The Seventeenth International Symposium on Artificial Life and Robotics 2012 (AROB 17th '12),
B-Con Plaza, Beppu, Oita, Japan, January 19-21, 2012

# Design of a User-Support System for Vision Information Using a Smart Phone

Hiroyuki Nishiyama[1] and Fumio Mizoguchi[1]

[1] Faculty of Sci. and Tech. Tokyo University of Science
2641 Yamazaki, Noda-shi, CHIBA, 278-8510, Japan
(Tel: +81-47122-1106; Fax: +81-47122-1106)
(hiroyuki@rs.noda.tus.ac.jp, mizo@wisdomtex.com)

**Abstract:** In this paper, we design a user-support system for vision information using smart phones. When the user takes a picture of a target using the smart phone camera, our system can recognize the image of the target and recommend a suitable service to the user. The system with executes simple image analysis and determines a suitable service for the target image. The simple image analysis can extract 13 parameters (e.g., color information (RGB), number of intersections, depth of intersection, line width, and line depth). We use machine learning to classify the parameters into each service. In our research, we use a Support Vector Machine (SVM) as a learning machine tool. Our system design provides user support for several services such as translation, barcode readers, and diagnosis of skin images and demonstrates the effectiveness of our research.

**Keywords:** User Support System, Data Mining, Support Vector Machine, Smart Phone

## 1 INTRODUCTION

Recently, many multifunctional cellular phone terminals, such as smart phones (e.g., Androids and iPhones), have been developed as a result of the evolution of the computer, network infrastructure, and lightweight battery technology. Thus, the number of users is rapidly increasing [5]. The smart phone is equipped with various sensor systems (e.g., an acceleration sensor, an infrared sensor, and a luminosity sensor), besides the camera function. Furthermore, easily programmed applications acquire input from each function. Various research activities and services performed with a smart phone using such features. Examples are the study of service applications such as translation from English to Japanese, barcode readers, recognition of billboards, recognition of facial images, and diagnosis of skin images using visual information from the camera of a smart phone. Examples are the study of the diagnosis service of skin image that recommends cosmetics that are appropriate for the skin condition by transmitting the skin image from the camera function to an analysis server [1, 2] To use these services, a user should select a service application and point the smart phone camera (take a picture) toward the subject of the service, and then receive the information service from the application. However, when the subject size or focus do not match, the user cannot access the service. The user must then retake the picture. Such mistakes are frequently made when a user first accesses an application.

Considering this background, we have designed a user-support system for vision information using smart phones. In our study, when the user takes a picture of a target using the smart phone camera, our system can recognize the image of the target and recommend a suitable service to the user. In addition, our system provides advice by voice and a user interface using a touch panel. When the target image conditions are unsuitable for the service, our system can advise the user directly. The user can then adjust the size of the image target area by operating the touch panel. In our design, we use a design model of a server-client system; a smart phone is used as a client because the smart phone's CPU is not powerful enough to achieve these services, so the services are implemented by server machines. The system with the smart phone as a client executes simple image analysis and determines a suitable service for the target image. To improve its judgments, this system can advise and request input via voice and text to adjust the area of the target image using the touch panel. The simple image analysis can extract 13 parameters (e.g., color information (RGB), number of intersections, depth of intersection, line width, and line depth). We use machine learning to classify the parameters into each service. In our research, we use a Support Vector Machine (SVM) [3] as a learning machine tool.

Our system design provides user support for several services such as translation, barcode readers, and diagnosis of skin images and demonstrates the effectiveness of our research.

## 2 SUPPORT VECTOR MACHINE

The SVM that is used in our study to determine is a classification technique that uses supervised learning and that can be applied to pattern recognition and regression analysis. The learning algorithm of the SVM is derived from optimization theory, and the learning method introduced by Vapnik [3] calculates a super-plane that separates positive and negative examples as training data. High identification performance can

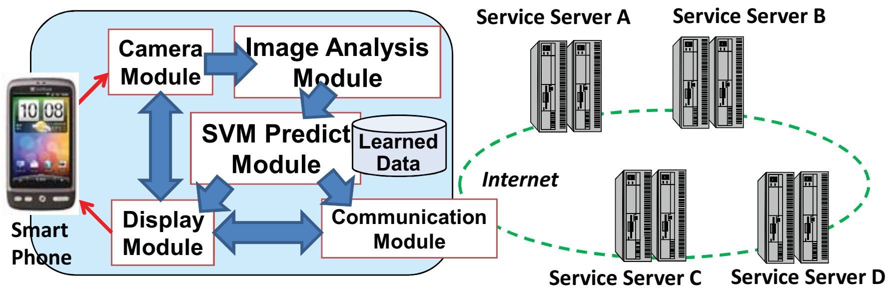


Fig. 1. Structure of the user-support system for vision information using smart phones

be demonstrated by adding the kernel function and margin maximization (e.g., mapping to a higher dimension space by using the nonlinear map, and introducing the kernel function to linear separate by the higher-dimension feature space). The pattern is mapped to the feature space of limited or infinite order, and it is possible to linear separate on the feature space. Thus, the SVM is used in various fields like categorizing text, image recognition, and bioinformatics, and was recently applied to Classification of Web services [4].

In our study, we use the Java version of libsvm [7] that can be implemented on an Android as the SVM and supports also multi-class classification.

## 3 SMART PHONE

A smart phone is similar to a small laptop computer. However, a smart phone has a telephone call function, and it is easy to carry because it is small and light. Its display is rather small, and it does not have a keyboard. Moreover, the development of an Android smart phone is comparatively easy, because the development language basically conforms to Java in the Android application (Android SDK[6]). In our study, the smart phone (Android) is an HTC Desire SoftBank X06HT (CPU: Qualcomm Snapdragon 1GHz, Internal memory: ROM 512MB/RAM 576MB, OS: Android 2.1 with HTC Sense).

The performance of the assumed smart phone (Android phone) in this study is as follows.

- Camera function

  A picture can be taken using the photography function of the Camera API (android.hardware.Camera) with which the Android phone is equipped. The camera focus when the picture is taken can also be checked using this API. The resolution of the picture changes with the performance of a smart phone; we use a size of 1216x2048 pixels in our research. (If the screen is too small, the accuracy will be affected; if it is too large, the processing speed of the smart phone will be affected.)

- Display function

  The picture and text can be displayed as output on the Android phone, which recognizes a finger touch and drag operation on a display as an input. These functions are provided to build the user interface of AndroidAPI (android.app.Activity).

Additionally, a smart phone is equipped with GPS and various sensor systems (e.g., acceleration sensor, brightness sensor, and temperature sensor) and can use voice annunciation through a speaker.

## 4 SYSTEM ARCHITECTURE

We have designed a user-support system for vision information using smart phones. When a user takes a picture of a target using the smart phone camera, this system executes simple image analysis and determines a suitable service for the target image. The system's simple image analysis can extract 13 parameters (e.g., color information (RGB), number of intersections, depth of intersection, line width, and line depth). We use machine learning to classify the parameters into each service. In our study, we use a Support Vector Machine (SVM) [3] as a learner, learn information collected with the smart phone beforehand, and judge the target image of newly collected information. The information is gathered and judged in a smart phone, and the learning process can be performed in a smart phone or a personal computer.

Figure 1 depicts the structure of the user-support system for vision information using smart phones. The modules in this system have the following roles.

- The camera module takes a picture using the camera function of the smart phone. This module judges the camera focus when photographing skin.

- The display module provides information to the user by displaying the input information and treats the operation on display as input from the user. This module displays the picture and text when the picture is taken, transmits

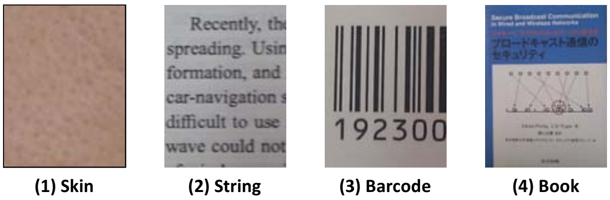


Fig. 2. Samples of parts of the image used by each service. (1) Skin image. (2) String image. (3) Barcode image. (4) Book image.

it to the service server, and provides the diagnostic result. The area of the picture to be transmitted to the service server can be selected using the touch-panel function of the smart phone.

- The image analysis module analyzes the image to extract the parameters from the image data. This image analysis extracts 13 parameters (e.g., color information (RGB), number of intersections, depth of intersection, line width, and line depth (image depth)). The image parameters are used by the SVM prediction module to determine a suitable service.
- The SVM prediction module judges the image parameters using learning data created by our learning tool (see the next section) and determines a suitable service for the target image.
- The communication module transmits information to and receives information from the suitable service server.

## 5 LEARNING BY SVM

In our study, the SVM creates the learning data of the image for which a suitable service is to be determined, and we implement the SVM prediction module in Fig. 1 by using the learning data. While learning, we collect image data that can be used by several services such as translation, barcode readers, and diagnosis of skin images; analyze each image and extract parameters; and execute parameter classification learning by SVM. We use the Java version of libsvm [7] that can be implemented on an Android as the SVM.

### 5.1 Extraction of Parameters by Image Analysis

To execute learning with an SVM, we implemented the function of image analysis to extract the parameters from the collected image data (such as Fig. 2) in the smart phone. This image analysis uses the same function implemented for the skin diagnosis service [1]. Processing was reduced to enable analyzing the image more quickly in the smart phone. The

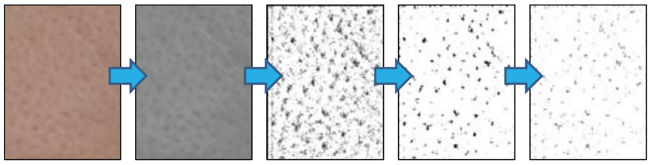

Fig. 3. Image analysis using a skin image. (A) Original image. (B) Gray-scale image. (C) Binary image. (D) Noise reduction. (E) Line image.

| Judge | ID0 | ID1 | ID2 | | ID12 |
|---|---|---|---|---|---|
| 1 | 0:489 | 1:471 | 2:154 | ··· | 12:128 |
| 1 | 0:704 | 1:676 | 2:143 | ··· | 12:124 |
| 2 | 0:1076 | 1:1051 | 2:146 | ··· | 12:120 |
| 3 | 0:744 | 1:726 | 2:147 | ··· | 12:130 |
| 3 | 0:877 | 1:862 | 2:170 | ··· | 12:128 |
| 4 | 0:1009 | 1:968 | 2:129 | ··· | 12:152 |
| ⋮ | | | | ⋮ | |

← 13 parameters → ; ↕ The number of images

Fig. 4. Parameters used for SVM learning (before scaling)

image analysis process (Fig. 3) first extracts 240 to 320 pixels of the central portion of the acquired image (Fig. 3(A)). Next, the function converts the image into a gray-scale image (Fig. 3(B)), prepares a binary image (Fig. 3(C)), performs noise reduction (Fig. 3(D)), and creates a line image (Fig. 3(E)). This image analysis extracts 13 parameters (e.g., color information (RGB), number of intersections, depth of intersection, line width, and line depth (image depth)). (The skin diagnosis service extracts 45 parameters to recognize the state of the skin image in detail.) The learning data are created by learning these parameters using the SVM.

### 5.2 Creating Learning Data

To create learning data by applying SVM learning, we collected the images that became training data. There were 100 data images, 25 each for skin, string, barcode, and book images. Next, we extracted the parameters from the collected image data using the image analysis function.

Figure 4 presents data when collected parameters are learned by the SVM. The parameters of one line represent one image. The left-hand numbers (1 to 4) indicate whether the image is skin (1), string (2), barcode (3), or book (4). After scaling, these parameters can be learned by the SVM, and learning data is created. We used the Java version of libsvm [7] that can be implemented on an Android as the SVM and that also supports multiple classifications. We used the Gaussian kernel as a kernel function for SVM learning. Cross validation indicated a learning accuracy of 89.0%.

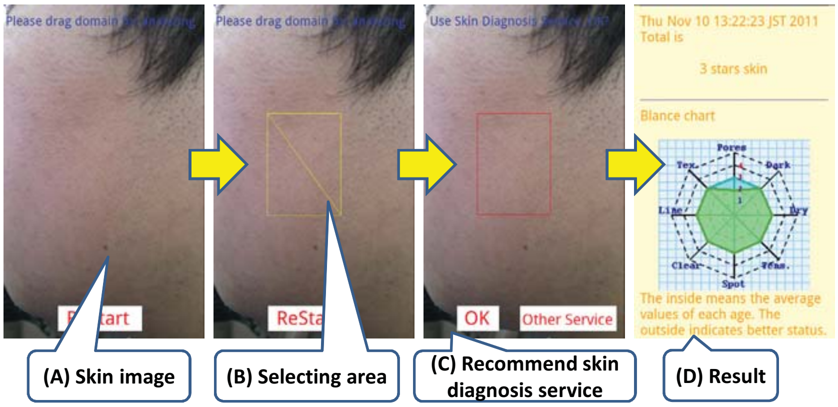


Fig. 5. Process of the user-support system for vision information using the smart phone

## 6 USER-SUPPORT SYSTEM

In our study, we implemented a user-support system that enabled service judgment based on the created learning data for vision information using smart phones (Fig. 1). We used HTC Desire X06HT (described in paragraph 3) as an Android smart phone. In our study, when the user takes a picture of a target using the smart phone camera, our system can recognize the target image and recommend a suitable service to the user. In addition, our system provides advice by voice and a user interface using a touch panel.

The process flow of the user-support system is presented in Fig. 5. In this figure, this user wants to use a skin diagnosis service. First, when the user photographs his/her own skin using the smart phone via the camera module, our system displays the skin image via the display module (Fig. 5(A)). Next, the user selects the area of the image (Fig. 5(B)) for which the service is to be used. The image analysis module then analyzes the image and extracts the parameters, and the SVM prediction module judges a suitable service. In Fig. 5(C), our system recommends the skin diagnosis service wanted by the user. When the user pushes the "OK" button (Fig. 5(C)), our system transmits the selected image data to the skin diagnosis service server via the communication module. Finally, the result is displayed (Fig. 5(D)). In these processes, our system can advise the user via text and voice while the user is taking a picture, for example, and provide operational advice based on the situation.

With our system, the user can take a picture of a target using the smart phone camera and receive a suitable service recommendation from our user-support system.

## 7 CONCLUSION

In this paper, we designed a user-support system for vision information using smart phones. When the user takes a picture of a target using the smart phone camera, our system can perform simple image analysis, recognize the target image, and recommend a suitable service to the user. To improve its judgments, this system can advise and request input via voice and text to adjust the area of the target image using the touch panel. We use machine learning with a Support Vector Machine (SVM) to classify the parameters extracted by the simple image analysis into each service. We implemented a user-support system that enabled the judgment based on the created learning data for vision information using smart phones and demonstrated the system's effectiveness.

## REFERENCES

[1] H. Hiraishi, F. Mizoguchi: A Cellular Telephone-Based Application for Skin-Grading to Support Cosmetic Sales, IAAI 2003, pp. 19-24, 2003.

[2] H. Nishiyama and F. Mizoguchi: Cognitive Support by Smart Phone - Human Judgment on Cosmetic Skin Analysis Support - , The 10th IEEE International Conference on COGNITIVE INFORMATICS & COGNITIVE COMPUTING: ICCI*CC 2011, 2011.

[3] V. Vapnik: The Nature of Statistical Learning Theory, Springer-Verlag, 1995.

[4] H. Wang, Y. Shi, X. Zhou, Q. Zhou, S. Shao and A. Bouguettaya: Web Service Classification using Support Vector Machine, The Proc. of the 22nd International Conference on Tools with Artificial Intelligence, pp. 3-6, 2010.

[5] Mobile Content Forum: *K-tai white paper 2010*, Impress R&D, 2009. [in Japanese]

[6] Android SDK Guide, http://www.android.com/

[7] C.-C. Chang and C.-J. Lin: LIBSVM: a library for support vector machines, http://www.csie.ntu.edu.tw/ cjlin/libsvm/