Object Detection (due Saturday 3/9/2019)

In this assignment, you will develop an object detector based on gradient features and sliding window classification. A set of test images and ***hogvis.py*** are provided in the Canvas assignment directory

---

Name: Neeraj Dharmadhikari

SID: 95579195

In [3]:
```
import numpy as np
import matplotlib.pyplot as plt
```

1. Image Gradients [20 pts]

Write a function that takes a grayscale image as input and returns two arrays the same size as the image, the first of which contains the magnitude of the image gradient at each pixel and the second containing the orientation.

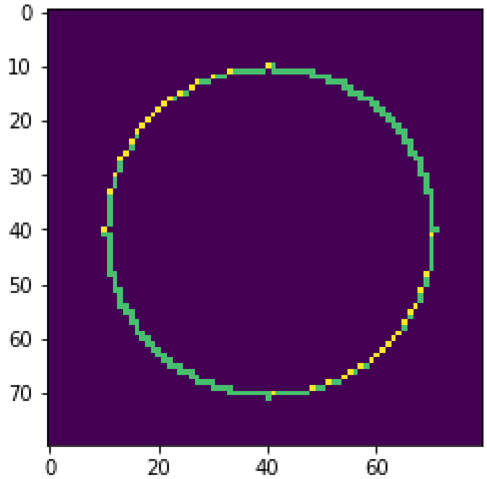
Your function should filter the image with the simple x- and y-derivative filters described in class. Once you have the derivatives you can compute the orientation and magnitude of the gradient vector at each pixel. You should use ***scipy.ndimage.correlate*** with the 'nearest' option in order to nicely handle the image boundaries.

Include a visualization of the output of your gradient calculate for a small test image. For displaying the orientation result, please uses a cyclic colormap such as "hsv" or "twilight". (see https://matplotlib.org/tutorials/colors/colormaps.html (https://matplotlib.org/tutorials/colors/colormaps.html))

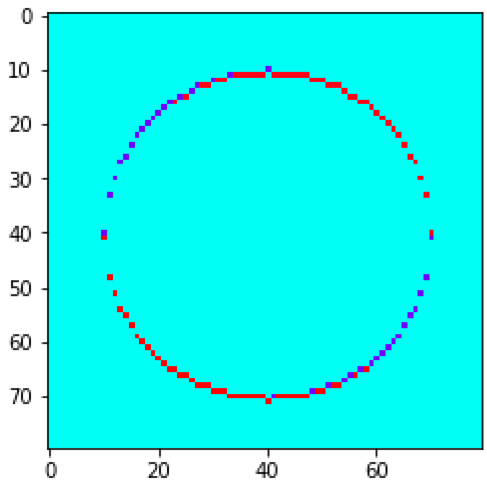
In [5]:
```
#we will only use:  scipy.ndimage.correlate
from scipy import ndimage as nd
import numpy as np

def mygradient(image):
    """
    This function takes a grayscale image and returns two arrays of the
    same size, one containing the magnitude of the gradient, the second
    containing the orientation of the gradient.


    Parameters
    ----------
    image : 2D float array of shape HxW
         An array containing pixel brightness values

    Returns
    -------
    mag : 2D float array of shape HxW
        gradient magnitudes

    ori : 2Dfloat array of shape HxW
        gradient orientations in radians
    """
    #create [2,1] & [1,2] filters that go through the array
    x_filter = [[-1,1]]
    y_filter = [[-1],[1]]
    x_diff = nd.correlate(image, x_filter, mode='nearest')
    y_diff = nd.correlate(image, y_filter, mode='nearest')

    x_diff_sqrd = x_diff*x_diff
    y_diff_sqrd = y_diff*y_diff

    sumOfXYDiffSqrd = x_diff_sqrd + y_diff_sqrd

    mag = np.sqrt(sumOfXYDiffSqrd)

    #mag = generic_gradient_magnitude(a, sobel)
    '''
    h,w = image.shape
    dx_image = image[:,1:w-1] - image[:,0:w-2]
    dy_image = image[1:h-1,:] - image[0:h-2,:]
    '''
    ori = np.arctan(y_diff/(x_diff+0.00001))

    # your code goes here

    return (mag,ori)
```

In [6]:

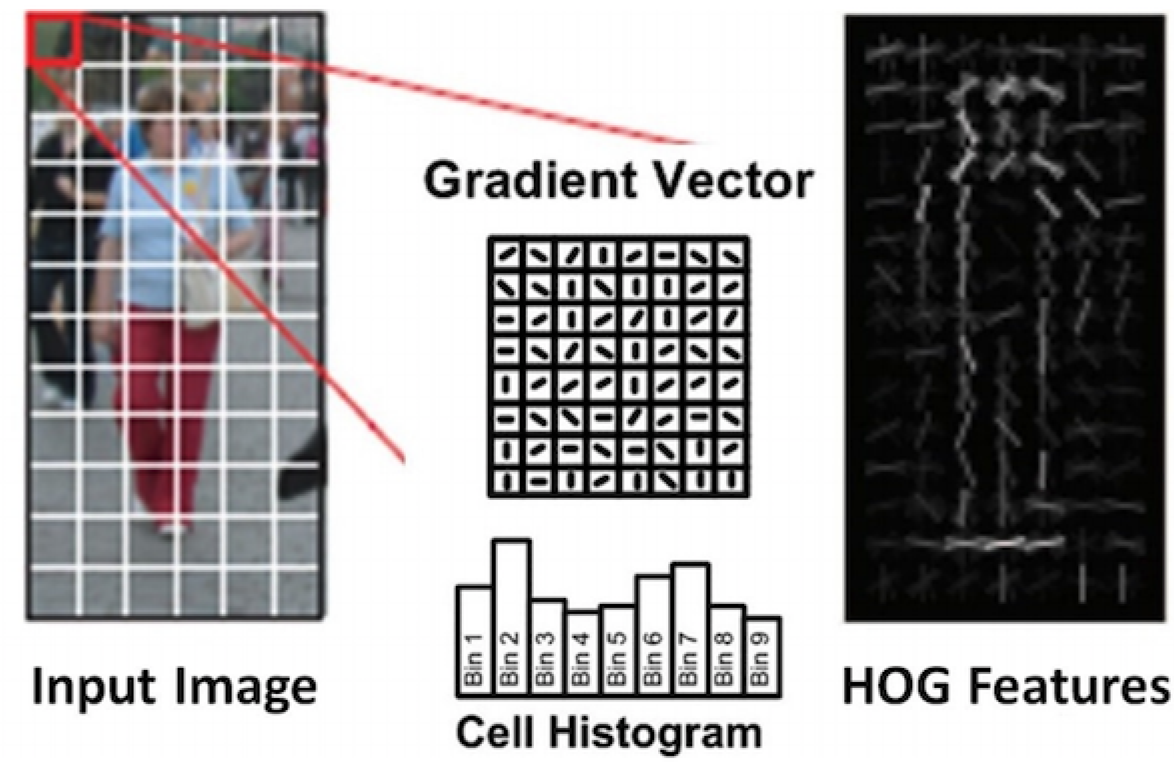
```
#
# Demonstrate your mygradient function here by loading in a grayscale
# image, calling mygradient, and visualizing the resulting magnitude
# and orientation images. For visualizing orientation image, I suggest
# using the hsv or twilight colormap.
#
import matplotlib.pyplot as plt
import numpy as np

#coloredImg = plt.imread("assignment4_files/images/faces/faces1.jpg")
'''
if (coloredImg.dtype == np.uint8):
    coloredImg = coloredImg.astype(float) / 256

image = (coloredImg[:,:,0] + coloredImg[:,:,1] + coloredImg[:,:,2])/3
'''
#plt.imshow(coloredImg)
#plt.show()

[yy,xx] = np.mgrid[-40:40,-40:40]
im = np.array((xx*xx+yy*yy<=30*30),dtype=float)

(mag,ori) = mygradient(im)

plt.imshow(mag)
plt.show()

plt.imshow(ori, cmap=plt.cm.hsv)
plt.show()

#visualize results.
```



0
10
20
30
40
50
60
70
0
20
40
60



0
10
20
30
40
50
60
70
0
20
40
60



2. Histograms of Gradient Orientations [25 pts]

Write a function that computes gradient orientation histograms over each 8x8 block of pixels in an image. Your function should bin the orientation into 9 equal sized bins between -pi/2 and pi/2. The input of your function will be an image of size HxW. The output should be a three-dimensional array ***ohist*** whose size is (H/8)x(W/8)x9 where ***ohist[i,j,k]*** contains the count of how many edges of orientation k fell in block (i,j). If the input image dimensions are not a multiple of 8, you should use ***np.pad*** with the ***mode=edge*** option to pad the width and height up to the nearest integer multiple of 8.

To determine if a pixel is an edge, we need to choose some threshold. I suggest using a threshold that is 10% of the maximum gradient magnitude in the image. Since each 8x8 block will contain a different number of edges, you should normalize the resulting histogram for each block to sum to 1 (i.e., ***np.sum(ohist,axis=2)*** should be 1 at every location).

I would suggest your function loops over the orientation bins. For each orientation bin you'll need to identify those pixels in the image whose magnitude is above the threshold and whose orientation falls in the given bin. You can do this easily in numpy using logical operations in order to generate an array the same size as the image that contains Trues at the locations of every edge pixel that falls in the given orientation bin and is above threshold. To collect up pixels in each 8x8 spatial block you can use the function ***ski.util.view_as_windows(..., (8,8),step=8)*** and ***np.count_nonzeros*** to count the number of edges in each block.

Test your code by creating a simple test image (e.g. a white disk on a black background), computing the descriptor and using the provided function ***hogvis*** to visualize it.

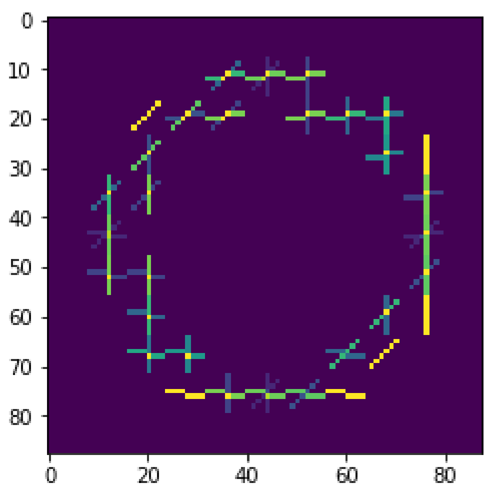
Note: in the discussion above I have assumed 8x8 block size and 9 orientations. In your code you should use the parameters ***bsize*** and ***norient*** in place of these constants.

In [7]:
```
#we will only use:  ski.util.view_as_windows for computing hog descriptor
import skimage as ski
import math
def hog(image,bsize=8,norient=9):

    """
    This function takes a grayscale image and returns a 3D array
    containing the histogram of gradient orientations descriptor (HOG)
    We follow the convention that the histogram covers gradients starting
    with the first bin at -pi/2 and the last bin ending at pi/2.

    Parameters
    ----------
    image : 2D float array of shape HxW
         An array containing pixel brightness values

    bsize : int
        The size of the spatial bins in pixels, defaults to 8

    norient : int
        The number of orientation histogram bins, defaults to 9

    Returns
    -------
    ohist : 3D float array of shape (H/bsize,W/bsize,norient)
        edge orientation histogram

    """

    # determine the size of the HOG descriptor

    (h,w) = image.shape
 #   print("h:", str(h))
  #  print("w:", str(w))
    h2 = int(np.ceil(h/float(bsize))) #height of the returning 3d array rounde
d UP to the nearest int
    w2 = int(np.ceil(w/float(bsize))) #width of the returning 3d array rounded
UP to the nearest int
    ohist = np.zeros((h2,w2,norient)) #this will just be height x width x 9
   # print("h2:", str(h2))
  #  print("w2:", str(w2))

    # pad the input image on right and bottom as needed so that it
    # is a multiple of bsize
    #pw = (0,...)
    pw = (0, (w2*bsize)-w)
   # print("pw:", str(pw))
    #ph = (0,...)
    ph = (0, h2*bsize-h)
   # print("ph:", str(ph))

    #image = np.pad(image,(ph,pw),'symmetric')
    image = np.pad(image,(ph,pw),'edge')
    #print("image shape:")
    #print(image.shape)
    #print("h2xbsize")
```

```
    #print(h2*bsize)
    #print("w2xbsize")
    #print(w2*bsize)
    # make sure we did the padding correctly
    assert(image.shape==(h2*bsize,w2*bsize))

    # compute image gradients
    (mag,ori) = mygradient(image)

    # choose a threshold which is 10% of the maximum gradient magnitude in the
image
    maxGradMag = np.amax(mag)
    #print("max gradient magnitude in image: ", str(maxGradMag))
    #thresh = ...
    thresh = 0.10*maxGradMag #thresh is now a float that is 10% of the the max
imum gradient magnitude in our image
    #print("10% of that magnitude", str(thresh))

    # separate out pixels into orientation channels, dividing the range of ori
entations
    # [-pi/2,pi/2] into norient equal sized bins and count how many fall in ea
ch block
    # as a sanity check, make sure every pixel gets assigned to at most 1 bin.
    bincount = np.zeros((h2*bsize,w2*bsize))
    for i in range(norient):
        #create a binary image containing 1s for pixels at the ith
        #orientation where the magnitude is above the threshold.

        B = np.zeros(mag.shape)
        #B = ...
        #we need to check if the orientation is in the right bin here?
        for j in range(len(mag)):
            for k in range(len(mag[0])):
                if mag[j,k] > thresh:
                    if(i == 0 and ori[j,k] == ((-math.pi)/2)):
                        B[j,k] = 1
                    if ori[j,k] > ((-math.pi)/2)+((math.pi)/9)*i and ori[j,k]
<= ((-math.pi)/2)+((math.pi)/9)*(i+1):
                        B[j,k] = 1



        #sanity check
        bincount = bincount + B
        #print("bincount:")
        #print(bincount)

        #pull out non-overlapping bsize x bsize blocks
        chblock = ski.util.view_as_windows(B,(bsize,bsize),step=bsize)
        #print(chblock)

        #sum up the count for each block and store the results
        #...
        #ohist[:,:,i] = ...
        temp = np.zeros((h2,w2))
        for m in range(chblock.shape[0]):
            for n in range(chblock.shape[1]):
```

```
                temp[m,n] = np.sum(chblock[m,n])

        ohist[:,:,i] = temp

    #print(ohist)
    assert(np.all(bincount<=1))

    # lastly, normalize the histogram so that the sum along the orientation di
mension is 1
    # note: don't divide by 0! If there are no edges in a block (i.e. the sum
 of counts
    # is 0) then your code should leave all the values as zero.
    for i in range(ohist.shape[0]):
        for j in range(ohist.shape[1]):
            sumOfHist = 0
            for k in range(ohist.shape[2]):
                sumOfHist += ohist[i,j,k]
            for k in range(ohist.shape[2]):
                if sumOfHist == 0:
                    #do nothing
                    ohist[i,j,k] = 0
                else:
                    ohist[i,j,k] = ohist[i,j,k]/sumOfHist

    #print("ohist")
    #print(ohist)

    #...


    assert(ohist.shape==(h2,w2,norient))

    return ohist
```

In [8]:
```
#provided function for visualizing hog descriptors
import hogvis as hogvis

#
# generate a simple test image... a 80x80 image
# with a circle of radius 30 in the center
#
[yy,xx] = np.mgrid[-44:44,-44:44]
im = np.array((xx*xx+yy*yy<=30*30),dtype=float)

hogDescriptor = hog(im)
#
# display the image and the output of hogvis
#
#... = hogvis.hogvis(...)
img = hogvis.hogvis(hogDescriptor)

plt.imshow(img)
plt.show()
```



3. Detection [25 pts]

Write a function that takes a template and an image and returns the top detections found in the image. Your function should follow the definition given below.

In your function you should first compute the histogram-of-gradient-orientation feature map for the image, then correlate the template with the feature map. Since the feature map and template are both three dimensional, you will want to filter each orientation separately and then sum up the results to get the final response. If the image of size HxW then this final response map will be of size (H/8)x(W/8).

When constructing the list of top detections, your code should implement non-maxima suppression so that it doesn't return overlapping detections. You can do this by sorting the responses in descending order of their score. Every time you add a detection to the list to return, check to make sure that the location of this detection is not too close to any of the detections already in the output list. You can estimate the overlap by computing the distance between a pair of detections and checking that the distance is greater than say 70% of the width of the template.

Your code should return the locations of the detections in terms of the original image pixel coordinates (so if your detector had a high response at block **[i,j]** in the response map, then you should return **(8i,8j)** as the pixel coordinates).

I have provided a function for visualizing the resulting detections which you can use to test your detect function. Please include some visualization of a simple test case.

In [9]:
```
#we will only use:  scipy.ndimage.correlate
from scipy import ndimage

def detect(image,template,ndetect=5,bsize=8,norient=9):

    """
    This function takes a grayscale image and a HOG template and
    returns a list of detections where each detection consists
    of a tuple containing the coordinates and score (x,y,score)

    Parameters
    ----------
    image : 2D float array of shape HxW
         An array containing pixel brightness values

    template : a 3D float array
        The HOG template we wish to match to the image

    ndetect : int
        Number of detections to return

    bsize : int
        The size of the spatial bins in pixels, defaults to 8

    norient : int
        The number of orientation histogram bins, defaults to 9

    Returns
    -------
    detections : a list of tuples of length ndetect
        Each detection is a tuple (x,y,score)

    """

    # norient for the template should match the norient parameter passed in
    assert(template.shape[2]==norient)

    fmap = hog(image,bsize=bsize,norient=norient)


    #cross-correlate the template with the feature map to get the total respon
se
    resp = np.zeros((fmap.shape[0],fmap.shape[1]))
    for i in range(norient):
        #resp = resp + ndimage.correlate(...)
        resp = resp + ndimage.correlate(fmap[:,:,i],template[:,:,i], mode = 'n
earest')

    # sort the values in resp in descending order.
    # val[i] should be ith largest score in resp
    # ind[i] should be the index at which it occurred so that val[i]==resp[ind
[i]]
    #
    #val = ...  #sorted response values
    flat_resp = resp.flatten()
    val = np.sort(flat_resp, axis=None)[::-1]
```

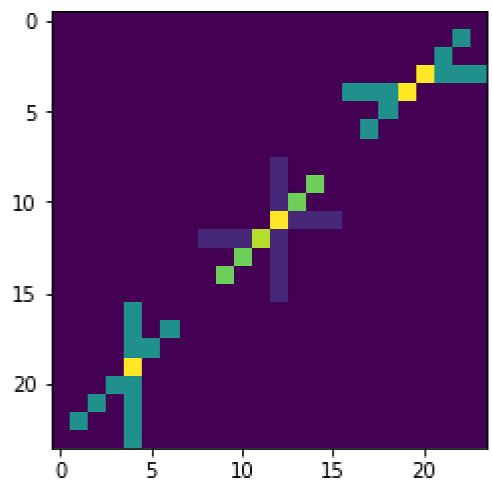
```
    #ind = ...  #corresponding indices
    ind = np.argsort(flat_resp, axis=None)[::-1]


    #work down the list of responses from high to low, to generate a
    # list of ndetect top scoring matches which do not overlap
    detcount = 0
    i = 0
    detections = []
    while ((detcount < ndetect) and (i < len(val))):
        # convert 1d index into 2d index
        yb = ind[i]//resp.shape[1]
        xb = ind[i]%resp.shape[1]

        assert(val[i]==resp[yb,xb]) #make sure we did indexing correctly

        #covert block index to pixel coordinates based on bsize
        xp = xb*bsize
        yp = yb*bsize

        #check if this detection overlaps any detections that we've already ad
ded
        #to the list. compare the x,y coordinates of this detection to the x,y
        #coordinates of the detections already in the list and see if any over
lap
        #by checking if the distance between them is less than 70% of the temp
late
        # width/height

        #...
        overlap = 0
        for j in range(detcount):
            xdiff = abs(detections[j][0] - xp)
            ydiff = abs(detections[j][1] - yp)
            #print("xdiff", xdiff)
            #print("ydiff", ydiff)
            if xdiff < 0.30*template.shape[1]*bsize or ydiff < 0.30*template.s
hape[0]*bsize: #then we have an overlap
                #print("overlap set: ", str(i))
                overlap = 1
                break


        #if the detection doesn't overlap then add it to the list
        if not overlap:
            detcount = detcount + 1
            detections.append((xp,yp,val[i]))

        i=i+1

    if (len(detections) < ndetect):
        print('WARNING: unable to find ',ndetect,' non-overlapping detections'
)

    return detections
```

In [ ]:

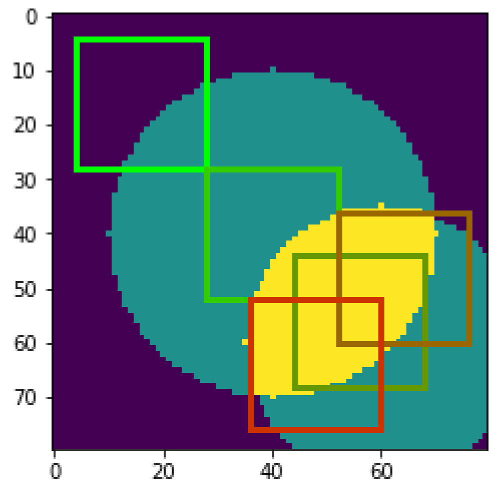
In [10]:
```
import matplotlib.patches as patches

def plot_detections(image,detections,tsize_pix):
    """
    This is a utility function for visualization that takes an image and
    a list of detections and plots the detections overlayed on the image
    as boxes.

    Color of the bounding box is based on the order of the detection in
    the list, fading from green to red.

    Parameters
    ----------
    image : 2D float array of shape HxW
         An array containing pixel brightness values

    detections : a list of tuples of length ndetect
        Detections are tuples (x,y,score)

    tsize_pix : (int,int)
        The height and width of the box in pixels

    Returns
    -------
    None

    """
    ndetections = len(detections)

    plt.imshow(image)
    ax = plt.gca()
    w = tsize_pix[1]
    h = tsize_pix[0]
    red = np.array([1,0,0])
    green = np.array([0,1,0])
    ct = 0
    for (x,y,score) in detections:
        xc = x-(w//2)
        yc = y-(h//2)
        col = (ct/ndetections)*red + (1-(ct/ndetections))*green
        rect = patches.Rectangle((xc,yc),w,h,linewidth=3,edgecolor=col,facecol
or='none')
        ax.add_patch(rect)
        ct = ct + 1

    plt.show()
```

In [11]:
```
#
# sketch of some simple test code, modify as needed
#


#create a synthetic image
[yy,xx] = np.mgrid[-40:40,-40:40]
im1 = np.array((xx*xx+yy*yy<=30*30),dtype=float)
[yy,xx] = np.mgrid[-60:20,-60:20]
im2 = np.array((xx*xx+yy*yy<=25*25),dtype=float)
im = 0.5*im1+0.5*im2

#compute feature map with default parameters
fmap = hog(im)

#extract a 3x3 template
template = fmap[1:4,1:4,:]
img = hogvis.hogvis(template)
print("HOG descriptor version of template:")
plt.imshow(img)
plt.show()
print("Detections:")

#run the detect code
detections = detect(im,template,ndetect=5)

#visualize results.
plot_detections(im,detections,(24,24))

# visually confirm that:
#   1. top detection should be the same as the location where we selected the
 template
#   2. multiple detections do not overlap too much
```

```
HOG descriptor version of template:
```



```
Detections:
```



---

4. Learning Templates [15 pts]

The final step is to implement a function to learn a template from positive and negative examples. Your code should take a collection of cropped positive and negative examples of the object you are interested in detecting, extract the features for each, and generate a template by taking the average positive template minus the average negative template.

In [12]:
```
def learn_template(posfiles,negfiles,tsize=np.array([16,16]),bsize=8,norient=9
):
    """
    This function takes a list of positive images that contain cropped
    examples of an object + negative files containing cropped background
    and a template size. It produces a HOG template and generates visualizatio
n
    of the examples and template

    Parameters
    ----------
    posfiles : list of str
         Image files containing cropped positive examples

    negfiles : list of str
        Image files containing cropped negative examples

    tsize : (int,int)
        The height and width of the template in blocks

    Returns
    -------
    template : float array of size tsize x norient
        The learned HOG template

    """

    #compute the template size in pixels
    #corresponding to the specified template size (given in blocks)
    tsize_pix=bsize*tsize

    #figure to show positive training examples
    fig1 = plt.figure()
    pltct = 1

    #accumulate average positive and negative templates
    pos_t = np.zeros((tsize[0],tsize[1],norient),dtype=float)
    for file in posfiles:
        #load in a cropped positive example
        img = plt.imread(file)

        #convert to grayscale and resize to fixed dimension tsize_pix
        #using skimage.transform.resize if needed.

        if (img.dtype == np.uint8):
            img = img.astype(float) / 256

        img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3

        img = ski.transform.resize(img, tsize_pix)

        #display the example. if you want to train with a large # of examples,
        #you may want to modify this, e.g. to show only the first 5.
        ax = fig1.add_subplot(len(posfiles),1,pltct)
        ax.imshow(img,cmap=plt.cm.gray)
        pltct = pltct + 1
```

```
        #extract feature
        fmap = hog(img,bsize,norient)

        #compute running average (just sum)
        pos_t = pos_t + fmap

    pos_t = (1/len(posfiles))*pos_t
    fig1.show()

    # repeat same process for negative examples
    fig2 = plt.figure()
    pltct = 1
    neg_t = np.zeros((tsize[0],tsize[1],norient),dtype=float)
    for file in negfiles:

        img = plt.imread(file)
        #convert to grayscale and resize to fixed dimension tsize_pix
        #using skimage.transform.resize if needed.
        if (img.dtype == np.uint8):
            img = img.astype(float) / 256
        img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3
        img = ski.transform.resize(img, tsize_pix)

        ax = fig2.add_subplot(len(negfiles),1,pltct)
        ax.imshow(img,cmap=plt.cm.gray)
        pltct = pltct + 1

        #extract feature
        fmap = hog(img,bsize,norient)

        #compute running average (just sum)
        neg_t = neg_t + fmap


    neg_t = (1/len(negfiles))*neg_t
    fig2.show()
    plt.show()

    # add code here to visualize the positive and negative parts of the templa
te
    # using hogvis. you should separately visualize pos_t and neg_t rather tha
n
    # the final tempalte.
    imgp = hogvis.hogvis(pos_t)
    print("HOG descriptor version of pos_t:")
    plt.imshow(imgp)
    plt.show()
    imgn = hogvis.hogvis(neg_t)
    print("HOG descriptor version of neg_t:")
    plt.imshow(imgn)
    plt.show()

    # now construct our template as the average positive minus average negativ
e
    template = pos_t - neg_t
```

```
    img = hogvis.hogvis(template)
    print("HOG descriptor version of template:")
    plt.imshow(img)
    plt.show()


    return template
```

In [ ]:

```
```

5. Experiments [15 pts]

Test your detection by training a template and running it on a test image.

In your experiments and writeup below you should include: (a) a visualization of the positive and negative patches you use to train the template and corresponding hog feature, (b) the detection results on the test image. You should show (a) and (b) for ***two different object categories***, the provided face test images and another category of your choosing (e.g. feel free to experiment with detecting cat faces, hands, cups, chairs or some other type of object). Additionaly, please include results of testing your detector where there are at least 3 objects to detect (this could be either 3 test images which each have one or more objects, or a single image with many (more than 3) objects). Your test image(s) should be distinct from your training examples. Finally, write a brief (1 paragraph) discussion of where the detector works well and when it fails. Describe some ways you might be able to make it better.

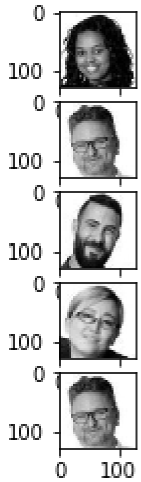
NOTE 1: You will need to create the cropped test examples to pass to your ***learn_template***. You can do this by cropping out the examples by hand (e.g. using an image editing tool). You should attempt to crop them out in the most consistent way possible, making sure that each example is centered with the same size and aspect ratio. Negative examples can be image patches that don't contain the object of interest. You should crop out negative examples with roughly the same resolution as the positive examples.

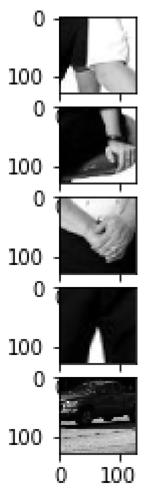
NOTE 2: For the best result, you will want to test on images where the object is the same size as your template. I recommend using the default ***bsize*** and ***norient*** parameters for all your experiments. You will likely want to modify the template size as needed

Experiment 1: Face detection

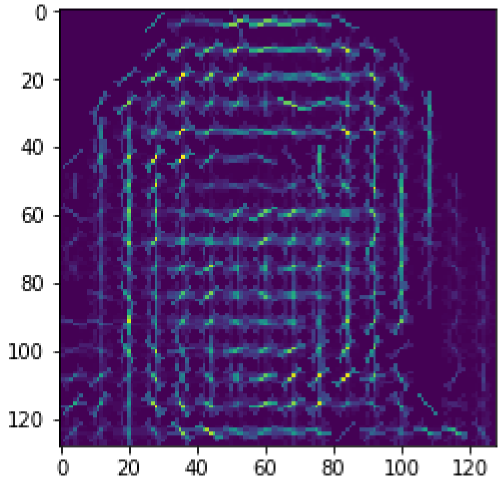
In [13]:
```
# assume template is 16x16 blocks, you may want to adjust this
# for objects of different size or aspect ratio.
# compute image a template size
bsize=8
tsize=np.array([16,16])  #height and width in blocks
tsize_pix = bsize*tsize  #height and width in pixels
posfiles = ('pos1.jpg','pos2.jpg','pos3.jpg','pos4.jpg','pos5.jpg')
negfiles = ('neg1.jpg','neg2.jpg','neg3.jpg','neg4.jpg','neg5.jpg')


# call learn_template to learn and visualize the template and training data
template = learn_template(posfiles,negfiles,tsize=tsize)

# call detect on one or more test images, visualizing the result with the plot
_detections function

img = plt.imread('assignment4_files/images/faces/faces3.jpg')
if (img.dtype == np.uint8):
    img = img.astype(float) / 256

    img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3

detections = detect(img, template, 4)
plot_detections(img,detections,tsize_pix)
```

```
C:\Users\neera\Anaconda3\lib\site-packages\skimage\transform\_warps.py:105: U
serWarning: The default mode, 'constant', will be changed to 'reflect' in ski
mage 0.15.
  warn("The default mode, 'constant', will be changed to 'reflect' in "
C:\Users\neera\Anaconda3\lib\site-packages\skimage\transform\_warps.py:110: U
serWarning: Anti-aliasing will be enabled by default in skimage 0.15 to avoid
aliasing artifacts when down-sampling images.
  warn("Anti-aliasing will be enabled by default in skimage 0.15 to "
C:\Users\neera\Anaconda3\lib\site-packages\matplotlib\figure.py:445: UserWarn
ing: Matplotlib is currently using module://ipykernel.pylab.backend_inline, w
hich is a non-GUI backend, so cannot show the figure.
  % get_backend())
```

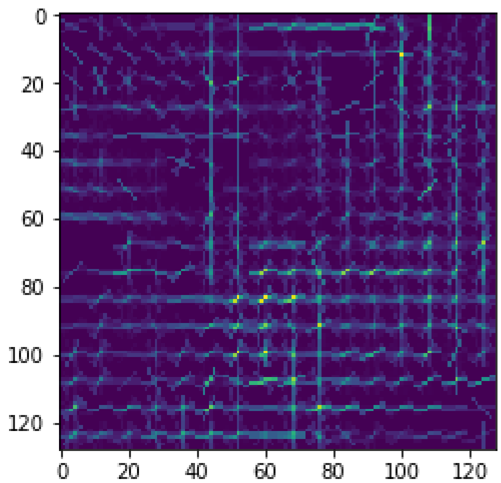




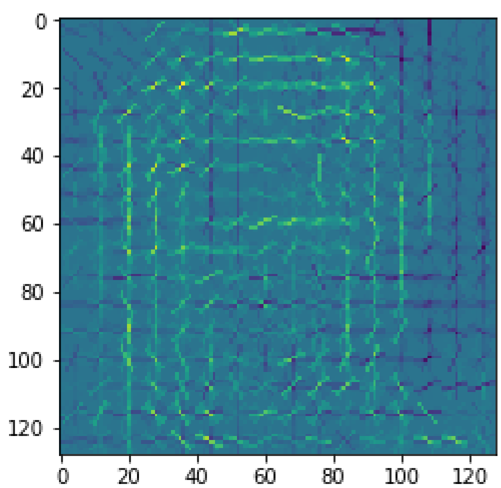
```
HOG descriptor version of pos_t:
```

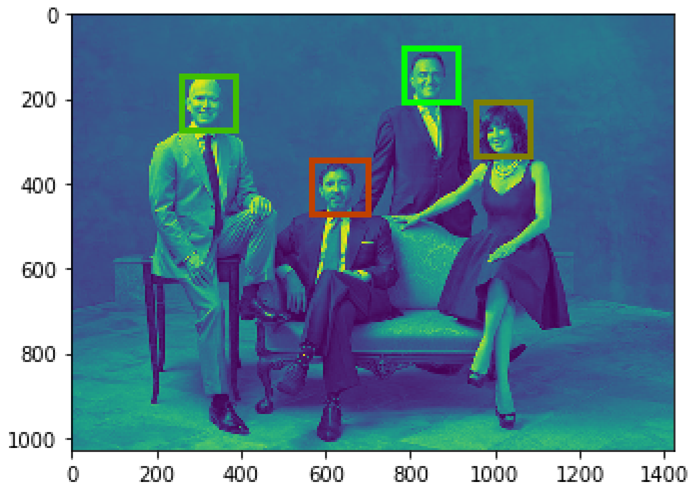


HOG descriptor version of neg_t:



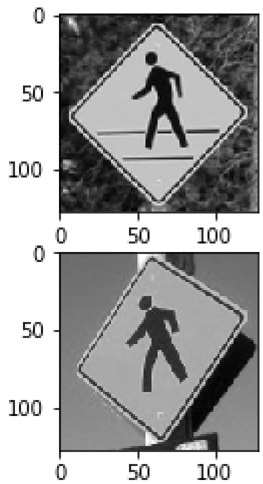
HOG descriptor version of template:

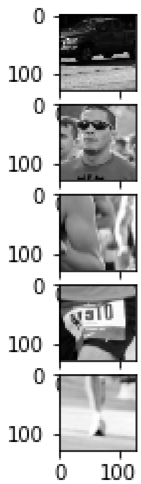




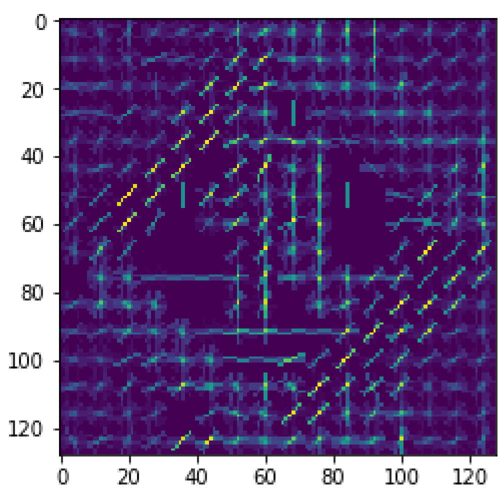
Experiment 2: ??? detection

In [14]:
```
# assume template is 16x16 blocks, you may want to adjust this
# for objects of different size or aspect ratio.
# compute image a template size
bsize=8
tsize=np.array([16,16])  #height and width in blocks
tsize_pix = bsize*tsize  #height and width in pixels
posfiles = ('pos6.jpg','pos7.jpg')
negfiles = ('neg6.jpg','neg7.jpg','neg8.jpg','neg9.jpg','neg10.jpg')


# call learn_template to learn and visualize the template and training data
template = learn_template(posfiles,negfiles,tsize=tsize)

# call detect on one or more test images, visualizing the result with the plot
_detections function

img = plt.imread('assignment4_files/images/signs/test0.jpg')
if (img.dtype == np.uint8):
    img = img.astype(float) / 256

    img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3

detections = detect(img, template, 1)
plot_detections(img,detections,tsize_pix)

img = plt.imread('assignment4_files/images/signs/test4.jpg')
if (img.dtype == np.uint8):
    img = img.astype(float) / 256

    img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3

detections = detect(img, template, 1)
plot_detections(img,detections,tsize_pix)

img = plt.imread('assignment4_files/images/signs/test3.jpg')
if (img.dtype == np.uint8):
    img = img.astype(float) / 256

    img = (img[:,:,0] + img[:,:,1] + img[:,:,2])/3

detections = detect(img, template, 1)
plot_detections(img,detections,tsize_pix)
```

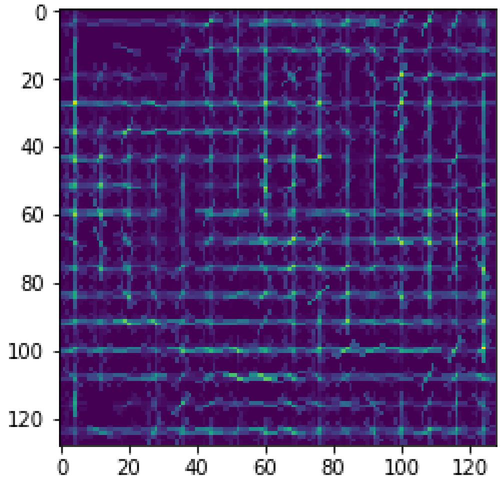




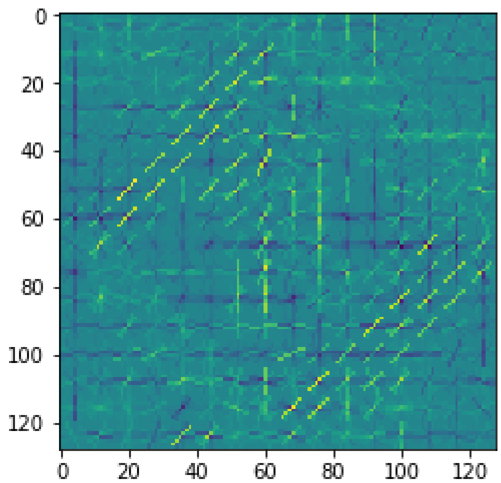

```
HOG descriptor version of pos_t:
```

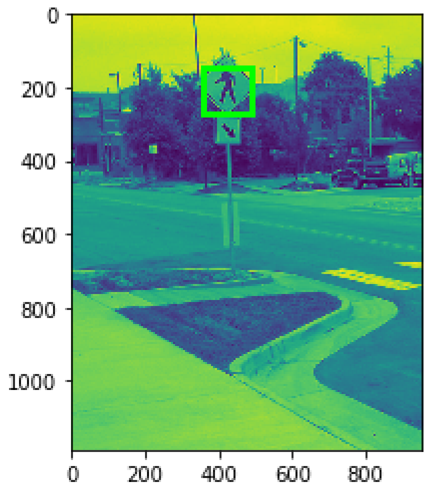


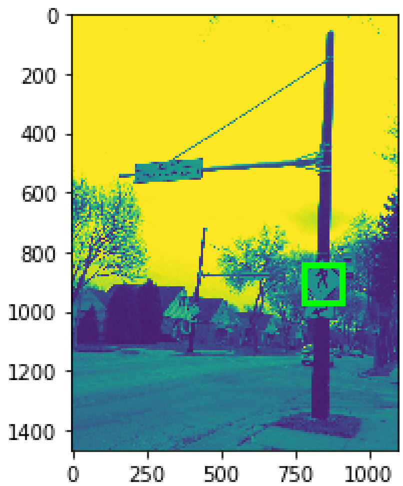

```
HOG descriptor version of neg_t:
```



HOG descriptor version of template:




0
200
400
600
800
1000
0
200
400
600
800


0
200
400
600
800
1000
1200
1400
0
250
500
750
1000

In [217]:
```
'''
Finally, write a brief (1 paragraph) discussion of where the detector works we
ll and when it fails.
Describe some ways you might be able to make it better.

ANS:
After executing the test with multiple photos, I've come to realize that the d
etector works better the more positive
and negative images that I give it to create a template. It also fails if I gi
ve it only two positive images, both
containing the pattern of certain size and the actual image containing a small
er and a bigger pattern. For example, I
gave the pos6.jpg and pos7.jpg the way I did in the last example above. When w
e pass in the image containing two signs
however, (the last image), we notice that it only identified the larger sign i
n the photo and the smaller one didn't get
detected. So we need to make sure that the positive images of the pattern are
 of different sizes and contain the pattern
at different distances. Other things we can modify consist of: threshold of th
e overlap and the threshold of the max
gradient magnitude that we had set to 70% and 10% respectively. Modifying thos
e depending on the pattern to detect and the
images we are given can also change detections. For example, if we are trying
 to detect faces that are all really close to
each other, then the 70% overlap requirement might make us miss the faces that
are really close. Adjusting the threshold of
10% of max gradient magnitude could also help us create better HOG descriptors
for various patterns, and that could help
better detect images.

'''
```

Out[217]:
```
"\nFinally, write a brief (1 paragraph) discussion of where the detector work
s well and when it fails. \nDescribe some ways you might be able to make it b
etter.\n\nANS:\nAfter executing the test with multiple photos, I've come to r
ealize that the detector works better the more positive\nand negative images
that I give it to create a template. It also fails if I give it only two posi
tive images, both \ncontaining the pattern of certain size and the actual ima
ge containing a smaller and a bigger pattern. For example, I \ngave the pos6.
jpg and pos7.jpg the way I did in the last example above. When we pass in the
image containing two signs\nhowever, (the last image), we notice that it only
identified the larger sign in the photo and the smaller one didn't get\ndetec
ted. So we need to make sure that the positive images of the pattern are of d
ifferent sizes and contain the pattern \nat different distances. Other things
we can modify consist of: threshold of the overlap and the threshold of the m
ax \ngradient magnitude that we had set to 70% and 10% respectively. Modifyin
g those depending on the pattern to detect and the\nimages we are given can a
lso change detections. For example, if we are trying to detect faces that are
all really close to\neach other, then the 70% overlap requirement might make
us miss the faces that are really close. Adjusting the threshold of\n10% of m
ax gradient magnitude could also help us create better HOG descriptors for va
rious patterns, and that could help\nbetter detect images.\n\n"
```