Gottfried Wilhelm Leibniz Universität
Hannover, Germany

Studienarbeit

Statistical Monitoring of Image Data

Xilin Huang

Mentor: M. Sc.
Examiner: Prof. Dr.
Examiner: Prof. Dr.

September 13, 2021

protocol of statistical monitoring of image data

The purpose of this research is to propose a feasible method to monitor the change image data with the help of statistical method and control chart.

The goal of this thesis is to find out a statik to accuratlly represent to state of image data.

Methodology

Algorithm: statistical based defect detection , e.g., maximun variance of sliding windows, Discrete wavelet transform,etc...
Framework: control chart e.g. Shewhart control chart, χ^2 control chart, Hotelling T^2 control chart.
Programming language: Matlab.

Tasks with Tentative Schedule

- The first and the second months: literature review and data collect
- The second to the forth months: empirical study and hypotheses propose
- In the third month: mid-term reflection
- The forth and the fifth month: test and modify the proposed hypotheses
- The sixth month: writing and presenting thesis
- The end: evaluation of performance (including thesis and presentation)

Statutory Declaration

I, Xilin Huang, declare that this Studienarbeit, and the work indicated herein have been composed by myself, and any sources have not been used other than those specified. All the consulted published or unpublished work of others have been clearly cited. I additionally declare that the work and student's thesis have not been submitted for any other previous degree examinations.

Xilin Huang Hannover, September 13, 2021

Eidesstattliche Erklärung

Ich, Xilin Huang, erkläre hiermit, die Studienarbeit selbstständig verfasst zu haben und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt zu haben. Alle Stellen der Arbeit, die wörtlich oder sinngemäß aus anderen Quellen übernommen wurden, habe ich als solche kenntlich gemacht. Die Arbeit wurde in gleicher oder ähnlicher Form noch keiner Prüfungsbehörde vorgelegt.

Xilin Huang Hannover, September 13, 2021

Acknowledgements

Here you can write your acknowledgements.

Abstract

Here write your abstract for your thesis. Please keep it accurate and clear within one page.

Contents

List of Figures

1 Introduction

In recent decades, statistical process control (SPC) charts are widely used in manufacturing processes to detect assignable causes of variability and keep the process in control. In manufacturing processes, quality characteristics like geometry can be measured by hand and controlled by traditional control charts. However, some sensory quality characteristics like surface appearance are often monitored by machine vision. Machine vision inspection is recently widespread used in manufacturing processes along with the wide application of image processing and sensor technology. Also, it can provide much relevant information in the process of production, such as surface defectives. Thus, machine vision system (MVS) is suitable for inspecting the products whose quality characteristics can be retrieved from images. In this study, the control charts are proposed to combine with machine vision system to monitor the surface quality of mobile phone cover.

In the manufacturing process of mobile phone accessories, quality inspection is an important part for ensuring product quality. Some traditional manufacturers still perform human inspections on products after production and processing. Despite some similarities between human and machine vision, there are significant differences between them. As Zuech (2000) stated, current machine-vision systems are primitive compared with the human eye-brain capability because current MVSs are susceptible to variations in lighting conditions, reflection, and minor changes in texture, among other variations, to which the human eye can easily adjust and compensate. In some cases, the use of MVSs is cheaper than the use of human inspectors, and it is expected that the cost of MVSs will continue to decrease over time [Megahed et al., 2011]. Thus, it is necessary to develop machine vision based verification techniques to improve the efficiency of faults detection during product quality verification processes.

The aim of this study is to proposed a framework which combines SPC and MVS to achieve high efficiency and high accuracy defects detection on mobile phone cover. The main procedures of this study is organized as follow:

- samples collect
- relative paper review
- methods propose
- experiments conduct and compare
- result analyze and improve

The proposed framework combines feature extract methods and statistical process control monitoring techniques to gradually infer a high-level statistic value in the cover of mobile phone image from the low-level representation of the scratches and stains and monitor the image base on statistic value. Firstly, the mobile phone cover surface texture properties such as scratch and stain are decomposed into so-called statistical characteristics by mean of sliding-window method introduced at section 4.1 and the wavelet transform introduced at section 4.3. Then statistical approach, i.e., Shewhart control chart and Hotelling T^2 control chart, are utilized respectively to monitor mean statistic value and the mean statistic vector of a univariate and multivariate process, which can be used to judge the existence of scratch defects in the sample image. The performance comparison of two combinations will be illustrated in chapter 5.

The remainder of this paper is organized as follows. The related work of this paper will be introduced in chapter 2. Chapter 3 introduce the background knowledge of image, wavelet decomposition and control chart. Chapter 4 explain the principle and algorithms of the proposed method in detail. Chapter 5 provides an experiment to apply the proposed methods in an industrial environment to evaluates the performance of the proposed methods. Finally, the conclusions and directions of future research are presented.

2 Related Work

3 Background

3.1 Discrete wavelet decomposition

The Discrete wavelet transform (DWT) foundations go back to 1976 when Croiser, Esteban, and Galand devised a technique to decompose discrete-time signals . In 1983, Burt defined a very similar technique to subband coding and named it pyramidal coding, also known as multiresolution analysis. Later in 1989, Vetterli and Le Gall made some improvements to the subband coding scheme, removing the existing redundancy in the pyramidal coding scheme.

DWT is any wavelet transform for which the wavelets are discretely sampled. As with other wavelet transforms, a key advantage over Fourier transforms (FT) is temporal resolution: it captures both frequency and location information (location in time), Fig. 3.1 show that a Sinusoidal signal (S) is composed of four frequency components at 30 Hz, 20 Hz, 10 Hz and 5 Hz. Fig. 3.2 and Fig. 3.3 show the transformed signal of Fourier transform and Continuous Wavelet Transform, respectively.

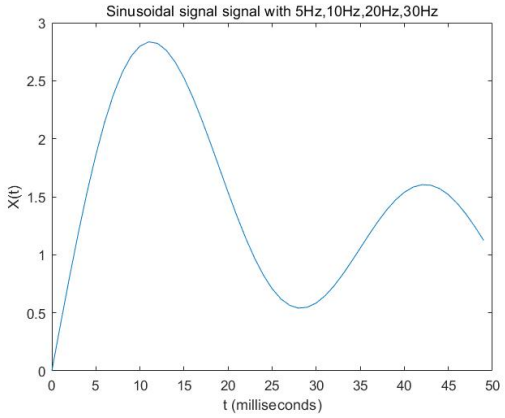

Figure 3.1: Sinusoidal signal composed of four frequency components at 30 Hz, 20 Hz, 10 Hz and 5 Hz

In Fourier transforms, we cannot know what spectral component exists at any given time instant. The best we can do is to investigate what spectral components exist at any given interval of time. This is a problem of resolution, and it is the main reason we use wavelet transform (WT) instead of FT. FT gives a fixed resolution at all times, whereas WT gives a variable resolution as follows:

higher frequencies are better resolved in time, and lower frequencies are better resolved in frequency. This means that a particular high-frequency component can be located better in the time domain (with less relative error) than a low-frequency component. On

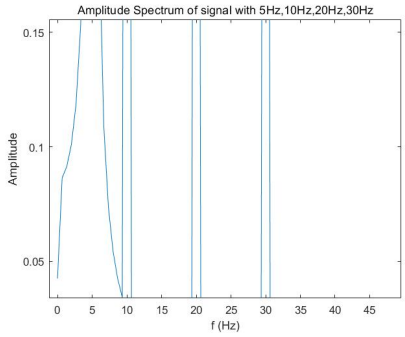

Figure 3.2: Sinusoidal signal after Fourier Transform

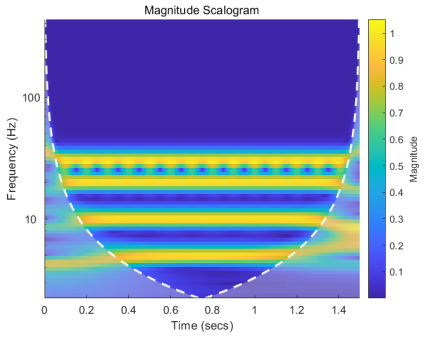

Figure 3.3: Sinusoidal signal after Continuous Wavelet Transform

the contrary, a low-frequency component can be better in the frequency domain than a high-frequency component [Polikar et al., 1996].

We can visualize the resolution problem at figure 3.4- 3.5:

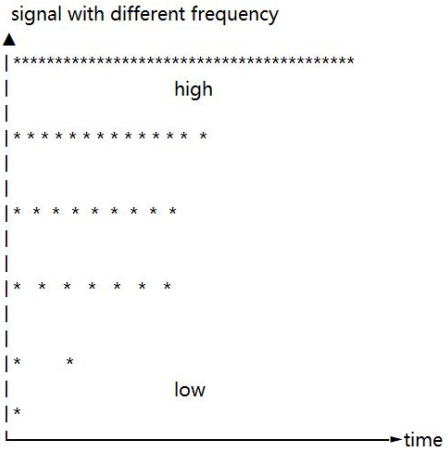

Figure 3.4: Illustration of signals with different frequency in time domain

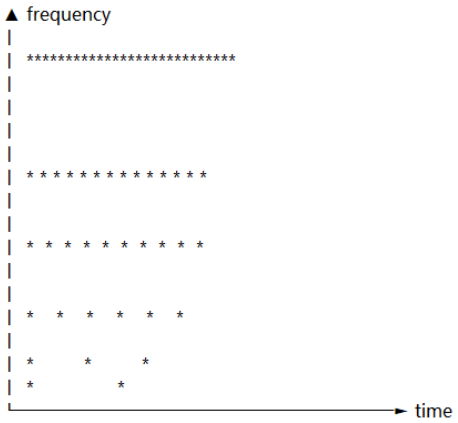

Figure 3.5: Illustration of different frequency components resolved in frequency domain

in Figure 3.4 the top row shows that at higher frequencies, we have more samples corresponding to smaller intervals of time. In other words, higher frequencies can be resolved better in time, which mean we can easily find out higher frequency signal in the time domain. As the frequency decrease, there are fewer points to characterize the signal. Therefore, low frequencies have poor resolution in time domain.

In Figure 3.5, the time resolution of the signal works the same as shown in figure 3.4, but now, the frequency information has different resolutions at every stage too. We can note that the space between subsequent frequency components increases as frequency increases. Thus lower frequencies are better resolved in frequency, whereas higher frequencies are not.

Multiresolution analysis (MRA), also known as Discrete Wavelet Transform, is a framework proposed to analyzes the signal at different frequencies with different resolutions. MRA is designed to give reasonable time resolution at high frequencies and good frequency resolution at low frequencies. This approach makes sense, especially when the signal at hand has high-frequency components for short durations and low-frequency components for long durations. Fortunately, most signals that are encountered in practical applications are often of this type.

The DWT procedure starts with passing the discrete signal $x[n]$ through a half band digital lowpass filter with impulse response $h[n]$. Filtering a signal corresponds to the mathematical operation of convolution of the signal with the impulse response of the filter. The convolution operation in discrete time is defined as follows:

$$y[n] = x[n] * h[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n-k] \,. \tag{3.1}$$

After passing the signal through a half-band lowpass filter, signals with frequencies larger than f/2 (f is the maximum frequency of the filtered signal) have disappeared. Since half the signal frequencies have now been removed, half the samples can be discarded according to Nyquist's rule [Shannon, 1949], that states if a function $x(t)$ contains no frequencies higher than B hertz, for a given sample rate f_s, perfect reconstruction is guaranteed possible for a bandlimit $f_s/2 > B$, which can be achieved by subsampling the signal by two, and the signal will then have half the number of points. The resolution of this signal, which is a measure of the amount of detailed information of signal, is halved by the lowpass filter operation. While the scale, which is the inverse of frequency, of the signal is now doubled.

In order to better illustrate the parameter scale, we can imagine it as the scale used in maps: high scales in case of map correspond to a non-detailed global view (of the signal), and low scales correspond to a detailed local view. Similarly, in terms of frequency, low frequencies (high scales) correspond to global information of a signal (that usually spans the entire signal), whereas high frequencies (low scales) correspond to detailed information of a hidden pattern in the signal.

One thing worth paying particular attention to is that the subsampling operation after filtering does not affect the resolution since removing half of the spectral components from the signal (in which the number of samples stays unchanged) makes half the number of samples redundant. Thus, half the samples can be discarded without any loss of information. In summary, the lowpass filtering halves the resolution but leaves the scale unchanged. The signal is then subsampled by two since half of the number of samples is redundant, which doubled the scale.

This procedure can mathematically be expressed as

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[2n-k]\,. \tag{3.2}$$

For the principle of decomposes the signal with DWT, we can draw the following conclusions: The DWT analyzes the signal at different frequency bands with different resolutions by decomposing the signal, which is achieved by filtering the time domain signal successively with highpass and lowpass. The signal filtered by highpass and lowpass constitutes one level of decomposition and can mathematically be expressed as follows:

$$\begin{aligned} y_{\text{low}}\,[n] &= \sum_{k=-\infty}^{\infty} x[k]g[2n-k]\,, \\ y_{\text{high}}\,[n] &= \sum_{k=-\infty}^{\infty} x[k]h[2n-k]\,, \end{aligned} \tag{3.3}$$

where $y_{\text{high}}\,[n]$ and $y_{\text{low}}\,[n]$ are the outputs of the highpass and lowpass filters after subsampling by two, respectively, which correspond to detailed information and coarse approximation.

After decomposition, the time resolution of the signal is halved, while the frequency resolution is doubled, this is because only half of the number of samples can represent the entire signal, and the frequency band of the signal only spans half of the previous frequency band. The above process is also called subband coding and can be repeated for further decomposition. At each level, filtering and subsampling will cause the number of samples to be halved (thus, the time resolution is halved), and the spanned frequency band is halved (thus, the frequency resolution is doubled). This procedure is shown in Figure 3.6, in which $x[n]$ is the original signal, and $h[n]$ and $g[n]$ are lowpass and highpass filters, respectively. f represents the bandwidth of each level.

3.2 Control chart

Statistical process control (SPC) is a powerful collection of problem-solving tools useful in achieving process stability and improving capability through the reduction of variability. SPC is one of the greatest technological developments of the twentieth century because it is based on sound underlying principles, is easy to use, has a significant impact, and can be applied to any process [Montgomery, 2020]. Its seven major tools are

1. Histogram or stem-and-leaf plot
2. Check sheet
3. Pareto chart
4. Cause-and-effect diagram
5. Defect concentration diagram

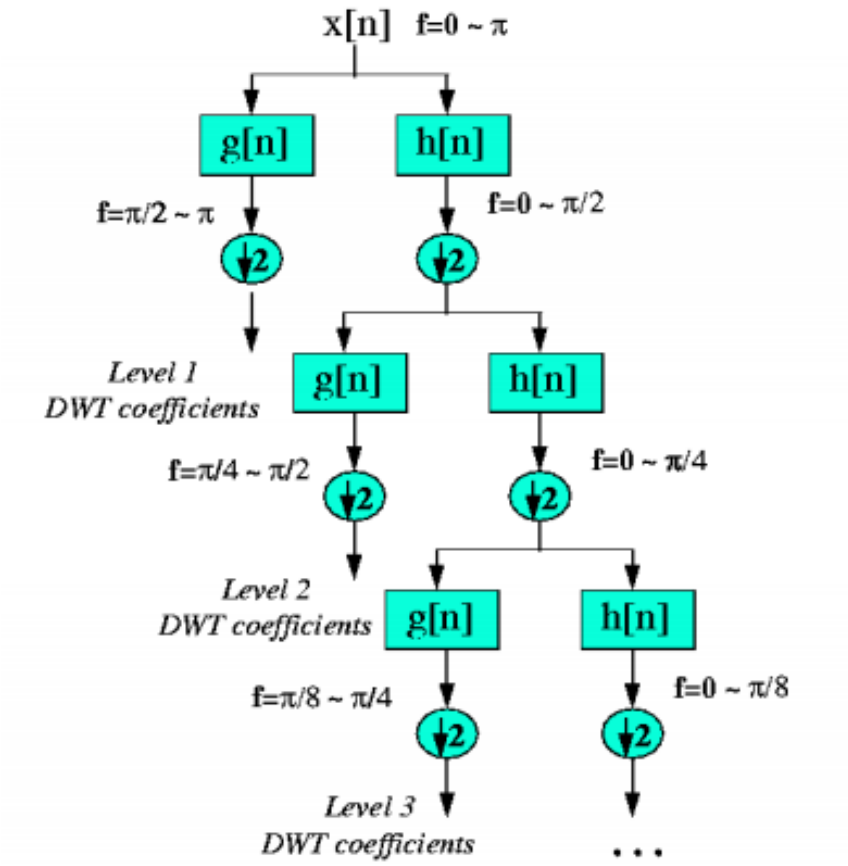

Figure 3.6: A schematic illustration of DWT decomposition. At the first beginning, the signal is decomposed into different coefficients by the highpass and lowpass filters, at each subsequent level the result of the lowpass filter is again decomposed by the highpass and lowpass filters, until the final level is decomposed.

6. Scatter diagram
7. Control chart.

The control chart is probably the easiest yet effective tool to analyze the process stability of these seven tools. To best understand the concepts of the Control chart, the classification of the control chart needs to be clarified. Depending on the number of process characteristics to be monitored, there are two primary control charts. The first, referred to as a univariate control chart, is a graphical chart of one quality characteristic, in which one is interested in monitoring changes in the parameter of an underlying univariate distribution over time. The second is a multivariate control chart, a graphical chart of a statistic that fuses more than one quality characteristic to monitor simultaneous changes in the parameter vector of an underlying multivariate distribution over time.

More specifically, the control chart is a graphical display, which plots the value of the quality characteristic that has been measured or calculated from a sample versus the sample number or versus time. Normally there are three lines in a control chart: a centerline corresponding to the mean value for the in-control process. Two other horizontal lines are called the upper control limit (UCL) and the lower control limit (LCL). These control limits are chosen so that almost all data points will fall within these limits as long as the process remains in control. A typical control chart is shown in Fig. 3.7.

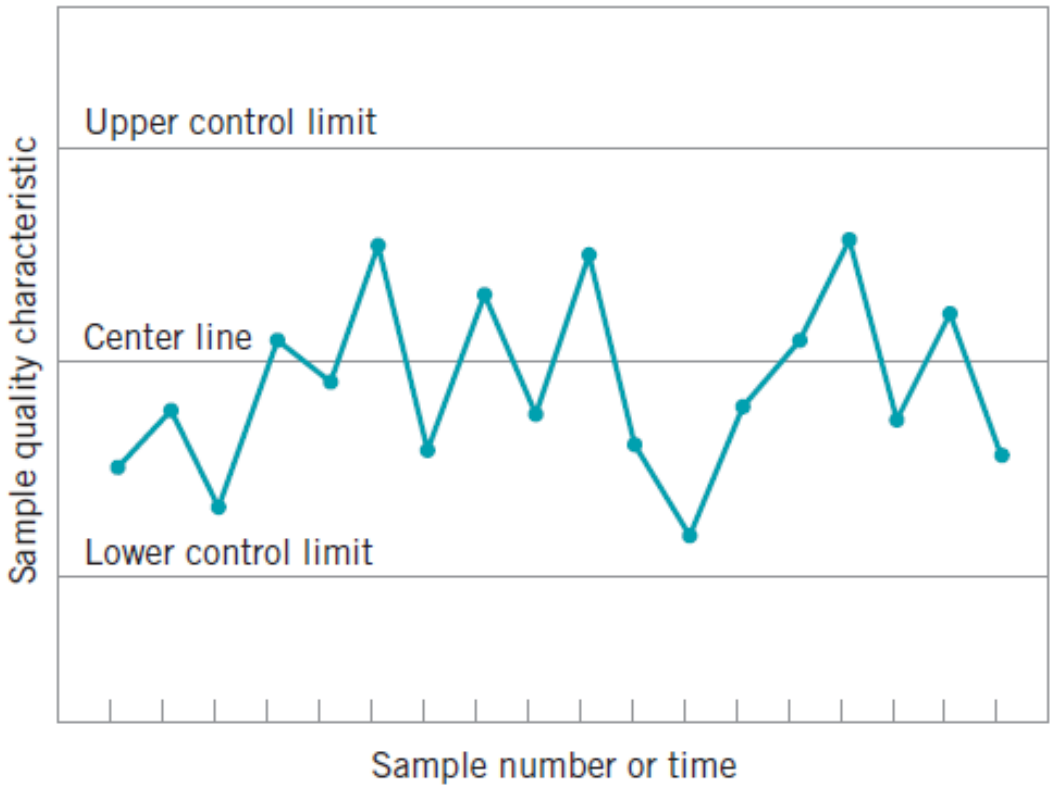

Figure 3.7: A schematic illustration of the control chart. A center line that correspond to the mean value for the in-control process. Two other horizontal lines, called the upper control limit (UCL) and the lower control limit (LCL). The scope between the two lines indicate the in-control area of process. Figure is adapted from Montgomery (2020)

3.2.1 Shewhart $\bar{X}$ control chart

In statistical quality control, the $\bar{X}$ and s chart is a type of control chart used to monitor variables data when samples are collected at regular intervals from a business or industrial process [Heckert et al., 2002].

In general, $\bar{X}$ and s control chart has following advantages:

1. The sample size (n) is relatively large ($n > 10$).

2. The sample size is variable.

The chart consists of two individual control charts: One ($\bar{X}$ control chart) to monitor the process mean and the other (s control chart) to monitor the process standard deviation.

During the 1920s, Dr. Walter A. Shewhart first proposed a general model for control charts as follows [Adams and Butler, 1999]:

Let s be a sample statistic that measures some continuously varying quality characteristic of interest (e.g., diameter), and suppose that the mean of s is μ_s, with a standard deviation of σ_s. Then the centerline (CL), the upper control limit (UCL), and the lower control limit (LCL) are

$$UCL = \mu_s + k\sigma_s \,, \tag{3.4}$$

$$CL = \mu_s \,, \tag{3.5}$$

$$UCL = \mu_s - k\sigma_s \,, \tag{3.6}$$

where k is the distance of the control limits from the centerline, expressed in standard deviation units. When k is set to 3, we speak of 3-sigma control charts. Historically, k=3 has become an accepted standard in the industry.

The centerline is the process mean, which is unknown, so as the σ. We need to replace them with some confident values. e.g., the average of all the data and the average standard deviation, respectively.

There exist two distinct phases of the control chart [Bersimis et al., 2007]:

- Phase I: charts are used for retrospectively testing whether the process was in control when the first subgroups were being drawn. In this phase, the charts are used as aids to the practitioner in bringing a process into a state where it is statistically in control.
- Phase II: control charts are used for testing whether the process remains in control when future subgroups are drawn. In this phase, the charts are used as aids to the practitioner in monitoring the process for any change from an in-control state.

In short, phase I deals with estimating process parameters to ensure process stability using historical data, and phase II pertains to signal any out-of-control condition or shifts in the process parameters.

Heckert et al (2002) introduce the X control chart in detail. It is assumed that the probability distribution of the characteristic is the normal distribution. In phase I, m

samples are required, let d_i be the parameter (s) of the ith sample. Then the average of the parameter is

$$\mu_s = \frac{1}{m}\sum_{i=1}^{m} d_i \,. \tag{3.7}$$

If σ^2 is the unknown variance of a probability distribution, then an unbiased estimator of σ^2 is the sample variance is

$$s^2 = \frac{\sum_{i=1}^{m}(d_i - \bar{x})^2}{m-1} \,. \tag{3.8}$$

However, s, the sample standard deviation, is not an unbiased estimator of σ. If the underlying distribution is normal, then s is actually

$$s = c_4 \times \sigma \,, \tag{3.9}$$

where c_4 is

$$c_4 = \sqrt{\frac{2}{m-1}}\frac{\left(\frac{m}{2}-1\right)!}{\left(\frac{m-1}{2}-1\right)!} \,, \tag{3.10}$$

in which m is the sample size.

Finally, the standard deviation of the sample standard deviation is

$$\sigma_s = \sigma\sqrt{1-c_4^2} \,. \tag{3.11}$$

3.2.2 Hotelling T^2 control chart

In fact, most data in the industry (especially in chemistry and manufacturing)are naturally multivariate. Hotelling (1947) introduced a statistic that uniquely lends itself to plotting multivariate observations. This statistic, appropriately named Hotelling's T^2, is a scalar that combines information from the mean of several variables. We can image this scalar (so-called Hotelling T^2 distance) as a statistic describing the distance away from the mean. The derivation of Hotelling T^2 distance is as follows:

In univariate statistical quality control, we generally use the normal distribution to describe a continuous quality characteristic behavior. The univariate normal probability density function is

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty \,. \tag{3.12}$$

The mean of the normal distribution is μ and the variance is σ. The term (apart from the minus sign) in the exponent of the normal distribution can be written as follows:

$$(x-\mu)\left(\sigma^2\right)^{-1}(x-\mu)\,. \tag{3.13}$$

This quantity measures the squared standardized distance from x to the mean μ, whereby the term "standardized" means that the distance is expressed in standard deviation units. This same approach can be employed in the multivariate normal distribution case.

Montgomery (2020) introduce the Hotelling T^2 control chart in detail. It is assumed that the joint probability distribution of the p quality characteristics is the p-variate normal distribution. Suppose we have p variables, given by $x_1, x_2, \ldots, x_p$. These variables are arranged in a p-component vector $\mathbf{x}' = [x_1, x_2, \ldots, x_p]$. Let $\mu' = [\mu_1, \mu_2, \ldots, \mu_p]$ be the vector of the means of the x's, and let the variances and covariances of the random variables in $\mathbf{X}$ be contained in a pp covariance matrix Σ. The main diagonal elements of Σ are the variances of the x's and the off-diagonal elements are the covariances. Now the squared standardized distance from $\mathbf{x}$ to μ is

$$(\mathbf{x}-\mu)'\Sigma^{-1}(\mathbf{x}-\mu)\,. \tag{3.14}$$

Since we extend the concept of mu and sigma from univariate parameter into multivariate parameters, then our next step is naturally to figure out how to construct sample mean vector ($\overline{\mathbf{x}}$) and sample covariance matrix ($\mathbf{S}$) so that they can achieve the same effect as Shewhart $\bar{X}$ control chart.

Suppose that we have a random sample from a multivariate normal distribution

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\,, \tag{3.15}$$

where the ith sample vector contains observations on each of the p variables $x_{i1}, x_{i2}, \ldots, x_{ip}$. Then the sample mean vector ($\overline{\mathbf{x}}$) is

$$\overline{\mathbf{x}} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{x}_i\,, \tag{3.16}$$

and the sample covariance matrix ($\mathbf{S}$) is

$$\mathbf{S} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\mathbf{x}_i-\overline{\mathbf{x}}\right)\left(\mathbf{x}_i-\overline{\mathbf{x}}\right)'\,. \tag{3.17}$$

That is, the sample variances on the main diagonal of the matrix S are computed as

$$s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)^2 \,, \tag{3.18}$$

and the sample covariances are

$$s_{jk} = \frac{1}{n-1}\sum_{i=1}^{n}(x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) \,. \tag{3.19}$$

Now suppose that $\mathbf{S}$ from equation (3.17) is used to estimate Σ and that the vector $\overline{\overline{\mathbf{X}}}$ is taken as the in-control value of the mean vector of the process. The Hotelling T^2 statistic is

$$T^2 = n(\overline{\mathbf{x}} - \overline{\overline{\mathbf{x}}})'\mathbf{S}^{-1}(\overline{\mathbf{x}} - \overline{\overline{\mathbf{x}}}) \,. \tag{3.20}$$

In which n is the subgroup size of each sample.

In some industrial settings (e.g., chemical and process industries) the subgroup size is naturally $n = 1$. Then the Hotelling T^2 statistic in equation (3.20) becomes

$$T^2 = (\overline{\mathbf{x}} - \overline{\overline{\mathbf{x}}})'\mathbf{S}^{-1}(\overline{\mathbf{x}} - \overline{\overline{\mathbf{x}}}) \,. \tag{3.21}$$

The phase II control limits for this statistic are

$$\begin{aligned} \text{UCL} &= \frac{p(m+1)(m-1)}{m^2 - mp} F_{\alpha,p,m-p} \,, \\ \text{LCL} &= 0 \,, \end{aligned} \tag{3.22}$$

in which m is the sample size, p is the number of quality characteristics, α is confident level.

4 Methodology

4.1 Maximum variance of sliding-windows

We assume that the pixels on the surface of the phone's cover are homogeneous. Therefore, an idea of using the variance of pixel values to extract features representing the state of the surface is proposed.

The control charts are used spatially by moving a mask (or window) across the image and then calculating and plotting a statistic each time the mask is moved. The size of the mask depends on the expected size of the defects to be detected, with smaller defective regions requiring smaller mask sizes [Megahed et al., 2011]. Inspired by this view, we move a ten by ten window (The size of the window is obtained by measuring the smallest size of defective region) across the image and calculate the variance of the pixel value each time the window is moved. The value with the largest variance among all windows in this image is taken as the desired statistic (maximum-variance) describing this image.

4.2 Sliding-windows based $\bar{X}$ control chart

In this section we combine Shewhart $\bar{X}$ control chart with sliding-window approach to perform statistical monitoring of image data. In section 4.1 a statistic "maximum-variance" is retrieved from a image. Thus in Phase I n standard sample images are used as Phase I data to retrieve the mean value ($\overline{m}$)

$$\overline{m} = \frac{1}{n}\sum_{i=1}^{n} m_i \,, \tag{4.1}$$

and sample standard deviation (σ_s) according to equation(3.8)-(3.11)

$$\sigma_s = \frac{s}{c_4}\sqrt{1-c_4^2}\,, \tag{4.2}$$

of the samples' statistic (maximum variance). In phase II, by selecting $k = 3$, Eqaution 3.4 become

$$UCL = \mu_m + 3\sigma_s \,, \tag{4.3}$$

and LCL according to Eqaution 3.6 is then

$$UCL = \mu_m - 3\sigma_s \,. \tag{4.4}$$

They will form a qualified range for the maximum-variance of the sample. we will monitor whether the maximum-variance of the incoming sample is within this range. If it is the case, this sample will be considered qualified. Otherwise, it is unqualified.

4.3 Haar Wavelet decomposition

As mentioned in the section 3: The DWT [Fig. 4.1] analyzes the signal at different frequency bands with different resolutions by decomposing the signal into a coarse approximation and detail information, which is associated with low-pass and high-pass filters, respectively. In our case we employ the Haar Wavelet Transform (HWT) as the wavelet basic function to perform image decomposition so that an original image is decomposed into four coefficient matrices: one low-pass filtering coefficient matrix (approximation coefficient matrix A) and three high-pass filtering coefficient matrices (detail coefficient matrices, containing the horizontal (H), vertical (V), and diagonal (D) detail coefficient matrix) at each level. HWT is one of the simplest yet powerful transformations. It is first proposed by Hungarian mathematician Alfréd Haar (1910) . The base transformation of HWT in the multiple level scaling space is

$$\begin{aligned} I_i^j &= \frac{1}{2}\left(I_{2i}^{j+1} + I_{2i+1}^{j+1}\right), \\ D_i^j &= \frac{1}{2}\left(I_{2i}^{j+1} - I_{2i+1}^{j+1}\right), \end{aligned} \tag{4.5}$$

in which j is the level of scaling space, I_i is $i-th$ value of discrete signal I. I_i^j is the average of $j+1$ level's values of two contiguous pixels, which correspond to coarse approximation of level j, while D_i^j is the difference of $j+1$ level's pixels value of two contiguous pixels corresponding to detail information of level j.

In this study the format of signal is the RGB-image of $(M \times N \times 3)$. We use haar transform to decompose the image to finest scale so that high-level statistics are retrieved to represent the characteristic of image. This procedure is also known as dimensionality reduction.

Since the analyzed images are in the form of 2-D, we need to perform the 2-D Haar wavelet transform by applying 1-D wavelet transform first on rows and then on columns. Based on the transfer concept of the one dimensional space, the Haar wavelet transform can process a two-dimensional image of $(M \times N)$ pixels each level in the following way:

$$\begin{aligned} \mathbf{L} &= \frac{1}{2} * \left(\mathbf{I}\left(:, 1:2:N\right) + \mathbf{I}\left(:, 2:2:N\right)\right), \\ \mathbf{H} &= \mathbf{I}\left(:, 2:2:N\right) - \mathbf{I}\left(:, 1:2:N\right), \end{aligned} \tag{4.6}$$

where $\mathbf{I}$ is the $M \times N$ matrix represent 2-D image, the bold letter with parentheses at $\mathbf{I}\left(a:b, c:d\right)$ represent matrix slicing and the range of slicing is from row a to b, column c to d. $\mathbf{I}\left(:, 1:2:N\right)$ extract all the odd column elements of matrix $\mathbf{I}$, $\mathbf{I}\left(:, 2:2:N\right)$ extract

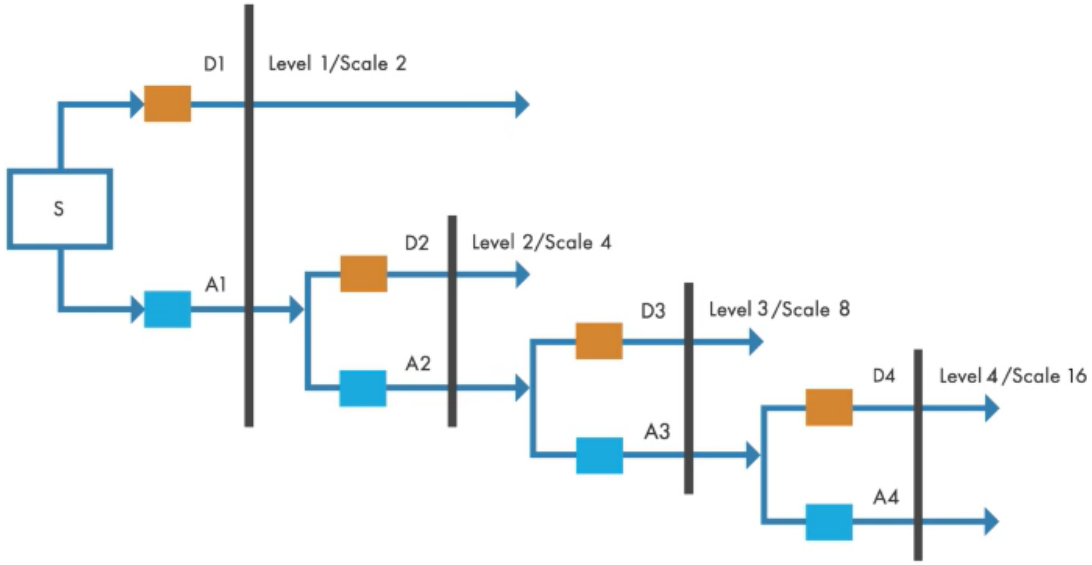

Figure 4.1: A general structure of DWT. The orange cube represent high-pass filter, the blue cube represent low-pass filter.

all the even column elements of matrix **I**. The dimension of **L** and **H** is now $M \times \frac{N}{2}$. Symbols $+$, $-$ and $*$ here are matrix addition operation, which add the corresponding elements in two matrices together, matrix subtraction operation, which subtract corresponding elements in two matrices, and matrix scalar multiplication, in which the scalar is a constant and multiply every elements in the matrix, respectively. The prerequisites of matrix $\pm$ operation are that the order of the matrices must be the same, while by scalar multiplication the matrix can be any order.

$$\begin{aligned} \mathbf{A} &= \mathbf{L}(2:2:M,:) + \mathbf{L}(1:2:M,:)\,, \\ \mathbf{V} &= \mathbf{L}(2:2:M,:) - \mathbf{L}(1:2:M,:)\,, \\ \mathbf{H} &= \frac{1}{2} * (\mathbf{H}(2:2:M,:) + \mathbf{H}(1:2:M,:))\,, \\ \mathbf{D} &= \frac{1}{2} * (\mathbf{H}(2:2:M,:) - \mathbf{H}(1:2:M,:))\,, \end{aligned} \tag{4.7}$$

$\mathbf{L}(1:2:N,:)$ and $\mathbf{H}(1:2:N,:)$ extract all the odd row elements of matrix L and H respectively. After the matrices operation, The dimention of the coefficient matrix A, V, H, and D is then $\frac{M}{2} \times \frac{N}{2}$. Coefficient matrix **A** can be considered as a smooth subimage and represents the coarse approximation of the image **I**, while coefficient matrix **H**, **V**, and **D** correspond to detail subimages and represent the horizontal, vertical and diagonal directions of the image **I** at corresponding level, respectively. The number of coefficients for matrices A H V, and D is halved each time the level increase.

The lowest level (L_l) that can be obtained by the HWT is

$$L_l = \log_2(min(r, c)) \tag{4.8}$$

If the row or column dimension of data is even, but not a power of two, the lowest level (L_l) that can be obtained by the HWT transformation is

$$L_l = \lfloor \log_2 \left(\frac{min(r, c)}{2} \right) \rfloor \tag{4.9}$$

where r, c is the row and column dimension of image, respectively, and the symbol $\lfloor E \rfloor$ rounds element E to the nearest number.

4.4 Wavelet decomposition based Hotteling T^2 control chart

In order to monitor the surface quality of the mobile phone's cover, feature statistics are needed to characterize the quality of the cover. Lin (2007) used wavelets and multivariate statistical approaches, including the Hotelling $T2$ control charts, to detect ripple and other types of defects in electronic components, particularly surface barrier layer (SBL) chips of ceramic capacitors. The specific approach first divides a grayscale image of (256×256) pixels into a set of equal-size sub-images of (4×4) pixels. The multivariate statistic (e.g. T^2) integrates the multiple wavelet characteristics (h,d,v) into a statistic value for each sub-image. This statistic value can be regarded as a distance value of the sub-image. The larger the statistic value, the larger the difference between the region and the normal area, the more that region can be judged as a defect region. Finally, a T^2 spatial distribution map of the image combined with T^2 statistic of sub-images is generated. The localization of the defect of the product can be achieved by finding the high value of the T^2 spatial distribution map.

Since our goal is to monitor whether an RGB image with three frames: Red (R), Green (G), and Blue (B) frame is qualified or not, the location of the defect is not our concern. We apply Haar wavelet transform on each frame simultaneously. The coefficients of the final level (L_l) have L times filtered by the high pass filter, which means the amplitude of coefficients contains the information of the high-frequency signal in the original image. The higher the coefficients, the more likely the signal will be abrupt. The abrupt in the signal then correspond to the defects in the monitored products.

According to Equation(4.6)-(4.9), an image can be decomposed into horizontal (H), vertical (V), diagonal (D), and approximation (A) coefficient matrix ($S \times S$) for each frame at final level (L_l), each coefficient matrix have S^2 coefficients,an image sample can have $4 \times S^2$ coefficients.

Montgomery (2020) present tables indicating the recommended number of quality characteristics p = 2, 3, 4, 5, 10, and 20. Here we use p = 3 characteristics retrieved from diagonal coefficient matrix, since the coefficients of the diagonal coefficient matrix can

best reflect the surface defects of various shapes by experiment 5.2.2. We absolute all coefficients of D matrix

$$[D] = \text{ absolute[D]} \tag{4.10}$$

which turn negative coefficients into positive coefficients without changing the value itself.

After that, we take the maximal coefficient among the diagonal matrix and consider them the desired statistical characteristics (d1,d2,d3) of the corresponding frame.

$$[\text{d1,d2,d3}] \ = \ \max[\text{ D }] \tag{4.11}$$

Further, three coefficients retrieved from three frames compose the multiple wavelet characteristic vector $\mathbf{x}$.

$$\mathbf{x} = \ [\text{d1,d2,d3}] \tag{4.12}$$

These procedures are illustrated in Figure 4.2. Finally, the Hotelling T^2 statistic

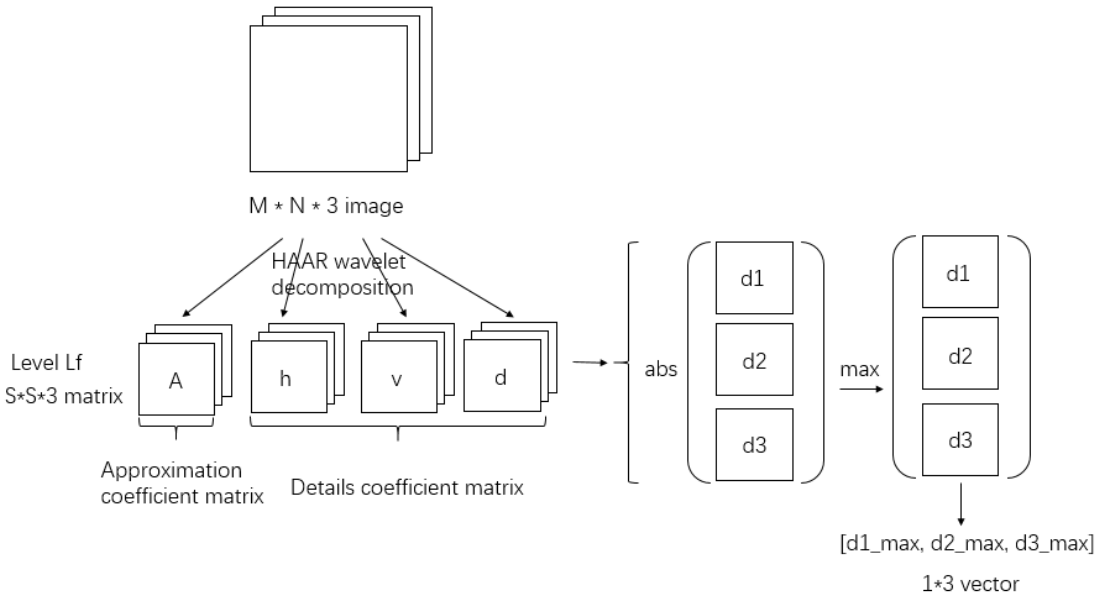

Figure 4.2: Decomposition of RGB image into 1×3 wavelet characteristic vector

$$T^2 = (\mathbf{x} - \overline{\mathbf{x}})' \mathbf{S}^{-1} (\mathbf{x} - \overline{\mathbf{x}}) \tag{4.13}$$

in which $\overline{\mathbf{x}}$ and $\mathbf{S}$ be the sample mean vector and covariance matrix, respectively, of these observations, integrates the multiple wavelet characteristics $\mathbf{x}$ into a statistic value T^2 for

each sample image. If this statistic value is larger than UCL (Equation 3.22), then we are $(1-\alpha)$ confident that this sample is out of control. Vice versa, we are $(1-\alpha)$ confident that this sample is in control.

The output of the phase I (the sample mean vector $\overline{\mathbf{x}}$ and covariance matrix $\mathbf{S}$ of standard images as well as UCL) is used as the input of phase II. The images in phase II are first decomposed by Haar wavelet transform into 3×1 characteristic vector $\mathbf{x}$. Then by using Equation 4.13 we can calculate the Hotelling T^2 statistic of this sample and compare it with UCL to judge if the sample is in control.

5 Empirical study

5.1 Dataset

In order to verify the feasibility of the proposed approach, authentic mobile phone cover images are acquired on a colour CMOS (Complementary metal–oxide–semiconductor) camera. Each coloured RGB image has 100×100 pixels in size, three-colour bands per pixel, and eight bits of intensity per colour band. The database consists of 235 mobile-phone cover images, of which 195 have no defects and 40 have various defects. All programming is done in Matlab.

A mobile-phone cover image shown in Fig. 5.1 is used for experiment. Firstly, the background, camera hole, and logo are subtracted, and then the original image is resized to several 100×100 pixels sample images, which are called standard images, shown in Fig. 5.2. The defect image samples are acquired with the same procedure as standard image samples, shown in Fig. 5.2.

Figure 5.1: Mobile-phone cover image

5.2 Experiment

5.2.1 Determine the size of the moving window

An experiment is conducted to find out the suitable size of the sliding window. Three different windows with size 5×5 pixels, 10×10 pixels, and 20×20 pixels are applied on several defect images (Fig. 5.2). After the specific size of window go through the image, the maximum variance among all the windows is considered as the desired statistic of this image by this specific size of the window. The window's size with the largest variance on

the same image will be the most suitable size of the sliding window for detecting defect in this image.

The result is shown in table 5.1. From the table 5.1, we can note that the largest maximum variance of samples a and c appear at 10×10 windows. Although the maximum value of sample b appears in the 5×5 window, consider its value is close to the second largest value, and the value itself is relatively small, we still believe 10×10 is the most appropriate size for the sliding window.

Sample	5×5 pixels	10×10 pixels	20×20 pixels
surface dent	220.96	494.57	191.27
surface scratch	71.12	48.75	45.47
surface swelling	1.17e+03	1.33e+03	9.20e+02

Table 5.1: The maximum variance of three size windows of different defect samples (see Fig. 5.2). As we can observe, the size of the sliding window is depend on the size of the defect area, thus we choose three sizes of window to compare their performance on different defects.

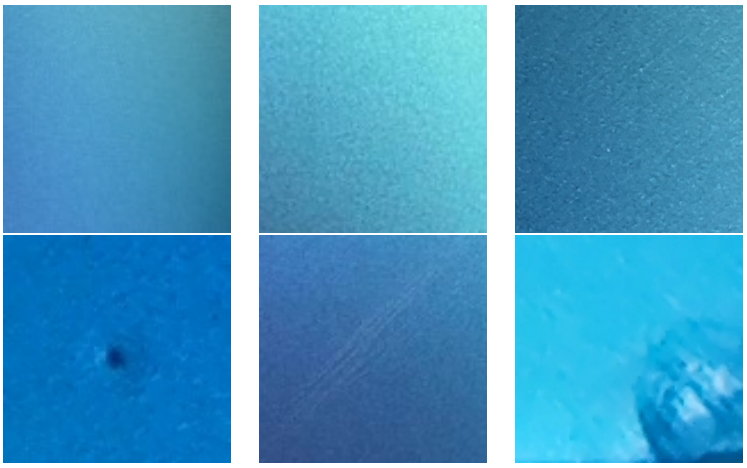

Figure 5.2: Illustration of standard sample images and defect images. The first row is the standard images, and the second row is images with different defects: From left to right are surface dent, surface scratch and surface swelling.

5.2.2 Sensitivity test of detail coefficient matrices to defects

This section conducts an experiment to test and compare the sensitivity of three detail coefficient matrices H, V, and D to various defects. The experiment will be performed in the following way: several test images with various defects shown in Fig. 5.2 are decomposed by HWT (section 4.3) to the final level. The decomposed detail matrix H, V, and D are compared to find out which matrix among the horizontal, vertical and diagonal matrices can best reflect defects in the corresponding image.

Since the decomposed detail matrices at the final level have a few coefficients and our concern is to find out the most sensitive matrix to defect, we calculate the maximum coefficient of H, V, and D by using equation (4.10) and (4.11) and compare the maximum coefficient H_max, V_max, and D_max of the same defect image. The matrix containing the maximum coefficient is considered the most suitable candidate among detail matrices.

Sample	H_max	V_max	D_max
surface dent	134.50	226.69	151.62
surface scratch	173.59	102.38	68.19
surface swelling	579.09	501.88	171.19

Table 5.2: The maximum coefficient among three detail matrices of different defect samples (see Fig. 5.2).

Before the experiment is conducted, we assume that the diagonal matrix is sensitive to the diagonal pixel value variation and the difference of horizontal and vertical direction since the diagonal matrix is the result of two repeated differences of pixel value (first by horizontal, then vertical) according to equation (4.6)-(4.7). While in Table 5.2 we can find out that the maximum coefficient of images with diagonal scratch (e.g.surface scratch) does not appear in the D matrix, which fails our assumption at the first place. Thus the following conclusion is made: Instead of using one specific detail matrix of HWT to retrieve the image characteristics, we determine to use all three matrices to characterize the sample images.

5.2.3 Performance comparison of different methods

In this section three aspects of method I (maximum variance based $\bar{X}$ control chart) and method II (wavelet based Hotelling T^2 control chart) will be compared, namely average run length (ARL), false detection rate and alarm strength. Here we use average run length (ARL)

$$ARL = \frac{1}{P}\,, \tag{5.1}$$

where P denoting the probability of an observation plotting outside the control limits to characterize the false alarm in phase I. The ARL tells us, for a given situation, how long on the average we will plot successive control charts points before we detect a point beyond the control limits [Heckert et al., 2002].

False detection rate (FDR) can be mathematically represent as

$$FDR = \frac{n_f}{n}\,, \tag{5.2}$$

in which n_f is the number of false detection in phase II while n is the total sample number in phase II, FAR is a percentage value describing the percentage of error results in the total sample. And the average alarm strength (ASS) is

$$ASS = \frac{\sum_{i=1}^{n} \frac{S_i}{UCL}}{n}, \tag{5.3}$$

where S_i is i_{th} sample's statistic, whose value is larger than UCL, to quantify the corresponding aspect of different methods. ASS is the average ratio of statistic to UCL, which is used to measure how sensitive the methods are to defects.

Jensen et al. (2006) point out that even larger sample sizes are required to ensure that the phase II average run length performance will actually be close to the anticipated values, thus in phase I, we first use 100 samples in phase I to retrieve the in-control parameters and the UCL of control chart, then apply the in-control parameters in phase II (120 samples). The result is shown in table 5.3. It turns out that the ARL in phase I and FDR in phase II are not satisfactory. Thus a calibration step of is needed.

method	phase I false alarm number	phase I ARL	phase II FDR
method I	31	3.22	9.1%
method II	8	12.50	8.3%

Table 5.3: The performance of control charts before calibration.

By calibration on the one hand we replace the false alarm samples in phase I with standard samples in phase II, on the other hand we collect another 50 standard samples in phase I to increase the sample size to improve the performance of both methods. Thus we use 150 standard sample images to retrieve the in-control state of the process by using maximum variance combining the $\bar{X}$ control chart and Haar wavelet decomposition combining Hotelling T^2 control chart, separately.

According to equation (4.1)-(4.4), the mean value ($\overline{m}$) of maximum variance of method I in phase I is 15.538, the standard deviation (σ_s) is 0.5879, thus the UCL and LCL is 17.302 and 13.775 respectively. By method II in phase I the sample mean vector of Haar wavelet decomposition characteristic is

$$\overline{\mathbf{x}} = \begin{bmatrix} 50.897 & 97.984 & 19.974 \end{bmatrix}, \tag{5.4}$$

and the sample covariance matrix is

$$\mathbf{S} = \begin{bmatrix} 448.7801 & 986.332 & -16.474 \\ 986.332 & 5251.402 & -139.947 \\ -16.4746 & -139.947 & 42.3404 \end{bmatrix}, \tag{5.5}$$

by using equation (3.22), we can calculate the UCL of method II is 8.157. The control chart of phase I of method I and method II is shown in fig 5.3 and fig 5.4. The ARL of method I and method II according to equation (5.1) is 3.125 and 30, respectively.

In phase II, 85 samples are arranged in a way that the first 40 images are negative, while the rest (45 images) are positive. The control charts of method I and method II are shown in fig 5.11 and 5.12 respectively. According to equation (5.2) and (5.3) the false detection rate of method I and method II is almost the same, their value is 5.83% and 1.18%, and the ASS of method I and method II is 27.47 and 67.12, respectively. The higher the ASS value, the larger the out-of-state amplitude of this method. The average out-of-state amplitude of method II is almost twice as that of method I, thus we can come to a conclusion that although there is almost no difference between method I and method II in terms of false detection rate in phase II, method II is more sensitive than method I in defects detection.

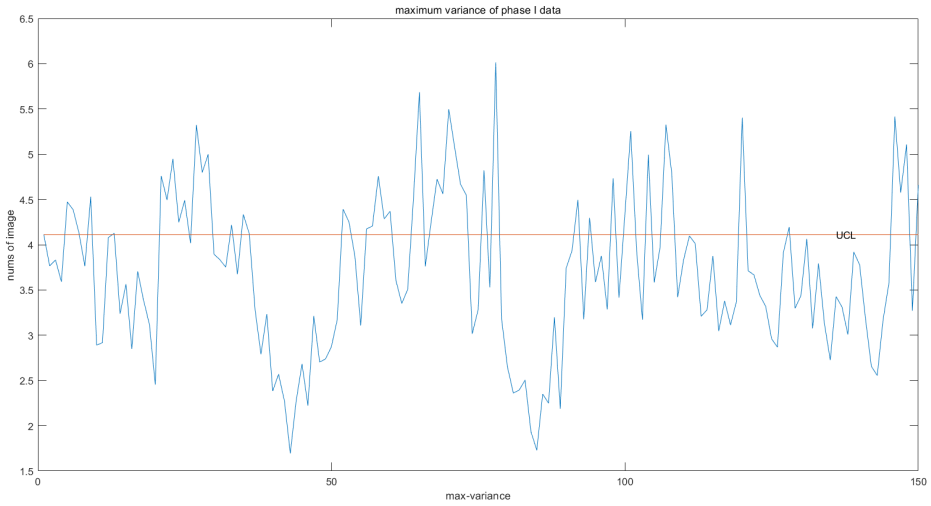

Figure 5.3: The phase I maximum variance based $\bar{X}$ control chart, where all the samples are qualified and there are 48 false alarms in the control chart. The orange line is the UCL of $\bar{X}$ control chart

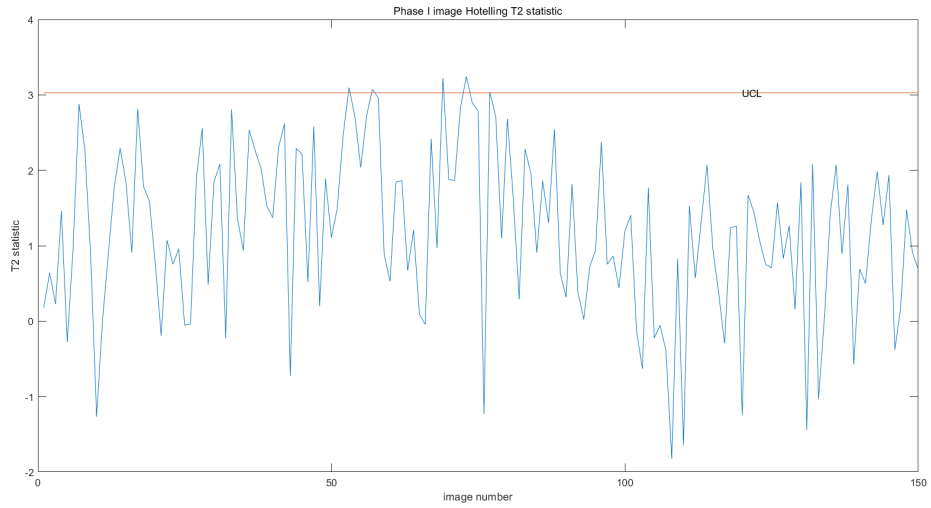

Figure 5.4: The phase I wavelet decomposition based Hotelling T^2 control chart, where all the samples are qualified and there are 5 false alarms in the control chart. The orange line is the UCL of Hotelling T^2 control chart

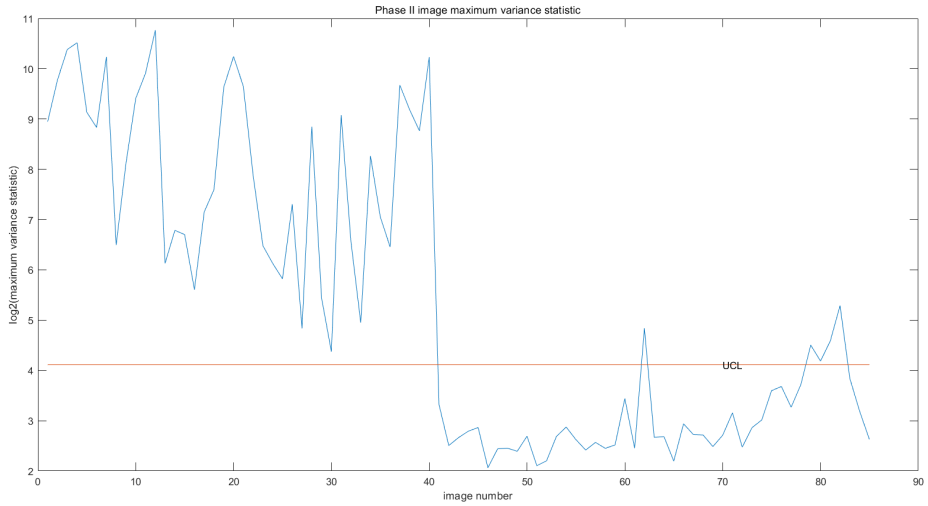

Figure 5.5: The phase II maximum variance based $\bar{X}$ control chart, where the first 40 images are sample with defects, other 45 are qualified samples. The orange line indicate the UCL and there are 5 false detection samples in the control chart.

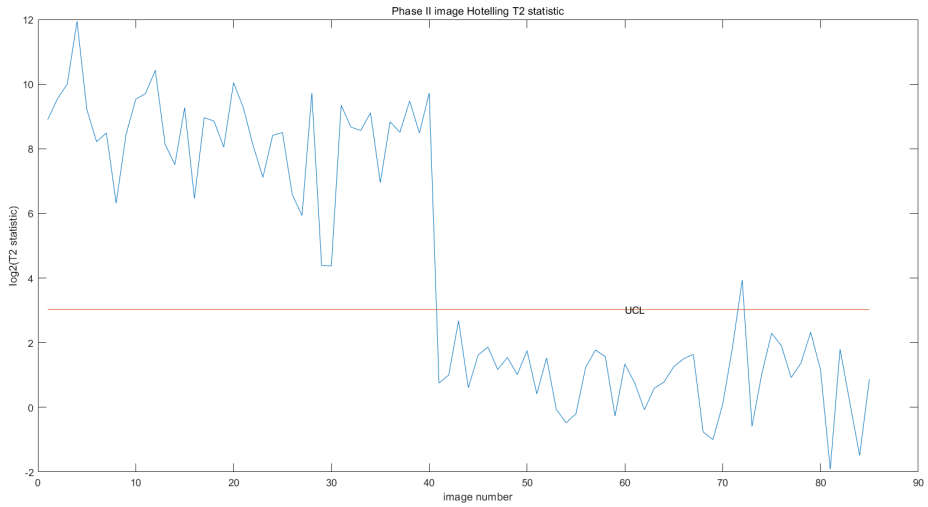

Figure 5.6: The phase II wavelet decomposition based Hotelling T^2 control chart, where the first 40 images are sample with defects, other 45 are qualified samples. The orange line indicate the UCL and there are 1 false detection sample in the control chart.

5.2.4 Application of two methods on different colour images

This study aims to verify if the above mentioned two methods are suitable to detect defects without considering the image colour. To change the colour, we first convert an RGB image into an HSV (hue, saturation, and value) image, an $M \times N \times 3$ numeric array with values in the range [0, 1]. The three dimensions of HSV (Fig 5.7) defines the hue, saturation, and brightness value for each pixel, respectively. The dimension hue describes value from 0 to 1 corresponding to the colour's position on a colour wheel. As hue increases from 0 to 1, the colour transitions from red to orange, yellow, green, cyan, blue, magenta, and finally back to red. As for saturation, it represents the amount of hue or departure from neutral. 0 indicates a neutral shade, whereas 1 indicates maximum saturation. And value works in conjunction with saturation and describes the brightness or intensity of the color, from 0 to 100 percent, where 0 is completely black, and 100 is the brightest and reveals the most color. We modify the value of the V frame to change the colour of the image from blue to red, as shown in 5.8.

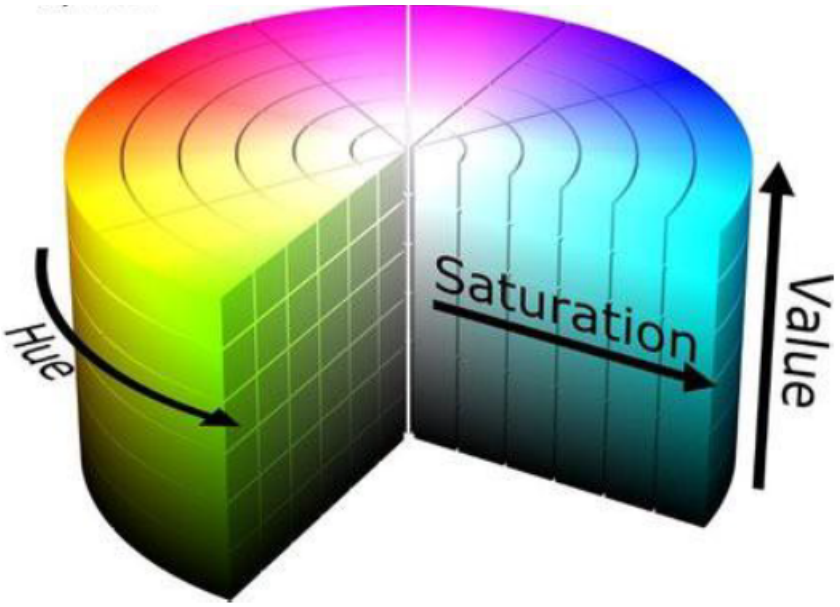

Figure 5.7: An Illustration of HSV dimention of an image.

To test the performance of two methods on different colour images, we first modify the brightness value of all samples and then repeat phase I and phase II procedure on all colour changed images as we have done in section 5.2.3. The control charts of phase I are shown in fig 5.9 and 5.10.

The mean value ($\overline{m}$) of maximum variance of changed colour image in phase I is 3.123e-04, the standard deviation (σ_s) is 1.09e-04, thus the UCL and LCL is 3.735e-04 and 2.51e-04 respectively. By method II in phase I the sample mean vector of Haar wavelet decomposition characteristic is

$$\overline{\mathbf{x}} = \begin{bmatrix} 0.229 & 0.414 & 0.085 \end{bmatrix}, \tag{5.6}$$

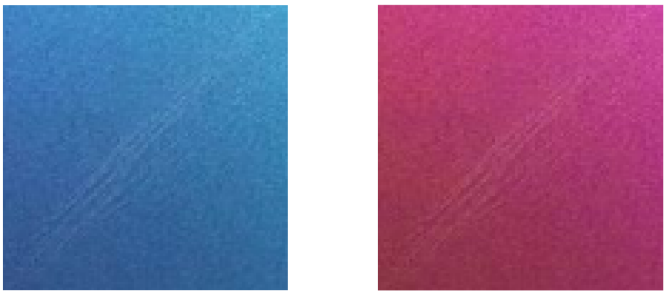

Figure 5.8: An example of defect image in different color. The left is the original image, the right one is the color changed image after HSV modification.

and the sample covariance matrix is

$$\mathbf{S} = \begin{bmatrix} 0.011 & 0.023 & -0.00019 \\ 0.02303 & 0.0868 & -0.00164 \\ -0.00019 & -0.00164 & 0.000639 \end{bmatrix}, \tag{5.7}$$

the UCL of colour changed image by method II is the same as that by original data, namely 8.157. The ARL of method I and method II according to equation (5.1) is 4.286 and 25, respectively.

The control charts of phase II data of two methods are shown in fig 5.11 and 5.12. There are 4 and 2 false detections in method I and method II respectively, the corresponding false detection rate of method I is 4.7%, for method II 2.35%.

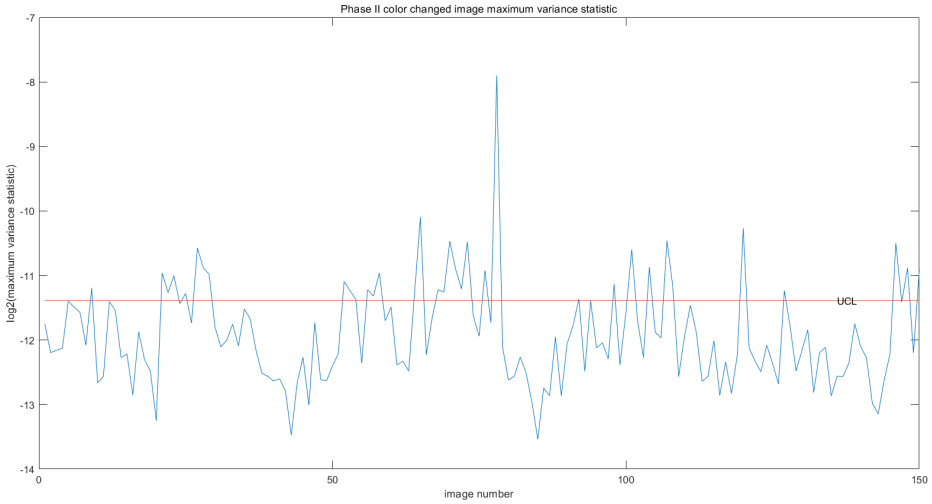

Figure 5.9: The phase I maximum variance based $\bar{X}$ control chart of changed colour images, where all the samples are qualified and there are 35 false alarms in the control chart. The orange line is the UCL of $\bar{X}$ control chart

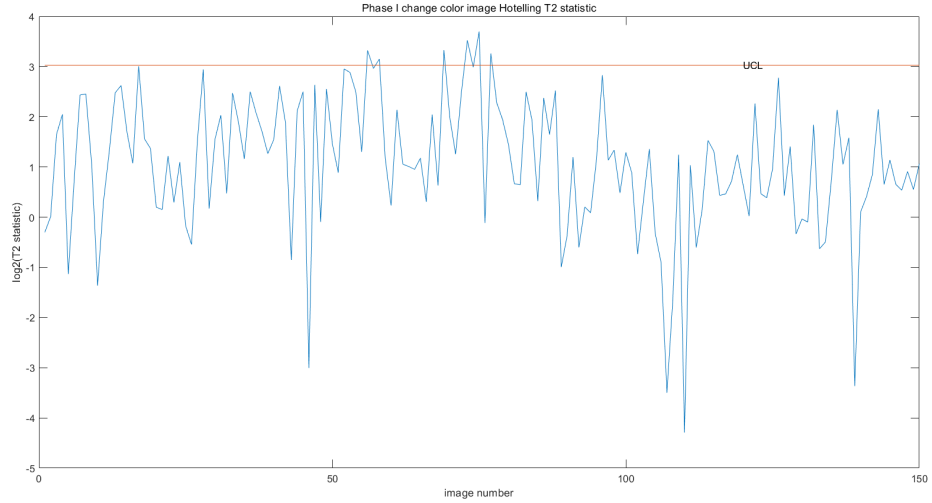

Figure 5.10: The phase I wavelet decomposition based Hotelling T^2 control chart of changed colour images, where all the samples are qualified and there are 6 false alarms in the control chart. The orange line is the UCL of Hotelling T^2 control chart

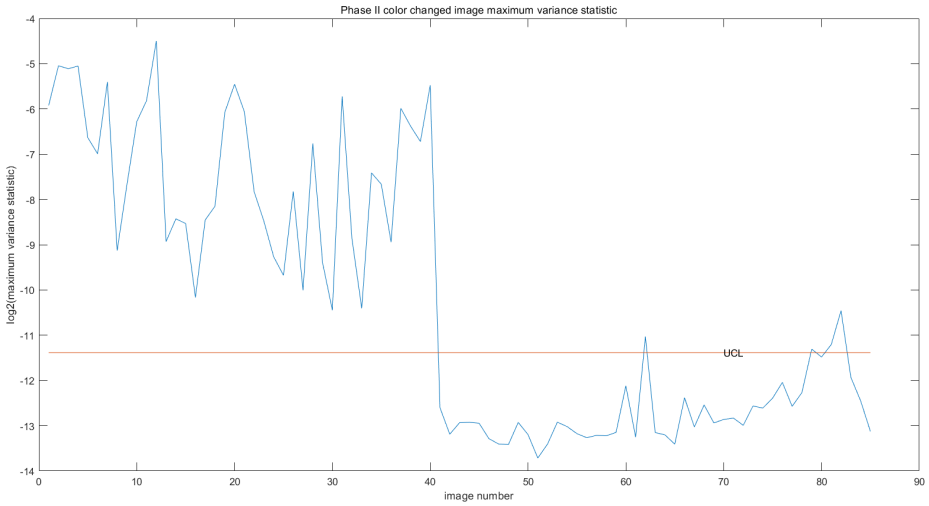

Figure 5.11: The maximum variance based $\bar{X}$ control chart of changed colour images in phase II, where the first 40 images are sample with defects, other 45 are qualified samples. The orange line indicate the UCL and there are 4 false detection samples in the control chart.

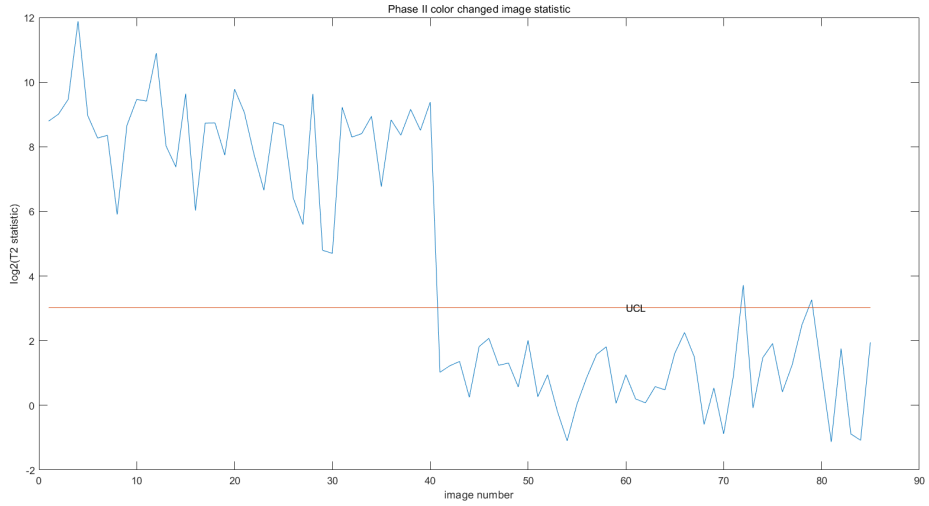

Figure 5.12: The wavelet decomposition based Hotelling T^2 control chart of changed colour images in phase II, where the first 40 images are sample with defects, other 45 are qualified samples. The orange line indicate the UCL and there are 2 false detection samples in the control chart.

5.3 conclusion

In this section we will summarize the results of empirical study and draw the conclusion. The result of the empirical study is shown in table 5.4

	phase I ARL	phase II ASS	phase II FDR
method I	3.13	27.42	5.8%
method II	30	67.12	1.1%
method I (different colour)	4.28	27.9	4.7%
method II (different colour)	25	62.7	2.3%

Table 5.4: The summary result of control charts in different methods. The table shows the different performance variable of two methods.

As is shown in the table, the ARL of method II is ten times that of method I, while the FDR of method II is five times smaller than that of method I, which lead to the conclusion that the method II achieve better performance in defects detection. With respect to the ASS, method II is almost twice as much as that of method I, which indicate that method II is more sensitive to defects than that of method I. By comparing the ARL, FDR, and ASS of colour images and normal images, we can see that there are slight difference between two forms of image, from which we can draw the conclusion that the colour of the image won't affect the performance of control chart by faults detection, which is a good news in our case by different colour handy cover defects detection.

6 Summary and Discussion

Please write down the findings for the experiments that you have been doing. It is also better to describe the results in figures or tables than only plain text. But again, you also need to have adequate text to explain the underlying meaning for the results.

Bibliography

[A. Crosier and Galand, 1976] A. Crosier, D. E. and Galand, C. (1976). Perfect channel splitting by use of interpolation/decimation/tree decomposition techniques. In *Proc. Int. Symp. on Info., Circuits and Systems,(Patras, Greece)*.

[Adams and Butler, 1999] Adams, S. B. and Butler, O. R. (1999). *Manufacturing the future: a history of western electric*. Cambridge University Press.

[Bersimis et al., 2007] Bersimis, S., Psarakis, S., and Panaretos, J. (2007). Multivariate statistical process control charts: an overview. *Quality and Reliability engineering international*, 23(5):517–543.

[Burt and Adelson, 1983] Burt, P. and Adelson, E. (1983). The laplacian pyramid as a compact image code. *IEEE Trans. Commun.*, 31:532–540.

[Haar, 1910] Haar, A. (1910). Zur theorie der orthogonalen funktionensysteme. *Mathematische Annalen*, 69(3):331–371.

[Heckert et al., 2002] Heckert, N. A., Filliben, J. J., Croarkin, C. M., Hembree, B., Guthrie, W. F., Tobias, P., and Prinz, J. (2002). Handbook 151: Nist/sematech e-handbook of statistical methods.

[Hotteling, 1947] Hotteling, H. (1947). Multivariate quality control, illustrated by the air testing of sample bombsights. *Techniques of statistical analysis*, pages 111–184.

[Jensen et al., 2006] Jensen, W. A., Jones-Farmer, L. A., Champ, C. W., and Woodall, W. H. (2006). Effects of parameter estimation on control chart properties: a literature review. *Journal of Quality technology*, 38(4):349–364.

[Lin, 2007] Lin, H.-D. (2007). Automated visual inspection of ripple defects using wavelet characteristic based multivariate statistical approach. *Image and Vision Computing*, 25(11):1785–1801.

[Megahed et al., 2011] Megahed, F. M., Woodall, W. H., and Camelio, J. A. (2011). A review and perspective on control charting with image data. *Journal of Quality Technology*, 43(2):83–98.

[Montgomery, 2020] Montgomery, D. C. (2020). *Introduction to statistical quality control*. John Wiley & Sons.

[Polikar et al., 1996] Polikar, R. et al. (1996). The wavelet tutorial.

[Shannon, 1949] Shannon, C. E. (1949). Communication in the presence of noise. *Proceedings of the IRE*, 37(1):10–21.

[Vetterli and Le Gall, 1989] Vetterli, M. and Le Gall, D. (1989). Perfect reconstruction fir filter banks: Some properties and factorizations. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(7):1057–1071.

[Zuech and Zuech, 2000] Zuech, N. and Zuech, N. (2000). Understanding and applying machihe vision.