Autonomous Drone Follow

Unmanned Aerial Vehicles

Task

- Accurate depth calculation
- Minimum Computation
- Maximum FPS

Literature Survey

Measuring Distance with Mobile Phones Using Single-Camera Stereo Vision

Abstract—Computer stereo vision is an important technique for robotic navigation and other mobile scenarios where depth perception is needed, but it usually requires two cameras with a known horizontal displacement. In this paper, we present a solution for mobile devices with just one camera, which is a first step towards making computer stereo vision available to a wide range of devices that are not equipped with stereo cameras. We have built a prototype using a state-of-the-art mobile phone, which has to be manually displaced in order to record images from different lines of sight. Since the displacement between the two images is not known in advance, it is measured using the phone's inertial sensors. We evaluated the accuracy of our single-camera approach by performing distance calculations to everyday objects in different indoor and outdoor scenarios, and compared the results with that of a stereo camera phone. As a main advantage of a single moving camera is the possibility to vary its relative position between taking the two pictures, we investigated the effect of different camera displacements on the accuracy of distance measurements.

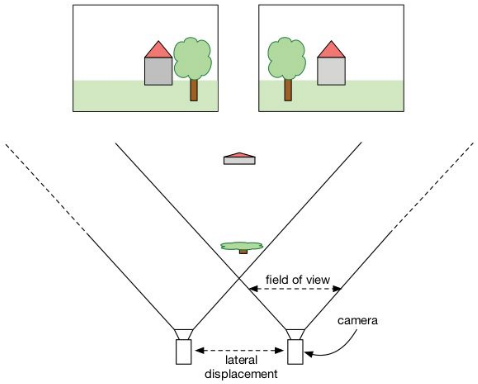
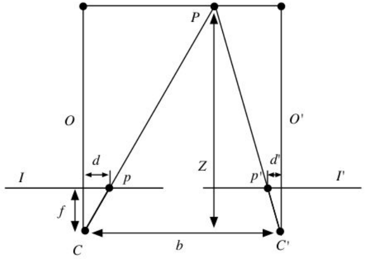
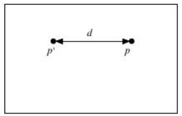


Figure 1. Two images recorded from a laterally displaced camera. The displacement of the objects provides relative depth information.



(a) Geometry of stereo triangulation.



(b) Overlay of the images shows the difference between the image locations p and p'.

Maths Used

Robert Collins
CSE486, Penn State

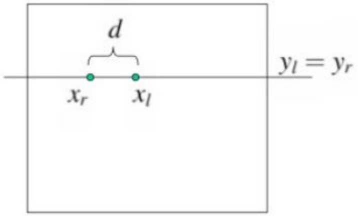
Recall: Stereo Disparity

Left camera

$$x_l = f\frac{X}{Z} \qquad y_l = f\frac{Y}{Z}$$

Right camera

$$x_r = f\frac{X - T_x}{Z} \qquad y_r = f\frac{Y}{Z}$$



Stereo Disparity

$$d = x_l - x_r = f\frac{X}{Z} - \left(f\frac{X}{Z} - f\frac{T_x}{Z}\right)$$

$$d = \frac{f\,T_x}{Z}$$

depth baseline

$$Z = \frac{f\,T_x}{d}$$

disparity

Important equation!

Triangle Similarity for Object/Marker to Camera Distance

In order to determine the distance from our camera to a known object or marker, we are going to utilize *triangle similarity*.

As I continue to move my camera both closer and farther away from the object/marker, I can apply the triangle similarity to determine the distance of the object to the camera:

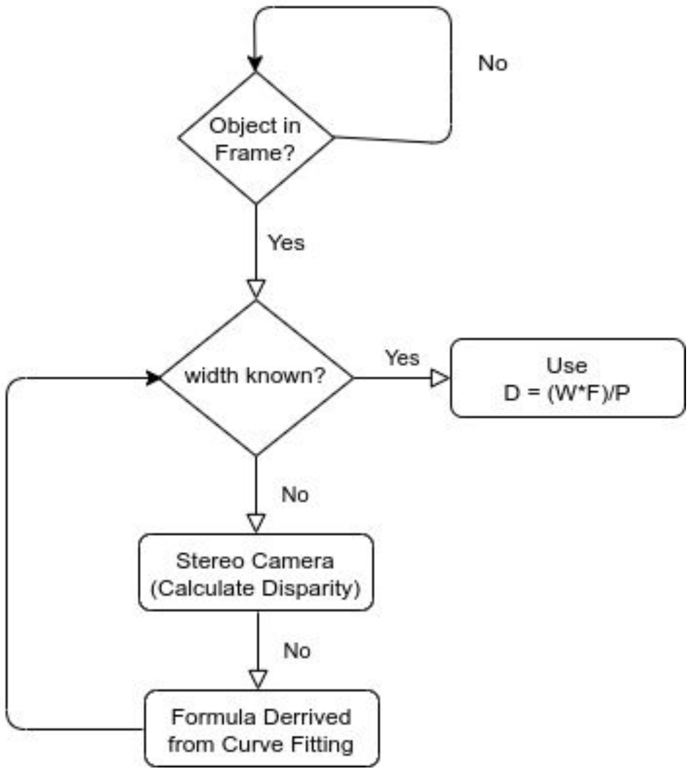
D' = (W x F) / P

1.1 What is the range of the Depth measured by the camera?

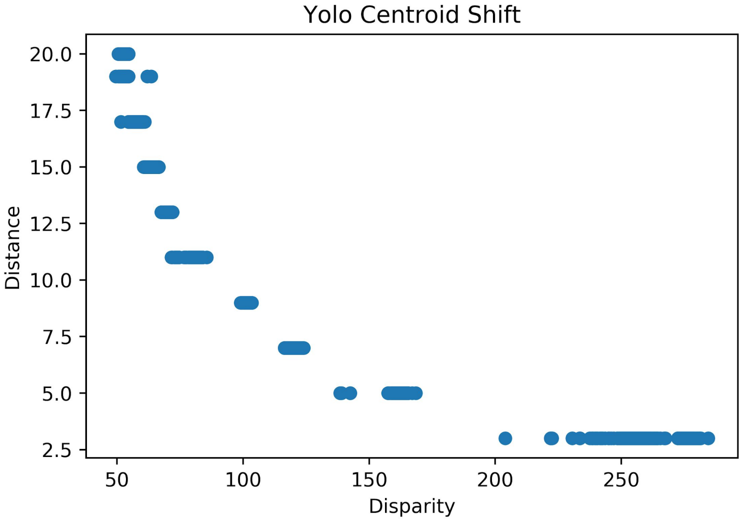
Depth Range is from **50 cm to 300cm**. The maximum working distance varies with the lens. With our default lens, the maximum working distance is about 300cm, beyond which the depth accuracy drops down

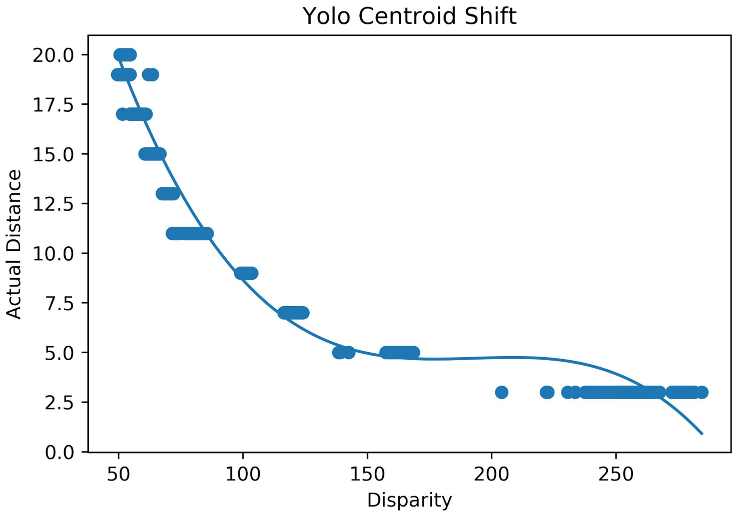
Approach

- Object detection with single camera
- Stereo vision camera for depth and width calculation
- Size based depth calculation (if object width known)

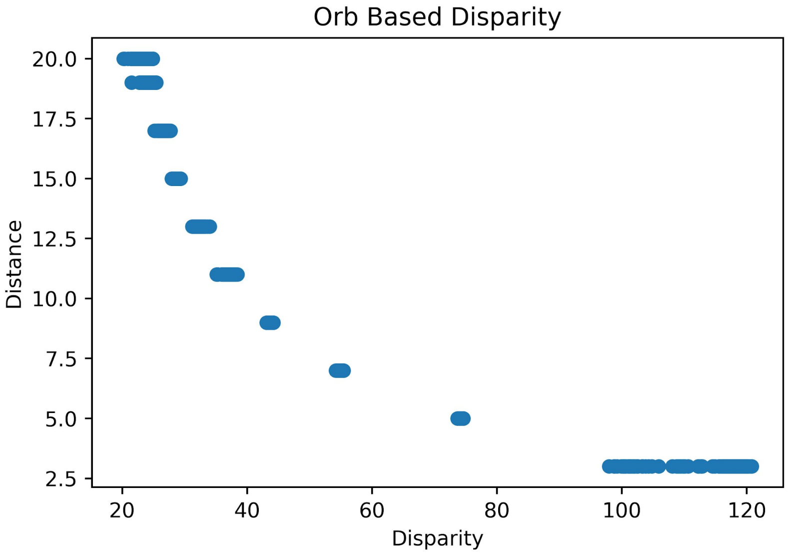


Results

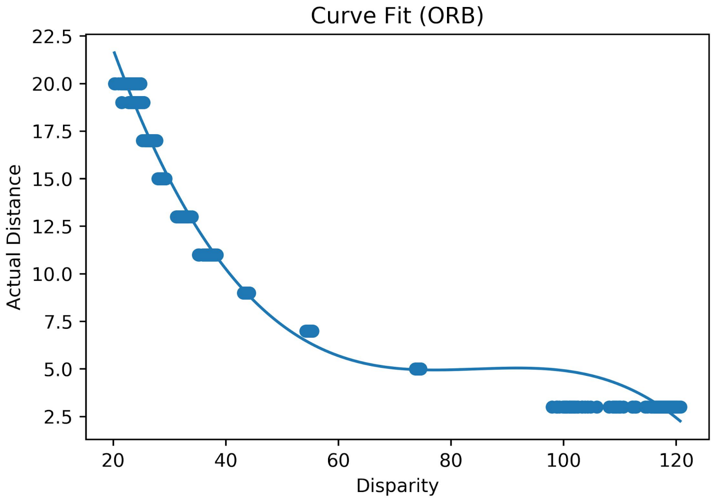

Yolo Centroid Shift
Distance
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
50
100
150
200
250
Disparity


Yolo Centroid Shift
Actual Distance
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
0.0
50
100
150
200
250
Disparity

Results



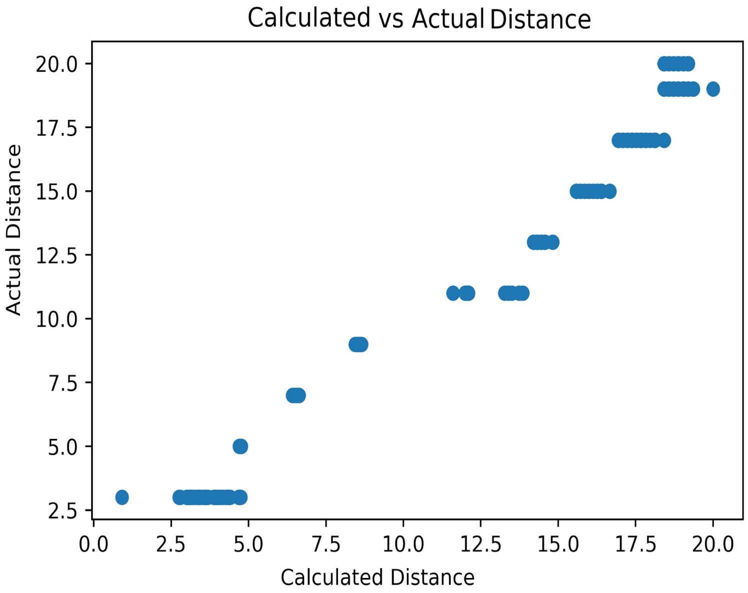
Orb Based Disparity
Distance
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
20
40
60
80
100
120
Disparity



Curve Fit (ORB)
Actual Distance
22.5
20.0
17.5
15.0
12.5
10.0
7.5
5.0
2.5
20
40
60
80
100
120
Disparity

Comparison

Yolo Shift



ORB Disparity