Three Dimensional Motion Tracking from Stereo Video

Samuel E. Mermall
Physics Department, The College of Wooster, Wooster, Ohio 44691, USA
(Dated: May 9, 2012)

The motivation for this work was the desire to track the motion of a fluttering streamer and analyze its motion to determine whether it was chaotic or non-chaotic and, if chaotic, determine at what point the motion of the streamer became non-chaotic. Toward that goal, a computer model was developed in *Mathematica* for tracking an object in 3-dimensional space using two sets of coordinates from two different reference points in a stereoscopic camera system. Additionally, a physical mirror apparatus was designed and implemented for the creation of a stereoscopic camera system using a single camera. Three-dimensional trajectories were thereby reconstructed from stereo video.

PACS numbers: 42.15.-i, 42.15.Eq, 07.68.+m, 47.27.-i, 47.52.+j

INTRODUCTION AND THEORY

Stereoscopic cameras allow for the calculation of the depth of a point in space when given two horizontal and two vertical positions via triangulation [1]. By tracking the motion of the ends of a streamer in three dimensions as it rotates rapidly about a central axis, analyses of its chaotic fluttering, twisting, and bending can be performed. This experiment, constrained by time and technology, is less focused on the chaotic motion of the streamers and is more focused on the tracking of a particle moving in three dimensions with only one camera. A system utilizing a series of mirrors and one high speed camera was constructed in order to create a stereoscopic perspective. This allowed for the eventual representation of the optical data collected by one camera to be translated into the path of an object moving in 3D space.

In order to track objects in 3D space from video, a stereoscopic camera system is necessary. A stereoscopic system is able to determine depth because of the known differences in perspective between each camera involved. In a two camera system, as seen in Fig. 1, the location of an object in three dimensional space is a matter of relatively simple geometry and vector addition if the cameras are assumed to be pinhole cameras, wherein there is no lens and a small aperture.

In reference to Fig. 1, the coordinates for an object in three dimensions can be calculated by incorporating the camera separation, δ, the object's position on the x-axis as seen by the right and left cameras, the average position on the y-axis as seen by the right and left cameras, the focal length f of the cameras, and the magnification constants of the cameras' lenses. Using these values and the assumption of a pinhole camera, the position of the object can be calculated by the geometric principal of similar triangles, where the triangle made up of sides x_L and f as seen in Fig. 1 is similar to the leftmost triangle projected from the pinhole plane upon the pixel plane and the large triangle with sides X and Z. Likewise, the triangle with sides x_R and f is similar to the rightmost triangle projected from the pinhole plane upon the pixel plane and the grey triangle with sides Z and $X - \delta$ as seen in Fig. 1.

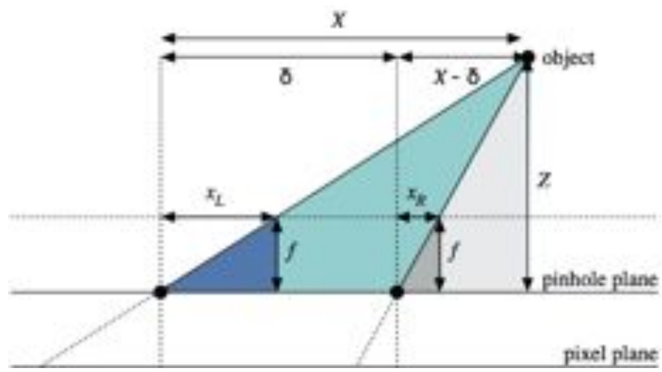

FIG. 1: This is a diagram of the geometric system used to calculate an object's position in three-dimensinal space, where X is the x-coordinate of the object, Z is the z-coordinate of the object, f is the focal length of the pinhole cameras, δ is the separation between the two cameras, x_L is the observed x-coordinate of the object from the left camera, and x_R is the observed x-coordinate from the right camera.

By using similar triangles, the parallax (apparent object displacement after a shift in perspective) is used to calculated the object's location in space [1]. In Fig. 1, x_L is the x-coordinate as seen by the left camera, x_R is the x-coordinate as seen by the right camera, δ is the camera separation, $\bar{y}$ is the average of the y-values taken from the left and right cameras, and f is the focal length of the cameras. The principal of similar triangles shows that

$$\frac{x_R}{X-\delta} = \frac{x_L}{X}. \quad (1)$$

Similarly,

$$\frac{f}{x_L} = \frac{Z}{X}. \quad (2)$$

By solving for X and Z (the y-coordinate will be the same for both cameras if both cameras have identical y-coordinates), the object's coordinates can be solved for individually assuming a common origin is set for both perspectives. The object's x-coordinate, or X in Fig. 1 and Eq. 3, in three dimensional space can be calculated by the simple algebraic manipulation of Eq. 1. Whereby,

$$X = x_L \frac{\delta}{x_L - x_R}. \quad (3)$$

The object's y-coordinate, Y, in three dimensional can be calculated by

$$Y = \bar{y} \frac{\delta}{x_L - x_R}. \quad (4)$$

The object's z-coordinate in three dimensions, represented by Z in Fig. 1, Eq. 5, and Eq. 6, can be calculated by the simple algebraic manipulation of Eq. 2. Whereby,

$$Z = f \frac{X}{x_L}. \quad (5)$$

By plugging in the value of the x-coordinate found in Eq. 3,

$$Z = f \frac{\delta}{x_L - x_R}. \quad (6)$$

When looking at a stereoscopic system utilizing one camera rather than two, something must be done so that one camera is able to receive light from two different vantage points at the same time. In order to do this, a series of mirrors must be used. To best fit the two-camera model as outlined above, one camera must face the corner created by two mirrors which are perpendicular to one another at their respective edges. Parallel to each perpendicular mirror must be another mirror, as seen in Fig. 2. These mirrors must be parallel due to the law of reflection, which states that the angle of incidence from the normal at which light strikes a mirror is equal to the angle of reflection from normal at which the light leaves the mirror ($\theta_i = \theta_r$) [2].

Since the the angle of incidence of the reflecting light is equal to the angle of reflection, even if the light does not strike the mirrors at a 45° (which is the vast majority of the light) the light coming from the streamers will still enter the camera. The camera will see a split image. The left half of the image will be the perspective from the left parallel mirror, and the right half will be the perspective from the right parallel mirror, which can be seen in Fig. 3. By implementing this system of mirrors, it is possible to create a stereoscopic system with a single camera as long as a common faraway origin can be found and set from both perspective. This eliminates synchronization issues, focal length differences, and lens magnification differences of a setup using two cameras.

EXPERIMENTAL

Pre-Testing

The initial step in the experiment was the derivation of the equations for an object's position in three dimensional space (Eq. 3, Eq. 4, and Eq. 6). Once these equations had been derived, and the parameters required for object tracking had been identified, a program was written on Wolfram's *Mathematica* software. This program read two sets of x and y data that was either user-inputed or read from a text file. This data was then centered about the origin by subtracting each x and y value by their respective averages. The average of all the x-coordinate data points from the left camera was subtracted from each individual x-coordinate measured from the left camera. The average of all the y-coordinate data from the left camera was subtracted from each individual y-coordinate measured from the left camera, just as with the x-coordinate data. The same thing was done for the data collected from the right camera. Since the cameras were considered to be at an identical height, the y-coordinate data for each individual point collected by the left and right cameras were added and then divided by two, giving the average y-coordinate of the object. Finally, the equations for the coordinates of an object in three dimensional space were incorporated and the parameters of focal length f, camera separation δ, and magnification constant m were initialized. These equations read the centered x-coordinates from both cameras and the average y-coordinates and plotted the resulting points in space.

Initial Program Testing

Once the initial framework of the program had been developed, the testing of simple motions took place in order to determine the program's accuracy and calibrate the parameters m and f. Using Vernier *Logger Pro v. 3.8.4* software, the tracking of an object through two dimensional space is possible as long as the software is given an origin and a scale. Two high-definition iMac desktop integrated cameras with a camera separation $\delta = 52.75$ cm were used as stereocameras for filming. In order to test the object-tracking capabilities of the software, a piece of bright red tape attached to a bicycle wheel with a radius of 30 cm oriented 45° from the vertical axis leaning away from the cameras. The wheel was held in place by a ringstand and several clamps. The wheel was spun and video was taken of several revolutions. Once the videos from both the left and right iMac cameras were synchronized and analyzed in *Logger Pro*, one revolution of the piece of red tape was analyzed as it moved with the wheel. This test was performed several times in order to ensure that the data collected was accurate.

When the data taken from the bicycle wheel tracking was inserted into the *Mathematica* program, the results were promising. In *Mathematica*, the motion of the tape as it moved with the wheel was mostly circular. Slightly jagged irregularities observed in the shape of the tape's motion were products of frame-skipping within the iMac cameras during filming. With these tests, it was determined that an appropriate focal length for the iMac cameras (under the assumption that both cameras acted as pinhole cameras) was $f = 30$ cm and the magnification constant to be $m = 0.26$. With this calibration complete, testing and analyzing the streamers' motion could begin.

Initial Streamer Testing

An Ealing Air Gyroscope was placed at the center of a large, level table. A thin wooden rod of length 27.94 cm was attached to the top of the gyroscope, laying perfectly horizontally. Attached to each end of this wooden rod was a paper towel streamer (with color-coded ends for video tracking), each being 15 cm in length and 3.5 cm in width. The streamers were 0.1 cm from each end of the rod. Two iMac cameras, just as with the bicycle wheel, were positioned to film the streamers. The gyroscope, resting on a cushion of compressed N_2 gas, was able to rotate without being acted on by a large amount of friction. When the gyroscope rotated with a high enough angular velocity, the streamers would flutter. This fluttering was filmed by the two cameras.

Not far into testing, it became apparent that this method of tracking the streamers was ineffective. The frame rate of the iMac high definition cameras is 30 fps (frames per second). This is not a high enough frame rate. Even with the addition of two bright lamps with 150W lightbulbs, the motion of the streamers as they fluttered was extremely blurred and choppy. Per frame, the streamers would rotate too much, making even the best plot of their motion bumpy and polygonal. The solution was to film the streamers with a faster camera with a higher frame rate. The caveat was that only one high speed camera was available, a Casio EX-FH20 with its tripod. In order to track motion in three dimensional space, two cameras (or perspectives) are necessary.

Mirror Apparatus

The next step in the experiment was determining how to film an object moving in three dimensional space from two different perspectives with one camera. The solution was to split the image received by the camera in half. The left half of the frame would be from a perspective to the left of the streamers, and the right half of the frame would be from a perspective to the right of the streamers. In order to accomplish this, an apparatus consisting of four mirrors was constructed. Two of the mirrors were placed perpendicular to one another at their edges, forming a corner. Two other mirrors were placed in parallel with the mirrors making up the corner. With a single camera facing directly into the corner, the image observed would be what ever the parallel mirrors were facing.

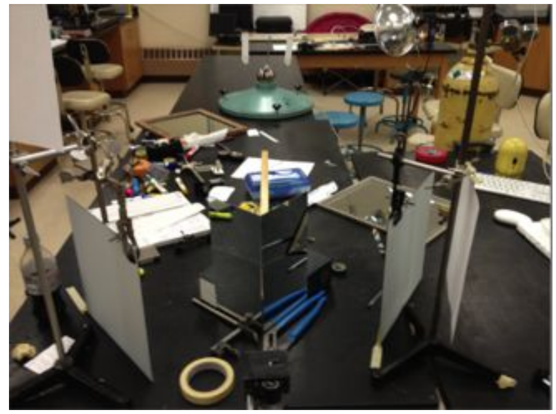

FIG. 2: This is a picture of the final mirror apparatus constructed from mirrored tiles. This apparatus is much improved over the previous versions. Its optical components are aligned very well, making the data collected while using it much more precise.

The first mirror apparatus was initially constructed out of two similar mirrors (making up the parallel mirrors) and two non-similar mirrors which formed the corner. The apparatus was held together with a large number of clamps, ring stands, and lab-jacks. The issue with this apparatus was that proper alignment of the mirrors was very difficult. This became an especially large problem when the interior corner mirrors turned out to be too large, blocking part of the streamer apparatus when the exterior parallel mirrors were properly positioned.

In order to solve this problem, a second mirror apparatus was built. This apparatus was assembled using three 30.48 cm $\times$ 30.48 cm mirrored tiles, two ring stands, two clamps, and one hardcover book (preserving the 90° angle between the corner mirrors). The corner mirrors were made up of one half of a full mirrored tile, each measuring 15.24 cm $\times$ 30.48 cm and oriented vertically. Since these mirrors were all either similar or equal in shape, they could be more easily aligned and adjusted. The final mirror apparatus, as see in Fig. 2, is much neater and less prone to error simply because there is less that can go wrong or be maladjusted during data collection. At this point, data collection using the high speed camera could begin.

Streamer Testing with High Speed Camera

Once the new mirror apparatus had been set up, precise data collection could begin. With the new mirror apparatus, the parallel mirrors could actually be placed in parallel, and a hardcover book ensured that a true right angle was preserved at the corner made by the perpendicular mirrors. The gyroscope was brought up to speed and the streamers were filmed at 200 fps five separate times. The data was then analyzed in *Logger Pro*. The origin used for both sets of data was the blue center of a bulls-eye drawn on a dry-erase board against the rear of the laboratory, as seen in Fig. 3.

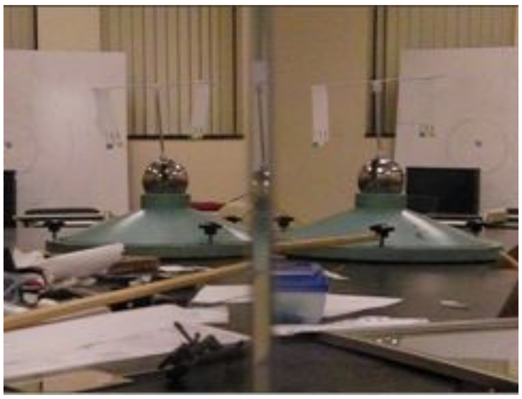

FIG. 3: This is the view from the Casio EX-FH20 high speed camera used in the experiment as it looks straight into a corner made by two perpendicular mirrors. The streamers can be seen from the perspectives of both the left parallel mirror and the right parallel mirror. The origin used for analysis in both the left and right perspective is the blue dot at the center of a bulls-eye against the back wall.

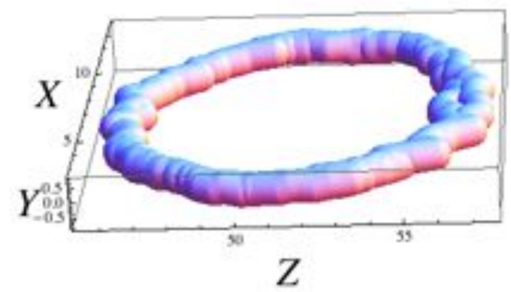

FIG. 4: This a screenshot from the *Mathematica* tracking the path of the one point on one streamer as it does two revolutions with angular velocity $\omega \approx 13.09$ rad/s.

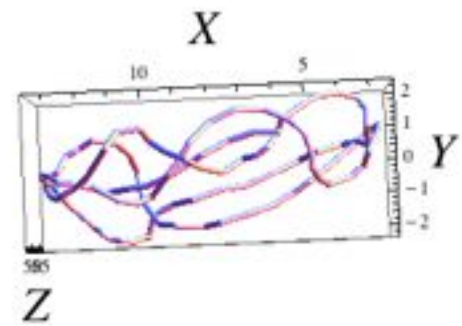

FIG. 5: This a screenshot from the *Mathematica* tracking the path of the one point on one streamer as it does two revolutions with angular velocity $\omega \approx 8\pi$ rad/s.

RESULTS

The precision of the improved mirror apparatus combined with the *Mathematica* program produced beautiful results. *Logger Pro* analyses were performed upon five sets of data. However, the most impressive are the multi-revolution data sets. For these multi-revolution data sets, one point at the tip of the streamer was tracked as it completed two full rotations about the gyroscope'saxis of rotation. One set of data was taken when the gyroscope was moving with a slow rotational velocity, such that the streamers were not fluttering but dragging slightly behind the rod. This motion is most definitely non-chaotic. Another set of multi-revolution data was analyzed when the streamer was rotating very quickly, such that the streamers were fluttering wildly.

For the first set of multi-revolution data, when the streamers were moving at a slow rotational velocity ($\omega \approx 13.09$ rad/s), the motion of the streamer through space can be seen in Fig. 4. For the second set of multi-revolution data, when the streamers were rotating with a much greater rotational velocity ($\omega \approx 8\pi$ rad/s), the motion of the streamer through space can be seen in Fig. 5.

CONCLUSION

In conclusion, I am glad that I was able to develop a model for tracking motion of an object in three dimensional space. I am especially pleased that I was able to do it with one camera, as well. Having one camera eliminates error from synchronization and differences in lenses. The results from the model not only shows how well the *Mathematica* program works, but how well the mirror apparatus works as well. While it is unfortunate that I was unable to actually analyze the data for chaotic motion as I had first intended upon doing, I still feel like I accomplished a lot. The techniques I performed over the course of this experiment will be valuable for anybody wanting to track an object in 3D space or working with a one camera stereoscopic system.

[1] Cheon, L., Hyok, S., Choi, B., Ho, Y. *3D Scene Capturing using Stereoscopic Cameras and a time-of-flight Camera*. IEEE Transactions on Consumer Electronics Aug2011, Vol. 57 Issue 3, p1370-1376.

[2] Heath. *A history of Greek mathematics. Volume II: From Aristarchus to Diophantus*. 1981.