Detection of Surface Defects
on
Polymer Sheets
using
Image Processing Techniques

Derek Kiernan, B.E. (Electronic)

Submitted to the National Council for Educational Awards in fulfilment of the requirements for the degree of Masters of Engineering

Project Supervisors
Dr. Eliathamby Ambikairajah
Mr. Paul Rowan
October 1996

College Director
Dr. David Fenton
TABLE OF CONTENTS

Acknowledgments ............................................................... iv
Declaration ................................................................. v
Abstract ................................................................. vi

1.0 Introduction .......................................................... 1

2.0 Description of Defects and Illumination ...................... 4
  2.1 Description of Surface Defects .................................. 4
  2.2 Sheet Illumination and Image Capture ......................... 8

3.0 Spatial Domain Techniques for Image Processing .......... 10
  3.1 Introduction ........................................................ 10
  3.1 Optimal Thresholding ............................................. 12
  3.2 Dynamic Thresholding ............................................ 17
  3.3 Histogram Equalization .......................................... 20
  3.4 Illumination Cancellation ....................................... 24

4.0 Frequency Domain Techniques for Image Processing ........ 26
  4.1 Introduction ........................................................ 26
  4.2 Homomorphic Filtering .......................................... 26
  4.3 A Novel Homomorphic Filtering Technique using DCT ..... 30

5.0 Neural Network based Classification System ................. 33
  5.1 Introduction to Neural Networks ............................... 33
  5.2 Neural Model .................................................... 33
  5.3 Types of Activation Functions .................................. 35
5.4 Network Architectures ............................................. 37
  5.4.1 Single-Layer Feedforward Networks .......................... 37
  5.4.2 Multi-layer Feedforward Networks ........................... 38
5.5 Back Propagation Training ....................................... 39
  5.5.1 Training with Momentum ...................................... 41
  5.5.2 Adaptive Learning Rate ...................................... 41
5.6 Neural Network Defect Classifier ................................ 42
  5.6.1 Construction of Neural Network Input Vector .............. 43
  5.6.2 Training and Testing Neural Network ....................... 46

6.0 The Wavelet and Wavelet Packet Transforms ..................... 49
  6.1 Introduction .................................................. 49
  6.2 The One-Dimensional Wavelet Transform ....................... 49
  6.3 Implementation of the Wavelet Transform ..................... 51
    6.3.1 Filter Bank Implementation ............................... 52
    6.3.2 'Algorithme à trous' ....................................... 55
  6.4 Two-Dimensional Wavelet Transform ........................... 56
  6.5 Wavelet Packets ............................................ 60
    6.5.1 The Search Criterion ..................................... 61
    6.5.2 The Search Method ....................................... 62
    6.5.3 One-Dimensional Wavelet Packet Example ................ 62
    6.5.4 Two-Dimensional Wavelet Packet Example ................ 64
  6.6 Image Compression using Wavelets ............................. 67
    6.6.1 Image Compression using Wavelet Transform .............. 68
    6.6.2 Image Compression using Wavelet Packets ............... 72

7.0 Classification based on Wavelet Zero-Crossings .................. 74
  7.1 Introduction ................................................ 74
  7.2 Wavelet Zero-Crossings ....................................... 74
  7.3 Ridge Detection within One-Dimensional Signals ............. 76
  7.4 Ridge Detection within Images ................................ 79
  7.5 Classification of Defects .................................... 84
Abstract

This thesis describes the development of several image processing techniques for the detection of surface defects on extruded polymer sheets. Two types of extruded polymer sheets are considered, embossed and non-embossed, and four types of defects, lake, bank, hot cylinder and streak marks. Several image processing techniques are developed, including spatial-domain, frequency-domain, and space/frequency-domain methods.

In the case of embossed sheets, a spatial-domain technique known as the dynamic thresholding method is described. This as a localised automatic thresholding method which differentiates between the image of a sheet with defects and the image of a sheet with no defect.

The classification of defects in non-embossed sheets forms the majority of the work described in this thesis. This involves enhancement of images of non-embossed sheets with the aim of providing data to a classifier which can distinguish between the four defects mentioned above. Enhancement methods considered include spatial and frequency-domain techniques. In particular, a novel homomorphic filtering technique involving the discrete cosine transform is described, and integrated with an artificial neural network to provide a viable classification system.

In addition, space/frequency-domain techniques using the Wavelet Transform are investigated. Central to this technique is the use of a wavelet which is the second derivative of a particular smoothing function. For such a wavelet, there exists a useful relationship between the zero-crossings of the wavelet and sharp variations in the image under analysis. This recently emerged technique has been applied to defect classification with promising results.
1.0 Introduction

Image processing is a subset of a larger research area commonly referred to as machine vision. The objective of a machine vision system is to capture, understand and interpret visual information. In achieving this objective machine vision systems bring together imaging devices, computers, and sophisticated algorithms for solving problems in areas such as industrial inspection for quality control, medicine, document analysis and remote sensing. Emerging areas for the application of machine vision is visual communication for human-computer interaction, multimedia, virtual reality, and intelligent vehicle guidance[Marion, 1991] [Schalkoff, 1989]. This thesis is concerned with an application in the field of industrial inspection, namely detection and classification of defects on extruded polymer sheets.

The machine vision system can be subdivided into the following three activities:

- Obtaining a digital representation of a scene (Image Acquisition)
- Employing computational techniques to process or modify the image data (Image Processing)
- Analysing and using the results of the processing to meet some goal, e.g. guiding a robot. (Image Analysis)

Image Acquisition requires that the scene be illuminated and a representation (image) of the scene be made based on the reflected illumination from the scene. The method of illumination used by a machine vision system can determine the success or failure of a project and reduce the amount of processing required by the system [Gonzalez, 1991]. The design of the illumination depends largely on the particular application as does the method of achieving a representation of the scene. The reflected light from the scene is converted to a digital representation by a sensor, a common sensor used to achieve this representation is the Charge-Coupled Device (CCD). This sensor will
provide a two-dimensional matrix containing values (pixels) directly proportional to the reflected light from each portion of the scene.

Image Processing involves the creation of a new image matrix by altering the original representation of the scene in such a way that the features of interest are enhanced and the effects of noise and remaining features are reduced, while Image Analysis is concerned with the extraction and the symbolic description of these features so that the machine vision system can interpret the images. There are numerous algorithms already in use which can be applied to both processing and analysis of images. These algorithms can generally be grouped into two sections, spatial-domain and frequency-domain techniques [Galbiati, 1990].

Spatial-domain techniques involve direct computation on the image using either convolution of the image with a filtering kernel, or mapping of the individual image values via a predefined function to create a new image.

Frequency domain techniques require that the original image is transformed to provide a new representation of the original image. This new representation of the image contains values (transform coefficients), which are a measure of the similarity of the image to a set of basis signals.

Image processing alters the coefficients of such a transformation, usually by multiplication, to increase or decrease the images similarity to individual basis signals. An enhanced image is provided by the implementation of an inverse transformation of the altered coefficients. Image analysis in the frequency domain extracts a description of an image from the transformations coefficients. This description is then used to interprets the image. There are numerous transformations which can be applied by frequency domain techniques including the Fourier Transform, the Direct Cosine Transform and the emerging Wavelet Transform [Mallat, 1989], which provides a time (or space) - frequency description of the image.
This thesis examines the development and application of several image processing techniques for the task of detection and classification of surface defects in images of extruded polymer sheets. The techniques considered include those which operate in the spatial-domain and the frequency domain, as described above. In addition, the development of a novel technique based on the wavelet transform is proposed. As well as image processing techniques, a classification system for surface defects on extruded polymer sheets based on neural networks are investigated.
2.0 Description of Defects and Illumination

2.1 Description of Surface Defects

The extruded polymer sheets are formed by compressing material from an extruder die between two rollers. This production method can be used to form two types of extruded polymer sheets, embossed and non-embossed. Embossed sheets have a pattern imprinted onto the sheet by the rollers while non-embossed sheets are smooth. This production procedure can result in a number of surface defects, the surface defects include ‘lake’ marks, ‘bank’ marks, hot cylinder marks and colour regression (streaks).

The cause of these surface defects maybe summarised as follows, lake marks are caused by a lack of material supplied to the rollers from the extruder die, while bank marks are caused by excess material being supplied. Hot cylinder marks are caused by the sheet sticking to an over heated roller, while streaks are caused by a different colour dye remaining after earlier production. Each of these defects results in different visible characteristics.

Lake marks show different characteristics on the embossed and the non-embossed sheets. On an embossed sheet the lake mark exhibits a higher reflectance than the remainder of the sheet, while the lake mark on the non-embossed sheet results in a shallow depression. Figure 2.1 shows a typical lake mark on an embossed sheet while Figure 2.2 shows a typical lake mark on a non-embossed sheet.

Bank marks are characterised by a periodic thicken and slimming of the sheet. On embossed sheets this results in a higher reflectance where the sheet slims in comparison to the remainder of the sheet. For non-embossed sheets there results an uneven surface. Figure 2.3 shows a bank mark on a non-embossed sheet.
Hot cylinder marks on non-embossed are indicated by horse shoe shaped swallow depressions. Figure 2.4 shows a typical hot cylinder mark on a non-embossed sheet.

Streaks, colour regression, are unlike any of the previously described defects, in that for embossed sheets the reflectance difference is largely influenced by the colour of the defect. Also there is no uneven surface for non-embossed sheets. Figure 2.5 shows a typical streak defect for an embossed sheet.

Figure 2.1: Lake Mark on an Embossed Sheet
Figure 2.2: (a) Lake Mark on a Non-embossed Sheet
(b) Enhanced Image of (a)

Figure 2.3: (a) Bank Mark on a Non-embossed Sheet
(b) Enhanced Image of (a)
Figure 2.4: (a) Hot Cylinder Marks on a Non-embossed Sheet
          (b) Enchanced Image of (a)

Figure 2.5: Streak on an Embossed Sheet
2.2 Sheet Illumination and Image Capture

The illumination method used with a machine vision system can determine how much success the system can achieve as it directly affects the quality of the input image (raw image). As two different types of polymer sheets are to be examined, two different methods of illumination are used in a manner which will exploit the characteristics of each. Both systems use a camera (Pulnix TM-460) which provides a 256 gray level image with 512-by 512 pixels during image acquisition. This camera was connected to a PC via the Philips Imaging System (SBIP2) and algorithms implemented using the MATLAB Computing Environment. The block diagram shown in Figure 2.6 shows the configuration used to implement the algorithms explained in the thesis. This involves capture of the image via a solid state camera which converts the image to a video signal. The resultant video signal is captured by the Philips Imaging System and stored on the PC hard disk. This file is loaded into the MATLAB Computing Environment where it can be used as input to algorithms developed in the MATLAB Computing Environment. This configuration and explained further in Appendix 5.

The surface defects on embossed sheets exhibit's a different reflectance than the remainder of the sheet. This characteristic can be exploited efficiently if the sheet is illuminated evenly by a normal incident white light source with the camera also positioned perpendicular to the polymer sheet, i.e. same as light source. This illumination method was used in the capture of the images shown in Figures 2.1 and 2.5.

Non-embossed sheets with surface defects are mainly characterised by uneven surfaces. This different characterisation requires a different method of illumination. Figure 2.7 shows the method used, the low angle of the incident light causing shadows to be cast by the uneven surface hence making the defects more visible. This method of illumination was applied to the sheets shown in Figures 2.2, 2.3 and 2.4. As can be seen from these figures the defects contained within the raw image have poor visibility emphasising the need for enhancement of such raw images.
Figure 2.6: Block Diagram of Equipment Configuration

Figure 2.7: Non-embossed Sheet Illumination with Oblique Lighting
3.0 Spatial Domain Techniques for Image Processing

3.1 Introduction

Spatial domain techniques involve direct computation on the image to obtain a new image. These techniques can be categorised into two sections, single point operations and multiple point operations.

Single point operations involve the generation of a new image (output image), by modifying the pixel value at a single location based on a global rule applied to every pixel in the original image (input image). The process involves obtaining a pixel value of a given location from the input image, modifying it by a specified function and placing the new pixel value in the corresponding location of the output image [Marion, 1991]. This procedure is demonstrated by Equation 3.1 and Figure 3.1 [Galbiati, 1990],

\[ g(x,y) = T[f(x,y)] \]  

in Equation 3.1 \( f(x,y) \) is a two-dimensional matrix (input image), where \( x, y \) denote the spatial co-ordinates and \( f \) at any location \( (x,y) \) is proportional to the brightness (or gray level) of the image at that point. The elements of such a matrix are referred to as pixels. \( T \) is a predefined operation and \( g(x,y) \) the resulting output image.

![Fig 3.1: Single point operation](image-url)
Multiple point operations use essentially the same procedure except that the output image can be generated in one of two ways[Galbiati, 1990]. For the first method, the value of each pixel in the output image is dependent on a combination of values, from separate images, of corresponding pixel location. Image averaging is a procedure which uses this methodology to reduce the noise effect in the output image. The second method, referred to as spatial filtering, convolutes a spatial filter with the image. A spatial filter is usually a two-dimensional matrix whose elements are referred to as filter coefficients. The centre of the spatial filter is moved from pixel to pixel, starting at the top left corner and applying an operation at each location \((x,y)\) to yield \(g\) at that location. A common operation is to sum the products between the filter coefficients and the pixel values under the spatial filter at a specific location in the image. Figure 3.2 demonstrates this process for the location \((x,y)\) [Veron, 1991]. Appendix 1 gives a number of spatial filters which can be used in this manner to obtain different effects.

\[
\begin{align*}
g(x,y) &= W_1 \times f(x-1,y-1) + W_2 \times f(x-1,y) + W_3 \times f(x-1,y+1) \\
&+ W_4 \times f(x,y-1) + W_5 \times f(x,y) + W_6 \times f(x,y+1) \\
&+ W_7 \times f(x+1,y-1) + W_8 \times f(x+1,y) + W_9 \times f(x+1,y+1)
\end{align*}
\]

**Figure 3.2 : Spatial Filtering**
3.2 **Optimal Thresholding**

Thresholding is a common single point operation used to segment an image [Teuber, 1993]. In this process the pixels having a gray level at or below a given threshold value are given a zero value and all above the threshold are set to one. This is shown in Equation 3.2;

\[
g(x,y) = \begin{cases} 
1 & \text{if } f(x,y) > T \\
0 & \text{if } f(x,y) \leq T 
\end{cases} \tag{3.2}
\]

in Equation 3.2, \( T \) is the threshold value, \( f(x,y) \) is the input image and \( g(x,y) \) is the resultant output image. This method of thresholding is referred as gobal thresholding [Gonzalez, 1992], a disadvantage of this method is the location of the pixel is not taken into account only the value of the pixel. This can cause a segment to be fragmented on the application of a gobal threshold. For our application a value for \( T \) is required that will segment the defect from the remainder of the sheet in an effective manner.

Consider an image that contains only two principal regions, hence the overall density function is the sum or mixture of two unimodal densities, one for the light and one for the dark regions, Figure 3.3 shows the histogram of such an image, i.e. using the image of the lake mark shown in Figure 2.1.
If the lower mode represents background and the higher mode represents the defect, on the application of a threshold there is the possibility that pixels on the object will be classified as background and pixels on the background will be classified as object pixels [Low, 1991]. Optimal thresholding describes a technique which minimises these incorrect classified pixels, error pixels.

The most common method for calculating the optimal threshold is to use estimates for the means and variances of the two modes and use these parameters to calculate the optimal threshold. This method is not straightforward and may require some iterative procedures [Gonzalez, 1992]. Another approach, developed by the author, is to use the gradient characteristics of the image’s histogram to determine an estimate of the optimal threshold.

Assuming that the optimal threshold occurs at the intersection of the modes in the histogram, a change in polarity of the gradient can be detected at this intersection.
point and used to estimate the optimal threshold. Figure 3.4 shows the gradient polarity characteristics for a bi-modal histogram of similar nature to the histogram shown in Figure 3.3. The optimal threshold is estimated as the value where the gradient makes the transition from a negative value to a positive value.

![Figure 3.4: Bi-modal Histogram Gradient Characteristics](image)

As seen from Figure 3.3 real histograms are not so well formatted for this type of analysis, hence there is a need for pre-processing. The pre-processing involves smoothing the histogram and some form of interpolation to ensure zero values in the histogram don’t affect the result [Clarkson, 1995]. For the histogram shown in Figure 3.3 an optimal threshold of corresponding to a gray level of 177 was calculated using this method. Figure 3.5 shows the different stages of this method as applied to the histogram shown in Figure 3.3.
Fig 3.5(a) shows the histogram form Figure 3.3 after pre-processing, (b) shows the gradient for the histogram in (a).
The optimal thresholding method works well on embossed sheets where the difference between the reflectance is sufficiently large and the area of the defect is also sufficiently large. Figure 3.6 shows the resulting binary image after applying the optimal thresholding procedure to the image shown in Figure 2.1.

Figure 3.6: Lake Mark Segmented using Optimal Threshold Procedure

There are cases when the optimal thresholding method will not segment the defect from the remainder of the sheet. This occurs when the histogram mode indicating the defect is swamped by the mode created by background, causing the histogram to appear uni-modal and alter the gradient characteristics indicated in Figure 3.4. Two likely causes for this occurrence, when the defect area is small relative to the area of the image covered by the background and when two colours with similar reflectance components cause a streak defect. Figure 3.7 shows the histogram of the image shown in Figure 2.5 where the later case occurs. To overcome these problems a procedure referred to as dynamic thresholding was developed by the author based on similar procedures explained in [Gonzalez, 1992] and [Schalkoff, 1989].
3.3 Dynamic Thresholding

This section describes a procedure developed to segment a defect in cases where the optimal thresholding procedure fails. This process consists of dividing the image into a number of sub-images, i.e. divide the image into 64 sections, using an 8 by 8 grid, each section is 64 by 64 pixels. The histogram of each section is examined to determine it’s type, i.e. is it unimodal or bi-modal. If the histogram is bi-modal the optimal threshold for the section is calculated using the procedure described in Section 3.2. These thresholds are then used to calculate thresholds for the remaining sections of the image using the Equations 3.3 and 3.4. These formulas were developed during the course of the research by the author.

\[
T = \frac{\sum_{i=1}^{N} (10 - \text{dis}(i)) \text{thres}(i)}{\sum_{i=1}^{N} (10 - \text{dis}(i))} \tag{3.3}
\]
\[ \Delta T = \frac{1}{N} \sum_{i=1}^{N} \frac{\text{dis}(i) (T - M)}{10} \]  

(3.4)

where \( N \) is the number of sections with a bi-modal histogram, \( \text{dis}(i) \) is the distance of the current section to the threshold \( i \), \( \text{thres}(i) \) is the value of the thresholds, \( M \) is the mean of the current section, \( T \) is the Base Threshold and \( \Delta T \) is the Threshold Enhancer.

The threshold for the remaining sections is got by summing \( T \) and \( \Delta T \). The closer the section is to a calculated optimal threshold the more that threshold influences \( T \) i.e. if the nearest threshold is large, \( T \) will be larger than if the nearest threshold was small. \( \Delta T \) is based on the distance form the thresholds, \( T \) and the mean (\( M \)) of the section. The further away from the thresholds the larger the magnitude of \( \Delta T \). The sign of \( \Delta T \) depends on \( T \) and the mean of the section, these also influence the magnitude of \( \Delta T \).

Consider a dark object on a bright background. The thresholds will be less than the means of sections containing mainly background and greater than the means for sections consisting mainly of object. Hence if a section of the image is at a relative large distance from the thresholds and it’s mean is greater than \( T \) it is likely that most or all of this section is background and \( \Delta T \) will be negative and large. Applying a threshold of \( T + \Delta T \) to this section increases the chances of the pixels been classified as background. If the distance from the thresholds is decreased the magnitude of \( \Delta T \) also decreases and increasing the probability that pixels will be classified as part of the object.

Figure 3.8 shows results of applying this method to the streak defect shown in the Figure 2.5 and the lake defect originally shown in Figure 2.1. This shows the ability of this procedure to segment an image when the defect mode of the histogram is swamped, as is the case for the streak defect, and when the histogram is clearly bi-modal, as is the case for the lake defect.
Figure 3.8: (a) Dynamic Thresholding applied to Streak defect
(b) Dynamic Thresholding applied to Lake defect
### 3.4 Histogram Equalisation

The technique of maximising contrast is known as histogram equalisation. Here, we perform a global operation, with a uniform level distribution as the intended result [Gonzalez, 1992]. Let's say the transformation used is,

\[ s = T(r) \quad 0 \leq r \leq 1 \quad (3.5) \]

where \( r \) is the input level and \( s \) is the resultant output level. The value of \( s \) will also be in the range 0 to 1. In both cases 0 represents black and 1 white and intermediate values different shades of grey. The inverse transformation from \( s \) to \( r \) is denoted by,

\[ r = T^{-1}(s) \quad (3.6) \]

Firstly assume the variables \( r \) and \( s \) are continuous. These can be characterised by their probability density functions, \( p_r(r) \) and \( p_s(s) \). We can state that,

\[ p_s(s) = \left[ p_r(r) \frac{dr}{ds} \right]_{r = T^{-1}(s)} \quad (3.7) \]

Consider transformation function given by Equation 3.8,

\[ s = T(r) = \int_{0}^{r} p_s(w)dw \quad (3.8) \]

where \( w \) is a dummy variable. From Equation 3.8 we can state that,

\[ \frac{ds}{dr} = p_s(r) \quad (3.9) \]

Hence substituting equation 3.9 into equation 3.7 we obtain

\[ p_s(s) = \left[ p_r(r)/p_s(r) \right] = 1 \]
which is a uniformly distributed function. Consider applying the transformation of given by Equation 3.8 to the probability density function given in Figure 3.9 and Equation 3.10 [Gonzalez, 1992]

\[
\text{Pr}(r) = \begin{cases} 
-2r + 2 & 0 \leq r \leq 1 \\
0 & \text{elsewhere}
\end{cases}
\]  
(3.10)

\[
s = T(r) = \int_0^r (-2w + 2) \, dw
\]

\[
\Rightarrow s = r^2 + 2r
\]

From this we can obtain the inverse transformation function

\[
r = T^{-1}(r) = 1 - \sqrt{1-s}
\]

\[
ps(s) = (-2r + 2) \frac{dr}{ds}
\]

\[
= \int 2\sqrt{1-s} \, d\left[1 - \sqrt{1-s}\right] / ds
\]

\[
= 1 \quad 0 \leq s \leq 1
\]
In this example the resulting probability density function is uniform. Figure 3.10 shows the transformation function $T(r)$ and the probability density function $p_s(s)$.

\[
\begin{align*}
\text{(a)} & \quad s = T(r) \\
\text{(b)} & \quad p_s(s)
\end{align*}
\]

**Fig 3.10**: (a) shows the transformation function $T(r)$
(b) shows the probability density function $p_s(s)$

In practice the probability density functions are replaced with histograms, which are not continuous but discrete, hence we replace the integration in Equation 3.8 with a summation to obtain the discrete transformation equation to achieve histogram equalisation [Jain, 1989].

\[
s = T(r) = \sum_{r < s} p_r(r) \quad 0 \leq r \leq 1
\]

Figure 3.11 shows an example of histogram equalisation for
A flat resultant histogram should not be expected from the histogram equalisation procedure. This can be contributed to the discontinuity of the transformation function [Gonzalez, 1992], $T(r)$, as demonstrated in Figure 3.11. Figure 3.12 shows the result of applying the histogram equalisation method to a non-embossed sheet with a bank mark, the image shown in Figure 2.3. The defect is enhanced but so is the uneven illumination caused by the oblique illumination of the sheet. This enhancement of the uneven illumination affects the detection methods applied, hence requiring some method of illumination cancellation prior to histogram equalisation.
Illumination Cancellation

The image, \( f(x,y) \), is formed as a direct result of light incident on some form of sensor. Light is a form of energy \( f(x,y) \) must be non-zero and finite. The basic nature of \( f(x,y) \) may be characterised by two components: (1) the amount of source light incident on the scene being viewed called illumination and denoted by \( i(x,y) \). (2) the amount of light reflected by the objects in the scene, called reflectance and denoted by \( r(x,y) \). These functions \( i(x,y) \) and \( r(x,y) \) combine as a product to form \( f(x,y) \), as displayed in Equation 3.11 [Gonzalez, 1992].

\[
    f(x,y) = i(x,y) r(x,y) \quad (3.11)
\]

The function \( i(x,y) \) ranges from zero to infinity, while the function \( r(x,y) \) ranges from zero to one. The parameter \( i(x,y) \) is determined by the light source used to illuminate the scene and \( r(x,y) \) is determined by the characteristics of the objects in the viewed scene. For a reflectance of 0 all light is absorbed and the object appears black, while a reflectance of 1 indicates that all the light incident on the object is reflected.

Uneven illumination, \( i(x,y) \), associated with the image capture of non-embossed letters is slowly varying, while the surface defects on the sheet cause sharp variations, \( r(x,y) \). Hence an approximation of \( i(x,y) \) can be obtained by filtering the image of a test, \( f(x,y) \), with a low-pass spatial filter [Marion, 1991]. Any of the low-pass filters in Appendix 1 can be used for this purpose. The procedure here repeatedly filters the image with a 3-by-3 averaging filter, for all examples demonstrated in this section filtering occurs four times, to obtain an illumination approximation. The reflectance component is then obtained by dividing the original image, \( f(x,y) \), by the passed filtered image, i.e. the illumination approximation \( i(x,y) \) [Jähne, 1991]. Figure 3.13 shows the result of this process applied to the image after contrasting shown in Figure 2.3.
The image shown in Figure 3.13 has a uni-modal histogram both globally and locally, hence the procedures described in Sections 3.2 and 3.3 would not be applicable here. As can be seen from this image the noise contained within the original image is also enhanced which would affect any detection method applied. To combat this enhancement of noise we turn to the frequency domain where more selectivity over the range of frequencies enhanced is achievable. Figure 3.14 shows the illumination cancellation method applied to the raw images in Figures 2.2 and 2.4.

Figure 3.13: Spatial Enhancement of Bank Mark

Figure 3.14: (a) Enhancement of Lake Mark
(b) Enhancement of Hot Cylinder Marks
4.0 Frequency Domain Techniques for Image Processing

4.1 Introduction

The foundation of frequency domain techniques is the convolution theorem, where the convolution process proves the basic mechanics for spatial filtering [Gonzalez, 1992]. This theorem states that convolution in the spatial domain is equivalent to multiplication in the frequency domain.

Figure 4.1 shows the general process implemented by frequency domain techniques. The input image \( f(x, y) \) is transformed to the frequency domain to form a frequency representation \( F(u, v) \), where \( u \) is the frequency variable in the x-direction and \( v \) is the frequency variable in the y-direction. The Fourier transformation is commonly used for this purpose. The frequency representation \( F(u, v) \) is then multiplied by a two-dimensional matrix, \( H(u, v) \), known as the transfer function or filter matrix. This yields the frequency representation \( G(u, v) \), on which applying the inverse transformation results in the desired image, \( g(x, y) \) [Jain, 1989].

![Figure 4.1: Frequency Domain Technique](image)

### Homomorphic Filtering

Homomorphic filtering is a process which improves the appearance of an image by simultaneous brightness range compression and contrast enhancement [Gonzalez, 1992].
From the discussion in Section 3.5, an image can be expressed in terms of its illumination and reflectance components, i.e.

\[
f(x,y) = i(x,y)r(x,y)
\]  \hspace{1cm} (4.1)

We cannot use Equation 4.1 directly in order to operate separately on the frequency components of illumination and reflectance because the Fourier transform of the product of two functions is not separable [Gonzalez, 1992], i.e.

\[
\mathcal{F}\{f(x,y)\} \neq \mathcal{F}\{i(x,y)\} \mathcal{F}\{r(x,y)\}
\]

Suppose, however, that we define \( z(x,y) \) as the natural logarithm of the input image \( f(x,y) \) then:

\[
z(x,y) = \ln[f(x,y)] = \ln[i(x,y)] + \ln[r(x,y)]
\]

Then,

\[
\mathcal{F}[z(x,y)] = \mathcal{F}[\ln[f(x,y)]] = \mathcal{F}[\ln[i(x,y)]] + \mathcal{F}[\ln[r(x,y)]]
\]

or

\[
Z(u,v) = I(u,v) + R(u,v)
\]

where \( I(u,v) \) and \( R(u,v) \) are the Fourier transforms of \( \ln[i(x,y)] \) and \( \ln[r(x,y)] \), respectively.

If we process \( Z(u,v) \) by means of a filter matrix \( H(u,v) \) then

\[
S(u,v) = H(u,v)Z(u,v) = H(u,v)I(u,v) + H(u,v)R(u,v)
\]
Then
\[
\begin{align*}
  s(x,y) &= \mathcal{F}^{-1}\{S(u,v)\} \\
         &= \mathcal{F}^{-1}\{H(u,v)I(u,v)\} + \mathcal{F}^{-1}\{H(u,v)R(u,v)\}
\end{align*}
\]
By letting
\[
  i'(x,y) = \mathcal{F}^{-1}\{H(u,v)I(u,v)\}
\]
and
\[
  r'(x,y) = \mathcal{F}^{-1}\{H(u,v)R(u,v)\}
\]
Then
\[
  s(x,y) = i'(x,y) + r'(x,y)
\]

As \(z(x,y)\) was formed by taking the logarithm of the original image \(f(x,y)\), the inverse operation yields the desired enhanced image \(g(x,y)\), i.e.
\[
  g(x,y) = \exp[s(x,y)]
\]
\[
  = \exp[i'(x,y)] \cdot \exp[r'(x,y)]
\]
\[
  = i_0(x,y) r_0(x,y)
\]

where \(i_0(x,y)\) and \(r_0(x,y)\) are the illumination and reflectance components of the output image. Figure 4.2 shows the process of homomorphic filtering [Gonzalez, 1992].

![Fig 4.2 : Showing the Process of Homomorphic Filtering](image)

The key to this approach is the separation of the illumination and reflectance components. Then the homomorphic filter matrix, \(H(u,v)\), can operate separately on each of these components.
The illumination component of an image is generally characterised by slow spatial variations, while the reflectance component tends to vary rapidly. These characteristics lead to associating the low frequencies of the logarithm of an image with illumination and the high frequencies with reflectance.

A good deal of control can be gained over the illumination and reflectance components with a homomorphic filter. This control requires specification of the filter matrix $H(u,v)$ that affects the low and high frequencies in different ways. Figure 4.3,[Gonzalez, 1992] shows a cross section of such a filter matrix. A complete specification of $H(u,v)$ is obtained by rotating the cross section $360^\circ$ about the vertical axis. If the parameters $\alpha_l$ and $\alpha_h$ are chosen so that $\alpha_l < 1$ and $\alpha_h > 1$ this process tends to decrease the low frequencies and amplify the high frequencies, resulting in simultaneous dynamic range compression and contrast enhancement [Banks, 1990].

![Fig 4.3: Showing the cross section of a homomorphic filter matrix](image)

Figure 4.4 shows the homomorphic filtering process being applied to the bank image shown in Figure 3.2, $\alpha_l = 0.5$ and $\alpha_h = 2.0$. Here we can see the enhancement of the defect but also the enhancement of the noise.