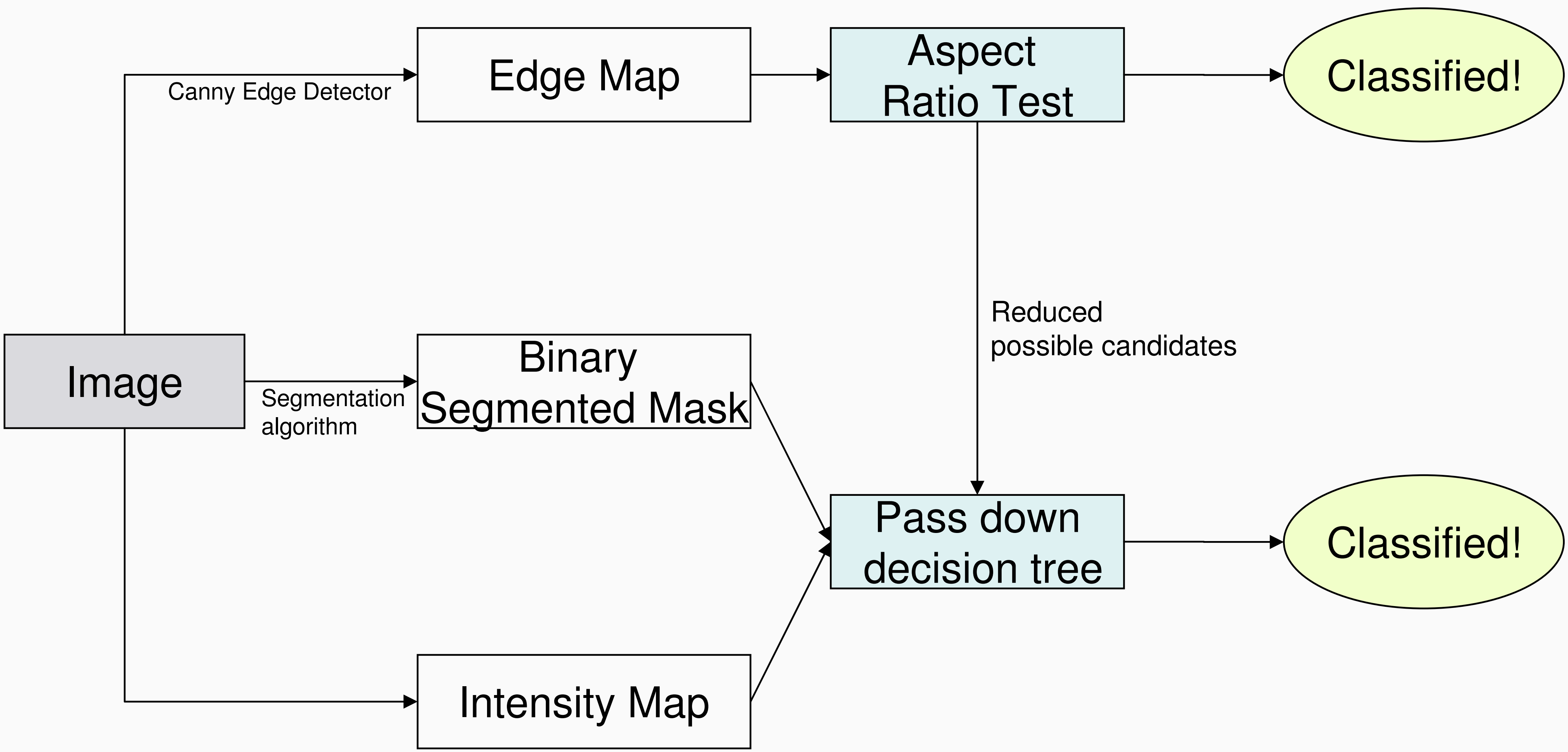
Object Recognition

Lin Qiuyu
Chow Ying Yi
He Congfu Joshua

System Overview



Segmentation

Purpose of Segmentation:

1. Create a binary segmentation mask of the input image. This is used in the Appearance Classifier as one of the features.
2. Used to group detected keypoints together as belonging to one connected component. (Object)
3. Increases the robustness of the system as keypoint-outliers are disregarded by the system.

Combination of 2 Algorithms is used:

1. Patch based Hue and Hue Variation Classifier
2. Pyramid Segmentation

Segmentation

Patch based Hue and Hue Variation Classifier

To distinguish between background and objects:

1. Background is divided up into patches of 11 x 11 pixels
2. The mean hue of the patch is extracted and tested if it is within a trained threshold value
3. If Hue of the patch exceeds a trained amount of threshold variation, the patch is disregarded and considered an object.

Segmentation

Patch Hue Classifier deciding on what regions are background

- Green: Background
- Blue: Hue test fails
- Red: Hue Variation too high

Observation:
Method is relatively accurate by with many false negatives.

Segmentation

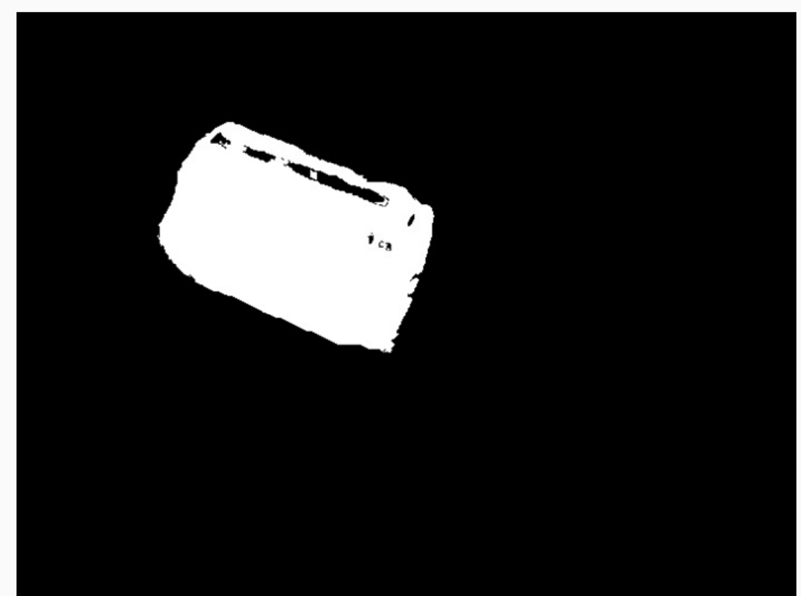
Pyramid Segmentation

Pyramid Segmentation is done on the image revealing connected regions of the image.

Observation: Good segmentation of Image, However, Background was segmented into 3 segments.

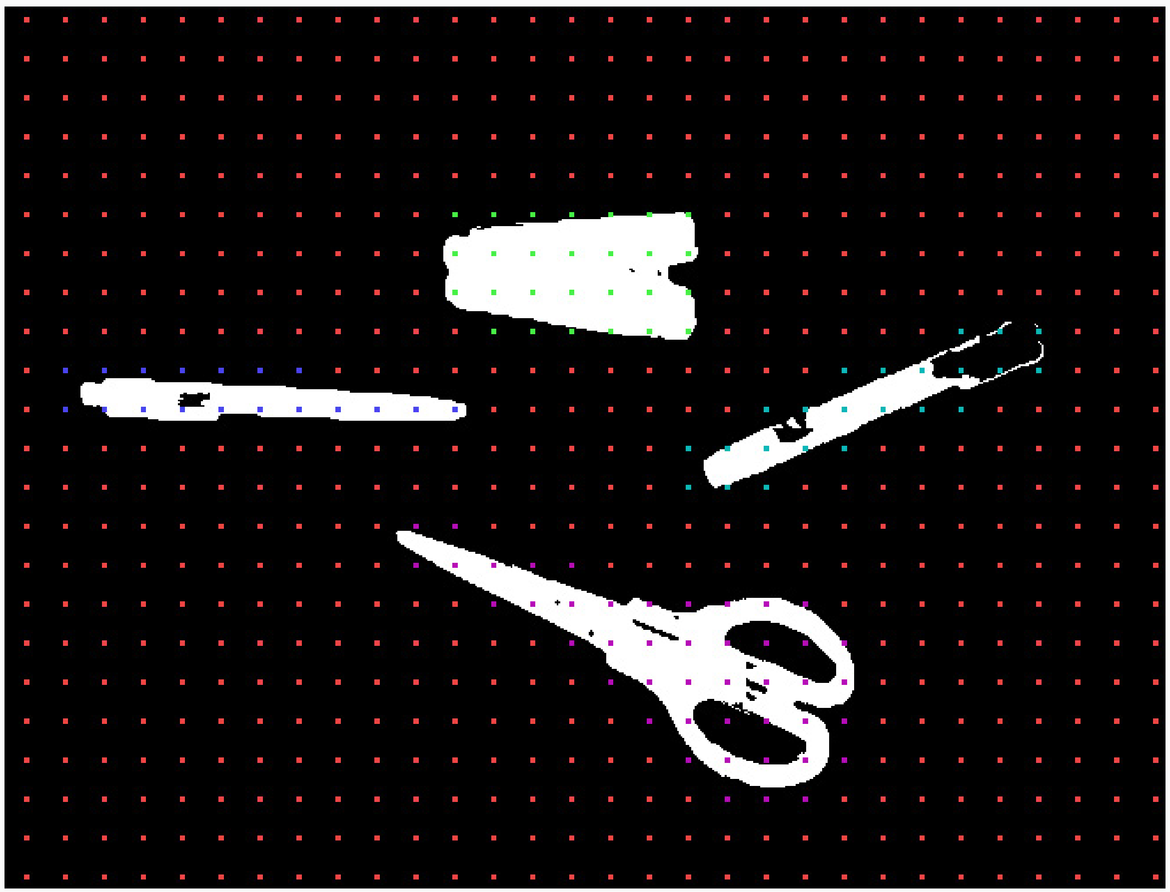
Combining 2 Algorithms

By Combining the 2 algorithms using simple voting of patches, a good binary Segmentation mask is achieved.



Segmentation

Multiple Objects Labeling



Using a connected component search on the segmentation mask created, connected regions are labeled. This is used to group keypoints together so each group of keypoints are evaluated independently.

Left: Connected regions are labeled the same colour, purple, blue, cyan and green. Red refers to background.

Aspect Ratio Classifier

Observe:

- Some objects have different aspect ratio between them.

E.g. pen/ruler vs. can/handphone

Using this, we can increase the accuracy and speed of detection by reducing the number of possible candidates.

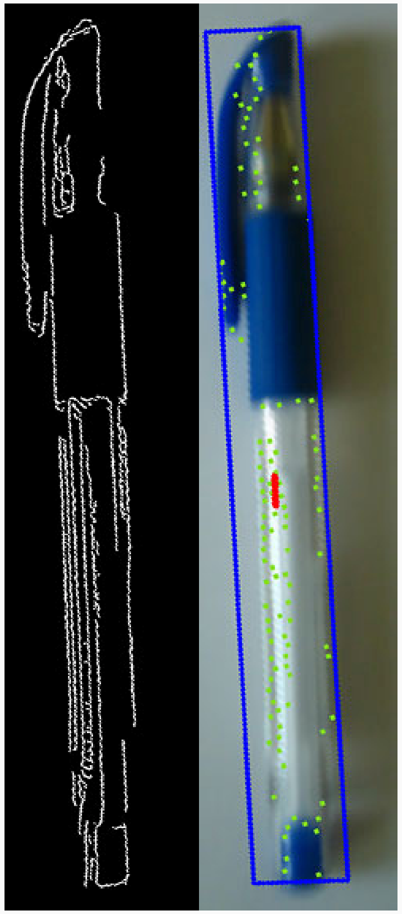
Aspect Ratio Classifier

Compute ratio (width/height) of the object's bounding box.

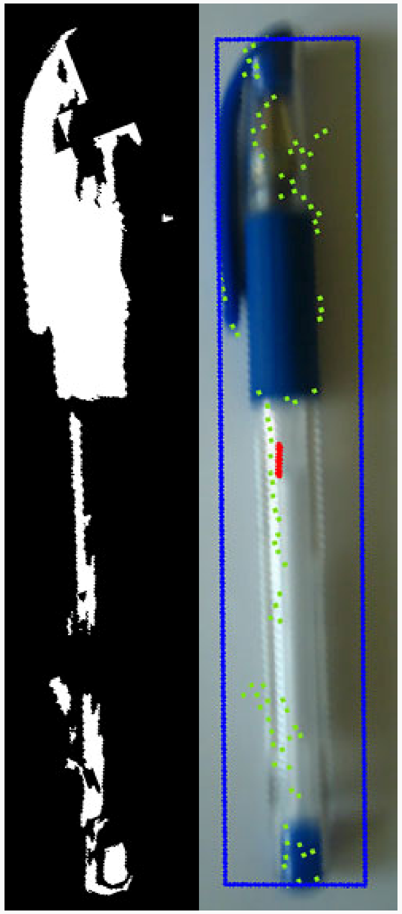
- Canny edge detection to obtain edge map
- Tomasi Algorithm on Canny Image
- Use PCA to get object orientation

Reasons for using Canny edge map over using segmentation mask is because the latter sometimes cannot show a complete binary image of the object

As seen from the image on the right, the bounding box obtained using keypoints detected on edge map is more well-defined than that using the segmentation mask.



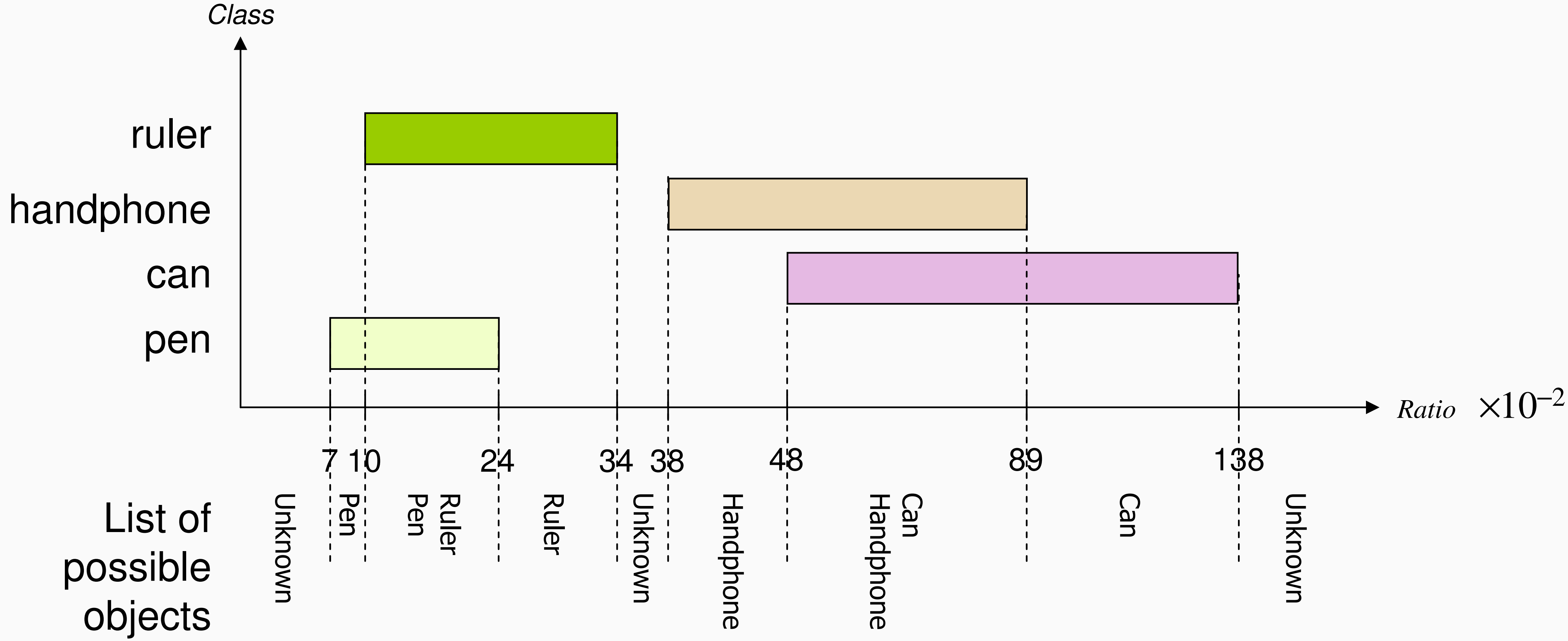
Use Edge Map



Use Seg. Mask

Aspect Ratio Classifier

Compare ratios to obtain list of possible objects.



Aspect Ratio Classifier

If possible candidates only has 1 object, it is immediately classified as that object.

Otherwise,

$$P_{aspect}(l,k)=\begin{cases}1 & \text{if keypoint } k \text{ is possible to belong to class } l\\ 0 & \text{if keypoint } k \text{ is unlikely to belong to class } l\end{cases}$$

The probability P_{aspect} is passed to the decision tree in the next step of our algorithm.

Appearance Classifier - Decision Tree

The appearance Classifier uses 2 Features

1. Grayscale Intensity patches of 11x11 pixels selected after keypoint detection from Tomasi Interest Point detector.
   - Gives accuracy when tested on trained objects

2. Binary Segmentation patches of 11x11 pixels that give the localised "shape" of the object at that point.
   - Increase generality, especially when detecting objects which shape remains constant by appearance changes e.g.. Cans with different patterns.

Appearance Classifier - Decision Tree

1. Patch-based approach used for appearance classifier
   - Increase generality
   - Consistent even when the object is partially occluded by foreign unknown objects (e.g. hand)
2. Decision tree built to obtain the probability of a patch being in a class *c*
3. Multiple decision trees used
   - Resolve the problem of intractability when the number of training images and classes increases
   - Increase generality by avoiding the training of details only specific to the trained images.

Appearance Classifier - Decision Tree

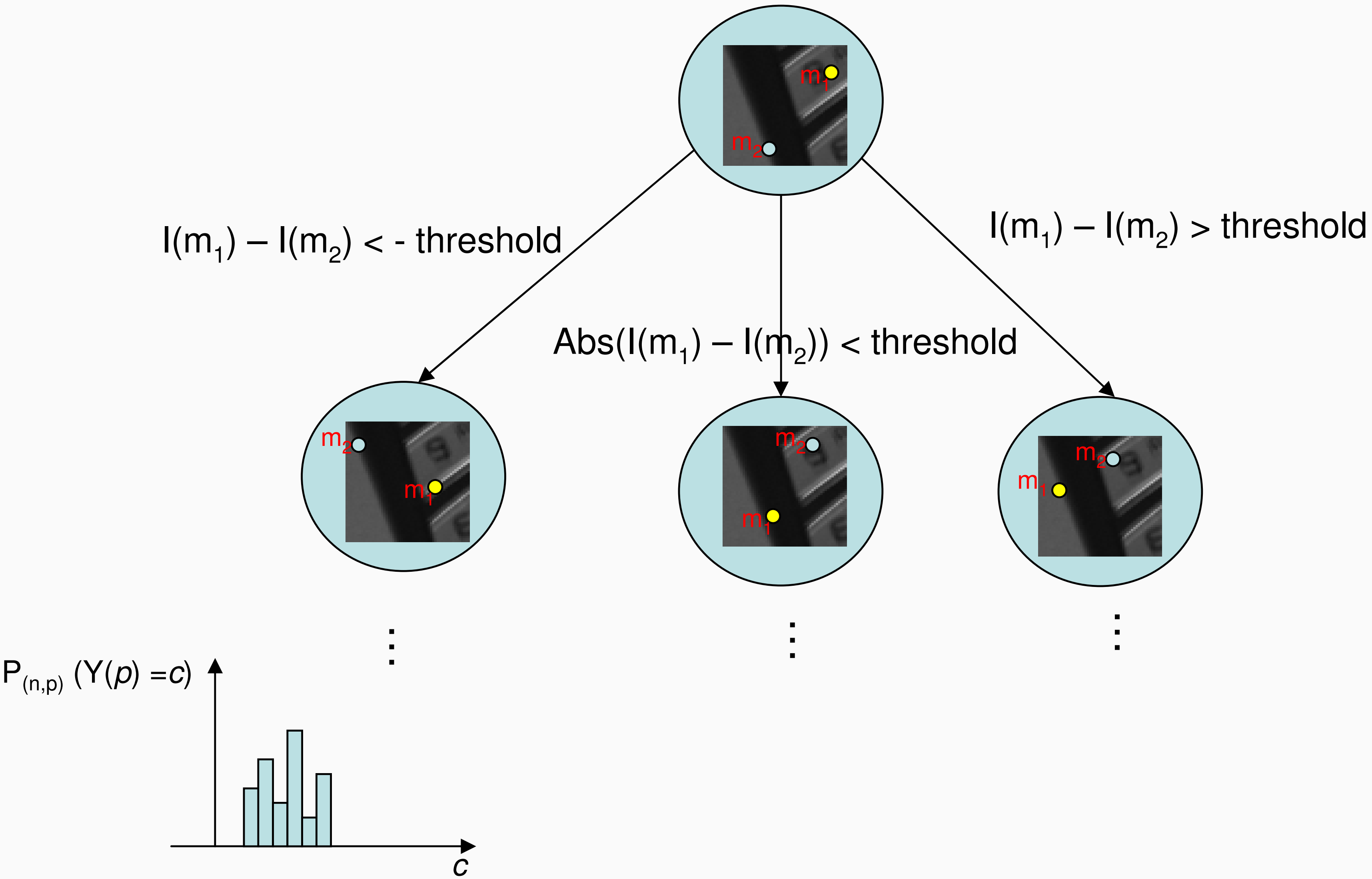
1. Decision tree nodes contain tests comparing intensity difference between two pixels in the keypoint neighborhood (patch)
2. The leaves contain the posterior probabilities over the different classes :

$$P_{\eta(l,\ p)}\ (Y(p) \ = c), c \in \{c_1, \ldots c_l, \ldots\}$$

Where

c is the object class,
l is the classes reduced by aspect ratio test
n is the tree number and p is the patch

Structure of Decision Tree



Appearance Classifier

Decision Tree Training

1. Random Affine transformations are applied to images
   $I' = R_{\theta} R_{\Phi}^{-1} S R_{\Phi}(I) + T$
2. Tomasi algorithm to detect keypoints on the original image
3. Patches extracted accordingly on each keypoint and rotated using the PCA so that the intensity change direction is the same for all patches (as shown in the image below).

Appearance Classifier

Decision Tree Training

4. Randomly choose shape or appearance test at each node
5. Use binary mask for shape test and gray scale image for appearance test to compare the pixel intensity
6. Choose 2 pixels in a patch and according to the intensity difference, split the tree using the standard greedy entropy based algorithm

Appearance Classifier

Decision Tree Testing

1. Tomasi Algorithm is run on testing images to detect interest Points
2. Patches of constant size are extracted and PCA is used to orientate extracted patches

Appearance Classifier

Decision Tree Testing

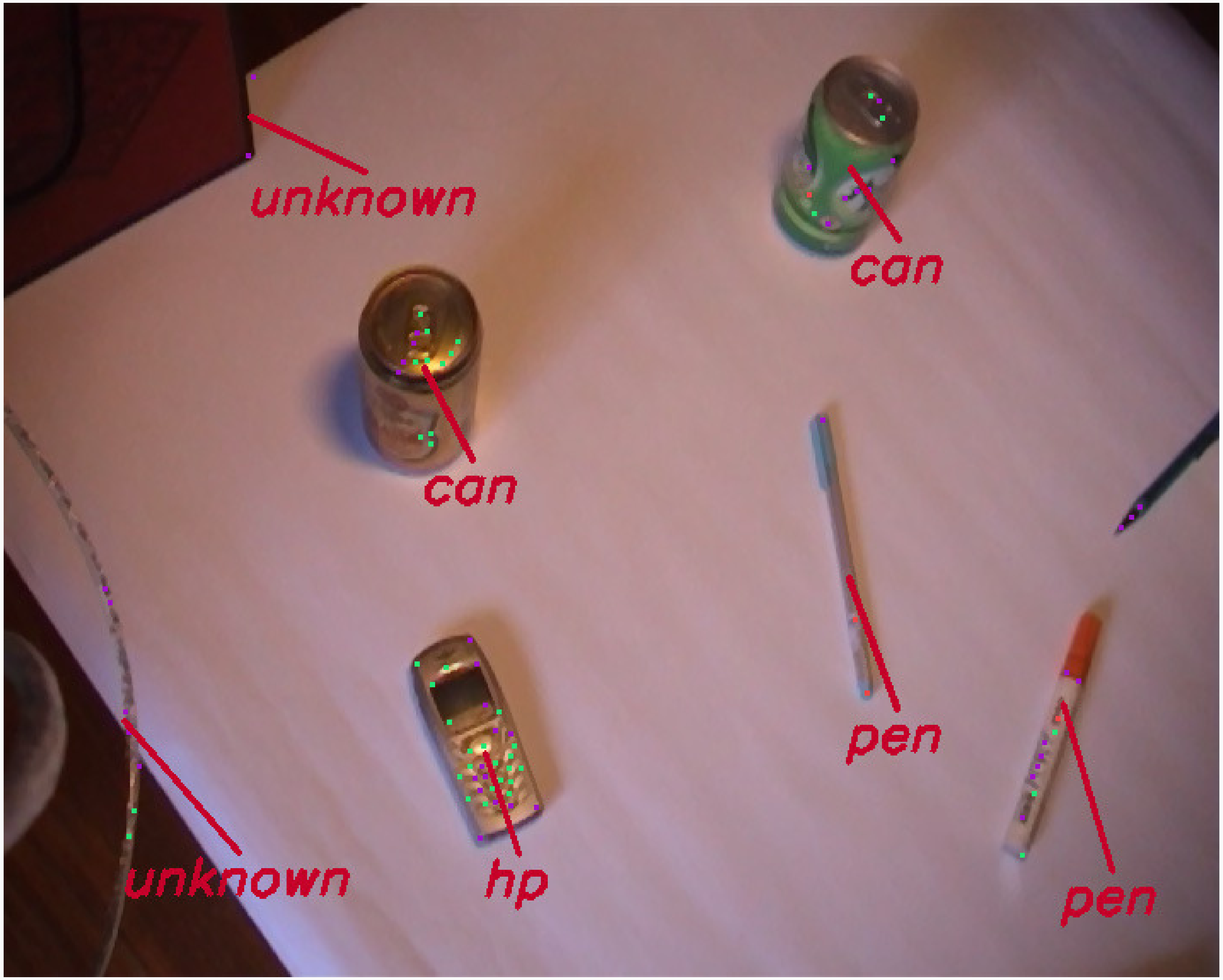
3. Pass down multiple decision trees to decide on the class with highest average probability:

$$\hat{Y}(p) = \arg\max_{c}(p_c(p)) = \frac{\arg\max_{c} \frac{1}{n} \sum_{1\ldots n} P_{\eta(l,\, p)}\,(Y(p) = c), c \in \{c_1, \ldots c_l, \ldots\}}{P(c \in \{c_1, \ldots c_l, \ldots\})}$$

where l is a class from a reduced set $\{c_1, \ldots c_l, \ldots\}$

provided by the aspect ratio classifier

Putting them Together



References

Patch-based Appearance Classifier using random, multiple decision trees
Towards Recognizing Feature Points using Classification Trees
Vincent Lepetit, Pascal Fua 2004 in EPFL Technical Report IC/2004/74

Ideas for system
Object Recognition at a Glance
John Winn and Antonio Criminisi 2006

Algorithms
OpenCV Version 1.0