Computer Science
SWJTU JOINT SCHOOL OF LEEDS

UNIVERSITY OF LEEDS

Intermediate Report

Investigation of Image Processing Techniques for Metal Detection in Water

Kang Liu

Submitted in accordance with the requirements for the degree of BSc Computer Science

Type of project: Exploratory Software

The candidate confirms that the work submitted is their own and the appropriate credit has been given where reference has been made to the work of others.

I understand that failure to attribute material which is obtained from another source may be considered as plagiarism.

Kang Liu

Contents

Chapter 1

Introduction

1.1 Project Background

Due to environmental concerns, it has become significant to accurately detect and monitor heavy metals in the industrial processes and environment. However, most of industrial techniques require complicated conditions and cost a lot. Thurs, there is a high demand for simple, efficient and portable detection methods.

1.2 Project Aims

This project focus on metal detection in water through image processing techniques. The main aim is to investigate, design and implement an image processing algorithm to achieve metal detection. This algorithm will use an image token by the user as input, it will then analyse and extract the features in the image to get the concentration of the metal. In addition, the evaluation of the algorithms is also important.

Alongside the main aim of the project, the secondary aim of the project is to develop and provide a generic mobile application platform. And a web service for the mobile application to access the image processing algorithm is also needed. This is to demonstrate the feasibility of the detection methods.

1.3 Objectives

1. Develop a base mobile application that allows users to upload data and get results.
2. Develop a web service that receives users' requests and sends the results back.
3. Design and evaluate image processing algorithms that detect the concentration of metals in water.
4. Reflect on limitations and possible future improvements to the mobile application, web service and image processing algorithms.

1.4 Deliverables

1. An android application developed by Android Studio.
2. A web service developed by Python Flask and deployed in Pythonanywhere.
3. Two image processing algorithms have been implemented and compared for evaluation.
4. Future improvements discussed in this report.

1.5 Initial Project Plan (Gantt Chart)

As is shown below in Figure 1.1, it is a Gantt Chart showing the initial project plan. The time line is arranged that there is one meeting with the supervisor every two weeks before the winter vacation and every one week after the vacation. The task: design and implementation of algorithms cost most of the time in the project, because it is the core aim of the project.

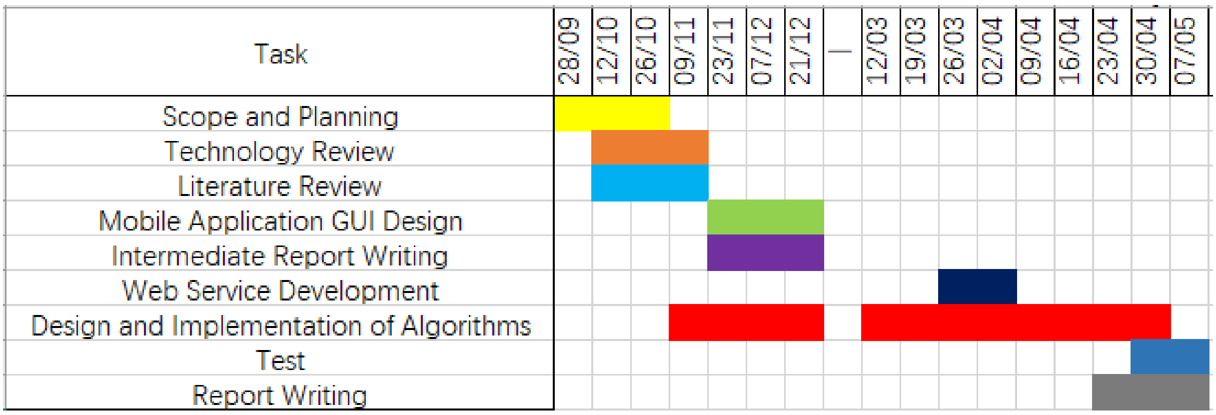

Figure 1.1: Gantt Chart for the project planning.

1.6 Version Control

This project use Git to control the version and manege the development of project. Git is also used for storing the code and important files in the project. The git repository url is shown in the Appendix A.

1.7 Legal, Ethical, Social and Professional Issues

There are actually no particular legal, ethical or social issues to be addressed in the project. The data involved in the project are experimental results and no any privacy or secure data exists. And the mobile application actually doesn't obtain or store any users' data or operations, while the server is the same. As the mobile application in the project is provided for workers or simulators in the factory, the only professional issue is that an expert is needed to maintain and update the algorithms in the server.

Chapter 2

Background Research

2.1 Introduction

Rapid development of industrial factories and progress in related technologies are causing increasing environmental problems [3, 7]. Most of heavy metals are non-biodegradable and can easily accumulate in ecological systems which makes them one of the most problematic pollutants. Heavy metals can also be harmful to public health. For example, cadmium (Cd) can accumulate in kidney and liver for over 10 years and affects physiological functions of a human's body eventually [8]. As a result, it has become significant to accurately detect and monitor heavy metals in the industrial processes and environment.

2.2 Current Detection Methods of Heavy Metals

Many techniques have been developed for the detection of heavy metals, including inductively coupled plasma mass spectrometry (ICP-MS), inductively coupled plasma-atomic/optical emission spectrometry (ICP-AES/OES), electrothermal atomic adsorption spectrometry (ETAAS), and so on. Many of these techniques have high sensitivity, specificity and accuracy. But these methods all require complicated facilities, professionals, and complex operation. Therefore, in developing countries and regions with a lack of sufficient infrastructure, professional experts, and appropriate environmental treatment, there is a high demand for simple, efficient and portable detection methods [6].

Considering above concerns, paper-based analytical devices (PADs) is one of the most promising solutions, which is cheap, simple, sensitive, specific, accurate, user-friendly, and environmental-friendly [6]. Thurs, we choose to use the results of PADs in this project.

2.3 Paper-based Analytical Devices and Sample Images

There are three main steps in the metal detection through PADs. First, the water sample is dipped on filter paper. Second, a specific reagent is applied to know the existence of particular metal. Third, the application of specific reagent to water on filter paper will produce only a single color showing the existence of particular metal and its concentration. There exist special types of chromogenic chemicals known as reagents which are reactant to a particular metal. For instance for cobalt(Co) and zinc(Zn), chromogenic reagents are 4-(5-bromo-2-pyridylazo)-N,N-diethyl-3-hydroxy-aniline (5-Br-PADAP) and dithizone (H2DZ), respectively.

Due to the limitation of the experimental conditions and available data, we focus on three metals' detection: cobalt(Co), hydrargyrum(Hg), and nickel(Ni). Figure 2.1 below shows examples of three metals' detection results, ppm (parts per million) is the unit of concentration which is the mass fraction of solute per million.

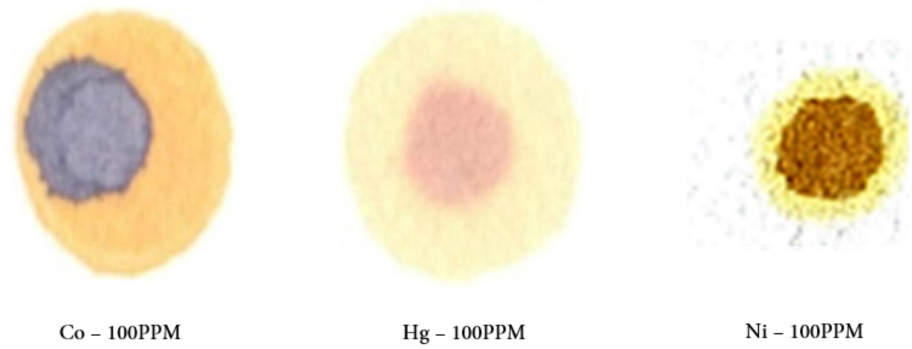

Figure 2.1: Examples of three metals' appearances on PADs with concentration 100PPM.

Figure 2.2 below shows a group of examples of cobalt(Co). It can be indicated that the concentration of a metal has relationship with the color in the PADs test result.

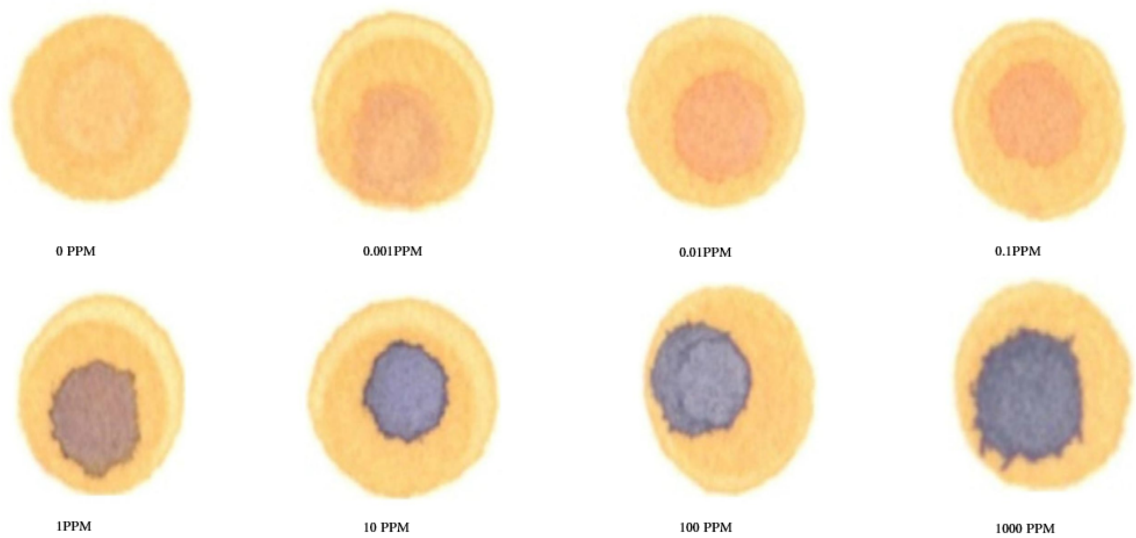

Figure 2.2: A group of examples showing cobalt(Co)'s appearances on PADs with different concentration ranging from 0 to 1000PPM.

2.4 Color Detection

Color detection is an important step in image processing. It is to classify pixels into different color regions from a given image [12]. Specifically, the aim of color detection in this project is to find the area showing the existence of particular metal and its intensity which is also called region of interest. As the results of this project are based on chemical reaction, environmental factors can easily cause decline in image quality like background noise, motion blur, or shadowing effects. So, image pre-processing needs to be carried out to partially reduce the impact of environmental factors on the images.

2.5 Color Space

There are many kinds of color spaces (also called color models) used for image representation, each with their own strengths and weaknesses [12].

First, RGB color space is the most frequently-used color model in computer storage. It has three color channel red, green and blue corresponding to the intensity of each primary color. The value of each color channel usually ranges from 0 to 255 in computer. The main advantage to use RGB color space is its simplicity which is easy to store, configure and display the image in all kinds of devices. However, some important color properties, such as brightness and purity, are embedded within the RGB color channels. So, it can be difficult to determine the specific colors and their reliable working ranges [12].

The other commonly used color spaces for image processing are the hue-based color models. They usually include: Hue-Chroma-Value (HCV), Hue-Chroma-Lightness (HCL), Hue-Saturation-Value (HSV) and Hue-Saturation-Lightness (HSL). In these models, hue represents the color angle, chroma or saturation represents the color purity, value or lightness represents the color brightness. Hue-based color models tend to be a favourable choice for colour detection because they separate important properties such as brightness and purity from the image. This allows better separation between different colours and thus they are the preferred colour models for various colour detection applications [11].

2.6 Color Distance in RGB Space

Color distance is a value to measure the difference between colors. It has different forms in different color space. According to the Euclidean distance between points in space, the fundamental form in RGB color space can be represented as:

$$\Delta E_{RGB} = \sqrt{\Delta R^2 + \Delta G^2 + \Delta B^2} \tag{2.1}$$

However, the human eye has a different sensitivity to the changes in the light intensity of the primary colors. These changes in the intensity of the colors cause different sensations [9]. This is taken into account by the following formula with the sensitivity coefficients to compensate for different color components of the eye. It has the form:

$$\Delta E_{RGB}{}' = \sqrt{3\Delta R^2 + 4\Delta G^2 + 2\Delta B^2} \tag{2.2}$$

2.7 Image Histogram Matching

Image histogram matching is a simple and basic method for image processing. It is usually used to compare, modify or select an image with a similar one [4]. This method changes the image into histogram by calculating the frequency of the pixels in grey scale or color distance in different color space. The histogram can show some features through the distribution of colors. As the histogram is easy to handle and analyse, the method is common in the pre-processing for data preparation in the bigger algorithms.

2.8 K-Means Algorithm

The K-Means algorithm is a clustering algorithm usually used in unsupervised learning. The algorithm is to partition a data set into k clusters, which groups all the data into K sets or classes. The method is to find k centers in the data set and assign other items to the nearest cluster which is represented by its center. After several iterations, when the centers have no changes larger than the error required, the algorithm finishes and the clusters determined by the last k centers are the objective [2].

In image processing, the K-Means algorithm can be used to classify the pixels into several groups with required features.

2.9 Linear Regression

Linear regression models are widely used to estimate the importance of parameters and predict the response variables at any point in the parameter space. [5]. There are many forms of linear regression model, one of the simpler forms of such models is:

$$y = wx_i + b \tag{2.3}$$

To solve the model, least square method is used to minimise the square loss while m is the size of the training set:

$$E_{(w,b)} = \sum_{i=1}^{m}(y_i - wx_i - b)^2 \tag{2.4}$$

And the solution is:

$$w = \frac{\sum_{i=1}^{m} y_i(x_i - \bar{x})}{\sum_{i=1}^{m} x_i^2 - (\sum_{i=1}^{m} x_i)^2/m} \tag{2.5}$$

$$b = \frac{1}{m}\sum_{i=1}^{m}(y_i - wx_i) \tag{2.6}$$

2.10 Existing Applications in Past Papers

There are two existing applications similar to this project in past papers. The first one is an on-site detection with a smartphone to detect free chlorine, hydrogen sulfide and formaldehyde in wastewater [1]. The other one is an on-site detection of heavy metals in wastewater which is the same as this project [10].

These two applications both use mobile application to take photo of paper test strip and get the result through on-site detection service. Their detection methods are also the same. They both use linear regression to get the linear model on a certain range of concentration and achieve a high precision and accuracy. The main advantage of this method is low cost and requirement to obtain a high accuracy.

These two applications have inspired me of the design in image processing algorithms and the demonstration application.

References

[1] S. Arsawiset and S. Teepoo. Ready-to-use, functionalized paper test strip used with a smartphone for the simultaneous on-site detection of free chlorine, hydrogen sulfide and formaldehyde in wastewater. *Analytica Chimica Acta*, 1118, 2020.

[2] A. Bansal, M. Sharma, and S. Goel. Improved k-mean clustering algorithm for prediction analysis using classification technique in data mining. *International Journal of Computer Applications*, 157(6):35–40, 2017.

[3] M. Bhuiyan, S. B. Dampare, M. A. Islam, and S. Suzuki. Source apportionment and pollution evaluation of heavy metals in water and sediments of buriganga river, bangladesh, using multivariate analysis and pollution evaluation indices. *Environmental monitoring and assessment*, 187(1):4075, 2015.

[4] T. M. Chan and J. Zhang. *An Improved Super-Resolution with Manifold Learning and Histogram Matching.* Advances in Biometrics, 2005.

[5] P. J. Joseph, K. Vaswani, and M. J. Thazhuthaveetil. Construction and use of linear regression models for processor performance analysis. In *Twelfth International Symposium on High-performance Computer Architecture*, 2006.

[6] Y. Lin, D. Gritsenko, S. Feng, Y. C. Teh, X. Lu, and J. Xu. Detection of heavy metal by paper-based microfluidics. *Biosensors and Bioelectronics*, 83:256–266, 2016.

[7] Y. Lu, S. Song, R. Wang, Z. Liu, J. Meng, A. J. Sweetman, A. Jenkins, R. C. Ferrier, H. Li, W. Luo, and T. Wang. Impacts of soil and water pollution on food safety and health risks in china. *Environment International*, 77:5–15, 2015.

[8] A. L. Marzo, J. Pons, D. A. Blake, and A. Merko?I. All-integrated and highly sensitive paper based device with sample treatment platform for cd2+ immunodetection in drinking/tap waters. *Analytical Chemistry*, 85(7):3532–3538, 2013.

[9] W. S. Mokrzycki and M. Tatol. Color difference delta e - a survey. *Machine Graphics and Vision*, 20(4):383–411, 2011.

[10] S. Muhammad-Aree and S. Teepoo. On-site detection of heavy metals in wastewater using a single paper strip integrated with a smartphone. *Analytical and Bioanalytical Chemistry*, 412(6):1395–1405, 2020.

[11] P. Sebastian, V. V. Yap, and R. Comley. The effect of colour space on tracking robustness. In *IEEE Conference on Industrial Electronics and Applications*, 2008.

[12] F. Su, G. Fang, and J. Zou. A novel colour model for colour detection. *Journal of Modern Optics*, pages 1–11, 2016.