

## Implementation of English-Chinese Translation Based on Android OCR

Hao Rao<sup>1,a\*</sup>, Yuman Lin<sup>2</sup>, Xiaofeng Chen<sup>1</sup>, Tianyu Liu<sup>2</sup>

<sup>1</sup>Department of Computer Science, Shaoguan University, Shaoguan, 512005, China

<sup>2</sup>Department of Instruction Technology, Shaoguan University, Shaoguan, 512005, China

<sup>a</sup>gdrh@tom.com

**Keywords:** OCR, English-Chinese translation, Android

**Abstract.** Realization of English-Chinese Translation on Smart phone platform has been discussed. Recognition processing mainly uses Google OCR technology. Image binarization is used to increase recognition accuracy. Realization of English-Chinese translation depends on local lexicon realized through SQLite database, or using Youdao online translation interface.

### Introduction

Due to the rapid development of communication technology and mobile network, it is very necessary to realize English-Chinese translation by smart phone. Rely on Google's OCR technology, the image can be converted into English texts. And then, English-Chinese translation can be accomplished by querying the local lexicon database, or using Youdao online translation interface.

### Get image

Take a picture by camera of Android smart phone. Display the picture by SurfaceView. Define MyPictureCallBack class Inherited from PictureCallback. Get the image by MyPictureCallBack. The effect of image is shown as Fig.1.

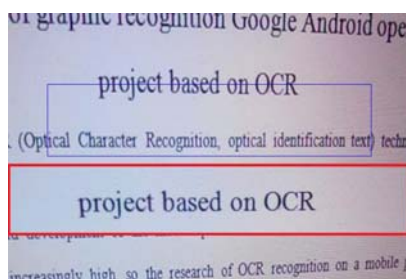


Figure 1. Get image by camera

### Image binarization

Image binarization process by the following steps: Transform color image into grayscale image. Calculate the maximum and minimum values of grayscale. Calculate threshold by iterative method. Method of image binarization is as following.

```
private static Bitmap binarization(int T) {
    for (int i = 0; i < imgHeight; i++) {
        for (int j = 0; j < imgWidth; j++) {
            int gray = imgPixels[i * imgWidth + j];
            if (gray < T) {imgPixels[i * imgWidth + j] = Color.rgb(0, 0, 0);}
            else{imgPixels[i * imgWidth + j] = Color.rgb(255, 255, 255);}
        }
    }
    Bitmap result = Bitmap.createBitmap(imgWidth, imgHeight, Config.ARGB_565);
    result.setPixels(imgPixels, 0, imgWidth, 0, 0, imgWidth, imgHeight);
    return result;
}
```

Result of image binarization is shown as Fig.2.

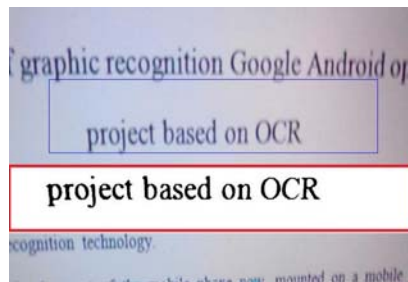


Figure 2. Result of image binarization

## Recognition process of Google OCR

TessBaseAPI class is an important class in the process of OCR recognition. Firstly the compiler environment must be configured. Secondly com.googlecode.tesseract.android.TessBaseAPI should be included. And then call an object. The object is an instance of TessBaseAPI class. The English texts can be recognized by calling methods of the object. Method of conversion is as following.

```
import android.graphics.Bitmap;
import com.googlecode.tesseract.android.TessBaseAPI;
public class OCR {
    public static final String TESSBASE_PATH = "/mnt/sdcard/tesseract/";
    public static final String DEFAULT_LANGUAGE = "eng";
    public static final String EXPECTED_FILE = TESSBASE_PATH + "tessdata/";
    public static String getImageString(Bitmap bitmap)
    { String content=null;
      bitmap = bitmap.copy(Bitmap.Config.ARGB_8888, true);
      TessBaseAPI baseApi = new TessBaseAPI();
      baseApi.setDebug(true);
      baseApi.init(TESSBASE_PATH, DEFAULT_LANGUAGE);
      baseApi.setImage(bitmap);
      content = baseApi.getUTF8Text();
      baseApi.clear();
      baseApi.end();
      return content; }
}
```

Method init(String datapath, String language) is responsible for OCR initialization. First parameter "datapath" indicates the path. Second parameter "language" represents the language used by the recognition process. If OCR initialization is successful "init" will return "true". Otherwise will return "false".

Method setImage (Bitmap bmp) is responsible for setting the image to be identified. Method getUTF8Text() in order to get the text extracted from image, and return the result with UTF-8 standard. Method clear() is used to clear settings for the object.

## English-Chinese translation

It is a good choice to using SQLite database for English-Chinese translation. AHibernate general framework supports the following functions.

Setting table automatically. It can set tables automatically according to comments, and support operations as increase, delete, modify. Needn't to write these code repeatedly in SQLite. Simplify the workload of programmers.

Flexible query method. It supports query method for SQL and Android method. Query results can be encapsulated in object, just as Hibernate frame. Results can be shown as

List<Map<String,String>>. This method is very practical and efficient in actual project development.

English-Chinese translation can also be accomplished by using Youdao online translation interface. The XML data format is as following.

http://fanyi.youdao.com/openapi.do?keyfrom = <keyfrom> & key = <key> & type = data & doctype = xml & version= 1.1 & q = This is Youdao API

```
<?xml version="1.0" encoding="UTF-8"?>
<youdao-fanyi>
<errorCode>0</errorCode>
<!-- Youdao online translation -->
<query><![CDATA[This is Youdao API]]></query>
<translation>
<paragraph><![CDATA[Here is the youdao translation API]]></paragraph>
</translation>
</youdao-fanyi>
```

The json data format is as following.

http://fanyi.youdao.com/openapi.do?keyfrom = <keyfrom> & key = <key> & type = data & doctype = json & version = 1.1 & q = good

```
{
  "errorCode":0
  "query":"good",
  "translation":[" //Chinese spelling
  "],
  "basic":{
    "phonetic":"gǒd"
    "uk-phonetic":"gǒd"
    "us-phonetic":"gǒd"
    "explains":[" //Chinese meaning
    "]}
  }
```

The jsonp data format is as following.

http://fanyi.youdao.com/openapi.do?keyfrom = <keyfrom> & key = <key> & type = data & doctype = jsonp & callback = show & version = 1.1 & q = API

```
show({
  "errorCode":0
  "query":"API",
  "translation":["API"],
  "basic":{ // The basic vocabulary
  "explains":[
  "abbr. Application Program Interface; ..."
  ]},
  "web":[
  { "key":"API",
  "value":["// Web Interpretation
  "]} ] }
);
```

## Design of Android form

Realization of the Android form is based on an interface called WindowManager. WindowManager has a lot of use, managing the properties of the window, adding or deleting the view, collecting message, processing the order of forms. Instance of WindowManager can be generated by Context.getSystemService (Context.WINDOW\_SERVICE). WindowManager has

three important methods for form management call `addView ( )`, `updateViewLayout( )`, `removeView( )`.

XML language is applied to the layout of the interface. `RelativeLayout` is applied to the design of the main interface. So the application can be compatible with a variety of smart phones

## Conclusions

This application is developed by Android smart phone platform. Java and C++ have been used with Eclipse. The software can be compatible with all Android platforms above Android4.0. The solution of the application depends on the following aspects. Recognition processing mainly uses Google OCR technology. Image binarization is used to increase recognition accuracy. Realization of English-Chinese translation depends on local lexicon realized through SQLite database, or using Youdao online translation interface.

## Acknowledgment

This paper supported by Foundation of Guangdong Education Department(Grant No. 2013WYM\_0076) and Foundation of Guangdong Philosophy and Social Project(Grant No.GD13CJY07) and Foundation of Guangdong Undergraduate Science and Technology Innovation and Foundation of Guangdong Undergraduate Training Program for Innovation and Entrepreneurship(Grant No.201510576055).

## References

- [1]Chi Yuliang, Gan Xinling, Li Yong, "Design and realization of multi-people real time positioning and communication system based on Android platform," *Experimental Technology and Management*, vol.32,No.9, pp.160-162, Sept. 2015.
- [2]Ma Wenke, Liu Kun, "Design and Realization of Digital Campus Client Based on Android Mobile,"*Journal of Pingxiang College*, vol.31,No.3, pp.55-58, Mar. 2014.
- [3]NI Jian-xin, "Design and Implementation of Campus Information System Based on Android," *Modem Computer*, No.24, pp.93-96, Dec. 2013.
- [4]Xu Bingquan,Zhang Yuan,Yang Min, "Grantroid:A Method Supporting In-Context Permission Granting for Android,"*Computer Applications and Software*, vol.31,No.8, pp.232-236, Aug. 2014.
- [5]RUAN Wenhui, XUE Yadi, "Image Enhancement Algorithm Based on Android Platform," *Journal of Jilin University:Sci Ed*, vol.53,No.5, pp.1006-1012, May. 2015.
- [6]Li Hongsheng, "Personalized m-learning platform design and realization based on Android-App," *Net Security Technologies and Application*, No.12, pp.18-19, Dec. 2014.
- [7]YANG fei, MAO Hehua, WEN Yijun, "Design and Development of Personal Customer Information Management System Based on Android,"*Journal of Shazhou Polytechnical Institute of Technology*, vol.15,No.4, pp.11-16, Apr. 2012.
- [8]Ni Jianxin, "The Research and Implementation of a Android-based teaching resources mobile clients,"*Microcomputer & Its Applications*, vol.33,No.24, pp.16-17, Dec. 2014.
- [9]Zhu Sheng, Mou Xingliang, Shan Kangkang, "Based on the Android platform application development research," *Net Security Technologies and Application*, No.10, pp.46-47, Oct. 2013.
- [10]Song Kai,Li Lei,Yuan Chen, "Campus commercial service integration platform based on Android,"*Microcomputer & Its Applications*, vol.34,No.18, pp.92-94, Sept. 2015.