Gapp: An Image Capture Application For The Google Glass Framework

Oliver A. Nina
University of Central Florida
Orlando, FL
onina@eecs.ucf.edu

Roger D. Pack
FamilySearch
Provo, UT
roger.pack@familysearch.org

ABSTRACT

Traditional image capture devices for historical documents such as scanners or specialized cameras are not very portable and could be expensive. In many instances such devices require specialized software and people to make them function correctly.

In this paper we present Gapp, a genealogy application for the Google Glass framework. Gapp comes as an alternative to traditional image capture devices. Our application is designed to capture text images specifically of printed text.

Our application also allows capturing pictures on voice command and automatically uploading the picture onto a server. The server side application consists of a program that uploads the image from the device and runs the image through an OCR engine. The results of the engine then get displayed on a website that the user can access later and where the images taken will be automatically indexed, thus making the text in the images searchable.

By allowing pictures taken on voice command through Google Glass, we liberate the user from using other equipment or people that would be required for the taking of the pictures otherwise. The device also makes image preserving of text documents more portable and simpler than regular cameras.

Although, we experienced some issues with the low camera resolution and other issues due to the prototyping stage of the device, our initial results are encouraging and lay a foundation towards future similar applications that could be built for the Google Glass framework. In this paper we describe our application, results as well as drawbacks and issues of our system.

General Terms

Wearable Computing

Keywords

Image Capture, Google Glass, Historical Documents

Figure 1: The Google Glass Device

1. INTRODUCTION

With the advances of mobile computing and smaller processors, there is a new field of devices known as wearable computing devices. One recent development and release of one of such devices is the popularly acclaimed Google Glass or simply Glass which is a device composed of a CPU, camera, microphone and display in the form of regular glasses.

In this paper we introduce a new application for the Google Glass platform that allows a user to capture, index and store text images on voice command for later use. To our knowledge this is the first time an application used for genealogy purposes has been developed for the Glass framework. Although there has been some ideas for data acquisition for genealogy purposes such as in [4] [5] [6]. However, these methods usually address a much more specific problem such as capturing only headstones as in [4], or capturing documents with a mounted DSLR camera which requires many resources and is less ubiquitous such in [5] [6].

In this paper we describe our prototype, its features and drawbacks. We also discuss future features that would make the application the ultimate application for data acquisition.

2. BACKGROUND

Glass is a wearable computer that takes the form of regular glasses and is worn in the same way. Glass is composed of a small head mounted camera, a high definition display, a microphone and headphone as seen in Figure 1

Although Google has not released any specific details on the type of the processor used in Glass yet, it is believed that Glass runs on a processor similar to the dual-core OMAP 4430 with 1GB of RAM [1] and 16 GB of storage [2].

Glass, among other things is capable of recording videos and taking 5MB pictures through its mounted camera. Also, the android operating system (GDK) in Glass allows access

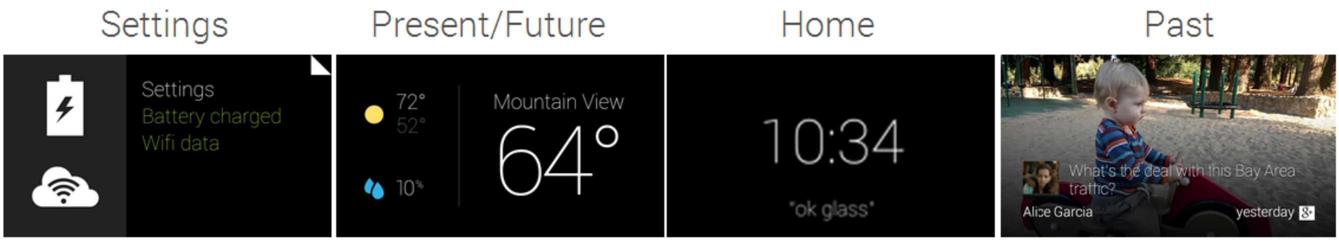


Figure 2: Timeline cards for Glass user interface

to an interface voice commands. Glass also allows connectivity to the internet through Wi-Fi and has pre-installed applications such as email, maps, navigation and Facebook. Google recently released an explorer edition of Glass for developers by invitation only to work on applications for the Glass.

Our application is a program that runs in the Glass OS Android system and allows the user to take a picture on voice command. We make use of two main components in the Glass Development Kit (GDK), timeline cards and voice commands.

2.1 Timeline Cards

In GDK terms, the timeline is the sequences of static or live cards that get displayed in the order they occur or on user demand.

We can think of each card as an application or glassware executed on the device. Figure 2 shows the timeline cards of a typical Glass device.

2.2 Voice Commands

There are three main ways to input commands into Glass such as with voice commands, touch gestures and head gestures. Our application is designed mainly to take input through voice commands, although the program can also get executed in the timeline through touch gestures.

3. OUR APPLICATION

To start the application the user is prompted first to the home card in the timeline. We do this by tapping the side of the Glass device which creates the event that takes us to the home card as seen in Figure 2.

Here we can see the current time of the system as well as the phrase "ok glass" which is the voice command for the contextual menu for executing the applications or glassware currently installed to work with the voice command interface.

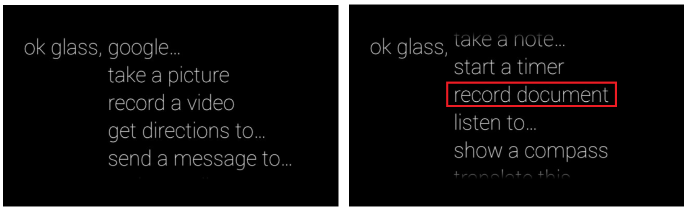
Once in the home timeline card, we activate the context menu of applications currently installed with the "ok glass" voice command. By prompting this command we can see all the glassware that accepts voice commands currently in the device. A list of such applications currently installed in our testing device can be seen in Figure 3.

The user then can access the Gapp application by saying: "record document" whereupon the device will take a picture of the current view of the camera.

After the picture is taken, the image is saved in the Glass device and is automatically uploaded to a server where it will also be saved and processed

The source code for this part of our application is made available as an open source project at:

`https://github.com/olivernina/gapp`



(a) Preinstalled Glassware (b) Command to Execute Gapp

Figure 3: Home Card Context Menu

3.1 Server Side Application

The server side application consists of a program that stores an image initiated by an upload Glass event. The upload time of each image to the server application is seemingly small (seconds) when connected trough a Wi-Fi connection.

After the consuming of the event, the server application will initiate the Optical Character Recognition (OCR) process performed by the Tesseract [3] OCR engine which is an open source engine.

Finally, the results of the OCR engine together with the image are displayed on our web application as seen in Figure 4.

4. RESULTS

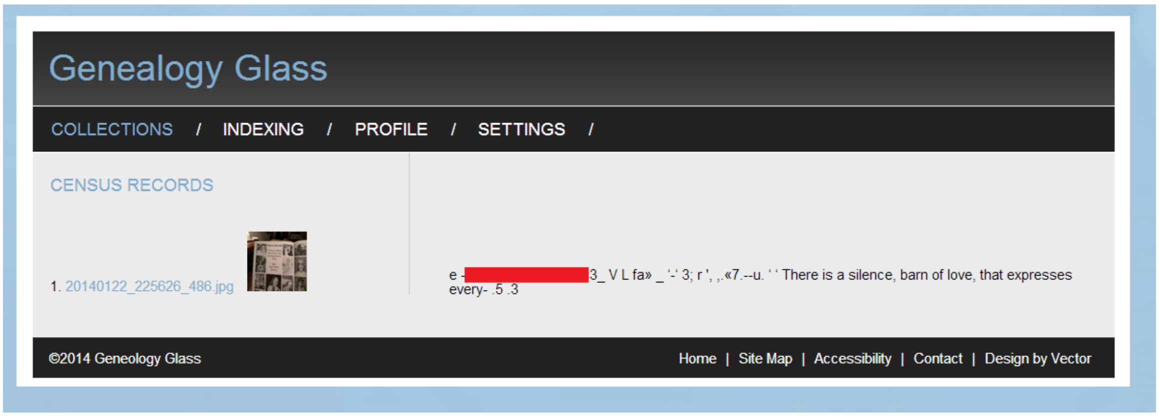
As we can see in Figure 5 we captured a page on a high school yearbook with Glass (faces and full names have been removed for privacy purposes). The captured image gets uploaded to the server application and gets indexed automatically.

The results of indexing the text in Figure 5 can be seen in Figure 4. Here, the name was correctly indexed and also the main phrase in the document. There were also false positives (represented by x) and false negatives such as the last three lines of text. The results of the OCR engine in the previous example is :"xx[Full Name]xx...xx There is a silence, barn of love, that expresses every- ."

5. ISSUES AND DRAWBACKS

As with any prototype application we have encountered certain issues with the Glass device and our application that makes our application a work in progress.

One of the main issues we encountered with the device is the low resolution of the camera that is not suitable for small text as seen in Figure 6. Hence the best test cases for our application are those where large text is present.


Genealogy Glass
COLLECTIONS / INDEXING / PROFILE / SETTINGS /
CENSUS RECORDS
Home | Site Map | Accessibility | Contact | Design by Vector

1. 20140122_225626_486.jpg

e -[redacted]3_ V L fa» _ '-' 3; r ' , ,.«7.--u. ' ' There is a silence, barn of love, that expresses every- .5 .3

©2014 Geneology Glass

Figure 4: Server Side Application

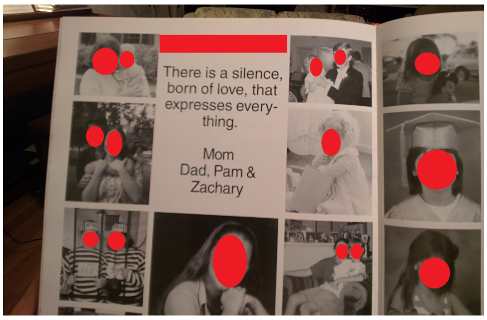

[redacted]
There is a silence, born of love, that expresses every-thing.
Mom
Dad, Pam &
Zachary

Figure 5: Image Captured From a High School Yearbook

Although, at the current stage of the Glass device, it is hard to correctly capture and indexed small text in the document, however, we expect that the camera resolution will improve in subsequent releases of the device which will allow sharper images and better OCR results.

Even in the cases where OCR fails due to the quality of the image, in many cases the images are still readable and are still saved and preserved with our application.

Other issues are for instance, no shaking stabilizer software for the device, so it is expected that the user does not move the device much to obtain a clear picture.

It is also important to note that the angle of view of the mounted camera does not negatively affect the system results in most cases. This is because the angle of view of the user's eyes usually coincide with that of the Glass camera. This indicates that if the user can read the text, then the text is at a good capture angle and can also be read by the system.

6. CONCLUSION

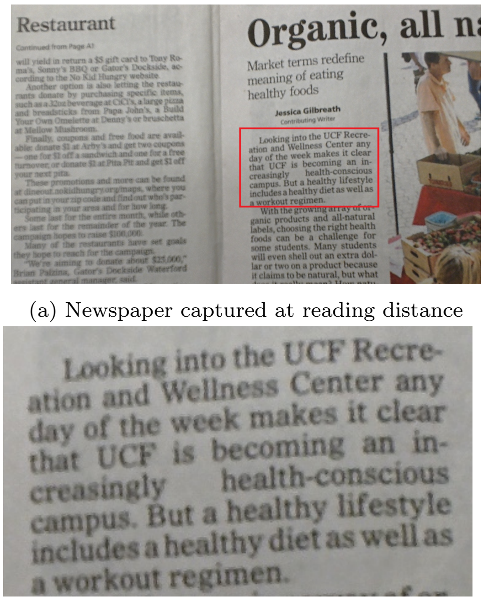
In this paper we present an application for the Google Glass device. Our application allows us to capture images on voice command, store and automatically index typewritten text images on a server where they are stored and can be accessed and queried at a later time.

Despite the issues we faced with low camera resolution, our initial experiments show promising results for future releases of Glass to capture and preserve text images while indexing them at the same time.

7. FUTURE WORK

Although Glass is still in prototype stage, we expect that it will improve its capabilities such as camera resolution in the near future.

Based on the future upgrades to Glass, we can also add more features to our application such as text detection that could allow us to detect the text and focus the camera based on the text present on the image thus aiding the OCR engine.


Restaurant
Organic, all na
Market terms redefine meaning of eating healthy foods
Jessica Gilbreath
Looking into the UCF Recreation and Wellness Center any day of the week makes it clear that UCF is becoming an increasingly health-conscious campus. But a healthy lifestyle includes a healthy diet as well as a workout regimen.

(a) Newspaper captured at reading distance

(b) Camera resolution too low for small text

Figure 6: Device Resolution Issues

Our system could be easily extended to android phones whose cameras are more advanced than Glass. Since the OS is the same, most functionality and code would work in similar way for phones and we can still take the advantages of portability with smart phones.

Also the ultimate recording application could use the same framework but instead of taking pictures it would record a video which is basically taking multiple pictures at a certain speed or frame rate. Thus such application could resemble a device with photographic memory which memorizes all images passed through the device to later be stored and indexed.

8. REFERENCES

[1] http://bit.ly/1eQmqsy.

[2] http://bit.ly/1j1tOZw.

[3] https://code.google.com/p/tesseract-ocr/.

[4] C. Christiansen and W. A. Barrett. Capstone: Gather valuable information found on headstones as an automated process. 2011.

[5] R. J. Laxman and B. R. Urry. Digital camera imaging of original documents. 2001.

[6] O. A. Nina and R. Pack. High dynamic range imaging (HDRI) for preservation of historical documents. 2011.