Person-Of-Interest Detection System Using Cloud-Supported Computerized-Eyewear

Xi Wang, Xi Zhao*, Varun Prakash+, Zhimin Gao+, Tao Feng, Omprakash Gnawali, Weidong Shi
Department of Computer Science, University of Houston, Houston, TX 77004, U.S.A
Email: {xiwang,tfeng3,larryshi,gnawali}@cs.uh.edu, xzhao21@central.uh.edu*, {vsprakash, zgao4}@uh.edu+

Abstract—Detecting Person-Of-Interest (POI), e.g., fugitives, criminals and terrorists in public spaces is a critical requirement of many law enforcers and police officers. In realty, most law enforcement personnel cannot effectively differentiate POIs from millions of faces and thus demand a portable assistant to recognize faces, in order to take the golden opportunity taking the POIs into immediate custody. Unfortunately, current face recognition systems are stationary and limited to a small scale of POI datasets. In this paper, we investigate a wearable computerized-eyewear based face recognition system. This system is a portable device which can accompany a police officer during patrolling or other tasks. The eyewear is connected to a cloud based face recognition system via wireless networks. Facial images captured by the mounted camera are sent to the cloud for identity retrieval. When the system finds a POI, it would alert officers via overlaying a virtual identity tag on the real POI's face on the transparent screen of the eyewear. We provide approaches to greatly minimize recognition time, including leveraging the large storage and high computational capacities provided by the cloud. The cloud enables nationwide POI database and supports parallel computing for face recognition.

Index Terms—POI detection, cloud computing, computerized-eyewear

I. Introduction

POI detection in the public places is performed widely by police and officers for purpose of law enforcement and homeland security. In many cases it has played a critical role in successfully identifying the POI, *e.g.,* criminals, terrorists or kidnapped and missing victims.

Nowadays, many biometrics techniques support for verifying or identifying POIs. They use unique bio-informatical features, e.g. face, iris, fingerprint and palm. However, it is not convenient for officers to use iris, fingerprint or palm modalities during patrolling. First, officers need the cooperation from people to obtain their iris or fingerprint scans. Second, it is time consuming and intrusive. Unlike using these modalities, face recognition is non-intrusive and able to detect multiple people at a time. Therefore, face recognition has become one of the most promising and powerful technological developments for the law enforcement agencies to detect potential POI. It is much more commonly-seen to deploy face recognition system than iris or fingerprint recognition system in public areas e.g. airport, subway, mall, street, bank and gas stations.

Face recognition techniques in public places are widely supported by surveillance cameras. But as reported in [1], surveillance cameras confront many problems. Firstly, they are normally installed at relatively undetectable spots, like on the ceiling or in the corner. Facial images caught from a top view angle increase the difficulty of face recognition. Secondly, they are usually fixed to one location. Criminals could study their locations, avoid appearing in the views of stationary cameras and plan their attacks accordingly. Thirdly, they are fairly easy to get destroyed. In addition, there are many simple ways to block the cameras. This gives criminals chances to completely evade or disable the face recognition system.

Considering the deficiencies of the current approaches, we propose a POI detection system using cloud-supported computerized-eyewears (e.g., Google glass). Specifically, the system consists of two major elements, a computerized eyewear as a client, cloud computation module as a server. The clients detect faces in front of an officer using the camera mounted on the eyewear, send them to the cloud [2],[3] where the faces are matched to the POI face database to identify and verify them. It will alert the officer when POI is detected, by overlaying information on the transparent screen of the eyewear. Essential features of our solution include: firstly, our system is wearable making it possible for the law enforcer to wear it during their daily tasks. Secondly, the cloud offers a large storage and high computational capacities. Our system can reduce the burden of memorizing large number of faces on law enforcers. Compared to largely-deployed stationary surveillance system, it is mobile and not limited to confined areas.

Overall, the major contributions of this study are following: (i) propose an advanced face recognition system using cloud supported computerized-eyewears; (ii) demonstrate the necessity of the computerized-eyewear aided face recognition for law enforcement via conducting a chasing experiment in a real-world scenario.

The rest of the paper is structured as follows: Section 2 introduces related works. Section 3 describes the system design, including the design of hardware and software components. Section 4 demonstrates the experimental results of the proposed system. Section 5 concludes the paper.

II. Related work

Our investigation of wearable POI detection system for law enforcement draws from researches in face recognition, its applications in law enforcement, Cloud computing and computerized-eyewear.

Face recognition: A face recognition system can automatically identify or verify a person from a digital image or a video frame from a video source, normally by comparing selected facial features from the image and a facial database [4]. Some face recognition algorithms identify facial features by extracting landmarks or features from an image of the subject's face. The algorithm, Dr. Bonsor introduced in [5], analyzes the relative position, size, and shape of the eyes, nose, cheekbones and jaw. These features are then used to search for other images with matching features [4]. Other algorithms normalize a gallery of face images and then compress the face data, only saving the data in the image that is useful for face recognition [4]. A probe image is then compared with the face data [5]. Popular recognition algorithms include Principal Component Analysis using eigenfaces [6], Elastic Bunch Graph Matching using the Fisherface algorithm [7], and Local Binary Pattern Histogram (LBPH) extraction of the face feature [8]. In addition, some companies or labs have built up face recognition systems, e.g. Lambda [9], Google's Picasa [10], Apple's iPhoto [11].

A newly emerging trend, claimed to achieve improved accuracies, is three-dimensional face recognition [4]. 3D face data contain more information of the face, like surface information, and are pose invariant [12]. For example, Prof. Ioannis, at University of Houston, has developed a fully automatic system which uses a composite alignment algorithm to register 3D facial scans with a 3D facial model, thus achieving complete pose-invariance [13]. Face recognition based on 3D data can achieve significantly higher accuracy than its 2D counterparts. But so far, 3D face recognition has not been widely utilized in POI detection due to the relatively high cost and the long computational time [14]. Meanwhile, one of the major disadvantages of 3D face recognition is to capture 3D face data. It requires cooperation from a subject. The subject is required to be at a certain distance from the sensor and completely immobile for a few seconds.

We choose LBPH [15] [16] to represent a facial image [17]. Because it is an illumination-robust feature [18] and thus can work reliably under different illuminations. Moreover, unlike some other methods e.g. Fisherface and Eigenface based methods [19], LBPH based method supports update of the face recognizer. Thus, it can save running time and memory space for training.

Face recognition in law enforcement: Face recognition is one of the most powerful new technologies being used by law enforcement [20]. In [21], James Wayman, the former director of the National Biometric Test Center at San Jose State University, described the past, present, and future of face recognition in law enforcement. In [22] and [23], discussions on face recognition in finding Boston bombers, show the important role of high quality of pictures and enough pictures in database in accuracy of face recognition. Moreover, a few news demonstrate the resolution police and FBI made to improve and expand the usability of face recognition technology. For example, Chicago police start using facial-recognition software to arrest suspects [24]. Police in 26 states use face recognition technology on driver's licenses and other IDs to aid crime solving [25]. And the federal government is making progress on developing a surveillance system that would pair computers with video cameras to scan crowds and automatically identify people by their faces [26].

Cloud computing: Cloud computing is a rapidly evolving information technology that is gaining remarkable success in recent years. It uses a shared pool of virtualized and configurable computing resources (both hardware and software) over a network to deliver services, such as to host and analyze large datasets immediately. As described in [27], one of the most popular programming models in the cloud is MapReduce [28], which is for distributed processing of large-scale data on clusters of commodity servers. Ananthanarayanan et al. [29][27] proposed an optimized cluster file system for MapReduce applications. They use metablock that is a consecutive set of blocks of a file that are allocated on the same disk instead of the traditional cluster file system. Apache Pig [30][27] is a platform for analyzing large data sets using MapReduce on the top of Hadoop.

computerized-eyewear: Steve Mann, the computerized eyewear pioneer, has been designing and wearing computerized eyewear for decades, the gear increasing markedly in sophistication over time. In [31], he introduced the development of a serial of his wearable computers in his life. Google Glass is a wearable computer with an optical head-mounted display (OHMD) that is being developed by Google [32]. It displays information in a smartphone-like hands-free format, that can communicate with the Internet via natural language voice commands [32]. The STAR 1200XL used in our system, is the third generation of the see-through augmented reality eyewear system. It enables you to see the real world directly through its transparent widescreen video displays. Computer content, such as text, images and video, are overlaid on the screen [33]. It is similar to Google Glass in terms of assembly, functionality and operation.

III. System Design

In this section, we present the hardware and software components of our system, as depicted in Fig.1.

A. Hardware

Our face recognizer cloud consists of over 20 virtual machines. The setup includes 64-bit Ubuntu OS, 2 GB of memory, 250 GB of hard disk and two virtual cores. Each virtual core is equivalent to a 1.6 2.0 GHz Xeon processor. The Apache Hadoop-0.20.203.0 and MapReduce service are provided by the cloud.

The eyewear functions as a data collection and also contains an information display unit. As depicted in Fig. 2, this Vuzix STAR 1200XL third generation augmented reality device comprises the following individual units, a forward pointing camera, an eyepiece including a LCOS (Liquid crystal on silicon) micro display and an interface/driver board that interfaces the components of the eyewear and an eyepiece adjustment mechanism to suit a variety of users. The driver
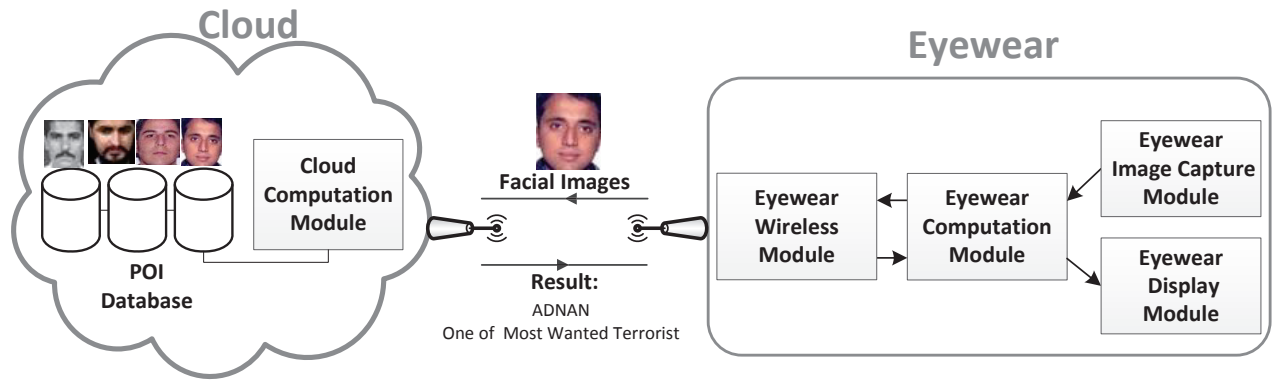
Fig. 1. The face recognition system overview. The system consists of a computerized-eyewear (the client) and the face recognition Cloud (the server). These two components are connected by a wireless network.
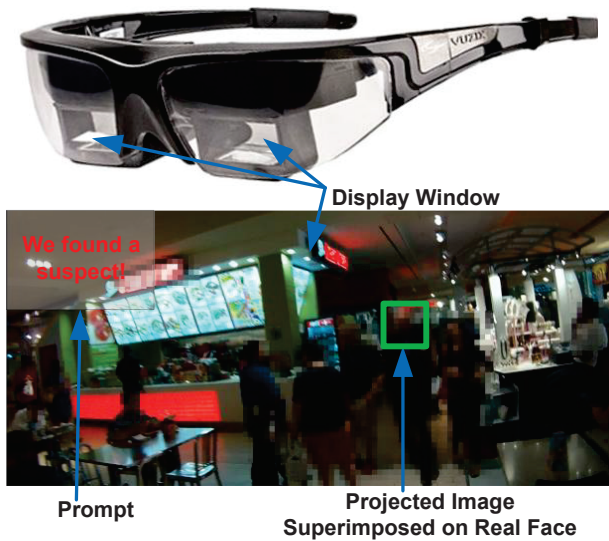

Fig. 2. The figure on top shows the location of the display window in the eyewear. The bottom figure demonstrates a sample view when the system has detected a suspect. We blurred people in the figures for privacy protection.

Fig. 3. The flow chart of the system software

board interfaces the camera, the LCOS microdisplay and the computation module of the eyewear in order to overcome the differences in formats of data transferred between these entities. The mounted camera records the video from the user's viewpoint and sends it to the driver board, which buffers data and sends all or a configured number of frames to the computation module.

B. Software

The software application used in the project consists of two components as shown in Fig.3: (1) configuration settings on eyewear and face recognizer training on Cloud in the offline phase, and (2) face recognition algorithm in the online phase. The two parts are explained as follows:

1) Offline: POI database consists of facial images of individuals who are wanted by law enforcement officials and organizations. First of all, user collects images of POIs from the existing databases, e.g. FBI Kidnappings and Missing Persons [34], FBI Most Wanted Terrorists List [35], U.S. Immigration and Customs Enforcement Most Wanted Fugi-
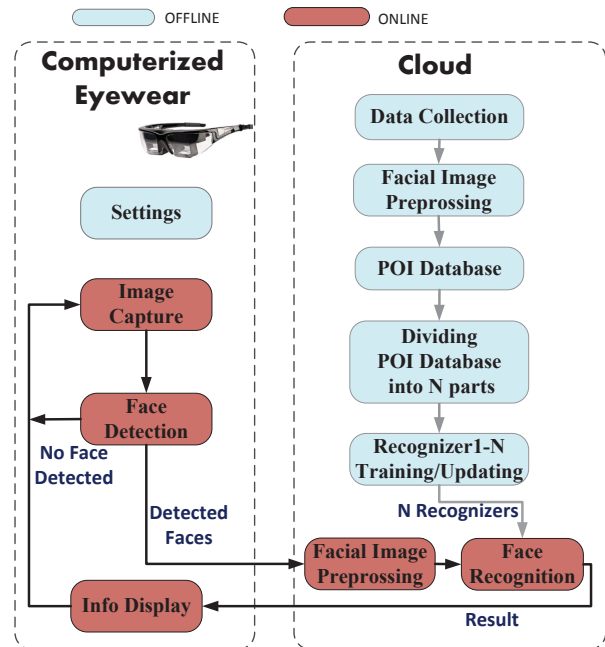
tives [36]. The images found in these data sets are complex images containing unnecessary background along with the foreground of the frontal faces. This complexity can lead to the inefficient performance of the training and recognition modules. Thus, it is important to prune images as per requirements. The facial areas are detected by Viola-Jones method [37] in images automatically. Then we align all facial images by normalizing their sizes into 100×100 pixels and converting color facial images to greyscale for simplicity and computation efficiency. User can add or edit extra information about a POI other than their identities, e.g. criminal record or missing people record, which are also saved in POI database along with the POI's preprocessed facial images. These extra information about the POI can also be displayed on eyewear when the user meets the POI.

In order to avoid extracting LBPHs from all facial images in a database every time, we normally store them in a recognizer (an xml format file). The recognizer training is the process

of facial feature extraction and storage in form of LBPHs. We conduct face recognition in the cloud which supports running the tasks in parallel. The whole database consisted of M images are divided into N sub-databases, each stored on one node respectively. N recognizers are trained on them respectively with roughly M/N images. When an new image arrives, its features are extracted and compared to different recognizer parallelly. We refer to dissimilarity between two faces in comparison as "distance". Each node calculates M/N distances between the new image and its sub-databases. In Fig.4, we can see the N recognizers located in N nodes. We will explain how the parallel face recognition algorithm on Cloud operates with the N recognizers in Subsection III-B2.

User can train a new recognizer with new data, associated labels and descriptive information, and replace the old ones if they exist. Furthermore, user can update the existing face recognizer by registering the new subject data into the existing recognizer, instead of retraining a new recognizer with both old and new data.

User can configure settings using the applications user interface on the eyewear: (1) minimum-window-size for face detection to set the minimum size of face which can be detected, i.e. the minimum resolution of facial image which would be fed into face recognition; (2) a verification threshold to set how easily an unknown face can be verified as a POI who appeared in the POI database; and (3) what information should be displayed on the eyewear in the online phrase when a particular POI's face is detected and recognized.

2) Online: The online stage consists of a cooperative computation task undertake by the eyewear and the central computation module in the cloud. The application on the eyewear detects faces on the image sent from the camera. If no face is detected in the current image, the application will simply discard that frame and continue its detection for the new incoming frames. If faces are detected and positions of them are located in the image, the application crops the detected faces. Then it performs a wireless transmission of the image of the face to the Cloud. Some of the preprocessing steps that has been discussed in the previous, offline stage will be applied to the image inorder to improve its recognisability.

Figure 4 is a pictorial representation of the distributed face recognition performed in the cloud. The well known MapReduce models are used to divide the tasks and merge the interim solutions to arrive at the final solution. The preprocessed query image is replicated and fed into N mapping instances at the same time. In each instance, the application extracts the features and matches them against the chunk of the database that is associated with the particular task. This chunk of the larger database contains M/N face images and thus will lead to the generation of M/N distances. The output of each map instances will be a particular face ID or the PID and a confidence value deduced by the distance. Among the many distances that have been obtained, the minimum distance after comparison or the comparison that instills the highest confidence is chosen and the respective PID is reduced as the result. This is to say that the image with the highest
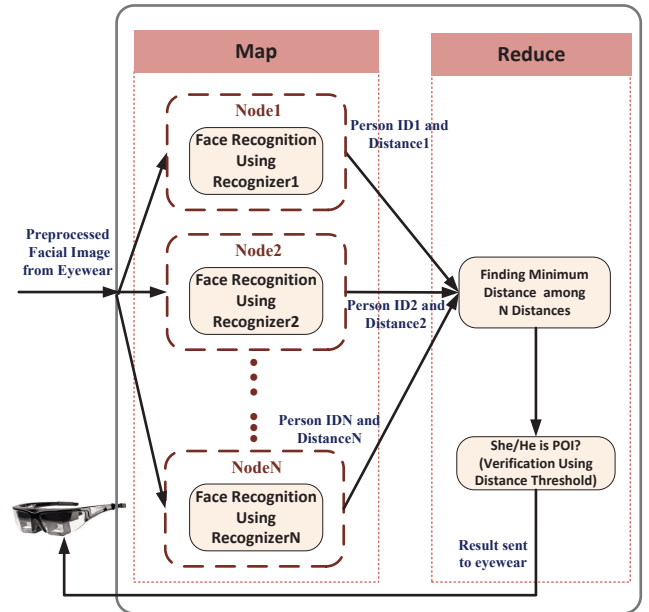

Fig. 4. The parallel, distributed face-recognition algorithm on Cloud. It can significantly reduce time cost of face recognition by parceling out work to different nodes.

confidence has the highest number of matched features and poses the highest probability that the query image belongs to a person who is wanted by the law enforcement agency. In order to to reduce the rate of false negatives, suitable thresholds are chosen to make sure that the system provides flexibility in stringency of comparison. The information relating to the matched POI is then sent to the eyewear unit and the information is displayed for the user.

IV. System Evaluation

In this section, we report and analyze experimental results of our prototyping system based on the approach detailed in the last section.

To demonstrate the necessity of the proposed system, we conducted an experiment in a real life scenario comparing the ability of recognizing strangers between human and machine. The experimental time was between 2pm to 3pm and the location was one level of a shopping mall in a major city of U.S.. Four participators were divided into two groups, two as pursuers and another two as evaders. Note that pursuers have never met evaders before the experiment and they were given eight facial images of eight strangers respectively, including the two evaders. Other than this information, pursuers have no information on the evaders (*e.g.,* gender, race or height). During the experiment, one pursuer and one evader participated in each round. The evader was walking around acting natural. Meanwhile the pursuer patrolled all the hallways and corridors until she/he finds the evader or time is up (30 minutes). The pursuer was requested to record the time when s/he found the evader. The whole process had been recorded by a head-mounted camera worn by the pursuer.

As seen in Table I, in the first round, the pursuer passed the evader four times without recognizing him while our system

|		Human	Machine	Note
First round	Evader appeared at 2:16pm as seen in Fig.5(a)	X	X	Both of human and machine did not detect evader. Machine failed because it mistook the evader for another evader.
	Evader appeared at 2:28pm as seen in Fig.5(b)	X	√	Human did not detect evader, while machine did.
	Evader appeared at 2:46pm as seen in Fig.5(c)	X	√	Human did not detect evader, while machine did.
	Evader appeared at 2:53pm as seen in Fig.5(d)	X	X	Both of human and machine did not detect evader. Machine failed because evader did not show face.
Second round	Evader did not appear	X	√	Human falsely recognized a shopper as the evader, while machine did not give false alarm.

X: Fail; √: Succeed.
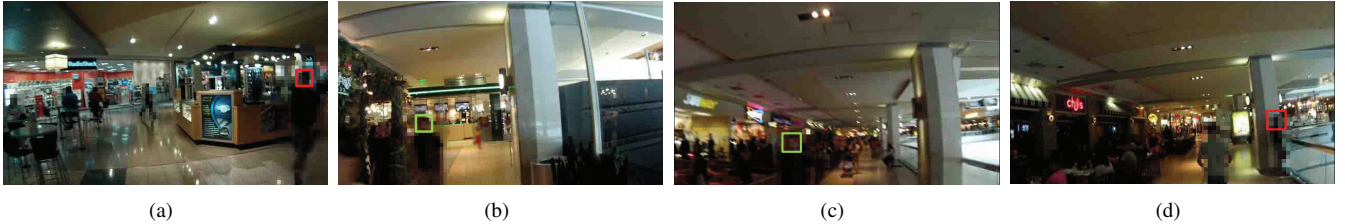
Fig. 5. In the first round, evader appeared in the video four times. Pursuer failed to find the evader, while our system found the evader twice as seen in (b),(c) where evader's face was tagged by green frame. Evader tagged by red frame in (a),(d) was not found. We blurred people in the figures for privacy protection.

succeeded in detecting evader twice. In the second round, the pursuer and the evader did not meet, but the pursuer falsely recognized a shopper as the evader once. In order to fairly compare the machine's face recognition ability to human's, we have built a recognizer trained from the eight photos exactly same as those given to the pursuers. It has been used to analyze the recorded videos and succeeded twice, as depicted in Fig. 5. Though the scale was limited in the experiments in terms of number of participants and gallery size, it still suggests the limited ability of human to recognize strangers from crowds. This gets even worse when officers try to identify thousands of POI in the lists like Most Wanted Criminals and Fugitives Lists [38]. Taking portable computerized devices to support face recognition is necessary. These portable devices can not only record faces with frontal pose more close to facial images on their identity cards or POI lists, but also cover all places where officers patrol. Thus, the computerized eyewears have their advantages over the surveillance cameras on identifying POIs. Cloud techniques and portable devices enable face recognition techniques on a large scale.

To evaluate the performance of proposed Cloud in the system, the video dataset named XD2013 is collected. It consists of 41 sessions of video data collected from 20 subjects. We

recorded these 41 videos with the camera mounted on the eyewear at a frame rate of 20 fps and a resolution of 640×480. Every video lasts from 20 to 40 seconds. We obtained the 10000 facial images with a resolution of 100×100 by cropping off facial images from frames of XD2013.

We deployed the same set of recognizers, which were extracted from the the 10000 facial images, to a cloud with different number of nodes from 1 to 20. We used the experiment conducted only on one node as reference. Figure 6 depicts the time cost of face recognition when the number of nodes used in the experiment increases from 1 to 20. It can be observed that with more nodes involved, the face recognition time is reduced from $33748ms$ to only $202ms$, i.e., 99.4% of time is saved.

V. CONCLUSIONS

This paper proposes a wearable face recognition system for law enforcement officials and organizations. Our system allows law enforcers to conduct POI detection anywhere he wears the computerized eyewear. Supported by the cloud computing technology, our system can store large POI database and conduct large scale POI detection. Time of face recognition can be minimized to 0.60% by employing parallel cloud based
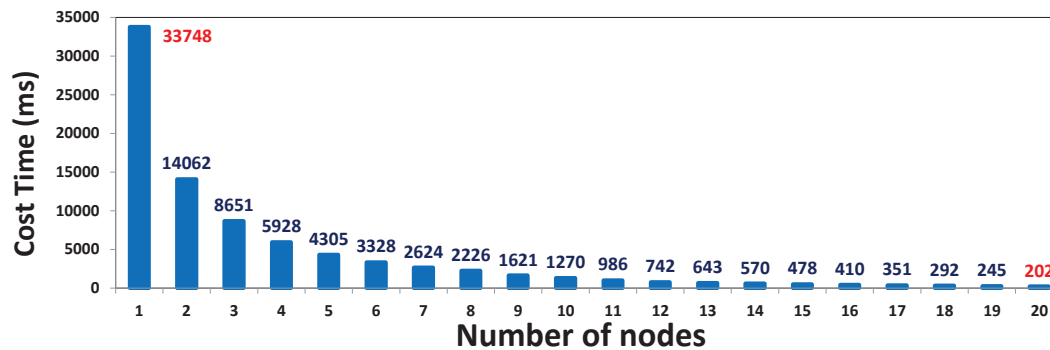
Fig. 6. Time cost of face recognition under different number of nodes in the Cloud. It depicts our parallel face-recognition algorithm in cloud can save up to 99.4% of time cost.

processing. Compared to humans, the system has the advantage of rapid detection of POIs from a large POI database. The experimental results demonstrate performance of the proposed technologies and the overall usability of the system.

REFERENCES

[1] UW Computer Security Reseach and Course Blog , "Surveillance cameras security reviews," January 2008.

[2] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, and M. Zaharia, "A view of cloud computing," *Communications of the ACM*, vol. 53, no. 4, April 2010.

[3] I. Foster, Y. Zhao, I. Raicu, and S. Lu, "Cloud Computing and Grid Computing 360-Degree Compared," in *Grid Computing Environments Workshop*, Austin,TX, November 2008, pp. 1–10.

[4] (2013, August) Facial recognition system. http://en.wikipedia.org/wiki/Facial_recognition_system.

[5] K. Bonsor. (2008, June) How Facial Recognition Systems Work.

[6] M. Turk and A. Pentland, "Eigenfaces for recognition," *Journal of Cognitive Neuroscience*, vol. 3, no. 1, pp. 71–86, December 1991.

[7] H.-J. Lee, W.-S. Lee, and J.-N. Chung, "Face recognition using Fisherface algorithm and elastic graph matching," in *International Conference on Image Processing*, vol. 1, Thessaloniki, October 2001, pp. 998–1001.

[8] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *ICCV 2001 Workshop on Statistical and Computation Theories of Vision*, vol. 28, no. 12, October 2001.

[9] Lambda Labs Face Recognition API. http://lambdal.com/docs/start#tab-Python.

[10] Picasa. http://picasa.google.com/.

[11] Iphoto. http://www.apple.com/ilife/iphoto/.

[12] A. F. Abate, M. Nappi, D. Riccio, and G. Sabatino, "2D and 3D face recognition: A survey," *Pattern Recognition Letters*, vol. 28, no. 14, October 2007.

[13] I. A. Kakadiaris, G. Passalis, M. N. Murtuza, Y. Lu, N. Karampatziakis, and T. Theoharis, "Three-Dimensional Face Recognition in the Presence of Facial Expressions: An Annotated Deformable Model Approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 4, April 2007.

[14] "Advantages and disadvantages of 3D face recognition," January 2008.

[15] T. Ojala, M. Pietikinen, and T. Menp, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002.

[16] T. Ojala, M. Pietikinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.

[17] T. Ahonen, A. Hadid, and M. Pietikainen, "Face recognition with Local Binary Patterns," in *European Conference on Computer Vision*, vol. 3021, May 2004, pp. 469–481.

[18] C. ho Chan, J. Kittler, and K. Messer, "Multi-scale local binary pattern histogram for face recognition," in *Proceedings of the 2007 international conference on Advances in Biometrics*, vol. 4647, 2007, pp. 809–818.

[19] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman, "Eigenfaces vs. fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, July 1997.

[20] Face Recognition in National Fusion Centers and Law Enforcement.

[21] S. Cherry, "Will Face Recognition Ever Capture Criminals?" *IEEE Spectrum*, May 2013.

[22] D. McCormick, "Face Recognition Failed to Find Boston Bombers," *IEEE Spectrum*, April 2013.

[23] J. C. Klontz and A. K. Jain, "A Case Study on Unconstrained Facial Recognition Using the Boston Marathon Bombings Suspects," pp. 1–8, May 2013.

[24] RT. (2013, July) Chicago police start using facial-recognition software to arrest suspects.

[25] A. Dugdale. (2013, June) Police Use Facial-Recognition Technology On Driver's Licenses And Other IDs To Aid Crime Solving.

[26] C. Savage. (2013, August) Facial Scanning Is Making Gains in Surveillance.

[27] Y. Wen, X. Man, K. Le, and W. Shi, "Forensics-as-a-Service (FaaS): Computer ForensicWorkflow Management and Processing Using Cloud," in *The Fifth International Conferences on Pervasive Patterns and Applications*, Spain, May 2013, pp. 1–7.

[28] J. Dean and S. Ghemawat, "Mapreduce: simplified data processing on large clusters," *Communications of the ACM*, vol. 50, no. 1, January 2008.

[29] R. Ananthanarayanan, K. Gupta, P. Pandey, H. Pucha, P. Sarkar, M. Shah, and R. Tewari, "Cloud Analytics: Do We Really Need to Reinvent the Storage Stack?" in *Proceedings of the 2009 conference on Hot topics in cloud computing*, CA,USA, 2009.

[30] (2013, April) Apache Pig. http://pig.apache.org//.

[31] S. Mann, "Steve Mann: My Augmediated Life," *IEEE Spectrum*, March 2013.

[32] (2013) Google glasses. http://en.wikipedia.org/wiki/Google_Glass.

[33] Wearable glasses technology. http://www.vuzix.com/augmented-reality/products_star1200xl.html.

[34] FBI Kidnappings and Missing Persons. http://www.fbi.gov/wanted/kidnap.

[35] FBI Most Wanted Terrorists List. http://www.fbi.gov/wanted/wanted_terrorists.

[36] U.S. Immigration and Customs Enforcement Most Wanted Fugitives. http://www.ice.gov/most-wanted/hsi.htm.

[37] P. Viola and M. Jones, "Robust real-time object detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 12, pp. 2037–2041, December 2006.

[38] American Screening Homeland Security Search. http://www.americanscreening.com/files/asset/mydocuments/HomelandSecurityMemo.pdf.