CSE468 Report

Submitted by

Omar Faruk Riyad

Arif Ur Rahaman Chowdhury Suhan

.

6 Falgun,1425

1   Introduction:

Human visualization is a sweet gift from The Almighty Creator.And it can't be perfectly replicated into a thing so that a machine can see.But the blind saw dream and we got hope.After the birth of Machine Learning, Computer Vision started a new era.

The report is focused on the feature extraction rather than developing a classification model.Using various kind of features like **Color Histogram**, **Histogram of Gradient** and **Scale Invariant Feature Transform** some collected datasets(images) are classified.

2   Literature Review:

2.1   1-NN:

The 1 Nearest Neighbor classifier is the simple and one of the oldest methods known.It's estimation is often low but the variance is high.The 1NN classifier idea is extremely simple: to classify $\mathbf{X}$ find its closest neighbor among the training points (call it $\mathbf{Y}$) and assign to $\mathbf{X}$ the label of $\mathbf{Y}$.

The good thing about 1NN is that it is conceptually simple and does not require learning model.Only few dataset is enough for it. Also it works for low dimensional complex decision surfaces.But for fixed,$\mathbf{K}$ it is asymptotically suboptimal,not by much.It's classification is slow and suffers a lot from the curse of dimensionality.

2.2   Metrics:

A matrics are 2D form of an entity composed of components arranged in rows and columns.

$$M = \begin{bmatrix} m_{00} & m_{01} & m_{02} \\ m_{10} & m_{11} & m_{12} \end{bmatrix} \tag{1}$$

Here M matrix contains 2 rows and 3 columns.
In computer vision, there are many types of matrics like:

- **Fundamental matrix:** A 3x3 matrix which relates corresponding points in stereo images in pixel coordinates.

- **Essential matrix:** A 3x3 matrix that also relates corresponding points but normalized pixel coordinates.

- **Camera matrix:** A 3x4 matrix describes the mapping of a pinhole camera from 3D points in the world to 2D points in an image.

2.3   Features:

A feature is a 1D vector that aims to represent an object or a part of an object in a compact. This means that features for different objects or object parts must be different and just by looking at the feature, one must be able to tell what object that is. That feature not change for different appearances of the object, which may be due to lighting, shadows, pose, etc. Also, one does a lot of computations with features, which is why it's size must be small.

Most algorithms in Computer Vision strive to identify features that can be used, for instance, for recognition or classification. A feature can also be a particular combination of simpler features, with some associated metrics and geometries. They can be used for :

- Interest point detection

- Descriptors and

- Matching of an image

3   Experimental setup:

For data set **CIFAR-10** is used.It consists of 60000 32x32 colour images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images.

The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly-selected images from each class. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class.

In the features, **Color Histogram** is faster but colors sometimes effect the classification.For example: white colored cat sometimes can't be classified if train dataset is blackish cat.

HOG and **SIFT** both are based on first order image gradients.But while HOG provides positional data, SIFT doesn't.From the point of descriptor, it can be said that HOG is a image descriptor but SIFT is keypoint one.

From performance perspective view, Color Histogram is much faster than HOG or SIFT.But for better dectection.

4 Results and discussion:

Dataset: Cifar 10			
Feature Name	Classifier Name	Accuracy	Comment
Color Histogram	1NN	0.5	—
HOG	1NN	NAN	—
SIFT	1NN	NAN	—

Color Histogram is basically on color frequency in range 0-255.It deals with image color.But it will not work always.Specially when Image with RGB value.For that case different object can have almost similliar type color histogram.Because for RGB histogram in feature extractor,the metrics is the average of the Red,Green and blue histogram.

For HOG, it is known that it can extract global feature.And SIFT is used for extracting local features.In global feature, a single metrics can give whole information of an image by using HOG. Scale and rotation invariant is followed by SIFT, HOG doesn't.Suppose our classifier wants to detect a walking human.If the human is moving left to right, then HOG can easily detect. But if the human comes near to camera and again back the starting position, then HOG is useless. SIFT can perform there better.

5 Conclusion:

As the reports shows , the classification is done on CIFAR-10 using just 1NN classifier. Instead of focusing learning model,it was always concern about the feature extraction. The best use of feature extraction varies dataset by dataset. For uniform color based dataset, color histogram will be better.Detecting object specially human if the size is invariant , then HOG is suitable. But when the object is moving or coming near and far, SIFT will be the perfect choice.