ECPE 293 Fruit Classification

Vivek Maheshwari

Abstract— The goal of this script is to identify and classify different fruits in picture. For this task the image is separated into foreground and background. Different fruits are treated as separate components, classified, and highlighted. For classification we use statistics of object or fruit which includes area and eccentricity etc.

I. Classification Algorithm

A. Handling Input

The process begins with a user selecting an image from their local device. The image used during development, Fig 1., is saved as the "Actual Image".

B. Foreground Separation

Separating the foreground and background layer from an image is a multistep process. Tried Utilizing Ridler Calvard for threshold value. However, it after some experimentation it was determined that these conveyor belt images of fruits use a 100 and 180 as threshold value for the foreground and background respectively. Then thresholded image goes through some rounds of erosion and dilation to get rid of any noise. This final image, Fig 2., is used for the Identification step

C. Component Identification

The goal of this step is to give discrete labels to different objects or components in the image. This process is carried using connected components with flood fill. It can be done using union find. While both methods mentioned above are available, the connected components are used in the figures throughout this report. Different components in the image are treated discrete with different labels. Connected Components image, Fig. 3.

D. Component Analysis & Classification

The process of identification and classification was two-step process. In the first step we calculate moments, central moments, eigenvalues, centroid, eccentricity and direction for a particular fruit or region. All this data is essential for classification as eigenvalues and direction help in calculating the major and minor axis of a fruit. Once these values are calculated we use unique factors for objects or fruits for classification. Experimentation determined that eccentricity greater than 0.5 and area greater than 3000 pixels is banana, whereas area greater than 5000 is orange, area greater than 2500 and eccentricity less than 0.5 is apple, and everything else is a stem. Obviously, this process is limited to camera, angle, picture, clean, and type of fruit. Fig.6 shows the tabular data for different fruits or objects.

E. Subcomponent Identification & Analysis

During the component analysis and classification phase, the stems of bananas were treated as a subcomponent. After confirming an object as a banana, a copy of the connected components image is subjected to erosion followed by dilation, effectively removing the stem. The image without stem is then compared against connected components image Any pixels that were part of the banana in the connected components image, but not in the stem_removed_image were candidates for the stem. Based on experimentation the pixel at furthest distance to the center of banana is the pixel of the stem because stem is at max distance from the center of the fruit in this case banana. Then we use that seed pixel and calculate nearest pixels distances and only account for those that fall under 40 pixels. That way pixels nearest to the seed pixel become part of the stem.This process is repeated for each detected banana. Overall, the classification methodically worked well for both our development image, fruit1.bmp and fruit2.bmp. Stem area was found to be between 250 and 300. The final images for both can be seen in Fig. 4 and Fig. 5.

F. Perimeter Detection & Presentation

We then apply wall following algorithm on final connected components image and store boundary or perimeter for a particular region or component in a list which is used to change the color of border pixels. Different components and subcomponents will have different colors, and all of this done in conjunction with plotting of the axes over each component, Fig. 4 and Fig.5 display the final output over different fruit images.

II. Results

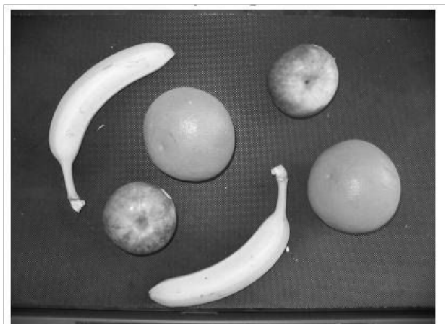

Fig 1. Input image(fruit1.bmp)

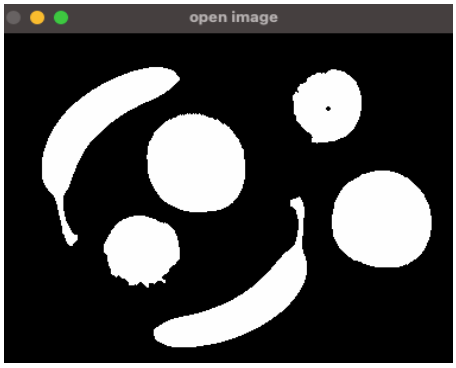

Fig 2. Final Thresholded Image

Fig 3.Connected Components Image

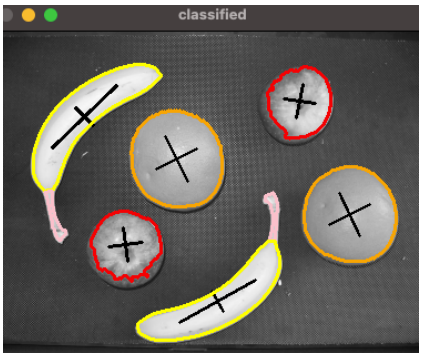

Fig 4. Output Image

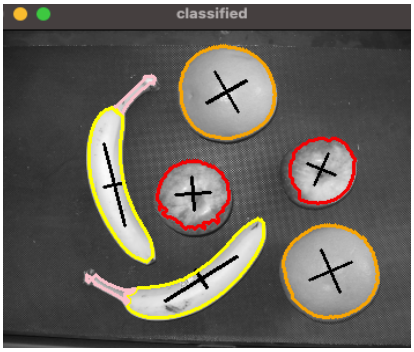

Fig 5. Fruit2.bmp Output Image

Fruit	Area	Eccentricity	Eigen 1	Eigen 2
Top Banana	4628	0.92	1649	133
Top Apple	2867	0.455	259	205
Middle Orange	5852	0.47	510	425
Right Orange	5785	0.37	479	422
Middle Apple	3026	0.255	248	235
Bottom Banana	4393	0.91	1572	115

Fig 6. Descriptive statistics for fruit1.bmp components