Abstract

In this part of the thesis, we present our initial proof of concept study towards the development of a low-cost 3-D printed smartphone add-on spectrometer. The study aimed at developing a cheap technology (less than 5 USD) to be used for detection of crop diseases in the field using spectrometry. Previously, we experimented with the problem of disease diagnosis using an off-the-shelf and expensive spectrometer (approximately 1000 USD). However, in real world practice, this off-the-shelf device can not be used by typical users (smallholder farmers). Therefore, the study presents a tool that is cheap and user friendly. We present preliminary results and identify requirements for a future version aiming at an accurate diagnostic technology to be used in the field before disease symptoms are visibly seen by the naked eye. Evaluation shows performance of the tool is better than random however below performance of an industry grade spectrometer.
7. A low-cost 3-D printed smartphone add-on spectrometer for diagnosis of crop diseases in the field

7.1 Introduction

The ability to detect crop disease in-field at a very early stage, especially amongst smallholder farmers, is still a big challenge. Usually disease identification happens at a stage when it is too late and little can be done to save the situation.

This research is an extension of our previous work in chapters 4 - 6 where diseases in Cassava crop have been identified at an early stage using spectral data acquired by a commercially available spectrometer.

Initially, we analysed spectral data from visibly diseased parts of a leaf as well as parts that appear visibly healthy (chapter 4). Findings of the study indicated that spectral data collected from asymptomatic parts of the leaf has implications for detecting disease in the plants before symptoms are visible. In chapter 5, we investigate spectral data in terms of matrix relevance learning. The ultimate aim was to identify a specific feature representation or particular features (i.e. wavelengths or ranges of wavelengths) which contain most information for classification and will facilitate technical solutions using simple sensors. Unlike in chapter 4 and 5, where the experiments considered mature plants aged 6 - 9 months and grown in open fields, experimental setup in chapter 6 was carried out in a screen house environment. The controlled setup rules out the influence of other diseases, pests or severe weather conditions. The results of the study indicate that the presence of the disease can be detected from leaf spectra several weeks before the appearance of visual symptoms.

Although good results have been attained with this approach, the ability to construct cheap and usable tools that help a smallholder farmer to diagnose crop diseases at an early stage is still vital. To tackle this problem, we propose the development of a 3-D printed smartphone add-on spectrometer that can determine the status of a crop from the leaves in field.

There have been some initiatives to build a small relatively cheap spectrometer to do these types of analyses e.g. (StellarNet Inc. 2019, Fennell et al. 2018). However, most of the studies on the application of this novel technology are still in the experimental stages, and are carried out in isolation with no comprehensive information on the most suitable approach. Also, most of the current options are expensive and require high interpretation power that makes the technology inappropriate for smallholder farmers who are the typical users in our context of the developing world.

The key contributions of this research can be seen in two areas. (i) To provide evidence of building a really low-cost technology that can be easily scaled even without a large production line. (ii). To develop an artifact that will actually be able to identify disease early such that smallholder farmers can cause an intervention to their
7.2. Materials and methods

We begin this section by describing the working of a commercial spectrometer that we used in our previous experiments and the customised spectrometer we constructed in this study. We then present the data and the pre-processing methods.

7.2.1 Commercial spectrometer

Previously, we experimented with a CI-710 spectrometer manufactured by (CID Bio-Science Inc 2010). The device consists of two modules, a leaf probe and a CCD-
based spectrometer, and is powered by a personal computer (PC) through a USB cable. A computer program, SpectraSuite, is used to set measurement parameters and display the spectral data. The leaf probe is equipped with a tungsten-LED dual light source that provides a broad range of wavelengths of light, suitable for visible and near infrared spectroscopy. The light then passes through a bifurcated fiber optic cable and connects to one of the two sampling light ports on the side of the leaf probe for absorption spectrum or transmission spectrum, as well as reflectivity measurements for reflectivity spectrum.

7.2.2 Customised spectrometer design

Light source & Diffraction grating

The customised setup in this experiment used a watch battery of capacity 1.5 V and white light diode as a source of light as a low cost alternative to optical fibre and more powerful light sources.

The heart of spectroscopy is found in the refracting element. This can be a prism or a diffraction grating. The component role is to split the incoming light beam into a spectrum of rays. Our design approach used diffraction grating (Young, 1802). We used a peeled digital versatile disc (DVD) piece as the diffraction grating medium. DVDs have been used as diffraction grating in other versions of DIY spectrometers (Public Lab 2019). Hence, the component separates white light from a light source into the underlying spectra as depicted in Figure 7.1. The current setup achieved measurements for absorption spectrum.

![Figure 7.1: Diffraction grating. The numbers near the screen are n values, the order of the image. Taken from (Burchill 2019).](image)

The digital display

We use a smartphone camera as a receiver. In principle, a transformation of light into spectrum measured by wavelength would be required. The light split by the DVD
7.2. Materials and methods

Figure 7.2: Architectural design

falling on the camera is captured as spectrum image, see Figure 7.5. The study also acquired spectral data using the AspectraMini Application (Google 2017) defined above. Both types of data have been presented for analysis in section 7.2.3. Our hypothesis in this study is that different diseases affect the metabolism of plants differently thus causing changes in light absorption properties.

Hardware design

We built a 3-D printed smartphone casing to integrate different components and our design aimed at the following: (i). To prevent much light scattering and splitting in order to retain enough information. (ii). To keep external light out of the experiment as to avoid information distortion. We also made the case adjustable to any phone controlling for where the camera is and the size of the phone. Figures 7.3 and 7.4 show a full setup and component designs respectively.

7.2.3 Methods

Dataset

Using the experimental setup described above, we acquired data of two types: (i). Color histograms as a transformation from color RGB spectra, see Figure 7.5 and 7.6. In section 6, we have recommended further transformation of light as measured by wavelength. (ii). We acquired spectral data using the already existing software (Aspectra mini Android application) (Google 2017), see Figure 7.7.
7. A low-cost 3-D printed smartphone add-on spectrometer for diagnosis of crop diseases in the field

![Image](image_url)

**Figure 7.3:** First prototype

![Image](image_url)

**Figure 7.4:** Adapter design for the 3D-printed smartphone case. Actual designs are available at [https://github.com/godliver/3-D-Printouts.git](https://github.com/godliver/3-D-Printouts.git).

The experiments in this chapter were performed in parallel with the previous chapter under the same setup. Guided by agricultural and bio-chemical experts from the Uganda National Crops Resources Research Institute (NaCRRI), cassava stems of variety Narocass 1 and NASE 14 were acquired from different fields and first tested to confirm that they were from healthy plants. These plants were grown in a screen house and at week four (4) of growth, plants were split into two (2) separate groups. The first one (10 plants) was reserved as a healthy control class (HC) and no disease inoculation was applied to the group. The second group of plants (17 plants) was infected with the cassava brown streak disease (CBSD) virus. CBSD virus was transmitted to these plants using a non-vector technique also known as grafting inoculation rated amongst the most efficient ways of inoculation (Rwegasira and MER 2015, Wagaba et al. 2013).

For experiments in this current study, we sampled five (5) plants from each category (HC and CBSD). For each week, three lower leaves on each plant were identified and tagged. Both types of data was captured for fifteen (15) consecutive weeks. In total, the number of samples collected per week, per class were 30 data points for HC and CBSD. Figure 7.5 and 7.7 illustrate sample data acquired in two forms.
7.2. Materials and methods

Figure 7.5: Spectral data in an image array form acquired with the setup in Figure 7.3

![Image](image.png)

Figure 7.6: Color histograms, a transformation from color RGB spectra in Figure 7.5

7.2.4 Data pre-processing

The study aimed at investigating the efficacy of the cheap spectrometer as a tool for in-field disease detection. Findings from this experimentation will be the basis for future improvements on the hardware and the data processing. Spectral data in Figure 7.7 was quite straightforward to process and we adopted pre-processing techniques that have been defined and described in previous chapters. This process included smoothing and standardising the data to reduce anomalies and to prepare data for the actual classification. For image data, we applied pre-processing computer vision techniques where the image was transformed from RGB to HSV color.
A low-cost 3-D printed smartphone add-on spectrometer for diagnosis of crop diseases in the field

Figure 7.7: Corresponding spectral data acquired with Aspectra mini application.

space. Here we attained histograms as image representation and presented this data for classification. Figure 7.8 shows the projection of the data points on the leading principal components obtained by standard PCA in the two sets of data.

Figure 7.8: Projection on eigenvectors of Principal components. On left panel are data points of color histograms. The right panel are data points acquired by the Aspectra mini application.

Table 7.1: Overall accuracy score Aspectra Mini vs. Color Histograms

<table>
<thead>
<tr>
<th>Classifier</th>
<th>Aspectra Mini</th>
<th>Color Histograms</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>0.627</td>
<td>0.521</td>
</tr>
<tr>
<td>Extra Trees</td>
<td>0.671</td>
<td>0.556</td>
</tr>
<tr>
<td>Linear SVM</td>
<td>0.612</td>
<td>0.503</td>
</tr>
<tr>
<td>GMLVQ</td>
<td>0.570</td>
<td>0.522</td>
</tr>
</tbody>
</table>


### 7.3 Results

We follow the same training and validation strategy as the previous study (chapter 6). We use Leave-One-Out cross-validation, where all data of a particular plant were disregarded in the corresponding training process.

Data for two classes: healthy (HC) and CBSD was grouped by unique plant labels in order to avoid training and testing on data from the same plant. During training, the partitioning was based on plant groups. We employed a validation scheme called Shuffle-Group(s)-Out cross-validation, also defined under (Pedregosa et al. 2011). Shuffle-split is an alternative to k-fold cross-validation that allows a finer control on the number of iterations and the proportion of samples on each side of the train / test split. Combined as Shuffle-Group(s)-Out cross-validation ensures that the same group is not represented in both testing and training sets. In our case, the groups were the plant ID since we obtained data from the same plant for consecutive weeks. For standard algorithms, we followed Scikit-learn (Pedregosa et al. 2011) implementation. In a similar way, this validation strategy was implemented for LVQ in MATLAB(R2016a) for the open source GMLVQ toolbox (Biehl 2017) that we employed for the GMLVQ algorithm. GMLVQ experiments were done with one prototype per class and batch gradient distance with adaptive step size control. If not specified otherwise, we used default parameters as suggested in the documentation of the toolbox (Biehl 2017).

The data was analysed by four classifiers: K-Nearest Neighbour (KNN), Extremely Randomized Trees, Linear SVM and GMLVQ defined previously (see chapter 5). Result Table 7.1 shows results for this pilot study. As indicated, we did not attain favourable results but this work serves as a baseline for future work. In Table 7.2 and 7.3, we expand on the results by presenting confusion matrices for Extremely Randomized Trees (Extra trees) algorithm applied on the datasets.

<table>
<thead>
<tr>
<th>Table 7.1: Overall accuracy score Aspectra Mini vs. Color Histograms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Classifier</strong></td>
</tr>
<tr>
<td>----------------</td>
</tr>
<tr>
<td>KNN</td>
</tr>
<tr>
<td>Extra Trees</td>
</tr>
<tr>
<td>LinearSVM</td>
</tr>
<tr>
<td>GMLVQ</td>
</tr>
</tbody>
</table>
Table 7.2: Confusion matrix for Aspectra Mini with Extra trees

<table>
<thead>
<tr>
<th></th>
<th>Healthy</th>
<th>CBSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>56.4</td>
<td>43.6</td>
</tr>
<tr>
<td>CBSD</td>
<td>44.9</td>
<td>55.1</td>
</tr>
</tbody>
</table>

Table 7.3: Confusion matrix for Color Histograms with Extra trees

<table>
<thead>
<tr>
<th></th>
<th>Healthy</th>
<th>CBSD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Healthy</td>
<td>59.5</td>
<td>40.5</td>
</tr>
<tr>
<td>CBSD</td>
<td>28.6</td>
<td>71.4</td>
</tr>
</tbody>
</table>

7.4 Discussion

We have presented an initial step towards the construction of an innovative low-cost spectrometer that can be used to diagnose disease in plants and infield. Our novel contribution in this area can be seen in the design of the prototype. Previous work (chapters 4 - 6) showed that spectrometry can gain the smallholder farmer an extra 8 weeks to apply an intervention before disease symptoms become visible. Our experiments in this study aimed at replacing the expensive spectrometer ($10K) with a cheap version to cost $5 – $8. While performance is clearly inferior to the one of the commercial spectrometer, we observe performance above mere guessing. This forms the basis for further improvements. One element to include in a future version is a powerful diffraction grating medium e.g. a prism. We will also experiment with different diodes. The explicit transformation of light emissions to actual spectrograms should facilitate further improvements. Literature (Public Lab 2019) also shows efficacy for this type of handcrafted cheap DIY tools. We intend to leverage on that success to provide a diagnostic tool to be used by smallholder farmers in developing countries.