

Development of “KAOLABO”, the Laboratory Web Site for the Face Search Technologies

Hajime TERAYOKO*, Yuuko ABE*, Tetsuya SAWANO*, and Norihisa HANEDA*

Abstract

We have released a website “KAOLABO” to promote our face search technologies. Internet users can personally test the face detection with some trial applications and the WebAPI on our website. In this paper, we report the approach to release our technology through the internet and its influence.

1. Introduction

In recent years, face detection technology rapidly gets a lot of attention. It is because this technology is generally mounted on digital cameras and used in services through cellular phones, and thus the technology and its effects are increasingly known by general public.

FUJIFILM has been addressing the research and development of the face detection technology for a long time thus far, and now, in response to abruptly increasing demand in these days, we broaden the range of applications of the technology, and set up an experimental public access website “KAOLABO” and disclosed our face detection technology on the internet for the purpose of develop new applications.

In this report, a detailed method of disclosing the technology in KAOLABO, a process flow performed in the internal system, and points we have exercised our ingenuity in providing the face detection technology as a Web service will be described. Also, the effects that we achieved as a result of technology disclosure will also be described.

2. What is Face Detection Technology?

Prior to the description of KAOLABO, a general description of the face detection technology and characteristics relating to the face detection technology of FUJIFILM will be briefly described.

The face detection technology is an image processing technology to automatically find a face area included in an image. There are many other image processing technologies relating to the face, such as a face recognition technology for recognizing the identical person from the similarity of the face, a technology to determine specific expression such as smile, and a technology to estimate the age or the gender of the person from the face. The face detection technology is a basic technology which forms the foundation of a series of face-related technologies.

A general method of face detection is for determining whether the object is a face or not by combining a number of simple determination units for detecting brightness of a rectangular area by machine learning (Fig. 1).

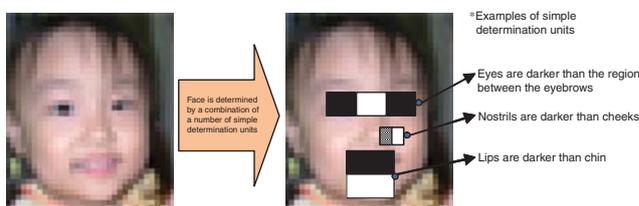


Fig. 1 Concept image showing face detection technology.

The face detection technology has been studied actively from 1980s' and since a technology of machine learning by AdaBoost combined with a rectangular characteristic referred to as Haar-Like feature became popular in around 2000, it has rapidly put into a practical use. It has stated to be mounted on accessible apparatuses such as digital cameras or cellular phones from around 2005, and is now widely used also in entertainment applications such as services of cellular phone or Website and computer games.

3. Face Detection Technology of FUJIFILM

FUJIFILM has started the research and development of the face detection technology from the age of analogue MINILAB in 1990s'. Thereafter, we made digitalization and have been accumulating the technologies for a long time. Since 2006, we applied this technology to our own products by, for example, mounting the face detection technology on the digital camera “FinePix”.

One of the characteristics of the FUJIFILM's face detection technology is “high robustness” which allows stable detection of the face irrespective of the conditions of the image. It is a substantial achievement of our continuing efforts in development of technology of face detection for a long time for general pictures which is likely to be taken under unstable conditions on the basis of ultimate purpose, “How to make pictures look prettier”.

Original paper (Received December 1, 2008)

* Internet Business Development Division

FUJIFILM Corporation

Nishiiazabu, Minato-ku, Tokyo 106-8620, Japan

For example, if it is face detection for a specific application such as a security system on the basis of the face recognition, we can target only images taken under limited conditions such as the size or orientation of the face or illumination. However, in the general pictures, these conditions enormously differ from one picture to another. Therefore, in order to detect the face stably in the general pictures, the high robustness which is unaffected by photographing conditions is required, and FUJIFILM has been focusing on this point most. Therefore, FUJIFILM's face detection technology realizes a high detection capability which is able to detect the face or faces irrespective of the orientation, inclination of the face or the number of people in the picture.

4. Face Detection Technology Experimental Public Access Website “KAOLABO”

FUJIFILM thought that it is necessary to appeal out the technology and develop new applications in association with increasing attention to the recent face detection technology. Therefore, we set up an experimental public access Website of face detection technology “KAOLABO” (<http://kaolabo.com/>) in Jan. 15, 2008 to disclose our face detection technology as a Web service (Fig. 2).



Fig. 2 Top page of “KAOLABO” website.

In KAOLABO, we disclose our technology in two modes; “Experience-Based Application” which allows the public to upload images from the Web browser and operate the face detection by GUI (Graphical User Interface), and “Face detection WebAPI” for allowing the public to use the face detection function from a program via Website. The greatest characteristic of KAOLABO is a point “Everyone can experience Operating Face detection”.

5. System Configuration of KAOLABO

Reviewing the system configuration of KAOLABO, detailed disclosed contents will be described.

The logical composition of KAOLABO system is as shown in Fig. 3, and brief summaries of the respective components will be given below.

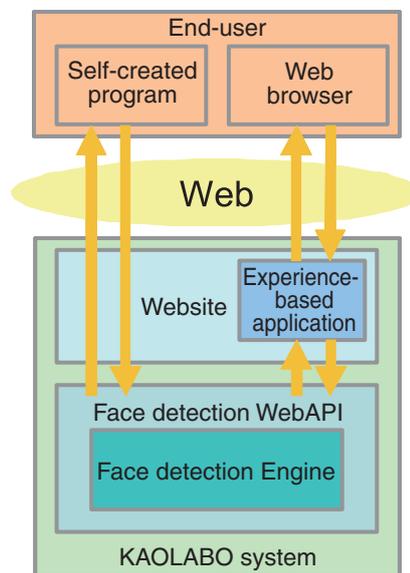


Fig. 3 System structure of “KAOLABO”.

5.1 Website

Website of KAOLABO serves as a front end for accepting public Website access with the system. Webpages in the site also include technical description about the face detection.

5.2 Experience-Based Application

The “Experience-Based Application” is a Web application implemented with Adobe Flash (hereinafter, referred simply as Flash), and is integrated in the Website of KAOLABO. The user uploads an image to a server, and is able to experience the face detection readily. For the moment, the following two experience-based applications are disclosed.

5.2.1 Face Secret (Fig. 4)

Processing such as Feathering, Eye-masking, Composite Picture with Stamps can be applied on the detected faces. This is an application created as a representative example of the practical applications.



Fig. 4 Trial application 1. “Face Secret”.

5.2.2 Face Bubble (Fig. 5)

We present a game such that if the user assigns a detected face, the face appears in a soap bubble blown by a boy. This is a representative example of entertainment applications, and appeals that “existence of the face” add a marked dramatic impact.

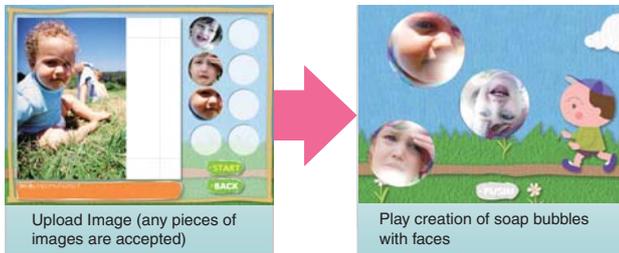


Fig. 5 Trial application 2. “Face Bubble”.

5.3 Face Detection WebAPI

This is a system which allows the user to call up the face detection function from a program in the form of a Web service. By uploading an image as a request to the server, we can get the result of face detection as a response. As the face detection is executed by calling up this WebAPI also from the experience-based applications, this is the pillar of the technology disclosure in KAOLABO (details will be described in the next paragraph).

5.4 Face Detection Engine

A face detection engine is a substance which performs the face detection in WebAPI, and cannot be accessed directly from the outside. Although this face detection engine is optimized according to the utilization on Web, it is actually the same as that mounted on the digital camera FinePix.

6. Face Detection WebAPI

Subsequently, “Face detection WebAPI”, an essential point of the face detection technology disclosure in KAOLABO, will be described further in detail.

The face detection WebAPI is so-called a REST type WebAPI on the basis of HTTPS protocol. When the user attaches a parameter to the request URL and sends the request, the user can receive result data of the face detection in XML form. Only JPEG images are supported.

It is necessary to register (free) to use the face detection WebAPI, and the user is required to register his/her mail address in the Website of KAOLABO in advance and acquire an authentication key.

6.1 Request

We present two modes for the face detection request with the convenience in mind, that is, methods which support POST and GET of HTTP. POST is a method of transmitting image data by uploading with request body. In contrast, GET

is a method of placing URL of image in a request parameter. With these methods, the face detection can be executed on the image resource on the Web. With this GET method, the face detection WebAPI can be executed from the Web browser without specific programming.

`https://kaolabo.com/api/detect?apikey=[AuthenticationKey]&url=[image URL]`

By entering the request URL as shown above to the browser, XML as shown in Fig. 6 is displayed. This is response data of face detection WebAPI.

```
<?xml version="1.0" encoding="UTF-8"?>
<results xmlns="http://xmlns.kaolabo.com/webapi" version="1.0">
  <faces>
    <face x="50" y="50" width="80" height="80" score="100">
      <left-eye x="70" y="70" />
      <right-eye x="110" y="70" />
    </face>
    <face x="603" y="418" width="100" height="100" score="30">
      <left-eye x="650" y="500" />
      <right-eye x="678" y="490" />
    </face>
  </faces>
</results>
```

Fig. 6 Response XML of face detection WebAPI.

6.2 Response

The face detection WebAPI returns the face data of respective faces detected from the image in XML as a response. Detailed contents of the detected face data are as follows (items within parentheses are corresponding XML elements and attributes)

- (1) Face position (face, x, y)
Coordinate of upper left of a rectangular area including the face.
- (2) Face Width (face, width) and Height (face, height)
Width and height of the rectangular area including the face.
- (3) Eye Coordinates (left-eye/right-eye)
Coordinates of left and right eyes. Definition of left and right is the left eye and the right eye of the face (not the left and right from the viewer).
- (4) Likelihood of Face (face, score)

“Likelihood of Face” is value indicating the likelihood that the detected object is the face, that is, the probability that the face detection engine determines the object as the face. It is used for filtering erroneous detect of an area which is not the face as the face area. Processing, clipping, and composition of face areas of the image can be performed by using the face data included in the XML data in the response.

As regards detailed specifications of Face detection WebAPI, see the following Website;

<http://kaolabo.com/webapi/spec>

7. Face Detection Processing in KAOLABO

How we disclose the face detection technology via KAOLABO has been described thus far. Then, how the face detection processing is performed in the interior of the KAOLABO system?

A flow of the face detection processing in KAOLABO will be described with the operation in the experience-based application as an example.

A flow from uploading an image until displaying the result of the face detection on the screen of the application by the experiment-based application is as shown in Fig. 7. The flow of the processing will be described below with reference to the drawing.

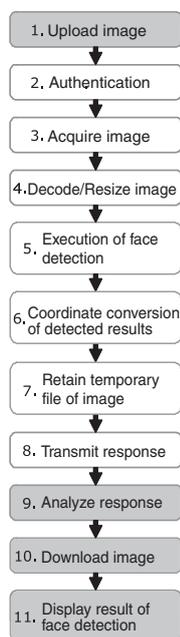


Fig. 7 Flow chart of face detection.

(1) Uploading of Image

First of all, the user uploads an image to a server from the GUI of the experience-based application (Flash). At this time, POST type is specified to the selected image and the face detection WebAPI is called in the interior of Flash.

The process up to here is to be done via the client side, and the process shown below is to be done by the server side.

(2) Authentication

KAOLABO server receives a face detection WebAPI request of the experience-based application, and starts processing. First of all, authentication is performed by the authentication key set in the request parameter.

(3) Acquisition of Image

When the authentication is successful, the image data specified in the request is acquired. When it is a call-up from the experience-based application, JPEG image data attached to the request body in the POST method is received.

(4) Decoding/Resizing of Image

The received JPEG image to be processed is decoded, and simultaneously, resized to a minimum possible size which allows the face detection engine to process without degrading the accuracy. This is part of improvement of process speed and reduction of required memory.

(5) Execution of Face detection

The RGB RAW data obtained by decoding is input to the face detection engine and the face detection processing is executed. The face detection engine outputs the face data as the detected result.

(6) Coordinate Conversion of Detected result

The image data used for face detection is resized data in advance, and therefore the coordinate conversion is performed for adopting the coordinates and the size of the output result to those before resizing.

(7) Temporary Retention of Image File

The uploaded original image file is retained in the server temporarily for referencing from the Flash experience-based application.

(8) Transmission of Response

The face detection result after having subjected to the coordinate conversion and the URL of the temporary file retained in the server are modified into XML data, and is transmitted as a WebAPI response.

The face detection process up to here is to be done via the server side, and the process shown below is shifted to the client side again.

(9) Analysis of Response

The experience-based application receives the response XML of the face detection WebAPI, analyzes the received response, and acquires face data and URL of the temporary file.

(10) Downloading of Image

Subsequently, the user accesses the URL of the temporary file retained in the server and downloads the image.

(11) Display of the Face detection Result

Flash reads out the downloaded image, and draws a frame line and displays at a position of the face using the face data. These processes are performed on memory by Flash.

With the procedure thus far, the image uploaded via the experience-based application is subjected to the face detection and the result is displayed on the application image. KAOLABO system executes the face detection process in this manner.

8. Optimization for Web Service

In order to disclose the technology of KAOLABO, it is necessary to optimize the technology for the usage on the Web, and various ingenuities are exercised as follows.

8.1 Tuning of Face Detection Engine

FUJIFILM's face detection engine allows a number of parameters to be set, such as the size of the face to be detected, or whether the orientation of the image is set in advance or not. The speed and accuracy can be adapted to the objective by adjusting the parameters adequately according to the characteristic of the input image. In general, the speed and accuracy are in "trade-off" relation, that is, when the parameters are tuned to get higher speed, the accuracy is sacrificed and, in contrast, when the accuracy is secured, the speed tends to be lowered. Therefore, how to balance the speed and the accuracy according to the objective is the point of the engine tuning. In the case of the KAOLABO, we made a basic policy to give preference to the speed over the accuracy in order to secure the operation with grate agility on the Web. It is because we need to avoid speed bottlenecks in transmission of images caused by the network.

On the other hand, there was an aspect that the accuracy was difficult to secure according to the conditions of the input image because the images uploaded by the users could not be specified. Therefore, we prepared a relatively wide variety of input image samples, and tuned the face detection engine while seeking the best possible balance which demonstrates the superiority of the detection accuracy even with the images under various conditions while giving preference to the speed.

Anticipating the flock to the site, we inspected the performance while applying a load to the system and derived a performance which could resist a certain degree of heavy traffic.

In this manner, the face detection engine was optimized to achieve an operation without stress and demonstrate high detection accuracy on the Web as well.

8.2 Security Guard

The following measures are taken for handling the image data including faces of human beings securely on the Web.

8.2.1 Authentication

As described before, we employed a pre-registration system for the usage of the WebAPI, and issued an authentication key to each registrant. Authentication is performed on the request base using the issued authentication key. With this system, operation control is performed according to the state of usage while recognizing the user.

8.2.2 Encryption of Communication via SSL

The entire communication with the WebAPI is encrypted by SSL. Although the convenience on the calling side is impaired a little thereby, we put priority on security.

8.2.3 Automatic Deletion of Temporary File

In the experience-based application, the image uploaded by the user is retained in the server as a temporary file.

However, in order to avoid retention of users' images more than necessary, the experience-based application is adapted to delete these temporary files automatically after elapse of certain period.

9. Effects of Technology Disclosure

In KAOLABO, we have been disclosed the face detection technology on the Web in this manner.

Accordingly, the face detection technology is now available for all users as the Web service. Therefore, in addition to the conventional applications such as integration in the apparatus or utilization in an independent system, utilization in various implementations are available now, and the range of application of the technology is dramatically widened. A mashup utilizing the face detection WebAPI was created by a general user with his unique idea. This is an example of the effects of the technology disclosure. Mashups applied by the users are partly introduced in "Application Showcase" in KAOLABO site. (See <http://kaolabo.com/wabapi/mashup>).

We received a number of approaches from enterprises interested in KAOLABO saying that they want to use the face detection technology in business application. Among them a number of kinds of industries and applications we did not supposed so far were included, so that we believe that we could develop new applications of this technology.

Furthermore, provision of an occasion to allow the public to try the actual operation has contributed to introduce the performance of this technology to the outside world. In particular, in the early days after having set up the Website, it was spotlighted and was reported in many mass media and personal weblogs.

Disclosure of the technology through KAOLABO produced substantial results.

10. Conclusion

Disclosure of the face detection technology in KAOLABO was a sort of risky experiment to provide FUJIFILM's core technology as Web service and bring our technology to the attention of the world as an advanced internet technology, and we could achieve enlargement of the range of application, development of new applications, and appeal of the technology.

FUJIFILM is considering that we disclose a new technology other than the face detection technology in KAOLABO in the same manner, also in the future.

(In this report, "Adobe Flash" is a registered trademark of Adobe Systems Inc. "KAOLABO" and "FinePix" are registered trademarks of FUJIFILM Corporation.)