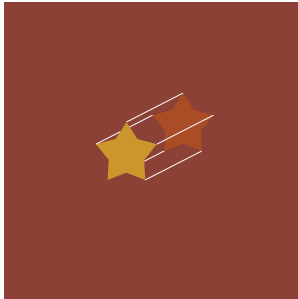We have an exaggerated example of three images where the object appears shifted

The program calculates the shift of the object...

and places the image in a bigger frame, with the shift corrected

Then, all the images are combined by the median value of each pixel