FACE AND VOICE RECOGNITION IN MOBILE DEVICES

Aishwarya Zamindar*1

*1IPS Academy, Institute Of Engineering And Science, Indore, India.

ABSTRACT

Face Recognition have become a new way for secure authentication for mobile phones. As mobile phones are becoming increasingly powerful, security of the data stored in mobile phones is a topic of concern; the data can be email addresses, sensitive and important documents, etc. Although, the current phones have password protection to provide security, a face recognition scheme is much more secure and flexible. In the face recognition feature, only by looking at the screen one can unlock their phone screen. Although figure print is the most used way of authentication in mobile phones, face detection is rapidly gaining popularity and acceptance. In this paper, we will focus on few algorithms such as color segmentation combined with template matching for face detection and KLT algorithm for face recognition. As we know that, face detection and face recognition mechanism depends on Biometrics system. Biometric is a measurable and unique characteristic of human being. It can determine both Physiological and Behavioral characteristics of human being. A person’s identity can be recognized and verified through this. The physiological biometric depends on the records obtained through the direct measurements of part of human body whereas the behavioral biometrics is based totally at the motion of human. A number of the physiological biometrics is facial recognition, finger-experiment, retina-experiment, iris-experiment, hand-experiment, and many others. And the special technologies associated with behavioral biometrics are voice-experiment, signature experiment, keystroke-scan, and so on. In latest networked international, cellular smartphone performs very vital function, it affects all components of human each day lifestyles. The need to keep the safety of information in mobile tool is becoming each increasingly important and more and more difficult. Some human features like finger prints, face, hand geometry, voice and iris are used to offer an authentication for protection structures to attain excessive security stage rather than traditional password primarily based systems. This paper presents a deployment of face popularity algorithms on mobile devices. Proposed approach uses pca set of rules with fpie and dcv on cellular device. On this paper all calculations performed on a mobile phone, in which a small quantity of pictures were used for trying out the device. Gadget accuracy is 92% for appropriately selected threshold, wherein the time taken to apprehend a face is approximately 0.35 sec and this can growth when database size increased.

Speech reputation (sr) is a area of pc technology that deals with recognition and conversion of spoken language into textual content by using computational gadgets. It carries research and methodologies from numerous other fields of computer technological know-how consisting of nlp, dsp. A few sr structures use schooling in which an man or woman speaker reads textual content or remoted vocabulary into the machine. The machine records this and analyzes that man or woman’s particular voice and uses it to first-class-track the popularity of that person’s speech, resulting in extended accuracy. Structures that do not use schooling are called “speaker impartial” systems whilst structures that use schooling are referred to as “speaker established”. Speech reputation has huge list of packages. One among its application is consumer input for cell applications. Google’s voice movements and iphone’s siri are programs that permit control of a cell cellphone the use of voice, including calling, businesses and contacts, sending texts and e mail, listening to song, browsing the net, and finishing common duties.

I. INTRODUCTION

In today’s networked world, mobile phone plays very important role, it affects all aspects of human daily life. The need to maintain the security of information in mobile device is becoming both increasingly important and increasingly difficult. Most of the current phones have security for Authentication. Authentication verifies that users or systems are who they claim to be, based on identity (e.g., username) and credentials (e.g., password). Mobile devices are easily lost or stolen; moreover Password could be easily hacked or detected. For that purpose high level of authentication for mobile devices is needed. The term Biometrics is becoming highly important in computer security world [1].The human physical characteristics like fingerprints, face, hand, geometry, voice and iris are known as biometrics[2]. These features are used to provide an authentication for
computer based security systems to reach high security level instead of traditional password systems [3][4][5]. Username and passwords can be replaced and/or provide double authentication by using any one of the biometric features. Face is the most suitable biometric that can be used for authentication in mobile devices since those devices usually have cameras. The incorporation of face recognition algorithms in mobile devices becomes a challenging problem due to the constraints on processing power and limited storage of mobile devices. In the field of pattern recognition; Dimensionality reduction is an important topic of research as in many practical technologies high dimensionality is a major cause of limitation. Also the large numbers of features degrade the performance of the used classifiers, especially when the size of the training set is small compared to the number of features.

In this work we have theorized the creation of an application that provides voice commands to call a person from contacts, send SMS messages, set alarms, turn on and off Bluetooth, Wi-Fi and GPS using Google speech recognition engine. The main goal of this application is to provide visually challenged persons with easy access to features of a smartphone. This will also set a precedent towards making smart phones completely touch free. In the Google API to be used in our application a large vocabulary speech recognition system is described that is accurate, has low latency, and yet has a small enough memory and computational footprint to run faster than real-time on an Android Smartphone. It employ a quantized Long Short-Term Memory (LSTM) acoustic model trained with connectionist temporal classification (CTC) to directly predict phoneme targets, and further reduce its memory footprint using an SVD based compression scheme. Additionally, it minimizes our memory footprint by using a single language model for dictation and voice command domains, constructed using Bayesian interpolation. Finally, in order to properly handle device-specific information, such as proper names and other context-dependent information, it injects vocabulary items into the decoder graph and bias the language model on-the-fly. This system achieves 13.5 percent word error rate on an open-ended dictation task, running with a median speed that is seven times faster than real-time.

1. BACKGROUND AND BASIC CONCEPTS

Face recognition is laptop/cellular based totally digital era and is been an active research location from the previous few years. To locate faces in a photo various strategies were proposed. Right here we've got noted template matching algorithm and color code segmentation using okay-mean algorithm for face detection and klt algorithm and viola-jones algorithm for face reputation. The following block diagram that is referred from face popularity in cellular phones through Guillaume Dave, Xing Chao, and Kishore Sriadibhatla shows the predominant steps in face popularity algorithm.

![Figure 1: Block diagram for face recognition system](image-url)
The facial recognition process normally has four interrelated phases or steps[10], which are:

- Face detection,
- Normalization,
- Feature extraction, and
- Face recognition.

These steps rely upon each different and regularly use similar techniques as proven in figure 2. They will also be described as separate components of a regular frs[11]. Detecting a face in a probe image can be a exceedingly simple project for people, however it is not so for a laptop. The laptop has to determine which pixels in the picture is part of the face and which are not. As soon as the face has been detected (separated from its historical past), the face needs to be normalized. Which means that the picture ought to be standardized in phrases of length, pose, illumination, and so on, relative to the photos in the gallery or reference database. After the face photo has been normalized, the feature extraction and recognition of the face can take place. In characteristic extraction, a mathematical representation known as a biometric template or biometric reference is generated, which is saved within the database and will form the premise of any recognition mission. Facial recognition algorithms range within the way they translate or transform a face photograph (represented at this point as grey scale pixels) into a simplified mathematical illustration (the functions) on the way to carry out the popularity venture.

**Figure 2**: Steps in face recognition

A. Android speech popularity

The speech recognition for the utility will be carried out with the aid of the usage of the magnificence furnished in android sdk, class speech recognizer. This class affords get admission to the speech reputation service. This carrier permits access to the speech recognizer. The elegance is instantiated with the aid of calling create speech recognizer (context). This elegance's techniques are invoked simplest from the principle application thread. This API isn’t always supposed to be used for non-stop reputation, which consumes a vast amount of battery and bandwidth. This API calls for to have report audio permission to use this elegance[2]. The magnificence recognizer intent is used for buying the speech to text conversion end result through intents. This newsletter is then used for in addition processing. The given “speech recognition” module is completed in five steps. The preliminary block is to take enter from user in the Shape of voice instructions. Then this is passed to the second block that is the API that converts speech into text. The Output of conversion block is the command in textual shape which is given as an input to the next block. Here enter is in comparison to the predefined listing of instructions from the database and the correct one is chosen. This command is then done inside the subsequent block. Inside the final output block the Command given via the person is saved as a log entry inside the Database so as to be later utilized by the machine getting to know module.
Steps in face popularity set of rules are as follows:-

1) Enter photograph: if the photo is taken successfully, the few assumptions can be made
   - The face is centralized and takes a large a part of the photograph, for the reason that photograph is shot closely.
   - The lighting situations are accurate.
   - The user have to face the camera.

2) Face detection: for an enter photo the first actual phase is the face detection phase. So, this segment has four steps that are as follows:
   - x8 down sampler or sampling technique.
   - color segmentation.
   - publish processing or morphological processing.
   - template matching.

K-mean Algorithm

Input
K: number of desired cluster
D: {d1, d2...dn} a data set containing n objects.

Output
A set of k cluster as specified in input Method

1) Arbitrarily choose k data item from D dataset as initial cluster centroid;
2) Repeat
3) Now, assign each data item di to the cluster to which object is most similar based on the mean value of the object in cluster;
4) Calculate the new mean value of the data items for each cluster and update the mean value;
5) Until no change. The third stage of face detection is morphological processing or post processi

In sampling system stage, the photograph sensor produces output as analog signal, it’s far not possible to do digital picture processing on that signal. So it is essential to convert analog signal to virtual signal. For this purpose sampling technique is executed at the input Picture. In color segmentation level, we use shade segmentation process to discover skin pixels of photo using k-mean set of rules which Solves clustering problem. The following information supplied underneath depicts the concept of the k-suggest set of rules which is being Referred from a review of ok-suggest set of rules via Jyoti Yadav, Monika Sharma which was published in worldwide magazine of Engineering tendencies and technology (ijett) – quantity four on 7- july 2013.
The “gadget studying” module presents a couple of functionalities. The primary and primary is the automated execution of very common commands given by the consumer with users permission. The module makes use of the command log generated by the speech reputation module to search for patterns, regulations and learn frequently used instructions. It’ll then ask consumer permission to execute them routinely. Instance is that if a person sets an alarm for 6:00 am frequently via voice commands, then the module acknowledges this as a frequent project and could now ask the user to set the alarm next time automatically. Some other functionality is to rely the stairs taken through consumer inside a specific time period using hobby popularity. This time period is defined by way of person enter . It takes start voice command from the user and prevents with the users stop voice command. This calls for sensor input from accelerometer. The drift of manage is depicted in fig. This module also notifies the user of buddies in the area. This is carried out by using attempting to find available paired bluetooth gadgets in users environment. This could require telling the gadget about the paired device. This information may be stored inside the software database.

B. Facial popularity in smartphones

Facial reputation is becoming the de-facto well known for unlocking telephones and sluggish removal of fingerprint sensors in Smartphones. The stairs it consists of are as follows:

All identity or authentication technology operate the usage of the following 4 stages:

A. Seize:
A behavioral or physical pattern is captured via the gadget throughout enrollment.

B. Extraction:
A template is created by extracting the unique information from the sample this is being captured and saved inside the phones database.

C. Evaluation:
Now, a new sample is in comparison with the template.

D. Suit/non-fit:
Right here, the machine comes to a decision if the features which can be extracted from the brand new samples are a healthy or non-suit.

Figure 4: Flowchart of k means algorithm
II. CONCLUSION

Face recognition has been taken into consideration as a secured although expensive application. Today because the center technology are evolving daily, the fee of the equipment's also are getting reduced dramatically because of integration and growing processing power. Sure programs of face reputation generation are now reliable, fee powerful and exceptionally correct. Face popularity duties has been proposed. This technique unifies the functionality of fuzzy set theory to gain the degree of belonging of different pixels of a face picture to distinctive lessons. Commonplace vector technique is obtained to lessen the range of samples utilized in education then conventional pca has been used for recognition challenge. This suggests the vast improvement in category accuracy and popularity time. Destiny paintings will beautify this approach by way of detecting outlier’s pixels to exclude distorted pictures from mastering technique to get higher recognition accuracy and performance. This paper theorizes the development of an android plat-form based clever voice popularity machine to perform multiple android apps with easy voice commands. This technology may be carried out in a user pleasant and compact tool. For the reason that software will perform in both on line and offline modes, it has a decided advantage over different such programs. This venture has the functionality of present day smart speech recognition software to boom independence for folks with disabilities. Main purpose of this paintings is to offer a device so that the visually disabled populace can without problems control many functions of a clever phone thru voice. The gadget could be very beneficial for the overall populace as nicely. Users can command a mobile tool to do something thru voice such as calling, texting, placing alarms. These instructions are then without delay completed. The utility is also the use of gadget gaining knowledge of concepts to execute frequent voice commands automatically. It is also keeping song of steps taken by way of person and coupled Bluetooth devices in place.

III. FUTURE SCOPE

The face recognition concept has an extensive scope within the destiny of era. It could be majorly used in security in addition to in advertising and marketing quarter. This app makes it feasible to improve first-class of looking in numerous methods. For instance, looking an image via someone’s name or the call of the location. Tagging an image at the time it is taken simplest as soon as, and then from next time whilst the photo of equal individual or factor is taken, tagging can be completed mechanically. It helps in saving time and gives a less difficult get right of entry to for looking images. The “device mastering” module presents more than one functionalities. The first and predominant is the automatic execution of very common instructions given by using the person with users permission. The module uses the command log generated by using the speech popularity module to search for styles, guidelines and examine frequently used instructions. It will then ask person permission to execute them mechanically. Instance is if a user sets an alarm for six:00 am often via voice
commands, then the module recognizes this as a frequent mission and could now ask the consumer to set the alarm subsequent time routinely.

IV. REFERENCES

[1] https://www.thenewsminute.com/article/over-billion-smartphones-have-facial-recognition-2020-research-76140