Eindhoven University of Technology

MASTER

Sorting of plastics by means of near-infra red imaging spectrometry

Ossewaarde, J.

Award date:
1998

Link to publication

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain

7609

Sorting of Plastics

by means of

Near-Infra Red Imaging Spectrometry

J. Ossewaarde

Master Thesis carried out at the Department of Electrical Engineering of Eindhoven University of Technology, The Netherlands

December 1997 - September 1998

Commissioned by

Prof. Dr. Ir. P.P.J. van den Bosch
Measurement and Control Group
Eindhoven University of Technology, TUE

Under Supervision of

Mountside Software Development
J. Backus, Sales Engineer
P.O. Box 2263, 5500 BG Veldhoven, The Netherlands

The Department of Electrical Engineering of the Eindhoven University of Technology accepts no responsibility for the contents of MSc. Theses or reports on practical training periods.

Coaches

MSc. H. Liu
Measurement and Control Group
Eindhoven University of Technology, TUE

Dr. Ir. J.A. Hegt
Electronic Circuit Design Group
Eindhoven University of Technology, TUE

7609

Samenvatting

Dit rapport vat het afstudeerwerk samen dat gedurende negen maanden verricht is in opdracht van Mountside Software Development BV. In de eerste drie maanden van deze periode is een haalbaarheidsstudie naar het recyclen van plastics uit een (huishoud) afvalstroom uitgevoerd. Plastics zijn hoogwaardige materialen en het sorteren ervan met de hand is erg onplezierig, onnauwkeurig, duur, en bovendien erg onhygiënisch. Een automatische sorteeroplossing moet daarom uitkomst bieden. Hiervoor zijn twee sorteerstappen gedefinieerd:

1. Sorteerstap I – het onderscheiden van plastics en niet-plastics.
2. Sorteerstap II – het scheiden van plastics onderling.

Op het moment lijkt alleen Sorteerstap II economisch haalbaar, en daarom wordt in het tweede deel van dit werk onderzocht of het gebruik van nabije infrarood spectrometrie (NIRIS) een goede identificatiemethode kan zijn. Het haalbaarheidsonderzoek gaf duidelijk aan dat NIRIS vele voordelen heeft ten opzichte van andere technieken. NIRIS is een robuuste, goedkope, en niet-destructieve techniek met vele andere toepassingsgebieden in de landbouw, farmaceutische industrie, en procesbeheersing. Met NIRIS is het niet mogelijk om zwarte en transparante objecten te herkennen.

Dit rapport moet antwoord geven op de vraag of het mogelijk is plastics te sorteren in het spectrale bereik van 950 – 1700nm. In dit bereik kan namelijk gebruik gemaakt worden van een zeer snel meetinstrument.

Voordat we deze vraag konden beantwoorden, hebben we eerst alle systeemvereisten op papier gezet. Daarna is de meettechniek en de manier van dataverwerking bestudeerd. Vervolgens is op basis van 'Principal Component Analysis (PCA)' de relevante informatie uit de spectra gehaald. Tenslotte hebben we op basis van de vereisten, een 'Probabilistic Neural Network' geselecteerd als de beste oplossing. De oplossing voldoet theoretisch gezien aan alle vereisten, maar door tijdsgebrek is het niet mogelijk geweest alle eisen te toetsen aan de werkelijkheid. De gepresenteerde oplossing kan ten minste zeven plastics (L/LDPE, HDPE, PVC, PP, ABS, Polyester, en PA4_6/ PA6) sorteren met een nauwkeurigheid van 100%, waarbij ongeveer 78% van de plastics wordt herkend. Het neurale netwerk is in staat nieuwe plastics te leren herkennen. De combinatie van PCA en het neurale netwerk maakt het automatisch extraheren van relevante data mogelijk.

Helaas is het sorteersysteem niet in staat gemengde plastics te sorteren. Het effect van sorteerfouten op de kwaliteit van de plastics dient nog onderzocht te worden. Misschien is het mogelijk door middel van kleine aanpassingen de kwaliteitseisen te halen. Maar eerst zouden we een betere indruk van de sorteerkwaliteit moeten krijgen door meer spectra te classificeren. Tenslotte zal het gebruik van een nog onder ontwikkeling zijnde meetinstrument in het spectrale domein van 900 tot 2400nm onderzocht moeten worden. Hoogstwaarschijnlijk zullen dan alle sorteerfouten worden opgelost.

Summary

This report summarises the work developed during a ninth month long Master Thesis Trainee Ship at Mountside Software Development BV (The Netherlands).

In the first three months a feasibility study directed to the recycling of plastics from (municipal) household waste streams has been completed. Plastics are a high-value material, but sorting by hand is unpleasant, not accurate, expensive, and above all unhygienic. Therefore, an automatic sorting concept is advocated. Two sorting steps have been defined:

1. *Sorting Step I* – Discrimination between plastics and non-plastics
2. *Sorting Step II* – Discrimination between plastics mutually.

At the moment only Sorting Step II seems feasible in economical sense, and the second part of the work must investigate the technical feasibility of the use of Near Infrared Spectrometry (NIRIS) as an identification method. The feasibility study outlined the many advantages of NIRIS compared to other techniques. NIRIS is a robust, a cheap, and a non-destructive technique with many applications in other areas such as agriculture, pharmacy, and process control. The main disadvantage of NIRIS are black and transparent objects, they can not be identified.

The report must answer one main question, is it possible to identify plastics by means of NIRIS with a fast measuring device in the spectral range from 950 – 1700nm?

Before we were able to answer this question, we put a list with requirements together. We studied the near infrared measuring technique, and the way in which we should process the spectral images to get all essential information. Thirdly, we extracted most relevant features by means of principal component analysis (PCA). Finally, we selected the Probabilistic Neural Network (PNN) as the most appropriate solution to meet all requirements.

Theoretically, the solution meets all requisites, but due to lack of time not all of them could be verified. The presented solution is able to classify at least seven plastic groups (L/LDPE, HDPE, PVC, PP, ABS, Polyester, and PA4_6/ PA6) with an accuracy of approximately 100% and a recognition rate of 78%. The classifier recognises new yet unknown plastics, and is able to learn them. The combination of the PCA algorithm and the classifier allows optimal feature extraction. Unfortunately, the sorting system is not able to classify mingled plastics correctly. The effect of misclassifications on quality should be investigated more thoroughly. For instance, what quality is acceptable for certain plastics, and is it possible to tune the classifier a little bit more to meet these quality measures? Furthermore, more spectra should be measured to obtain more statistical evidence on sorting accuracy. Finally, an identical device in the spectral region from 900 to 2400nm will be available very soon, and its use will most likely alleviate the problem of misclassifications.

Table of Contents

1 Introduction

1.1 Problem

Nowadays, public has become aware of the preservation of our environment. Natural resources are becoming scarce, landfills are growing, and the green house effect is frequently on the agenda of world top conferences. We now seem to understand that our environment is highly related to our own lives and environmental measures have to be taken quickly.

Reducing the amount of material for manufacturing processes is one of these measures. Nevertheless, re-using and recycling materials seem to have a much larger impact. The duration of our natural resources is simply extended. However, the re-use and recycling of materials is not always beneficial. The environmental and energy costs associated with collecting and transporting small amounts of materials to be recycled can exceed any environmental benefits. Sorting jointly collected material into separated material stream at one distribution point is therefore advocated [38].

Up till now the sorting of several material streams was too expensive, but new techniques made the separation of (high-valuable) materials such as aluminium, iron, tin, etc. viable. Also plastic is a high-valuable product massively used by consumers and industry. Recycling and re-using them will have a great impact on our environment.

Plastics can be re-used and recycled in different ways. For instance, one can collect plastics separately. High-value and easy to handle types of plastic products like PET bottles - which are recollected and re-used in a collection system with deposit - are an example of this. But as stated before, this is not an overall solution. Besides this example of mechanical (material) recycling, also chemical recycling and energy recovery are recycling methods.

Since large scale jointly collecting waste material is advocated, plastics have to be isolated from this waste stream. Since sorting by hand is unpleasant, not accurate and expensive, an automatic sorting concept is advocated. Besides the mechanical handling of plastics, the identification should also take place by some alternative means. Near Infrared Imaging Spectrometry (NIRIS) seems a very promising technique in technical and economical sense.

1.2 Assignment

A three months long feasibility study outlined that the use of Near Infrared Imaging Spectrometry (NIRIS) for the identification of plastics from municipal household waste is technically very appealing. Unfortunately, a potential market in this area is not very prospective. Namely, governmental policy advocates the combustion of waste to produce energy in the form of electricity and heat. Large waste-to-energy plants have been built in the last years and their depreciation terms are based on 25-year contracts with municipals.

All the more, one should consider very carefully how to process the plastics once extracted from the waste stream. The diversity of plastics is very large. A plastic type like PVC exists of as many as six to eight different resins and thousands of different resin grades, blends, additives, fillers, flame retardant, and composites. Mixing two apparently similar plastics can lead to an inferior or even useless plastic for which no market exists. Unfortunately, the market of recycled plastics is not well established yet and good quality measures are not available nowadays.

Nevertheless, one clear conclusion can be drawn: in case that one really wants to get a recycled plastic with the characteristics of its virgin counterpart, only chemical recycling is the appropriate solution. In this process a homogenous plastic stream is turned into monomers – the basic building blocks of plastics. Because this method is too expensive these days, methods like NIRIS are much more appealing. The quality level as would be obtained with chemical recycling methods is not obtained, but the price of the recycled plastics is much lower and therefore more competitive in a potential market.

We distinguish two sorting steps:
3. *Sorting Step I* – Discrimination between plastics and non-plastics
4. *Sorting Step II* – Discrimination between plastics mutually.

The first sorting step is needed for (municipal) household waste streams, in which no separation of plastics took place. Sorting Step II can readily performed on homogenous waste streams from Sorting Step I, and of course waste stream from municipals and industry in which Sorting Step I already took place. For example, in Germany, plastics are collected separately by means of subsidiaries. The market of recycled plastics becomes in this way much more interesting, since politics stimulates new recycling methods.

Summarising, the market for Sorting Step II is more prospective than for Sorting Step I, and NIRIS becomes an interesting recycling method for this market. Above all, the technology and knowledge concerning NIRIS is also very suitable for the pharmaceutical, agricultural and food processing industry. These reasons are the motivation of investigating the feasibility of *Sorting Step II* in the succeeding six months of this Master Trainee Ship.

1.3 Near Infra-Red Imaging Spectrometry (NIRIS)

Near Infrared Imaging Spectrometry (NIRIS) is a commonly used non-destructive identification technique, which finds its merits in agricultural science, pharmaceutical science, and process control [33]. NIRIS allows on-line monitoring structural and molecular properties of several materials under investigation.

In the spectral range from 900 to 2500nm plastics absorb and reflect near infrared light differently. Since every plastic has a unique spectrum, all plastics can be identified on this 'fingerprint' by means of a classifier algorithm. Preliminary studies show that the measured NIR spectra of a wide range of plastics are clearly different than those of non-plastics. Distinguishing between plastics and non-plastics in Sorting Step I is therefore

easy. Unfortunately, black and transparent plastics cannot be identified. Black plastics absorb the whole spectrum and thus do not return a measurable spectrum. Also transparent plastics cause difficulties. The spectrum of the conveyor belt is measured as the optical near infrared signal passes through the object under measurement [55].

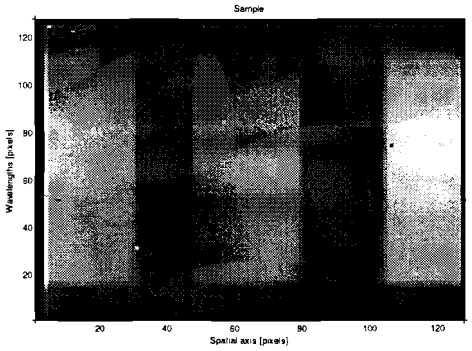


Figure 1 Recorded Spectral Image from Three Samples

In Figure 1 a recorded spectral image from three different plastic samples is shown. In the figure the spatial axis (horizontal) versus the spectrum information has been recorded. The grey level represents the measured (near infrared) light intensity. We now want to know what types of plastic are present there. This problem definition can be divided into three basic blocks, which we will refer to throughout the report (Figure 2):

1. *Image Processing* – This block retrieves the spectral information from the image. The information is processed in such a way it hardly depends on changes in ambient conditions like non-uniform lighting, noise, and spikes. But one could also think about irregularities in the samples themselves, like for example the elimination of baseline shift due to variation in sample thickness. Namely, we want to classify plastics on type and not on the variations within the same type.
2. *Feature Extraction* – In this block the most important characteristics from the pre-processed spectral information are extracted to allow optimal discrimination between plastics. The classifier will compare these features with the characteristic features of each plastics group. Namely, the classifier decides to which plastics group an unknown sample belongs on similarities and dissimilarities between these features.
3. *Classification* – In the final block the decision to which plastics group a sample should be categorised is made. One should keep in mind that classification is only possible if the information of different plastic types is mutually exclusive. Hence, if the information of two distinct plastics is similar, no classifier can do a correct classification.

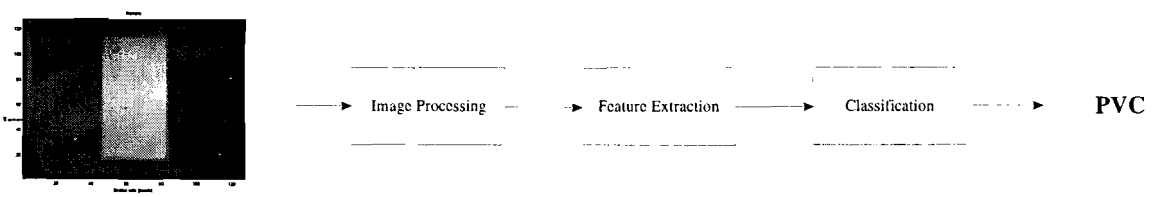


Figure 2 Three Basic Blocks of the Identification Unit

1.4 Requirements

For the identification unit the next requirements have been derived from previously performed feasibility studies. Although not all requirements should and can be solved in one particular block of the identification unit, we have categorised them for reasons of simplicity.

Globally, the first five requirements must be dealt with in the image processing part:

1. The sorting system should be insensitive to environmental influences such as ambient light and variations in temperature.
2. Insensitive to filler, fibers, colours, flame retardant, adhesives, dyes, pigments, and other additives.
3. Insensitive to dust, paint, labels, food and drink residue and other surface defects.
4. Insensitive to plastic objects composed out of more than one plastic such as a soft drink bottle composed of a bottom, a flake, and a cap. But also objects like milk flakes consisting of a combination of cardboard and plastic foil (laminated packaging).
5. Plastics with irregular dimensions and geometric should be handled.
6. Plastics with different weights and physical properties should be handled.
7. Black and transparent plastics should be excluded in this processing step. Namely, other research groups encountered severe problems when trying to classify these spectra [56].

For the feature extraction part the next requirement is important:

8. The features between distinct plastic groups should be as mutually exclusive as possible. If they are not, no classification can take place.

For the classifier the next requirements are of interest:

9. The identification unit should be able to predict to which plastics group a new as yet unseen sample belongs.
10. In case that this yet unseen sample does not belong to an existing group, a new plastics group should be made. In other words, the sorting system should respond to a new type of plastic, which may be introduced into the market in the future. Secondly, within a plastics group the variety can be so large that subclasses must be made, whereas one group does not represent the variety within the group.
11. Mingled plastics – composed out of more than one type of plastic - like for example 'Syntal' should not confuse the sorting system.
12. The sorting system should be insensitive to environmental influence such as ageing of the measurement equipment. Also, as a result of continuous changes in (new) manufacturing processes, legislation, and cost aspects, identification should be flexible to slight changes in characteristics of samples within one particular plastics group.

Finally, the following requirements deal with the economical aspects of the sorting system:

13. On-line identification and sorting should be done at very short time intervals to meet market demands. Typically, 1.5 – 2.5 tons of plastic per hour should be processed. Because this figure depends strongly on the mechanical configuration of the plastics handling, and market demands, it is difficult to translate this number exactly into an amount of identifications per second. Nevertheless, from previous investigations we can derive a maximum identification time of 15 milliseconds per sample [57].
14. It should be possible to achieve sorting rates with an accuracy of more than 99% - naturally depending on market demands.
15. The sorting system should operate fully automatic with little downtime.
16. Easy handling by unskilled people.
17. The sorting system should be economically interesting. As guidance, the complete sorting system including the mechanical plastics handling should be available for less than 195.000 DM in comparison with the equipment manufactured by Ti-Tech [52].
18. Because the measurement set-up with camera connected to a prism is the fastest, cheapest, and most robust in comparison with other equipment, we will use this set-up in our investigation. The camera set-up is fast and robust because it makes use of a prism to disperse the incident (infrared) light into a continuous spectrum. Figure 3 shows this principle for one point. The disadvantage of the equipment lies in its limited spectral range of 900 – 1700nm. Other camera set-ups make use of an Acoustic Optical Tuneable Filter (ATOF) or a fixed mechanical set-up of filters in the range from 900 – 2500nm. In a nutshell the camera set-up with an AOTF works as follows. Acoustic and optical waves move through a tellurium dioxide (TeO_2) crystal. The crystal is cut in such a way that resonation takes place in response to the acoustic waves. The resonation on its turn changes the refractive index of the crystal and the wavelength changes accordingly. Due to its sweeping time across all frequencies, the AOTF is quite slow compared with the prism.

Figure 3 Prism

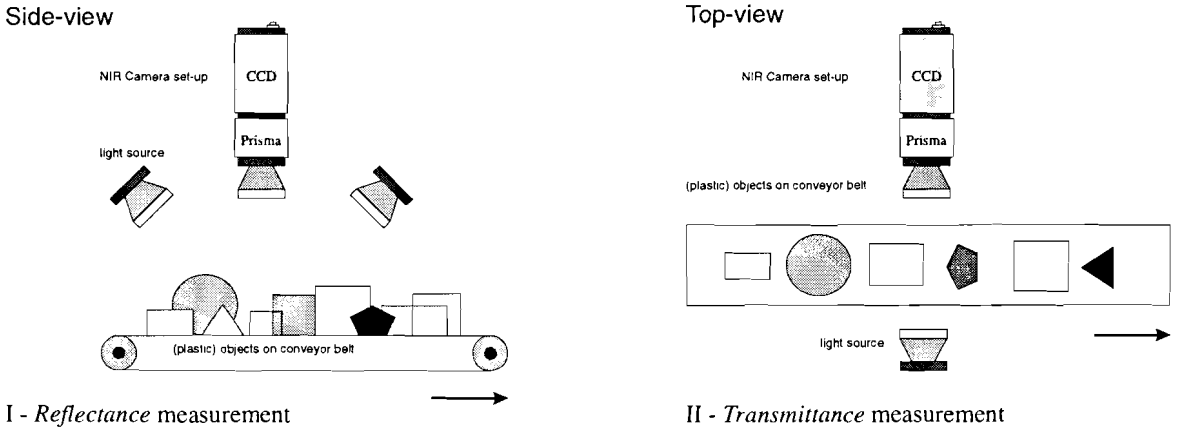
2 Near Infrared Imaging Spectrometry

2.1 Introduction

Before we can start implementing solutions for our problem definition, we must get a good understanding of the relationships between the plastic objects on the conveyor belt and their recorded spectra. With this knowledge we can implement the image processing and feature extraction part of the identification unit. But first we will discuss the basic technique of near infrared imaging spectrometry in more detail. Secondly, we will explain the choices we made in selecting a proper experimental set-up. In the next chapter we will explain what experiments we did to get basic knowledge about the characteristics of spectra.

2.2 Measurement Techniques

Generally speaking, two optical near infrared measurement methods are possible: *reflectance* measurement and *transmittance* measurement. Both methods are demonstrated in Figure 4.



I - *Reflectance* measurement

II - *Transmittance* measurement

Figure 4 Reflectance and Transmittance Measurement

The names of the methods refer to the way the light travels from the light source to the camera. In case of reflectance measurement the light from the sample under investigation is reflected into the camera, and in case of transmittance measurement the light is transmitted through the sample into the camera.

Selecting the camera set-up for reflectance measurements has two great benefits:

1. Thin plastics that are lying flat on the conveyor belt are still measurable.
2. The plastics do not have to be queued in a single stream to pass the camera set-up, resulting in higher processing speeds.

Although we only treat sorting step II (discrimination between plastics mutually), we remark that the mechanical set-up of Sorting Step I should allow reflectance measurements for the same reasons.

Reflectance measurement is based on the interaction between the molecules of the sample under investigation and the light radiated by the light source(s). Shortly, four types of interaction between the incident light (photons) on the sample and its molecules are possible (Figure 5):

I. *Absorption* – photons are completely absorbed by the molecules. The energy level of the molecules rises. Typically, this interaction occurs with black opaque samples. The energy of the incident light is converted into heat.

II. *Nothing* – the incident photons pass through the sample and no interaction with the molecules takes place. Typically, transparent samples exhibit this behaviour.

III. *Reflection* – the surface acts as a mirror and incident light is reflected.

IV. *Scattering* – photons interact with the molecules and are emitted (scattered) in random directions.

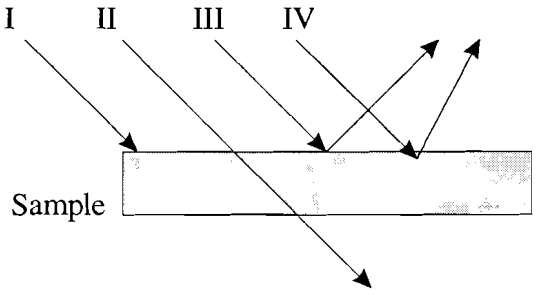


Figure 5 Four Types of Interaction between a Sample and Incident Light

Identification of molecules is based on the scattered radiant energy – in chemistry literature called overtones. For example, at the wavelength of 1145nm a typical CH-overtone is present. By measuring these overtones and relating them to the molecules of plastics, identification is possible. Generally, in the region between 1600 and 1800nm methylene C-H stretches are found, and in the region between 2100 and 2500nm combinations of C-H and O-H vibration bands [33]. With our choice of the camera in combination with the prism we limited ourselves to the first region.

2.3 Experimental Set-up

In Figure 6 the experimental set-up is shown. The camera and light source are illustrated with points. The measurement plane is a piece of conveyor belt that was black and made of plastic. On top of this piece of conveyor belt we marked three locations in the measurement plane. The spatial axis starts at the side of location 1 ($x = 0$) and ends just outside location 3 ($x = l$). At location 1 we put the sample under investigation, at location 2 we always put the same polypropylene (PP) sample, and at location 3 we always put a standard reference sample. The reason for doing this will be explained later.

In reflectance measurement the intensity of light received from a sample ($\mathbf{I}_{\text{sample}}$) is measured and expressed in proportion to the reference standard ($\mathbf{I}_{\text{reference}}$):

$$\mathbf{I}_{actual}(\lambda, x) = \frac{\mathbf{I}_{sample}(\lambda, x) - \mathbf{I}_{dark}(\lambda, x)}{\mathbf{I}_{reference}(\lambda, x) - \mathbf{I}_{dark}(\lambda, x)} \qquad \text{Equation 1}$$

where the intensity of the light $\mathbf{I}(\lambda,x)$ is expressed as a function of x and λ. The matrix $\mathbf{I}$ is a m x n pixel photodiode matrix in which each pixel intensity is stored as a grey level of k bits. For these experiments a photodiode matrix with a size of 128 x 128 pixels, and a grey level representation of 8 bits has been used. However, each recorded image $\mathbf{I}_{sample}(\lambda,x)$ is the sum of the reflected light from the sample and the dark current $\mathbf{I}_{dark}(\lambda,x)$. This dark current is caused by the heat radiation of the photodiode matrix. Because the dark current does not give useful information about the amount of light received from the sample image, we subtract it from all subsequent images. After the calculation of Equation 1 the light intensity of the sample $\mathbf{I}_{actual}$ is related to the reference standard at a linear scale between 0 and 1. In Figure 7 the sample image has a reflectance intensity of 40% compared with the standard reference.

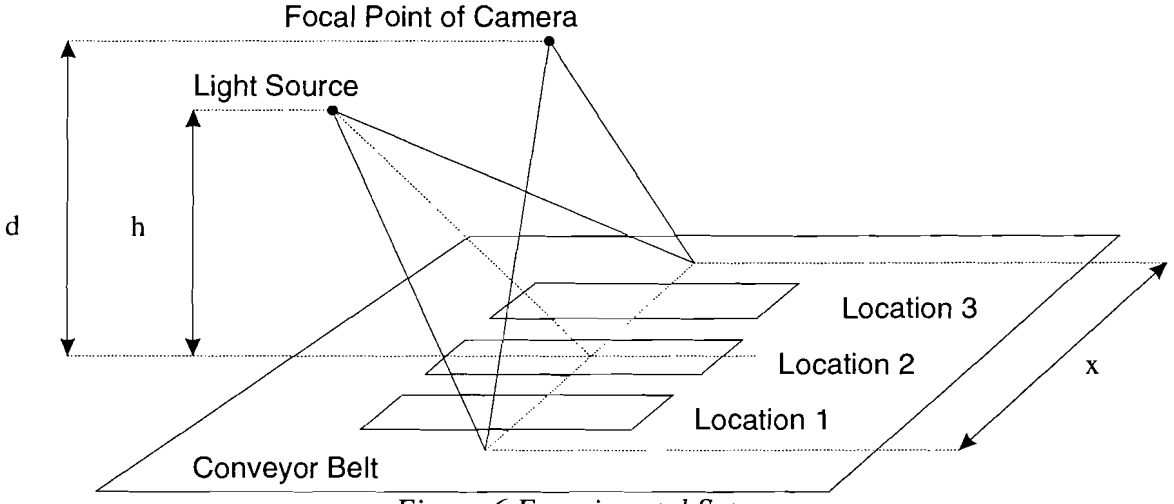


Figure 6 Experimental Set-up

Typically for infrared cameras, the dark current reduces the dynamic range of the detector with about 20%, resulting in an actual grey level of 6 bits.

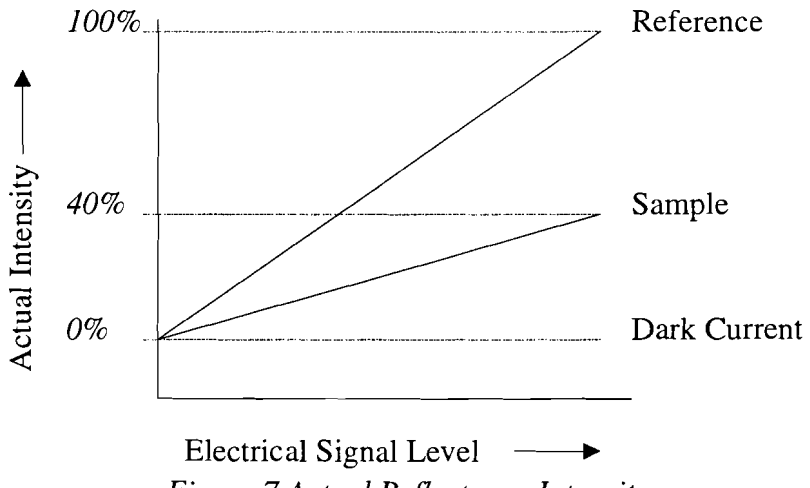


Figure 7 Actual Reflectance Intensity

The dark current image and the white reference image have to be measured only once, - of course - unless the behaviour of the photodiode matrix is varying in time. The dark current image is measured by simply closing the shutter of the camera and recording the spectrum. The standard reference image is taken by placing the white reference on the location where the actual sample would be measured as well.

In Figure 1 the intensity is represented as a grey colour map, the vertical axis is the spectral wavelength region. The wavelength region is typically calibrated with a lamp filled with Neon, Argon or Mercury gas. From each gas the relationship between spectral peaks and their corresponding wavelengths are known. By finding these peaks in a recorded image from the calibration lamp, we determine which pixels in the image are related to particular peaks in the calibration gas. In the recorded image Figure 8 the (four) highest peaks of these gasses correspond with the wavelengths 1014, 1129, 1357, and 1530nm.

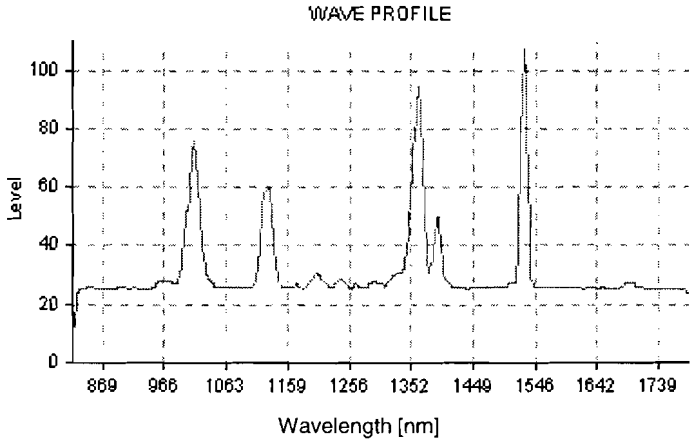


Figure 8 Spectrum from the Calibration Lamp

Because some spectral peaks are missing in this image, we conclude that the spectral range of the image acquisition device does not range from 826 to 1790nm, but from 950 to 1700nm. So the prism does work in the spectral range 826 – 1790, but the camera does not. This means that the first 17 x 128 pixels and the last 12 x 128 pixels of the photodiode matrix do not contain meaningful information, and should be ignored.

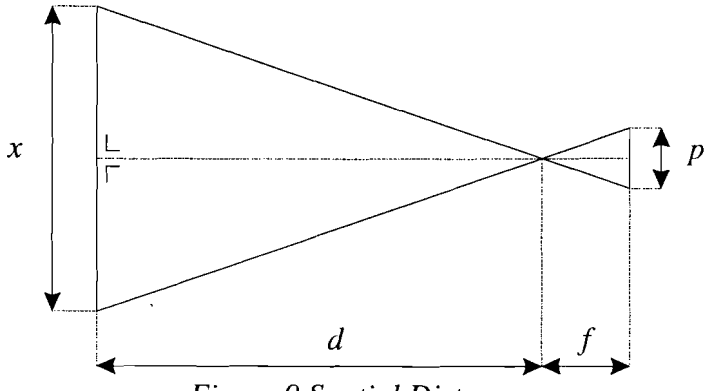


Figure 9 Spatial Distance

The image acquisition device scans a spatial line on the conveyor belt. The width of this line (spatial distance x) is directly related to the focus of the lens f, the sensor size p, and the distance d between the prism and the area of measurement as graphically illustrated in Figure 9. Theoretically, the spatial distance x is calculated as follows:

$$x = p \cdot \frac{d}{f} \qquad \text{Equation 2}$$

The sensor area consists of 128 x 128 pixels with a pitch of 60 µm in both directions. This results in a sensor size p of 128 x 60 µm = 7.68 mm. The focal length f of the lens is 25 mm. The lens is mounted (d =) 46.5 cm above the conveyor belt. Hence, by substituting all values in Equation 2 we find a spatial distance x of approximately 14.3cm. This distance can also be found by applying a more practical method. The start and endpoints of the spatial distance can be found by moving a white object under the camera and watching the effect on a video screen at the same time.

2.4 Measurement conditions

All experiments took place in a special temperature controlled darkroom at constant room temperature of typically 22 °C and a humidity of 36%. The integration time of the camera was 16ms.

By including the standard reference and PP sample in each measurement we are able to monitor changes in the measurement conditions. Namely, the standard reference sample gives information about changes in the light source intensity, and the PP sample about changes in the wavelength region.

During the measurements the intensity level should not exceed the grey level of 230 (range 0-255, 8 bits), whereas the detector device does not behave linearly for higher values. A standard reference sample of 75% has been chosen to increase the dynamic range of the intensity level. The light source intensity was set in such a manner that the value 230 was just reached with the reference sample of 75%. This intensity corresponds with a light source power of about 95 watts. To have optimal forward infrared radiation a halogen light source with pure gold plated coating has been used.

At the beginning of each experiment we measured a calibration image from the calibration lamp, a dark current image, and standard reference images of all three locations. Since we did not have a large reference sample we had to take these images by placing the standard reference sample sequentially at location 1, 2, and 3. Of each sample three images were taken by moving the sample slightly between each measurement. Each sample number has been written on the sample to allow identification afterwards. All samples are listed in Appendix C.

3 Image Processing

3.1 Introduction

Processing the raw image data must lead to spectra, which are not or only slightly dependent on ambient (measurement) conditions. By doing experiments we want to gain knowledge about the influence of ambient conditions on the spectra. We decided to split the experiments into two groups to reduce the risk and cost related with this project. In the first batch of experiments we want to find out what image processing must be done to extract useful features. The second batch will consider real-world circumstances under which the sorting concept must still meet our requirements.

The first and second batches contain 308 and 116 samples respectively. Generally speaking, the more spectra are measured from distinct plastics and non-plastics, the more accurate and consistent our knowledge will be.

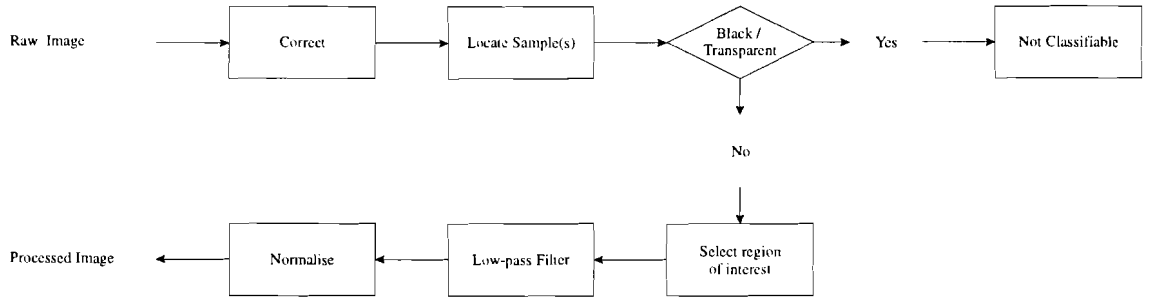


Figure 10 Image Processing Part

In the next paragraphs we will shortly treat the processing steps that resulted from the experiments. Figure 10 shows the chain of operations that must be performed on a raw image. In the first process step, the raw image is corrected with the dark current image $\mathbf{I}_{dark}$ and divided by the standard white reference $\mathbf{I}_{reference}$ as described with Equation 1. This correction is also very helpful in eliminating artefacts in the photodiode matrix. Pixels that always represent a grey level of 255 or 0 are eliminated. In case division by zero would take place in Equation 1, the actual pixel intensity is defined as zero. Therefore, the value zero indicate pixels with an artefact. Those pixels do not contain any information that could be useful for identification purposes.

3.2 Localising Samples

3.2.1 Problem

In the second process step of Figure 10 the relevant spectral regions must be localised in the image. Only, from these regions we get spectral information of the plastics under investigation. When two plastic objects are situated next to each other on the conveyor belt, we want to be able to identify each object individually. In situations in which an object on the conveyor belt is a composition of more than one type of plastic, we want to

be able to identify the plastics separately, too. Of course, this completely depends on market demands. For example, a milk bottle made of HDPE with a plastic cap of PP could be considered as a useless recycled plastic. Eventually, the localised spectral regions are used to eject the samples from the conveyor belt.

3.2.2 Solution

One way to find the relationship between the spectra and their locations on the conveyor belt experimentally is using a triangle as illustrated in Figure 11. At the indicated locations we recorded the spectral information along the *y* direction. By measuring the distance of the recorded object and by looking at its corresponding pixels in the image, we know location *x* exactly. Of course, instead of this triangle, individual samples of different sizes can be used as well.

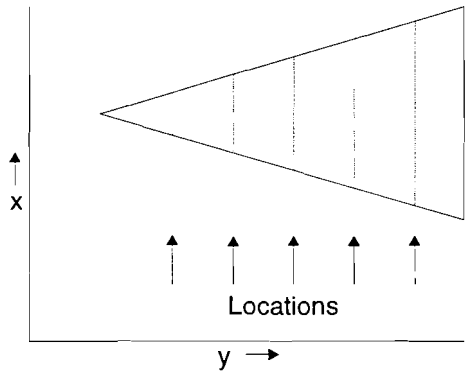


Figure 11 Measurement of Locations at Particular Intervals

The result of measuring three samples at the same time is shown in Figure 12. We did not use the triangle here because we want to explain some more characteristics. The raw image data is shown because the image could not be corrected with Equation 1. Namely, the width of the standard reference sample was only 4cm. The width of the line array of the camera is about 14.3cm for 128 pixels, resulting in a resolution of about 0.112 cm/pixel.

Although three samples were present at the measurement plane of Figure 12, only two are visible here. The first peak between pixel 48 and pixel 80 is the response of the PP sample was always present during the measurements. The second peak between pixel 106 and 127 is from the white standard reference. Between pixel 10 and 28 a black plastic was present. Its spectrum is similar to the spectrum of the (black) conveyor belt (pixels 80 – 106). Apparently, the spectra of a black sample and the conveyor belt are identical. This behaviour is also found for transparent plastics.

From the image we extracted a sample size of 3.58cm for the PP sample. This is in good approximation with the 3.56cm found by measuring its size directly with a calliper rule. The size of the standard reference is 2.90cm. This value is in good correspondence with the 2.92cm found with a calliper rule. (Do not confuse the 'size' of the standard reference with its real size of 4cm. In this image only about 75% of the reference standard had been

recorded. The standard reference was located at the border of the line array as can be seen in Figure 6).

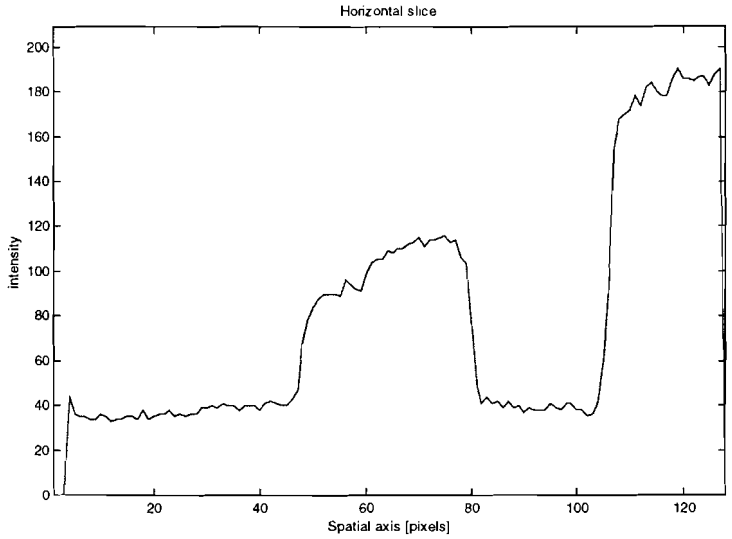


Figure 12 Horizontal slice of a Raw Image

Let have a closer look at the black PP sample. After correcting the region in which the black PP sample is located (Equation 1) the vertical slice at pixel 13 of Figure 12 is taken. This spectrum has been corrected with Equation 1. In Figure 13 we see a spectrum with a very low signal-to-noise ratio. The actual maximum intensity $\mathbf{I}_{actual}$ does not exceed the ratio 0.06. Hence, the maximum amount of light received from the black sample is less than 4.5% (0.06 x 75%). Verifying this ratio for 20 black and transparent plastics results in a maximum intensity ratio of 0.10. Hence, the light intensity level never exceeds 7.5% for black and transparent plastics.

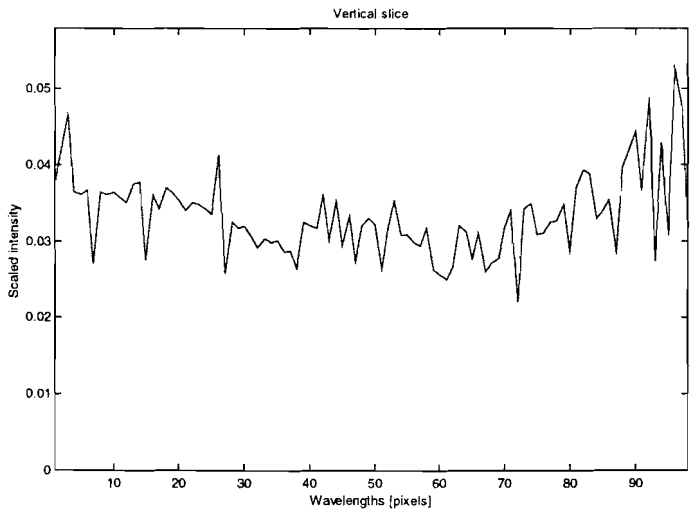


Figure 13 Vertical Slice of the Black PP Sample Region

3.2.3 Conclusion and Recommendations

It is proven here that the size and location of the samples can be derived from the image data by looking in the spatial direction. Black and transparent plastics cannot be located because they are not distinguishable from the conveyor belt.

Because we were restricted in the amount of measurements, samples that are composed of more than one type of plastics could not be measured. We recommend doing these measurements in future. Also an algorithm to determine the location and size of each sample could be implemented then. In these experiments we know the exact locations of the samples because they were always put exactly at location 1.

The algorithm could for example detect changes in the slope in the spatial direction. Probably, those slopes are also present for combinations of plastics. For instance, assume that the PP sample and the standard reference sample in Figure 12 are located side by side. A change in intensity is also present now. One should keep in mind that the artefacts of the photodiodes in this uncorrected image might influence the effect of the algorithm negatively. Namely, artefacts are masked with a zero and therefore result in sharp peaks when the derivative is calculated. The intensity ratio of 0.10 could be very useful as a threshold to alleviate calculating the derivative at pixels with an artefact.

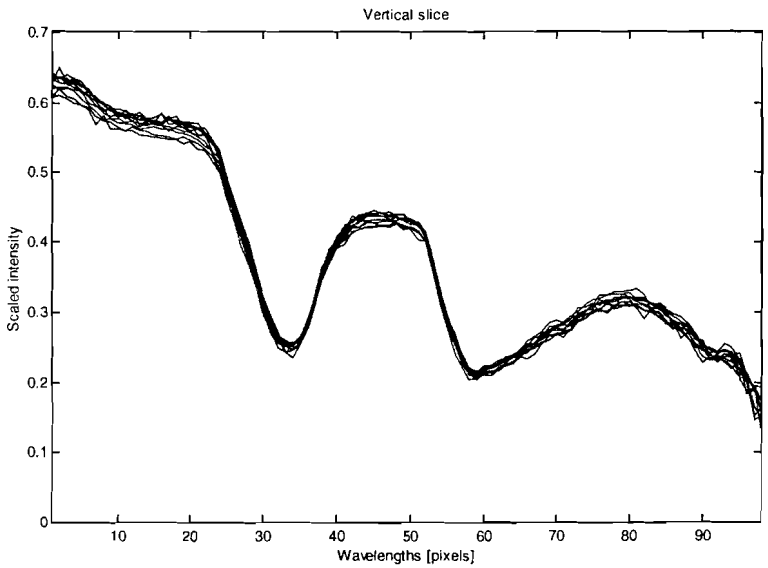


Figure 14 Wavelength Regions between Pixel 48 and 80 of the PP Sample

3.3 Selecting a Wavelength Region

3.3.1 Problem

After the location of the samples is found we want to select one particular wavelength region that contains the spectral information of the sample. Basically, this is the same procedure as used to slice Figure 12 in the vertical direction, but which wavelength region would be the best selection?

3.3.2 Solution

In Figure 14 all wavelength regions between pixel number 48 and 80 of the PP sample of Figure 11 are plotted. The wavelength regions have been cropped to the region between 950 and 1700nm. Namely, in the calibration phase we found that the camera does not have a spectral response outside this region. Although, all wavelength regions are taken from the same sample - but at intervals about 0.1cm away from each other -, they differ slightly in intensity.

We could select one of the ten wavelength regions or take the average of all these wavelength regions. In the latter case, we will have an average wavelength region of 98 pixels that will represent the whole set. The main advantage of using the average is removing uncorrelated disturbances (additional noise) in pixels. In contrary, the random selection of one wavelength region would considerably increase processing speed.

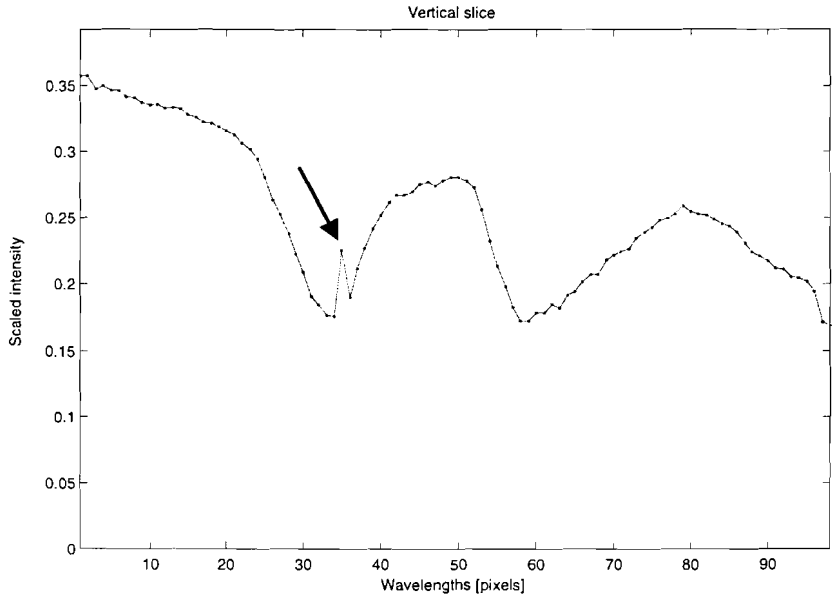


Figure 15 A PP Sample with a Spike at Pixel 35

3.3.3 Conclusion

We choose to average over the wavelength regions from pixel 48 to pixel 80, instead of selecting arbitrarily one wavelength region. In this stage of the project we want to verify the possibility to classify plastics, and we are less concerned about speed. In a latter stage of this project more time could be spend in investigating what effect selecting one of the ten wavelength regions would have had. If classification would still be possible, skipping the average calculation will increase processing speed.

3.4 Low-Pass Filtering

3.4.1 Problem

Photodiodes with a not consistent (linear) behaviour, generate spikes in the spectra. For instance, the PP sample in Figure 15 has a spike at pixel 35. Because these spikes might have a negative contribution to the classification process we want to remove them

beforehand. The information of the spectrum is not stored in these spikes or in the additional noise around it. Low-pass filtering should preserve this information while eliminating spikes and noise.

It appears that the spikes do not have a regular behaviour. Only samples with a higher intensity than 0.17 (at pixel 35) exhibit this artefact. Since we averaged over the spatial direction to obtain this spectrum, we have to conclude that this error is not related to one pixel, but to all pixels over which averaging took place.

3.4.2 Solution

The simplest solution seems omitting this pixel (or better pixels), but examining more spectra indicate that we are not only concerned with these pixels, but with approximately 20 pixels more. Above all, eliminating these pixel makes our image processing completely dependent on the camera we used. Changing the camera would demand revising all locations where spikes could occur. Also a pixel that starts malfunctioning during operation will not be noticed, and in the worst case lead to classification errors. Therefore, we eliminate the pixels by low-pass filtering.

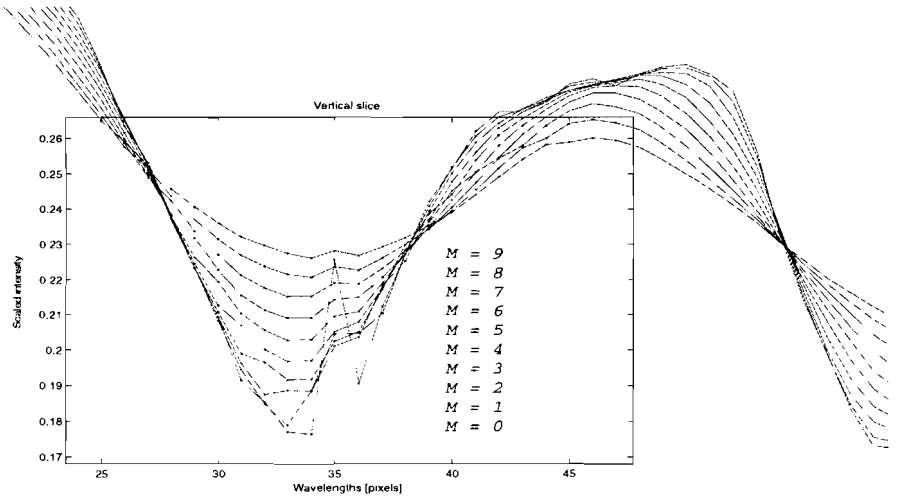


Figure 16 Moving Average Filter applied on the PP Sample

The moving average filter is a frequently used low-pass filter in image processing. The output y of the filter at a certain pixel x is the average value of the pixels around it:

$$y[n] = \frac{1}{2M+1} \sum_{m=-M}^{M} x[n-m] \qquad \text{Equation 3}$$

where M indicates the amount of values taken from the 'past' and from the 'future'. In Figure 16 the effect of moving average filtering with M = 0, 1, 2, .., 9 is shown. The moving average filter not only reduces the size of the peak in pixel 35, but the whole spectrum becomes flatter. When we look at the scale of the vertical axis we see that this effect is not tremendously yet, but despite spectral information is lost.

The conservation of information - with exception of pixel 35 - is calculated and plotted in Figure 17. The data conservation is nothing more than subtracting the RMS error introduced by filtering from the original data. The effect of reducing the spike with an M-moving average filter is expressed as:

$$r = 100 - \frac{i(p)}{(i(p-1)+i(p+1))/2} \ [\%] \qquad \text{Equation 4}$$

where the peak reduction ratio r is assumed to be optimal (i.e. 100%) in case the intensity of the peak p is equal to the average of its neighbours.

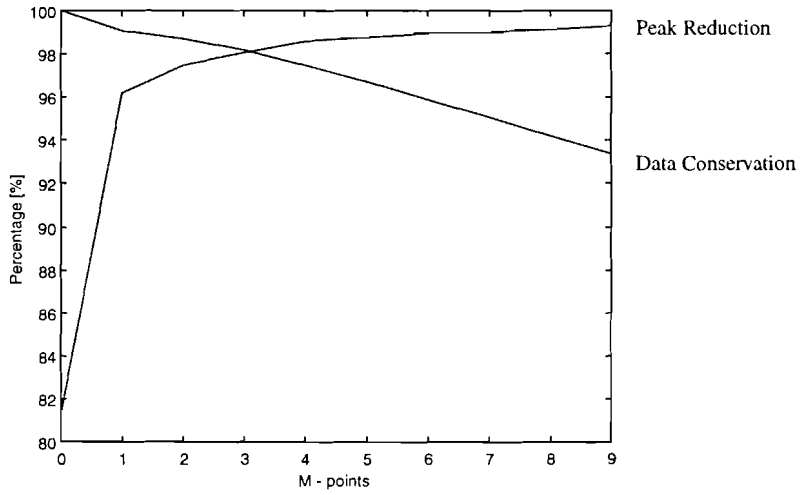


Figure 17 Effect of applying a M-Point Moving Average Filter

The optimum M-point moving average filter is a trade off between the reduction of the spike p and the conservation of the original spectrum. For this example, a moving average filter of 4 points would be most appropriate (Figure 17). Increasing the filter range does not improve the peak reduction any more, and 97% of the original spectrum is maintained. Performing this optimum search with 10 more spectra results in an average of M = 3.7. Hence, a 4-points moving average filter will be used.

3.4.3 Conclusions and Recommendations

The use of a 4-points moving average filter is a good trade off between conservation of the spectrum and the reduction of spikes.

Eliminating the pixels from the spectrum that cause the spikes is not a good solution, it is not flexible enough to be used under operation conditions. Of course, a feature extraction, and classifier part which are both insensitive to these spikes would be the simplest solution. Nevertheless, at this moment the moving average filter is a good alternative. In the future more time could be spend investigating the effects of not using a moving average filter. For now, noise and spikes have been removed to get a smooth and more reliable spectrum.

3.5 Normalisation

3.5.1 Problem

In Figure 18 the reflectance intensity of three polypropylene samples with thickness 1.70, 3.17 and 3.94mm versus the wavelength region (962 - 1692nm) are plotted. These spectra have been processed according to the chain of process steps - except of the normalisation process - in Figure 10. Although the plastics are made of the same material, their intensity response is different. The feature and classification processes would be much easier to implement, if the spectra of one plastics group would be more or less the same.

Superfluously, we mention here that the information of the spectrum that tells us what type of plastic is examined, is more related to its general shape than its intensity level [33]. By normalising the spectra we could emphasise this characteristic. In this process the intensity (or baseline) shift that is less related to the type of plastic will be eliminated and the basic shape will be enhanced. The question is, how do we eliminate the shift that is related to the variation in thickness?

3.5.2 Solution

In literature we found that the intensity of the spectra vary more or less in reverse proportion with the material thickness [42]. Unfortunately, this relationship does not hold exactly for the samples shown here. The discrepancy is clarified by dividing all recorded samples into two sets. One set contains samples that are thinner than 0.5mm and the second set consists of samples that are thicker than 0.5mm. It seems that the set of samples with a thickness smaller than 0.5mm approximates this relationship, but the samples of the other set do not.

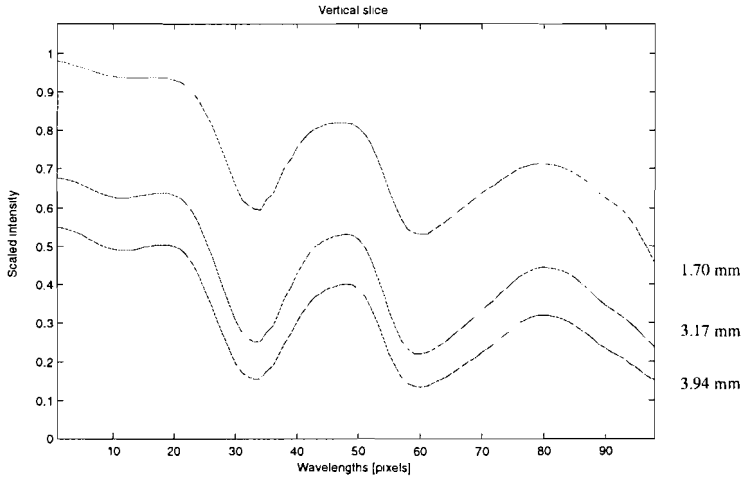


Figure 18 PP Samples with Thickness of 1.70, 3.17, and 3.94mm

Hence, we could use this reverse proportional relationship to normalise the set with thin samples, but the set with thicker samples would not be normalised correctly.

An alternative approach to eliminate the baseline shift is using a first derivative. Differentiation would eliminate the base line shift, but small fluctuations in the spectra - that could not be removed with the low-pass filter – are blown up enormously. A practical very useful technique is the calculation of the Standard Normal Variate (SNV). The SNV normalises the spectrum in such a manner that the mean is 0 and the standard deviation is 1. SNV is a combination of mean centring (correction for baseline shift), and normalising the standard deviation (the standard deviation represents the surface under a spectrum) [63].

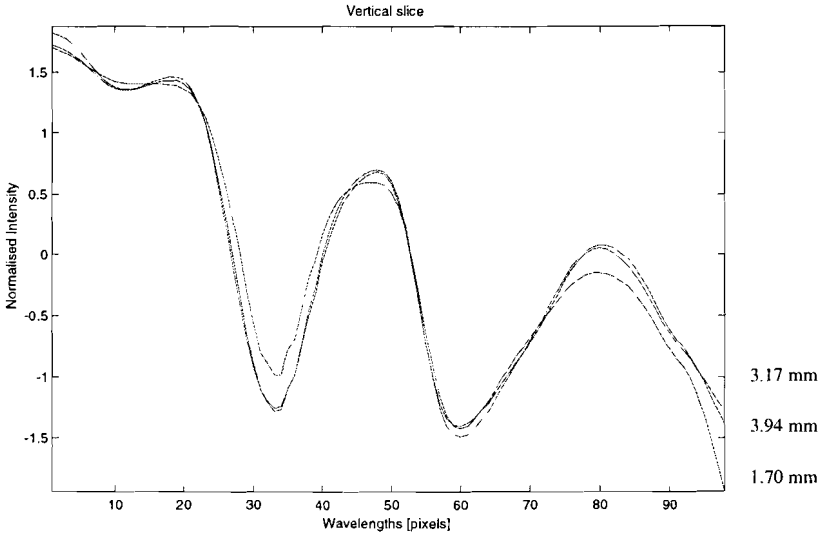


Figure 19 Result of the Normalisation Process

The result of the normalisation process for the three PP samples of Figure 18 is plotted in Figure 19. The amplitudes are not unlimited as would be the case with the first derivative, and the shapes of the three samples are more or less identical. Each of the samples could now be used to represent the set of three PP samples. The process has been repeated with 25 samples, and proven to be successful.

3.5.3 Conclusion and Recommendation

The Standard Normal Variate is a good technique to eliminate the shift in intensity level that is caused by variation in sample thickness. After all processing steps the spectra are more or less the same within a plastics group. This will greatly simplify the feature extraction and classification part.

Although not presented here, an experiment in which the light source intensity was varied between 8 and 80 Watt leads to similar results. The relationship between the intensity and the amount of light source power was linear, but the spectra became flatter when light intensity dropped. This means that in future the use of the SNV could be very appropriate to eliminate the effect of (small) changes in the intensity of the light source, too. Of course, one should keep in mind that in case the light intensity becomes too small, the spectral information of a plastic gets lost.

3.6 Conclusions and Recommendations

3.6.1 Image Processing Part

In this chapter we have seen how the raw image data is processed to get a consistent spectrum for each plastic sample within a particular plastics group. Figure 10 illustrates all these process steps. First, the raw image data is corrected for consistent camera artefacts and non-uniform lighting in the scene. Then the wavelength regions of the interest are located in the image by looking at the slope. In this process black and transparent plastics are automatically omitted, but it could be needed to help the process a little bit with a threshold. After that the region of interest is selected (by averaging or maybe by randomly selecting one), the wavelength region is filtered with a 4-points moving average filter. This filter eliminates spikes that are present due to unstable (i.e. defect) photodiodes. In future, the effect of not using this filter could be investigated to maintain more spectral information and to speed up the image processing process. Finally, the spectrum is normalised to get consistent representations of each type of plastic. The normalisation process enhances the shapes meanwhile eliminating the intensity (or baseline) shift due to variation in thickness and light source intensity.

In total, 480 spectra have been processed completely (Table 1). For the first seven classes a spectrum has been plotted in Figure 20. From each sample three spectra were measured, so actually, only one third of the spectra are really taken from different samples. Therefore, we will only use one third of the first six classes (372/3 spectra) for classification purposes. Only six classes have been selected, because performing statistics on classes with only one or two samples is not well conditioned. Nevertheless, these samples can be used to test the classifier for unknown classes. The classifier should be able to detect that a yet unseen class is present at the input. The spectra from class 9 to class 13 can be used to analyse the performance of the classifier in case a combination of two or more types of plastic is measured.

Table 1 Spectra after Image Processing

Class number	*Label*	**Amount**
1	PVC	63
2	PP	63
3	POLY.EST.	63
4	ABS	63
5	PA4_6	60
6	HDPE	60
7	L/LDPE	36
8	PA6	18
9	ABS/PA6	15
10	SMA/ABS	12
11	SMA/ABS/PMMA	9
12	SMA/PMMA	6
13	ABS/PMMA	6
14	TH.PLAST.ELAST	3
15	EPDM	3
	Total	480

3.6.2 Second Batch of Experiments

The extraction of features from the first batch of experiments seems to be successful. If good classification results are achieved, it is strongly recommended to repeat the whole process of image processing, feature extracting and classification with more realistic measurement conditions. The next paragraphs could be useful in setting up the experiments for this second batch.

3.6.2.1 Non-Plastics

Although Sorting Step II is meant for plastics only, this cannot be fully guaranteed .For instance, according to information from the German DSD-system, it is proven that public is simply not impeccable in sorting their household waste into separated waste streams [3]. Thus, accidentally also pieces of glass, metals, wood, textiles, paper, etc. will be on the conveyor belt, and it makes sense to check their effect on the whole sorting process.

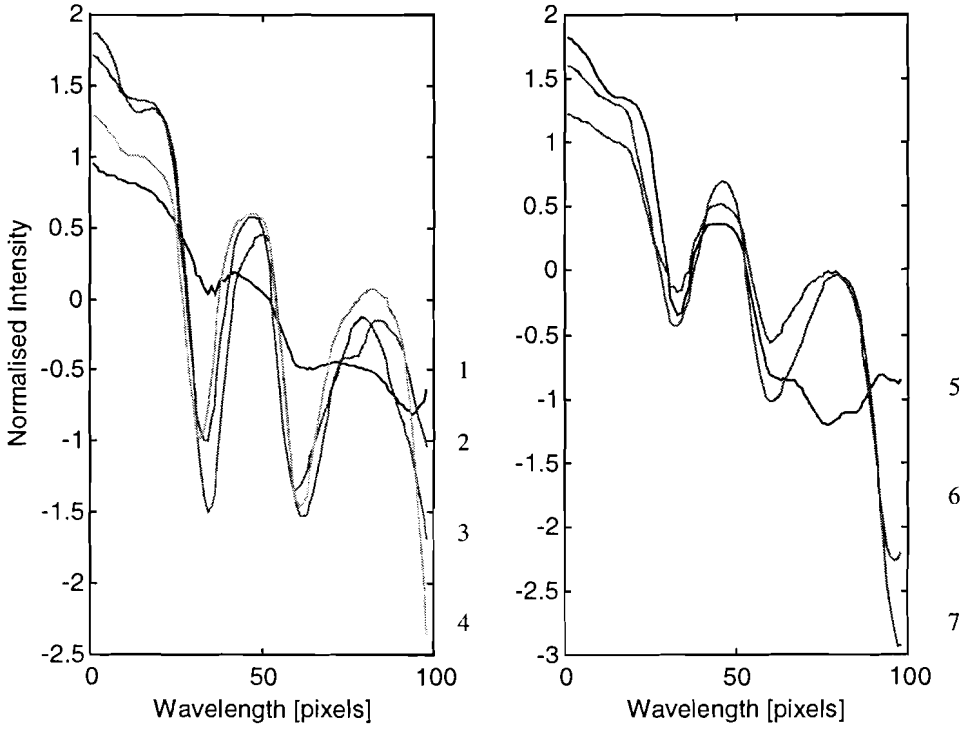


Figure 20 Seven Spectra after the Image Processing Process

Because we do not know exactly what other materials than plastics are in a common (municipal) household waste stream, we suggest measuring at least the items listed in Appendix D. (Some already have been measured). At this point, the only thing we know is that spectra from other materials than plastics are normally extremely flat [64].

3.6.2.2 Mixed Plastics

Products like 'Syntal' which are a composition of plastics belonging to distinct plastics groups can lead to misclassifications. Namely, it is unclear to which plastics group they should be categorised. We recommend investigating whether these plastics might confuse the classifier. According to Table 1 we already have measured some of these samples.

3.6.2.3 Compounds of Plastics and Non-Plastics

Some consumer products are compositions of plastics and non-plastics. Take for example a milk flake composed of a combination of cardboard and plastic foil (laminated packaging) or a soft drink bottle consisting of three parts: a bottom, flake and cap. We recommend investigating what type of spectra we can expect from these materials. Do the non-plastics or plastics have the upper hand in a single wavelength region or is the recorded spectrum a mix of both? Theoretically, we have seen that it is possible to distinguish different plastics in one single image, but what about compounds? Summarising, a lot of questions still miss an answer.

3.6.2.4 Water/ Moisture/ Dirt Conditions

Solid (municipal) household waste is normally stored outside. This means that the waste is exposed to regional weather conditions. Typically, in the north of Europe this means cold, water and moisture conditions and in the south sun and heat. Furthermore, plastics are contaminated with dirt so that in a real application no pure plastic material is measured. We recommend doing research on the quality of the spectra when the surface conditions change.

Moisture conditions are typical for waste in which green material is included. By sprinkling water - with a very fine nozzle -, moisture conditions could be imitated.

In case of wetness and heat, we can use other means to simulate these conditions. Water can simply be sprinkled on the objects under investigation and temperature can be increased with a fan. Since heating a surface is much easier than cooling it down, we suggest assuming that in case of differences, this effect in the spectra can be generalised for temperatures below room temperature.

Literature tells us that spectra in the (near) infrared region are independent of dirt conditions (food and drink residue) at the surface layer [56]. But bottles, containers, and other items with contents do not posses this advantage [41]. Which statement is true or what makes the difference between them? Namely, both could have a tremendous effect on the costs of our sorting concept. Imagine that the waste stream must be washed before classification can proceed or that the waste stream has to be pulverised into small pieces!

3.6.2.5 Non-uniform Illumination (Shadow)

Non-uniform illumination leads to a variation in the actual light intensity as formulated in Equation 1. Investigations of other research groups [62] tell us that spectra under shadow conditions (cast shadow (umbra) and self-shadow (penumbra)) lead to severe misclassifications.

4 Feature Extraction

4.1 Introduction

Now the raw image data has been processed, we are ready to start analysing the spectral information. If we want good classification results, we must have features that discriminate between plastics mutually. Similarities in spectra do not distinguish plastics, dissimilarities do. Two applications that rely on efficient data representation are pattern recognition and data compression. In pattern recognition, a classifier decides to which class an observed pattern belongs. The patterns within a class may vary considerably so that matching the observation with some characteristic template may be an inefficient classification procedure. Extracting measurements that are invariant or insensitive to the variations within each class is therefore desirable. This process is called feature extraction. We already have optimised the spectra in the image processing part, and now want to find dissimilarities and similarities in the spectra. Features, which are constant within a class but different between classes, make good classification possible. Typically, the classifier follows the feature extractor in a classification set-up. Principal Components Analysis (PCA) seems a very useful technique to emphasise dissimilarities meanwhile suppressing similarities in data.

4.2 Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a classical technique for analysing the covariance structure of multivariate statistical data. PCA derives the most important linear features of the random observation vector $\boldsymbol{x}$. PCA can be considered as a feature extraction or as a data compression technique if many variables can be represented by a few (principal) components. PCA is also closely related to least-squares techniques in estimation theory, the Karhunen-Loève (KL) transformation in time series and image processing theory, and the singular value decomposition (SVD) in numerical analysis. Also the Hotelling transform is frequently used as a synonym in literature [11].

Suppose that we have a population of M observations on n variables (x_1, x_2, x_3, .., x_n) each. Each observation is represented as a vector of the form (the superscript T denotes transposition):

$$\mathbf{x}_k = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & .. & x_n \end{pmatrix}_k^T \qquad \text{Equation 5}$$

Figure 21 illustrates an example of a population with two dimensions. A common way to predict a variable is by its mean, and in case of a multivariable population with M vectors we can calculate the mean vector as follows:

$$\mathbf{x} = E\{\mathbf{x}\} = \begin{bmatrix} E\{\mathbf{x}_1\} & .. & E\{\mathbf{x}_n\} \end{bmatrix}^T = \frac{1}{M}\sum_{k=1}^{M} \mathbf{x}_k \qquad \text{Equation 6}$$

If we calculate the mean for both variables in Figure 21, we see that for both variables the mean – in case of the original data - is a poor predictor. In contrary, if we rotate the data about 45 clockwise we see that the observations on variable x_2 do not deviate much from its mean $E\{x_2\}$. Hence, for this variable, the mean is a reasonable good predictor. In case of data reduction we store this commonly shared mean instead of all individual variables.

On the other hand, for classification purposes it would be better to select variables that are very distinct. For example, if we would like to discriminate with one variable (x_1 or x_2) all projected data, we should select variable x_1. Variable x_2 would not contribute much to the performance of the classifier, and omitting it would be advantageously for the reduction in data overhead.

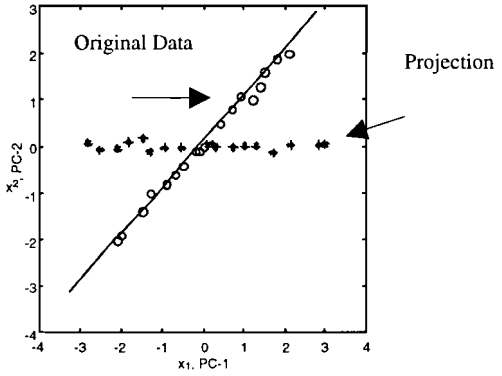


Figure 21 *Example of a PCA Analysis*

Instead of looking at the observations on one particular variable solely, we can also look at the relationships between variables. If we know the relationship between variable x_i and x_j then we can predict x_i on the knowledge of x_j. Similarly, the reverse holds: if we have knowledge of x_j we can predict x_i. Hence, for each pair of variables we calculate the covariance:

$$c_{ij} = E\left\{(x_i - \mathbf{x}_i)^T (x_j - \mathbf{x}_j)\right\} \quad i,j \in n \qquad \text{Equation 7}$$

If x_i is strongly correlated with x_j the covariance-coefficient c_{ij} is one, and x_i can be predicted exactly by x_j or x_j can be predicted exactly by x_i. Thus $c_{ij} = c_{ji}$. On the other hand, if the variables x_i and x_j are uncorrelated, their covariance is zero: $c_{ij} = c_{ji} = 0$. For the whole population we get the $n \times n$ covariance matrix $\mathbf{C}_x$:

$$\mathbf{C}_x = \begin{bmatrix} E\left\{(x_1 - \mathbf{x}_1)^T (x_1 - \mathbf{x}_1)\right\} & E\left\{(x_1 - \mathbf{x}_1)^T (x_2 - \mathbf{x}_2)\right\} & .. & E\left\{(x_1 - \mathbf{x}_1)^T (x_n - \mathbf{x}_n)\right\} \\ E\left\{(x_2 - \mathbf{x}_2)^T (x_1 - \mathbf{x}_1)\right\} & .. & .. & .. \\ .. & .. & .. & .. \\ E\left\{(x_n - \mathbf{x}_n)^T (x_1 - \mathbf{x}_1)\right\} & .. & .. & E\left\{(x_n - \mathbf{x}_n)^T (x_n - \mathbf{x}_n)\right\} \end{bmatrix} \qquad \text{Equation 8}$$

Except for a given variable x_i predicted from a single variable x_j, the variable x_j can also be predicted by a combination of some (or all other variables). Hence, reduction in data is obtained by employing these relationships. PCA performs this reduction by seeking a new orthonormal basis with a minimum number of p orthonormal vectors to describe the original set of variables as accurate as possible, so that $p << n$. The transformation error between both spaces should be kept as small as possible. If we have found this new orthonormal basis every individual observation can be transformed to a lower dimension, thereby keeping most of its essential information. The classifier would perform best, if the PCA selects the data with the largest variances between classes. How this is implemented with a PCA analysis, will be explained now.

The knowledge of the relationships between the variables is present in the covariance matrix $\mathbf{C_x}$. Creating a new space by finding a set of n orthonormal eigenvectors is always possible [11]. Let $\boldsymbol{e}_i$ and λ_i, $i = 1, 2, .., n$ be the eigenvectors and corresponding eigenvalues of $\mathbf{C_x}$ ($\mathbf{C_x}\boldsymbol{e}_i = \lambda_i\boldsymbol{e}_i$), arranged in descend order so that $\lambda_j \geq \lambda_i + 1$ for $j = 1,2, .., n - 1$. Let $\mathbf{A}$ be a matrix in which the rows correspond with the eigenvectors of $\mathbf{C_x}$, ordered so that the first row of $\mathbf{A}$ is the eigenvector corresponding with the largest eigenvalue, and the last row is the eigenvector corresponding with the lowest eigenvalue. Then the (Hotelling) transformation from $\boldsymbol{x}$ to $\boldsymbol{y}$ is denoted as:

$$\mathbf{y} = \mathbf{A}(\mathbf{x} - \bar{\mathbf{x}}) \qquad \text{Equation 9}$$

The mean of the $\mathbf{y}$ vectors resulting from this transformation are zero ($E\{\mathbf{y}\} = 0$), and the covariance matrix of the $\mathbf{y}$'s can be obtained with $\mathbf{C_y} = \mathbf{AC_xA}^T$. Furthermore, $\mathbf{C_y}$ is a diagonal matrix whose elements along the main diagonal are the eigenvalues of $\mathbf{C_x}$:

$$\mathbf{C_y} = \begin{bmatrix} \lambda_1 & .. & .. & 0 \\ .. & \lambda_2 & .. & .. \\ .. & .. & .. & .. \\ 0 & .. & .. & \lambda_n \end{bmatrix} \qquad \text{Equation 10}$$

The off-diagonal elements of the covariance matrix are zero, thus the elements of the $\mathbf{y}$ vectors are uncorrelated. Hence, the effect of the Hotelling transform is the establishment of a new coordinate system whose origin is at the centroid of the population $\mathbf{x}$, and whose axes are in the directions of the eigenvalues of $\mathbf{C_x}$. The first axes or principal components reflect the variables with the largest value, and the last principal components, the variance with the smallest variance. If the variance in the last $n - p$ axes are zero, these dimensions do not improve the performance of the classifier at all, and could be omitted. The error made with the rejection of $n - p$ dimensions, is calculated as follows. The root mean square (RMS) error between vector $\mathbf{x}$ and its reconstructed vector $\tilde{\mathbf{x}}$ is expressed as:

$$e_{rms} = \sqrt{(\mathbf{x} - \tilde{\mathbf{x}})^T (\mathbf{x} - \tilde{\mathbf{x}})} \qquad \text{Equation 11}$$

Because the rows of $\mathbf{A}$ are orthonormal vectors: $\mathbf{A}^{-1} = \mathbf{A}^T$, and the vector $\mathbf{x}$ can be reconstructed from its corresponding vector $\mathbf{y}$ by using the relation:

$$\mathbf{x} = \mathbf{A}^T\mathbf{y} + \bar{\mathbf{x}} \qquad \text{Equation 12}$$

In case that we reduced the dimensionality of the space to K, the transformation matrix $\mathbf{A}$ becomes of order K x n, where K corresponds with the K largest eigenvalues. The $\mathbf{y}$ vector will also be a K-dimensional vector, and the reconstruction to vector $\mathbf{x}$ will no longer be exact. The reconstructed vector $\tilde{\mathbf{x}}$ is:

$$\tilde{\mathbf{x}} = \mathbf{A}_k^T\mathbf{y}_k + \bar{\mathbf{x}} \qquad \text{Equation 13}$$

It can be shown that the RMS error between $\mathbf{x}$ and $\tilde{\mathbf{x}}$ is given by [11]:

$$e_{rms} = \sum_{j=1}^{n} \lambda_j - \sum_{j=1}^{K} \lambda_j = \sum_{j=K+1}^{n} \lambda_j \qquad \text{Equation 14}$$

which is zero when $K = n$. Because the λ_i ($i \in 1,2, .., n$) decrease monotonically, the error can be minimised by selecting the K eigenvectors with the largest eigenvalues.

The transformation to the new space is based on the population **x**. The PCA analysis will seek for dissimilarities and similarities for this population only. For this reason, an unknown vector **v** can be transformed wrongly. Namely, the PCA analysis omitted some dimensions that distinguish vector **v** from population **x**. Unfortunately, the opposite does not hold. A vector **v** that is transformed correctly can be correlated with population **x**. Hence, the PCA transformation does not always notice that a vector different than those in the population is applied. All the more, the PCA analysis is not optimal for new applied vectors of the latter case. In the first case, the vector **v** is detected and a new PCA analysis could be performed. This vector can be added to population **x**, and after a new PCA analysis the new decorrelated space will be optimal again. In Figure 22 the process of verifying the transformation to the new space is illustrated.

The decision boundary between a good and bad projected vector is set by *Error*. If the value of *Error* is too small, the PCA analysis is repeated for almost every new applied vector. If the *Error* is set too large, the PCA analysis is not able to detect unknown vectors, and the new space does not represent the population **x** very well.

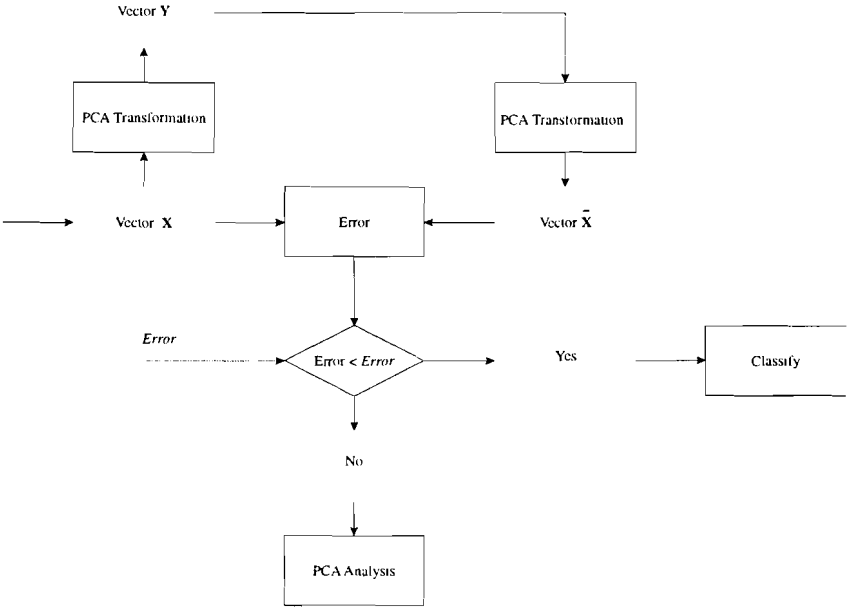


Figure 22 Feature Extraction Part

4.3 Results

The PCA analysis has been performed on the fifteen types of plastics from Table 1. In Figure 23 the features for the first two principal components are plotted. These two principal components represent about 70% of the data. We see that not all classes can be distinguished clearly with only these two components. A lot of classes overlap. Based on these two principal components no good classification results will be achieved.

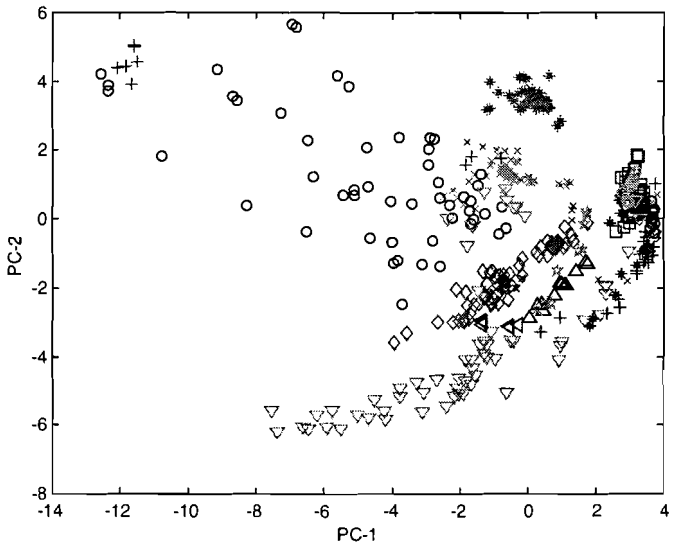


Figure 23 First Two Principal Components

In Figure 24 we see the effect of selecting more principal components. The data is very similar for all classes: selecting more than 10 principal components will not improve the ability to discriminate the data very much. These first 10 principal components represent the data for about 99.75%.

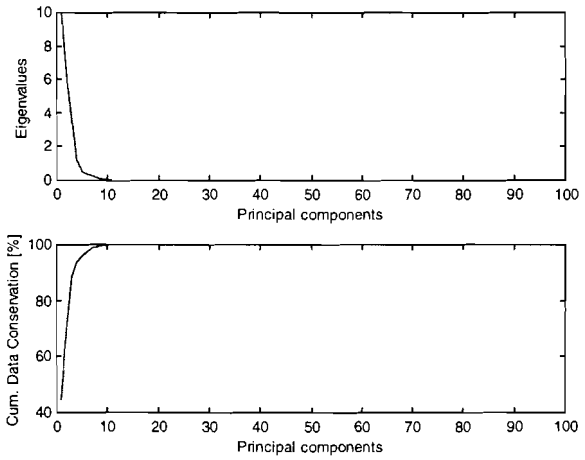


Figure 24 Results of PCA Analysis

For a value of *Error* = 1%, the verification process of Figure 22 does not reject the vectors from its own population anymore.

4.4 Conclusions and Recommendation

The amount of essential information that is present in the spectra is poor. With only 10 principal components it is possible to describe 99.75% of the original data (98 dimensions). Whether the spectra of different plastics can be separated on this information has to be verified yet.

By performing a back transformation, the PCA transformation to the new space can be verified. In case that a spectrum was not transformed correctly, the PCA analysis has to be repeated. This could be an indication of a vector of an yet unknown class, but it could also be a vector correlated with population **x**. Since the reverse does not hold, it is recommended to initialise the PCA analysis with at least some groups of plastics. Namely, if a vector is a linear combination of the vectors in the population, the vector will be transformed well. So no new PCA analysis will be performed. This is a disadvantage in case that the vector belongs to an unknown class. We would like the PCA analysis to emphasise differences between classes, but for such a vector it does not. Hence, the PCA analysis is only optimal if we include all vectors that must be classified. This might be a disadvantage if we want to classify new unknown spectra, but at this stage of the project we are only interested in the quality of the spectra. Are we able to classify them or not? Nevertheless, we suggest a solution to the problem. By using other knowledge that is able to tell whether an unknown vector is present, we are able to include this vector in the PCA analysis after classification took place.

For speed purposes the PCA analysis could be implemented in parallel hardware as found in literature [11]. We think that this would not be necessary. Namely, after a while sufficient spectra will have been added to the PCA analysis, so that less PCA analyses have to be repeated.

5 Classification

5.1 Introduction

In this paragraph we repeat the main requirements upon which we must implement a proper solution for the classification task. Moreover, we have to include the findings from the image processing and feature extraction part.

The requirements from paragraph 1.4 are:

9. The identification unit should be able to predict to which plastics group a new as yet unseen sample belongs.
10. In case that this yet unseen sample does not belong to an existing group, a new plastics group should be made. In other words, the sorting system should respond to a new type of plastic, which may be introduced into the market in the future. Secondly, within a plastics group the variety can be so large that subclasses must be made, whereas one group does not represent the variety within the group.
11. Mingled plastics – composed out of more than one type of plastic - like for example 'Syntal' should not confuse the sorting system.
12. The sorting system should be insensitive to environmental influence such as ageing of the measurement equipment. Also, as a result of continuous changes in (new) manufacturing processes, legislation, and cost aspects, identification should be flexible to slight changes in characteristics of samples within one particular class.

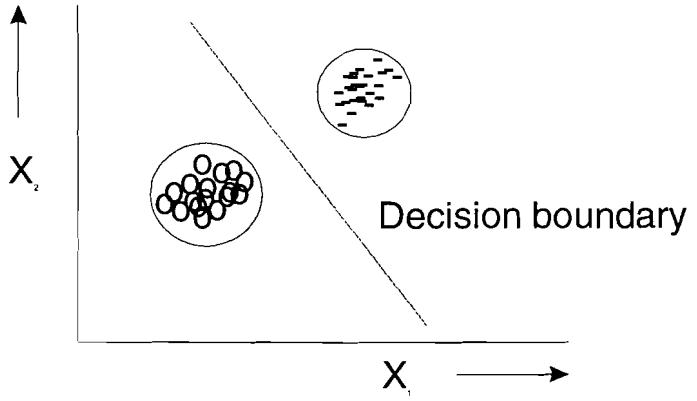


Figure 25 Example of Linearly Separable Data

Including the main economical aspects of the sorting system:

13. On-line identification and sorting should be done at very short time intervals to meet market demands. Typically, 1.5 – 2.5 tons of plastic per hour should be processed. Because this figure depends strongly on the mechanical configuration of the plastics handling, and market demands, it is difficult to translate this number exactly into the amount of identifications per second. Nevertheless, from previous investigations we can derive a maximum identification time of 15 milliseconds per sample [63].

14. It should be possible to achieve sorting rates with an accuracy of more than 99% - naturally depending on market demands.
15. The sorting system should operate fully automatic with little downtime.
16. Easy handling by unskilled people.

From the feature extraction part we derive the next requirements:

19. The classifier should be able to handle feature data with complex decision boundaries. The classification of the feature data would be easy if the decision regions would be linearly separable as illustrated in Figure 25. In this case it would be simple to define rules on which decisions could be based. However, the decision boundaries of the features are complex, and therefore we must find other alternatives to 'describe' these spaces.
20. The classifier must be able to relearn a set of features after the feature extraction part has been repeated. Namely, the PCA analysis will be repeated in case transforming a vector **x** to the new *p*-dimensional space is inaccurate. Since the vector **x** is included in this new PCA transformation, the generated features will be different now. The decision boundaries should be redefined to map the new features.
21. The architecture of the classifier must be able to handle features of arbitrary dimension. All spectra are now describable with 10 principal components, but by adding new vectors this number could increase and decrease. The largest amount of dimensions that could be reached is 98 (the maximum number of principal components, so no data reduction is obtained).
22. We would like that the classifier is able to tell the PCA analysis algorithm to repeat its procedure with the vector that was transformed correctly, but according to the classifier does not belong to a yet known class. This would alleviate the problem that the data is not decorrelated optimally in all cases.

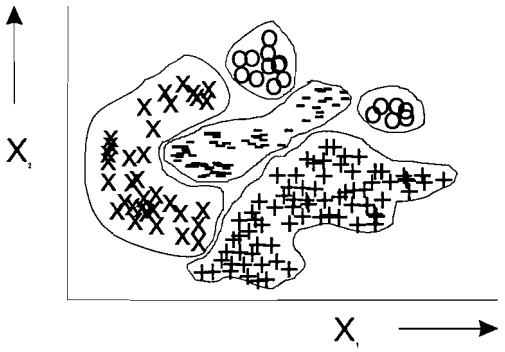


Figure 26 Example of Complex Decision Boundaries

5.2 Solution

Although we are mainly concerned with the question whether we can classify the spectra, we do base our classifier on all requirements.

The simplest view of the task of classification is that a specific pattern of observations is associated with a specific class. This is the pattern recognition perspective [40]. If the database was limited, each pattern could be stored in a table, and for a given pattern of observations one would simply look up in the table the corresponding class that had

previously been associated with it. Unfortunately, in our case the number of possible combinations of samples is very huge, thus even the most systematic and long-term record keeping is unlikely to cover all the possible combinations. Moreover, when the database with samples is relatively small, there may always be new, as yet unknown samples that will invalidate some of the defined rules of classification, and require their revision. The goal of our system is prediction on new cases, and not solely discrimination between existing samples of the database. We must build a classifier that is capable of generalising new, as yet unseen cases. Unfortunately, complications arise because classification and prediction tasks are often fraught with uncertainty. It is quite possible that two similar or even identical samples of prior cases may fall into different classes – that is, there may be ambiguity within our sample set. If many of our samples are ambiguous for a given set of features, we must conclude that these features have poor predictive power, and no good solution to the classification problem may be possible with them alone. For instance, classification that would only include the first two principal components is an example of this ambiguity (Figure 21).

Heuristic methods like expert and fuzzy logic systems based on implemented decision rules, and multivariate statistical data analyses such as LDA, and QDA are all useless. We have to rely on methods, which are able to interpret the complex data for us. Making use of Artificial Neural Networks (NN) seems very promising.

Neural Networks [59] store knowledge and use it to make decisions on new inputs. They can generalise, giving the correct output despite of minor variations in the input vector. NN acquire knowledge through training from a training set composed of actual observations of the physical world, rather than being formed by human opinions used for fuzzy-logic systems (or expert) systems. Neural networks have the great advantage that it is not always necessary to define rules on how clustering and classification should take place. Based on the type of network, relationships are created to map inputs (spectra) on outputs (classes). How this mapping is done by the network structure is often unclear. In other words, the neural network lets the data speak for itself.

Unfortunately, this can be a major disadvantage: the training set must be adequate to fully represent the domain of interest, otherwise the neural network can make decisions in areas where it has no experience. In both humans and machines this can lead to serious errors. Because our sample database is relatively small, the chance that this will happen is quite large. If we would be able to estimate the accuracy of classifications we could keep track of performance, and alleviate this problem.

Above all, a measure of accuracy is very helpful in accomplishing market demands (for example: only PVC of 99% purity). Also keeping track of the accuracy can help in determining whether we should add a new class or not. For instance, if the classifier indicates that a sample can not be categorised because no associated class is present, we could implement an algorithm that automatically creates a new class for this sample.

The sorting system should operate fully automatic which means that no human intervention is allowed. Unfortunately, we can already conclude in advance that this

requirement cannot be met when we have to create new classes on-line. Namely, our sample database is too limited to include all representations of plastics, and over a period of time new yet unknown classes could be lying on the conveyor belt. Thus, regardless of our ability to make a smart algorithm to add new plastics groups, we still need to label these groups. Even if the classifier could automatically label the new classes with labels like 'unknown_1', 'unknown_2', or 'probably_pp', 'probably_pvc', etc. we still have to verify these labels manually. The classifier simply has not enough knowledge to tell what plastic it really is. It can only tell it does not belong to any of the plastics it already knows.

Summarising, the use of a neural network that is able to keep track of its performance is advocated. In the next paragraphs we will select the NN that meets all requisites best. We concluded that human intervention is always needed in areas where the classifier has no knowledge.

5.3 Artificial Neural Networks

5.3.1 Basic Architecture

Basically, neural networks (NN) do nothing more than mapping input vectors into output vectors. Since the way mapping takes place is often not known, neural networks are often described as 'black boxes'. The basic construction of a NN consists of three building blocks:

1. *Architecture.* Internal connections of signal paths.
2. *Training or learning algorithm.* Determining the weights of the connections.
3. *Activation function.* The behaviour of the neurons to an input activation.

5.3.2 Learning Methods

There are many learning methods for neural networks by now. Nobody knows exactly how many. New ones (or at least variations of existing ones) are invented every week. Globally, learning methods for NN can be classified into three basic types: *supervised*, *reinforcement*, and *unsupervised* learning.

In *supervised* learning a teacher is assumed to be present during the learning process. After that an input vector has been applied, the teacher compares the produced output vector of the neural network with the correct output vector. The teacher calculates the error between these vectors and the error is then used to change the weights of the neural network and thereby improving its performance. Hence, supervised learning demands for a sample database with their corresponding categories.

In *reinforcement* training, a teacher is also assumed to be present, but instead of calculating the error rate and back-propagating it to adjust the weights, the network is only told whether the output vector it computes is right or wrong. The neural network must then use this information to improve its performance. Typically, the weights of neurons, which give the right answer, are reinforced and the weight values on those neurons giving the wrong answer are reduced.

In *unsupervised* learning, the network has no feedback on the desired or correct output. There is no teacher to present target vectors. Therefore, the system must learn by discovering and adapting to structured features in the input vectors, that is, by adapting to statistical regularities or clustering vectors from the input training samples. For instance, the following structures might be discovered in the sample data:

1. Grouping or clusters of closely related patterns.
2. Frequencies of occurrence of groups or patterns.
3. Relative orderings (e.g. length) among the input vectors.
4. Correlation of the patterns (for example, the greatest variance among the variables like in PCA)
5. Mapping which transforms input vectors to a lower dimensional space.

5.3.3 Performance Indication

Neural networks suffer from a lack of perceived reliability. They can provide useful answers in the vast majority of cases, and then without any warning produce a totally incorrect answer. This can usually be traced back to an inadequate training set. Still, there is often no way to determine when the output will be incorrect. Even worse, it is usually impossible to determine how an artificial neural network made an incorrect decision, so that the problem can be corrected. This situation arises from the fundamental structure of the network: knowledge distribution over the entire pattern of weights, thus a large amount of knowledge may be involved with one decision. The only way to fully characterise the response of the neural network is to apply all possible inputs, and see what happens. Unfortunately, this is in many cases impossible due to the complexity of the problem. With neural networks one must be satisfied with a probabilistic statement of reliability. This is achieved by splitting the data into a training set and a test set and evaluating its performance.

Fortunately, some neural networks - Probabilistic Neural Network (PNN), Radial Basis Function Network (RBF), and General Regression Neural Network (GRNN) - provide an indication of the correctness by means of probabilities. The selection of the PNN network is easy: GRNN networks are meant for function approximation rather than on classification tasks. RBF networks require a radial symmetric distributed feature space. The latter does not hold in our application.

We will not discuss all types of network topologies here, since many of them are problem related. Whereas the literature on the Back-Propagation neural network (BP) is enormously we will discuss this topology as a reference. We only want to check whether the PNN meets all requirements.

5.3.4 Back-Propagation Neural Network

Although the Back-Propagation topology is investigated most among the NN, this should not imply that it is the best solution to all practical problems [59]. Its primary use is to perform supervised training on multi-layer networks. With suitable training it may be used for either continuous mapping or classification. Back-propagation suffers from a variety of ills. Maybe the worst of these is its very long training time. For this reason various methods have been developed to increase learning speed. Despite that these

solutions can reduce training time enormously, their performance is highly problem dependent. Another problem results from the training method – gradient descent. In a nutshell gradient descent means that in the training phase the error of classifications descents to an optimum value by adjusting the network weights. After a long training time the network can be trapped in a local minimum, thus, no (optimal) solution may be obtained. Finally, back-propagation is subject to paralysis. Paralysis happens when a neuron is saturated and its output cannot be changed. If this occurs, training virtually stops and nothing else can be done then to repeat the training process with different initial conditions. Many of the problems of BP can be avoided by using special training procedures. However, most of them focus only on one particular problem.

5.3.5 Probabilistic Neural Network

5.3.6 Advantages

Compared with BP the Probabilistic Neural Network (PNN) offers the following major advantages:

1. Rapid training. The PNN is as much as thousand times faster than BP. The days or weeks of iterative training of a BP are replaced by simply reading in the training set once. This is very useful when the PCA analysis has been repeated and the new feature data has to be mapped into the PNN network.
2. With enough training data a PNN is guaranteed to converge to a Bayesian classifier (the usual definition of optimality) [59]. This irrespective of the complexity of the data. This cannot be guaranteed with BP. Long training times can terminate in a local optimum that may be an unsatisfactory solution or even in the worst case no solution can be obtained.
3. The PNN algorithm allows data to be added or deleted without lengthy retraining, whereas any modification to a BP training set will require a repetition of the entire training process. This characteristic of PNN makes it more compatible with many real-world applications. As with human experience, network learning is often a continuous process. Moreover, adding new input data during operation improves the performance of the classifier.
4. The topology of a PNN is fixed and is thus independent of the problem that has to be mapped into it. The number of layers and neurons in a BP depend completely on the complexity of the data. Commonly, the topology of a BP is determined experimentally. So an arbitrarily dimensional space can be mapped immediately.
5. The PNN provides an output indicating the amount of evidence, which is used to make its decision. A BP does not give such a confidence indication. A BP will always give an answer without telling how certain this answer really is.
6. Finally, a BP can give multiple answers. In case happens you cannot decide how a sample should be classified without a confidence indication. A PNN estimates for all classes the probability, and based on this information we can decide to categorise the sample or to reject it. For instance, if a sample belongs 50% to class A, and belongs 50% to class B, it would be better to reject the sample. Rejection could also mean that the samples are rejected from the conveyor belt for closer investigation (for instance by hand).

5.3.7 Common Characteristics

Summarising, the PNN overcomes many of the disadvantages of BP, and retains with it the most important characteristics:

1. *Learning*. The network is capable of learning any arbitrarily complex relationship between training data and their corresponding classes. Arbitrarily complex non-linear (and linear) decision boundaries in a space of any dimension can be extracted without human intervention.
2. *Generalisation*. The network is able to classify correctly input vectors that are similar, but not identical to those that were in the training set. Erroneous, noisy, or incomplete training sets do not have large effects on classification accuracy.
3. *Concurrency*. As with many NN, the structure of the algorithm allows efficient use of multiprocessor systems. The classification time is almost exactly inversely proportional to the number of processors. Speed improvements are obtained by adding parallel hardware. Modular VLSI implementations provide systems with high performance and an acceptable price [59].

5.3.8 Disadvantage

The main disadvantage of PNN is that the computational load is transferred from the learning phase to the classification step. The basic PNN requires one node or neuron for each training pattern. This means that training is extremely fast, but classifying a large number of new patterns can be slow because the amount of computations is proportional to the number of training vectors mapped in the network. The use of parallel hardware alleviates this problem, and increases classification speed enormously. Secondly, by clustering the training data less training patterns will be needed. A single or some more cluster centres will represent each set of training patterns.

5.4 Conclusion

PNN uses supervised learning and trains virtually instantaneously. Training of a PNN is much faster than BP. It is capable of learning arbitrarily complicated classification tasks with an accuracy that approaches that of a Bayesian classifier (the standard of optimality). The PNN algorithm allows data to be added or to be deleted from the training set without lengthy retraining. Finally, PNN provides information on the evidence upon which it makes a decision. It indicates how accurate it performs a classification task. The PNN classifier is able to tell that no knowledge is present to perform a correct classification. This indication could be used to add additional knowledge to the classifier at any time.

Apparently, a practical problem of a PNN is its use of the entire training set for each classification. This demands large storage capacity and lengthens classification times. Fortunately, employing clustering techniques reduce the size of the training set, and moreover, special purposed parallel hardware can be used to speed up the classification process [59]. Summarising, the main disadvantage of PNN is solved with additional hardware and smart programming, and therefore the PNN classifier turns out to meet our requirements best.

6 The Probabilistic Neural Network (PNN)

6.1 Introduction

The Probabilistic Neural Network (PNN) is actually nothing more than a Bayesian classifier. It has been successfully used to solve a diverse group of classification problems [7].

In a nutshell, the operation of Bayesian classifiers can be described as follows [35]. Let **y** be an n-dimensional input vector, characterising the features of a plastic sample, which belongs to one of the k possible plastics groups. Let $f_1(\mathbf{y})$, $f_2(\mathbf{y})$, $f_3(\mathbf{y})$, .., $f_k(\mathbf{y})$ be conditional probability density functions for the k-class population, and let $p_1, p_2, .., p_k$ be the a priori probabilities that a feature vector **y** belongs to its corresponding class k. For classification purposes we want to derive a decision function $d(\mathbf{y}) = C_i$, $i = 1, 2, .., k$, that classifies **y** as belonging to class C_i with minimal risk of incorrect classification. Let L_1, L_2, .., L_k be loss (or cost) functions associated with making a wrong decision such that loss L_i is incurred whenever $d(\mathbf{y}) = C_i$, $i \neq j$ and $\mathbf{y} \in C_j$. The cost of loss of incorrect decision is taken to be zero. The Bayes' decision rule for this classification problem compares the product probabilities:

$$p_1 L_1 f_1(\mathbf{y}), p_2 L_2 f_2(\mathbf{y}), ..., p_k L_k f_k(\mathbf{y}) \qquad \text{Equation 15}$$

and selects the class corresponding to the largest product value or probability. So if the next relationship holds:

$$p_i L_i f_i(\mathbf{y}) > p_j L_j f_j(\mathbf{y}) \quad ; \text{ for } j = 1, 2, .., k \; ; \; i \neq j \qquad \text{Equation 16}$$

the decision rule assigns **y** to class C_i. In the calculation of the product values more selection criteria can be introduced in the form of cost functions L to penalise the choice of misclassifications.

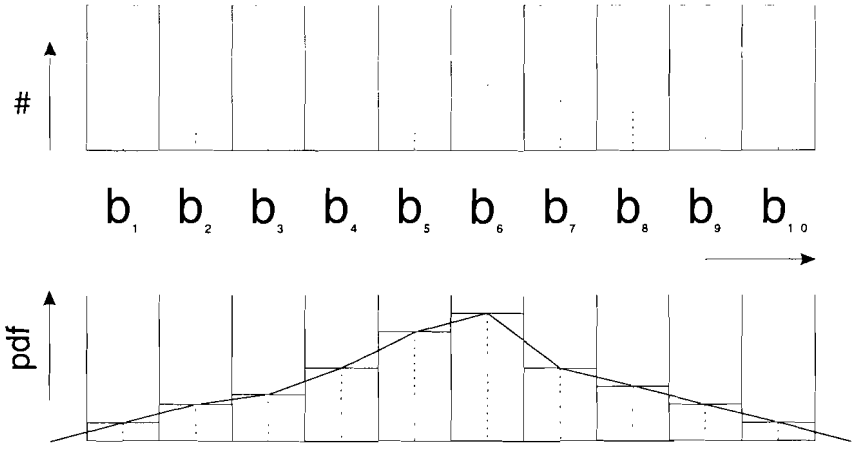


Figure 27 Estimate of a fictive PVC

One of the main criticism of Bayes' classification rule is the lack of information about the probability distributions $f_i(\mathbf{y})$ and p_i. The a priori probabilities p_i may be known or can easily be estimated from a (representative) data set. In case of a representative data set we

can simply derive the distribution by calculating the ratios between the number of samples of one particular class and the total number of samples in the data set:

$$p_i = \frac{\#\text{samples in class } C_i}{\#\text{ total number of samples}} = \frac{k_i}{k} \qquad \text{Equation 17}$$

Generally, the probability density functions (pdf's) $f_i(\mathbf{y})$ are more difficult to estimate. Of course one could assume some form of distribution function (i.e. normal distribution) and estimate the unknown parameters using standard statistics. But in our case the complexity of the feature space must be estimated with other techniques. First, we will estimate a (fictive) pdf for PVC to get a better understanding of its meaning.

Assuming that all PVC samples can be represented by a one-dimensional feature vector consisting out of ten binary bits $\mathbf{y} = (b_1, b_2, b_3, .., b_{10})$. For simplicity, let assume that the bits are mutual exclusive. That is to say, in a feature vector only one bit can be a logical one and all others are logical zeros. Over a large period of time (or from all feature vectors from a sample database) we count the number of times that a bit of the feature vector is a logical one. Figure 27 illustrates this procedure. For b_1 we counted three times a logical one and for b_2 five times, etc. The obtained histogram is then scaled by dividing each bin with the total number of samples that has been examined, so that the area under the pdf curve is unity. The result after smoothing and scaling this 'histogram' is the fictive (discrete) pdf of PVC. The more samples we add to the histogram, the more accurate the pdf will be.

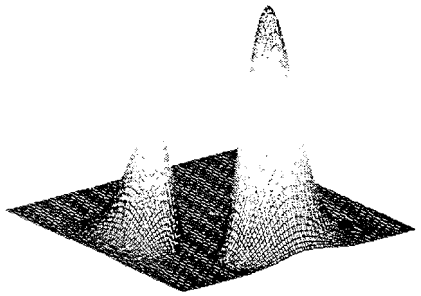


Figure 28 Two-dimensional Pdf

The probability density function (pdf) can be used as follows. For example, if we want to know with what probability a feature vector will have a logical one for the fourth, fifth, and sixth component, we calculate this probability by integrating the area under the pdf curve between bits b_4 and b_6.

In practice, we will have a feature vector with more than one dimension. So we must repeat the previous method of histograms for each dimension. Because we cannot visualise feature vectors of more than three dimensions we visualised a two-dimensional

pdf to get some feeling for multi-dimensional pdf functions (Figure 28). In practice, the method of histograms is very laboriously, and therefore estimation functions are used. One of such functions is the Parzen window estimate.

6.1.1 Parzen Windows

The Bayesian classifier requires a pdf $f_i(\mathbf{y})$ for each class C_i. In practice, it is often difficult to determine the pdf with high accuracy. Namely, the sample data set may be relatively small, incomplete or partially inaccurate. Fortunately, the method of Parzen windows allows the derivation of good probability density functions from sparse sample data [7].

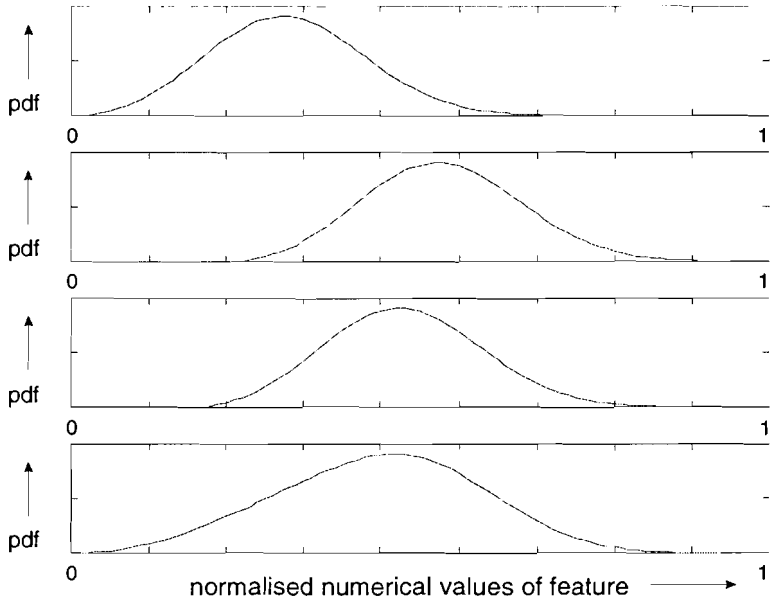


Figure 29 Illustration of Parzen Windows

Figure 29 illustrates the principle of Parzen windows. The horizontal axis represents the numerical values of a one-dimensional feature vector. For each set of feature vectors - within one specific class - a unit area Gaussian curve is drawn centred at the numerical value of each feature component. In Figure 29 this has been done for the first three samples of class C_i. All of the curves are then added to produce the composite curve in the lowest box. Parzen showed that with a large number of samples and suitable scaling, the composite curve approaches the true pdf very accurately [7]. Although, there is no general method that indicates the number of samples required to estimate the pdf with a certain degree of accuracy, good classification results are often obtained with a modest number of samples [59].

The estimator often used for Parzen windows is:

$$f_i(\mathbf{y}) = S\frac{1}{k_i}\sum_{p=1}^{k_i}\exp\left(-\frac{(\mathbf{y}-\mathbf{y}_{ip})^T(\mathbf{y}-\mathbf{y}_{ip})}{2\sigma^2}\right) \qquad \text{Equation 18}$$

which constructs the pdf for a n-dimensional feature space of training vectors. The vector $\mathbf{y}_{ip}$ is the p^{th} training vector belonging to class C_i and the (input) vector $\mathbf{y}$ has to be classified and is unknown. The coefficient σ is a smoothing parameter, which must be determined experimentally. The scaling factor S scales the sum of multivariate Gaussian distributions to unity:

$$S = \frac{1}{(2\pi)^{n/2}\sigma^n} \qquad \text{Equation 19}$$

In Figure 30 an unscaled one-dimensional estimated is plotted. The conditional probability $f_i(\mathbf{y})$ will be larger as the distance between the unknown vector $\mathbf{y}$ and the training patterns $\mathbf{y}_p$ becomes smaller. Likewise, a high pdf output $f_i(\mathbf{y})$ indicates that the vector $\mathbf{y}$ is highly associated with class C_i.

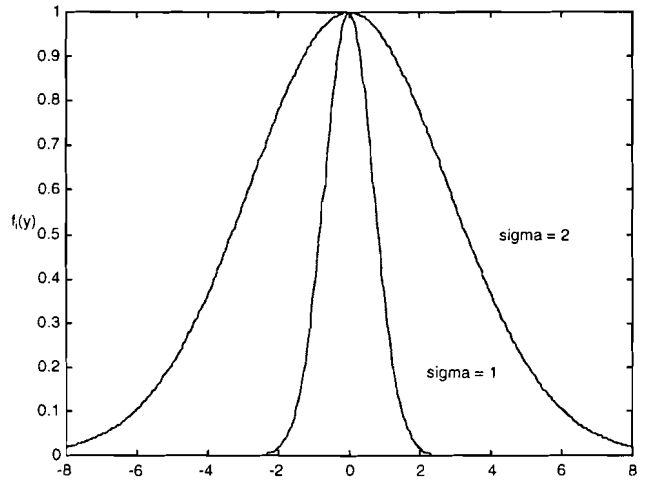


Figure 30 *One-dimensional Parzen Window Estimate*

With the smoothing parameter σ (the standard deviation of the Gaussian), the width of the curve is adjusted. If σ is small ($<<1$), each Gaussian is sharply peaked, and a large σ will smooth the pdf. Generally speaking, a σ that approaches zero, approaches a nearest neighbourhood classifier, and a σ that goes to infinity approaches decision boundaries in the form of hyperplanes [7]. In the latter the classifier is able to generalise outside the decision boundaries which are defined by the training set. In case high accuracy is needed, σ should be small. Nevertheless, real world experiments show that the choice of the value σ is not very strict, but the data set should be representative.

6.1.2 The PNN Architecture

The PNN can be implemented easily in a neural network topology [59]. We will show how to derive the equations for the network structure. The decision rule for a multi-category problem is defined by Equation 16. For simplicity let assume that the cost functions to penalise a wrong decision are equivalent for all classes. Substituting Equation 18 in Equation 16 gives the relationship:

$$\sum_{p=1}^{k_i} \exp\left(-\frac{(\mathbf{y}-\mathbf{y}_{ip})^T(\mathbf{y}-\mathbf{y}_{ip})}{2\sigma^2}\right) > \sum_{p=1}^{k_j} \exp\left(-\frac{(\mathbf{y}-\mathbf{y}_{jp})^T(\mathbf{y}-\mathbf{y}_{jp})}{2\sigma^2}\right) \qquad \text{Equation 20}$$

for $j = 1, 2, ..k$, and $i \neq j$. For a two category problem the decision rule $d(\mathbf{y}) = C_l$ between the classes C_1 and C_2 becomes:

$$\text{if} \sum_{p=1}^{k_1} \exp\left(-\frac{(\mathbf{y}-\mathbf{y}_{1p})^T(\mathbf{y}-\mathbf{y}_{1p})}{2\sigma^2}\right) > \sum_{p=1}^{k_2} \exp\left(-\frac{(\mathbf{y}-\mathbf{y}_{2p})^T(\mathbf{y}-\mathbf{y}_{2p})}{2\sigma^2}\right) \text{ then } \mathbf{y} \in C_l \qquad \text{Equation 21}$$

The architecture of Figure 31 implements this classification task. The unknown vector $\mathbf{y} = (y_1\ y_2\ ..\ y_n)$ is applied at the distribution layer, which merely distributes the input values to the second layer and does not compute anything.

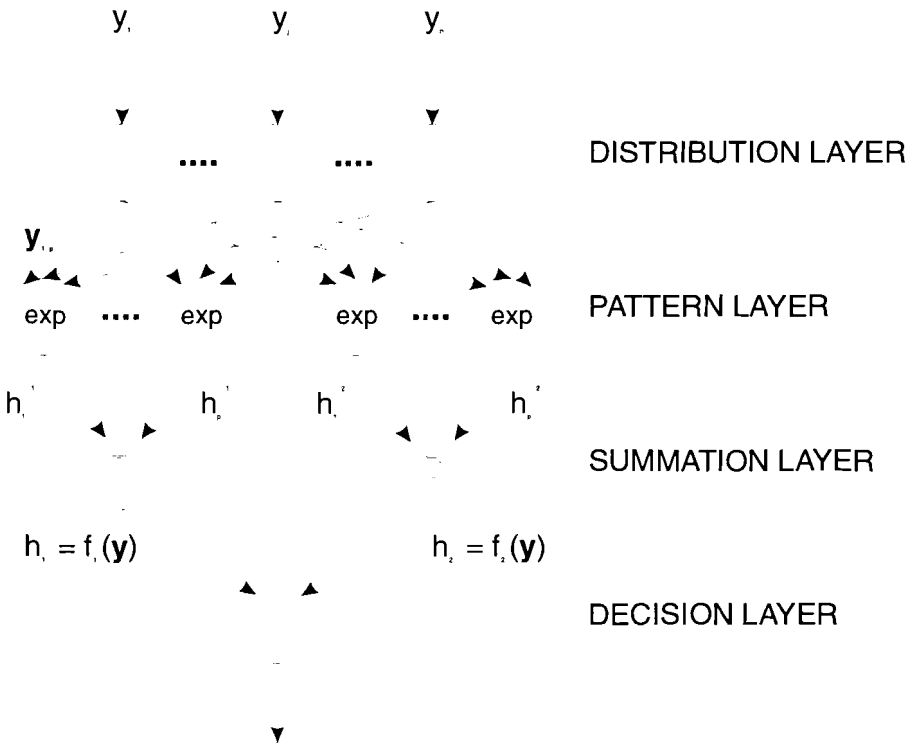


Figure 31 Two Category Probabilistic Neural Network Architecture

The weights connected in the second (pattern) layer represent the values of each component of the training vector $\mathbf{y}_{ip} = (y_{p1}\ y_{p2}\ ..\ y_{pn})$. These vectors have been mapped in the network during the training phase. Actually, the training phase is nothing more than reading the training set: storing the vectors $\mathbf{y}_{ip}$ into the weights of the pattern layer, and making connections to the summation layers. The neurons in the pattern layer all calculate the (Euclidean) distance between the training vectors and the unknown input vector. In the next step the exponential function is applied, and their results are summed in the summation layer. Now, the conditional probability functions $f_1(\mathbf{y})$ and $f_2(\mathbf{y})$ have been calculated, in the decision layer the largest probability is taken indicating the class

with the highest probability of occurrence. The decision layer is often called a winner-take-all-circuit or competition layer.

Because we want to have an indication of the confidence by which a decision is made, we must omit the decision layer and implement a slightly different equation in the network. We want the network to generate the probabilities that vector **y** belongs to class C_i:

$$p(i \mid \mathbf{y}) = \frac{p(\mathbf{y} \mid i)p_i}{p(\mathbf{y})} = \frac{f_i(\mathbf{y})p_i}{p(\mathbf{y})} \qquad \text{Equation 22}$$

where p_i is the probability of class C_i occurring, and $p(\mathbf{y})$ the probability of an input vector with value **y**. Substituting $f_i(\mathbf{y})$, which is approximated by the Parzen estimator gives:

$$p(i \mid \mathbf{y}) = \frac{p_i S}{p(\mathbf{y})k_i} \sum_{p=1}^{k_i} \exp\left(-\frac{(\mathbf{y} - \mathbf{y}_{ip})^T (\mathbf{y} - \mathbf{y}_{ip})}{2\sigma^2} \right) = \frac{p_i S}{p(\mathbf{y})k_i} \sum_{p=1}^{k_i} h_p^i \qquad \text{Equation 23}$$

where $h_p{}^i$ is the output of the pattern layer neuron belonging to training vector $\mathbf{y}_{ip}$. The probability p_i is estimated by its relative frequency in the training set as:

$$p_i = \frac{k_i}{k} \qquad \text{Equation 24}$$

Applying Bayes' rule to calculate the conditional probability **y** for each class, and then summing over all classes results in:

$$p(\mathbf{y}) = \sum_{p=1}^{k_i} p(\mathbf{y} \mid i)p_i = \frac{S}{k} \sum_{i=1}^{k_c} \sum_{p=1}^{k_i} h_p^i \qquad \text{Equation 25}$$

where k_c represents the total number of classes. By counting all hidden neuron outputs $h_p{}^i$, the double summation sign is eliminated. Hence:

$$p(\mathbf{y}) = \frac{S}{k} \sum_{i=1}^{k} h_i \qquad \text{Equation 26}$$

, where h_i is the sum layer output of class C_i. Finally, the probability that vector **y** is in class C_i is obtained by substituting Equation 24, and Equation 26 in Equation 23:

$$p(i \mid \mathbf{y}) = \frac{\sum_{p=1}^{k_i} h_p^i}{\sum_{i=1}^{k} h_i} \qquad \text{Equation 27}$$

6.1.3 Classification Accuracy

In case that the number of training vectors is limited, we must avoid that the classifier starts making decisions in regions where it has no knowledge. Such a region can be detected by looking at the hidden neuron outputs (h_p^i). Assume the two-category problem once more. For example, the classifier indicates that an unknown vector **y** belongs 40% to class C_1, and belongs 60% to class C_2 (Equation 21). It would not make sense to decide that **y** belongs to class C_2 in case that the hidden neuron outputs h_1 and h_2 would be $4 \cdot 10^{-6}$, and $6 \cdot 10^{-6}$, respectively. The decision that **y** belongs more to class C_2 than to class C_1 is correct, but the information on which the decision is made, is insufficient. Clearly, the decision boundaries are not well defined in the region where we want to classify **y**. As a consequence, it would be better to indicate that the decision could not be made.

In this case the vector **y** is rejected and human intervention is needed to categorise **y** into its proper class. The previously unknown vector could be used to extend the decision boundaries of the classifier. In case the previously unknown vector does not belong to an already existing class the pattern and a new class will have to be added to the network. Provided that the class of the new training vector already exists, only the pattern and the connections to that class have to be made.

The classifier can decide for a second reason not to classify vector **y**. Let us assume that the vector **y** lies within the decision boundaries of the classifier, and the neuron outputs for the two-category problem are $h_1 = 0.90$, and $h_2 = 0.95$ respectively. The classifier decides to classify vector **y** as belonging to class C_2 ($p(C_2|\mathbf{y}) = 51.4\%$) although the vector has an almost equal probability ($p(C_1|\mathbf{y}) = 48.6\%$) for class C_1. This means that the vector **y** is situated between the decision boundaries of class C_1 and C_2. Rejecting the vector **y** would be better and could make the PCA analysis optimising the dissimilarities and similarities between this vector and the training vectors.

Summarising, the classifier can decide to reject an unknown vector for two reasons:
1. The decision boundaries are not well defined.
2. The decision boundaries are defined, and the vector belongs to more than one class.

The first situation is detected by looking at the highest neuron output of the summation layer:

$$\text{if } \max(h_i) > T_d \text{ for } \forall i \in k \text{ then } \mathbf{y} \text{ is 'known'} \qquad \text{Equation 28}$$

The label 'known' means that the classifier has knowledge in this region. T_d is the threshold to reject a vector as 'unknown'. The relationship between the 'known' vectors (r_k) and the 'unknown' vectors (r_u) is defined as:

$$r_u[\%] + r_k[\%] = 100[\%] \qquad \text{Equation 29}$$

In case the highest neuron output does not meet Equation 28, no training vector is related to vector **y**.

In the second situation two neuron outputs will indicate that the vector **y** should be categorised into their corresponding class. It might be useful updating the PCA analysis with vector **y**. Namely, this might enhance the dissimilarities between vector **y**, and the training vectors. Providing that the decision boundaries are representative. This condition is formulated as:

$$\text{if } T_d < p(i|\mathbf{y}) - p(j|\mathbf{y}) \text{ then } \mathbf{y} \in C_i \quad (h_i \geq h_j) \qquad \text{Equation 30}$$

The number of vectors that is rejected by this criteria, is indicated with the rejection rate r_j, the number of correct classifications with the recognition rate r_c, and the number of misclassification with the error rate r_e. The relationship between these rates is:

$$r_u[\%] + r_c[\%] + r_e[\%] + r_j[\%] = 100\ [\%] \qquad \text{Equation 31}$$

Since we must obtain very accurate sorting rates, we must optimise the accuracy (r_a) of the classifier:

$$r_a = \frac{r_c}{r_c + r_e} = 100 - \frac{r_e}{100 - r_u - r_j} = 100 - \frac{r_e}{r_k - r_j}[\%] \qquad \text{Equation 32}$$

which signifies that the error rate r_e should be minimised.

In the PNN three parameters can be adjusted to meet this requirement:

1. σ - The smoothing parameter, which does not influence the performance of the classifier in case the training set is representative [7].
2. T_d - The threshold value to detect vectors lying outside the decision boundaries.
3. T_p – The threshold value to reject a vector as it is associated with more than one class.

The optimal smoothing parameter σ and the optimal threshold value T_d can be estimated by using the maximum likelihood that all patterns $\mathbf{y}_{pi}$ belong to class C_i [7]. The conditional probability $f_i(\mathbf{y})$ should be maximum when a training vector $\mathbf{y}_{pj}$ is applied. By using the leave-one-out method we avoid $f_i(\mathbf{y})$ having an artificial maximum of $\sigma = 0$. Namely, $f_i(\mathbf{y}) = 1$ holds for $\mathbf{y}_{pi} = \mathbf{y}_{pi}$ as can be seen in Equation 18. So the best σ can be found as the value that maximises the log likelihood (log LH):

$$\log LH = \sum_{p=1}^{k_i} \log \sum_{\substack{j=1 \\ p \neq j}}^{k_i} \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(-\frac{\left\| \mathbf{y}_{pi} - \mathbf{y}_{pj} \right\|^2}{2\sigma^2} \right) \qquad \text{Equation 33}$$

By substituting σ and calculating:

$$T_d = 1 - \sum_{\substack{j=1 \\ p \neq j}}^{k_i} \frac{1}{(2\pi)^{n/2}\sigma^n} \exp\left(- \frac{\left\| \mathbf{y}_{pi} - \mathbf{y}_{pj} \right\|^2}{2\sigma^2} \right)$$ Equation 34

the threshold value T_d is found. This method is useful in case that a vector **y** must be classified into one class of interest versus everything else. Namely, it may be impractical to get a sufficient number of training samples for 'everything else'. Unfortunately, for this multi-category problem the training patterns will be frequently updated. Especially, in case of a not representative training set this will lead to a lot of (unwanted) additional computational load in the start up phase.

For this reason (and lack of time) the optimum between these parameters will be determined on the basis of simulations, and the number of false-positive classifications (rejections) one is able to tolerate. The simulations will clarify the effect of selecting a certain parameter set (σ, T_d, T_p).

6.1.4 Tracking and Pruning

After the PNN network classifies the vector **y** into class C_i the most associated pattern vector $\mathbf{y}_{pi}$ is adjusted (tracked) into the direction of **y**. In this manner the classifier is less sensitive to small changes in the equipment (and perhaps the samples) over a long period of time. A simple updating algorithm for tracking the classifier could be:

$$\mathbf{y}_{pi,new} = a \cdot \mathbf{y}_{pi,old} + (1-a) \cdot \mathbf{y}$$ Equation 35

where a avoids rapid changes. By reconstructing all patterns $\mathbf{y}_{pi}$ back to $\mathbf{x}_{pi}$ before a (new) PCA analysis starts, the effect of tracking will also be taken into account in the PCA analysis.

Till this point we only increased the network size by adding new training vectors and new classes. Of course, the number of classes and training vectors can not grow continuously. After a while hardware and/or software will reach their limit and one should prune the network. By keeping for each pattern a record with the amount of classified vectors, the training patterns lately not involve in decision making can be removed. The same can be done for the classes. This record is also useful in estimating the conditional probability p_i of Equation 17.

6.1.5 Training and Validation Phase

The PNN network is trained by mapping the training set into the weights of the pattern layers and connecting these layers with the summation layers. Because the number of training vectors is limited we apply the leave-one-out validation method. Sequentially we leave one vector (test vector) out of the data set, and train the network with the rest of the data set (training set). After validation the network with the test vector, this vector is put back into the training set and a new vector is left out. Again the network is trained and validated until all vectors have been left out once.

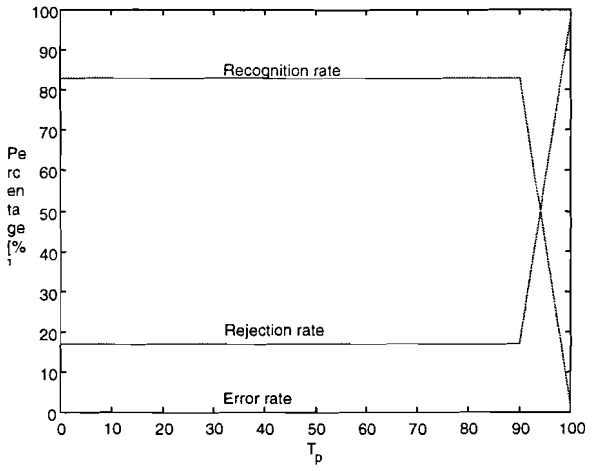


Figure 32 Parameter Set (0.81, 0.10)

Furthermore, we must avoid the classifier is trained and tested with the same data. Since each sample has been measured three times, the chance is high that a vector nearly identical to the test vector is present in the training set. Therefore, we divide the six largest classes into three subsets in which from each sample only one spectrum is present. Now, three sets of 124 samples are available for training and testing. We verified the network with one subset only because of lack of time. Furthermore, the spectra are more or less the same in each subset.

6.2 Results

6.2.1 Classification Accuracy

Unfortunately, we could not evaluate all possible combinations of σ, T_d, and T_p. Therefore, we decided to optimise the parameter set (σ, T_d, T_p) separately. In the first simulation we optimise the error rate r_e by roughly estimating the values for σ and T_d. The parameter T_p is set to zero, and T_d must be selected in such a manner that no vectors or a very few are outside the decision boundaries of the training set as defined by Equation 20. Of course, this assumption is a little bit dangerous in case outliners are introduced in the training data. But this can not be avoided, as outliners are difficult to trace in a sparse data set.

In the next simulation we optimised the parameter T_d, and finally the performance of the classifier has been investigated for yet unknown classes.

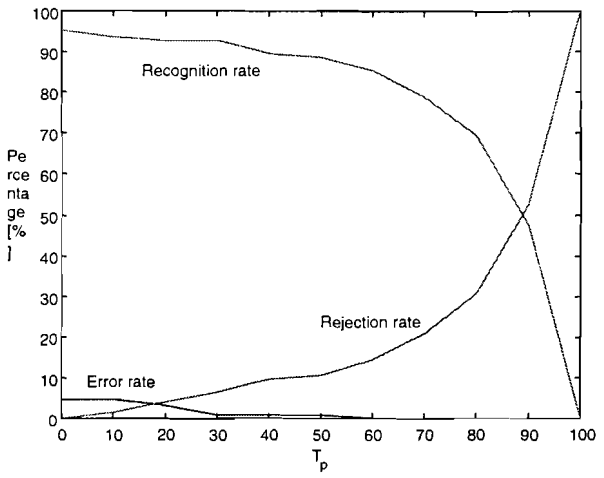


Figure 33 Parameter Set (1.23, 0.06)

In Table 2 the selected parameters from the simulation listed in Appendix D have been summarised. Although more than one set of parameters (σ, T_d) gives a minimum error rate $r_e = 0\%$ we selected three sets only:

1. The parameter set (0.81, 0.10) with maximal accuracy r_a, and maximal recognition rate r_c.
2. The situation (1.23, 0.06) for which the amount of 'unknown' vectors r_u was minimum, and the accuracy r_a maximum. Superfluously, 'unknown' means outside the decision boundaries.
3. The parameter set (1.11, 0.07) is the best trade-off between the number of 'unknown' vectors r_u and the number of wrong decisions r_e.

Table 2 Simulation $T_p = 0\%$

Set	σ	T_d	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	16.94	83.06	0.00	83.06	0.00	100.00
II	1.23	0.06	0	0.00	100.00	4.84	95.16	0.00	95.16
III	1.11	0.07	0	2.42	97.58	3.23	94.35	0.00	96.69

Parameter set (1.23, 0.06) seems the best choice since it meets the requirement of Equation 33 best. The decision boundaries must be clear, as each class exists of about 20 vectors and only 1 is left out for validation. On the other hand it also recognises vectors wrongly, which could indicate outliners.

Since we want to optimise accuracy the parameter set (0.81, 0.10) is best. Hence, we assume that the rejected vectors are probably outliners. Since 0.81 is relatively small we call this a nearest neighbourhood classifier for simplicity.

What effect has the parameter T_p on these results? Introducing the threshold $T_p \neq 0$ in the parameter set (0.81, 0.10) will never improve the accuracy of the classifier – its accuracy is already maximal. Adjusting $T_d \neq 0$ can only increase and decrease the rates r_j and r_c, respectively. Figure 32 shows the latter for $T_p > 90\%$ (rough estimate since T_p = 0, 10, 20, .., 100%). The rejection rate in the figures is defined as the sum of r_j and r_u. Clearly,

the decision boundaries are very sharp defined. Few vectors are rejected because they belong to more than one class (Appendix D).

What effect has T_p on the other parameter sets? In this situation the accuracy can be increased by increasing T_p. Vectors belonging to more than one class will be rejected more frequently (Equation 30). The effect is shown in Figure 33. For T_p = 60% the highest accuracy and highest recognition is reached. Practically the same holds for situation III. At T_p = 70% the highest accuracy and highest recognition is reached as well. The recognition rate of parameter set (1.11, 0.07) is even a slightly better than that of situation II. But this should not be considered too carefully, as the step size of T_p is quite large. The classifier with the largest smoothing parameter σ, and smallest threshold Td is best in generalising vectors outside its decision boundaries. It does not reject vectors as 'unknown', but decides on basis of T_d to reject a vector.

Table 3 Simulation $T_p \neq 0\%$

Set	σ	T_a	T_p	r_u	R_k	r_e	r_c	r_i	r_a
I	0.81	0.10	0	16.94	83.06	0.00	83.06	0.00	100.00
II	1.23	0.06	60	0.00	100.00	0.00	85.48	14.52	100.00
III	1.11	0.07	70	2.42	97.58	0.00	87.10	10.48	100.00

6.2.2 Provision for an Unknown Class

Table 2 lists the new parameter settings of the second simulation. With these parameters we investigated the behaviour of the classifier for the 'unknown' class L/LDPE. The data of L/LDPE from all three subsets can be used together now, and the leave-one-out method is not necessary as we do not train the network again.

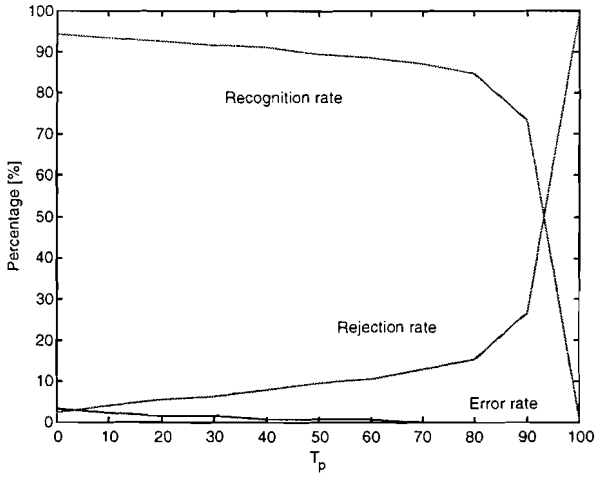


Figure 34 Parameter Set (1.11, 0.07)

We evaluated the effect of classifying the unknown vectors of L/LDPE in correlation with the PCA analysis. Namely, the PCA analysis indicates vectors that could not be

transformed correctly to the new space. Six vectors with a transformation error of about 1.8% (instead of the maximum of 1%) were rejected by the PCA analysis. The other 36 vectors were transformed to the new space with an average accuracy of 0.12 %. Figure 35 shows the rejected vectors. The PCA analysis rejects data, which is not correlated with population **x**. The data which has been transformed correctly, is correlated with population **x**. In the worst situation the not rejected data might even be similar to already existing classes, and classification will never be possible.

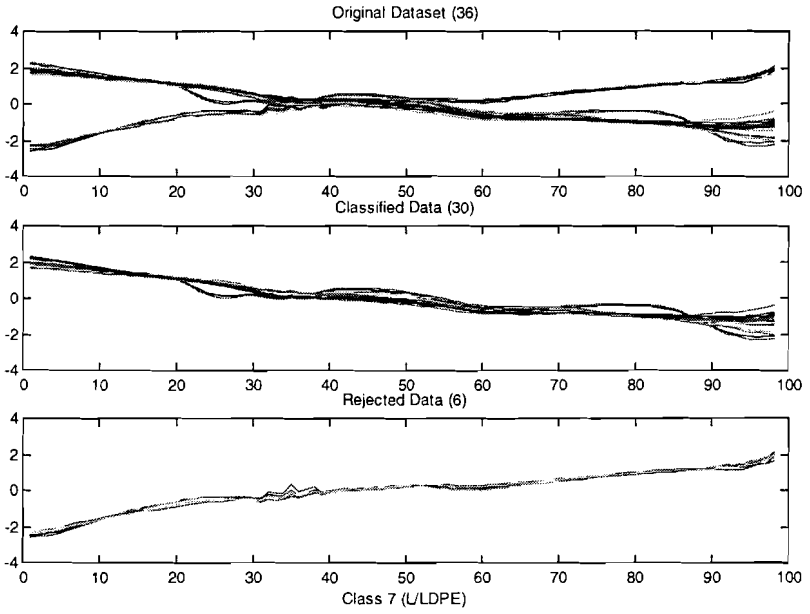


Figure 35 Rejected Vectors of the PCA Analysis

The next three situations were investigated:

1. The network classifies only *not* rejected vectors of the PCA analysis. This means that the PCA analysis of the first simulation is used without any modification.
2. The second case investigated the same for the rejected data.
3. Thirdly, the PCA matrix had been updated with the rejected data included in population **x**. The classification has been repeated with the rejected data. However, the training set was not updated.
4. Finally, the previous simulation was repeated with the vectors that were *not* rejected.

Table 4 Situation I - Not Rejected Vectors and No PCA Modification

Set	σ	T_o	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	96.67	3.33	3.33	0.00	0.00	-
II	1.23	0.06	60	0.00	100.00	10.00	0.00	90.00	-
III	1.11	0.07	70	0.00	100.00	13.33	0.00	86.67	-

In Table 4 the result of situation I is listed. Only the classifier with parameter set (0.81, 0.10) was able to reject most of the not rejected vectors as 'unknown'. The error rate is approximately 3.33%. The classifiers with the other parameter sets could not reject the

unknown vectors as being outside the decision boundaries. They rejected most of the vectors because they were belonging to more than one class. All classifiers were able to reject the rejected data of the PCA analysis (Table 5).

Table 5 Situation II - Rejected Vectors and No PCA Modification

Set	σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	100	0.00	0.00	0.00	0.00	-
II	1.23	0.06	60	100	0.00	0.00	0.00	0.00	-
III	1.11	0.07	70	100	0.00	0.00	0.00	0.00	-

Summarising, not all vectors could be recognised as 'unknown'. The nearest neighbourhood classifier (0.81, 0.10) scored best in rejecting 96.67% of the vectors as 'unknown'.

Now, let us look at the effect of updating the PCA analysis without updating the training set. All previously rejected vectors of the PCA analysis were rejected in the same way as the situation in which no PCA modification took place. After including one rejected vector in the training data all rejected vectors could be classified with 100% accuracy (Table 6). Superfluously, the number of principal components in the new space decreased from 10 to 9.

Table 6 Situation III - Rejected Vectors and PCA Modification

Set	σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	0.00	100	0.00	100.00	0.00	100.00
II	1.23	0.06	60	0.00	100	0.00	100.00	0.00	100.00
III	1.11	0.07	70	0.00	100	0.00	100.00	0.00	100.00

The performance of the classifiers (1.23, 0.06) and (0.11, 0.07) for the not rejected vectors does not change after the PCA modification (Table 7). The error rate of the nearest neighbourhood classifier increases. Most likely, the decision boundaries are only optimal for the PCA matrix, which had been calculated for the six classes. Apparently, after the PCA analysis the decision boundaries have changed slightly. The procedure of Equation 33 could be used to optimise the decision boundaries again, or a new optimum can be found experimentally. The PCA analysis updated with a previously rejected vector has no effect on the classification accuracy of the not rejected vectors.

Table 7 Situation IV - Not Rejected Vectors and PCA Modification

Set	σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	73.32	26.27	26.67	0.00	0.00	-
II	1.23	0.06	60	0.00	100.00	10.00	0.00	90.00	-
III	1.11	0.07	70	0.00	100.00	13.33	0.00	86.66	-

Namely, the PCA analysis rejects vectors, because they were uncorrelated with the vectors used to build the PCA matrix. The not rejected vectors had been transformed correctly, and the high rejection rate r_j indicates the correlation of these vectors with the

training data in the PCA population **x**. Performing a PCA analysis will only enhance the dissimilarities and similarities of the rejected data. The performance of the not rejected data does not change.

Summarising:

- All classifiers were able to reject the rejected vectors of the PCA analysis. No modification of the PCA matrix was needed for this purpose.
- Only the nearest neighbourhood classifier was able to detect most of the rejected vectors as 'unknown'. Again a PCA modification was not needed.
- The other classifiers classified the 'unknown' vectors as a combination of training vectors and rejected them for this reason.
- After modifying the PCA matrix and training set, all rejected vectors could be classified with an accuracy of 100%.
- The performance of the classifier for the not rejected vectors did not change after the PCA modification. Hence, the PCA modification only favours the rejected vectors.
- The error rate of the nearest neighbourhood classifier even increases after the PCA modification. Probably its parameter set (0.81, 0.10) is only well defined for the previous PCA matrix with six classes.

Table 8 Including an 'Unknown' Not Rejected Vector in the Training Set

Set	σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
No PCA modification	0.81	0.10	0	41.37	58.62	3.44	55.17	0.00	94.11
PCA modification	0.81	0.10	0	41.38	58.62	3.44	55.17	0.00	94.12

Till this point, we looked only at the overall performance of the classifier for a whole set of L/LDPE vectors. In contrary, in practice each vector arrives sequentially at the classifier input. A vector that is recognised as 'unknown' could then immediately serve to upgrade the decision boundaries of the network. Fewer vectors will be 'unknown' and the decision boundaries will be extended. The question is what would happen in case we add a not rejected L/LDPE vector to the training set? Probably, the rejection rates r_u and r_j decrease, and the recognition rate r_c increases.

But first, when do we have to add a test vector to the training set? Should we use r_u or r_j as indicator? According to our definition of 'unknown', only the vectors that do not have decision boundaries yet belong to this category. Vectors rejected by the threshold T_p are correlated with more than one class. Hence, their decision boundaries already exist. We select the nearest neighbourhood classifier because it is the only classifier, which is able to detect not rejected vectors of a yet unknown class by means of r_u.

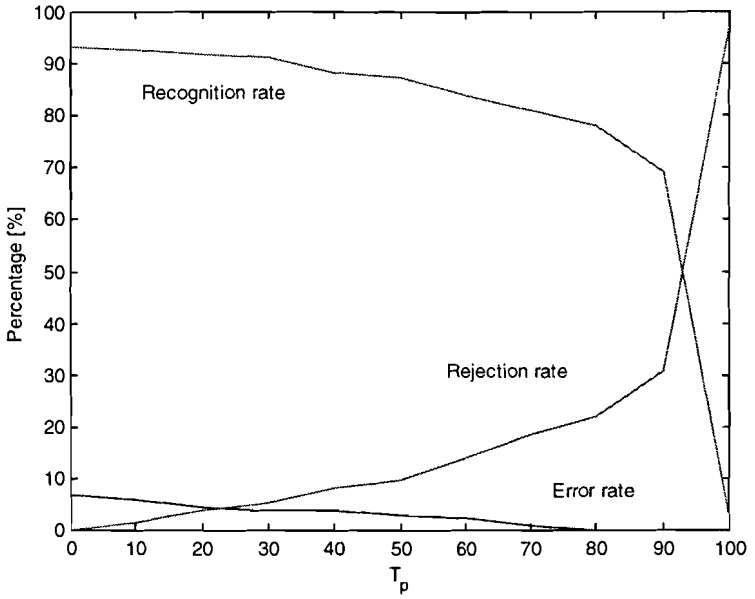


Figure 36 Classifier Performance for Seven Largest Classes

We selected arbitrarily one vector that was 'unknown' from the not rejected vectors, and we used it to upgrade the network. After verifying the classifier with the other 29 vectors, we see an immediately increase in accuracy (Table 8 versus Table 4). The error rate did not decrease. Probably, the unknown vector we selected did not produce previously an error. Those vectors that do are still in the test set. The recognition rate increased, although the decision boundaries are still not well defined. L/LDPE vectors are still rejected as 'unknown'. Modifying the PCA matrix with the new training vector has no effect (Table 8).

Furthermore, does this new training set effect the classification performance of the previous six classes? We verified this situation by applying the leave-one-out validation method, and found that the accuracy of the classifier was still 100% (Table 9). Hence, the new vector does not have a negative contribution to the classifier performance.

Table 9 New Training Set with Leave-one-out Validation

Set	σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
I	0.81	0.10	0	16.03	83.97	0.00	83.97	0.00	100.00
II	0.81	0.10	0	16.03	83.97	0.00	83.97	0.00	100.00

Before we start analysing the effect of new yet unknown classes on the classifier performance, we update the PCA analysis for the seven largest classes. After classification with the leave-one-out validation method we obtain Table 10.

Table 10 Classification of the Seven Largest Classes (0.81, 0.10)

σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
0.81	0.10	0	12.50	87.50	0.63	86.88	0.00	99.29
0.81	0.10	10	12.50	87.50	0.63	86.88	0.00	99.29
0.81	0.10	20	12.50	87.50	0.63	86.25	0.63	99.28
0.81	0.10	30	12.50	87.50	0.63	86.25	0.63	99.28
0.81	0.10	40	12.50	87.50	0.63	85.63	1.25	99.28
0.81	0.10	50	12.50	87.50	0.63	85.63	1.25	99.28
0.81	0.10	60	12.50	87.50	0.63	85.00	1.88	99.27
0.81	0.10	70	12.50	87.50	0.63	85.00	1.88	99.27
0.81	0.10	80	12.50	87.50	0.63	85.00	1.88	99.27
0.81	0.10	90	12.50	87.50	0.63	84.38	2.50	99.26
0.81	0.10	100	12.50	87.50	0.00	4.38	83.13	100.00

We repeat the search for an optimal parameter set (σ, T_d, T_p), since not all vectors belong to the training set anymore. (r_u = 12.50%). The parameter set (1.10, 0.005) turns out to be optimal. No vectors are rejected because they would be outside the boundaries of the training set. The error rate reaches zero for T_p = 80% and the recognition rate is about 78%. Provided that one would like to have higher throughput, selecting T_p = 70% still gives an acceptable accuracy of 99%, and about 80% of the plastics are recognisable.

Table 11 Classification of the Seven Largest Classes (1.10, 0.005)

σ	T_a	T_p	r_u	r_k	r_e	r_c	r_j	r_a
1.10	0.005	0	0.00	100.00	6.62	93.38	0.00	93.38
1.10	0.005	10	0.00	100.00	5.88	92.65	1.47	94.03
1.10	0.005	20	0.00	100.00	4.41	91.91	3.68	95.42
1.10	0.005	30	0.00	100.00	3.68	91.18	5.15	96.12
1.10	0.005	40	0.00	100.00	3.68	88.24	8.09	96.00
1.10	0.005	50	0.00	100.00	2.94	87.50	9.56	96.75
1.10	0.005	60	0.00	100.00	2.21	83.82	13.97	97.44
1.10	0.005	70	0.00	100.00	0.74	80.88	18.38	99.10
1.10	**0.005**	**80**	**0.00**	**100.00**	**0.00**	**77.94**	**22.06**	**100.00**
1.10	0.005	90	0.00	100.00	0.00	69.12	30.88	100.00
1.10	0.005	100	0.00	100.00	0.00	2.94	97.06	100.00

Finally, we simulated what effect the other yet unknown classes have on the performance of the classifier. We merged all subsets of the seven largest classes together into one set of 408 spectra. We did the same for the other classes and the simulation results have been listed Table 12.

Table 12 Classification with Parameter Set (1.10, 0.005, 80)

Class	*Label*	r_u	r_k	r_e	r_c	r_j	*Decision*
8	PA6	0.00	100.00	100.00	0.00	0.00	PA4_6
9	ABS/PA6	0.00	100.00	0.00	0.00	100.00	-
10	SMA/ABS	0.00	100.00	25.00	0.00	75.00	POLY_EST
11	SMA/ABS/PMMA	0.00	100.00	33.33	0.00	66.67	POLY_EST
12	SMA/PMMA	50.00	50.00	50.00	0.00	0.00	POLY_EST
13	ABS/PMMA	50.00	50.00	50.00	0.00	0.00	POLY_EST
14	TH.PLAST.	0.00	100.00	66.67	0.00	33.33	POLY_EST
15	EPDM	0.00	100.00	0.00	0.00	100.00	-

The plastic type PA6 is classified completely as PA4_6. If we compare the spectra of both classes we must conclude that PA6 is identical to PA4_6. However, we can not say whether these two plastics should be classified differently or not. If they should, we never will be able to classify them correctly, as the spectral information is not sufficient.

The combination of ABS/PA6 is completely rejected, since it could be classified as ABS and PA6 at the same time. Most likely, adding this class to the training set would make identification possible. However, due to lack of time we are not able to verify this. Furthermore, the statistical relevance of this simulation can be point of discussion. For example, for the class EPDM only 3 spectra could be verified, and the relationship with real world is not well conditioned.

Despite of this, if these spectra would represent real world, than the plastic type POLY_EST would be contaminated with at least five other types of plastics. Whether this is a negative or positive is not determinable without good knowledge of the recycling market.

6.3 Conclusion and Recommendations

Classification rates of 100% accuracy with a recognition rate of 78% are feasible for seven types of plastics. The plastic (polyester) will be contaminated with at least five other types of plastics. More research is needed in the area of quality measures before we can conclude what effect misclassifications have on quality. Above all, it is recommended to repeat the simulations with more spectra, and to find out exactly why particularly this plastic is contaminated and no other. Maybe defining a stricter decision boundary for this plastic solves the problem. Furthermore, a new identical device in the spectral range from 900 – 2400nm is under development. This device could allow us to extract more relevant spectral information. In the region between 1600 and 1800nm methylene C-H stretches are found, and in the region between 2100 and 2500nm combinations of C-H and O-H vibration bands [33].

Nevertheless, with regard to the classifier solution in relation with the PCA analysis we can conclude that the PCA analysis has to be updated only after the PCA analysis rejected a vector. In case the PNN network rejects a vector, updating the PCA analysis has little to no effect. The PCA analysis is not able to enhance the quality of the features anymore, since it is already nearly optimal. The capability of classifying unknown classes increases very rapidly after only adding one yet unknown class to the PNN structure.

Finally, although not all requirements have been verified, the PNN network meets most of our requirements right now.

7 Conclusion and Recommendations

Theoretically, our solution meets all requirements, but more investigation is needed in practice. Especially, concerning classification speed. In the image processing part six out of seven requirements could be met. The sorting system is insensitive to environmental influences such as ambient light, but on the other hand the effect of temperature has to be investigated yet. Generally, the spectra are insensitive to fillers, fibers, colours, flame retardant, adhesives, dyes, pigments, and other additives. Since we do not know exactly what the compounds of the plastics under investigation were, we can not fully guarantee this for other samples. The effect of dust, paint, labels, food and drink residue, and other surface defects has to be investigated yet. We recommend doing more real world experiments. The recommendations in Chapter 3 might be very useful in setting up these experiments. Plastic objects composed out of more than one plastic such as a soft drink bottle composed of a bottom, a flake, and a cap, could be measured then. According to our image analysis no troubles should be encountered, as the individual parts can be identified separately. All kinds of plastics with irregular dimensions, different weights can be handled. Finally, black and transparent plastics could not be identified. No reliable spectra could be measured. The black and transparent plastics are not forwarded to the feature and classifier part.

The classification part meets only two of the four requirements. Actually, three out of four requirements are met. Namely, the identification unit is able to predict to which plastics group a new as yet unseen sample belongs, and in case this yet unseen sample does not belong to an existing group, a new plastics group can be made. But the classifier can not tell exactly what plastic is classified here. Human intervention will always be needed to label this new class. Although not verified in practice, the sorting system is able to track small changes in equipment or samples over a period of time.

The main disadvantage of the sorting system is its incapability to sort mingled plastics (compositions of more than one type of plastic). The mingled plastics in this research were all classified as polyester. The other plastics (PVC, PP, ABS, PA6, PA4_6, and L/LDPE) were not contaminated, and could be sorted with an accuracy of 100%. No time was left to investigate whether with some additional tuning the purity of polyester could be increased. One should carefully keep in mind that the number of samples that could be sorted with 100% accuracy is very limited. We strongly recommend verifying the classification accuracy for more samples.

Finally, the economical aspects of the sorting system could not be evaluated. With regard to speed we conclude that the sorting speed was limited by the speed of the camera. A cheap low-resolution camera with an integration time of 16ms has been used. Secondly, the speed of the PCA analysis and the classifier was limited by laboratory conditions. We think the speed of the mechanical handling of the plastics and the camera integration time will limit the speed of the sorting system in future. Namely, the PCA analysis can be implemented in digital signal processors or like the classifier in special neural hardware. The latter will profit from high speed parallel processing.

Literature

[1] 'KOSTPRIJZEN AVI'S VERSCHILLEN ENORM', Milieumagazine, December 1997

[2] Amano, Toshio, (Optical Research Inc., Japan), 'NEW PRACTICAL IDENTIFICATION METHOD (AOTF-NIR)', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997

[3] APME (Association of Plastics Manufacturers in Europe), Avenue E van Nieuwenhuyse 4, Box 3, B-1160 Brussels, Belgium.

[4] Beale, Mark, 'NEURAL NETWORK DESIGN', Boston: PWS Publishing Company, 1995

[5] Burns, D.A., and E.W. Ciurczíak, 'HANDBOOK OF NEAR-INFRARED ANALYSIS', New York: Marcel Dekker Inc., 1992

[6] Chatfield, C., and A.J. Collins, 'INTRODUCTION TO MULTIVARIATE ANALYSIS', London: Chapman and Hall, 1980

[7] Chen, C.H., 'FUZZY AND NEURAL NETWORK HANDBOOK', London: McGraw-Hill, 1996

[8] Cooley, William W., and Paul R. Lohnes, 'MULTIVARIATE DATA ANALYSIS', New York: John Wiley & Sons, Inc., 1971

[9] Corbet, K.S., L. Jung, L. Dillon, and M.J. Olasavsky, 'PLASTIC RECYCLING IN THE ELECTRONICS INDUSTRY', Proceedings of the 1997 IEEE International Symposium on Electronics and the Environment, San Franciso, CA, USA, 5-7 May 1997, New York, NY, USA: IEEE, 1997, p. 297-302, ISBN: 0 7803 3808 1 (Cat. No. 97CH36035).

[10] De Groene, A., and J. Vermunt, 'BURNING OF WASTE: PROBLEMS OF CAPACITY OF BURNING INSTALLATIONS', In: Proceedings first international working seminar on reuse, Eindhoven: Eindhoven University of Technology, 11 December 1996.

[11] Diamantaras, K. I., and S.Y. Kung, 'PRINCIPAL COMPONENT NEURAL NETWORKS – THEORY AND APPLICATIONS', New York: John Wily & Sons, Inc., 1996

[12] Didde, R., 'GESLEEP MET AFVAL HOUDT AAN', Milieumagazine, July-August 1996

[13] Didde, R., 'NASCHEIDING GAAT HET MAKEN', Milieumagazine, July-August 1996

[14] Eisenreich, N., and Th. Rohe, 'IDENTIFIZIEREN VON KUNSTSTOFFEN', Kunststoffe, Vol. 86 (1996) 2, p. 222-224

[15] Eisenreich, N., H. Kull, and Th. Rohe, 'SCHNELLE KUNSTSTOFFIDENTIFIKATION MIT NAHINFRAROT-SPEKTROSKOPIE', Symposium ICT-Fraunhofer-Institut, Pfinztal (Berghausen), 1996

[16] Fausett, Laurence, 'FUNDAMENTALS OF NEURAL NETWORKS – ARCHITECTURES, ALGORITHMS, AND APPLICATIONS', London: Prentice-Hall International, Inc., 1994

[17] Gemperline, P.J., and N.R. Boyer, 'CLASSIFICATION OF NEAR-INFRARED SPECTRA USING WAVELENGTH DISTANCES: COMPARISON TO THE MAHALABIS DISTANCE AND RESIDU VECTOR', Analytical Chemist 67 (1995), p.160-166

[18] Hummel, Dieter D., 'POLYMERE: STRUKTUR UND SPEKTRUM', München: Hanser, 1978, 'Atlas der Polymer und Kunststoffanalyse: Bd.'

[19] Huth-Fehre, T., R. Feldhoff, Th. Kantimm, L. Quick, F. Winter, K. Cammann, W. Van den Broek, D. Wienke, W. Melssen, and L. Buydens, 'NIR - REMOTE SENSING AND ARTIFICIAL NEURAL NETWORKS FOR RAPID IDENTIFICATION OF POST CONSUMER PLASTICS', Journal of Molecular Structure 348 (1995) 143, pp. 143-146.

[20] Johnson, A. Richard, and Dean W. Wichern, 'APPLIED MULTIVARIATE STATISTICAL ANALYSIS', London: Prentice-Hall International, Inc., 1992

[21] Jolliffe, I.T., 'PRINCIPAL COMPONENT ANALYSIS', New York: Springer-Verlag, 1986

[22] Kelly, Erin Patrice, 'QUALITY CONROL METHODS ENSURE HIGH-QUALITY PLASTICS FOR RECYCLING', Resource Recycling, Washington: American Plastics Council, Inc., October 1996

[23] Klebba, R., 'PLASTICS IDENTIFICATION BY PYROLYSIS FT-IR', In: Automatic identification and sorting of plastics from different waste streams', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997

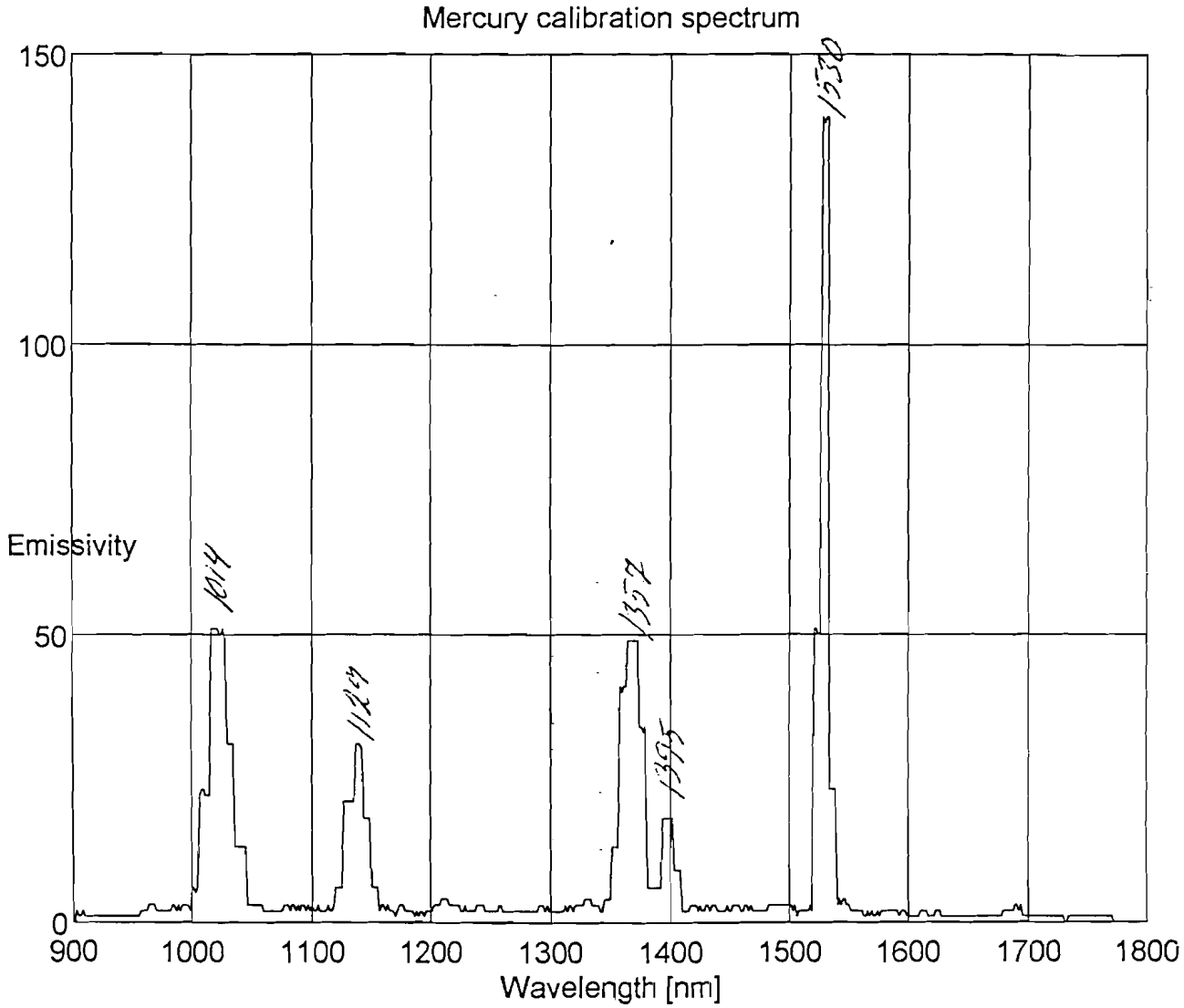
[24] Kosko, Bart, 'NEURAL NETWORKS AND FUZZY SYSTEMS', London: Prentice-Hall International, Inc., 1992

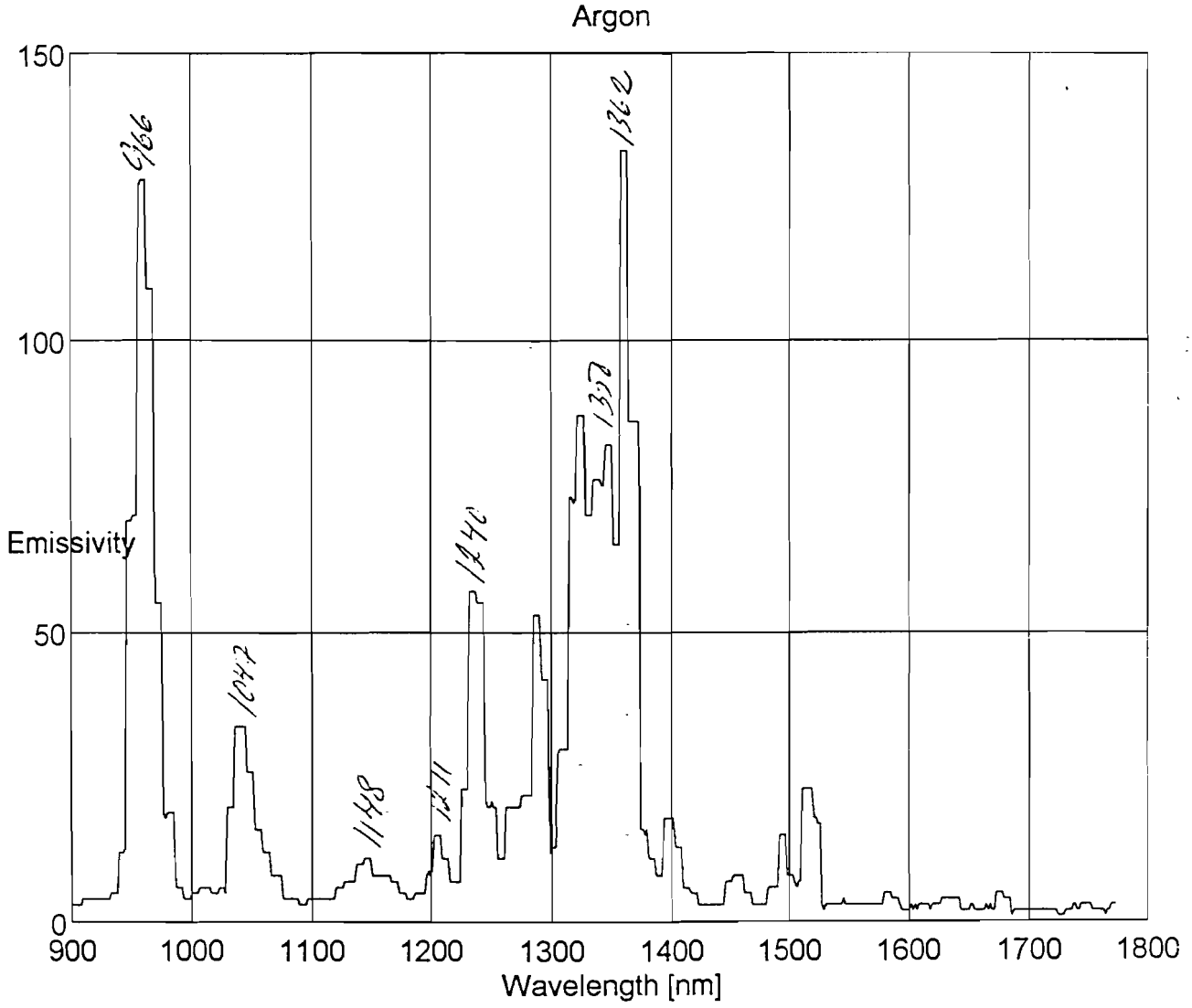
[25] Kreyszig, E., 'ADVANCED ENGINEERING MATHEMATICS', New York: John Wiley & Sons, Inc., 1993

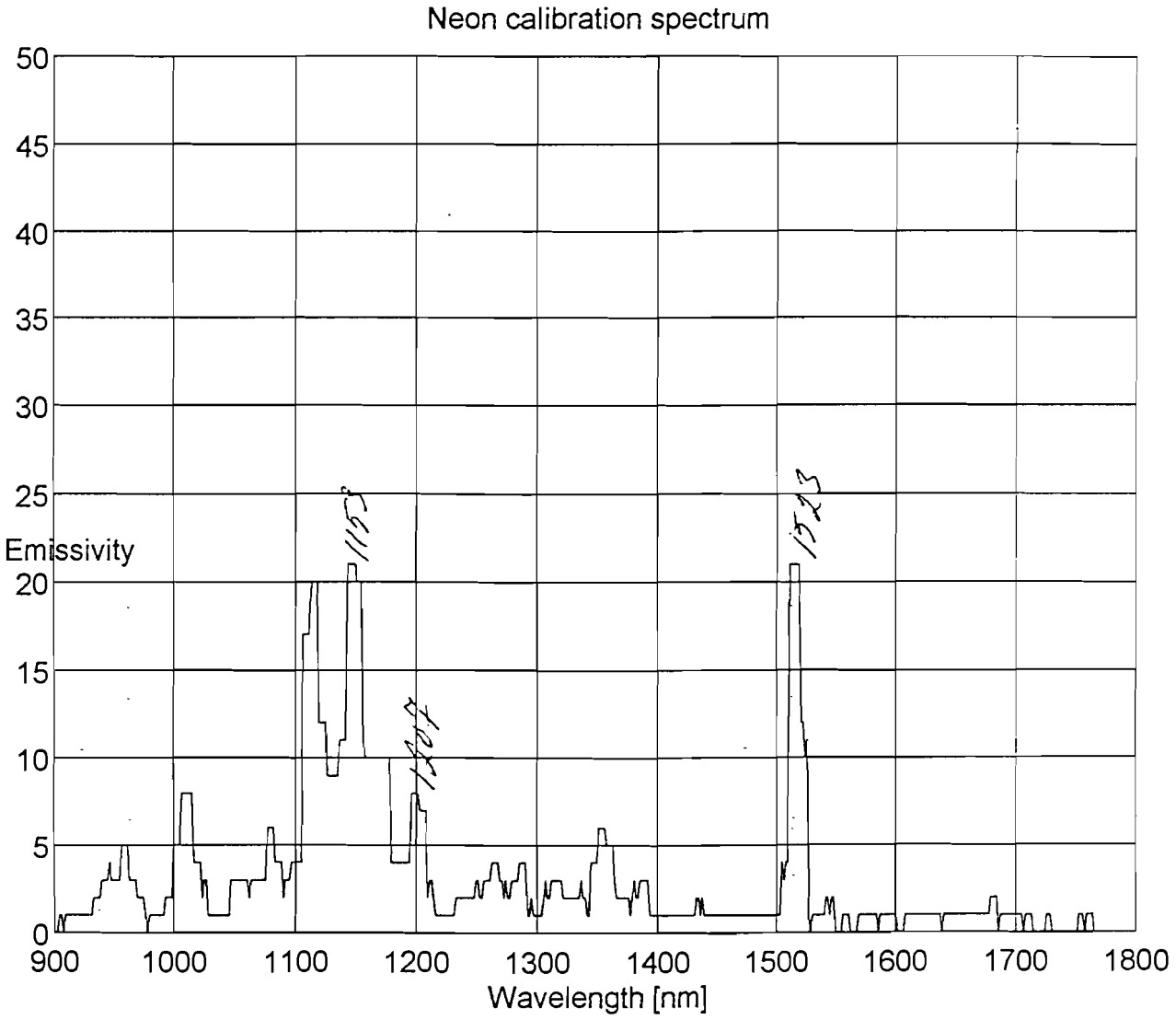
[26] Mark, Frank E. et al., 'ENERGY RECOVERY – ANALYSIS OF PLASTICS IN MUNICIPAL SOLID WASTE', Brussels: AMPE, February 1995
[27] Mark, Frank E., 'THE ROLE OF PLASTICS IN MUNICIPAL SOLID WASTE COMBUSTION', Brussels: European Centre for Plastics in the Environment (PWMI), 1995
[28] Marsfield, J.R., M.G. Sowa, G.B. Scarth, R.L. Somorjai, and M.H. Mantsch, 'ANALYSIS OF SPECTROSCOPIC IMAGING DATA BY FUZZY C-MEANS CLUSTER', Analytical Chemistry, Vol. 69 (1997) 16, p.3370-3374
[29] Massen, R., and M. Simnacher, 'OPTICAL SORTATION OF PLASTICS BY COLOUR AND COMPOSITION', Massen Machine Vision Systems GmbH, Am Seerhein 8, D-78467 Constance, Germany
[30] McClure, W.F., 'THE GIANT IS RUNNING STRONG', Anal. Chem. 66 (1994) 43A-53A.
[31] Murase, Atsuchi, (Toyota Central R&D Labs, Inc.), 'RAPID IDENTIFICATION OF PLASTICS BY PYROLYSIS IR AND HAND-HELD DIELECTRIC INSTRUMENT', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[32] Orr, Mark, J. L., 'INTRODUCTION TO RADIAL BASIS FUNCTION NETWORKS', Centre for Cognitive Science, University of Edinburgh, Scotland, April 1996
[33] Osborne, B.G. and T. Fearn, 'NEAR INFRARED SPECTROSCOPY IN FOOD ANALYSIS', 2nd impr., New York: Longman Scientic & Technical 1988
[34] Overton, B. W., 'THE HIGH SPEED AND POSITIVE IDENTIFICATION OF POLYMERS, AND VARIANTS OF POLYMERS, BY THE DETECTION OF FLUORESCENT TRACERS', In: Automatic identification and sorting of plastics from different waste streams', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[35] Patterson, Dan. W., 'ARTIFICIAL NEURAL NETWORKS – THEORY AND APPLICATIONS', London: Prentice-Hall International, Inc., 1996
[36] Feldhoff, R., D. Wienke, K. Cammann, and H. Fuchs, 'ON-LINE POST CONSUMER PACKAGE IDENTIFICATION BY NIR SPECTROSCOPY COMBINED WITH A FUZZYARTMAP CLASSIFIER IN AN INDUSTRIAL ENVIRONMENT', Applied Spectroscopy, Vol. 51 (1997), No.3, p.362-368.
[37] Radush, J.J., W. Engel, S. Naaga, and D.E. Merseburg, 'RAPID X-RAY IDENTIFICATION OF POLYMERS FOR THE PURPOSE OF RECYCLING', Fraunhofer-Institut für Chemische Technology, 27 June 1996, 87 p., Available: TIB Hannover, F96B1447
[38] Rashid Khan, M. 'CONVERSION AND UTILIZATION OF WASTE MATERIALS', USA: Taylor & Francis, 1996, ISBN 1-56032-382-5
[39] Rikken, B., 'ALLES IS TE SORTEREN', Afval, No. 6, August 1997
[40] Russ, John C., 'THE IMAGE PROCESSING HANDBOOK', London: CRC Press 1995
[41] Scott, D.M., 'A 2-COLOR NEAR-INFRARED SENSOR FOR SORTING RECYCLED PLASTIC WASTE', Measurement Science and Technology, Vol. 6 (1995) 2, p.156-159
[42] Scott, D.M., and R.L. Waterland, 'IDENTIFICATION OF PLASTIC WASTE USING SPECTROSCOPY AND NEURAL NETWORKS', Polymer Engineer and Science 35 (1995) 12, p. 1011-1015, Du Pont de Nemours, Centrum Resource & Development (USA)
[43] Scott, J.A.N., 'THE MECHANICAL RECYCLING OF MIXED PLASTICS WASTE – A THREE YEAR STUDY INTO THE PROCESSING AND PROPERTIES OF MECHANICALLY RECYCLED MIXED PLASTICS WASTE', Brussels: AMPE, October 1994
[44] SENTER, Grote Markstraat 43, Postbus 30732, 2500 GS Den Haag
[45] Siesler, H.W., and K.Holland-Moritz, 'INFRARED AND RAMAN SPECTROSCOPY OF POLYMERS', New York: Marcel Dekker Inc., 1980
[46] Sivia, D. S., 'DATA ANALYSIS – A BAYESIAN TUTORIAL', Oxford: Clarendon Press, 1996
[47] Slapak, M.J.P., and J.M.N. van Kasteren, 'SELECTION OF A RECYCLING ROUTE FOR HETEROGENEOUS PVC-WASTE' In: Proceedings first international working seminar on reuse, Eindhoven 11 December 1996, Eindhoven University of Technology
[48] Sofres Group, 'PLASTICS RECOVERY IN PERSPECTIVE – PLASTICS CONSUMPTION AND RECOVERY IN WESTERN EUROPE 1995', Brussels: AMPE, 1995
[49] Sommer, J. Edward, (National Recovery Technologies, Inc.), 'TECHNIQUES FOR RAPID IDENTIFICATION AND SORTING OF PLASTICS', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997

[50] Suykerbuyk, M.A.W., and A.M. Ransdorp, 'INVENTARISATIE PREVENTIEMAATREGELEN IN DE KUNSTSTOFVERWERKENDE INDUSTRIE', Rapport: Ministerie van VROM – Direktie afvalstoffen, Den Haag, September 1996, No.: 21709/206
[51] Ten Houten, A.Th, 'RECYCLING VAN BEDRIJFSAFVALSTOFFEN: NIEUWE TECHNIEKEN EN PROCESSEN OM OP ECONOMISCHE EN KOSTEN BESPARENDE WIJZE AFVALSTOFFEN OPNIEUW TE GEBRUIKEN', Amsterdam: WEKA Uitgeverij BV, 1997
[52] TiTech Autosort 700 KT Datasheet, IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[53] Uhlig, Dieter, E., 'PLASTIC WASTE SORTING IN DS AND AUTOMATIC IDENTIFICATION – DUALES SYSTEM DEUTSCHLAND', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[54] Van de Vusse, S.G., and F. Soomers, 'NASCHEIDING EN BE/VERWERKING VAN KUNSTOFFEN UIT HUISHOUDELIJK AFVAL', Den Haag: Rapport van Ministerie van Volkshuisvesting, Ruimtelijke Ordening, 1996
[55] Van den Broek, W.H.A.M., D. Wienke, W.J. Melssen, and L.M.C. Buydens, 'OPTIMAL WAVELENGTH RANGE SELECTION BY A GENETIC ALGORITHM FOR DISCRIMINATION PURPOSES IN SPECTROSCOPIC INFRARED IMAGING', Appl. Spectroscopy, Vol. 51 (1997), No.8, p.1210-1217.
[56] Van den Broek, W.H.A.M., D. Wienke, W.J. Melssen, C.W.A. de Crom, and L. Buydens, 'IDENTIFICATION OF PLASTICS AMONG NONPLASTICS IN MIXED WASTE BY REMOTE SENSING NEAR-INFRARED IMAGING SPECTROSCOPY. 1. IMAGE IMPROVEMENT AND ANALYSIS BY SINGULAR VALUE DECOMPOSITION', Anal. Chem. 67 (1995), p.3753-3759.
[57] Van den Broek, W.H.A.M., D. Wienke, W.J. Melssen, R. Feldhoff, T. Huth-Fehre, T. Kantimm, and L.M.C. Buydens, 'APPLICATION OF A SPECTROSCOPIC INFRARED FOCAL PLANE ARRAY SENSOR FOR ON-LINE IDENTIFICATION OF PLASTIC WASTE', Applied Spectroscopy, Vol. 51 (1997), No.6, p.856-865.
[58] Van den Heijden, F., 'IMAGE BASED MEASUREMENT SYSTEMS: OBJECT RECOGNITION AND PARAMETER ESTIMATION', London: John Wiley & Sons Ltd., 1994
[59] Wasserman, Philip D., 'ADVANCED METHODS IN NEURAL COMPUTING', New York: Van Nostrand Reinhold, 1993
[60] Weyts, Guy, 'HOW CAN THE IDENTIFICATION OF PLASTICS CONTRIBUTE TO MECHANICAL RECYCLING?', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[61] Wienke, D., W. van den Broek, and L. Buydens, 'IDENTIFICATION OF PLASTICS AMONG NONPLASTICS IN MIXED WASTE BY REMOTE SENSING NEAR-INFRARED IMAGING SPECTROSCOPY. 2. MULTIVARIATE IMAGE RANK ANALYSIS FOR RAPID CLASSIFICATION', Anal. Chem. 67 (1995), p.3760-3766.
[62] Wienke, D., W. Van den Broek, W. Melssen, L. Buydens, R. Feldhoff, T. Kantimm, T. Huth-Fehre, L. Quick, F. Winter, and K. Cammann, 'COMPARISON OF AN ADAPTIVE RESONANCE THEORY BASED NEURAL NETWORK (ART-2A) AGAINST OTHER CLASSIFIERS FOR RAPID SORTING OF POST CONSUMER PLASTICS BY REMOTE NEAR-INFRARED SPECTROSCOPIC SENSING USING AN INGAAS DIODE ARRAY', Analytica Chimica Acta 317 (1995) 1-16.
[63] Wienke, D., W.H.A.M. van den Broek, W. Melssen, and L.M.C. Buydens, 'NEAR-INFRARED IMAGING SPECTROSCOPY (NIRIS) AND IMAGE RANK ANALYSIS FOR REMOTE IDENTIFICATION OF PLASTICS IN MIXED WASTE', Fresnius' Journal of Analytical Chemistry 354 (1996) 7-8, p.823-828.
[64] Wienke, D., W.H.A.M. van den Broek, W. Melssen, and L.M.C. Buydens, 'NEAR-INFRARED IMAGING SPECTROSCOPY (NIRIS) AND IMAGE RANK ANALYSIS FOR REMOTE IDENTIFICATION OF PLASTICS IN MIXED WASTE', J. Anal. Chem. 354 (1996) 823.
[65] Willenberg, B., W. Kohn, and H. Ritzmann, 'IDENTIFICATION OF PLASTICS USING NIR', Kunstverarbeiter 44 (1993) 12, p.12-25.
[66] Willenberg, Bernd, (Brüker Analytische Messtechnik', 'MIR: A PROVEN TECHNOLOGY FOR IDENTIFICATION OF ENGINEERING PLASTICS', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997
[67] Yoneda, Kenichi, (Mitsubishi heavy Industries Ltd.), 'NIR AUTOMATIC IDENTIFICATION AND SORTING SYSTEM', IdentiPlast Conference, Brussels, Belgium, 27 – 28 October 1997

Appendix A – Calibration Lamp




Argon
150
100
50
0
Emissivity
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
Wavelength [nm]
966
1042
1148
1211
1340
1357
1362


Neon calibration spectrum
Emissivity
Wavelength [nm]
50
45
40
35
30
25
20
15
10
5
0
900
1000
1100
1200
1300
1400
1500
1600
1700
1800

Appendix B – Equipment

Light Source
Fabricator: Osram
Type: 64635 with pure gold plated coating.
Ratings: 150W/ 15V

Prism
This information is confidential.

Sensor
This information is confidential.

Appendix C - Samples

File Format
The images have been stored in MATLAB-format (filename written by 8 characters). Their filenames have been numbered to allow automatic data processing. The next code has been used:

[E] ee [S] s XXX

The two characters after [E] indicate the experiment number. The first letter after [S] indicates the number of measurement taken from the sample. The sample itself is indicated by the last 3 numbers (X).

For example, the first measurement taken from sample 001 in experiment 1 is denoted with E01S1001, the second measurement with E01S2001 etc. For experiment 2, the third measurement on sample 345 is denoted with E02S3345.

In the image files also basic information about the sample (colour and material type) and the measurement conditions (remarks) have been included. Each file can be extended with information on the ways that processing of the image took place. For example, the measurement of sample E01S1000 is stored in the MATLAB file E01S1000.mat and has the next structure:

E01S1000 =
sample_data: [128x128 double]
label: 'PC/ABS'
colour: 'BLACK'
remarks: 'Surface is smooth.'
actions: []

Samples
See next pages.

EXPERIMENT 1 – BLACK AND TRANSPARENT PLASTICS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E01S0000	3	PC/ABS	BLACK	GLANCE
E01S0001	3	ABS	BLACK	GLANCE
E01S0003	3	PP	BLACK	
E01S0004	3	PP	LIGHT GREY	
E01S0005	3	ABS/PA6	DARK GREY	
E01S0006	3	PVC	LIGHT BROWN	
E01S0007	3	TH. ELAST	BLACK	SOFT
E01S0008	3	PP	DARK BLUE	
E01S0009	3	PP	LIGHT GREY	
E01S0010	3	PP	BLACK	
E01S0011	3	PP	DARK BLUE	
E01S0012	3	PP	DARK BLUE	
E01S0011	3	PP	DARK BLUE	
E01S0012	3	PP	DARK BLUE	
E01S0013	3	PP	DARK BLUE	
E01S0014	3	PP	DARK BLUE	
E01S0015	3	PE	LIGHT GREY	GLANCE
E01S0050	3	PC	TR. ORANGE	GLANCE
E01S0051	3	PC	TR. RED	GLANCE
E01S0052	3	PC	TRANSPARENT	
E01S0053	3	L/LDPE	TRANSPARENT	COTTON
E01S0054	3	PC	TRANSPARENT	MILK
E01S0055	3	PC	TRANSPARENT	MILK
E01S0056	3	L/LDPE	TRANSPARENT	BAG
E01S0057	3	L/LDPE	TRANSPARENT	BAG
E01S0058	3	PP	TRANSPARENT	CHEWGUM
E01S0059	3	PP	TRANSPARENT	CHEWGUM
E01S0060	3	PET	TRANSPARENT	BOTTLE
E01S0061	3	PET	TRANSPARENT	BOTTLE
E01S0062	3	PET	TRANSPARENT	BOTTLE

EXPERIMENT 2 – GEOMETRICAL RELATIONSHIPS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E02S0001	3	POLY ESTHER	WHITE	
E02S0002	3	POLY ESTER	WHITE	
E02S0003	3	SMA/ ABS	WHITE	
E02S0004	3	CARD BOARD	LIGHT BROWN	
E02S0005	3	TEXTILE	WHITE	Calibration different!
E02S0006	3	WOOD	BROWN	Calibration different!
E02S0007	3	PE	WHITE	Calibration different!

EXPERIMENT 3 – VARIATION IN THICKNESS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E03S0001	3	PP	WHITE	1.70mm
E03S0002	3	PP	WHITE	3.17mm
E03S0003	3	PP	WHITE	3.94mm
E03S0004	3	HDPE	WHITE	1.15mm
E03S0005	3	HDPE	WHITE	1.66mm
E03S0006	3	HDPE	WHITE	1.75mm
E03S0007	3	HDPE	WHITE	1.6mm
E03S0008	3	HDPE	WHITE	3.2mm
E03S0009	3	HDPE	WHITE	4.8mm
E03S0010	3	HDPE	WHITE	6.54mm
E03S0011	3	ABS	LIGHT GREY	2.0mm
E03S0012	3	ABS	LIGHT GREY	4.0mm
E03S0013	3	ABS	LIGHT GREY	6.0mm
E03S0014	3	ABS	LIGHT GREY	8.4mm
E03S0015	3	PA4_6	WHITE/YELLOW	1.0mm
E03S0016	3	PA4_6	WHITE/YELLOW	2.0mm
E03S0017	3	PA4_6	WHITE/YELLOW	3.0mm
E03S0018	3	PA4_6	WHITE/YELLOW	4.02mm
E03S0019	3	LDPE	TRANSPARENT	10um
E03S0020	3	LDPE	TRANSPARENT	20um
E03S0021	3	LDPE	TRANSPARENT	30um
E03S0022	3	LDPE	TRANSPARENT	40um
E03S0023	3	LDPE	TRANSPARENT	50um
E03S0024	3	LDPE	TRANSPARENT	60um
E03S0025	3	LDPE	TRANSPARENT	80um

EXPERIMENT 4 – PLASTICS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E04S0001	3	L/LDPE	LIGHT GREEN	Equal to sample 04S0002
E04S0002	3	L/LDPE	LIGHT GREEN	Equal to sample 04S0001
E04S0003	3	L/LDPE	GREEN	Equal to sample 04S0004
E04S0004	3	L/LDPE	GREEN	Equal to sample 04S0003
E04S0005	3	L/LDPE	WHITE	
E04S0006	3	L/LDPE	WHITE	2 x thicker
E04S0007	3	L/LDPE	WHITE	2 x thicker
E04S0008	3	L/LDPE	WHITE	
E04S0009	3	L/LDPE	WHITE	
E04S0010	3	L/LDPE	WHITE	
E04S0011	3	L/LDPE	WHITE	
E04S0012	3	L/LDPE	WHITE	
E04S0013	0	L/LDPE		
E04S0014	0	L/LDPE		
E04S0015	0	L/LDPE		
E04S0016	0	L/LDPE		
E04S0017	0	L/LDPE		
E04S0018	0	L/LDPE		
E04S0019	0	L/LDPE		
E04S0020	0	L/LDPE		
E04S0021	3	HDPE	WHITE	
E04S0022	3	HDPE	WHITE	
E04S0023	3	HDPE	WHITE	
E04S0024	3	HDPE	WHITE	
E04S0025	3	HDPE	WHITE	

EXPERIMENT 4 – PLASTICS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E04S0030	3	HDPE	WHITE	
E04S0031	3	HDPE	WHITE	
E04S0032	3	HDPE	WHITE	
E04S0033	3	HDPE	WHITE	
E04S0034	3	HDPE	WHITE	
E04S0035	3	HDPE	WHITE	
E04S0026	3	HDPE	WHITE	
E04S0027	3	HDPE	WHITE	
E04S0028	3	HDPE	WHITE	
E04S0029	3	HDPE	WHITE	
E04S0036	3	HDPE	WHITE	
E04S0037	3	HDPE	WHITE	
E04S0038	3	HDPE	WHITE	
E04S0039	3	HDPE	WHITE	
E04S0040	3	HDPE	WHITE	
E04S0041	3	PP	YELLOW	
E04S0042	3	PP	YELLOW	
E04S0043	3	PP	YELLOW	
E04S0044	3	PP	YELLOW	
E04S0045	3	PP	YELLOW	
E04S0046	3	PP	YELLOW	
E04S0047	3	PP	YELLOW	
E04S0048	3	PP	YELLOW	
E04S0049	3	PP	BROWN	
E04S0050	3	PP	BROWN	
E04S0051	3	PP	RED	
E04S0052	3	PP	RED	
E04S0053	3	PP	RED	
E04S0054	3	PP	RED	
E04S0055	3	PP	RED	
E04S0056	3	PP	RED	
E04S0057	3	PP	RED	
E04S0058	3	PP	WHITE	
E04S0059	3	PP	ORANGE	
E04S0060	3	PP	ORANGE	
E04S0061	3	PP	BLUE	
E04S0062	3	PVC	BLUE	Equal to sample 04S0063
E04S0063	3	PVC	BLUE	Equal to sample 04S0062. "Layer"
E04S0064	3	PVC	GREEN	Equal to sample 04S0065
E04S0065	3	PVC	GREEN	Equal to sample 04S0064. "Layer"
E04S0066	3	PVC	GREEN	Equal to sample 04S0067. "Layer"
E04S0067	3	PVC	GREEN	Equal to sample 04S0066.
E04S0068	3	PVC	WHITE	Equal to sample 04S0069.
E04S0069	3	PVC	WHITE	Equal to sample 04S0068. "Layer"
E04S0070	3	PVC	YELLOW	Equal to sample 04S0071. "Layer"
E04S0071	3	PVC	YELLOW	Equal to sample 04S0070.
E04S0072	3	PVC	YELLOW	Equal to sample 04S0073.
E04S0073	3	PVC	YELLOW	Equal to sample 04S0072. "Layer"
E04S0074	3	PVC	WHITE	Equal to sample 04S0075.
E04S0075	3	PVC	WHITE	Equal to sample 04S0074.

EXPERIMENT 4 – PLASTICS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E04S0076	3	PVC	BROWN	Equal to sample 04S0077. "Layer"
E04S0077	3	PVC	BROWN	Equal to sample E04S0076.
E04S0078	3	PVC	BLUE	Equal to sample 04S0079. "Layer"
E04S0079	3	PVC	BLUE	Equal to sample 04S0080.
E04S0080	3	PVC	WHITE	
E04S0081	3	PVC	WHITE	
E04S0082	3	PVC	WHITE	
E04S0083	3	PA4_6	WHITE	
E04S0084	3	PA4_6	WHITE	
E04S0085	3	PA4_6	WHITE	
E04S0086	3	PA4_6	WHITE	
E04S0087	3	PA4_6	WHITE	
E04S0088	3	PA4_6	WHITE	
E04S0089	3	PA4_6	WHITE	
E04S0090	3	PA4_6	WHITE	
E04S0091	3	PA4_6	WHITE	
E04S0092	3	PA4_6	WHITE	
E04S0093	3	PA4_6	WHITE	
E04S0094	3	PA4_6	WHITE	
E04S0095	3	PA4_6	WHITE	
E04S0096	3	PA4_6	WHITE	
E04S0117	3	POLY. EST.	WHITE	
E04S0118	3	POLY. EST.	WHITE	
E04S0119	3	POLY. EST.	WHITE	
E04S0120	3	POLY. EST.	WHITE	
E04S0121	3	POLY. EST.	WHITE	
E04S0122	3	POLY. EST.	WHITE	
E04S0123	3	POLY. EST.	WHITE	
E04S0124	3	POLY. EST.	WHITE	
E04S0125	3	ABS	WHITE	
E04S0126	3	ABS	WHITE	
E04S0127	3	ABS	WHITE	
E04S0128	3	ABS	WHITE	
E04S0129	3	ABS	WHITE	
E04S0130	3	ABS	WHITE	
E04S0131	3	ABS	WHITE	
E04S0132	3	ABS	BROWN	
E04S0133	3	ABS	BROWN	
E04S0134	3	ABS	WHITE	
E04S0135	3	ABS	WHITE	
E04S0136	3	ABS	WHITE	
E04S0137	3	ABS	WHITE	
E04S0138	3	ABS	WHITE	
E04S0139	3	ABS	WHITE	
E04S0140	3	ABS	WHITE	
E04S0141	3	ABS	WHITE	
E04S0142	3	ABS	WHITE	
E04S0143	3	ABS	GREY	
E04S0144	3	ABS	GREY	
E04S0145	3	ABS	BROWN	
E04S0146	3	SMA/ABS	WHITE	
E04S0147	3	SMA/ABS	WHITE	
E04S0148	3	SMA/ABS	WHITE	
E04S0149	3	SMA/ABS	WHITE	
E04S0150	3	SMA/PMMA	WHITE	
E04S0151	3	SMA/PMMA	WHITE	
E04S0152	3	EPDM	LIGHT GREY	

EXPERIMENT 4 – PLASTICS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E04S0153	3	SMA/ABS/PMMA	WHITE	
E04S0154	3	SMA/ABS/PMMA	WHITE	
E04S0155	3	SMA/ABS/PMMA	WHITE	
E04S0156	3	TH.PLAST. ELASTO.	WHITE	
E04S0157	3	ABS/PA6	GREEN	
E04S0158	3	ABS/PA6	WHITE	
E04S0159	3	ABS/PA6	GREY	
E04S0160	3	ABS/PA6	WHITE	
E04S0161	3	ABS/PA6	WHITE	

EXPERIMENT 6 – MIXED PLASTIC OBJECTS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E06S0000	3	PP	BLUE	DRINK CUP. calibration different!
E06S0001	3	PP	ORANGE	DRINK CUP. calibration different!
E06S0002	3	PP	LIGHT GREEN	DRINK CUP. calibration different!
E06S0003	3	PP	DARK GREEN	DRINK CUP. calibration different!
E06S0004	3	PP	YELLOW	DRINK CUP. calibration different!
E06S0006	3	PE	LIGHT GREY	Parker box. calibration different!
E06S0007	3	PET	LIGHT BLUE	Telephone card. calibration different!
E06S0008	3	PLASTIC?	LIGHT BLUE	Telephone card. calibration different!
E06S0009	3	METAL/PLASTIC	BLACK	Razorblade
E06S0010	3	TEXTILE	YELLOW	
E06S0011	3	PAPER	WHITE	
E06S0012	3	TEXTILE/PLASTIC	WHITE	Plaster up.
E06S0013	3	SUGAR	WHITE	
E06S0014	3	METAL	METAL	Key
E06S0015	3	MEDICINE	WHITE/ROSE	Aspirin
E06S0016	3	MEDICINE/PLASTIC	WHITE/ROSE	Aspirin in package
E06S0017	3	TEXTILE/PLASTIC	WHITE	Plaster down.
E06S0018	3	HDPE	LIGHT YELLOW	Bottle
E06S0019	3	HDPE	LIGHT YELLOW	Bottle with label
E06S0020	3	PP	GREEN	Tap and bottle 06s0018
E06S0021	3	PE	RED	bottle
E06S0022	3	PE	RED	bottle with label.
E06S0023	3	PP?	RED	Tap and bottle

EXPERIMENT 6 – MIXED PLASTIC OBJECTS

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E06S0024	3	PE	PINK	Bottle
E06S0025	3	PE	PINK	Bottle with label.
E06S0026	3	PP	PINK	Tap and bottle 06s0024
E06S0030	3	PE	WHITE	Bottle
E06S0031	3	PE	WHITE	Bottle with label.
E06S0032	3	PP	WHITE	Tap and bottle 06s0032
E06S0033	3	PE	WHITE	Bottle
E06S0034	3	PE	WHITE	Bottle with label.
E06S0035	3	PP	WHITE	Tap of bottle 06s0033
E06S0036	3	PE	TRANSPARENT	Bottle
E06S0037	3	PE	TRANSPARENT	Bottle with label.
E06S0038	3	PP	WHITE	Tap of bottle 06s0036
E06S0036	3	PE	TRANSPARENT	Bottle
E06S0037	3	PE	TRANSPARENT	Bottle with label.
E06S0038	3	PP	WHITE	Tap and bottle 06s0036
E06S0039	3	PP	YELLOW	Cup (butter)
E06S0040	3	PP	YELLOW	Tap of cup 06s0039.
E06S0041	0			
E06S0042	3	PE	PURPLE	Bottle
E06S0043	3	PE	PURPLE	Bottle with label.
E06S0044	3	PP	PURPLE	Tap of bottle 06s0042
E06S0045	3	PET	WHITE	Bottle (deodorant stick)
E06S0046	3	PET	WHITE	Bottle with label.
E06S0047	3	PP	WHITE	Tap of bottle 06s0045
E06S0048	3	PE	BLUE	Tap of bottle (inside)
E06S0049	3	PP	LIGHT BROWN	With dirt (hand creme)
E06S0050	3	PET	RED	VTT
E06S0051	3	PET	RED	VTT
E06S0052	3	PET	DARK BLUE	VTT

EXPERIMENT 9 – VARIATION IN LIGHT SOURCE

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E09S1001	1	PP	BLUE	10.0V x 8.00A = 80.0W
E09S2001	1	PP	BLUE	8.99V x 7.50A = 77.5W
E09S3001	1	PP	BLUE	8.00V x 7.00A = 56.0W
E09S4001	1	PP	BLUE	6.90V x 6.50A = 45.5W
E09S5001	1	PP	BLUE	6.00V x 6.00A = 36.0W
E09S1002	1	PP	BLUE	10.0V x 8.00A = 80.0W. Without ink.
E09S2002	1	PP	BLUE	8.99V x 7.50A = 77.5W Without ink.
E09S3002	1	PP	BLUE	8.00V x 7.00A = 56.0W Without ink.
E09S4002	1	PP	BLUE	6.90V x 6.50A = 45.5W Without ink.
E09S5002	1	PP	BLUE	6.00V x 6.00A = 36.0W Without ink.
E09S1003	1	PP	YELLOW	10.0V x 8.00A = 80.0W
E09S2003	1	PP	YELLOW	8.99V x 7.50A = 77.5W
E09S3003	1	PP	YELLOW	8.00V x 7.00A = 56.0W
E09S4003	1	PP	YELLOW	6.90V x 6.50A = 45.5W
E09S5003	1	PP	YELLOW	6.00V x 6.00A = 36.0W
E09S1004	1	PP	YELLOW	10.0V x 8.00A = 80.0W. Without ink.
E09S2004	1	PP	YELLOW	8.99V x 7.50A = 77.5W Without ink.
E09S3004	1	PP	YELLOW	8.00V x 7.00A = 56.0W Without ink.
E09S4004	1	PP	YELLOW	6.90V x 6.50A = 45.5W Without ink.
E09S5004	1	PP	YELLOW	6.00V x 6.00A = 36.0W Without ink.

EXPERIMENT 9 – VARIATION IN LIGHT SOURCE

Sample Number	*Measurements*	*Label*	*Colour*	*Remarks*
E09S1005	1	PP	WHITE	10.0V x 8.00A = 80.0W
E09S2005	1	PP	WHITE	8.99V x 7.50A = 77.5W
E09S3005	1	PP	WHITE	8.00V x 7.00A = 56.0W
E09S4005	1	PP	WHITE	6.90V x 6.50A = 45.5W
E09S5005	1	PP	WHITE	6.00V x 6.00A = 36.0W
E09S1006	1	PP	WHITE	10.0V x 8.00A = 80.0W. Without ink.
E09S2006	1	PP	WHITE	8.99V x 7.50A = 77.5W Without ink.
E09S3006	1	PP	WHITE	8.00V x 7.00A = 56.0W Without ink.
E09S4006	1	PP	WHITE	6.90V x 6.50A = 45.5W Without ink.
E09S5006	1	PP	WHITE	6.00V x 6.00A = 36.0W Without ink.
E09S1007	1	PP	RED	10.0V x 8.00A = 80.0W
E09S2007	1	PP	RED	8.99V x 7.50A = 77.5W
E09S3007	1	PP	RED	8.00V x 7.00A = 56.0W
E09S4007	1	PP	RED	6.90V x 6.50A = 45.5W
E09S5007	1	PP	RED	6.00V x 6.00A = 36.0W
E09S1008	1	PP	RED	10.0V x 8.00A = 80.0W. Without ink.
E09S2008	1	PP	RED	8.99V x 7.50A = 77.5W Without ink.
E09S3008	1	PP	RED	8.00V x 7.00A = 56.0W Without ink.
E09S4008	1	PP	RED	6.90V x 6.50A = 45.5W Without ink.
E09S5008	1	PP	RED	6.00V x 6.00A = 36.0W Without ink.
E09S1009	1	PVC	YELLOW	10.0V x 8.00A = 80W
E09S2009	1	PVC	YELLOW	8.9V x 7.50A = 78W
E09S3009	1	PVC	YELLOW	8.0V x 7.00A = 56W
E09S4009	1	PVC	YELLOW	6.9V x 6.50A = 45W
E09S5009	1	PVC	YELLOW	6.0V x 6.00A = 36W
E09S6009	1	PVC	YELLOW	5.1V x 5.50A = 28W
E09S7009	1	PVC	YELLOW	4.3V x 5.00A = 22W
E09S8009	1	PVC	YELLOW	3.6V x 4.50A = 16W
E09S9009	1	PVC	YELLOW	2.8V x 4.00A = 11W
E09S0009	1	PVC	YELLOW	2.2V x 3.50A = 8W

Appendix D – Simulation Results

Table 13 Simulation T_p = 0%

σ	*Td*	*Tp*	*ru*	*rk*	*re*	*rc*	*rj*	*ra*
0.81	**0.09**	**0**	**16.94**	**83.06**	**0.00**	**83.06**	**0.00**	**100.00**
0.81	0.10	0	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.11	0	18.55	81.45	0.00	81.45	0.00	100.00
0.81	0.12	0	18.55	81.45	0.00	81.45	0.00	100.00
0.75	0.05	0	19.35	80.65	0.00	80.65	0.00	100.00
0.81	0.13	0	19.35	80.65	0.00	80.65	0.00	100.00
0.81	0.14	0	20.16	79.84	0.00	79.84	0.00	100.00
0.75	0.06	0	20.97	79.03	0.00	79.03	0.00	100.00
0.81	0.15	0	20.97	79.03	0.00	79.03	0.00	100.00
0.75	0.07	0	21.77	78.23	0.00	78.23	0.00	100.00
0.75	0.08	0	21.77	78.23	0.00	78.23	0.00	100.00
0.75	0.09	0	22.58	77.42	0.00	77.42	0.00	100.00
0.75	0.10	0	23.39	76.61	0.00	76.61	0.00	100.00
0.75	0.11	0	23.39	76.61	0.00	76.61	0.00	100.00
0.75	0.12	0	23.39	76.61	0.00	76.61	0.00	100.00
0.75	0.13	0	23.39	76.61	0.00	76.61	0.00	100.00
0.75	0.14	0	23.39	76.61	0.00	76.61	0.00	100.00
0.75	0.15	0	23.39	76.61	0.00	76.61	0.00	100.00
1.05	0.05	0	4.84	95.16	0.81	94.35	0.00	99.15
1.05	0.06	0	4.84	95.16	0.81	94.35	0.00	99.15
1.05	0.07	0	4.84	95.16	0.81	94.35	0.00	99.15
1.05	0.08	0	5.65	94.35	0.81	93.55	0.00	99.15
1.05	0.09	0	5.65	94.35	0.81	93.55	0.00	99.15
0.99	0.05	0	6.45	93.55	0.81	92.74	0.00	99.14
1.05	0.10	0	6.45	93.55	0.81	92.74	0.00	99.14
1.05	0.11	0	6.45	93.55	0.81	92.74	0.00	99.14
1.05	0.12	0	6.45	93.55	0.81	92.74	0.00	99.14
0.99	0.06	0	7.26	92.74	0.81	91.94	0.00	99.13
0.99	0.07	0	7.26	92.74	0.81	91.94	0.00	99.13
0.99	0.08	0	7.26	92.74	0.81	91.94	0.00	99.13
1.05	0.13	0	7.26	92.74	0.81	91.94	0.00	99.13
1.05	0.14	0	7.26	92.74	0.81	91.94	0.00	99.13
1.05	0.15	0	7.26	92.74	0.81	91.94	0.00	99.13
0.99	0.09	0	8.06	91.94	0.81	91.13	0.00	99.12
0.99	0.10	0	8.06	91.94	0.81	91.13	0.00	99.12
0.99	0.11	0	8.06	91.94	0.81	91.13	0.00	99.12
0.99	0.12	0	8.06	91.94	0.81	91.13	0.00	99.12
0.99	0.13	0	8.06	91.94	0.81	91.13	0.00	99.12
0.87	0.05	0	8.87	91.13	0.81	90.32	0.00	99.12
0.87	0.06	0	8.87	91.13	0.81	90.32	0.00	99.12
0.87	0.07	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.05	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.06	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.07	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.08	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.09	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.10	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.11	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.12	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.13	0	8.87	91.13	0.81	90.32	0.00	99.12
0.93	0.14	0	8.87	91.13	0.81	90.32	0.00	99.12
0.99	0.14	0	8.87	91.13	0.81	90.32	0.00	99.12

0.99	0.15	0	8.87	91.13	0.81	90.32	0.00	99.12
0.81	0.05	0	10.48	89.52	0.81	88.71	0.00	99.10
0.87	0.08	0	10.48	89.52	0.81	88.71	0.00	99.10
0.87	0.09	0	10.48	89.52	0.81	88.71	0.00	99.10
0.87	0.10	0	10.48	89.52	0.81	88.71	0.00	99.10
0.93	0.15	0	10.48	89.52	0.81	88.71	0.00	99.10
0.87	0.11	0	12.10	87.90	0.81	87.10	0.00	99.08
0.87	0.12	0	12.10	87.90	0.81	87.10	0.00	99.08
0.81	0.06	0	12.90	87.10	0.81	86.29	0.00	99.07
0.87	0.13	0	12.90	87.10	0.81	86.29	0.00	99.07
0.81	0.07	0	14.52	85.48	0.81	84.68	0.00	99.06
0.87	0.14	0	14.52	85.48	0.81	84.68	0.00	99.06
0.87	0.15	0	14.52	85.48	0.81	84.68	0.00	99.06
0.81	0.08	0	15.32	84.68	0.81	83.87	0.00	99.05
1.11	0.13	0	4.84	95.16	1.61	93.55	0.00	98.31
1.11	0.14	0	5.65	94.35	1.61	92.74	0.00	98.29
1.11	0.15	0	5.65	94.35	1.61	92.74	0.00	98.29
1.11	0.10	0	4.03	95.97	2.42	93.55	0.00	97.48
1.11	0.11	0	4.03	95.97	2.42	93.55	0.00	97.48
1.11	0.12	0	4.03	95.97	2.42	93.55	0.00	97.48
1.11	0.05	0	2.42	97.58	3.23	94.35	0.00	96.69
1.11	0.06	0	2.42	97.58	3.23	94.35	0.00	96.69
1.11	**0.07**	**0**	**2.42**	**97.58**	**3.23**	**94.35**	**0.00**	**96.69**
1.11	0.08	0	2.42	97.58	3.23	94.35	0.00	96.69
1.11	0.09	0	2.42	97.58	3.23	94.35	0.00	96.69
1.17	0.05	0	1.61	98.39	4.03	94.35	0.00	95.90
1.17	0.06	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.07	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.08	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.09	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.10	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.11	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.12	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.13	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.14	0	2.42	97.58	4.03	93.55	0.00	95.87
1.17	0.15	0	2.42	97.58	4.03	93.55	0.00	95.87
1.23	0.05	0	0.00	100.00	4.84	95.16	0.00	95.16
1.23	**0.06**	**0**	**0.00**	**100.00**	**4.84**	**95.16**	**0.00**	**95.16**
1.23	0.07	0	0.00	100.00	4.84	95.16	0.00	95.16
1.23	0.08	0	0.81	99.19	4.84	94.35	0.00	95.12
1.23	0.09	0	1.61	98.39	4.84	93.55	0.00	95.08
1.23	0.10	0	2.42	97.58	4.84	92.74	0.00	95.04
1.23	0.11	0	2.42	97.58	4.84	92.74	0.00	95.04
1.23	0.12	0	2.42	97.58	4.84	92.74	0.00	95.04
1.23	0.13	0	2.42	97.58	4.84	92.74	0.00	95.04
1.23	0.14	0	2.42	97.58	4.84	92.74	0.00	95.04
1.23	0.15	0	2.42	97.58	4.84	92.74	0.00	95.04
1.29	0.05	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.06	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.07	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.08	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.09	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.10	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.11	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.12	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.13	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.14	0	0.00	100.00	7.26	92.74	0.00	92.74

1.35	0.05	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.06	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.07	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.08	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.09	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.10	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.11	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.12	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.13	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.14	0	0.00	100.00	7.26	92.74	0.00	92.74
1.35	0.15	0	0.00	100.00	7.26	92.74	0.00	92.74
1.29	0.15	0	0.81	99.19	7.26	91.94	0.00	92.68

Table 14 Simulation $\sigma = 0.81$ and $T_d = 0.10$

σ	Td	Tp	ru	rk	re	rc	rj	ra
0.81	0.10	100.00	16.94	83.06	0.00	1.61	81.45	100.00
0.81	0.10	0.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	10.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	20.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	30.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	40.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	50.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	60.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	70.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	80.00	16.94	83.06	0.00	83.06	0.00	100.00
0.81	0.10	90.00	16.94	83.06	0.00	83.06	0.00	100.00

Table 15 Simulation $\sigma = 1.23$ and $T_d = 0.06$

σ	Td	Tp	ru	rk	re	rc	rj	ra
1.23	0.06	100.00	0.00	100.00	0.00	0.00	100.00	-
1.23	**0.06**	**60.00**	**0.00**	**100.00**	**0.00**	**85.48**	**14.52**	**100.00**
1.23	0.06	70.00	0.00	100.00	0.00	79.03	20.97	100.00
1.23	0.06	80.00	0.00	100.00	0.00	69.35	30.65	100.00
1.23	0.06	90.00	0.00	100.00	0.00	47.58	52.42	100.00
1.23	0.06	30.00	0.00	100.00	0.81	92.74	6.45	99.14
1.23	0.06	40.00	0.00	100.00	0.81	89.52	9.68	99.11
1.23	0.06	50.00	0.00	100.00	0.81	88.71	10.48	99.10
1.23	0.06	20.00	0.00	100.00	3.23	92.74	4.03	96.64
1.23	0.06	0.00	0.00	100.00	4.84	95.16	0.00	95.16
1.23	0.06	10.00	0.00	100.00	4.84	93.55	1.61	95.08

Table 16 Results $\sigma = 1.11$ and $T_d = 0.07$

σ	Td	Tp	ru	rk	re	rc	rj	ra
1.11	0.07	100.00	2.42	97.58	0.00	0.81	96.77	100.00
1.11	0.07	90.00	2.42	97.58	0.00	73.39	24.19	100.00
1.11	0.07	80.00	2.42	97.58	0.00	84.68	12.90	100.00
1.11	**0.07**	**70.00**	**2.42**	**97.58**	**0.00**	**87.10**	**10.48**	**100.00**
1.11	0.07	40.00	2.42	97.58	0.81	91.13	5.65	99.12
1.11	0.07	50.00	2.42	97.58	0.81	89.52	7.26	99.11
1.11	0.07	60.00	2.42	97.58	0.81	88.71	8.06	99.10
1.11	0.07	20.00	2.42	97.58	1.61	92.74	3.23	98.29
1.11	0.07	30.00	2.42	97.58	1.61	91.94	4.03	98.28
1.11	0.07	10.00	2.42	97.58	2.42	93.55	1.61	97.48
1.11	0.07	0.00	2.42	97.58	3.23	94.35	0.00	96.69