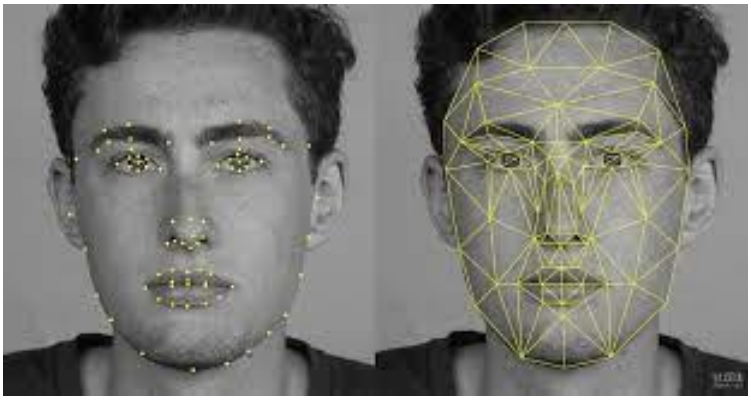
SEMINAR

WORKING BEHIND SNAPCHAT FILTERS

By *Adwitiya Mathur | B.Tech CS | 2016088*

Submitted To : Dr. Bharti Nathani

Working behind Snapchat's Lenses and explanation of Viola-Jones Algorithm



Index

Acknowledgement

In successfully completing this project, many people have helped me. I would like to thank all those who are related to this project.

I would like to express my gratitude to my teacher Dr Bharti Nathani who gave me the golden opportunity to present this seminar and write a report on the topic Working behind Snapchat Filters. It helped me in doing a lot of Research and I came to know about a lot of things related to this topic.

I am very grateful to my mentors and panel members for their constant support and guidance in completing this project.

Preface

At a point in our lives all of us must have been asked about the usage of these ‘weird - looking yet attractive’ Snapchat filters. It’s an instant hit in Gen Z and millennials equally. But have you ever wondered how this dog filter is detecting your face so perfectly?

History

Snapchat is an American multimedia instant messaging app and service developed by Snap Inc, originally Snapchat Inc.

Snapchat was created by Evan Spiegel, Bobby Murphy and Reggie Brown, former students at Stanford University. It has become known for representing a new, mobile-first direction for social media, and places significant emphasis on users interacting with virtual stickers and augmented reality objects.

Snapchat acquired a Ukrainian company called “Looksery” for this endless fun in 2014, for $150 million and launched filters itself in 2015 January, with the name “Lenses”.

Introduction

Snap.Inc ,the parent company of Snapchat, said that it's the multitude of Technology like artificial intelligence combined with machine learning techniques and Augmented reality that help create the filters.

Filters Snapchat uses AI powered lenses with a small machine learning model to detect facial features, differentiate the structure of features and then create a 3D model of face.

The process is completed using “Image processing and facial landmarks” - plotting points on face and comparing it with collected data sets and active shape modeling which creates a virtual face and aligns it with actual face.

Augmented reality is further able to create time machine filters or gender swapping filters etc..

Computer vision

Computer vision is an application that uses pixel data from a camera in order to identify objects and interpret space.

Computer vision is all about pattern recognition. One needs to feed thousands of images and let them analyze color, shapes, and distances so that it can identify a human face correctly.

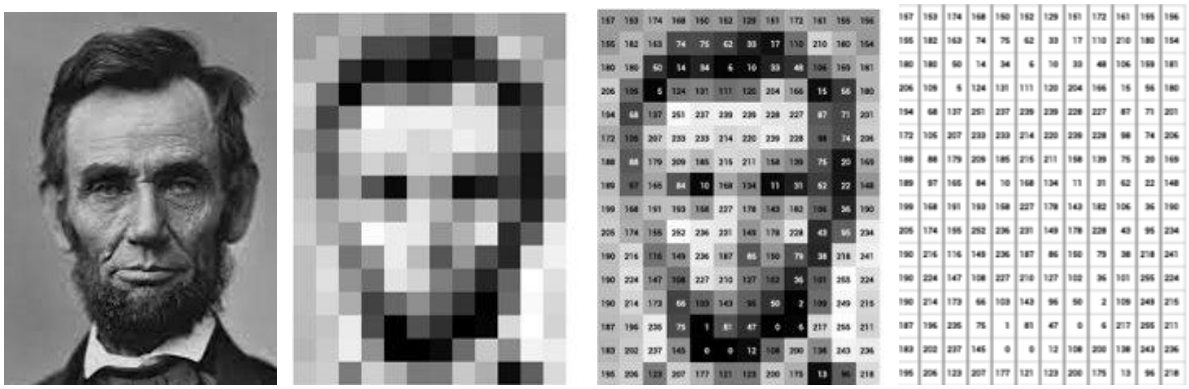
This is how google photos identify faces and sorts them into folders.

Pixelation

If we take an image and change the image resolution we can see pixels of an image.

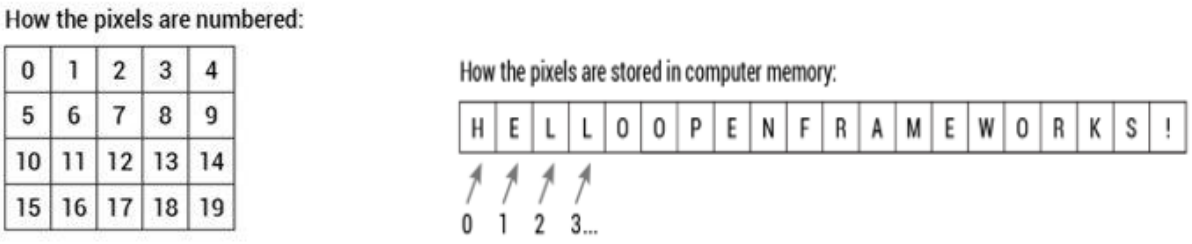
Pixel is the smallest unit of a digital image

Each pixel brightness is represented by a single 8-bit number which ranges from 0 to 255, this is how image data is formed and how the computer sees an image.



Pixel data

Storing data somewhat appears to be 2D when it's displayed but in computer memory it's stored in a one dimensional linear list with addresses.



Face detection

Image data is data to our eyes but it's a photograph to a computer. For example: if you look at the dark and light parts of an image we can see different brightness of Darkness and lightness.

The image is first converted into grayscale because it is easier to work and takes lesser data to process. This algorithm first finds the face on grayscale images then it finds the face location in the color image.

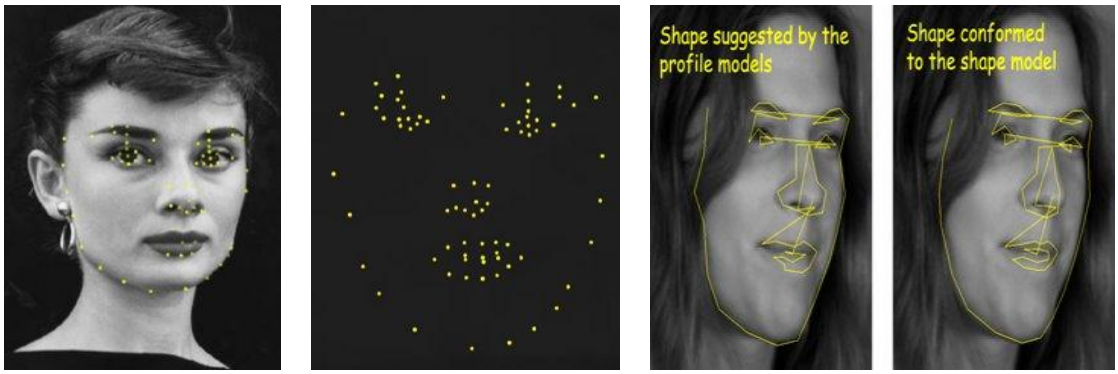
It outlines a rectangular box and searches for a face from the top-right corner.

This algorithm is designed to look for Frontal Faces**.** The performance will be poor when it comes to detecting side-ways, upwards, or downwards.

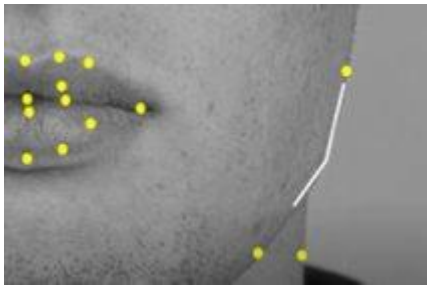
Active face model

Face model is a statistical model of a face trained by people who manually mark boundaries (points on picture) and then collect that data. This is done on hundreds and thousands of sample pictures.

Then an average face from this trained data is taken and is aligned with the face to perfectly find facial features. Then this suggested face shape with its facial features is made it resized to fit the face detected by the camera.



Analyzing pixel data these dots resize themselves to fit the facial features. The dots join edges to edges analyzing dark and bright spaces, creating a 3-D mesh like structure (a mask) for lenses to fit in.



Then finally the data comes in the form of video data where you can scan your face and put on lenses.

Viola-Jones Algorithm

This algorithm is the foundation of the OpenCV library. It was developed by Paul Viola and Michael Jones in 2001. It allows the detection of images in real-time. Algorithm works on grayscale images. It looks at smaller subregions and tries to find a face by looking for specific features. Later on, it check different positions and scales

There are two stages in this algorithm: 1. Training and 2. Detection.

Haar- like features

Haar-like features are digital image features used in object recognition.It is named after Alfred Harr, a Hungarian mathematician who developed the concept of Haar wavelets.

There are some similar features of human faces i.e. eye socket is darker area than adjacent area or the nose bridge is lighter region than adjacent area

In a simple way, we want to find the region which is darker or lighter and then sum the pixel value of both the region and compare.

There are two types of features: 1. Edge-like features 2. Line-like features and

Edge Features - It looks for one side that is light and another side that dark , for example - eyebrow

Line Features - It looks like a dark region surrounded by light region, for example - lips

Integral Images and Darkness Algorithm

The Integral Image concept helps to perform intensive calculations quickly.Each point in the integral image is a sum of the pixels above and left of the corresponding pixel in the source image.

We can also find average darkness 'A' by comparing 'A' of an area with the scale ; starting from 1 ranged to 10 with one being the lightest and 10 being the darkest.

Adaboost

The algorithm is setting a minimum threshold to determine whether something can be classified as a feature or not. But only some features are important in the face. Like eyes or lips or nose for the detection of facial features.

It finds the best threshold and classifies the face images into positive and negative. This training is done to detecting minimum error-rate which means they're features best classifying an image into face/ non-face image

We have an equation with f1,f2,f3 as features and a1,a1,a3 as their respective weights. F(x) is called a Strong Classifier. The single classifier will be weak, we get a strong classifier when we have a combination of two or three weak classifiers. This process is called an ensemble.

Cascading Classifiers

We take a subregion and the best feature , checking if it is present in the image or not.
If not we reject the subregion and move on to another subregion .
If it is present we look at the second feature , if it is not present we reject it and move on to another.

Cascading speeds up this process and the machine delivers results faster.

References

- **Computer vision**

https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e

- **Active shape model**

https://en.wikipedia.org/wiki/Active_shape_model

- **Facial detection**

https://medium.com/analytics-vidhya/creating-snapchat-like-filters-from-scratch-using-computer-vision-techniques-6374cde6a7db

- **Viola-Jones Algorithm**

https://www.mygreatlearning.com/blog/viola-jones-algorithm/#sh1

- **Adaboost**

https://hub.packtpub.com/implementing-face-detection-using-haar-cascades-adaboost-algorithm/

- **Cascading Classifiers**

https://www.sciencedirect.com/topics/computer-science/classifier-cascade