3.1 Filters

3.2 Histogram Processing

3.2.1 Histogram Equalization

3.2.2 Histogram of Oriented Gradient

3.3 Geometric Perspective Transformation

3.4 Thresholding

Image thresholding is the process of converting grayscale image to binary image. It scans all pixels in a given grayscale image, and decide whether each pixel should be converted to black or white.

Image thresholding could be divided into two categories, global thresholding and local thresholding. If the decision of pixel color depends only on the pixel value, it is called global thresholding. And if the decision of pixel color depends on the pixel value and other local properties (eg. The values of neighbor pixels), it is called adaptive thresholding.

3.4.1 Global Thresholding

Global thresholding is the simplest method of thresholding and image segmentation. It scans all image pixels, for each pixel if the pixel value is greater than or equal to the threshold value the pixel is labeled as white pixel, and if the pixel value is less than the threshold value the pixel is labeled as black pixel.

3.4.2 Adaptive Thresholding

Adaptive thresholding can be very useful in case of non-uniform lighting, where a global threshold value may not be the perfect choice in all image parts.

Adaptive thresholding find a threshold value for each pixel using its neighbor pixels. Two popular algorithms are finding the threshold value using the mean of neighbor pixels values and finding the threshold value using the weighted sum of neighbor pixels value where weights are a Gaussian window.

3.4.3 Otsu’s Thresholding

Otsu’s thresholding algorithm is a global thresholding algorithm used to automatically calculate the best threshold value especially in bimodal images (which have two peaks in their histogram).

Since the image histogram contains two peaks, Otsu’s thresholding scans all gray level values and find the value which minimize the weighted within-class variance which is the optimum threshold value.

The variance within the class defined as:

σ2 = ω0(t) σ02(t) + ω1(t) σ2(t)

Which is the weighted sum of variances of the two classes separated by threshold value. Where ω is the probability and σ is the variance of each class and t is the threshold value.

Otsu’s method can be very inaccurate in case of non-bimodal image. For example in case of small object in big background, or in the case when the object and background peaks in histogram are close to each other, so image enhancement techniques like histogram equalization is used before thresholding to minimize this issue.

3.5 Morphological Processing

Morphological operations are an image processing techniques related to the shape of image performed for the better representation of an image.

Morphological operations manipulate an image using another image called structuring element (or kernel).

Set theory is the language used by mathematical morphology to represent and manipulate images. A set contains tuples each represent one pixel, and the tuple contents are different for each image type. For example in binary images objects are represented by white in black background, the set contains tuples of pixels of the white object and each tuple contains x and y coordinates of the pixel. Also in grayscale images the tuple contains a third element for the value of the pixel.

There are two basic morphological operation, dilation and erosion. Other operations can be performed using these basic operations combined with other set’s operations like union, intersection, complement and difference.

Additional two set’s operations are reflections and translation.

Reflection of a set A denoted by Â is defined as:

Â = { b | b = -a for a ∈ A}

Translation of a set A by point z denoted by (A)z is defined as:

(A)z = { b | b = a + z for a ∈ A }

3.5.1 Dilation

Dilation of a set A by another set B as a structuring element denoted by A ⊕ B is defined as:

A ⊕ B = { z | [ ()z ∩ A ] ⊆ A}

It is obtained by reflecting B about the origin and translate it by z then find the points z in which the reflected translated B is contained in A.

In other words, the dilation of A by B is the points covered by B when the center of B moves inside A.

3.5.2 Erosion

Erosion of a set A by another set B as a structuring element denoted by A ⊖

B is defined as:

A ⊖ B = { z (B)z ⊆ A}

It is obtained by translating B by z then find the points z in which the translated B is contained in A.

In other words, the erosion of A by B is the points covered by the center of B when B moves inside A.

3.5.3 Opening and Closing

Opening of a set A by another set B as a structuring element denoted by A ∘ B is defined as:

A ∘ B = ( A ⊖ B ) ⊕ B

It is obtained by erosion of A by B then dilate the result by B.

Since dilation expands the image and erosion shrinks the image, opening can be used for filtering small objects in black background.

Closing of a set A by another set B as a structuring element denoted by A • B is defined as:

A • B = ( A ⊕ B ) ⊖ B

It is obtained by dilation of A by B then erode the result by B.

Since dilation expands the image and erosion shrinks the image, opening can be used for filtering black holes inside white objects.

3.6 Edge Processing

3.6.1 Image Derivative

3.6.2 Canny Edge Detection

3.7 Contours

3.8 Connected Components

Connected components is one of the most accurate algorithms used in image segmentation. A connected components is a group of pixels have the same value and connected to each other using 4-pixels or 8-pixels connections. Image is required thresholding before using this algorithm, because grayscale image of an object can’t be all in same gray level.

Connected components algorithm make use of morphological dilation and set’s intersection given a known pixel inside the object to be extracted. It can be described as:

Xk = ( Xk-1 ⊕ B ) ∩ A

Where A is the original image that it is desired to extract a component from, B is suitable structuring element and X is the new image containing only the extracted component.

It starts with the known pixel, a dilation is done to expand the component and intersection to eliminate parts of the component which is not exists in the original image. This operation is repeated until Xk is equal to Xk-1, and the result is an image with only one connected component.

Connected components algorithms is very useful in extracting an object in an image containing nested objects but not overlapping ones. But it is considered as time consumer due to high number of iterations in dilations and intersections. More knowledge about the object to be extracted can increase the speed of the algorithm, for example if objects in an image are not close to each other, then a larger structuring element can be used and that decrease the number of iterations required.

3.9 Image Correlation

Image correlation is a numeric representation of the similarity between two images. It describes how much an image A is similar to other image B.

The method of image correlation is the core concept of template matching algorithm. In template matching, a template image is moved across the original image while calculating the similarity in each point. This measure of similarity usually done using image correlation techniques.

3.9.1 Cross Correlation

Image cross-correlation is a numeric representation of the similarity between two images, and it is obtained by the sum of pairwise multiplications of corresponding pixel values of the images.

This can be described by:

Cross-Correlation (img1 , img2) =

This is the fundamental concept of image correlation, and it is not practically efficient due to lack of robustness. For example, the difference between image brightness may cause a bright image produce high value of correlation although the difference of the template to the original image.

3.9.2 Normalized Cross Correlation

Normalized cross-correlation is modified version of cross-correlation that overcomes some of the classical cross-correlation problems.

Normalized cross-correlation is obtained by the sum of pairwise multiplications of corresponding pixel values of the images after subtracting the mean value of each image from each pixel in that image. Then, the result is normalized by dividing by the multiplication of the standard deviation of each image and the number of pixels in them.

This is described by:

NCC (img1 , img2) =

This produce a more robust method that resists the changes in brightness to obtain more accurate results of the similarity between images.

The improvement of normalized cross-correlation over the classical cross-correlation can be viewed in two main points:

* The result is invariant to the global brightness of images.
* The result is normalized so that the output is within the interval [-1.0 , 1.0]. Where 1.0 is the result when comparing the image with itself, and -1.0 is the result when comparing the image with its negative.

3.10 Support Vector Machine

3.10.1 Overview

Traditional learning approaches suffer from difficulties in over-fitting of training data, and hence less generalization.

In techniques that use mapping of data in n-dimensional features space and uses boundaries to classify the input data, many decision boundaries can be used to perfectly classify the input data, but there is one optimum decision boundary that classify data with minimum error. To find that optimum decision boundary support vector machine algorithm is used.

Support vector machine is a supervised learning method used in data classification and regression for a linearly separable data by obtaining the optimum hyper-plane that classify the data.

In addition to performing linear classification, support vector machine can efficiently perform a non-linear classification using what is called the Kernel function, which maps inputs into high-dimensional feature spaces.

3.10.2 Inputs and Outputs

Inputs are sets of input-output training pairs, these represent a pair of features and result. The output is values of weights that describe how much each feature effects the result.

3.10.3 Hyper-plane

A hyper-plane of data that are mapped in two dimensional features space is a line that separate these data into given classes.

The further from the hyper-plane the data points lie, the more confident that they have been correctly classified. Therefore data points are desired to be as far away from the hyper-plane as possible, while still being on the correct side of it.

3.10.4 Support Vectors

Support vectors are the data points nearest to the hyper-plane, the points of a data set that, if removed, would alter the position of the dividing hyper-plane.

These are the data used to identify the hyper-plane since it must maximize the distance to both support vectors.

If the training data are linearly separable, two parallel hyper-planes can be selected to separate two classes of data, so that the distance between them is as large as possible. The region bounded by these two hyper-planes is called the margin, and the maximum-margin hyper-plane is the hyper-plane that lies halfway between them.

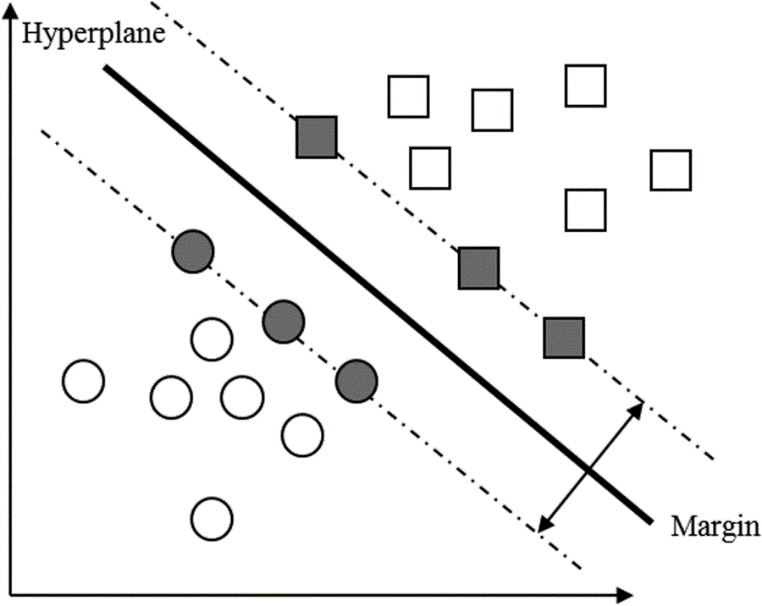


Figure 1

* + 1. Advantages
* It give the best results in classification where the result is more general and can be used for newly arrived data, and that increase the generalization and robustness of the system
* By using the kernel transformation the hyper-plane of non-linearly separable data can be obtained by finding the linear separation plane in higher coordinates.
* The uniqueness of solution is also one of the advantages of support vector machine, which ensure that the obtained solution is the optimum among all the available solutions.
* Support Vector Machine is a very powerful memory efficient algorithm since all data required to obtain the support vectors is in the training data itself.
  + 1. Limitations
* Since it seeks to identify the optimum hyper-plane, for fewer training data supplied to the system the result become less accurate because it is biased by error data.
* The algorithm is not time efficient. The time required for training the system using training data is long. So that the performance decrease. Resulting from computational complexity and seeking to generate the perfect solution.
* It can be used to classify data of two classes only. An extension to use SVM to classify multi-classes data is provided in many ways. The most common algorithm is to use (One vs the rest) where the hyper-plane separates the specific class from others in each iteration. By repeating this operation by the number of classes (n), a full classification of n-classes is resulted.

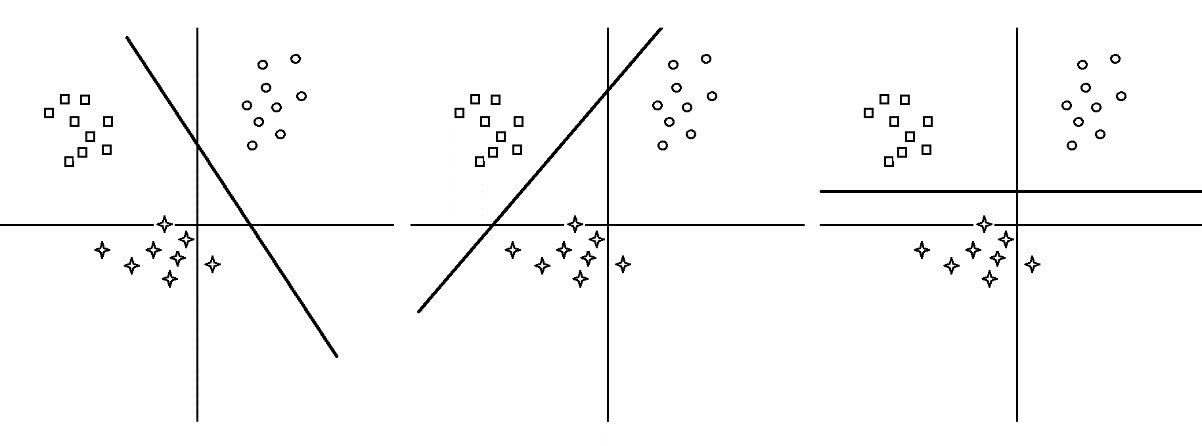


Figure 2

3.11 Implementation

3.11.1 Detection

3.11.2 Segmentation

3.11.3 Recognition