Location-based Face Recognition Using Smart Mobile Device Sensors

Nina Taherimakhsousi, Hausi A. Müller
Department of Computer Science
University of Victoria, Victoria, Canada
ninata@uvic.ca; hausi@uvic.ca

Abstract– Face recognition is a challenging computational task which assigns identity to detected faces. There has been prodigious improvement on face recognition in recent years. At first, researchers focused on face recognition under controlled conditions. Facilitated by classic datasets, researchers investigated face representations invariant to changes in pose, facial expression, and illumination. Despite their initial success, substantial face recognition research is required in less-controlled or uncontrolled conditions, such as in personal photo collections, web images and videos, or on-line social networks. Searching over a large set of images amplifies the need for more robust face recognition methods. In this paper, we propose a location aware face recognition application that alleviates the challenges of searching over a huge set of images in uncontrolled conditions. By using location sensor information of mobile devices, we create databases with location information. Furthermore, we propose a method to reduce the recognition search space, resulting in higher accuracy and real-time performance for identification.

Keywords: Context-Aware, Location-Aware, Face Recognition, Smart Mobile Device Sensors

1. Introduction

Face recognition is a task that humans accomplish customarily and effortlessly during social interactions. The ubiquity of personal computers and high powered embedded computing systems has created tremendous interest in automatic image and video processing in a variety of applications, such as biometric authentication, surveillance, recommendation systems, as well as human computer interaction. Over the last twenty years, face recognition has become a popular area of research in computer vision and one of the most successful applications of image analysis and understanding. However, it still faces many challenges. A single person can have many facial expressions which are difficult to match with different glasses and illuminations which can cause significantly different facial features. For a system that can answer a query of recognizing a person in an image, searching over databases and comparing faces will amplify all these difficulties and create face recognition problems.

Humans recognize faces in social interactions based on environment and social parameters subconsciously. The relevant context includes information related to the image of the scene surrounding a person. This includes camera context such as location information and image capture time as well as the social context that describes the interactions between people which helps us identify faces in daily social interactions. Without context, even humans may fail to recognize the observed face. Hence, an important related issue that should be addressed in automatic face recognition is how to take context information into account in real-world face recognition applications.

With the proliferation of smart mobile devices, such as phones and tablets, people are increasingly instrumented. These devices can track locations and moving directions of their users. Some even have sensors for temperature and humidity which in combination with microphones enables improved determination of locations and their surroundings. Furthermore, many of us use them as replacement for more sophisticated cameras because smart mobiles are easy to carry and convenient to use. A study showed that of the photos taken in 2011 smartphones took 27% and regular cameras 44% (Web-1). In Yahoo's Flickr, a place where people store and share their digital photographs, iPhone 5 was responsible

for more photos posted to this site than any other device in 2014 (Web-2). Also, growing online mobile photo-sharing services, such as Instagram, enable its users to take pictures and videos, and share them on a variety of social networking platforms using mobile devices. Photos taken with these smart devices include sensor-based context information, such as location, time, humidity, temperature, and acceleration pertinent to the moment when the photo was captured. The annotated measures of, smart mobile device can be exploited to sensors aid face recognition tasks.

In this paper, of all the available context information we concentrate on location context and investigate how to exploit location information in face recognition problems. Our approach is to reduce the search space of the recognition problem by discernment of the user location. If a user took photos at a certain location, there is high chance that the user will be in the image which is taken at this location. Consequently, when we want to recognize a face, we can save some effort and gain accuracy by only comparing faces which are taken at locations that where that face normally appears.

In this paper we present three main research contributions. First, we present the first treatment of location information as a facial attribute and propose a hybrid face recognition algorithm. Our approach includes a searching and matching method. If a search fails within the current location, then it will search over an extended database. Second, we show that our approach is surprisingly accurate by taking into account social network information. When we recognize a user appears in a particular location, her social friends will likely show up in that location too. Afterwards, social friends' images are used to train the face classifier for each location. Third, we utilize Future Internet nodes on the Smart Applications on Virtual Infrastructure (SAVI) network to categorise the database and dispatch the query which reduces network traffic and in turn the response time.

The remainder of this paper is organized as follows: Section 2 presents our location-based face recognition system as well as our location-based recognition method and algorithm. Section 3 explains our experimental setup and highlights selected accuracy critical parameters in the classification stages. Section 4 reports on our evaluation and result. Finally, Section 5 concludes the paper and discusses avenues for future research.

2. Location-based Face Recognition Approach

This section introduces our context-aware location-based face recognition system and provides a detailed description of our method for using location information within our proposed algorithm.

We propose location-centric image databases to recognize faces in images that have been taken at or nearby locations frequently visited by these individuals. This face recognition problem is defined as follows: Given a set of known images of faces for training and a set of the same set of people as a testing set, recognize each face in the testing set. In our proposed system, each face image associates with the location information. The system creates many clusters of locations from the training set. Each location cluster contains a set of users who have images in that location and images of their friends. The user can take an image and attach the location information, then send it to the system and query for the face in the photo. The system will answer the recognition question and return the identification to the faces in the image.

Important related issues and research questions that should be addressed in location-based face recognition include how to form the location categories, how to take location information into account in the feature extraction processes, how to search efficiently to recognize the faces and how to advance the recognition steps, and minimize response times by using advantage of Future Internet nodes such as SAVI core and edge nodes (Kang et al., 2014).

2. 1. Overview

Fig. 1 illustrates our location-based face recognition approach. Using a smart mobile device the user sends images and associated recognition queries to SAVI network nodes for processing (Web-3). The features extraction stage is performed on the user smart mobile device. Extracted features and location information are sent to the SAVI server for face recognition. On the SAVI server, we categorize the location-based database.
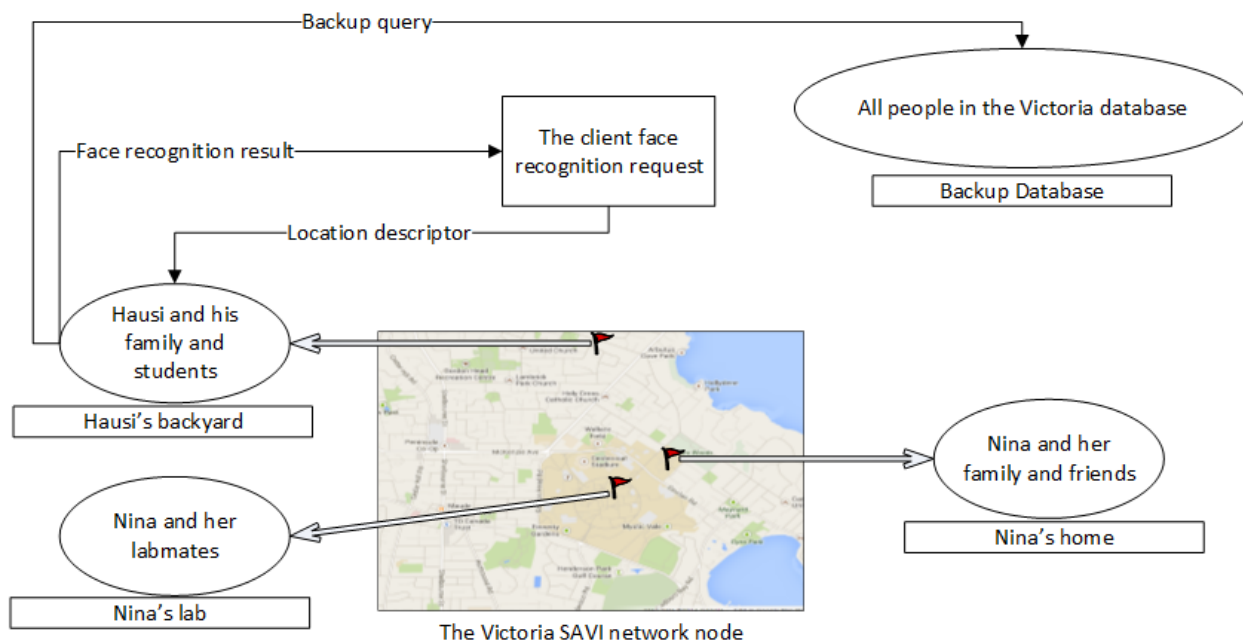
Fig. 1. Proposed location-based face recognition system approach.

The database at each location includes images of faces that have previously appeared at that location as well as social network friends' images. Because social friends are more likely to visit an individual in the database, we can insert images of those friends and associate them with the individual's location. We also maintain a backup database which represents a collection of all images in different databases on the SAVI network nodes as shown in Fig 2. Should the system fail at any location database, we active the backup database to recognize the face image.
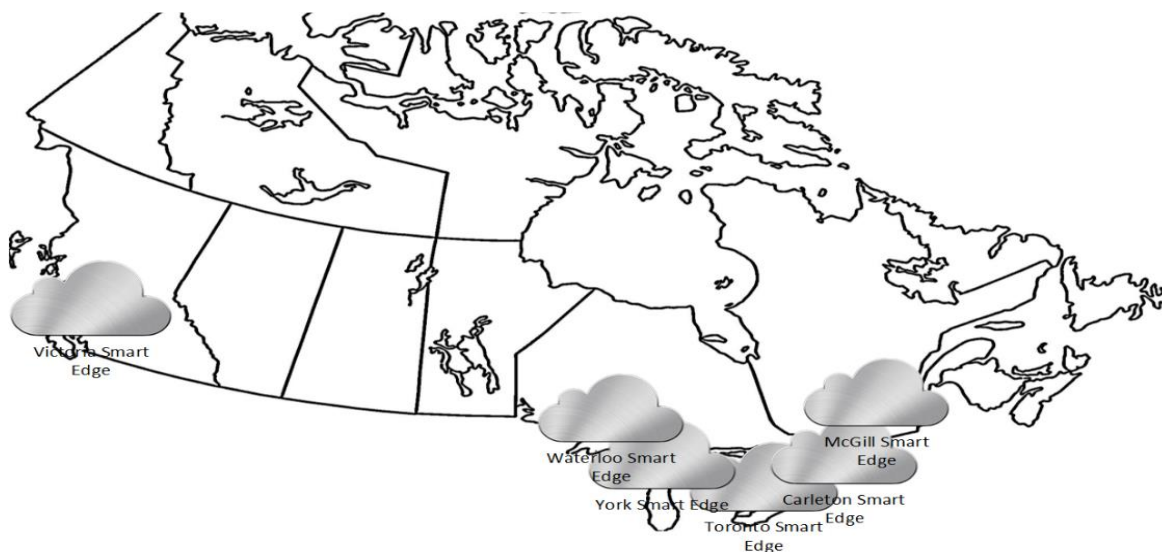

Fig. 2. Backup databases on smart nodes in the SAVI networks (Leon-Garcia, 2011).

2. 2. Categorizing Location Information

We have a database of training images based on location information which represent each location point with its longitude and latitude. Initially, the database contains *n* points corresponding to *n* different categories. At each step of database categorization, we merge the two categories if the distance between two categories is the minimum among all pairs of categories. The distance between two categories *X* and *Y* is defined in Eq. (1):

$$d(X,Y) = \frac{1}{|X||Y|} = \sum_{x \in X} \sum_{y \in Y} d(x, y) \tag{1}$$

Where $d(x, y)$ is the Euclidean distance between point *x* and *y*. We keep merging categories until the minimum distance in each iteration is above a threshold or the number of categories we want to obtain is reached (Li et al., 2004).

2. 3. Features Preparation

Features are associated with the location information for training at each SAVI network node classifier. We used the Viola-Jones face detection method which is an efficient face detection method suitable for real-time applications to detect faces from user images (Viola et al., 2004).

Detected faces are normalized to the same size. We employed the algorithm presented by Taherimakhsousi et al., to extract face feature from each face image which is efficient and suitable (Taherimakhsousi et al., 2009). Moreover, utilizing the method shown by Tan et al., we reduce the lighting effects on face features (Tan et al., 2010).

Table 1. Location-based Mixture of Experts (MoE) classifiers accuracy rates obtained through
5-fold cross validation at nine different locations.

Locations	Test 1	Test 2	Test 3	Test 4	Test 5
Location 1	0.889	0.901	0.891	0.875	0.846
Location 2	0.891	0.923	0.933	0.941	0.877
Location 3	0.965	0.938	0.942	0.968	0.989
Location 4	0.687	0.736	0.745	0.684	0.692
Location 5	1.000	1.000	0.975	1.000	1.000
Location 6	0.766	0.741	0.795	0.689	0.690
Location 7	0.625	0.632	0.687	0.692	0.627
Location 8	0.535	0.516	0.576	0.594	0.581
Location 9	0.715	0.735	0.749	0.754	0.752

3. Experiment Setup

Fig. 3 illustrates how we deployed our location-based face recognition system on the SAVI network which includes two parts, the smart mobile device user and the SAVI network edge.

Face features are extracted from images and trained with the Mixture of Experts (MoE) classifiers which is described in (Taherimakhsousi et al., 2009b) using the neuro fuzzy classifiers on the SAVI network node. Each location has its own expert classifier. When face features are received on the SAVI

network node, it rates the location of the received face features and searches the closest expert classifier in the database on the SAVI network. The accuracy of each location-based expert shows how well the face image is classified by MoE. If the accuracy is below a certain threshold, we send that face feature to the backup database to classify over the backup database on the other SAVI network nodes.

Our database contains 1,000 images from 100 labeled faces. The dataset has nine locations. Our location-based face recognition system includes the labeled faces and the social network relations as the training set. Further, each location is included with face images of friends from a user's social network. We also create a backup database containing images of those face images on the SAVI network nodes, which is exploited should the local SAVI network node fail to recognize the face based on recognition thresholds.
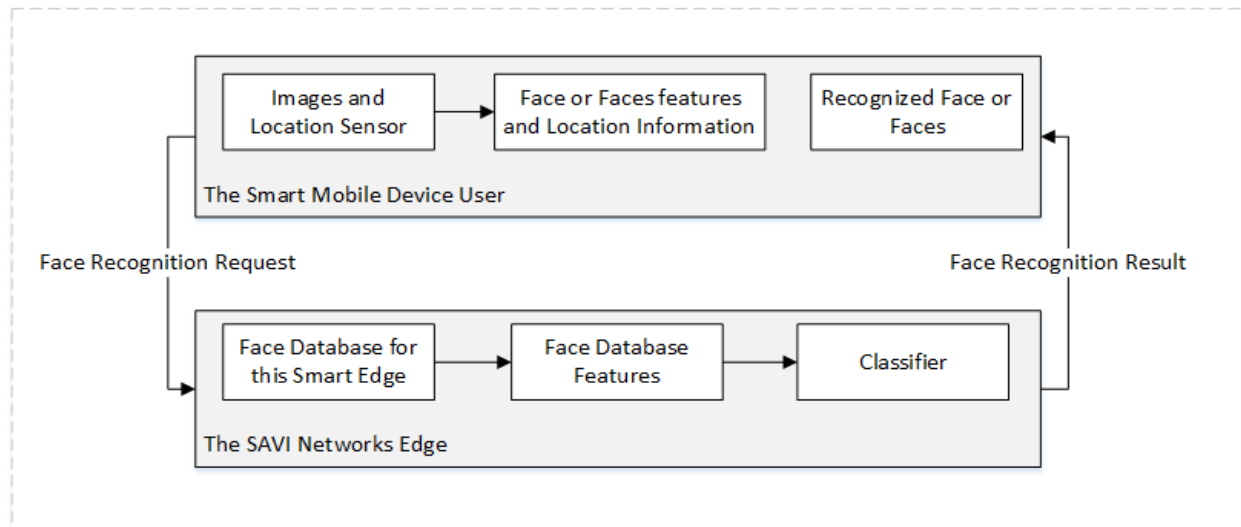

Fig. 3. The proposed location-based face recognition system the user and the SAVI network node.

4. Results and Discussion

We conducted a five-fold cross validation to test the proposed location-based method. In each test, 80% of the images ware used as the training set and 20% of the images as the testing set. We compared our approach with the MoE method which uses no context information. Our location-based method improves the accuracy from 0.635 to 0.885. The validation results are presented in Table 1.

The complexity of the recognition algorithm is $O(log\ k)$, where k is the number of locations in our location-based face recognition system. The complexity of recognizing the MoE is $O(n)$, where n is the number of known faces at each location-based database. The search space requirement of MoE only depends on the number of known faces in each database but not on the number of training set face images, which reduces the search space when the training set is big.

We have used publicly available face detection and representation algorithms in this research (Zhu et al., 2012). However, there is plenty of room for improvement. When only detecting frontal faces, many faces go undetected. This reduces the recognition rate. Variations in head pose and illumination could be addressed in the face recognition (Wagner et al., 2012). We will work taking other available context information into account. Additionally, if multiple photos of the same event are available, considering clothing appearance might also improve the recognition.

In addition to face recognition based on social networks we could recognize faces in photo albums. We believe that many of our observations also hold when using different recognition algorithms. It would be useful to investigate whether the insights gained from our experiments can be generalized to other face recognition approaches. In particular, it would be interesting to see how well these algorithms perform on an entire social network. We expect that the results of such a comparison will add to a scalable and accurate approach to recognize faces from on-line social networks.

5. Conclusion

Our location-based face recognition system on smart mobile devices improves classic face recognition without context in several ways. Applications such as smart video communication tools (Bergen et al., 2013) and personalized web tasking applications can become smarter using location-based face recognition (Taherimakhsousi et al., 2014). For future research, it would be interesting to build different context based ontologies and then investigate their relative merits with respect to face recognition speed and accuracy. Furthermore, it would be worthwhile investigating how to exploit social network information effectively.

Acknowledgements

References

Kang, J. M., Lin, T., Bannazadeh, H., and Leon-Garcia, A. (2014). Software-Defined Infrastructure and the SAVI Testbed. In Testbeds and Research Infrastructure: Development of Networks and Communities (TridentCom 2014), pp. 3-13. Springer.

Leon-Garcia, A. (2011) NSERC Strategic Network on Smart Applications on Virtual Infrastructure. University of Toronto.

Li, Y., Han, J., Yang, J. (2004). Clustering moving objects. In Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (SIGKDD 2009), pp. 617-622.

Viola, P., Jones, M. J. (2004). Robust real-time face detection. International Journal of Computer Vision, 57(2), pp. 137-154.

Taheriakhsousi, N., Ebrahimpour, R., Hajiany, A. (2009a). View-independent face recognition with biological features based on mixture of experts. In International Conference on Intelligent Systems Design and Applications (ISDA 2009), pp. 1425-1429, IEEE.

Tan, X., Li, Y., Liu, J., Jiang, L. (2010). Face liveness detection from a single image with sparse low rank bilinear discriminative model. In Computer Vision (ECCV 2010), pp. 504-517, Springer.

Taheriakhsousi, N., Ebrahimpour, R., Hajiany, A. (2009b). Face recognition based on neuro-fuzzy system. International Journal of Computer Science and Network Security (IJCSNS), 9(4), pp. 319-326.

Zhu, X., Ramanan, D. (2012). Face detection, pose estimation, and landmark localization in the wild. In IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2012), pp. 2879-2886.

Wagner, A., Wright, J., Ganesh, A., Zhou, Z., Mobahi, H., Ma, Y. (2012). Toward a practical face recognition system: Robust alignment and illumination by sparse representation. IEEE Transactions on Pattern Analysis and Machine Intelligence, 34(2), pp. 372-386.

Taherimakhsousi, N., Müller, H. A. (2014). Context-based face recognition for smart web tasking applications. In proceedings 1st Workshop On Personalized Web-Tasking (PWT 2014) At Tenth IEEE World Congress on Services (SERVICES 2014), pp. 21-23.

Bergen, A., Taherimakhsousi N., Jain P., Castaneda L., Müller, H. A. (2013). Dynamic context extraction in personal communication applications. In Centre for Advanced Studies Conference (CASCON), pp. 261-273.

Web sites:
Web-1: https://www.npd.com/wps/portal/npd/us/news/press-releases/pr_111222/ retrieved March 2015.
Web-2: http://www.freerepublic.com/focus/chat/3245617/posts consulted 11 Jan. 2015/ retrieved March 2015.
Web-3: http://www.savinetwork.ca/ retrieved March 2015.