability of mismatches seen in the data. The search range can also be adjusted dynamically by exploiting the fact that adjacent points are likely to have similar mismatches. Another class of optimization relies on the observation that the error function is perfectly smooth as a function of the horizontal deviation (from which the misalignment is determined), with a single and perturbed minimum.

Often a smooth error curve will find a minimum without an exhaustive search. For example, one can sample the error curve at a relatively small number of points and select the best point(s) for further refinement.

At ShapeUnderSampleTheOh youof a stereo system is shown and asin the pastalso it was saidToObtaining depth information or indeed didn't writeSpatial difference, need two images of the same scenewe have. In figure (2-7) two points andOO_rpositionTwo cameras StereoIn the scene, it shows that the role of these two cameras, likeTwo eyes are in the human visual system.SpaceTheYAmong thisTwo points, in the system HiDifferent stereo, different IsAnd it is known as the baseline.

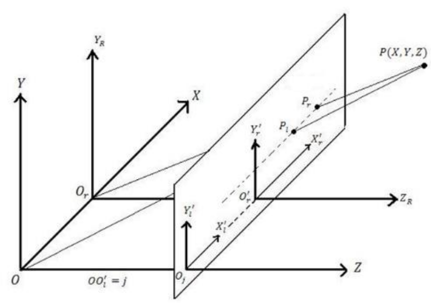


Fig 7. Basic stereo system geometry [1]

In the figure above, the point is in the three-dimensional scene and its projection on the left and right images formed in the cameras.$PP_l$for you.P_r

(2-15)$x_{pl} = \frac{f}{Z}X \quad x_{pr} = \frac{f}{Z}(X - b) \quad y_{pl} = y_{pr} = \frac{f}{Z}Y$

According to equation (2-15) the value$y$It is the same for both left and right images. Equation (2-16) can be used to obtain the value of spatial difference:

(2-16) $x_{pl} - x_{pr} = \frac{f}{Z}b$

The amount of spatial difference is called and denoted by The depth of each point can be obtained according to the value of its spatial difference according to the equation (2-17):$x_{pl} - x_{pr} d$

$$(2\text{-}17) d = x_{pl} - x_{pr} = \frac{f}{Z} b \rightarrow Z = \frac{f}{d} b$$

Therefore, the depth is proportional to the image of spatial difference. Meanwhile, in this equation f The focal length of the camera lens and the base line b [1].

2-18 Calculating the Three-Dimensional Coordinates of a Feature of the Object

If the pixel coordinates of a feature of the object in the image of the left camera are represented by and$x_{pl} y_{pl}$ And the coordinates of the same feature in the image of the right camerax_{pr} and show, according to relation (2-8), we will have:y_{pr}

$$(2\text{-}18) \begin{cases} x_{pl} = f_c(1) \times x_{dl}(1) + cc(1) \\ y_{pl} = f_c(2) \times x_{dl}(2) + cc(2) \end{cases}$$

$$(2\text{-}19) \begin{cases} x_{pr} = f_c(1) \times x_{dr}(1) + cc(1) \\ y_{pr} = f_c(2) \times x_{dr}(2) + cc(2) \end{cases}$$

Using the above relations, we can get the vectors and$x_{dl} x_{dr}$ calculate In these two values, there are error values that we remove these error values using numerical methods. After removing the error values, the 2x1 vectors,x_{nl} And it is achieved. According to the normalized values$x_{nr} x_{nl}$ and and using the values of and,$x_{nr} R_{stereo} T_{stereo}$ The pixel coordinates of the desired feature in the left camera coordinate device are as (20-20).

(20-20)

$$\begin{cases} M = R_{stereo}(0)x_{nl}(1)x_{nr}(1) + R_{stereo}(7)x_{nl}(2)x_{nr}(1) + R_{stereo}(8)x_{nr}(1) \\ \qquad -R_{stereo}(0)x_{nl}(1) - R_{stereo}(1)x_{nl}(2) - R_{stereo}(2) \\ N = T_{stereo}(0) - T_{stereo}(2)x_{nr}(1) \end{cases} =>$$

$$\begin{cases} x_L = x_{nl}(1)z_L \\ y_L = x_{nl}(2)z_L \\ z_L = M/N \end{cases}$$

Since we know the relative position of the left camera coordinate system and the reference coordinate system, we can obtain the coordinates of the desired feature in the reference coordinate system.

$$(2\text{-}21)\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = R_L^{-1}\left(\begin{bmatrix} x_L \\ y_L \\ z_L \end{bmatrix} - T_L\right)$$

2-18 Calculation the Three-Dimensional Coordinates of a Feature of the Object

Using the coordinates of three points of a three-dimensional object, we can define a coordinate system for it. If these three points,P_0 $P_1$and name$P_2 P_0$ We consider as the origin of the coordinates. Normal vectors $\overrightarrow{P_0P_1} = v_0$,$\overrightarrow{P_0P_1} \times \overrightarrow{P_0P_2} = v_1$ And$v_1 \times v_0 = v_2$ We define as the coordinate axes. A 4×4 matrix in the form (22-2) is defined for the position of the object.

$$(2\text{-}22)Loc = \begin{bmatrix} v_0 & v_1 & v_2 & P_0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

For the object in the reference position or zero measurement, we calculate the position matrix. When the object rotates or moves relative to this position, the amount of rotation and movement is calculated with equation (2-23).

$$(2\text{-}23)Trans_{4\times4} = Loc_{new} \times Loc_{zero}^{-1}$$

From the matrix, we can calculate 6 correction parameters. These parameters are: displacement along the axes and rotation around these axes. $Trans$X ،Y و Z

Camera

With Attention To Models geometric presentation done For Camera I see And Methods calibration That At last season To they Hint was, in the investigation of the cameras used in the beginning of the two model camerasPK-750MJFrom the brandA4TECHWas used,From benefits Camera item Use Can To cheapness, smallness size relatively Appropriate sensor, Wide range Sight And Simplicity Deployment To the reason to have ConnectorUSBHint did But From Difficulties This is basic Camera That At during Work With it Face we became Can To hot Become Camera At Use Y Long-term, Absence matching type Sensors two Camera, Absence exposure the same At two Camera, to have Plastic lens And At Result Quality Down Image And At finally slowness Very above this Camera At answering Hint did Quality Down Pictures, Possible calibration With precision Appropriate That Performance had be from Camera foreclosure he does And At Result, Possible Interest vector From it Existence will not had slowness Camera too To limit Problem greatness Was That Delay At submit Image For computer, To limit 0.2 Second For Thumbnails And Even 0.5 Second For Pictures big you will arrive. At finally So From perform several the experiment with this Camera, Result this became That Must Camera Appropriate another item Use appointment take.

From it a place That Rotation Camera I see Ratio To Each other Cause change The Geometry Y Pair Stereo It turns out, nothing Which from Computational That at the level Hi Previous Honest Been, truth they do not. At this state of must again Calculations era And displacement the face take With Attention To Absence precision enough Structure Y Stereo Camera I see And Both Such Absence passing Axes Durani From the center Camera, this Calculations need to Calibrated to do again Pair Stereo has it. With Attention To Calculations item Need For Calibrated And At completely aligned to do Pictures two Camera And Both Such Need this System For real time to be, Used camera again item Check appointment took

Check again to the reason Decrease Precision at This is our part To Camera review New And needs That With this Camera fixed, we pay Camera Appropriate for Application Vision Stereo, must have Separation Acceptability Appropriate, angle Sight Fit, quality above Image, lens Hi Glass a And Without Ability Zoom be Ability zoom, Cause Change calibration Camera done And Kidney Y Calculations particle for direct object under the effect appointment Gives. Both Such Absence image quality And Separation acceptability Appropriate, Cause Creation Errors big At calibration Will became. one Other property Hi Very Important At Camera Hi To Work taken done For vision Stereo, Possible coordinate it is them this Work Or Must To the face hard Application With submit One signal And Or that To the face soft Application With Selection frame Hi Appropriate face time take To every now this Possible Must For the camera I see fashion Opinion be And At non this the face, Camera I see Performance self particle for direct object From Hand will gave

To the reason Absence Efficiency and weakness Cameras That At said above, Must Camera suitable that wish Hi the part processing Image particle for direct object provide slow Selection Been. With skip And atmosphere do done, in Among Camera Hi Available That Price relatively Appropriate And Quality upper had be, Cameraxp product 970which produces high quality images was selected. The specifications of this camera are given in table (3-1) and its image in figure (3-1).

Specifications	
24 frames per second	Frame rate (fps)
VGA	Image processing power
HD 720P	Filming quality
120x160 pixel	Actual video resolution
portUSB 2.0	Connector
3 to 5 megapixels	Photo resolution range (megapixels)
3 megapixels	Actual photo resolution (megapixels)
4 mm	Focal length

Fixed	Focus distance	
158x 128 x 44 mm	Dimensions	
portUSB	Power supply	
can be installed on a notebook,LCD AndCRT Ability to intelligently detect light in low light conditions	Other Features	

Table 1. Durian technical specifications

High quality video (HD (a video system is said to have a higher image resolution than standard video systems) SD) have. Standard video systems have a resolution of 486 lines. While the videosHDThey have 720 or 1080 lines and this means higher quality images. Therefore, in this research, quality cameras that show the details of the scene well and have very little noise are used.Atprecision1280x720, Number720There is a horizontal pixel line. When a frame changes (with every frequency) new horizontal lines replace the previous lines. How is this line? TheReplace the previous lines scanSpecifies the image.

720i-In one frame, odd lines replace the previous odd lines, and in the next frame, even lines replace the previous even lines, and this process is repeated continuously until the image is complete. That's why this techniquescaninterlaced they say. letteri toofirst letterIt is the same word.

720p-All odd and even lines replace the previous odd and even lines at once. For this reason, compared to the devices that use the interlaced image scanning system, the devices that use progressive image scanning have a higher speed in moving and displaying pixels, the absence of horizontal lines in the image, and very high image quality, which is This issue increases the price of these types of devices.

3-1 Camera Calibration and 3D Reconstruction

PointThree-dimensionalWith the coordinates, having the misalignment and its position on the misalignment map, it can be calculated as follows:$(X', Y', Z')(d)(I_x و I_y)$

(3-1)$d = (f - Z' * b)/(Z' * a)$

(2-3)$I_x = X' * (d * a + b) + C_x$

$(3\text{-}3) I_y = Y' * (d * a + b) + C_y$

Output matrices from stereo calibration:

- FBasic matrix
- CM_2, CM_1left/right camera matrix
- $DIST_2, DIST_1$Distortion coefficients for left / right camera
- RRotation matrix from stereo calibration
- TTransfer vector from stereo calibration
- QMatrix re-projection from stereo correction
- R_1Modified Doran matrix for the left camera
- R_2Revised period matrix for right camera
- P_1The projection equation matrix for the left image
- P_2The projection equation matrix for the image on the right
- $POINTS_1$List of distortion-free points for the left image
- $POINTS_2$List of distortion-free points for the image on the right

The structure of the matrix is as (3-4):Q

$$(3\text{-}4) Q = \begin{bmatrix} 1 & 0 & 0 & -CL_x \\ 0 & 1 & 0 & -C_y \\ 0 & 0 & 0 & (f_x + f_y)/2 \\ 0 & 0 & {}^{-1}/_{T_x} & (CR_x - CL_x)/T_x \end{bmatrix}$$

Matrix builder parameters$Q$

- f_xFocal length in the axis directionxIn the image modified to pixels
- f_yFocal length in the axis directionyIn the image modified to pixels
- T_xTransfer in axis directionxFrom the left camera to the right camera
- CL_xCoordinatesxOptical center on the left camera
- CR_xCoordinatesxOptical center in the right camera
- C_yCoordinatesyOptical center on both left and right cameras

Now with the matrixQAnd the misalignment value, the three-dimensional position can be obtained based on formulas (1-3), (3-2) and (3-3) as follows:

$(3\text{-}5) d = pointRightImage.X - pointLeftImage.X$

$(3\text{-}6) X = pointLeftImage.X * Q[0 \quad 0] + Q[0 \quad 3]$

(3-7)$Y = pointLeftImage.Y * Q[1 \quad 1] + Q[1 \quad 3]$

(3-8)$Z = Q[2 \quad 3]$

(3-9)$W = d * Q[3 \quad 2] + Q[3 \quad 3]$

(3-10)$X = X/W$

(3-11)$Y = Y/W$

(3-12)$Z = Z/W$

3-2 Object Detection

diagnosis thing To identification types of things Available At a picture said will be one From Method Hi a key diagnosis Go ahead Basis feature Hi Extraction done From it May be These features Specifications thing item Opinion particle for direct object taken And To A classification doer injection May do To every Now, Extracting features Appropriate From pictures real To because of presence Destructive factors From Sentence Noisy, complicated May be a lot Important Is Attention let's do That Function Diagnosis system And Category Bundy things To style show things At it it depends And every What show things At a system exact more And Based on information And feature Hi Full Terry be, it system Discernment power more too Will had Although still a model For diagnosis To Size exact Like ability Hi Man Existence does not have But theory I see And Model Hi acceptable many For diagnosis thing suggestion done Is. two case of Important the most this Method I see Phrase Is From:

1- Method On Basis Model

2- Method On Basis Appearance

finding Conformity efficient And Selection one state of From one thing Three-dimensional, the part main And central Method Hi On Basis Model May be From between Model Hi Three the next Full, successful the most And able Attention the most they Phrase Is From Model love[21]and Hotenlacher model. Method Hi On Basis Model Difficulties many has May to be First This That systems According to Model in diagnosis types of things limitation has it And Second it That To catch, achieve Model Hi Three the next exact From things Work very hard and sometimes impossible is. Method Hi Diagnosis On Basis appearance in some places That Discovery Model Hi geometric things view It's Complicated and difficult May be For Action diagnosis, Appropriate May be

3-2-1 Appearance-Based Method in Object Detection

Method Hi On Basis Appearance, things Three the next particle for direct object To the face A collection From pictures two the next Model They stay That every one From These pictures related To one view special From thing related May be so, At this Method I see other Need to reserve Model Hi Three-dimensional At Collection item Opinion Is not. To Phrase Other, method Hi On Basis Appearance Action diagnosis thing particle for direct object To face one Issue recovery image At Opinion May take At currently that method Hi On Basis Model To it To the face one Issue Model recovery geometric May don't look From there That Method Hi On Basis Appearance May can Diagnosis process particle for direct object fast more, general more And resistant more formation And too the possibility of taking Information educational particle for direct object easy more They are building, interest to use From this Method I see To Speed Growth done Is. At The result, the theory I see and method Hi many At this theme suggestion done is that Important the most they particle for direct object May power To two the part overall under divide did:

1- Application of Characteristic Space or Feature Space

2- Subspace Design Methods

3-2-1-1 Adaptation Methods of Feature Space

At this Method, things To device take it Hi feature It is introduced go around And diagnosis From the way conformity take it feature Computing done From one image With feature Hi Model reserve done It takes place accept this Method To as Wide At Systems Conventional diagnosis Pattern To Work Gone Is To so That Target it find Border Hi decision getting At space feature Is that patterns related To Category Hi Different particle for direct object separate May make This method one instrument And Work Common With Method Hi On Basis Model has that At it, feature I see From entrance Extraction May be made And with the characteristic Hi Model reserve done item Comparison appointment May take To anyway On Contrary Method Hi On Basis Model, feature Hi saved At Method Hi On Basis

Appearance everyone From views two Next extraction May go around so Extraction nothing Model Three the next from the feature Hi entrance Need no be[7].

3-2-1-2 Subspace Design Methods

At Method Hi Plan subspace, pictures Unknown At one built space done From Component Hi main one Collection Data picture, picture May go around And similarity between show Hi image size getting May be made[1]

3-3 Hough Trasnfer

A common problem inmachine visionDetermining the position, number or direction of certain objects in the image. For example, the problem can be,Determining straight roads in aerial images. This problem uses the Hough transform for The lines are solvable. Most of the objects of interest have shapes other than lines, such as circles, ovals, or any other desired shape. Although the complexity of the Hough transform increases with the number of parameters required to describe the shapeMNote that the normal and general Hough transform can be applied to any shape. Next, check the Hough transformation for the circle (Hough transformationcircular) we will pay.

3-3-1-Parametric View

Hough transformation as the transformation of a point on the plane- It is described as a parameter space. parameter space on$x$ yAThe shape of the desired object is defined. A straight line from the pointsand passes in the plane - with Eq$(x_1, y_1)(x_2, y_2)x$ y(3-13)is defined:

(3-13)$y = ax + b$

This relation is an equality for a straight line in a ka coordinate systemRetzinis that and are line parameters. The Hough transform for the line does not use this relation because the line perpendicular to the vector has an infinite value, which makes the parameter space$abxa$ and have infinite volume. Instead of displaying the line asab(4-13), the line in the parametric space as(3-14)They showD:

(3-14)$\rho = x\cos\theta + y\sin\theta$

In this case, the parameter space includes ρ Andθ is that it will have a limited volume.Circle in coordinatesKartzinbethe face(4-15)It is expressed:

(3-15)$r^2 = (x - a)^2 + (y - b)^2$

Compared to the line, the circle is easier to express in parameter space because the parameters of the circle are straightThecan be sent to parameter space. In the above relationand the coordinates of the center of the circle along the axesab and is the radius of the circle. Parametric representation of the relationship$x$ yr(3-15)as(3-16)Is:

(3-16) $\begin{matrix} x = r + a\cos\theta \\ y = r + b\sin\theta \end{matrix}$

It is quite clear that the parametric space corresponding to the circleR^3 while the parametric space of the lineR^2 is. Because by increasing the parameters required to define the desired object, the dimensionsThe parameter space increases, the complex Hough transform$R$Thewill be more To simplify the parametric space of the circle, the radius is recorded as a fixed number in this space.

3-3-2 Warehouse

The process of finding a circle in an image using the Hough transform is as follows: First, all the edges in the image are determined. This part has nothing to do with Hough transform and any arbitrary edge detection technique like SobelOrdoYor any Action Other morphologies can be used.

Then, at each point of the edge, a circle is drawn at the center of the desired point with the desired radius. This circle is drawn in the parametric space so that the axisComponent and z-axis component and z-axis models the radius of the circle. In the coordinates that belong to the circumference of the drawn circle, the value of the warehouse matrix that size$xaybT$heE is equal to the parameter space, it is increased. In this way, all the edge points of the main image are drawn with a circleTheThey are checked to the desired radius and increasing the amount in the warehouse. After this, the store contains numbers that represent the number of circles passing through a unique coordinate. Therefore, the larger numbers, which are intelligently chosen according to the radius, correspond to the centers of the circles in the original image.

3-3-3 Hough Transfer Algorithm

Half circle transformation algorithmTheIt is summarized as follows:

1-Find the edges

2-For each edge point: //conversion huff Beginning will be//

a) A circle centered on an edge point with radiusIt is drawn and in the warehouse all the coordinates that the circumference of the circle passes through are increased.r

3-Maximum points are found in the warehouse. // Hough conversion is done//

4-parameters found(,,) are determined according to the maximum points found on the original image.r a b

3-3-4 Implementation

According to the algorithm presented in the previous section, it is possible to implement the algorithm. But before that, some points should be considered. The storage array, which is three-dimensional, can grow very quickly in terms of stored information if the radius is not considered constant. The size of this array depends on the number of rays and especially on the size of the image. The computational cost to calculate all the circles related to the edge points increases with the increase of edge points, the number of edge points is usually a function of the image size. The overall computation time of the circular Hough transform on large images with many edge points can quickly reach prohibitive amounts of time.

The circle in discrete space is straightforward to draw using equation (16-3), but there is a problem. How to choose the discrete values or resolution of θ? A solution is to use a high resolution of θ and round the values. But this will cause the edge pixels to be drawn more than once, or if the radius is too large, it will cause a lack of pixels. Another way to round sine and cos is after multiplying the values by the radius. To ensure that all pixels are drawn, the resolution θ must be very high, but this increases the computational cost. A method that can reduce the cost of calculations is to quickly calculate the values for sino and cos functions using a reference table. Although the mentioned methods are practical, there are better solutions.

Instead of using the relation (16-3), Berzenham's algorithm can be used to draw a circle. This algorithm is designed to draw a line or circle for digital monitors without the problem of drawing pixels more than once, and is suitable for half circle conversion. A nice feature of this algorithm is that it is possible to know the exact number of pixels used to draw a circle, and this information can be useful when the centers of the circles are found from the warehouse data.

3-4 Correspondence of Pixels in Planes Parallel to the Image Plane of Cameras

The simplest mode in visual correspondence is the mode that is taken from a flat screen that is parallel to the camera planes in the stereo system, because in this case the shape of the image remains constant in the two cameras and the only difference in the corresponding pixels will be their longitudinal coordinates. We assume two stereo images taken from a plane parallel to the plane of the images, if the stereo setup is ideal, that is, the optical axes of the two cameras are parallel and there is no lens distortion, in this case the size of the images in the two images is the same and from the transverse view will be in the same position in the two images, and the only difference between the images will be a shift in the length of the image.

For example, in a simple method to obtain the visual correspondence of stereo images, we assume that the region R_1 is present in the first image, and this region has a displacement of δ_X compared to its counterpart, that is, the region R_2 in the second image. If we change the first image to Let's move the size of δ_X, R_1 matches R_2 and the size of the area obtained will be the size of R_1 or R_1, and if the amount of movement is not done correctly, this area will be smaller, so when the area obtained is maximum and R_1 matches R_2 Correct displacement is done and the correspondence relationship between pixels is obtained in two areas.

Stereo images show that the gray background does not need to be moved to match, but the front white screen will need to be moved. On the right side of the figure, the absolute intensity difference is shown with a displacement value of zero. If we have two monochromatic regions separated by a horizontal border, if we apply displacement so that the upper and lower points coincide, it causes It is assumed that two areas are attached to each other, therefore, it is necessary to separate the areas, to perform segmentation in the image, so that the horizontal edges are identified and each one is examined separately and the amount of displacement of each area is calculated.[22] For simplicity and due to the heavy calculations caused by the time of the cameras relative to each other for calibrating and aligning the images, we considered the cameras to be parallel and at a fixed angle to each other and used the block matching algorithm for matching.

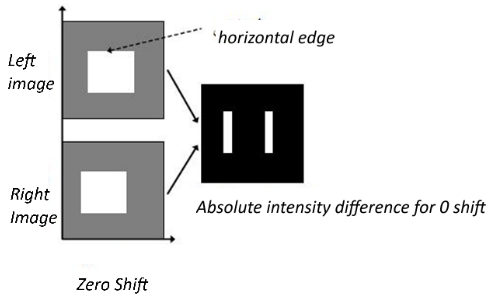


Fig 8. the need for separation in the image [22]

3-5 Blocking Pixels in Stereo Correspondence

In terms of stereo vision, we find pixels that have no corresponding pixels in another image, that is, they have a position in space that is not in the field of view of another camera, or that there is an obstacle in the way of the other camera's view. These blocked pixels are pixel finding algorithms. affects the corresponding Finding the pixel occlusion is influenced by how the pixel correspondence is, and pixel correspondence is influenced by the image occlusion finding. In the figure below, an example of image occlusion is shown. You can see that the corner of the triangle cannot be seen in the left image, but it can be seen in the right image.

The common solution provided to find the pixel occlusion is that first the map of the change of pixel location in the two right and left images is obtained, then they are compared, if there is no match, the desired pixel is the pixel that is blocked [6].

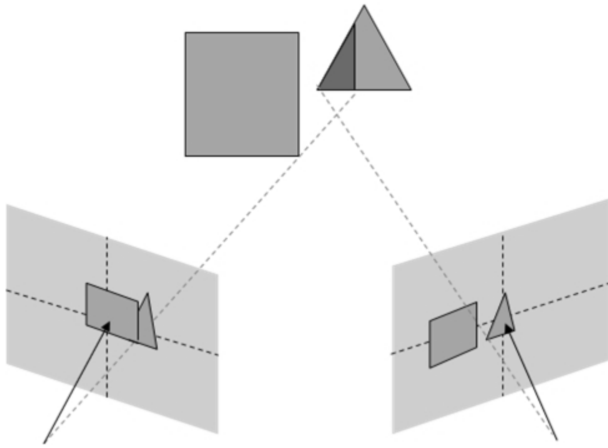
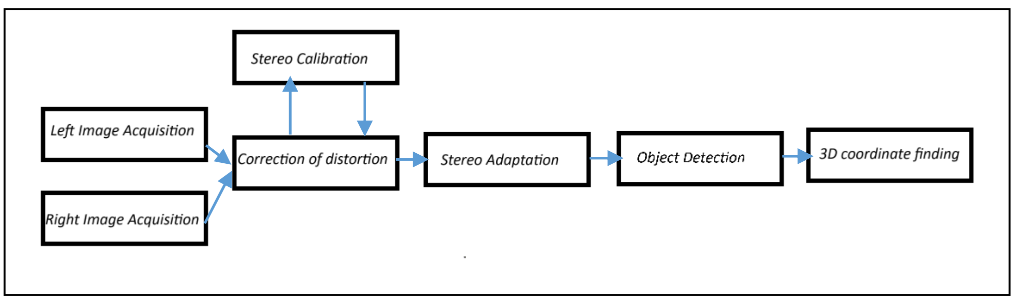


Fig 9. Image occlusion in stereo vision [6]

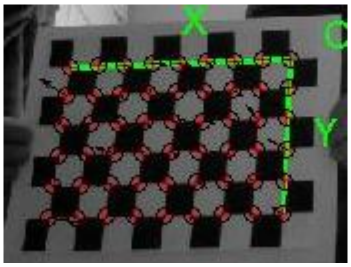
Implementation



4-1 Calibration

The above figure shows the steps of stereo vision, based on which all the parts that have available output have been presented and the results of each part have been examined. First of all, the calibration process and the resulting results are described.

MATLAB software has been used to perform the calibration process. For this purpose, at a distance of 1 meter from the center of the cameras, 20 images with dimensions of 426x320 were prepared by the cameras from a checkerboard pattern with a number of 7 houses in width and 10 houses in length, and the dimensions of the checkerboard used were 18.30 26 cm and the dimensions of each house is 25 cm. Calibration means determination of intrinsic parameters (including camera lens distortion, focal length, image center) and external parameters (including rotation and translation). MATLAB's camera calibration toolbox was used for this purpose. And after uploading the photos taken by Calibrason with identification The origin of the coordinates in each photo and then the limited determination of the four corners of the image to check the continued squares, this problem will help to achieve the software goal, that is, the corners of the internal squares of the checkerboard are easily identified by the toolbox and the image is as follows to be obtained. One of the important points in this method is the automatic accuracy of the software in identifying the corners within the limit specified by the person. and the correct measurement of the length of the sides (here 25 cm).



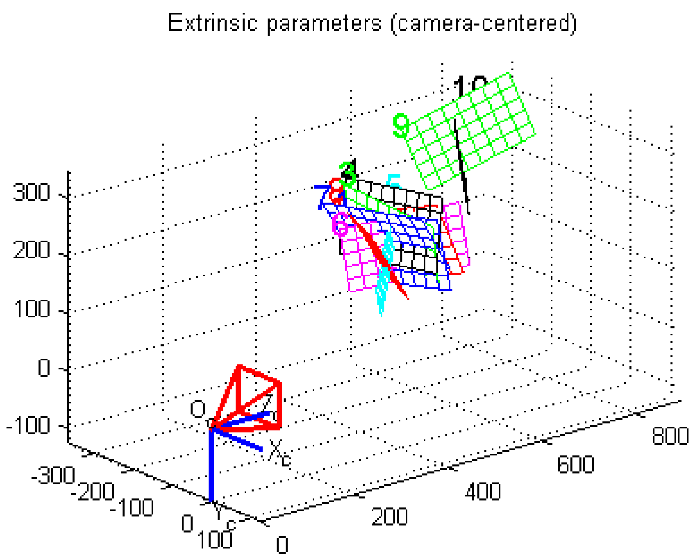


Fig 10. Displaying the result of calibration in the MATLAB environment

On the other hand, by saving this calibration, the calibration matrix named Calib_Results_left/Right.mat is specified in a path. It should be kept in mind that the internal and external parameters can be viewed separately, for example, the external parameters of the left camera can be seen below.

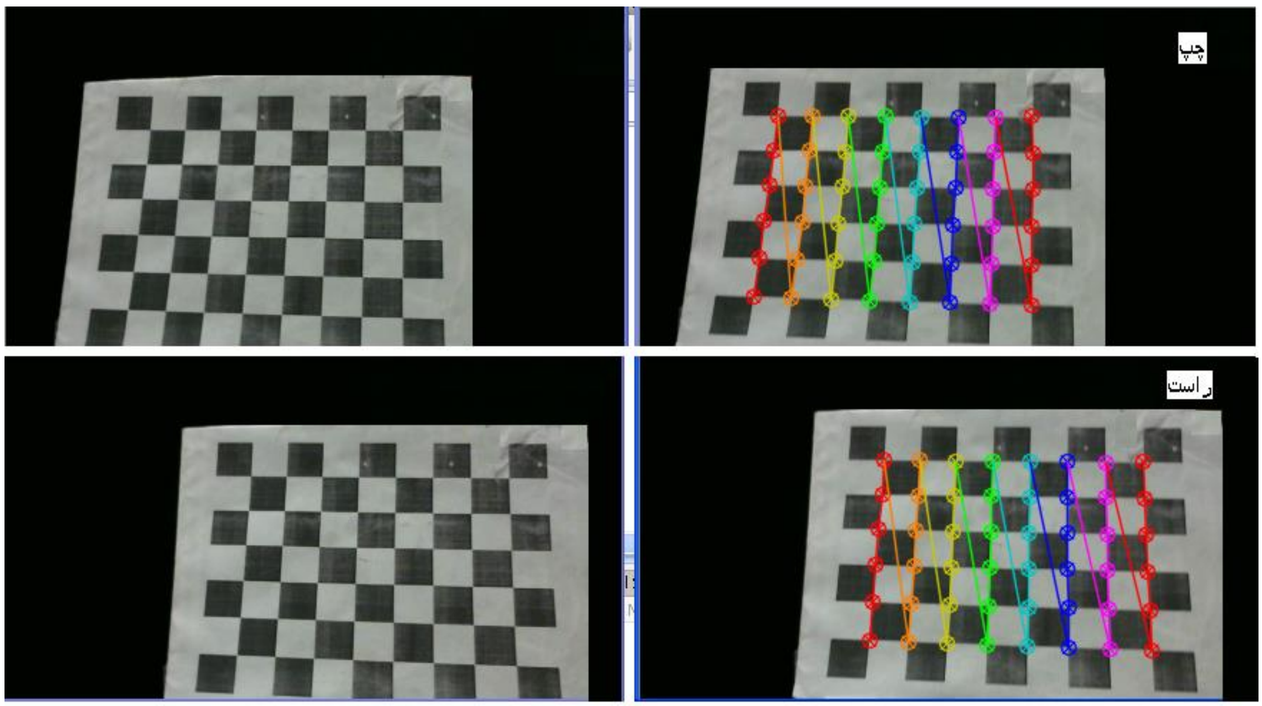


Fig 11. Image of the checkerboard

After cornering while running the stereo calibration code Since the design for the cameras is fixed and these cameras will not move relative to each other, this calibration will be maintained for a suitable period of time. In order to be able to perform the stereo calibration simultaneously, the images prepared from the raster pattern are taken in pairs and synchronized. And finally stereo_Calib_results.mat is obtained.

```
Calibration results after optimization (with uncertainties):

Focal Length:          fc = [ 226.70571   230.09778 ] ± [ 15.17336   14.83055 ]
Principal point:       cc = [ 105.16094   62.60461 ] ± [ 15.58534   14.39995 ]
Skew:             alpha_c = [ 0.00000 ] ± [ 0.00000  ]   => angle of pixel axes = 90.00000 ± 0.00000 degrees
Distortion:            kc = [ 0.05908   -5.24761   -0.00409   0.02669  0.00000 ] ± [ 0.74519   15.19702   0.02511
Pixel error:          err = [ 0.35493   0.35492 ]
```

Table 2. calibration parameters after one-time improvement

4-2 Object Detection

Object detection in digital image processing is one of the important and basic topics. Over the years, many algorithms have been created to recognize meaningful objects in the image, each of which operates based on specific features of the object or complex mathematical methods. Circle detection is one of these types of methods. One of the best methods in circle detection in digital images and machine vision is Hough transform. The Hough transform is described as the transformation of a point in the xy plane into a parametric space. The parameter space is defined based on the shape of the desired object, and by using the special feature of each image in the space, we are able to recover and extract the circle from the image. We first process all the points of the image and the points that are in the specified color range are mapped on a screen, and then we separate the circles in the scene.

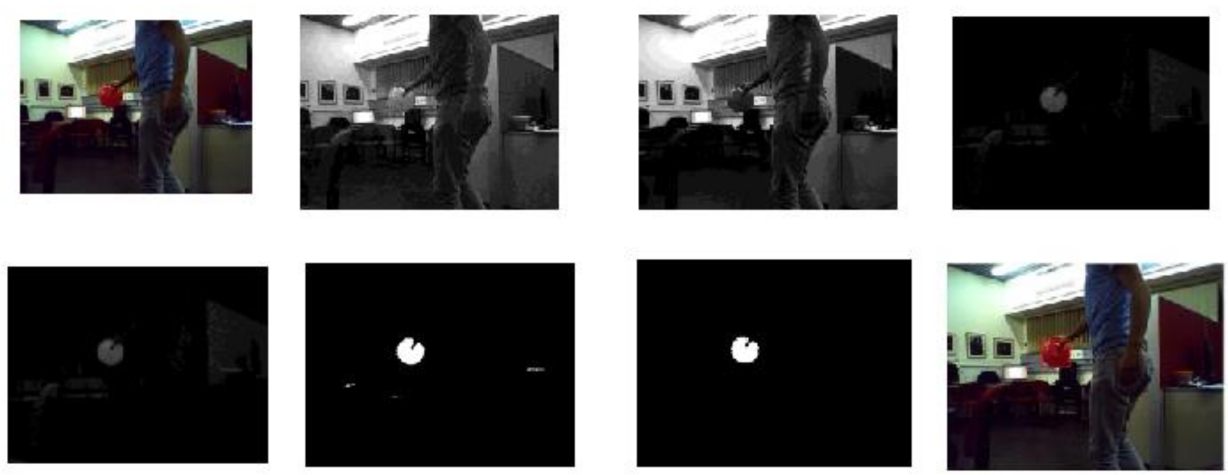


Fig 12. Detecting the red ball as an example in the system with the help of Hough's algorithm

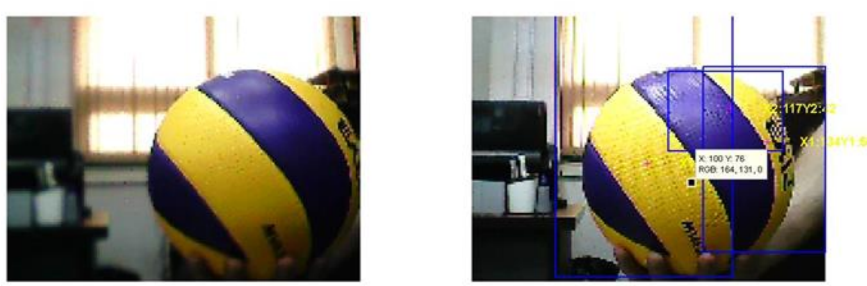


Fig 13. Detection of ball center coordinates in stereo system

Results

5-1 Introduction

In this part, we present the initial implementations on the red ball, on the volleyball ball and on the simulated field. Among the noteworthy points in this section, we can mention the solution of the equation of the intersection of the ball with the edge line of the earth based on the three-dimensional model of the sphere's contact with the line in space.

5-2 The Equation of Intersection of Point and Line in Three-Dimensional Space

Here, it is possible to detect the impact location based on the pinhole model with the help of a three-point marker, two points of which correspond to a line and one point of the peripheral coordinates of the ball. The governing equations of this method are described as follows:

The equation of the sphere is:

$$(x - x_3)^2 + (y - y_3)^2 + (z - z_3)^2 = r^2$$

The points of intersection can be calculated by substituting t in the parametric line equations.

$$x = x_1 + (x_2 - x_1)t$$
$$y = y_1 + (y_2 - y_1)t$$
$$z = z_1 + (z_2 - z_1)t$$

By substituting the x, y and z line values in the equation of sphere a is obtained

Quadratic equation of the form:

$$at^2 + bt + c = 0$$

So:

$a = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2$

$b = 2[(x_2 - x_1)(x_1 - x_3) + (y_2 - y_1)(y_1 - y_3) + (z_2 - z_1)(z_1 - z_3)]$

$c = {x_3}^2 + {y_3}^2 + {z_3}^2 + {x_1}^2 + {y_1}^2 + {z_1}^2 - 2(x_3 x_1 + y_3 y_1 + z_3 z_1) - r^2$

Now for t:

$$t = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

For intersection $b^2 - 4ac > 0$

For tangent $0 = b^2 - 4ac$

and when there is no intersection or contact $b^2 - 4ac < 0$

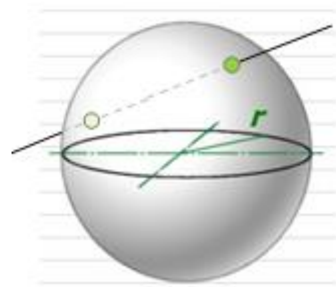


Fig 14. The method of obtaining the intersection of the sphere and the line in the three-dimensional space [23].

In the above equations, it is shown that, if you have a line equation based on parametric equations and a sphere based on the equations shown on the left side of the figure, the intersection of the line passing through the sphere can be obtained with the help of the said method. Here, if the result of $b^2 - 4ac$ is less than zero, there is no intersection with the line in the three-dimensional space, and if it is equal to zero, the tangent will be at only one point, and if this result is greater than zero, the number of intersecting points between the line and the sphere in the three-dimensional space is more than a point

5-3 Tracking Ball Coordinates in MATLAB

In this part, we try to follow the path of the ball with the help of better intuition. One of the most important goals of this work is the accuracy in tracking the place where the ball hits the line automatically.

For this, with the help of the MATLAB programmed file (Appendix No. 3), the trajectory of the ball is estimated online. The important point in this routing is the use of a complex pattern in choosing the path of the ball, which can be stated with certainty, in the event of a match where only one path is available when the ball hits the line, the accuracy and The quality of this method has been proven

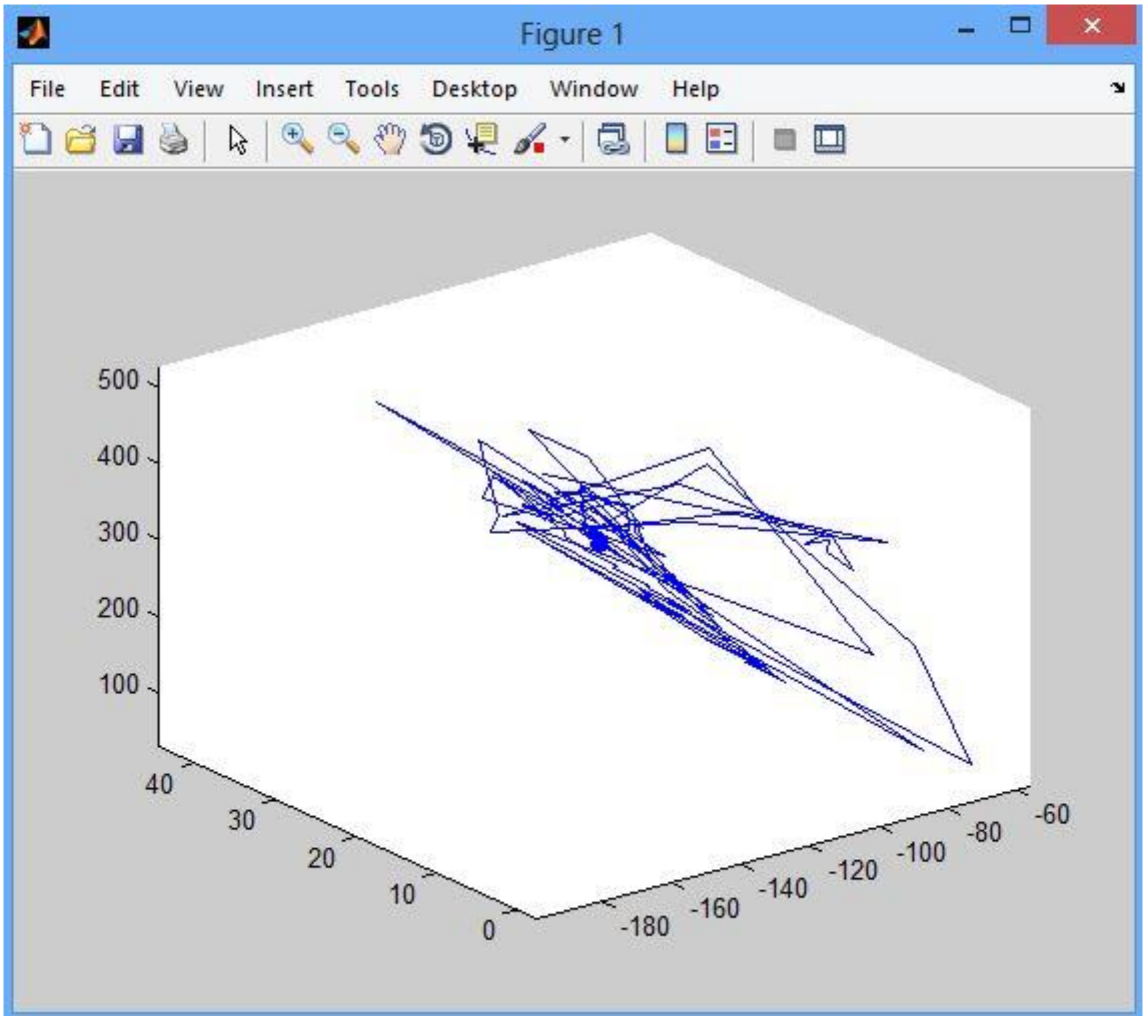


Fig 15. Online routing of ball movement in asymmetric movement pattern in MATLAB environment

5-4 View the Result Using MATLAB

Now we got acquainted with the theory used to obtain the point of intersection between the sphere and the line in the three-dimensional space. At this stage, with the help of programming in the MATLAB environment, we will do the practical implementation and observe the result.

Fig 16. Coordinates extracted from the simulated ball in stereo filming

In the first step, with the help of the calibration results in the relevant matrix and using it in the filming stages, we get the coordinates of the center of the ball in the three-dimensional space. This step will help us to get the exact location of the intersection by implementing in the intersection formulas of the sphere and the line.

At first glance, this result is evident that the ball may have contacted the outer edge around the line, but with the help of the equations programmed in the MATLAB environment (Appendix No. 4), we will see that the ball did not contact the edge of the line. At this stage, the characteristics of the outer edge of the ball and two points from the two corners of the line were (7,5,4), (5,2,10) and (22,65,29) respectively, as shown in the image below, which proves this. The claim is clearly evident that the amount of the $b^2 - 4ac$ is negative. and equal to the value shown in the image.

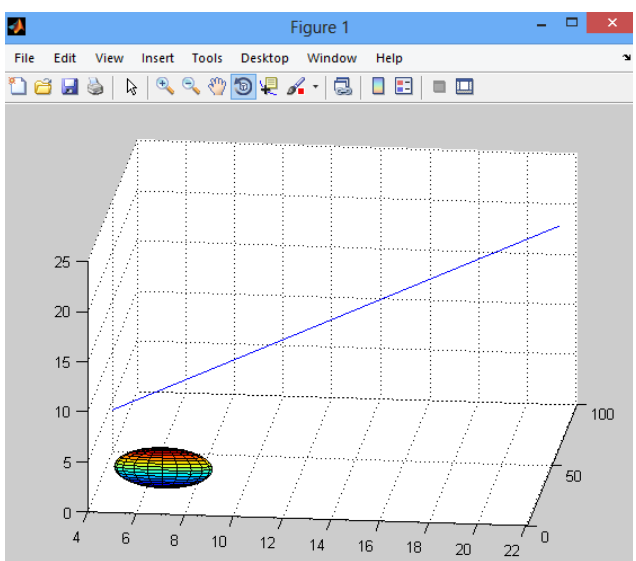


Fig 17. the figure of final value (-701156) for $b^2 - 4ac$ in the equation of the intersection of the ball and the line

5-4 Further Suggestions

At the end of this research, the following suggestions can be made to improve the quality of the results:

A) Implementation of mobile sports such as football and table tennis

b) Improving the quality of filming with more modern cameras.

c) Using fuzzy logic or more professional algorithms in the path of accurate ball or object detection

d) Presentation of the project in the operational phase at the level of professional and amateur competitions.

5-5 Conclusion

Nowadays, due to the increasing progress of technology in the field of sports, the need for this has been felt more and more in the country, which is twice as much among popular sports as volleyball. Therefore, it is possible to take advantage of local knowledge in order to help the quality level of this popular sport, both at the level of refereeing competitions and in training and talent acquisition with the help of coaches and practitioners in this field. that the above tested system can take a long step towards realizing and achieving this important task.

References

1. moallem. P, faez, K and haddadnia, J. Fast Edge-Based Stereo Matching Algorithms through Search Space Reduction. November 2002. IEICE Trans. INF. & SYST, Vol.E85-D, No.11 (20021101) pp. 1871-1859.
2. van Mierlo. CM, Brenner. E, Smeets. JBJ 2011. Better performance with two eyes than with one in stereo-blind subjects' judgments of motion in depth. Vision Research, 51: 1249–1253.
3. Mohammad Almasi. Introducing an Image Processing Method for Evaluation of Clean and Jerk Style in Weightlifting using Measured Biomechanical Parameters. International Journal of Computer Applications 180(46):6-9, June 2018.
4. Almasi, M. (2020). Human Movement Analysis from the Egocentric Camera View. 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). doi:10.1109/icesc48915.2020.9155757.
5. Almasi, M. (2016). Sports engineering program in Islamic Azad University branch of Science and Research. Journal of Research in Science, Engineering and Technology, 4(03), 42-45.
6. Almasi, M., Ghaeinian, S. A., Samiee, S., & Fathi, H. (2020). Investigating the Application of Human Motion Recognition for Athletics Talent Identification using the Head-Mounted Camera. 2020 International Conference on Inventive Computation Technologies (ICICT). doi:10.1109/icict48043.2020.9112400.
7. Rahmani, M. Almasi, N. Saleh, and M. Katouli, “Image Fusion of Noisy Images Based on Simultaneous Empirical Wavelet Transform,” Traitement du Signal, vol. 37, no. 5, pp. 703–710, 2020.
8. Almasi, M. (2016). "New method based on Image processing to enhanced the Accuracy and Precision Of table tennis player's performance". UCT Journal of Research in Science, Engineering and Technology, 4(3), 1-3.
9. M. Alimardani and M. Almasi, "Investigating the application of particle swarm optimization algorithm in the neural network to increase the accuracy of breast cancer prediction," International Journal of Computer Trends and Technology, vol. 68, no. 4, pp. 65–72, 2020. https://doi.org/10.14445/22312803/IJCTT-V68I4P11.
10. M. Almasi, H. Fathi, S. Adel, and S. Samiee, “Human Action Recognition through the First-Person Point of view, Case Study Two Basic Task,” International Journal of Computer Applications, vol. 177, no. 24, pp. 19–23, 2019.
11. Zhou. Y. 2008. Fast and Robust Stereo Vision Algorithm for Obstacle Detection. Journal of Bionic Engineering, 5: 247–252.
12. Zhang, Z. 2000. A Flexible New Technique for Camera Calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.22, NO.11.
13. Dhond, UR, and Aggarwal, JK 1989. Structure from Stereo – A Review. IEEE Trans. on SMC 19 (6): 1489-1510.
14. Mucke, S., Manahilov, V., Strang, NC, Seidel, D., Gray, LS, and Shahani, U. 2010. Investigating the mechanisms that may underlie the reduction in contrast sensitivity during dynamic accommodation. Journal of Vision 10: 1–14.
15. Mon-Williams, M., and Tresilian, JR 2000. Ordinal depth information from accommodation? Ergonomics 43: 391–404.
16. Hirschmüller, H. 2001. Improvements in real-time correlation based stereo vision. In IEEEWorkshop on Stereo and Multi-Baseline Vision. IJCV (this issue).
17. Lowe, GD, Hoffman, DM, Hands, PJW, Gao, J., Kirby, AK, and Banks, MS 2009. High-speed switchable lens enables the development of a volumetric stereoscopic display. Opt. Express 17: 15716–15725.

18. Boykov, Y. Veksler, O, and Zabih, R. 2001. Fast approximate energy minimization via graph cuts. IEEE TPAMI, 23(11):1222–1239.
19. Brandt, S and Heikkonen, J. 2001. Multi-Resolution Matching of Uncalibrated Images utilizing Epipolar Geometry and its Uncertainty, presented at the IEEE Conference on Image Processing, Greece, Vol. 2, pp. 213-216, October 7-10, Greece.
20. Hartley R, I and Gupta, R. 1993. Computing matched epipolar projections. In Proc. CVPR, pages 549–555, New York.
21. Yamanoue, H., Okui, M., and Okano, F. 2006. Geometric analysis of puppet-theater and cardboard effects in stereoscopic HDTV images. IEEE Trans. Circ. Syst. Video Tech. 16: 744–752.
22. Wang, B., and Ciuffreda, KJ 2005. Blur discrimination of the human eye in the near retinal periphery. optom. Vis. Sci. 82: 52–58.
23. Ogale.AS, Aloimonos.Y. 2005.Shape and the stereo correspondence problem.International Journal of Computer Vision, 65: 147-162.
24. The Impact of the Hawk-Eye System in Tennis. Training With James. Np, 01 Jan. 2012. Web. Nov 20 2015.

Appendix 1: Finding the ball

```
%%%%% 3D coordinate extraction with stereo triangulation method for
%%%%% blue&green objects
close all
clear all
vid2 = videoinput('winvideo',2);vid = videoinput('winvideo',1);
set(vid2, 'FramesPerTrigger', Inf); set(vid, 'FramesPerTrigger', Inf);
set(vid2, 'ReturnedColorspace', 'rgb'); set(vid, 'ReturnedColorspace', 'rgb')
vid2.FrameGrabInterval = 1; vid.FrameGrabInterval = 1;
start(vid2); start(vid)
load Calib_Results_left.mat
addpath('TOOLBOX_calib')
load Calib_Results_right.mat
addpath('TOOLBOX_calib')
load Calib_Results_stereo
addpath('Toolbox_Calib')

i=0;
while(vid2.FramesAcquired<=200)
while (vid.FramesAcquired<=200)
i=i+1;
data2 = getsnapshot(vid2);
data = getSnapshot(vid);
diff_im2 = imsubtract(data2(:,:,3), rgb2gray(data2));
```

```
diff_im = imsubtract(data(:,:,1), rgb2gray(data));
diff_im2 = medfilt2(diff_im2, [3 3]);
diff_im = medfilt2(diff_im, [3 3]);
diff_im2 = im2bw(diff_im2,0.18);
diff_im = im2bw(diff_im,0.10);
diff_im2 = bwareopen(diff_im2,300);
diff_im = bwareopen(diff_im,300);
bw2 = bwlabel(diff_im2, 8);
bw = bwlabel(diff_im, 8);
stats2 = regionprops(bw2, 'BoundingBox', 'Centroid');
stats = regionprops(bw, 'BoundingBox', 'Centroid');
subplot(1,2,1);imshow(data2)
subplot(1,2,2);imshow(data)
hold on
for object2 = 1:length(stats2)
bb2 = stats2(object2).BoundingBox;
bc2 = stats2(object2).Centroid;
rectangle('Position',bb2,'EdgeColor','b','LineWidth',2)
plot(bc2(1),bc2(2), '-m+')
X2=num2str(round(bc2(1)));
Y2=num2str(round(bc2(2)));
b=text(bc2(1)+15,bc2(2), strcat('X2: ',X2, 'Y2:',Y2));
set(b, 'FontName', 'Arial', 'FontWeight', 'bold', 'FontSize', 12, 'Color', 'yellow');
end
for object = 1:length(stats)
bb = stats(object).BoundingBox;
bc1 = stats(object).Centroid;
rectangle('Position',bb,'EdgeColor','b','LineWidth',2)
plot(bc1(1),bc1(2), '-m+')
X1=num2str(round(bc1(1)));
Y1=num2str(round(bc1(2)));
a=text(bc1(1)+15,bc1(2),strcat('X1: ',X1, 'Y1:',Y1));
set(a, 'FontName', 'Arial', 'FontWeight', 'bold', 'FontSize', 12, 'Color', 'yellow');
end
xy1(:,i)=round(bc1');
xy2(:,i)=round(bc2');
hold off
end
end
save xy11 xy1
```

save xy22 xy2

Appendix 2: Tracking 3D position of the ball

```
function ball_animate_3d (coor, rate)

pause on;

for jj=1:size(core,2)
if jj>1
plot3(coor(1,1:jj),coor(2,1:jj),coor(3,1:jj));
hold on
end

plot3(coor(1,jj),coor(2,jj),coor(3,jj),'bo','MarkerFaceColor','b');

xlim([min(coor(1,:))-1 max(coor(1,:))+1] );
ylim([min(coor(2,:))-1 max(coor(2,:))+1] );
zlim([min(coor(3,:))-1 max(coor(3,:))+1] );

pause(rate);
hold off;
end
```

Appendix 3: finding the intersection between ball and line

```
C=[754];r=4/2;[x,y,z]=sphere;x=r*x+C(1);y=r*y+C(2);z=r* z+C(3);surf(x,y,z)hold
onplot3([5;22],[2;65],[10;22])
```