

A Remote Image Collecting to Create Initiative Database with Indexing and Querying for Enhance Face Recognition

Phichaya Jaturawat, Pasinee Pongmanawut, Manop Phankokkrud*

Faculty of Information Technology
King Mongkut's Institute of Technology Ladkrabang
Bangkok, Thailand 10520
Email: manop@it.kmitl.ac.th

Abstract—Face recognition is the personal identification method that gaining the attention and popularity for many years. It can be seen from an increasing of applications and researches in this field. The user's face database is a very important part for enhancing the face recognition when it was applied in the actual system. Thus, the study focused on the method for collecting user images by using web application with WebRTC technology that helps the system to be able to use with any platforms without the additional installation. Additionally, XML is used to index the images and query them back by using Path_MD algorithm to make the system be more efficiency. The system implementation result shows that remote collecting images, image indexing with XML, and image querying using Path_MD algorithm are effective and can be developed in further studies.

Keywords- WebRTC; image collecting; image database; image indexing; image querying; XML; Path_MD; face recognition.

I. INTRODUCTION

Over the past decade, personal identification becomes a popular research topic and has been used in many systems. The organizations which concern about the security trying to apply the effective personal identification methods into their system for protected the confidential information or limited access. Personal identification can be achieved in several ways such as passwords, PIN, ID cards, etc. The biological characteristic is one of the ways that prevalent uses for personal identification, because of the uniqueness and arduous of characteristics were difficult to be duplicated, so we can identify a person precisely. The biological characteristics that commonly use are faces, fingerprints, iris, voice, signature, and keystroke. Thus these methods can solve the vulnerable of other personal identification methods. However, face recognition is the most outstanding and overcomes the other biological methods because it can be done passively without device participation from users. This is beneficial for security, surveillance, general identity verification, and the other purposes [1].

Nevertheless, there are two factors that affected the performance of face recognition; the face recognition algorithms and the quality of the image database. Many face recognition algorithms have been proposed and some of them

also used widely, but many researchers try to develop a new algorithm to use for face recognition by combining multiple algorithms together to improve the recognition performance which described by Teja [2] and Lone et al [3], and the research topic also includes the comparative study and the feasibility testing in terms of applications. Although, the algorithm is the key factor that affects the recognition performance, but the image database is an important factor as well. Not only the quality and quantity of sample images that affect the accuracy of the recognition [4], but the size of the image also affects the memory usage and execution time. The researchers use the constructed image databases that provided by the organization or the university, for the experiments. From the reasons as we explained about the relative of an image database and recognition performance, in practice, we have to collect many sample images several times in different face expressions and many constraints of environment to obtain a good image database and improve recognition performance. Thus, the construction of image database is difficult and some methods not flexible for increasing the number of user outright. Even though, the images of user were collected as well, but there are plenty of the images that need the good management to facilitating face detection and recognition process. The image indexing and retrieving process are the major parts of the images management that influence face detection and recognition, so the method that using for indexed image must provide a good performance in face recognition.

In this paper, we proposed the method for remote face images collecting using web application with WebRTC technology. WebRTC is a standard which supports for the web real-time communication. In recent years, the researchers mostly studies WebRTC in terms of applied in video conferencing and collaboration [5]. The distinctive point of this API enable web to access through the peripheral input devices such as microphone and web camera. Not only web real-time communication ability, but this technology also runs on any platform without need of installing additional plug-ins. Consequently, web application with WebRTC is very useful and helps the image capture process more simply and have an ability to verify the person in the image is the real user or not. Additionally, all of images would be clustering and stored

within the database for each user automatically after collecting. Then, it would be indexed in tree structure using XML representations. This method is capable of image indexing efficiently as described in Azzam et al [6], so in this study XML was decided to use for index and Path MD algorithm was used for efficient images querying. Due to these reasons, the system particularly helpful for creating an initiative database for face detection and recognition, and make it more convenience to apply in the online system later.

The rest of this paper is organized as follows. In section 2, the state of problems and system architecture would be described as well as WebRTC technology and image indexing using XML. In section 3, described about the implementation of the proposed system and the image quality assessment. Section 4, the results of the images that were collected are shown and discussed. Finally, the conclusion and the future works are discussed in section 5.

II. METHODOLOGY

In this section describes about the problem definition of face database construction, the architecture of the system that we designed with WebRTC technology, the system functionality, and the image indexing with XML.

A. Problem Definition

Face recognition is suitable for applying in the access control or surveillance system because it is a good method to identify a person that can be done in both direct and indirect ways. However, we need to create the user's face database before using the face recognition system because the greater number of images will affect the face recognition process to be more precise. The study found that creating an image database that can be done in several ways. There are advantages and constraints in each method as summarized in Table I.

TABLE I. COMPARISON OF IMAGE COLLECTING METHODS

Methods	Advantages	Constraints
Let the users come and take a photo in specific time and place.	- Good for a small number of users. - Easy to control the factors that affect the photo quality such as light, face angle and expressions.	- Not flexible for increasing the users.
Using previously taken user's photos	- Good for a large number of users - Not required user comes in person	- Unable to control the quality of pictures. - Difficult clustering if it is not digital files. - Increasing the number of users is not easy.
Using image collecting system.	- Good for large number of users. - Not required user comes in person. - Clustering images of each user automatically. - Flexible for increasing the users.	- Difficult to control user's posture and some factors that affect the photo quality

Accordingly, the development of the web application that can be used for collecting images in real-time is the most

challenge. The system must be able to provide the convenience of images collection even if the user was not in the same place and must be clustering and stored the images of each user separately to make it easier for using in the face detection process.

B. System Architecture

From the problem that was mentioned above, the remote user's image collecting system has been designed to solve the problem and provide more convenience for collecting user's images. This system has been developed in form of web application. It consists of two major parts that are client-side and server-side.

The client-side consists of two components that are camera and web browser. Camera is the first important thing for this system. In this case, web camera is a recommended tool because it is easy to connect to the computer. For the web browser, API called WebRTC technology which in the most of now available browser was decided to use in this study. The two main functions in WebRTC API that were used are `getUserMedia()` and `toDataURL()`. The `getUserMedia()` function using for request access to media devices (web camera) directly. The `toDataURL()` function is used for convert the captured image into the base64 DataURL.

The main component on server-side is the MySQL database that uses for contain the user information and the image database for storing the images that were collected from the system. We use `base64_decode()` function in PHP for decoding the DataURL into jpg image file.

From the Fig. 1, the system works following these steps. First of all create direct web camera connection via `getUserMedia()` function, this function allows to connect the web camera without additional installation. When users capture their image from the web camera, the image would be

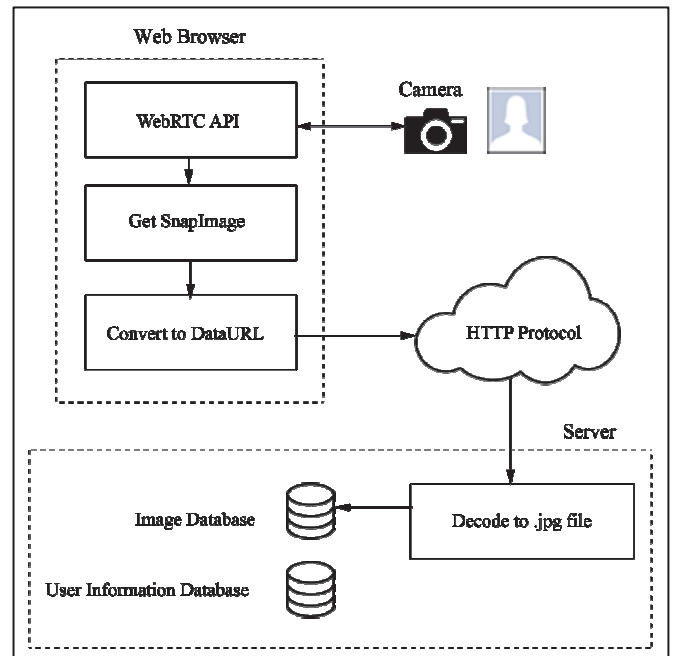


Figure 1. System Architecture

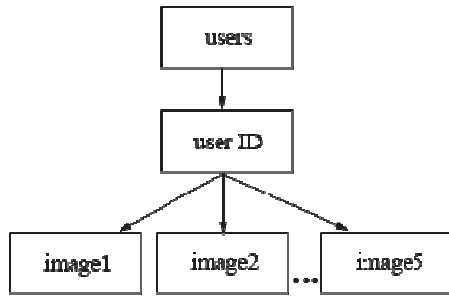


Figure 2. The XML tree

converting into base64 DataURL and sending through HTTP protocol to the server. Then, the DataURL would be decoding into jpg image file on the server-side before tagged user ID and storing in the database separately for each person.

C. WebRTC

WebRTC (Web Real-time Communication) [7] is an upcoming standard that aims to enable the real-time communication between web browsers to send and receive real-time media and data in a peer-to-peer (P2P). It can simplify the development of real-time communication between browser and other clients. In the common browsers, webRTC interactive component can be added to web application by using the common API.

The WebRTC architecture [8] involves at least three parts; one application provider and two peers. WebRTC helps the media and data flow directly between browsers without any intervening servers. For this ability, the web applications (client-side) that typically implement in combination of HTML, and JavaScript can interact with the local context (web browsers or mobile applications) through the WebRTC API. This API provides a set of function, such as a connection management, encoding/decoding capabilities, media control, firewall, etc., that is being implemented in JavaScript too. WebRTC API allowing the browsers and the scripting language to interact with media devices (microphones, web cameras, and speakers), processing devices, and transmission functions.

There are many advantages of WebRTC [9, 10]. First, it is a free open source API that existing in most all available browsers. Second is its platform independence, any browsers that support WebRTC with any operating systems can create a real-time voice or video connection to another WebRTC device or to a WebRTC media server. Third, WebRTC always on the data encryption, the Secure RTP protocol (SRTP) was used as key management for encryption and authentication of both voices and videos. The last important advantage of WebRTC is it able to working without the need of installing additional software or plug-ins.

For these advantages, the web application with WebRTC technology has been proposed to provide the convenience of images collecting. It can help to collect the user's face image anytime anywhere. So, the construction of user's face database would be easier.

```

<?xml version="1.0"?>
<users>
  <user id = 'user001'>
    <img1>filename1.jpg</img1>
    <img2>filename2.jpg</img2>
    <img3>filename3.jpg</img3>
    <img4>filename4.jpg</img4>
    <img5>filename5.jpg</img5>
  </user>
  <user id = 'user002'>
    <img1>filename1.jpg</img1>
    <img2>filename2.jpg</img2>
    <img3>filename3.jpg</img3>
    <img4>filename4.jpg</img4>
    <img5>filename5.jpg</img5>
  </user>
</users>

```

Figure 3. An example of XML code.

D. Image indexing

The web application with WebRTC technology that has the advantage as described in previously topic. The images management process, firstly, we are clustering image by user ID and then stored it separately by the user in the folder. Not only the image clustering but this application also needs an appropriate method to indexed images file for efficient retrieval in the system that developed for online using. Hence, the XML (Extensible Markup Languages) for image indexing was decided to use. XML is considered as a standard for exchange the information over the internet, because of its data self-describing and flexibility for representing many kinds of information by using a tree structure [11]. Because of these reasons, XML is the proper method for image indexing in the web application as well.

The XML tree structure was designed as depicted in Fig. 2; the user element has child elements, i.e., image1, image2, image3, image4, and image5 elements for image file names. As the example code in Fig. 3, user element has an attribute 'id' with a value 'user001'. It is an identification code that was defined for each user.

When the images were indexed, we have to choose the proper method to query the images for using in face detection and recognition process in next procedure. The popular XML query technique that widely used is PathStack [12] which has the advantage that requires only one scan of the relevant data to evaluate path queries with not-predicates and memory space requirement is bounded by the longest path in the input XML document. However, there are a lot of researches on the efficient of XML query to improve querying performance and makes it more efficiently for using in varies application. Bruno et al [13] introduced a new index-based technique to answer multiple XML path queries, namely Index-Filter. This technique takes advantage to pre-computed index over the input document, but can also build an index on the fly. Another interesting technique is multi-resolution structural index (MRI) [14], it facilitates fast retrieval of XML by assigns unique number to the elements on all possible paths. It can directly access the internal data via unique coding.

The method that we choose to implement has a similar concept as MRI but instead of assign unique number to elements, it uses multidimensional space to define the correspondence between the simple paths and disk blocks which the paths would be stored. This method names Path_MD. It was proposed by Musleh et al [15] to answers XML simple path query efficiently. When using this method, the XML data was stored as simple path using Dewey code numbering scheme. Each element of XML has its dimensional and index that would be computed based on its information. All paths are represented as a point in an N-dimensional space when N equals to the number of levels in XML tree. The multidimensional space would divide into a set of hyper rectangles, and their elements would be pointers to disk blocks. When querying, it significantly reduces hard disk access. That is the reason why this method can help to reduce execution time and increase accuracy for our system to query image paths. It is useful and suitable for using images in the next process.

III. IMPLEMENTATIONS

This system was implemented by using a combination of HTML and JavaScript for web application on the client-side, and using PHP with MySQL database for the system on the server-side. In the web application, WebRTC technology was mainly used in the image capture process. The important function is getUserMedia, this function is used for request web camera access from the user. In testing, only web camera on desktop PCs was using as a tool to collect the sample images in the laboratory without other optional devices. Before using this web application, all of users need to register and get username and password for their account before using. Afterwards, they can login to the system for capture and upload their photos by themselves. From the above method, the number of images can pre-define and flexible to request for additional image capture later.

VGA resolution (640*480 pixels) was decided to use for the image size to get the image that contains sufficient face features and have a good quality. Even though the smaller size is better in less memory usage, less network bandwidth for file transfer, and less execution time, but it might not has enough face features for use in the face recognition process. In spite of the bigger image contains more details but it tradeoff for resources usage too. All of the images will be stored in the database separately in the folder for each user and automatically record the taken time and date. Then, the image was indexed by using XML as a schema depicted in Fig. 2 that was described in previous section. The XML indexed file was kept for using in the face detection process for efficient querying with Path_MD algorithm. It would be measure performance in term of the number of disk access and execution time.

After collecting process, the images would be assessing and check for the quality following these criteria; the facial size in the images, the face straight, the face apparent without anything cross over or eyes closing, the image sharpen, the image brightness and light position, and face detection testing by using Haar cascade in OpenCV (Open Source Computer

Vision). This is a library of programming function mainly in real-time computer vision, thus it helps the development of face detection module easier. The evaluation would be taken by manual looking and test in simple face detection process to classify the image into three categories; good, fair, and bad quality. The good quality must achieve at least five in six of the criteria, four for fair quality and less than four for bad quality. However, this evaluation process is not very precision for image quality assessing but it can used as a part of criteria to assess the overall system as well. The system assessment did not only evaluate the result of the system that is the image quality, but also need usability testing and evaluation from actual user. Because this system provide in online environment that user must use it by themselves without trainer, so the system must be easy to use and easy to learn.

IV. RESULTS AND DISCUSSION

The result of the implementation obtained the 148 user's face database that contains about 8-10 images each person. The images were collected in the same laboratory environment but not have many constraints. Thus, the images are very difference. Mostly of the collected images have a good quality similar as showed in Fig. 4(a) and (b). It is a full straight face without obscured object and has enough brightness. Even the Fig. 4(b) the facial size is smaller than the expectation from the VGA quality. In Fig. 4(c) and (d) were classified into fair quality because the face was obscured by glasses, user's hand, and fringe but these images still have a good facial size and enough brightness, though the face in Fig. 4(c) cannot detected. Fig. 4(b), (c) and (e) are the example of the images that not have only one person.

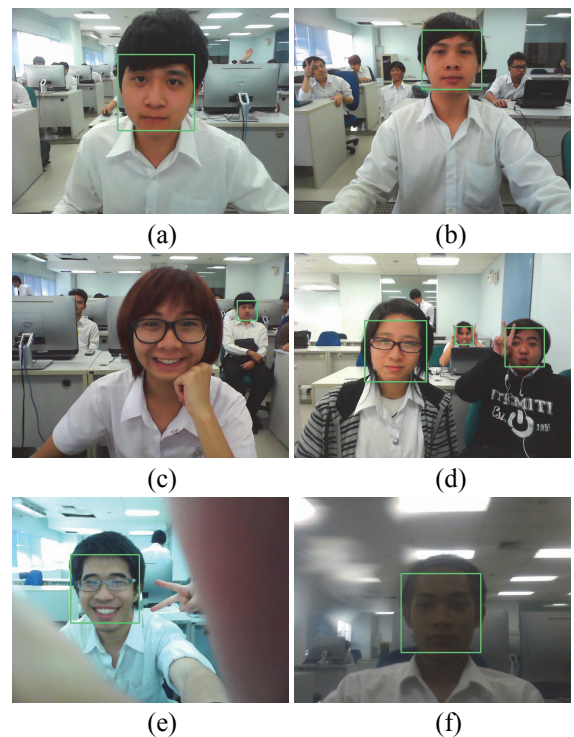


Figure 4. The result images.

In this case will assume the person who has a biggest facial size is the owner of the images. Although the images were collected in the same room, but the position was not the same. Fig. 5(f) is the example of the image that blurry and not has enough brightness even the face in the image can be seen and detected but the quality is not sufficient enough for use in recognition process, this image was classified into a bad quality.

The result of this attempt showed this way is a good way for collect the user's images because it takes a few times to captured the image and it can use by all users in the same time from any computers. So, this method can get enough user images in short time and this method can be applied into the online system that uses face recognition because it has flexibility to adding new user. Furthermore, this system can ensure that the person in the image is the real user because only user who has username and password can login to the system. The system was evaluated by the 148 representative users in the factor including ease of use and easy to learn. The result showed the most of users can understand and use the system correctly without assistance.

However this method has some limitations for using to collect images. First, the system cannot control some environment such as light illumination. Second, the other factors that affect to the quality of images such as user's postures and other people in the images. Finally, the web camera quality because it directly affects the quality of the image (image sharpen). According to these problems, some of the images might unavailable to use.

In the image indexing, XML is the proper method for indexed images to use in other procedures such as face detection. And the Path_MD is perform well for querying images especially in term of disk access and less execution time, but this method is not the best for this work because of the XML data have not many tree levels. Thus, the other simply method was able to use for querying.

V. CONCLUSION

This paper presented a new method for remote face images collecting using web application with WebRTC technology to improve image database creation more precisely and convenience. This is an effective method because it can help to capturing and sending image through the internet for storing in the database no matter where users are. Moreover, this system able to runs on any platforms without need of installing additional plug-in so, it capable to use anytime anywhere. Not only to improve image collecting and clustering process, but this system also design for more convenience image query by creating XML index. The implementation result showed Path_MD is a good XML query algorithm that performs well for retrieving images. Without doubt, presented system achieved the objective of image collecting and management.

In the future work will improve the capture image process by including the posture guideline and real-time face detection to validate the facial size. And find more suitable XML query algorithm for using in our face detection process.

ACKNOWLEDGMENT

This work was funded by Faculty of Information Technology, King Mongkut's Institute of Technology Ladkrabang.

REFERENCES

- [1] R. Jafri and H.R. Arabnia, "A survey of face recognition techniques," *Journal of Information Processing Systems*, vol.5, No.2, pp.41-68, 2009.
- [2] G. Teja and S. Ravi, "Face recognition using subspaces techniques," *Recent Trends in Information Technology (ICRTIT)*, pp.103-107, 2012.
- [3] M.A. Lone, S.M. Zakariya, and R. Ali, "Automatic face recognition system by combining four individual algorithms," *Computational Intelligence and Communication Networks (CICIN)*, pp.222-226, 2011.
- [4] M. Artiklar, A. Masadeh, M.H. Hassoun, and P. Watta, "The effect of expressions in a database of face images," *Circuits and Systems, Proceedings of the 43rd IEEE Midwest Symposium*, vol.2, pp.746-749, 2000.
- [5] A. Zeidan, A. Lehmann, and U. Trick, "WebRTC enabled multimedia conferencing and collaboration solution," *World Telecommunications Congress*, pp.1-6, 2014.
- [6] I. Azzam, A. Charlapally, C. Leung, J. Horwood, "Content-based image indexing and retrieval with XML representations," *Intelligent Multimedia, Video and Speech Processing*, pp.181-185, 2004.
- [7] S. Loreto and S. Romano, "Real-Time Communications in the Web: Issues, Achievements, and Ongoing Standardization Efforts," *IEEE Internet Computing*, Vol.16, pp.68-73, 2012.
- [8] C. Jennings, T. Hardie and M. Westerlund, "Real-time communications for the web," *IEEE Communications Magazine*, vol.51, no.4, pp.20-26, 2013.
- [9] G. Audin, "9 Advantages of WebRTC", Mar. 2014; <http://www.networkcomputing.com/unified-communications/9-advantages-of-webrtc/a/d-id/1113301>.
- [10] P. Edholm, "The Benefits of WebRTC", Dec. 2012; <http://www.webrtcworld.com/topics/from-the-experts/articles/319037-benefits-webrtc.htm>.
- [11] B. Hou, R. Zheng, G. Yang, "Quick search algorithms based on ethnic facial image database," *Software Engineering and Service Science (ICSESS), 2014 5th IEEE International Conference on*, pp.573-576, 2014.
- [12] Enhua Jiao, Tok Wang Ling, Chee-Yong Chan, "PathStack: A Holistic Path Join Algorithm for Path Query with Not-Predicates on XML Data", *DASFAA'05 Proceedings of the 10th international conference on Database Systems for Advanced Applications*, pp.113-124, 2005.
- [13] N. Bruno, L. Gravano, N. Koudas, and D. Srivastava, "Navigation- vs. index-based XML multi-query processing", *Data Engineering, 2003. Proceedings. 19th International Conference on*, pp.139-150, 2003.
- [14] A. Maghamez and Gongzhu Hu, "Multi-resolution indexing for XML data," *Software Engineering Research, Management and Applications, Third ACIS International Conference on*, pp.206-211, 2005.
- [15] D. Musleh, S. Adam, and M. Al-Mulhem, "Efficient Multidimensional Simple Path Query Processing Algorithm for XML data" *IEEE Information Science and Applications (ICISA)*, pp.1-5, 2013.