Creating Panoramic Images: A hardware comparison between Sony BRC-300 and EVI-HD1 Cameras

Chris Cornelius, Ian McGinnis, and Charles Nye

Abstract

We investigate the use of pan-and-tilt cameras for the capture of panoramic data sets, and discuss two specific systems. At the core of these systems are two different robotic cameras: the Sony BRC-300 and the Sony EVI-HD1. Through comparing and contrasting these devices, we enumerate four important factors involved in the construction of panoramic capture systems and desirable qualities for each. First, the optical quality of the cameras is important -- and the EVI-HD1 camera was found to give exemplary high-resolution images. Second, and most important to our procedure, the motion of both cameras is evaluated for precision and repeatability. Such accuracy is necessary if subsequent image processing will rely upon the position information of the camera to guide creation of the panorama. Further, we explored some simple techniques for determining the repeatability and tilt range of a camera. Two capture cards were also surveyed, a Blackmagic Design card for capturing High Definition (HD) images and a generic brand card for acquiring Standard Definition (SD) images. Each capture card used a different Application Programming Interface (API) which each had unique merits and downsides. We also discuss our choice of storage methods for images as well as explain our storage format for recording which section of the sky each of the images contains. Finally, we describe techniques for creating image stitching software which combines captured images into a unified panorama using advanced methods to avoid image processing artifacts.

We conclude with a theoretical description of an ideal device for use in a pan-tilt panoramic system. We realize that many developers may need to make choices based upon cost, and we therefore demonstrate that the most important characteristic of a panoramic camera for this application is its pan-tilt precision and accuracy. By maintaining high standards of precision this ideal camera would allow for the creation of composite panoramic images and minimize unwanted effects.

1. Introduction

Applications of digital images can be improved by gathering and retaining as much information as possible. While advances in technology continually increase the number of pixels in an image, the edges of an image continue to sharply and irrevocably limit the data in a picture. This loss of information is an ongoing problem in computer vision.

Panoramic images provide a possible solution to this trouble. A panorama can achieve a very large field of view by stitching individual frames into one large image. By combining many small images with high individual pixel count, panoramas can also achieve very high resolution. After capturing a small field of view at high resolution, an expert may stitch many of those small, but detailed, images together, increasing the field of view, while retaining the detail of the smaller images.

Here we investigate the way that two different video cameras behave when used to capture a panoramic image. Both cameras incorporate built-in pan-tilt functionality, allowing them to scan their environment using motors integrated into the camera body. Video capture cards allow a computer to read images from the cameras, and store them to disk. Custom software is then used to composite the individual frames of data into large panoramic images.

We hope that this paper will serve as a guide to those who wish to construct similar panoramic capture systems. To assist such advances we not only describe two such systems, but also generalize our experience and data to apply to other systems. Through these efforts, we attempt to indicate which characteristics of cameras, capture cards, and software are critical in the design and implementation of a successful computer vision project.

2. Related work

Our team is not the first at Saint Olaf to investigate robotic video cameras for use in computer vision. We are building on the work of past individuals including Daniel Wiebe, Michael Krahulec, and Nathaniel Meierpolys who pursued various methods for calibrating cameras in 2008.

This previous group mounted the camera, installed drivers for Video4Linux2 (V4L2) onto our computer, and ran cables from that computer to the camera. Further, we were able to adapt code from the earlier group's project, which was similar to examples in the V4L2 API.

We drew inspiration from existing papers on image stitching, and projections of spheres onto flat surfaces.

3. Creation of a satisfactory panoramic image

3.0 Overview

A panoramic image creation system will generally have three main pieces of hardware: a pan-tilt camera, a capture card to receive the video stream, and a computer to process the data from the capture card's buffer into composite images. We identified four critical sections in which design choice and equipment features can influence the course of a panoramic image project. These four areas are: the optical quality of the camera's lens, the movement precision of the camera's motors, the features of the capture card, and the image processing used to combine the raw images into panoramic image files.

3.1 Camera optics

An ideal camera has a high level of accuracy in its lens and color sensors, and records high-resolution images. While the Sony EVI-HD1 captures two million pixels and the BRC-300 only three hundred thousand, the amount of data recorded does not tell the full story. The overall image quality of the camera can be better measured by identifying distortion to the shape and color of the image caused by the camera lens.

One way to study the optical distortions in a camera lens is to photograph a "chessboard" of alternating black and white squares, held perpendicular to the axis of the camera's view (Fig. 1.). When the Sony BRC-300 camera views this set of straight lines, the lines near the edges appear to bulge out, away from the center of the image. This bowing indicates that this camera suffers from "barrel" distortion, which can create poor quality panoramas (as seen in Fig. 2.). If the lines bowed inward, the camera would instead be said to suffer from "pincushion" distortion. In general, "barrel" distortion will be less noticeable at greater zoom levels. Applications demanding low distortion may benefit from sampling an image at high zoom, or from the application of distortion correction in software. However, the best choice is to choose a camera such as the EVI-HD1 which exhibits less pronounced spatial distortion.

The quality of color in an image can also be affected by chromatic aberration. Chromatic aberration occurs when different wave-lenghts of light bend less or more as they pass through the camera's lens. Frequently this is seen in a red "fringe" along one side of a high-contrast region, and a symmetrical blue "fringe" on the other side. This undesirable effect can be seen in Fig. 3. In images from the EVI-HD1 camera, we did not find evidence of high levels of chromatic aberration.

The lens and zoom level also affect the number of pixels in an image that correspond to a given field of view in degrees, a quantity essential for creation of a panoramic image. Knowledge of the field of view contained by an image is crucial in determining whether the pixels of the image lie along a given ray from the center of the camera. In search of this data, some of our colleagues whose studies focus on analysis of the geometric precision of camera images computed the fields of view of each of the cameras using their own highly accurate methods (Magee et al, 2012). This information, recorded for both cameras at several zoom levels, was recorded and used in the image stitching process.

3.2 Camera movement

Both the BRC-300 and EVI-HD1 support Sony's VIdeo System Control Architecture (VISCA) interface. VISCA is a serial connection, based on the RS232 protocol, which allows various commands sent by the computer to control the movement of the camera. VISCA only controls camera movement, and does not control image capture. Our system used adaptor cables to convert Sony's 8-pin mini-DIN VISCA connector into the 9-pin serial standard, to the computer via a Keyspan USB-to-serial converter. Software control was based on libvisca, a freeware open-source C++ library written by Damien Douxchamps. The libvisca API manages the sending and receiving of RS232 packets, and freed us to focus on the creation of software structures to manage the cameras. Each Sony camera has a different set of VISCA commands, so careful attention to Sony's white-papers is crucial. The end result was a control system which gave good performance and simple functions to control power, pan, tilt, zoom, shutter, iris, and gain in both cameras.

The motors that move the camera provide the final component of accuracy in the capture of images. Cameras must demonstrate acceptable performance in the resolution, accuracy, and repeatability of their motion. In this respect, the BRC-300 offers a much greater resolution for movement, which can be expressed as the number of discreet data values into which one degree of motion is subdivided. The BRC-300 uses values from -35416 to +35416 to describe pan positions from -170 to +170 degrees, while the EVI-HD1 uses a range of -1440 to +1440 over a pan range of -100 to +100 degrees, making the former a much more precise instrument.

To more easily visualize accuracy of motion we attached a laser pointer to the BRC-300. We then moved the camera to a given pan and tilt, moved it to a variety of new points, and from each moved back to its original location. We compared the location of the laser from the original and final locations. This experiment revealed that at a range of 10meters the camera could return to a previous location with less than two centimeters of error. This means that if we taped a nickel to a wall at 10 meters, and centered the camera on the nickel, we could arbitrarily return to those coordinates and be certain that the laser dot would land again on the nickel (see Fig. 4). This computes to an accuracy of +/- 0.05 degrees, which is below two percent of the width of an image at a zoom ratio of 10x on BRC-300 camera.

Although the camera is able to return to a position, we needed to perform another test to investigate whether the camera actually turned the amount we desired. To discover this we tilted the camera to its polar coordinates. We then centered a CD under the camera and marked the middle of the CD. This meant that when the camera paned, the mark in the center of the CD rotated, but didn't appear to move. This was key to ruling out the possibilities that the center of rotation might not be aligned with the camera's center, or that the camera might not tilt down to exactly 90 degrees (Figure 5). Although the positive and negative angles don't appear to have a common focus, they do show perfect angular accuracy. The only places in which the cameras do not more as much as they claim is at +/-170 and 0. These areas are edge cases which do not fit the function to convert VISCA signals into degrees. Considering that these inaccuracies are most likely due to our program rather than a flaw in the camera, we conclude that the BRC-300 is highly accurate in moving to the pan location demanded of it.

3.3 Capturing data

Data from the camera must be saved as image files on the computer, and this bridge of video formats is accomplished by the video capture card. The BRC-300 camera was already configured with a generic brand card designed for S-Video, an analog format, while the EVI-HD1 used the High Definition Serial Digital Interface (HD-SDI) and a specialized capture card made by Blackmagic Design, a professional video hardware manufacturer.

The generic capture card exhibited several undesirable characteristics. This card was unable to capture accurate data for a significant amount of time after the start of capture, with the resulting first few images from a capture stream having over-bright luminance and distorted color information. The generic card also used a non-standard byte order when it stored data in memory, reversing the green and blue information in a standard RGB image, which required customization of our software in order to correctly save the images to disk. In contrast, the Blackmagic card supported multiple video encodings, and greater bit depth (ten bits per channel instead of eight) for higher quality.

Both cards use specialized APIs for developing programs to interact with the data they obtain from the cameras. Blackmagic provided excellent developer support and a very useful Software Developer's Kit (SDK) for public use, while the generic card was accessible using the open-source framework Video 4 Linux 2 (V4L2). Future investigators may be unwise to rely too heavily on V4L2 because development stopped in 2008.

After accessing the desired image information from the capture card's buffer we chose to store the information in PPM files, a standard image format which stores three-channel RGB information in plaintext format with minimal overhead. We created a new file type, .pano, to store information regarding the spatial location of the captured image and its This file contained pertinent data about each image such as the file path, the center of the image in terms of latitudinal degrees (phi) and longitudinal degrees (theta) as well as the width of the image in phi and theta.

3.4 Creation of panoramic images

The process of combining multiple discreet images into one panoramic view, known as compositing, was the final step in our workflow. Following techniques of Richard Szeliski, we were able to create and iteratively improve a compositing program which allows the extraction of data from our .pano files over any section of the observed field.

Each pixel in the destination image is accessed based on its angles of reference -- in effect extracting the information along a certain vector from the origin in the center of the sphere of images captured from the camera. Our method offers an advantage in that the source images are already well-aligned due to accurate information about the angle of view at which they were taken. In addition, the accuracy of the camera servomotors means that pixels within the image can be referenced to the know ray which points from the image center of the camera to the center of the image. Requested pixels are computed by blending the relevant pixels from every source image which lies along the query ray, which allows for interpolation between pixel values. This means that images can be extracted at any resolution desired. Further, storing source data in standalone .ppm files offers an additional advantage because source data is preserved permitting new or improved stitching algorithms to be applied to old data.

Because most errors in image stitching are readily apparent to an observer, evaluation of a method's quality is a relatively simple process. Our methods encountered three errors as they blended source images, namely ghosting from motion of objects within the scene, visible seams on the edges between source images, and blurred sections due to imperfect alignment of source images. An averaging method, which renders each output pixel as the mean value of all contributing source pixels, is simple to implement but results in visible seams, much ghosting, and heavily blurred sections of the image where source images do not line up perfectly. Using an image mask greatly improves the result by reducing the weight of pixels near the edge of source images so that they contribute less to the final image.

4. Conclusions

Many factors of hardware and software must be considered in the design of a pan-tilt-panoramic capture system. Our creation of two end-to-end panoramic systems, based around the Sony BRC-300 and EVI-HD1 cameras, allowed us to compare the most and least desirable aspects of each system. Although the EVI-HD1 offers excellent image quality, its shortcomings in pan and tilt resolution indicate that an ideal pan-tilt camera would take the best of each of the two cameras we considered. Such a camera would incorporate a sensor and lens like that of the

EVI-HD1 but also a pan-tilt platform with highly accurate servomotors with a wide range of motion and a fully digital data path such as HD-SDI.

However, we recognize that such an ideal device may not exist. Many developers of panoramic systems may be forced to choose a camera based on budgetary concerns. We therefore recommend that developers prioritize the features of accurate, precise, and wide-ranging pan and tilt in their selection of a camera, because these features cannot be corrected for in software if they are lost in hardware.

5. Acknowlegements

We would like to thank our Professor, Olaf Hall-Holt, who guided us through our moments of doubt and pain. Devin Lackie, for welcoming us into his stockroom, assisting us with technical issues, and always being ready with a relevant story to drive home his advice. Dr. David Nitz, for lending us his personal laserpointer when we couldn't find one anywhere else. Blackmagic Tech Support, for providing us with personalized technical assistance and technical specs to ensure that we could use the full power of their capture cards. Dana Thompson, who found a conversion cable for our cameras when we had no other option.

6. References

Exposure Fusion Tom Mertens, Jan Kautz, Frank Van Reeth.
Image Alignment and Stitching: A Tutorial Richard Szeliski.
Refinement of Plane Based Calibration though Enhanced Precision of Feature Detection Rogan Magee, Leah Roth, Jared Brown.

7. Appendix

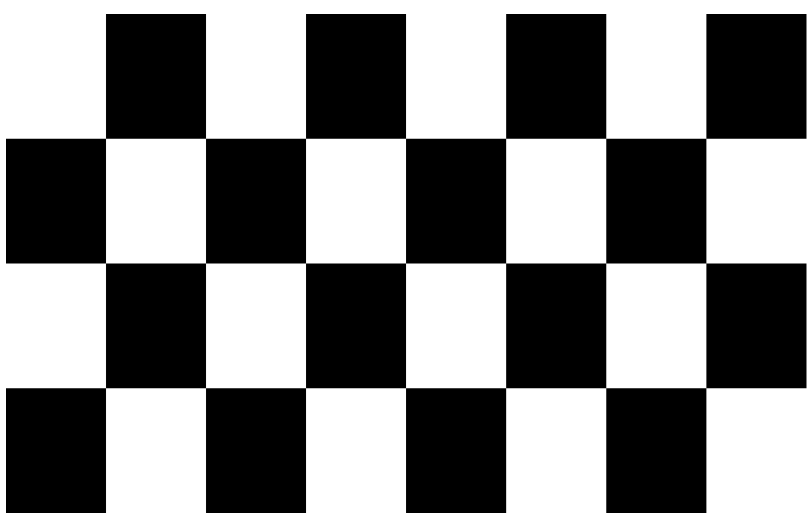

Figure 1: the source image at half size.

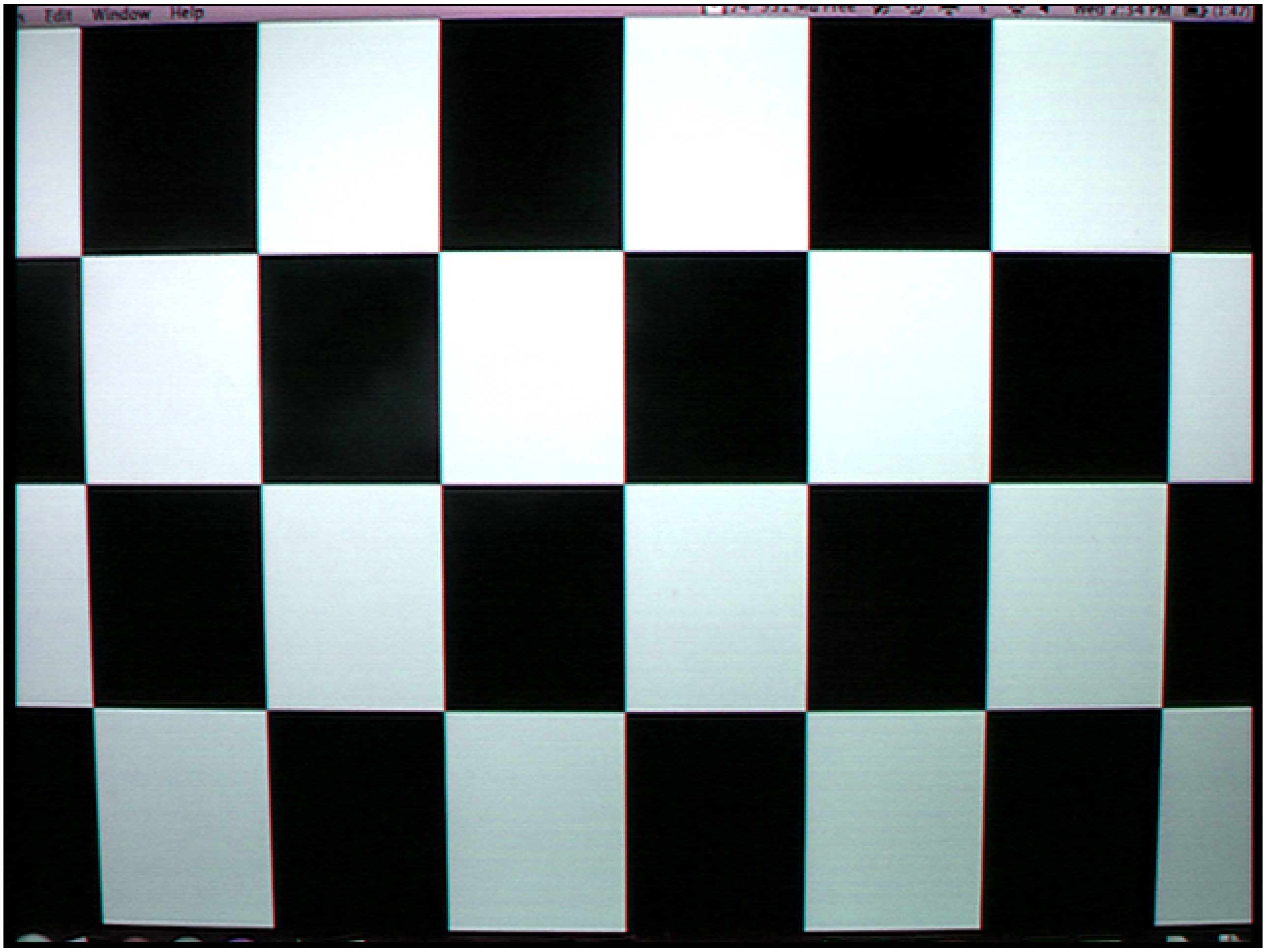

Figure 2: A capture from the BRC-300 camera, exhibiting strong barrel distortion.

Figure 3: Horizontal and Vertical chromatic aberration on a chessboard, visible at 4 times the image's original size.

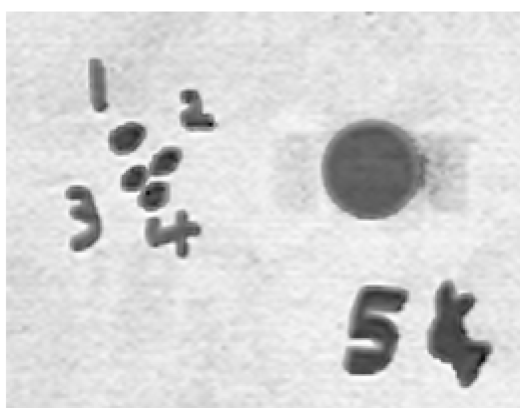

Figure 4: A fully zoomed and enlarged image showing that after seeking and returning 20 times the camera returned to one of four points 10 meters away. As can be seen from the nickel on the wall next to the points all of these points could be covered with a single nickel, indicating high repeatability.

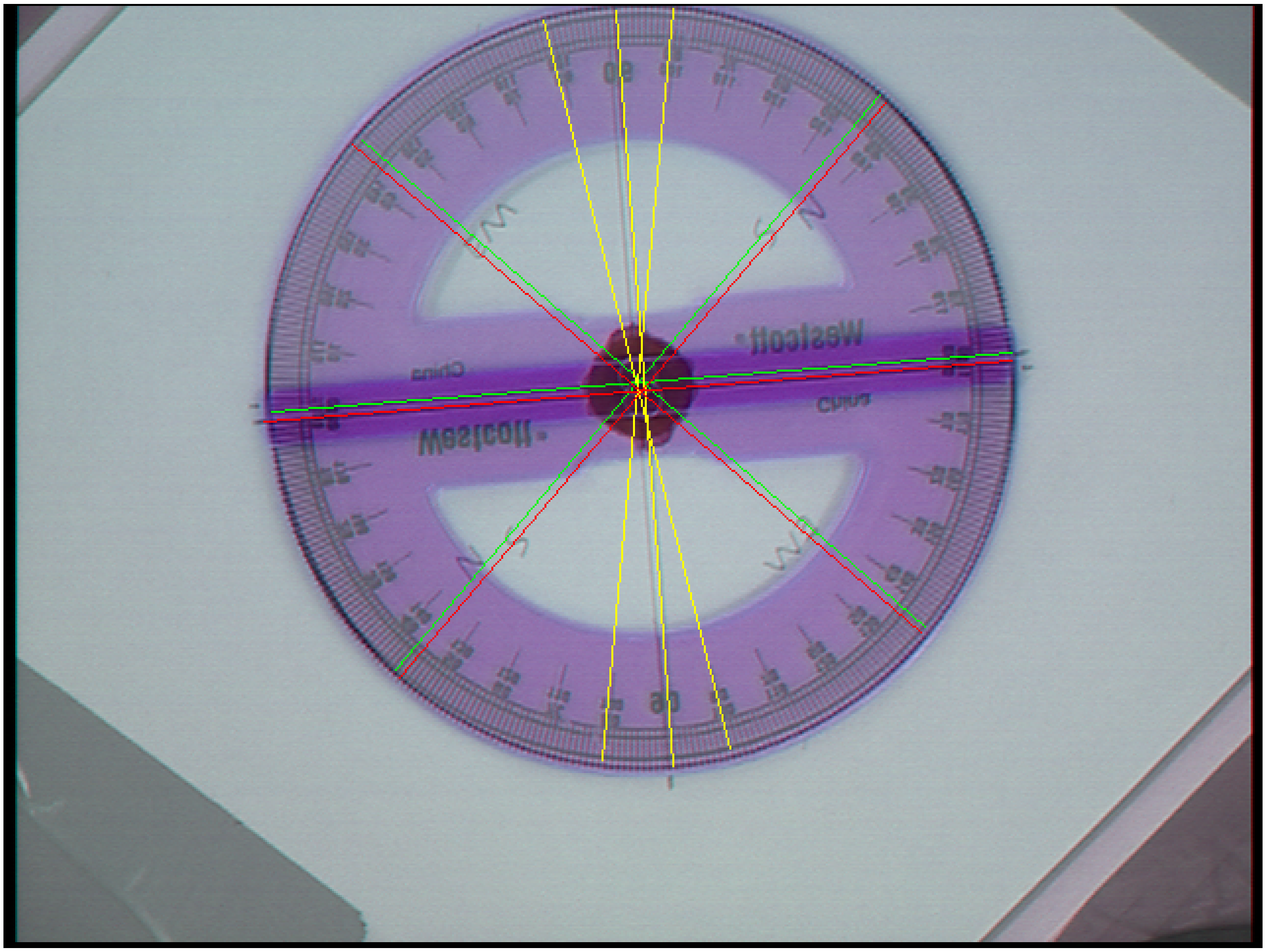

Figure 5: Precision of pan operations on the Sony BRC-300. Our function to control rotation breaks down in the edge cases of 0, +170 and -170 as seen in yellow. However, because the red lines and green lines intersect and perfect 45 degree angles the camera has excellent accuracy in rotating the amount desired by the user.

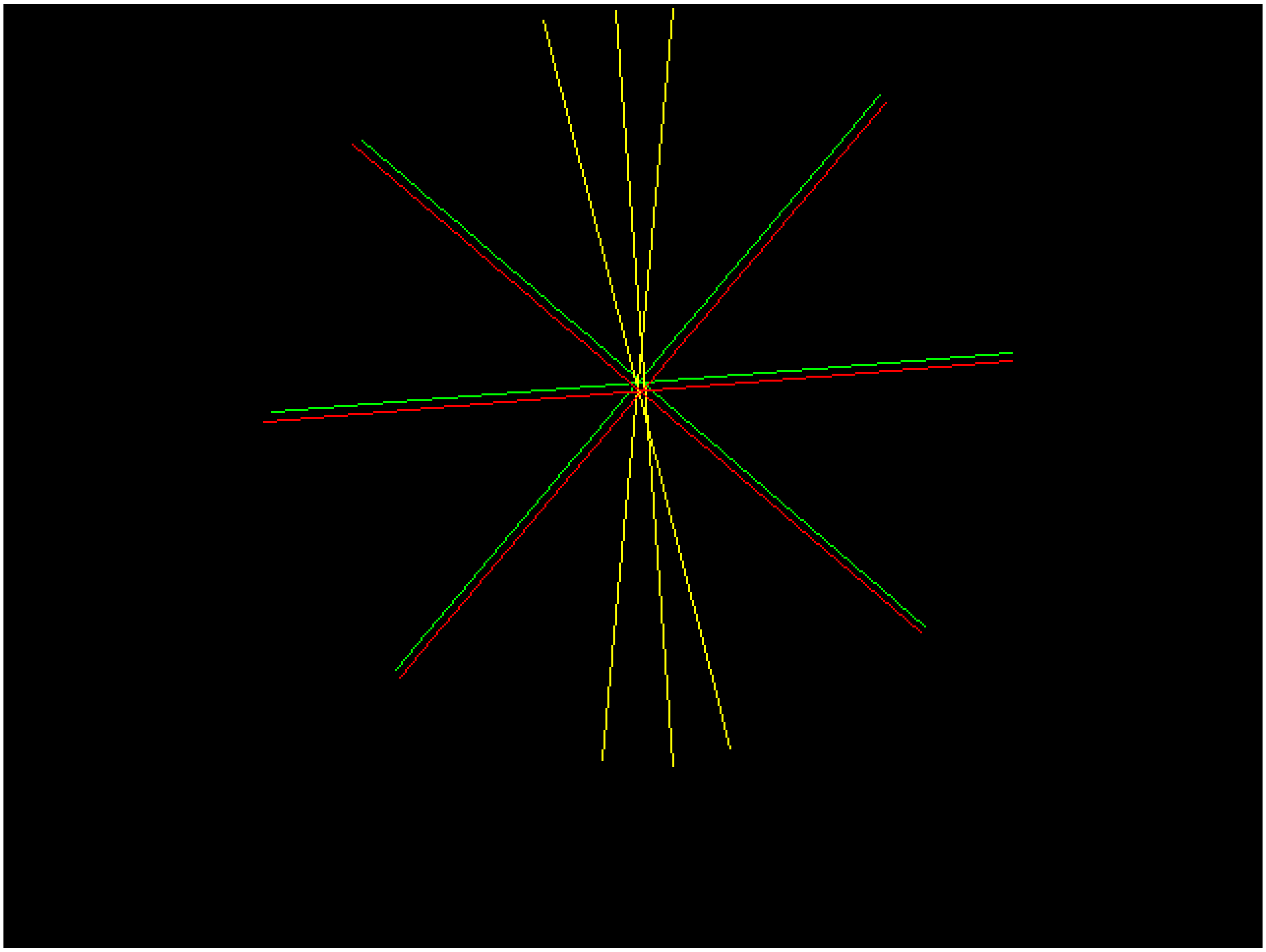

Figure 6: A copy of Figure 5, simplified to show the perfect 45 degree angles in the positive(green) and negative (red) directions.