EDGE DETECTION USING DISCRETE WAVELET TRANSFORM

BY

AZZAM ABDULAZIZ ALFARRAJ

A Thesis Presented to the

DEANSHIP OF GRADUATE STUDIES

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS

DHAHRAN, SAUDI ARABIA

In Partial Fulfillment of the
Requirements for the Degree of

MASTER OF SCIENCE

In

MATHEMATICS

December 2017

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
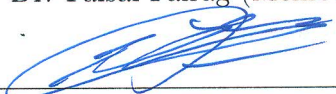## DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **AZZAM ABDULAZIZ ALFARRAJ** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN MATHEMATICS**.

Thesis Committee

_____
Dr. Mohamed El-Gebeily (Adviser)

_____
Dr. Faisal Fairag (Member)

_____
Dr. Abdelmalek Zidouri (Member)

_____
Dr. Husain Al-Attas
Department Chairman

_____
Dr. Salam A. Zummo
Dean of Graduate Studies

_____
Date

KING FAHD UNIVERSITY OF PETROLEUM & MINERALS
DHAHRAN 31261, SAUDI ARABIA

DEANSHIP OF GRADUATE STUDIES

This thesis, written by **AZZAM ABDULAZIZ ALFARRAJ** under the direction of his thesis adviser and approved by his thesis committee, has been presented to and accepted by the Dean of Graduate Studies, in partial fulfillment of the requirements for the degree of **MASTER OF SCIENCE IN MATHEMATICS**.

Thesis Committee

_____
Dr. Mohamed El-Gebeily (Adviser)

_____
Dr. Faisal Fairag (Member)

_____
Dr. Abdelmalek Zidouri (Member)

_____
Dr. Husain Al-Attas
Department Chairman

_____
Dr. Salam A. Zummo
Dean of Graduate Studies

_____
Date

الإهداء

إلى جدي..
محمد بن حمد الفراج..
الأستاذ الأول في حياتي..

Dedication

To my grandfather..
Mohammad Hamad Alfarraj..
My first mentor..

ACKNOWLEDGMENTS

TABLE OF CONTENTS

LIST OF FIGURES

THESIS ABSTRACT

NAME: Azzam Abdulaziz Alfarraj

TITLE OF STUDY: Edge Detection Using Discrete Wavelet Transform

MAJOR FIELD: Mathematics

DATE OF DEGREE: December 2017

The edge detection problem plays an important role in many applications. It helps in extracting the main features of an image and specifying its constituent parts. In the literature, this problem has been tackled using different mathematical approaches such as gradient-based detectors and minimization-based approach. Also, the continuous wavelet transform approach is involved in many edge detection algorithms. The cornerstone in all approaches is to look for points where the intensity of the image has a jump discontinuity. Applying the discrete wavelet transform in two dimensions to an image will result in large detail coefficients at the parts that have edges. The objective of this thesis is to propose a rigorously analysed algorithm for edge detection using discrete wavelet transform in two dimensions. To the best of our knowledge, this algorithm has not been treated with rigor in the literature.

الخلاصة

مسألة كشف حواف الصورة تلعب دورا مهما في تطبيقات عديدة، حيث تساعد في استخلاص الخواص الأساسية في الصورة وتحديد المكونات الرئيسة فيها. في أدبيات هذه المسألة، سبق أن عولجت باستخدام عدة مقاربات رياضية مثل الكاشفات المعتمدة على مؤثر التدرج و المعتمدة على التقليل. كذلك استُخدمت مقاربة تحويل المويجات المتصل في عدد من خوارزميات كشف الحواف. حجر الزاوية في كل تلك المقاربات هو البحث عن النقاط التي تتغير عندها شدة الصورة بشكل حاد. لو طبقنا تحويل المويجات المتقطع لبعدين على صورة ما فإننا سنحصل على معاملات تفاصيل كبيرة في الأجزاء التي تحوي حوافًا. الهدف في هذه الرسالة هو تقديم خوارزمية محللة بصرامة لكشف الحواف باستخدام تحويل المويجات المتقطع لبعدين. حسب أقصى ما توصلت إليه فإن هذه الخوارزمية لم يسبق أن عُملت بصرامة في أدبيات هذا البحث.

CHAPTER 1

INTRODUCTION

The edge detection problem plays an essential role in many applications. In fact, it is a field on its own in image processing [1]. Moreover, it is the first step in many computer vision applications, [2]. It is used mainly for feature extraction [3] or for image segmentation [4], [5]. Edges identify the significant information in an image, which enable the machine to separate its constituent parts. Additionally, it reduces the amount of data stored and preserves the main structural properties of the image for further processing [1], [6]. Other applications include medical imaging [7], [8], face recognition, automatic control of traffic systems, obstacle recognition, etc. [1], [2]

Before we start, it is worth noting that, mathematically speaking, an image is a function of two variables $u(x, y)$ defined on a rectangle. If grey-scale images are considered, as in this work, $u(x, y)$ is a grey-scale level that represents the intensity of the image color at the point (x, y), rescaled to be a value in an interval $[a, b]$ for prespecified real numbers a and b. This rescaling process is admissible since

images are bounded functions [9].

The cornerstone in edge detection techniques and algorithms is to look for points where the image intensity changes sharply. An edge is a sign for discontinuity [4]. Roughly speaking, it is a segment of connected pixels separating two different regions in the image [1], [4]. Indeed, this definition of an edge is not mathematically rigorous and it may lead to detecting false edges as in the case of textures [9]. This is why in the very early work on this problem, researchers only considered homogeneous (without texture) images [10].

The main challenges in edge detection include detecting false edges, missing edges in low contrast images, high computational time, sensitivity of noise, missing edges of objects with small change of intensity [2]. Amazingly, some detectors face problems in detecting edges of high-definition images. The algorithm proposed in this thesis performs better than most of the edge detectors when applied to high-definition images.

The discrete wavelet transform can be applied using plenty of wavelet families such as Haar, Daubechies, Mexican hat, etc. In this thesis we will focus on the Haar wavelet which produced the best results in detecting edges. This is because the abrupt change in its definition matches the discontinuity represented by an edge in an image. Furthermore, it does not mollify the image as other smoother wavelets do.

CHAPTER 2

LITERATURE REVIEW

2.1 Gradient-based Methods

Gradient-based methods for detecting edges were commonly used in applications since 1960's. Roberts [11] presented an operator in his Ph.D thesis where he used the following two 2×2 kernel matrices

$$D_x = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \qquad\qquad D_y = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

to compute the spatial gradient components of the image G_x and G_y. G_x and G_y stand for the difference in the x-direction and y-direction respectively. These components are used to get the gradient magnitude and angle as follows:

$$|G| = \sqrt{G_x^2 + G_y^2} \qquad\qquad \theta = \arctan(\frac{G_x}{G_y}) - \frac{3\pi}{4}.$$

Then some threshold T is set to mark the edge pixels wherever $|G| > T$. The angle θ represents the direction perpendicular to the edge line.

For further improvements, Sobel [12] and Prewitt [13] presented two kernels as 3×3 matrices as follows:

Sobel kernel matrices:

$$D_x = \begin{pmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix} \qquad D_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

Prewitt kernel matrices:

$$D_x = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix} \qquad D_y = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{pmatrix}$$

The gradient magnitude is computed as in Robert's while the gradient angle is

$$\theta = \arctan(\frac{G_x}{G_y}).$$

Anyway, all three operators were sensitive to noise.

In 1980, Marr and Hildreth [14] introduced an operator that convolves the image u with a Gaussian function G and then applies the Laplacian operator ∇^2 to

4

the convolved image. The algorithm then looks for the points where $\nabla^2(u * G) = 0$ to identify points of local extrema of the gradient. The idea was dominating at that time [1].

In 1986, John Canny presented an approach where edges are marked as maxima in gradient magnitude of a Gaussian-smoothed image [15]. The image u is convolved with a Gaussian function G then the gradient magnitude $|\nabla(u * G)|$ is computed to detect edges. Bergholm [16] used Canny's operator and he succeeded to differentiate between shadow contours and perfect ones [17]. Due to its great advantages, Canny edge detector is still used widely after some improvements in nowadays image processing techniques [1].

2.2   Minimization-based Methods

Mumford and Shah introduced a different approach to image segmentation in 1989 [18]. It is based on a minimization problem. They look for a pair (u, K) that minimizes the following functional

$$F(u, K) = \int_{\Omega \setminus K} (u - u_0)^2 dx + \alpha \int_{\Omega \setminus K} |\nabla u|^2 dx + \beta \int_K d\sigma$$

where $\Omega = [0, 1]^2$, u_0 the initial image, $K \subset \Omega$ is a set of discontinuities and α and β are non-negative parameters. In fact, if we drop the second term and reject

the trivial solution, we would get the reduced Mumford and Shah functional

$$E(u, K) = \int_{\Omega \backslash K} (u - u_0)^2 dx + \beta \int_K d\sigma.$$

This particular case gained some interest and Mumford and Shah [18], Morel and Solimini [19] [20], and Massari and Tamanini [21] devoted some work to it. Indeed, this approach will segment the image into its constituent pieces and K will present the edges in the image.

2.3   Wavelet Transform-based Methods

Stephane Mallat *et al.* [22] implemented Canny's approach with the conjugate of a *convolution kernel* $\theta(x, y)$, instead of a Gaussian function. Practically, it was computed with two wavelets ψ^1 and ψ^2 that are partial derivatives of the *convolution kernel* θ [9]

$$\psi^1 = -\frac{\partial \theta}{\partial x} \qquad \text{and} \qquad \psi^2 = -\frac{\partial \theta}{\partial y}.$$

The algorithm computes the wavelet transform components of an image and the result is proportional to the gradient sought after. More explicitly,

$$Wu(x, 2^j) = 2^j \nabla (u * \bar{\theta}),$$

where $Wu(a, 2^b)$ is the continuous wavelet transform of a function f at scale b and location a.

In 2016, Monika published a paper [23] discussing image edge detection using the discrete wavelet transform. The paper was descriptive and it was not clear what algorithm was used. Indeed, she talked about a threshold to get the high detail coefficients but she never specified the threshold or even the way how to arrive at it.

In 2005, Chaganti [17] described how to use four kinds of wavelets, Haar, Daubechies, Coifman and biorthogonal, to detect edges. His thesis was also descriptive, as he was an engineering student. Moreover, he did not state a clear algorithm.

CHAPTER 3

PRELIMINARIES

In this chapter, we will go through many notions and results needed in this thesis. First, we will give the definition of a wavelet. Then, we will give some results of vector valued inner products. Also, we will go through some function spaces needed in our work and the convergence of sequences of functions in these spaces. After that, we will give a brief about the Haar system since our work depends heavily on the Haar system. Finally, we will talk about the notion of multiresolution analysis.

3.1   Wavelet Theory

Definition 3.1 (Wavelet) *A function $\psi(x)$ is called a wavelet if it satisfies the following:*

$$\int_{\mathbb{R}} |\psi(t)|^2 dt < \infty \tag{3.1}$$

$$\int_{\mathbb{R}} \psi(t)dt = 0 \tag{3.2}$$

3.2   Vector Valued Inner Products

In this section we let $\mathbf{H}$ be a Hilbert space. We will show some definitions and properties of inner products of vectors in $\mathbf{H}$.

Now we give the definition of vector valued inner products in $\mathbf{H}$.

Definition 3.2 *Let $\mathbf{F}$ and $\mathbf{G}$ be two vectors of elements in $\mathbf{H}$ defined as follows:*

$$\mathbf{F} = \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \quad , \quad \mathbf{G} = \begin{pmatrix} g_1 \\ g_2 \\ \vdots \\ g_m \end{pmatrix}$$

where $f_i, g_j \in \mathbf{H}$ for $1 \leqslant i \leqslant n$ and $1 \leqslant j \leqslant m$, then

$$\langle \mathbf{F}, \mathbf{G} \rangle = \begin{pmatrix} \langle f_1, g_1 \rangle & \langle f_1, g_2 \rangle & \cdots & \langle f_1, g_m \rangle \\ \langle f_2, g_1 \rangle & \langle f_2, g_2 \rangle & \cdots & \langle f_2, g_m \rangle \\ \vdots & \vdots & \vdots & \vdots \\ \langle f_n, g_1 \rangle & \langle f_n, g_2 \rangle & \cdots & \langle f_n, g_m \rangle \end{pmatrix} \tag{3.3}$$

In fact, (3.3) is obtained by forming (formally) the matrix $\mathbf{F}\mathbf{G}^*$ and then applying the inner product to each element of the resulting matrix.

Lemma 3.1 *The vector valued inner products have the following three properties:*

(1) $\langle \mathbf{F} + \mathbf{E}, \mathbf{G} \rangle = \langle \mathbf{F}, \mathbf{G} \rangle + \langle \mathbf{E}, \mathbf{G} \rangle$

(2) $\langle \mathbf{AF}, \mathbf{BG} \rangle = \mathbf{A} \langle \mathbf{F}, \mathbf{G} \rangle \mathbf{B}^*$, *where* $\mathbf{A}$ *and* $\mathbf{B}$ *are complex matrices and* $\mathbf{B}^*$ *is the conjugate transpose of* $\mathbf{B}$

(3) $\langle \mathbf{F}, \mathbf{G} \rangle^* = \langle \mathbf{G}, \mathbf{F} \rangle$

Definition 3.3 *Given a vector* $\mathbf{G} = \{g_1, g_2, ..., g_n\}$, *we call* $\langle \mathbf{G}, \mathbf{G} \rangle$ *the Grammian of* $\mathbf{G}$.

Theorem 3.1 *Suppose* $\mathbf{G} = \{g_1, g_2, ..., g_n\}^T$ *is a set of vectors in an inner product space* $\mathbf{H}$. *Then* $\langle \mathbf{G}, \mathbf{G} \rangle$ *is invertible iff* $g_1, g_2, ..., g_n$ *are linearly independent.*

Proof. ($\Longrightarrow$:) Assume that $\langle \mathbf{G}, \mathbf{G} \rangle$ is invertible and $g_1, g_2, ..., g_n$ are not linearly independent.

$\therefore$ $\quad$ $\exists$ a vector c $\neq 0$ such that $c^* \mathbf{G} = 0$

$\Rightarrow \langle \mathbf{G}, c^* \mathbf{G} \rangle = 0$

$\Rightarrow \langle \mathbf{G}, \mathbf{G} \rangle c = 0$

$\Rightarrow c = 0$ which contradicts the assumption.

$\quad$ ($\Longleftarrow$:) Assume that $g_1, g_2, ..., g_n$ are linearly independent and $\langle \mathbf{G}, \mathbf{G} \rangle$ is not invertible.

$\therefore$ $\quad$ $\exists$ a vector c $\neq 0$ such that $\langle \mathbf{G}, \mathbf{G} \rangle c = 0$

$\Rightarrow \langle g_i, c^* \mathbf{G} \rangle = 0, \qquad i = 1, 2, ..., n$

$\Rightarrow c^* \mathbf{G} \perp \text{span}(\mathbf{G})$. On the other hand, $c^* \mathbf{G} \in \text{span}(\mathbf{G})$

$\therefore$ $\quad$ $c^* \mathbf{G} = 0$

Then either $c^* = 0$ or $\mathbf{G}$ is not linearly independent, which both contradict the assumption. ∎

3.3 Function Spaces

In this thesis we will come across many function spaces and I will give a brief definition of each one of them.

Definition 3.4 *The space of all functions $f : \Omega \longrightarrow \mathbb{R}$; such that f is continuous on Ω is denoted by $C(\Omega)$ and called the space of continuous functions.*

Definition 3.5 (See [24]) *For $1 \leqslant p < \infty$, we define $L^p(\Omega)$ to be the collection of all functions f for which*

$$\int_\Omega |f(x)|^p dx < \infty.$$

Moreover, the L^p-norm of a measurable function f is defined by

$$\|f\|_p = \left(\int_\Omega |f(x)|^p dx \right)^{1/p}$$

Definition 3.6 (See [24]) *A function f is called essentially bounded if there is some $M \geqslant 0$, called the essential upper bound for f, for which*

$$|f(x)| \leqslant M \text{ for almost all } x \in \Omega.$$

Definition 3.7 (See [24]) *We define $L^\infty(\Omega)$ to be the collection of all essentially*

bounded functions. The L^∞-norm of a measurable function $f(x)$ is defined by

$$\|f\|_\infty = \operatorname{ess\,sup}\{|f(x)| : x \in \Omega\}.$$

The space $L^2(\Omega)$ is a Hilbert space equipped with the inner product

$$\langle f(x), g(x) \rangle = \int_\Omega f(x)\overline{g(x)}dx, \qquad f, g \in L^2(\Omega)$$

where $\overline{g(x)}$ is the complex conjugate of $g(x)$.

Definition 3.8 (Support of a Function) *Suppose that $f : \Omega \longrightarrow \mathbb{R}$, then the support of f(x), denoted by $\operatorname{supp}(f)$, is the closure of the set of all vectors x such that $f(x) \neq 0$. i.e., $\operatorname{supp}(f) = \overline{\{x \in \Omega : f(x) \neq 0\}}$.*

Definition 3.9 (Compact Support) *A function f is said to be of compact support if its support can be included in a closed and bounded domain.*

Definition 3.10 *A function f(x) defined on an Ω is C^n on Ω with $n \in \mathbb{N}$ if it is n-times continuously differentiable on Ω. C^0 on I if it is continuous on Ω. f(x) is C^∞ on Ω if it is C^n on Ω for all $n \in \mathbb{N}$. A function f(x) is C_c^n on Ω if it is C^n on Ω and compactly supported.*

Theorem 3.2 (See [25]) *If $f \in L^2(\mathbb{R})$ and $\epsilon > 0$, then there exists a function g and $R > 0$, such that $\int_\mathbb{R} |f - g|^2 dx < \epsilon$ with $g(x) = f(x)\chi_{[-R,R]}(x)$.*

This means that most of the energy of f is in a finite interval.

3.4 Convergence of Sequences of Functions

Definition 3.11 (Pointwise convergence) *Given a function f(x) and a sequence of functions $\{f_n(x)\}_{n\in\mathbb{N}}$ defined on an interval I, we say that the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges pointwise to f(x) if for every $x_0 \in I$ the sequence $\{f_n(x_0)\}_{n\in\mathbb{N}}$ converges to $f(x_0)$. We write $f_n(x) \to f(x)$ pointwise as $n \to \infty$.*

Definition 3.12 (Uniform (L^∞) convergence) *Given a function f(x) and a sequence of functions $\{f_n(x)\}_{n\in\mathbb{N}}$ defined on an interval I, we say that the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges uniformly to f(x) if for every $\epsilon > 0$, there exists a number $N > 0$ such that if $n > N$, then $|f_n(x) - f(x)| < \epsilon$ for every $x \in I$. We write $f_n(x) \to f(x)$ uniformly as $n \to \infty$.*

Definition 3.13 (Mean-square (L^2) convergence) *Given a function f(x) and a sequence of functions $\{f_n(x)\}_{n\in\mathbb{N}}$ defined on an interval I, we say that the sequence $\{f_n(x)\}_{n\in\mathbb{N}}$ converges in the mean-square to f(x) if*

$$\lim_{n\to\infty} \int_I |f_n(x) - f(x)|^2 dx = 0.$$

We write $f_n(x) \to f(x)$ in the mean-square as $n \to \infty$.

Theorem 3.3 (See [25]) *Given that I is a finite interval,*

(a) $L^\infty(I) \subset L^2(I)$.

(b) if $f_n(x) \to f(x)$ uniformly on I, then $f_n(x) \to f(x)$ in the mean-square on I.

Theorem 3.4 (See [25]) *Given that I is a finite interval, if $f_n(x) \to f(x)$ in $L^\infty(I)$ or $L^2(I)$, then $f_n(x) \to f(x)$ pointwise on I.*

3.5 The Haar System

Alfred Haar, the Hungarian mathematician, itroduced the Haar function in his Ph.D thesis in 1910 [26]. In this section, we will discuss the Haar function and Haar scaling function with some notions, lemmas and theorems that relate to them. First, we will introduce the notion of dyadic intervals.

3.5.1 Dyadic Intervals

Definition 3.14 *For each pair $j, k \in \mathbb{Z}$, the interval $I_{j,k} = [2^{-j}k, 2^{-j}(k+1)]$ is called a dyadic interval. The collection of such intervals is known as the collection of dyadic sub-intervals of $\mathbb{R}$. The parameter j will be referred to as the scale level or scale parameter, or simply "scale".*

Note that if $j \neq j_1$ or $k \neq k_1$ then either

(i)
$$I_{j,k} \cap I_{j_1,k_1} = \phi \tag{3.4}$$

(ii)
$$I_{j,k} \subseteq I_{j_1,k_1} \quad or \quad I_{j_1,k_1} \subseteq I_{j,k} \tag{3.5}$$

Definition 3.15 *Given a dyadic interval $I_{j,k}$ at scale j, we write $I_{j,k} = I_{j,k}^l \cup I_{j,k}^r$, where $I_{j,k}^l$ and $I_{j,k}^r$ denote the left half and right half of the interval $I_{j,k}$ and they are dyadic intervals at scale $j+1$. Indeed, $I_{j,k}^l = I_{j+1,2k}$ and $I_{j,k}^r = I_{j+1,2k+1}$.*

3.5.2 The Haar Scaling Function and the Haar Wavelet Function

The Haar scaling function is defined as

$$\varphi^H(x) := \chi_{[0,1)}(x) = \begin{cases} 1 & if \quad 0 \leqslant x < 1 \\ 0, & otherwise. \end{cases}$$

Figure 3.1 shows the graph of the Haar scaling function.
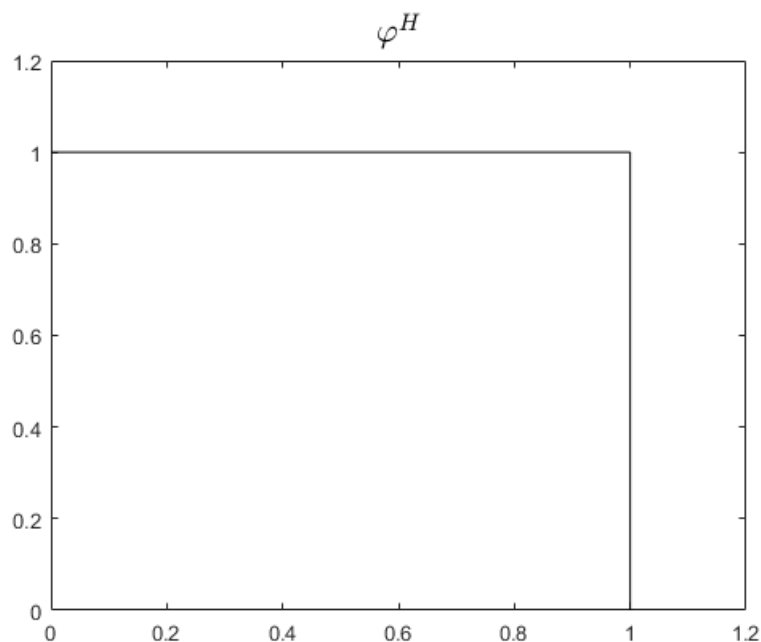
Figure 3.1: The Haar scaling function

Definition 3.16 *Let $\varphi_{j,k}^H(x) = 2^{j/2}\varphi^H(2^j x - k)$ with $j,k \in \mathbb{Z}$. The collection $\{\varphi_{j,k}^H(x)\}_{j,k\in\mathbb{Z}}$ is called the Haar system of scaling functions. The collection $\{\varphi_{j,k}^H(x)\}_{k\in\mathbb{Z}}$ is called the system of Haar scaling functions at scale j.*

Figure 3.2 shows the graphs of the Haar scaling functions at level $j = 2$.
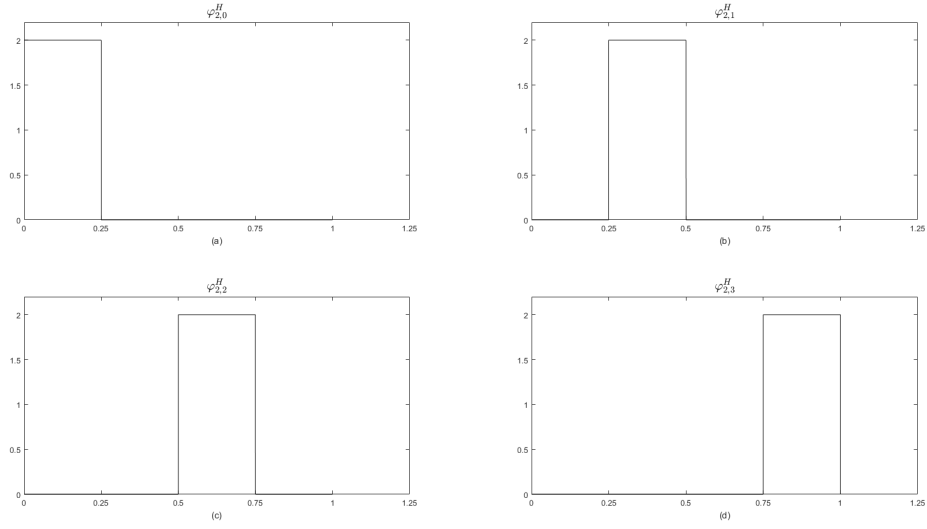
Figure 3.2: (a) The Haar scaling function at level $j = 2$ and $k = 0$. (b) $j = 2$ and $k = 1$. (c) $j = 2$ and $k = 2$. (d) $j = 2$ and $k = 2$

Lemma 3.2 *For all $j, k \in \mathbb{Z}$,*

$$\|\varphi_{j,k}^H(x)\|_2 = \|\varphi^H(x)\|_2.$$

Proof.

$$\|\varphi_{j,k}^H(x)\|_2^2 = \int_{-\infty}^{\infty} |\varphi_{j,k}^H(x)|^2 dx = \int_{-\infty}^{\infty} [2^{j/2}\varphi^H(2^j x - k)]^2 dx$$

Let $t = 2^j x - k$, $dt = 2^j dx$ then

$$\|\varphi_{j,k}^H(x)\|_2^2 = \int_{-\infty}^{\infty} 2^j [\varphi^H(t)]^2 2^{-j} dt = \int_{-\infty}^{\infty} [\varphi^H(t)]^2 dt = \|\varphi^H(x)\|_2^2$$

16

The system of Haar scaling functions satisfies the following [25] [27]:

(i) $\varphi_{j,k}^H(x)$ is supported on the interval $I_{j,k}$.

(ii) For every $j, k \in \mathbb{Z}$, $\int_{\mathbb{R}} |\varphi_{j,k}^H(x)|^2 dx = 1$.

(iii) For every $j, k \in \mathbb{Z}$, $\int_{\mathbb{R}} \varphi_{j,k}^H(x) dx = 2^{-j/2}$.

Lemma 3.3 *If $\varphi_{j,k}^H(x) = 2^{j/2}\varphi^H(2^j x - k)$ with $j, k \in \mathbb{Z}$, and $\mathrm{supp}(\varphi^H) = [a, b]$ then $\mathrm{supp}(\varphi_{j,k}^H) = I_{j,k} = [\dfrac{a+k}{2^j}, \dfrac{b+k}{2^j}]$*

Proof. If $\mathrm{supp}(\varphi^H) = [a, b]$, then we will look for values of x that will make the expression $(2^j x - k)$ fall within the interval [a,b].

$$2^j x - k \geqslant a \qquad \Rightarrow \qquad x \geqslant \frac{a+k}{2^j}$$

$$2^j x - k \leqslant b \qquad \Rightarrow \qquad x \leqslant \frac{b+k}{2^j}.$$

Then, $\mathrm{supp}(\varphi_{j,k}^H) = I_{j,k} = [\dfrac{a+k}{2^j}, \dfrac{b+k}{2^j}]$.

Moreover, $|I_{j,k}| = \dfrac{b-a}{2^j}$

We now address the question of which Haar scaling functions have support in [0,1].

Lemma 3.4 *The Haar scaling functions at scale j with support in $[0, 1]$ are exactly $\{\varphi_{j,k}^H(x)\}_{k=0}^{2^j-1}$.*

Proof. First, keep in mind that in the interval $[0, 1]$ we have $I_{j,k} = 2^{-j}[k, k+1]$

$$2^{-j}k \geqslant 0 \qquad \Rightarrow \qquad k \geqslant 0$$

$$2^{-j}(k+1) \leqslant 1 \qquad \Rightarrow \qquad (k+1) \leqslant 2^j \qquad \Rightarrow \qquad k \leqslant 2^j - 1$$

∎

The Haar wavelet function (or the Haar function) is defined as

$$\psi^H(x) = \chi_{[0,1/2)}(x) - \chi_{[1/2,1)}(x) = \begin{cases} 1 & if \quad 0 \leqslant x < 1/2 \\ -1 & if \quad 1/2 \leqslant x < 1 \\ 0, & otherwise. \end{cases}$$

Definition 3.17 *The collection $\{\psi_{j,k}^H(x)\}_{j,k \in \mathbb{Z}}$ is called the Haar wavelet system.*

The collection $\{\psi_{j,k}^H(x)\}_{k \in \mathbb{Z}}$ is called the system of Haar wavelet functions at scale j.

Figure 3.3 shows the graph of the Haar wavelet function.

ψ^H
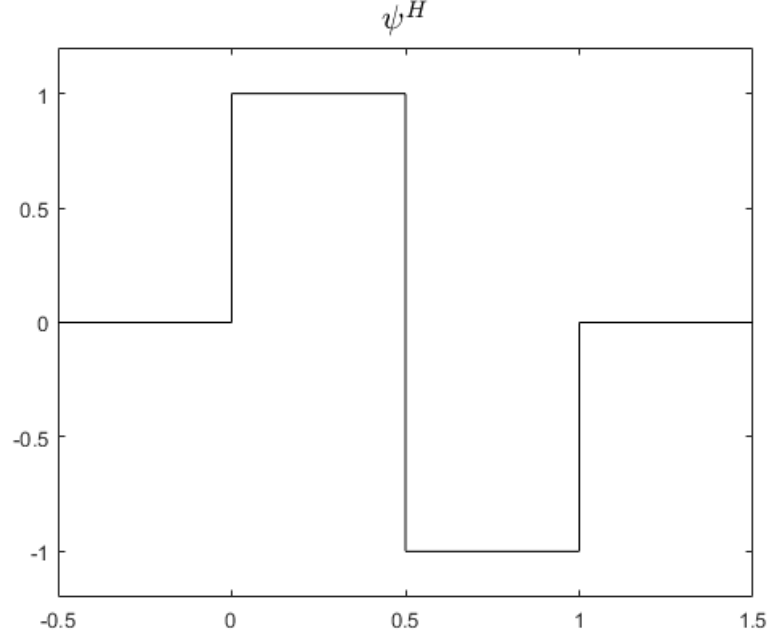
Figure 3.3: The Haar wavelet function

Note that lemmas 3.2, 3.3 and 3.4 also hold for Haar wavelet functions.

The Haar wavelet system satisfies the following [25] [27]:

(i) $\psi_{j,k}^H(x)$ is supported on the interval $I_{j,k}$.

(ii) For every $j, k \in \mathbb{Z}$, $\int\limits_{\mathbb{R}} |\psi_{j,k}^H(x)|^2 dx = \int\limits_{I_{j,k}} |\psi_{j,k}^H(x)|^2 dx = 1$.

(iii) For every $j, k \in \mathbb{Z}$, $\int\limits_{\mathbb{R}} \psi_{j,k}^H(x) dx = \int\limits_{I_{j,k}} \psi_{j,k}^H(x) dx = 0$.

(iv) $\psi_{j,k}^H(x) = 2^{j/2}(\chi_{I_{j,k}^l}(x) - \chi_{I_{j,k}^r}(x))$.

3.5.3   Orthogonality of The Haar System

Definition 3.18 (Orthogonal Basis and Orthonormal Basis) *A basis $\{g_n(x)\}_{n=1}^{\infty}$ of a Hilbert space X is orthogonal if*

$$\langle g_n, g_m \rangle = 0 \qquad for \qquad n \neq m.$$

An orthogonal basis is called orthonormal if

$$\langle g_n, g_m \rangle = \delta_{nm},$$

where δ_{nm} is the kronecker δ; $\delta_{nm} = \begin{cases} 1 & n = m \\ 0, & n \neq m. \end{cases}$

Theorem 3.5 *The system of Haar wavelet functions(or the Haar System) $\{\psi_{j,k}^{H}(x)\}_{j,k \in \mathbb{Z}}$ is orthonormal on $\mathbb{R}$.*

Proof.   The inner product we have is

$$\langle \psi_{j,k}^{H}, \psi_{j_1,k_1}^{H} \rangle = \int\limits_{-\infty}^{\infty} \psi_{j,k}^{H}(x)\psi_{j_1,k_1}^{H}(x)dx. \tag{3.6}$$

We show the orthonormality within a scale j first. The properties of dyadic intervals tell us that if $k \neq k_1$ then $\psi_{j,k}^{H}\psi_{j,k_1}^{H} \equiv 0$, and thus

$$\langle \psi_{j,k}^{H}, \psi_{j,k_1}^{H} \rangle = 0.$$

If $k = k_1$, then

$$\langle \psi_{j,k}^H, \psi_{j,k_1}^H \rangle = \int_{-\infty}^{\infty} \psi_{j,k}^H(x)\psi_{j,k_1}^H(x)dx = \int_{I_{j,k}} |\psi_{j,k}^H(x)|^2 dx = 1$$

Now, we show the orthonormality in case of different scales. Without loss of generality, say that $j > j_1$, then the properties (3.4) and (3.5) give us the following three possibilities:

(i) $I_{j,k} \cap I_{j_1,k_1} = \phi$ which results in

$$\langle \psi_{j,k}^H, \psi_{j_1,k_1}^H \rangle = \int_{-\infty}^{\infty} \psi_{j,k}^H(x)\psi_{j_1,k_1}^H(x)dx = 0.$$

(ii) $I_{j,k} \subseteq I_{j_1,k_1}^l$. In this case, $\psi_{j_1,k_1}^H(x)$ is identically 1 on I_{j_1,k_1}^l. This results in

$$\langle \psi_{j,k}^H, \psi_{j_1,k_1}^H \rangle = \int_{I_{j,k}} \psi_{j,k}^H(x)\psi_{j_1,k_1}^H(x)dx = \int_{I_{j,k}} \psi_{j,k}^H(x)dx = 0.$$

(iii) $I_{j,k} \subseteq I_{j_1,k_1}^r$. In this case, $\psi_{j_1,k_1}^H(x)$ is identically -1 on I_{j_1,k_1}^r and on $I_{j,k}$. This results in

$$\langle \psi_{j,k}^H, \psi_{j_1,k_1}^H \rangle = \int_{I_{j,k}} \psi_{j,k}^H(x)\psi_{j_1,k_1}^H(x)dx = -\int_{I_{j,k}} \psi_{j,k}^H(x)dx = 0.$$

$\blacksquare$

Theorem 3.6 *For any $j \in \mathbb{Z}$, the system of Haar scaling functions at scale j is an orthonormal system on $\mathbb{R}$*

21

Proof. First, If $k = k_1$ then

$$\langle \varphi_{j,k}^H, \varphi_{j,k}^H \rangle = \int\limits_{-\infty}^{\infty} \varphi_{j,k}^H(x)\varphi_{j,k}^H(x)dx = \int\limits_{-\infty}^{\infty} |\varphi_{j,k}^H(x)|^2 dx = 1.$$

Second, if $k \neq k_1$ then the properties of dyadic intervals tell that $\varphi_{j,k}^H \varphi_{j,k_1}^H \equiv 0$, and thus

$$\langle \varphi_{j,k}^H, \varphi_{j,k_1}^H \rangle = \int\limits_{-\infty}^{\infty} \varphi_{j,k}^H(x)\varphi_{j,k_1}^H(x)dx = 0.$$

$\blacksquare$

Theorem 3.7 ([25]) *For any $J \in \mathbb{Z}$, the collection $\{\varphi_{J,k}^H(x), \psi_{j,k}^H(x) : j \geqslant J, k \in \mathbb{Z}\}$ is an orthonormal system on $\mathbb{R}$*

3.6 Multiresolution Analysis (MRA)

We saw in Section 3.5 that the collection $\{\psi_{j,k}^H(x)\}_{j,k\in\mathbb{Z}} = \{2^{j/2}\psi^H(2^j x - k)\}_{j,k\in\mathbb{Z}}$ forms an orthonormal system on $\mathbb{R}$. In this section, we will see how to generalize this construction. In particular, we will see a general framework for constructing wavelet functions $\psi(x)$ such that the collection

$$\{\psi_{j,k}(x)\}_{j,k\in\mathbb{Z}} = \{2^{j/2}\psi(2^j x - k)\}_{j,k\in\mathbb{Z}}$$

is an orthonormal system on $\mathbb{R}$. Such collection $\{\psi_{j,k}(x)\}_{j,k\in\mathbb{Z}}$ is called a *wavelet orthonormal basis* on $\mathbb{R}$.

Definition 3.19 (Uniformly stable basis) *For a given scale j, the basis $\Phi_j =$*

22

$\{\varphi_{j,k}(x) : k \in \mathbb{Z}\}$ *is a uniformly stable basis for the space* S_j *if there exist two real numbers* α *and* β *independent of* j *and* $\alpha, \beta > 0$ *such that*

$$\alpha\|c\|_2 \quad \leqslant \quad \|\sum_{k\in\mathbb{Z}} c_{j,k}\varphi_{j,k}\|_2 \quad \leqslant \quad \beta\|c\|_2$$

Definition 3.20 (Multiresolution analysis MRA [28]) *A multiresolution analysis is a sequence* $\mathcal{S} = \{S_j\}_{j\in\mathbb{Z}}$ *of subspaces of* $L^2(\mathbb{R})$ *that satisfies the following properties:*

(i) *The spaces are nested. i.e.* $S_j \subset S_{j+1}$ *for every* $j \in \mathbb{Z}$.

(ii) *Their union is dense in* $L^2(\mathbb{R})$. *i.e.* $\overline{\bigcup_{j\in\mathbb{Z}} S_j} = L^2(\mathbb{R})$.

(iii) *Their intersection is trivial. i.e.* $\bigcap_{j\in\mathbb{Z}} S_j = \{0\}$.

(iv) *There is a function* φ *such that each* $\Phi_j = \{\varphi_{j,k}(x) : k \in \mathbb{Z}\}$ *is a uniformly stable basis for* S_j.

(v) *The spaces arises by scaling. i.e.* $\varphi \in S_j \Rightarrow \varphi(2\cdot) \in S_{j+1}$.

(vi) *They are shift-invariant. i.e.* $\varphi \in S_0 \Rightarrow \varphi(\cdot - k) \in S_0, k \in \mathbb{Z}$.

Definition 3.21 ([28]) *For a MRA* $\mathcal{S}$, *any function* $\varphi \in L^2(\mathbb{R})$ *that satisfies*

$$\varphi(x) = \sum_{k\in\mathbb{Z}} a_k\varphi(2x - k) \tag{3.7}$$

is called a scaling function of the MRA $\mathcal{S}$. *Moreover, the equation (3.7) is called*

the refinement equation and the sequence $\{a_k\}_{k \in \mathbb{Z}}$ is called the refinement coefficients (or refinement mask).

CHAPTER 4

DISCRETE WAVELET TRANSFORM IN ONE AND TWO DIMENSIONS

In this chapter, we will give the definitions and properties of spaces and operators needed in discrete wavelet transform in one dimension and then we will build on the one-dimension case to go to the two-dimension case. We will discuss the notions of approximation space and detail space and the operators associated with them. Indeed, our interest is in discrete wavelet transform in two dimensions since the images are two dimensional functions. All the work done in this chapter is over the interval [0,1] rather than $\mathbb{R}$ since our scope is images which are bounded in domain. Throughout this chapter, we will consider only the Haar system.

4.1 Discrete Wavelet Transform in One Dimension

Let $S_j = \text{span}(\{\varphi_{j,k}(x)\}_{k\in\mathbb{Z}})$ with $\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)$. In discrete wavelet transform in one dimension, we project $L^2[0,1]$ functions into the approximation space S_j.

$$\mathbf{P}_j : L^2[0,1] \longrightarrow S_j$$

$$\langle \mathbf{P}_j f, g \rangle = \langle f, g \rangle \qquad \forall g \in S_j \tag{4.1}$$

We can easily show the following properties of the operator $\mathbf{P}_j$

(i) $\mathbf{P}_j$ is linear. i.e. $\mathbf{P}_j(\alpha f + h) = \alpha \mathbf{P}_j f + \mathbf{P}_j h$.

(ii) $\mathbf{P}_j$ is an orthogonal projection. i.e. $\langle g, \mathbf{P}_j f - f \rangle = 0 \qquad \forall g \in S_j$

We will now show how to derive an explicit formula for the operator $\mathbf{P}_j$. Given that $\mathbf{\Phi}_j = \{\varphi_{j,k}(x)\}_{k\in\mathbb{Z}}$ is a basis for the space S_j, we have the following

$$\mathbf{P}_j f = \mathbf{c}^* \mathbf{\Phi}_j$$

$$\langle f, \mathbf{\Phi}_j \rangle = \langle \mathbf{P}_j f, \mathbf{\Phi}_j \rangle = \langle \mathbf{c}^* \mathbf{\Phi}_j, \mathbf{\Phi}_j \rangle = \mathbf{c}^* \langle \mathbf{\Phi}_j, \mathbf{\Phi}_j \rangle$$

So, since $\langle f, \mathbf{\Phi}_j \rangle = \mathbf{c}^* \langle \mathbf{\Phi}_j, \mathbf{\Phi}_j \rangle$

then, using Theorem 3.1 we get $\mathbf{c}^* = \langle f, \mathbf{\Phi}_j \rangle \langle \mathbf{\Phi}_j, \mathbf{\Phi}_j \rangle^{-1}$.

If $\mathbf{\Phi}_j$ is an orthogonal basis, as in the case of Haar system, then

$$\langle \mathbf{\Phi}_j, \mathbf{\Phi}_j \rangle^{-1} = \mathbf{I} \qquad \text{and} \qquad \mathbf{c}^* = \langle f, \mathbf{\Phi}_j \rangle.$$

Definition 4.1 (The approximation operator) *For every $j \in \mathbb{Z}$, the approximation operator $\mathbf{P}_j$ of a function $f \in L^2[0,1]$ is*

$$\mathbf{P}_j f = \langle f, \mathbf{\Phi_j} \rangle \mathbf{\Phi_j} \qquad \qquad \mathbf{\Phi}_j \text{ is orthogonal}$$

If we write the refinement coefficients in Definition 3.21 in a matrix $\mathbf{a}$ with elements defined as follows:

$$a_{k,m} = \frac{1}{\sqrt{2}} a_{m-2k} \tag{4.2}$$

we will be able to go from a finer level to a coarser level of approximation subspaces by the relation

$$\mathbf{\Phi}_j = \mathbf{a} \mathbf{\Phi}_{j+1}. \tag{4.3}$$

In the following proposition, we will show more properties of the operator $\mathbf{P}_j$.

Proposition 4.1 *The operator $\mathbf{P}_j$ given in Definition 4.1 has the following properties:*

(i) $\mathbf{P}_j^2 = \mathbf{P}_j$.

(ii) $\mathbf{P}_{j+1}\mathbf{P}_j = \mathbf{P}_j\mathbf{P}_{j+1} = \mathbf{P}_j$

(iii) $\langle \mathbf{P}_j f, g \rangle = \langle f, \mathbf{P}_j g \rangle \qquad \forall f, g \in L^2[0,1]$

Proof.

(i)

$$\mathbf{P}_j^2 f = \mathbf{P}_j \mathbf{P}_j f = \mathbf{P}_j \langle f, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \mathbf{P}_j \boldsymbol{\Phi}_j = \langle f, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j$$

(ii)

$$\mathbf{P}_{j+1} \mathbf{P}_j f = \mathbf{P}_{j+1} \langle f, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j = \langle f, \boldsymbol{\Phi}_j \rangle \mathbf{P}_{j+1} \boldsymbol{\Phi}_j$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \langle \boldsymbol{\Phi}_j, \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_{j+1}$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \langle \mathbf{a} \boldsymbol{\Phi}_{j+1}, \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_{j+1} \qquad \text{using the refinement relation (4.7)}$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \mathbf{a} \langle \boldsymbol{\Phi}_{j+1}, \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_{j+1}$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \mathbf{a} \boldsymbol{\Phi}_{j+1} = \langle f, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j = \mathbf{P}_j f$$

$$\mathbf{P}_j \mathbf{P}_{j+1} f = \mathbf{P}_j \langle f, \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_{j+1} = \langle f, \boldsymbol{\Phi}_{j+1} \rangle \mathbf{P}_j \boldsymbol{\Phi}_{j+1}$$

$$= \langle f, \boldsymbol{\Phi}_{j+1} \rangle \langle \boldsymbol{\Phi}_{j+1}, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j$$

$$= \langle f, \boldsymbol{\Phi}_{j+1} \rangle \langle \boldsymbol{\Phi}_{j+1}, \mathbf{a} \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_j \qquad \text{using the refinement relation (4.7)}$$

$$= \langle f, \boldsymbol{\Phi}_{j+1} \rangle \langle \boldsymbol{\Phi}_{j+1}, \boldsymbol{\Phi}_{j+1} \rangle \mathbf{a}^* \boldsymbol{\Phi}_j$$

$$= \langle f, \boldsymbol{\Phi}_{j+1} \rangle \mathbf{a}^* \boldsymbol{\Phi}_j = \langle f, \mathbf{a} \boldsymbol{\Phi}_{j+1} \rangle \boldsymbol{\Phi}_j$$

$$= \langle f, \boldsymbol{\Phi}_j \rangle \boldsymbol{\Phi}_j = \mathbf{P}_j f$$

(iii)

$$\langle \mathbf{P}_j f, g \rangle = \langle \langle f, \mathbf{\Phi}_j \rangle \mathbf{\Phi}_j, g \rangle$$

$$= \langle f, \mathbf{\Phi}_j \rangle \langle \mathbf{\Phi}_j, g \rangle$$

$$= \langle f, \langle g, \mathbf{\Phi}_j \rangle \mathbf{\Phi}_j \rangle$$

$$= \langle f, \mathbf{P}_j g \rangle$$

∎

Definition 4.2 (The detail operator) *For every $j \in \mathbb{Z}$, the detail operator $\mathbf{Q}_j$ is defined as*

$$\mathbf{Q}_j = \mathbf{P}_{j+1} - \mathbf{P}_j. \tag{4.4}$$

Moreover, let $W_j = \mathrm{Im}(\mathbf{Q}_j)$ and it is called the detail subspace.

In the following proposition, we will show some properties of the operator $\mathbf{Q}_j$.

Proposition 4.2 *The operator $\mathbf{Q}_j$ given in Definition 4.2 has the following properties:*

(i) $\mathbf{Q}_j^2 = \mathbf{Q}_j.$

(ii) $\mathbf{Q}_j \mathbf{P}_j = \mathbf{P}_j \mathbf{Q}_j = 0$

(iii) $\langle \mathbf{Q}_j f, g \rangle = \langle f, g \rangle \qquad \forall g \in W_j$

(iv) $\langle \mathbf{Q}_j f, g \rangle = \langle f, \mathbf{Q}_j g \rangle \qquad \forall f, g \in L^2[0,1]$

(v) $\mathbf{Q}_j g = g \qquad \forall g \in W_j$

Proof.

(i)

$$\mathbf{Q}_j^2 = (\mathbf{P}_{j+1} - \mathbf{P}_j)^2 = \mathbf{P}_{j+1}^2 - \mathbf{P}_{j+1}\mathbf{P}_j - \mathbf{P}_j\mathbf{P}_{j+1} + \mathbf{P}_j^2$$

$$= \mathbf{P}_{j+1} - \mathbf{P}_j - \mathbf{P}_j + \mathbf{P}_j \qquad \text{using Proposition 4.1(ii)}$$

$$= \mathbf{P}_{j+1} - \mathbf{P}_j = \mathbf{Q}_j$$

(ii) $\mathbf{Q}_j\mathbf{P}_j = (\mathbf{P}_{j+1} - \mathbf{P}_j)\mathbf{P}_j = \mathbf{P}_{j+1}\mathbf{P}_j - \mathbf{P}_j^2 = \mathbf{P}_j - \mathbf{P}_j = 0$

using Proposition 4.1(ii)

(iii) Let $g = \mathbf{Q}_j h$ for some $h \in L^2[0,1]$, and let $f \in L^2[0,1]$

$$\langle \mathbf{Q}_j f, g \rangle = \langle \mathbf{Q}_j f, \mathbf{Q}_j h \rangle$$

$$= \langle (\mathbf{P}_{j+1} - \mathbf{P}_j)f, \mathbf{Q}_j h \rangle$$

$$= \langle \mathbf{P}_{j+1} f, \mathbf{Q}_j h \rangle - \langle \mathbf{P}_j f, \mathbf{Q}_j h \rangle$$

$$= \langle f, \mathbf{Q}_j h \rangle - \langle f, \mathbf{P}_j \mathbf{Q}_j h \rangle \qquad \text{using Proposition 4.1(iii)}$$

$$= \langle f, \mathbf{Q}_j h \rangle = \langle f, g \rangle$$

(iv)

$$\langle \mathbf{Q}_j f, g \rangle = \langle (\mathbf{P}_{j+1} - \mathbf{P}_j) f, g \rangle$$

$$= \langle \mathbf{P}_{j+1} f, g \rangle - \langle \mathbf{P}_j f, g \rangle$$

$$= \langle f, \mathbf{P}_{j+1} g \rangle - \langle f, \mathbf{P}_j g \rangle \qquad \text{using Proposition 4.1(iii)}$$

$$= \langle f, (\mathbf{P}_{j+1} - \mathbf{P}_j) g \rangle = \langle f, \mathbf{Q}_j g \rangle$$

(v) Let $f \in L^2[0,1]$

$$\langle \mathbf{Q}_j g, f \rangle = \langle g, \mathbf{Q}_j f \rangle \qquad \text{by property (iv)}$$

$$= \langle g, f \rangle \qquad \text{by property (iii)}$$

The proposition 4.2 shows that the operator $\mathbf{Q}_j$ shares the same the properties as $\mathbf{P}_j$.

Lemma 4.1 *The approximation space S_{j+1} can be decomposed as follows:*

$$S_{j+1} = S_j \oplus W_j \tag{4.5}$$

Proof. First, we show that

$$S_{j+1} = S_j + W_j.$$

31

$(\subseteq :)$

Let $f \in S_{j+1}$, then

$$f = P_{j+1}f = (P_{j+1} - P_j)f + P_j f = Q_j + P_j f$$

then $S_{j+1} \subseteq S_j + W_j$.

$(\supseteq :)$

Since $S_j \subset S_{j+1}$ and $W_j \subset S_{j+1}$,

then $S_j + W_j \subseteq S_{j+1}$.

So, $S_{j+1} = S_j + W_j$.

Now, we show that $S_j \perp W_j$

$$\langle f, g \rangle = \langle \mathbf{P}_j f, g \rangle = \langle \mathbf{P}_j f, \mathbf{Q}_j g \rangle = \langle \mathbf{Q}_j \mathbf{P}_j f, g \rangle = 0.$$

Thus,

$$S_{j+1} = S_j \oplus W_j$$

∎

Now, if we keep applying the decomposition relation 4.5 to each approximation subspace $S_l, \quad 0 \leqslant l \leqslant j$ we get

$$S_{j+1} = S_0 \oplus W_0 \oplus W_1 \oplus \cdots \oplus W_j \tag{4.6}$$

The key idea in discrete wavelet transform is that any function $f \in S_{j+1}$ can

be decomposed into two parts in a coarser level, an approximation part in S_j and a detail part in W_j, as in (4.5). Moreover, it can be decomposed into an approximation part in the coarsest level S_0 and details in many levels from 0 to j as shown in (4.6). This decomposition process helps in many applications such as denoising, compression, singularity detection, etc.

At this stage, the following question arises: Is there a function ψ such that $\{\psi_{j,k}(x) : k \in \mathbb{Z}\}$ spans the detail subspace W_j? The following proposition gives the answer.

Proposition 4.3 ([28]) *Suppose that the scaling function φ has a finite refinement mask $\{a_k\}_{k\in\mathbb{Z}}$ and generates an orthonormal MRA $\mathcal{S}$. Let $b_k = (-1)^k a_{1-k}$ then $\psi(x) = \sum_{k\in\mathbb{Z}} b_k \varphi(2x - k)$ is a wavelet such that $\{\psi_{j,k}(x) : k \in \mathbb{Z}\}$ is an orthonormal basis for W_j. i.e. for every $j \in \mathbb{Z}$, the detail operator $\mathbf{Q}_j$ on a function $f \in L^2[0,1]$ is given by*

$$\mathbf{Q}_j f = \langle f, \mathbf{\Psi}_j \rangle \mathbf{\Psi}_j = \mathbf{d}_j^* \mathbf{\Psi}_j$$

where $\mathbf{d}_j^* = \langle f, \mathbf{\Psi}_j \rangle$.

If we write the coefficients $\{b_k : k \in \mathbb{Z}\}$ in Definition 4.3 in a matrix $\mathbf{b}$ with elements defined in the same way that we did in (4.2), we get the coarse-fine relation

$$\mathbf{\Psi}_j = \mathbf{b}\mathbf{\Phi}_{j+1}. \tag{4.7}$$

4.1.1 Fast Wavelet Transform in One Dimension

Now, we are going to introduce the fast wavelet transform used in the decomposition and the reconstruction of any function $f \in L^2[0,1]$. This transform uses the refinement equations introduced previously in (4.2) and (4.7) to get the approximation coefficient $\mathbf{c}_j$ and the detail coefficient $\mathbf{d}_j$ at the coarser levels from the approximation coefficient $\mathbf{c}_{j+1}$ at the finer level and vice-versa.

Decomposition

Given a function $f \in L^2[0,1]$, the projection of f into the space $\mathbf{S}_j$ can be written as follows:

$$\mathbf{P}_j f = \mathbf{P}_{j-1} f + \mathbf{Q}_{j-1} f$$

$$\langle f, \mathbf{\Phi}_j \rangle \mathbf{\Phi_j} = \langle f, \mathbf{\Phi}_{j-1} \rangle \mathbf{\Phi}_{j-1} + \langle f, \mathbf{\Psi}_{j-1} \rangle \mathbf{\Psi}_{j-1}. \qquad \text{See Lemma 4.1}$$

Let $\mathbf{c}_j^T = \langle f, \mathbf{\Phi}_j \rangle$ and $\mathbf{d}_j^T = \langle f, \mathbf{\Psi}_j \rangle$, we have:

$$\mathbf{c}_{j-1}^T = \langle f, \mathbf{\Phi}_{j-1} \rangle = \langle f, \mathbf{a}\mathbf{\Phi}_j \rangle$$

$$= \langle f, \mathbf{\Phi}_j \rangle \mathbf{a}^T = \mathbf{c}_j^T \mathbf{a}^T$$

$$\therefore \quad \mathbf{c}_{j-1} = \mathbf{a}\mathbf{c}_j \qquad (*)$$

$$\mathbf{d}_{j-1}^T = \langle f, \mathbf{\Psi}_{j-1} \rangle = \langle f, \mathbf{b}\mathbf{\Phi}_j \rangle$$

$$= \langle f, \mathbf{\Phi}_j \rangle \mathbf{b}^T = \mathbf{c}_j^T \mathbf{b}^T$$

$$\therefore \quad \mathbf{d}_{j-1} = \mathbf{b}\mathbf{c}_j \qquad\qquad (**)$$

The equations $(*)$ and $(**)$ show how to obtain the coarser approximation and detail coefficients from the finer one. This is a decomposition process. Repeating the decomposition process we get the decomposition diagram shown in Figure 4.1.
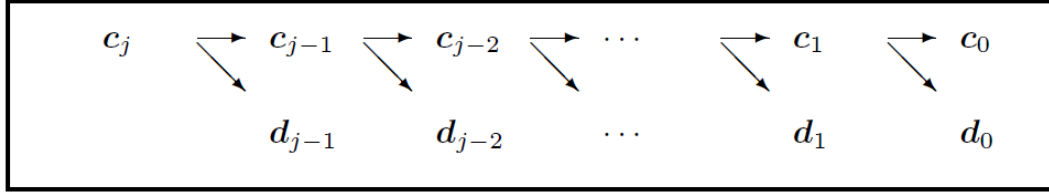
Figure 4.1: The decomposition of an input vector $\mathbf{c}_j$.

Reconstruction

This process is the reverse of decomposition and hence we keep the same definitions of $\mathbf{c}_j$ and $\mathbf{d}_j$. Recall that

$$\mathbf{P}_{j+1}f = \mathbf{P}_j f + \mathbf{Q}_j f$$

$$\langle f, \mathbf{\Phi}_{j+1} \rangle \mathbf{\Phi}_{j+1} = \langle f, \mathbf{\Phi}_j \rangle \mathbf{\Phi}_j + \langle f, \mathbf{\Psi}_j \rangle \mathbf{\Psi}_j$$

Then

$$\mathbf{c}_{j+1}^T \mathbf{\Phi}_{j+1} = \mathbf{c}_j^T \mathbf{\Phi}_j + \mathbf{d}_j^T \mathbf{\Psi}_j$$

$$\implies \quad \mathbf{c}_{j+1}^T \mathbf{\Phi}_{j+1} = \mathbf{c}_j^T \mathbf{a} \mathbf{\Phi}_{j+1} + \mathbf{d}_j^T \mathbf{b} \mathbf{\Phi}_{j+1}$$

$$\implies \quad \mathbf{c}_{j+1}^T \mathbf{\Phi}_{j+1} = (\mathbf{c}_j^T \mathbf{a} + \mathbf{d}_j^T \mathbf{b}) \mathbf{\Phi}_{j+1}$$

$$\implies \quad \mathbf{c}_{j+1}^T = \mathbf{c}_j^T \mathbf{a} + \mathbf{d}_j^T \mathbf{b}$$

$$\implies \quad \mathbf{c}_{j+1} = \mathbf{a}^T \mathbf{c}_j + \mathbf{b}^T \mathbf{d}_j.$$

The last equation shows how to reconstruct $\mathbf{c}_{j+1}$ from $\mathbf{c}_j$ and $\mathbf{d}_j$. We can start from the coarsest level of approximation and details $\mathbf{c}_0, \mathbf{d}_0, \mathbf{d}_1, ..., \mathbf{d}_j$ and recursively reconstruct the approximation coefficient $\mathbf{c}_{j+1}$ as depicted in Figure 4.2.
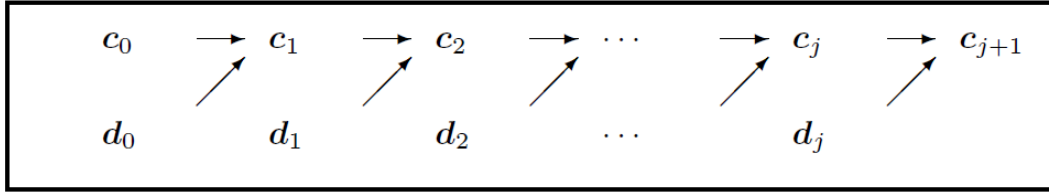

Figure 4.2: The reconstruction of a vector $\mathbf{c}_{j+1}$.

The reconstruction process is also known as the inverse discrete wavelet transform.

Operation Count

In this section, we will calculate the number of operations needed to perform the fast wavelet transform and compare it to the fast Fourier transform (FFT).

Assume that the input vector c_j is of length N. To get the vector c_{j-1} using the refinement equation (4.2) we need $(l_2 - l_1 + 1)\dfrac{N}{2}$ multiplication operations and $(l_2 - l_1 + 1)\dfrac{N}{2} - 1$ addition operations. In total, the calculations needed for c_{j-1} is $(l_2 - l_1 + 1)N - 1$. The same number of operations is needed for d_{j-1}. So, to get the approximation and detail coefficients at level $j - 1$ we need $2(l_2 - l_1 + 1)N - 2$ operations. Continuing by induction, we get that the total number of operations needed is $2(l_2 - l_1 + 1)N(1 + \dfrac{1}{2} + \dfrac{1}{4} + \ldots + \dfrac{1}{2^{j-1}}) - 2j < 4(l_2 - l_1 + 1)N$

So, the number of operations needed is $\mathcal{O}(N)$. On the other hand, the (FFT) needs $\mathcal{O}(NlogN)$ operations and obviously the fast wavelet transform is optimal [28].

4.2 Discrete Wavelet Transform in Two Dimensions

As we mentioned in the introduction of this thesis, images are functions of two variables. In this section, we will extend the notions and results shown in the previous section to the case of two dimensions. We will build scaling function and associated wavelet functions in two dimensions. Also, we will introduce the approximation subspace and the detail subspaces in three directions, horizontal, vertical and diagonal. The two-dimensional space will be thought of as the tensor product of two one-dimensional spaces.

Definition 4.3 (The tensor product) *Let V and W be two subspaces of*

$L^2[0,1]$, *with bases* E *and* F *respectively. The tensor product of* V *and* W *is defined as*

$$V \otimes W = \text{span}\{f(x)g(y) : f \in E, g \in F\}. \tag{4.8}$$

A typical element in $V \otimes W$ is of the form

$$\sum_{i=1}^{m}\sum_{j=1}^{n} \alpha_{ij} f_i(x) g_j(y). \tag{4.9}$$

Clearly, $V \otimes W \subseteq L^2[0,1]^2$

Theorem 4.1 *Let* U, V, W *be subspaces of* $L^2[0,1]$, *then*

$$(U + V) \otimes W = U \otimes W + V \otimes W$$

Proof. Let $f \in U, g \in V$ and $h, r \in W$, then

$(\subseteq :)$

$$(f(x) + g(x))h(y) = f(x)h(y) + g(x)h(y)$$

$$\therefore \quad (U + V) \otimes W \subseteq U \otimes W + V \otimes W$$

$(\supseteq :)$

Let $F \in U \otimes W$ and $G \in V \otimes W$ such that

$F = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} f_i(x) g_j(y)$ and $G = \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_{ij} h_i(x) g_j(y)$

$$F = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} f_i(x) g_j(y) = \sum_{i=1}^{m} \sum_{j=1}^{n} \alpha_{ij} (f_i(x) + 0) g_j(y) \qquad \in (U + V) \otimes W$$

$$G = \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_{ij} h_i(x) g_j(y) = \sum_{i=1}^{r} \sum_{j=1}^{s} \beta_{ij} (0 + h_i(x)) g_j(y) \qquad \in (U + V) \otimes W$$

$$\therefore \qquad F + G \in (U + V) \otimes W$$

∎

Corollary 4.2 *Let U, V, W, Y be subspaces of $L^2[0,1]$, then*

$$(U + V) \otimes (W + Y) = U \otimes W + U \otimes Y + V \otimes W + V \otimes Y$$

Definition 4.4 *Let $A_{m \times n}$ and $B_{r \times s}$ be two matrices, then*

$$A \otimes B = \begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ a_{21}B & a_{22}B & \dots & a_{2n}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1}B & a_{m2}B & \dots & a_{mn}B \end{pmatrix}$$

Theorem 4.3 *If E is an orthonormal basis for V and F is an orthonormal basis for W, then $E \otimes F$ is an orthonormal basis for $V \otimes W$.*

Proof. Suppose that $E = \{\xi_i\}_{i \in \mathbb{Z}}$ and $F = \{\eta_j\}_{j \in \mathbb{Z}}$, then $E \otimes F = \{\xi_i \eta_j\}_{i,j \in \mathbb{Z}}$. First, we show that any function in $V \otimes W$ can be represented as a linear combination of functions in $E \otimes F$.

39

Let $G \in V \otimes W$, then

$$G = \sum_{i=1}^{k} \sum_{j=1}^{l} \alpha_{ij} f_i g_j,$$

such that
$$f_i = \sum_{r \in \mathcal{I}_i} \beta_r \xi_r \qquad \text{and} \qquad g_j = \sum_{s \in \mathcal{I}_j} \beta'_s \eta_s$$

so
$$f_i g_j = \sum_{r \in \mathcal{I}_i, s \in \mathcal{I}_j} \beta_r \beta'_s \xi_r \eta_s$$

$$= \sum_{r \in \mathcal{I}_i, s \in \mathcal{I}_j} \beta_{rs} \xi_r \eta_s$$

so
$$G = \sum_{i=1}^{k} \sum_{j=1}^{l} \sum_{r \in \mathcal{I}_i, s \in \mathcal{I}_j} \alpha_{ij} \beta_{rs} \xi_r \eta_s$$

$$= \sum_{r \in \mathcal{I}_i, s \in \mathcal{I}_j} \beta'_{rs} \xi_r \eta_s$$

hence
$$G = \beta'(E \otimes F).$$

Now, we show the orthonormality of the elements of $E \otimes F$

$$\langle \xi_i \eta_j, \xi_k \eta_l \rangle = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \xi_i(x) \eta_j(y) \xi_k(x) \eta_l(y) dx dy$$

$$= \int_{-\infty}^{\infty} \xi_i(x) \xi_k(x) dx \int_{-\infty}^{\infty} \eta_j(y) \eta_l(y) dy$$

$$= \langle \xi_i \xi_k \rangle \langle \eta_j, \eta_l \rangle$$

$$= \delta_{i,k} \delta_{j,l} = 1 \qquad \text{when} \qquad i = k, j = l$$

$$\therefore \qquad \text{the elements of} \quad E \otimes F \quad \text{are orthonormal.}$$

40

4.2.1 MRA in $L^2[0,1]^2$

Let φ generate an orthonormal MRA $\mathcal{S} = \{S_j\}_{j \in \mathbb{Z}}$ in $L^2[0,1]$. Now, we want to build an MRA in $L^2[0,1]^2$.

$$\text{Let} \quad S_j^2 := S_j \otimes S_j$$

$$= (S_{j-1} \oplus W_{j-1}) \otimes (S_{j-1} \oplus W_{j-1})$$

$$= (S_{j-1} \otimes S_{j-1}) \oplus \underbrace{(S_{j-1} \otimes W_{j-1}) \oplus (W_{j-1} \otimes S_{j-1}) \oplus (W_{j-1} \otimes W_{j-1})}_{W_{j-1}^2}$$

$$=: S_{j-1}^2 \oplus W_{j-1}^2,$$

where S_{j-1}^2 is the approximation space at the coarser level $j-1$,

$S_{j-1} \otimes W_{j-1}$ is the subspace of vertical details,

$W_{j-1} \otimes S_{j-1}$ is the subspace of horizontal details and

$W_{j-1} \otimes W_{j-1}$ is the subspace of diagonal details.

Orthonormal bases of these spaces are as follows: (See Theorem 4.3)

For S_j^2, a basis is $\{\varphi_{j,k}(x)\varphi_{j,k}(y) : k, l \in \mathbb{Z}\} = \Phi_j \otimes \Phi_j = \Phi_j^2$,

and for $S_{j-1} \otimes W_{j-1}$, a basis is $\{\varphi_{j,k}(x)\psi_{j,k}(y) : k, l \in \mathbb{Z}\} = \Phi_j \otimes \Psi_j = \Psi_j^{21}$,

and for $W_{j-1} \otimes S_{j-1}$, a basis is $\{\psi_{j,k}(x)\varphi_{j,k}(y) : k, l \in \mathbb{Z}\} = \Psi_j \otimes \Phi_j = \Psi_j^{22}$,

and for $W_{j-1} \otimes W_{j-1}$, a basis is $\{\psi_{j,k}(x)\psi_{j,k}(y) : k, l \in \mathbb{Z}\} = \Psi_j \otimes \Psi_j = \Psi_j^{23}$.

Let $\Psi_j^2 = \begin{pmatrix} \Psi_j^{21} \\ \Psi_j^{22} \\ \Psi_j^{23} \end{pmatrix}$.

41

The projections from $L^2[0,1]^2$ to the spaces S_j^2 and W_j^2 are defined as follows:

$$\mathbf{P}_j^2 : L^2[0,1]^2 \longrightarrow S_j^2$$

$$\mathbf{P}_j^2 = \langle f, \boldsymbol{\Phi}_j^2 \rangle \boldsymbol{\Phi}_j^2$$

$$\mathbf{Q}_j^2 : L^2[0,1]^2 \longrightarrow W_j^2$$

$$\mathbf{Q}_j^2 = \langle f, \boldsymbol{\Psi}_j^2 \rangle \boldsymbol{\Psi}_j^2$$

$$= \langle f, \begin{pmatrix} \boldsymbol{\Psi}_j^{21} \\ \boldsymbol{\Psi}_j^{22} \\ \boldsymbol{\Psi}_j^{23} \end{pmatrix} \rangle \begin{pmatrix} \boldsymbol{\Psi}_j^{21} \\ \boldsymbol{\Psi}_j^{22} \\ \boldsymbol{\Psi}_j^{23} \end{pmatrix}$$

$$= (\langle f, \boldsymbol{\Psi}_j^{21} \rangle \quad \langle f, \boldsymbol{\Psi}_j^{22} \rangle \quad \langle f, \boldsymbol{\Psi}_j^{23} \rangle) \begin{pmatrix} \boldsymbol{\Psi}_j^{21} \\ \boldsymbol{\Psi}_j^{22} \\ \boldsymbol{\Psi}_j^{23} \end{pmatrix}$$

$$= \langle f, \boldsymbol{\Psi}_j^{21} \rangle \boldsymbol{\Psi}_j^{21} + \langle f, \boldsymbol{\Psi}_j^{22} \rangle \boldsymbol{\Psi}_j^{22} + \langle f, \boldsymbol{\Psi}_j^{23} \rangle \boldsymbol{\Psi}_j^{23}.$$

Let $\mathbf{c}_j^T = \langle f, \boldsymbol{\Phi}_j^2 \rangle, \quad \mathbf{d}_j^{1T} = \langle f, \boldsymbol{\Psi}_j^{21} \rangle, \quad \mathbf{d}_j^{2T} = \langle f, \boldsymbol{\Psi}_j^{22} \rangle, \quad \mathbf{d}_j^{3T} = \langle f, \boldsymbol{\Psi}_j^{23} \rangle$

and $\mathbf{d}_j = \begin{pmatrix} \mathbf{d}_j^1 \\ \mathbf{d}_j^2 \\ \mathbf{d}_j^3 \end{pmatrix}.$

Refinement Equations

Recall that $\mathbf{\Phi}_j = \mathbf{a}\mathbf{\Phi}_{j+1}$ and $\mathbf{\Psi}_j = \mathbf{b}\mathbf{\Phi}_{j+1}$, where $\mathbf{a}$ and $\mathbf{b}$ are given by $a_{k,m} = \frac{1}{\sqrt{2}}a_{m-2k}$ and $b_{k,m} = \frac{1}{\sqrt{2}}b_{m-2k}$, hence

$$
\begin{aligned}
\mathbf{\Phi}_j^2 = \mathbf{\Phi}_j \otimes \mathbf{\Phi}_j &= \mathbf{a}\mathbf{\Phi}_{j+1} \otimes \mathbf{a}\mathbf{\Phi}_{j+1} \\
&= (\mathbf{a} \otimes \mathbf{a})(\mathbf{\Phi}_{j+1} \otimes \mathbf{\Phi}_{j+1}) \\
&= \underbrace{(\mathbf{a} \otimes \mathbf{a})}_{\mathbf{a}^2}\mathbf{\Phi}_{j+1}^2 =: \mathbf{a}^2\mathbf{\Phi}_{j+1}^2 \qquad (*)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{\Psi}_j^2 = \begin{pmatrix} \mathbf{\Psi}_j^{21} \\ \mathbf{\Psi}_j^{22} \\ \mathbf{\Psi}_j^{23} \end{pmatrix} = \begin{pmatrix} \mathbf{\Phi}_j \otimes \mathbf{\Psi}_j \\ \mathbf{\Psi}_j \otimes \mathbf{\Phi}_j \\ \mathbf{\Psi}_j \otimes \mathbf{\Psi}_j \end{pmatrix} &= \begin{pmatrix} (\mathbf{a} \otimes \mathbf{b})\mathbf{\Phi}_{j+1}^2 \\ (\mathbf{b} \otimes \mathbf{a})\mathbf{\Phi}_{j+1}^2 \\ (\mathbf{b} \otimes \mathbf{b})\mathbf{\Phi}_{j+1}^2 \end{pmatrix} \\
&= \underbrace{\begin{pmatrix} \mathbf{a} \otimes \mathbf{b} \\ \mathbf{b} \otimes \mathbf{a} \\ \mathbf{b} \otimes \mathbf{b} \end{pmatrix}}_{\mathbf{b}^2} \mathbf{\Phi}_{j+1}^2 \\
&=: \mathbf{b}^2\mathbf{\Phi}_{j+1}^2. \qquad (**)
\end{aligned}
$$

Equations $(*)$ and $(**)$ will play an important rule in the decomposition and reconstruction processes.

4.2.2 Fast Wavelet Transform in Two Dimensions

After we showed the fast wavelet transform in one dimension, we will now show the fast wavelet transform in the case of two dimensions. The derivation is the

same.

Decomposition

For any function $f \in L^2[0,1]^2$, we get the approximation coefficient in a coarser

level from the finer one as follows:

$$\mathbf{c}_j^T = \langle f, \mathbf{\Phi}_j^2 \rangle = \langle f, \mathbf{a}^2 \mathbf{\Phi}_{j+1}^2 \rangle$$

$$= \langle f, \mathbf{\Phi}_{j+1}^2 \rangle \mathbf{a}^{2T}$$

$$= \mathbf{c}_{j+1}^T \mathbf{a}^{2T}$$

$$\therefore \qquad \mathbf{c}_j = \mathbf{a}^2 \mathbf{c}_{j+1}$$

and we get the detail coefficient as follows:

$$\mathbf{d}_j^T = \langle f, \mathbf{\Psi}_j^2 \rangle = \langle f, \mathbf{b}^2 \mathbf{\Phi}_{j+1}^2 \rangle$$

$$= \langle f, \mathbf{\Phi}_{j+1}^2 \rangle \mathbf{b}^{2T}$$

$$= \mathbf{c}_{j+1}^T \mathbf{b}^{2T}$$

$$\therefore \qquad \mathbf{d}_j = \mathbf{b}^2 \mathbf{c}_{j+1}$$

This is a decomposition process. Repeating the decomposition process we get the

decomposition diagram shown in Figure 4.1.

Reconstruction

The following equation shows how to reconstruct c_{j+1} from c_j and d_j. We can start from the coarsest level of approximation and details $c_0, d_0, d_1, ..., d_j$ and recursively reconstruct the approximation coefficient c_{j+1} as depicted in Figure 4.2.

$$\mathbf{c}_{j+1}^T \mathbf{\Phi}_{j+1}^2 = \mathbf{c}_j^T \mathbf{\Phi}_j^2 + \mathbf{d}_j^T \mathbf{\Psi}_j^2$$

$$= \mathbf{c}_j^T \mathbf{a}^2 \mathbf{\Phi}_{j+1}^2 + \mathbf{d}_j^T \mathbf{b}^2 \mathbf{\Phi}_{j+1}^2$$

$$= (\mathbf{c}_j^T \mathbf{a}^2 + \mathbf{d}_j^T \mathbf{b}^2) \mathbf{\Phi}_{j+1}^2$$

$$\therefore \qquad \mathbf{c}_{j+1} = \mathbf{a}^2 T \mathbf{c}_j + \mathbf{b}^2 T \mathbf{d}_j$$

CHAPTER 5

EDGE DETECTION

Discrete wavelet transform has many applications, including but not limited to compression, denoising and singularity detection. In this chapter, we are interested in image edge detection which can be defined as the process of locating and identifying the sharp discontinuities in an image [17]. First, we are going to consider singularity detection in one dimension. Then, we will extend the work to the case of discontinuities in two dimensions, i.e. discontinuities in images. After that, we will give a brief overview of Canny edge detector and compare the results of our algorithm to Canny's. In both detectors, we will add noise and observe the effect of noise on edge detection. Furthermore, we will experiment with our algorithm using Haar wavelet, Daubechies wavelets db2 and db3 to illustrate the distinction of the Haar wavelet in edge detection. Then, we will apply our algorithm to colored images and compare the results to Canny's.

5.1 Detecting Singularities of Functions in One Dimension

Detecting singularities of functions in one dimension can be done using the discrete wavelet transform by looking for high magnitude detail coefficients. In this work, we are going to use the Haar wavelet because it matches well abrupt changes in signals. Theorem 5.1 will show that the detail coefficients are small when a function f is smooth. As we shall see, it provides the justification for using wavelets in edge detection.

Definition 5.1 (Vanishing moment) *The function $f : \Omega \longrightarrow \mathbb{R}$ is said to have d vanishing moments if*

$$\int_\Omega x^r f(x) dx = 0, \qquad r = 0, 1, \ldots, d-1 \qquad (5.1)$$

Theorem 5.1 ([25]) *For $n \in \mathbb{N}$, let the function g be $C^n(\mathbb{R})$ and $g^{(n)}(x)$ be $L^\infty(\mathbb{R})$. Suppose that the function $\psi(x)$ has compact support and n vanishing moments and that $\int_\mathbb{R} |\psi_{j,k}(x)|^2 dx = 1$ for all $j, k \in \mathbb{Z}$. Then there exists a constant $C > 0$ that depends only on n and g such that*

$$|\langle g, \psi_{j,k} \rangle| \leqslant C 2^{-jn} 2^{-j/2}.$$

Proof. First, suppose that $\psi(x)$ is supported in the interval $I = [0, a]$ for some $a > 0$. Recall that $\psi_{j,k}(x)$ is supported in the interval $I_{j,k} = [2^{-j}k, 2^{-j}(k+a)]$

and $|I_{j,k}|$ is $2^{-j}a$. Denote the center of the interval $I_{j,k}$ by $\overline{x}_{j,k}$ and note that

$$\overline{x}_{j,k} = 2^{-(j+1)}a + 2^{-j}k.$$

Since $\psi_{j,k}(x)$ has n vanishing moments, given any polynomial $p(x)$ of degree no greater than $n-1$ we have

$$\int_{\mathbb{R}} p(x)\overline{\psi_{j,k}(x)}dx = 0. \tag{5.2}$$

Since $g(x)$ is $C^n(\mathbb{R})$, for every $j, k \in \mathbb{Z}$, $g(x)$ can be expanded about the point $\overline{x}_{j,k}$ in a Taylor expansion as follows:

$$g(x) = g(\overline{x}_{j,k}) + (x_{j,k} - \overline{x}_{j,k})g'(\overline{x}_{j,k}) + \ldots + \frac{1}{(n-1)!}(x_{j,k} - \overline{x}_{j,k})^{n-1}g^{(n-1)}(\overline{x}_{j,k}) + r_n(x),$$

$$\tag{5.3}$$

where

$$r_n(x) = \frac{1}{n!}(x_{j,k} - \overline{x}_{j,k})^n g^{(n)}(\xi)$$

for some number ξ between $\overline{x}_{j,k}$ and $x_{j,k}$.

If $x \in I_{j,k}$, then $x_{j,k} - \overline{x}_{j,k} \leqslant 2^{-(j+1)}a$, which implies that

$$|r_n(x)| \leqslant \frac{1}{n!}2^{-n(j+1)}a^n \max_{x \in I_{j,k}} |g^{(n)}(x)|. \tag{5.4}$$

Now, we apply (5.2) to (5.3) and compute

$$
\begin{aligned}
\langle g, \psi_{j,k} \rangle &= \int_{\mathbb{R}} g(x) \overline{\psi_{j,k}(x)} dx \\
&= \int_{\mathbb{R}} \left(\sum_{l=0}^{n-1} \frac{1}{l!} (x_{j,k} - \overline{x}_{j,k})^l g^{(l)}(\overline{x}_{j,k}) + r_n(x) \right) \overline{\psi_{j,k}(x)} dx \\
&= \left(\sum_{l=0}^{n-1} \frac{1}{l!} g^{(l)}(\overline{x}_{j,k}) \int_{\mathbb{R}} (x_{j,k} - \overline{x}_{j,k})^l \overline{\psi_{j,k}(x)} dx \right) \\
&\quad + \int_{\mathbb{R}} r_n(x) \overline{\psi_{j,k}(x)} dx \\
&= \int_{I_{j,k}} r_n(x) \overline{\psi_{j,k}(x)} dx, \qquad\qquad \text{by (5.2).}
\end{aligned}
$$

Then we apply (5.4) and the Cauchy-Schwarz inequality to get

$$
\begin{aligned}
|\langle g, \psi_{j,k} \rangle| &= \left| \int_{I_{j,k}} r_n(x) \overline{\psi_{j,k}(x)} dx \right| \\
&\leqslant \frac{1}{n!} 2^{-n(j+1)} a^n \max_{x \in I_{j,k}} |g^{(n)}(x)| \int_{I_{j,k}} |\psi_{j,k}(x)| dx \\
&\leqslant \frac{1}{n!} 2^{-n(j+1)} a^n \max_{x \in I_{j,k}} |g^{(n)}(x)| |I_{j,k}|^{1/2} \left(\int_{I_{j,k}} |\psi_{j,k}(x)|^2 dx \right)^{1/2} \\
&= \frac{1}{n!} \max_{x \in I_{j,k}} |g^{(n)}(x)| 2^{-n(j+1)} a^n 2^{-j/2} a^{1/2} \\
&= 2^{-nj} 2^{-j/2} \left(\frac{1}{n!} 2^{-n} a^{n+1/2} \max_{x \in I_{j,k}} |g^{(n)}(x)| \right) \\
&= 2^{-nj} 2^{-j/2} \left(\frac{1}{n!} 2^{-n} a^{n+1/2} ||g^{(n)}||_\infty \right).
\end{aligned}
$$

with $C = \frac{1}{n!} 2^{-n} a^{n+1/2} ||g^{(n)}||_\infty$, Theorem 5.1 is satisfied. ▌

The above theorem suggests that the magnitude of the detail coefficients that correspond to the smooth parts of a function will be very small compared with

the ones corresponding to the non-smooth parts of a function. This result has a crucial implication to edge detection in images, since edges can be interpreted as non-smooth parts of the image and hence can be detected by looking for high detail coefficients. The following example will substantiate this statement.

Example 5.2 *Consider the function f defined as follows:*

$$f(x) = \begin{cases} 1 & if & 0 \leqslant x < 1/3 \\ \\ 0 & if & 1/3 \leqslant x < 1. \end{cases}$$

It is clear that f is continuous and continuously differentiable infinitely many times on $(0, \frac{1}{3})$ and $(\frac{1}{3}, 1)$ and it is discontinuous at $x = \frac{1}{3}$. Now, we will calculate the detail coefficients at level $j = 1$. Then, we will give a general formula for calculating $d_{j,k}$ for all values of j and k. This will show that, in the subinterval that has the discontinuity, the detail coefficient is nonzero while, in the subintervals that have smooth parts the detail coefficients are zero.

Recall first that $\psi^H_{j,k}(x) = 2^{j/2}\psi(2^j x - k)$

$d_{1,0} = \langle f, \psi^H_{1,0} \rangle = \int\limits_0^{1/4} \sqrt{2}dx + \int\limits_{1/4}^{1/3}(-\sqrt{2})dx + \int\limits_{1/3}^{1/2} 0dx = (\frac{1}{4})\sqrt{2} - (\frac{1}{12})\sqrt{2} + 0 = \frac{\sqrt{2}}{6}$

$d_{1,1} = \langle f, \psi^H_{1,1} \rangle = \int\limits_{1/2}^{3/4}(0)\sqrt{2}dx + \int\limits_{3/4}^{1}(0)(-\sqrt{2})dx = 0 + 0 = 0$

Now, we are going to deduce a general formula for calculating $d_{j,k}$. First, let us

determine the dyadic subinterval $I_{j,k'} = [\frac{k'}{2^j}, \frac{k'+1}{2^j}]$ that contains the value $x = \frac{1}{3}$.

$$\frac{1}{3} \in [\frac{k'}{2^j}, \frac{k'+1}{2^j}] \qquad \Rightarrow \qquad \frac{1}{3} \geqslant \frac{k'}{2^j} \qquad \text{and} \qquad \frac{1}{3} \leqslant \frac{k'+1}{2^j}$$

$$\frac{1}{3} \geqslant \frac{k'}{2^j} \qquad \Rightarrow \qquad k' \leqslant \frac{2^j}{3}$$

$$\frac{1}{3} \leqslant \frac{k'+1}{2^j} \qquad \Rightarrow \qquad k' \geqslant \frac{2^j}{3} - 1$$

$$\therefore \qquad k' = \lfloor \frac{2^j}{3} \rfloor$$

So, for any value of j we have three cases for k. We will show the value of the detail coefficients in each case.

Case I: $k < \lfloor \frac{2^j}{3} \rfloor$

In this case, $x < \frac{1}{3}$ which implies that $f(x) = 1$. So

$$d_{j,k} = \langle f, \psi_{j,k}^H \rangle = \int_{I_{j,k}} f\psi_{j,k}^H dx = \int_{I_{j,k}} \psi_{j,k}^H dx = 0$$

Case II: $k > \lfloor \frac{2^j}{3} \rfloor$

In this case, $x > \frac{1}{3}$ which implies that $f(x) = 0$. So

$$d_{j,k} = \langle f, \psi_{j,k}^H \rangle = \int_{I_{j,k}} f\psi_{j,k}^H dx = \int_{I_{j,k}} (0)\psi_{j,k}^H dx = 0$$

Case III: $k = k' = \lfloor \frac{2^j}{3} \rfloor$

In this case, $\frac{1}{3} \in I_{j,k'}$ so we have two subcases:

Case III(a): $\dfrac{1}{3} \in I^l_{j,k'}$, which gives

$$d_{j,k'} = \langle f, \psi^H_{j,k'} \rangle = \int_{I_{j,k'}} f\psi^H_{j,k'}dx = \int_{\frac{k'}{2^j}}^{1/3} \psi^H_{j,k'}dx$$

$$= \int_{\frac{k'}{2^j}}^{1/3} 2^{j/2}dx = \left(\frac{1}{3} - \frac{k'}{2^j}\right)(2^{j/2}) > 0$$

$$\text{since} \quad \frac{1}{3} > \frac{k'}{2^j}$$

Case III(b): $\dfrac{1}{3} \in I^r_{j,k'}$, which gives

$$d_{j,k'} = \langle f, \psi^H_{j,k'} \rangle = \int_{I_{j,k'}} f\psi^H_{j,k'}dx = \int_{I^l_{j,k'}} \psi^H_{j,k'}dx + \int_{\frac{2k'+1}{2^{j+1}}}^{1/3} f\psi^H_{j,k'}$$

$$= 0 + \int_{\frac{2k'+1}{2^{j+1}}}^{1/3} 2^{j/2} = \left(\frac{1}{3} - \frac{2k'+1}{2^{j+1}}\right)(2^{j/2}) > 0$$

$$\text{since} \quad \frac{1}{3} > \frac{2k'+1}{2^{j+1}}$$

Figure 5.1 shows the function $f(x)$ and the detail coefficients for the levels $j = 1, 2, 3$. You can see clearly that the coefficient of the subinterval that contains the discontinuity is high while all other coefficients are zeros.
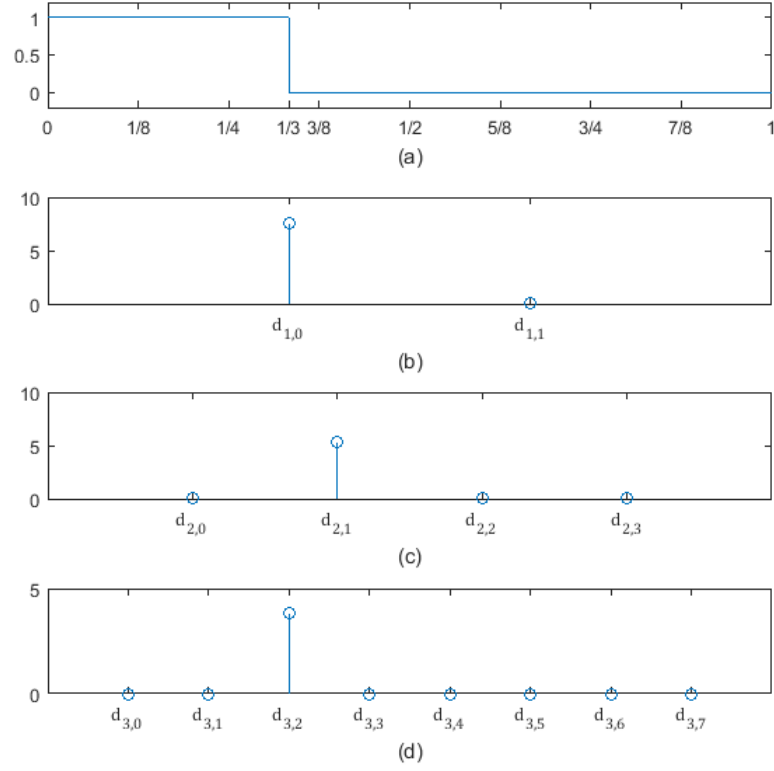
Figure 5.1: (a) The function $f(x)$, (b) The detail coefficients for $j = 1$, (c) The detail coefficients for $j = 2$ and (d) The detail coefficients for $j = 3$

The proposed algorithm in one dimension

Now, we present the most important part in this section which is an algorithm to detect singularities using discrete wavelet transform.

Suppose that we have a function $f \in L^2[0,1]$, and we want to detect singularities in it. The algorithm proposed is as follows:

Algorithm 5.3 (Wavelet-based singularity detection algorithm)

1) Apply discrete wavelet transform in one dimension at scale j, chosen a priori, which produces the approximation and detail coefficient vectors $\mathbf{c}_j$ and $\mathbf{d}_j$.

2) Discard the approximation coefficient vector $\mathbf{c}_j$.

3) Identify the detail coefficient $d_{j,k}$ that has the maximum modulus in $\mathbf{d}_j$ and call it m.

4) Choose a threshold T. Here we take $T = \dfrac{m}{10}$.

5) Discard all the detail coefficients that are less than T in modulus.

6) The remaining coefficients specify the dyadic subintervals where the singularities are located. The higher the scale j, the more precise the singularity is bracketted since $I_{j,k} = [2^{-j}k, 2^{-j}k + 1]$.

The reason for step 2 is that the approximation coefficient $\mathbf{c}_j$ contains the smooth part of the signal, since S_j contains all polynomials $\mathcal{P}_d$ of degree $d - 1$, where d is the vanishing moment of ψ. Singularities are manifested only in the detail coefficients since W_j is orthogonal to these polynomials. Figure 5.2 shows the flow chart of this algorithm.
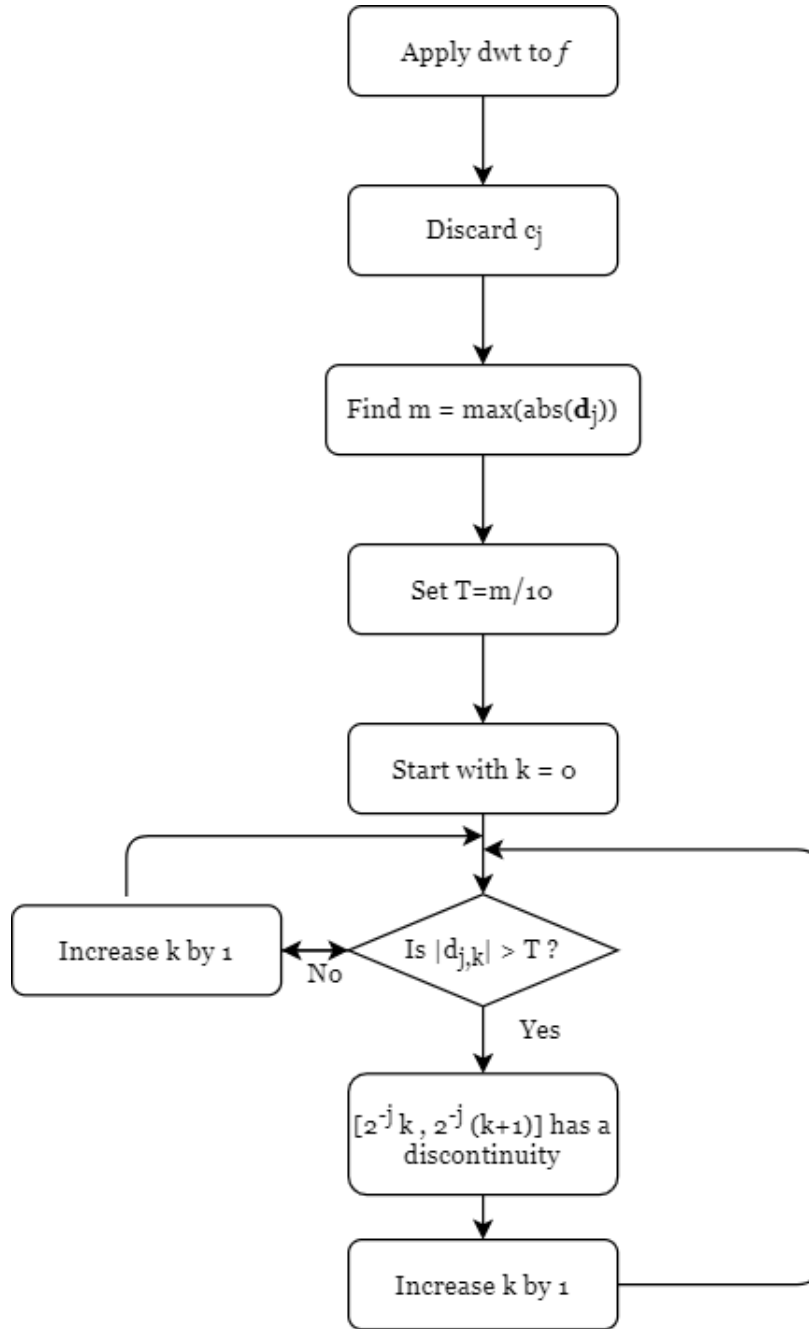
Figure 5.2: Flow chart for wavelet-based singularity detection algorithm

5.2 Edge Detection Using Discrete Wavelet Transform in Two Dimensions

It is worth noting that images are functions of two variables [29]. Moreover, they are bounded and at the same time have bounded variations [9]. Discrete wavelet transform in two dimensions has many applications in image processing such as compression and denoising. Most importantly for us, it enables us to detect edges by tracking the high detail coefficients in vertical, horizontal and diagonal directions. In this section, we will present an algorithm for image edge detection. In fact, our algorithm is analogous to the one we proposed in the previous section. Edge detection in the presence of noise will also be tried. Images considered in this work, are all high-definition square images of size 2048×2048. Indeed, our algorithm outperformed the Canny edge detector, which is the most popular one used in the literature. The highest scale j that can be reached is

$$log_2(\sqrt{\text{size of image}}) - 1 \qquad\qquad (*)$$

.

The proposed algorithm in two dimensions

Suppose that we have an image $u \in L^2[0,1]^2$ and we want to detect the edges in u. The algorithm proposed is as follows:

Algorithm 5.4 (Wavelet-based edge detection algorithm)

1) Convert the image to a black-and-white image.

2) Apply discrete wavelet transform in two dimensions at scale j according to $(*)$ to produce the approximation and detail coefficient vectors $\mathbf{c}_j$ and $\mathbf{d}_j$. Recall that $\mathbf{d}_j = \begin{pmatrix} \mathbf{d}_j^1 & \mathbf{d}_j^2 & \mathbf{d}_j^3 \end{pmatrix}^T$. (See 4.2.1)

3) Discard the approximation coefficient vector $\mathbf{c}_j$.

4) Identify the detail coefficient $d_{j,k}^l$ that has the maximum modulus for every l with $1 \leqslant l \leqslant 3$ and call them m_1, m_2 and m_3, respectively. Then choose the maximum one among them and call it m.

5) Let the threshold be $T = \dfrac{m}{10}$.

6) Discard all the detail coefficients in the three directions that are less than T in modulus.

7) Plot the location of detail coefficients after thresholding to get the image edges.

It is worth noting that this algorithm is applied using MATLAB software. Figure 5.3 shows the flow chart of this algorithm.
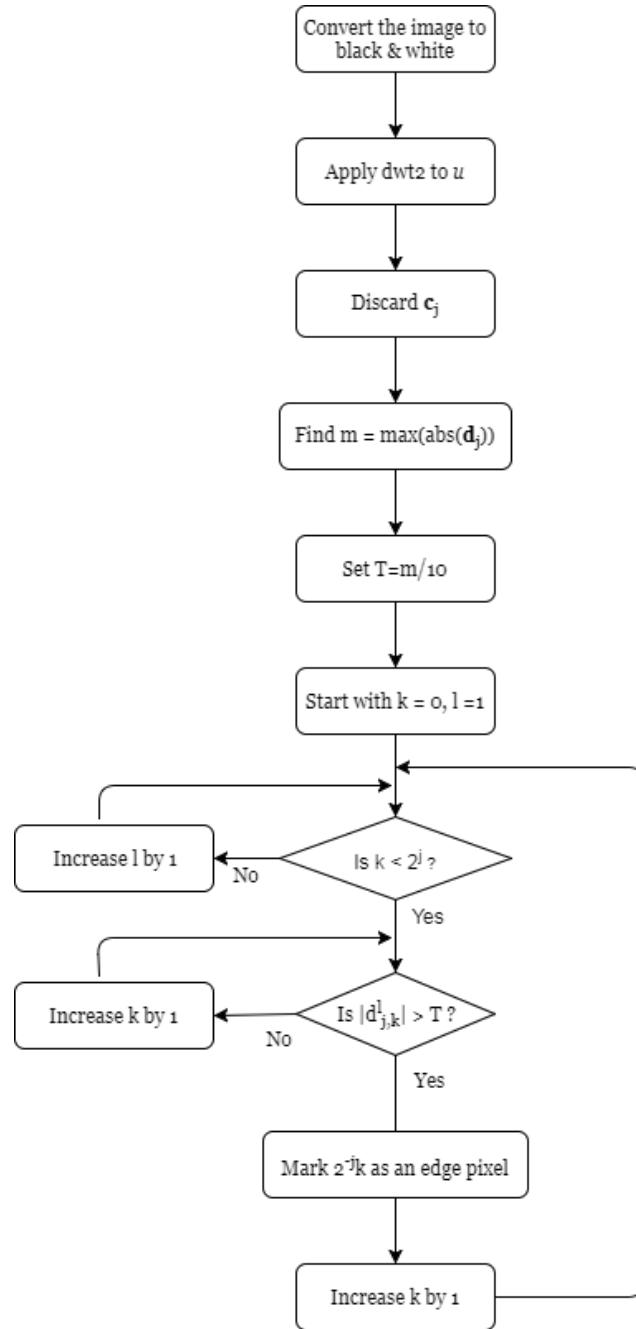
Figure 5.3: Flow chart for wavelet-based edge detection algorithm

5.3   Edge Detection Using Canny edge detector

John Canny proposed his detector in 1986 [15]. In this section, we will give a short description of Canny edge detector. His detector is dominating nowadays due to its great advantages [1]. Given an image u, the algorithm that Canny proposed is as follows:

Algorithm 5.5 (Canny edge detector algorithm)

1) Convert the image to a black-and-white image.

2) Convolve the image u with a Gaussian function G. This step is done to smooth the image. i.e. $w = u * G$.

3) Take the modulus of gradient of the convolved image. i.e. Find $|\nabla w| = |\nabla(u * G)|$

4) Threshold by looking for pixels of high modulus gradient $|\nabla w| \geqslant T$.

5) Suppress any pixel that is not maximum in its neighborhood. Indeed, here it only compares the pixel with the two neighboring pixels that are in the direction of the tangent $\theta = \arctan \dfrac{w_y}{w_x}$. This step is called non-maximum suppression to thin the edges.

6) Double threshold by setting two thresholds T_{high} and T_{low} and mark the pixels above T_{high} as *strong*, pixels between T_{high} and T_{low} as *weak* and suppress pixels below T_{low}.

7) Track the true edges by comparing the *weak* pixels to the 8 neighboring pixels. If a *weak* pixel is connected to a *strong* one, then it is preserved. Otherwise, it is considered false edge and, therefore, suppressed. Note that, the Canny edge detector is applied to images using the MATLAB command "edge(I,'canny')".
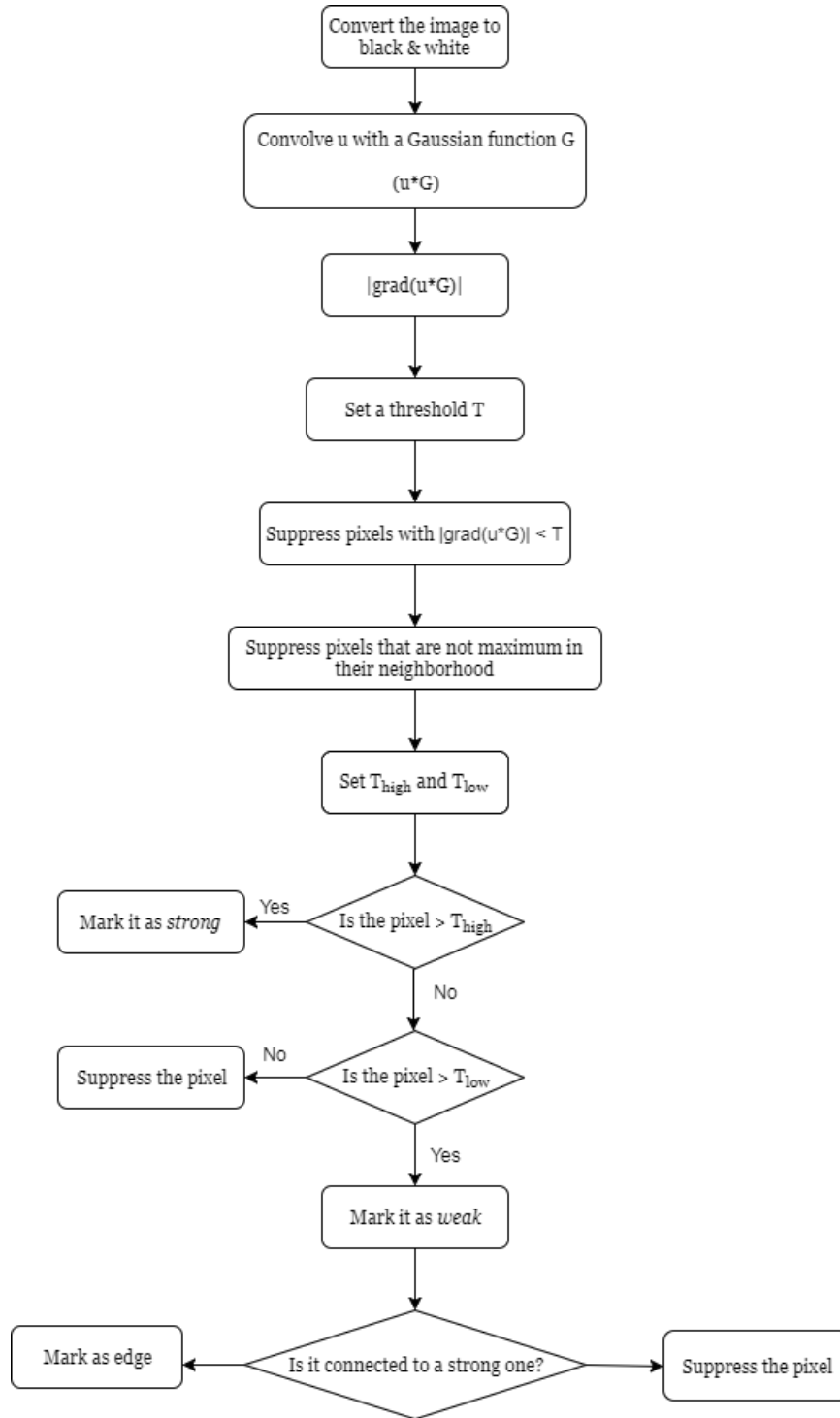
Figure 5.4: Flow chart for Canny edge detector algorithm

Figure 5.4 shows the flow chart of this algorithm.

5.4 Results and Comparison of Images Without Noise

In this section, we compare edge detection by wavelets to that with Canny. We begin by five images without noise. Then we add noise and observe its effect. The first image shows a Pepsi can with a blurred background (Figure 5.5). It is clear that our algorithm outperforms Canny's. Using our algorithm, we only miss some circles with weak boundaries in the background. For Canny algorithm, the true edges are not clear and it produces many false edge pixels.

Figure 5.5: Pepsi can image
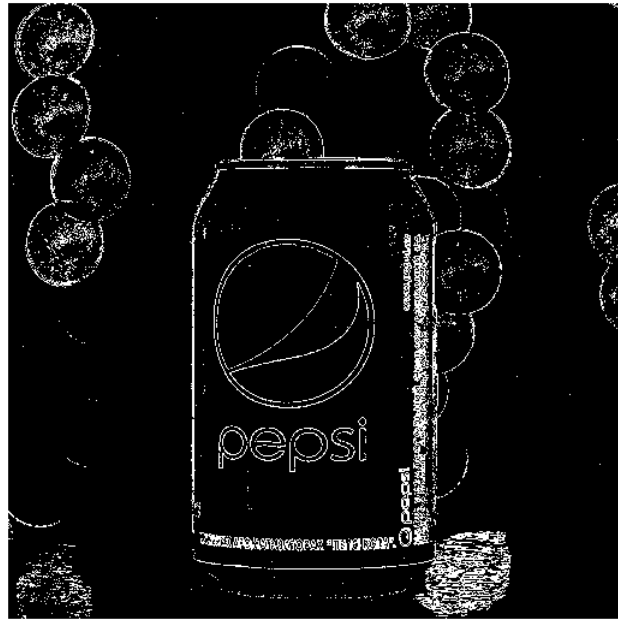
Figure 5.6: Pepsi can edges using dwt2 with Haar

Figure 5.6 shows the Pepsi can edges using Haar wavelet.

Figure 5.7: Pepsi can edges using Canny edge detector

Figure 5.7 shows the Pepsi can edges using Canny edge detector.

The second image that we are going to take is an image of the famous Times Square in New York (Figure 5.8). Looking at the two edge images it is easily seen that our algorithm outperforms Canny's. Using our algorithm, the edges are much clearer and the constituent parts in the image are easy to distinguish. For example, you can easily see the edges of the cars and also read the word "Kodak" written vertically which is not the case with Canny's.

Figure 5.8: Times square image
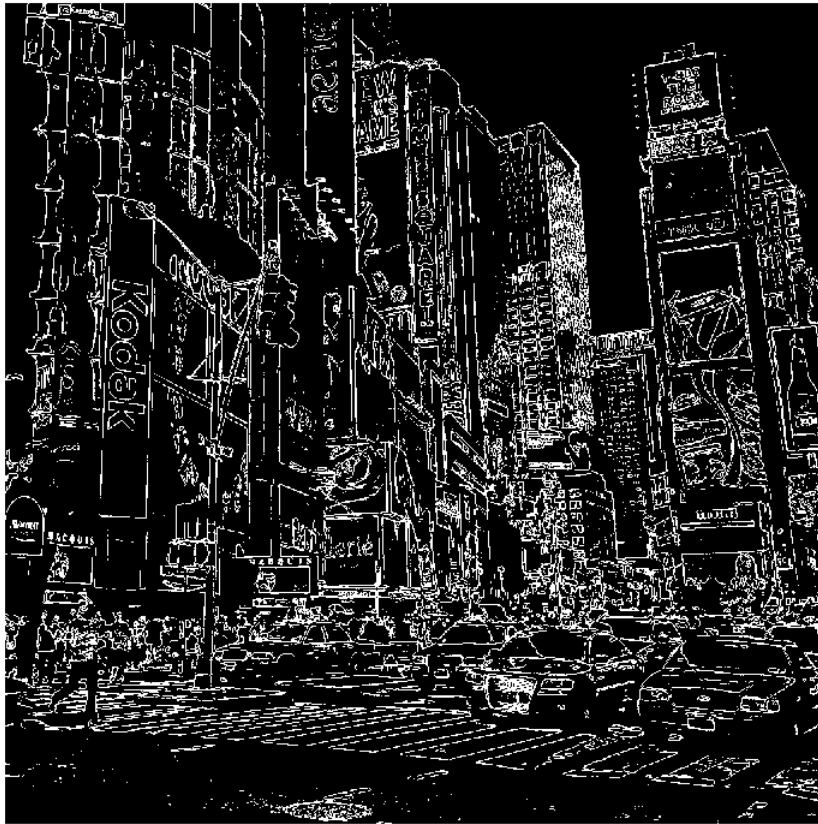
Figure 5.9: Times square edge image using dwt2 with Haar

Figure 5.9 shows the edge image of the Times Square using Haar wavelet.
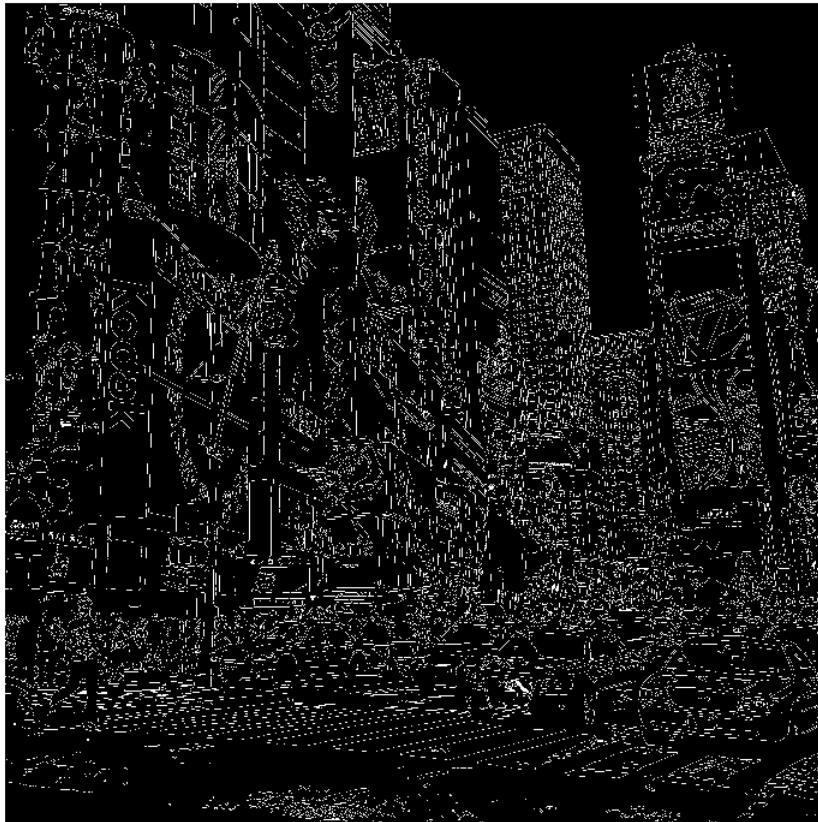
Figure 5.10: Times square edge image using Canny edge detector

Figure 5.10 shows the edge image of the Times Square using Canny edge detector.

The third image is of a bird with a blurred background (Figure 5.11).

Figure 5.11: Bird image

Figure 5.12: Bird edge image using dwt2 with Haar

Figure 5.12 shows the edge image of the bird image using Haar wavelet.

Figure 5.13: Bird edge image using Canny edge detector

Figure 5.13 shows the edge image of the bird image using Canny edge detector.

The fourth image is of a Dr Pepper can. (Figure 5.14).

Figure 5.14: Dr Pepper can image
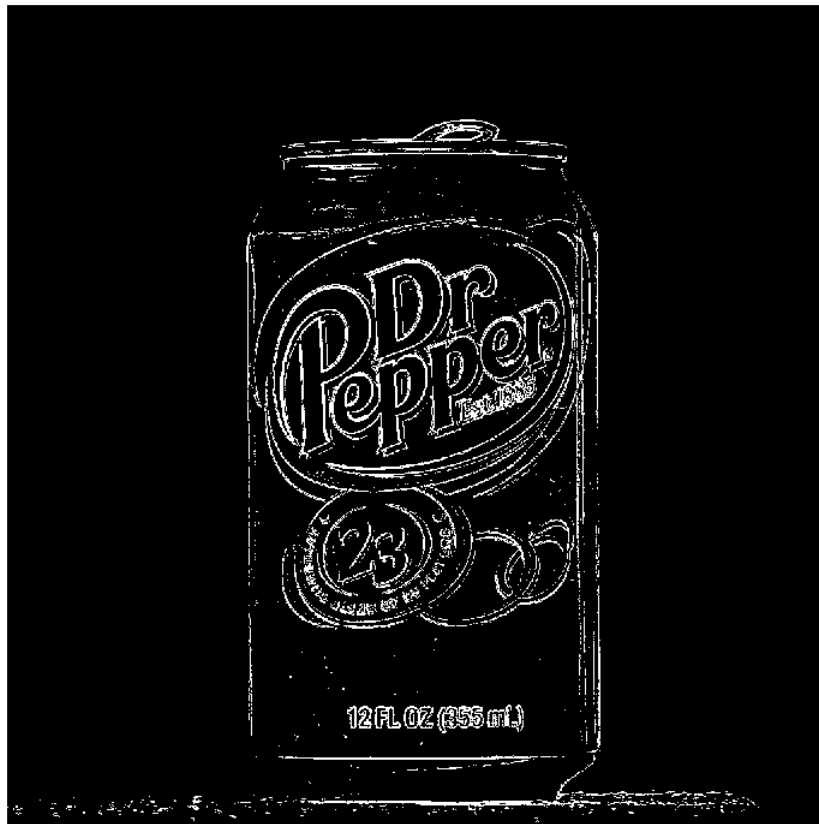
Figure 5.15: Dr Pepper can edge image using dwt2 with Haar

Figure 5.15 shows the edges of the Dr Pepper can image using Haar wavelet.

Figure 5.16: DrPepper can edge image using Canny edge detector

Figure 5.16 shows the edges of the Dr Pepper can image using Canny edge detector.

The fifth image is an image of a toys room (Figure 5.17).

Figure 5.17: Toys room image
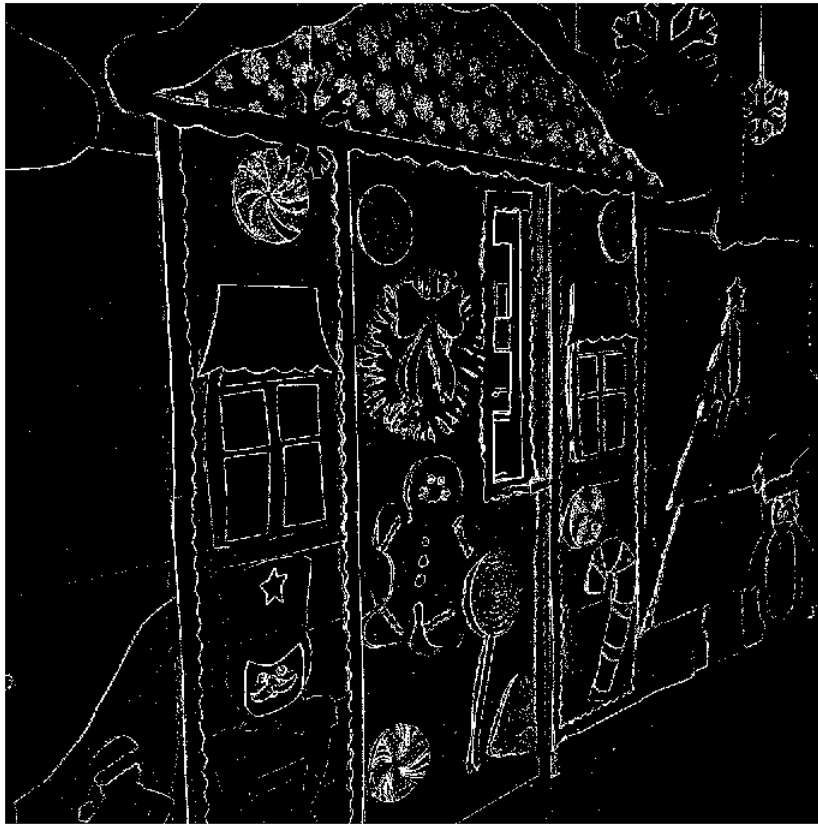
Figure 5.18: Toys room edge image using dwt2 with Haar

Figure 5.18 shows the edges of the toys room image using Haar wavelet.
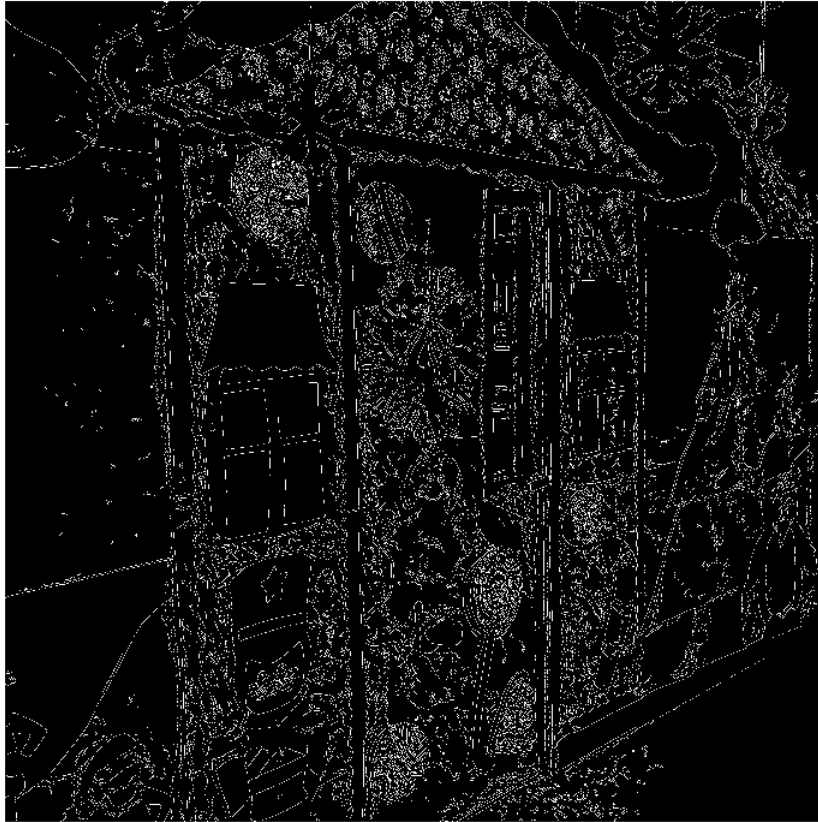
Figure 5.19: Toys room edge image using Canny edge detector

Figure 5.19 shows the edges of the toys room image using Canny edge detector.

5.5 Results and Comparison of Images with Noise

Now we will see the effect of adding the *salt and pepper* noise to the Pepsi can (Figure 5.20) and Times Square (Figure 5.23) images. Here we use the same

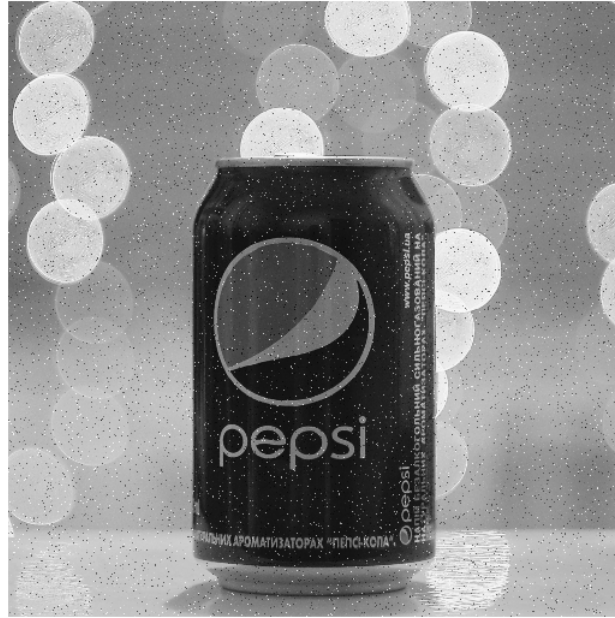algorithm proposed but we lower the threshold to $T_{noise} = \dfrac{m}{20}$, instead of $T = \dfrac{m}{10}$.

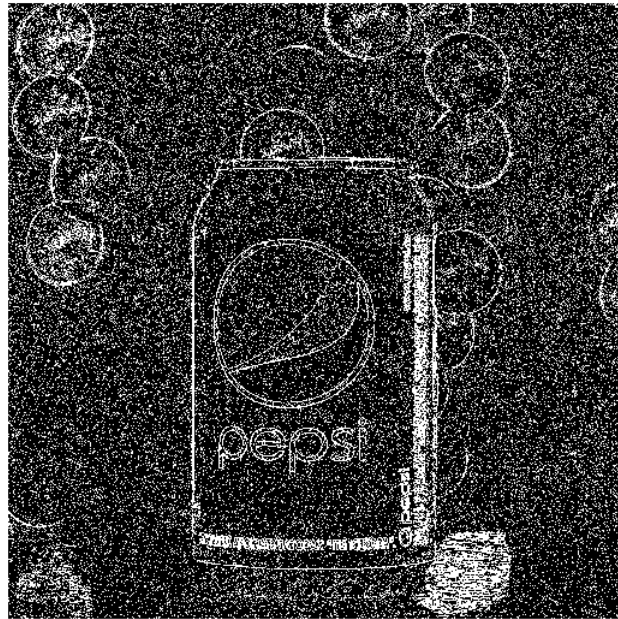Figure 5.20: Noised Pepsi can image

Figure 5.21: Noised Pepsi can edge image using dwt2 with Haar

Figure 5.21 shows the edges of the Pepsi can noised image using Haar wavelet.
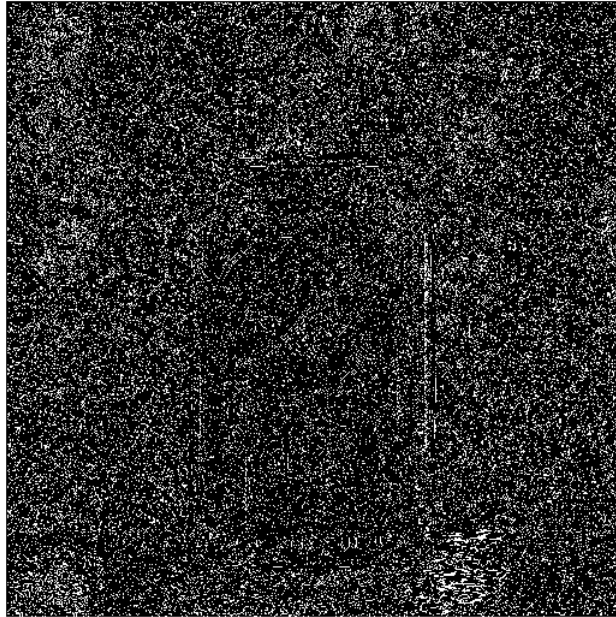
Figure 5.22: Noised Pepsi can edge image using Canny edge detector

Figure 5.22 shows the edges of the Pepsi can noised image using Canny edge detector.

Figure 5.23: Noised Times square image

Figure 5.24: Noised Times square edge image using dwt2 with Haar

Figure 5.24 shows the edges of the Times Square noised image using Haar wavelet.
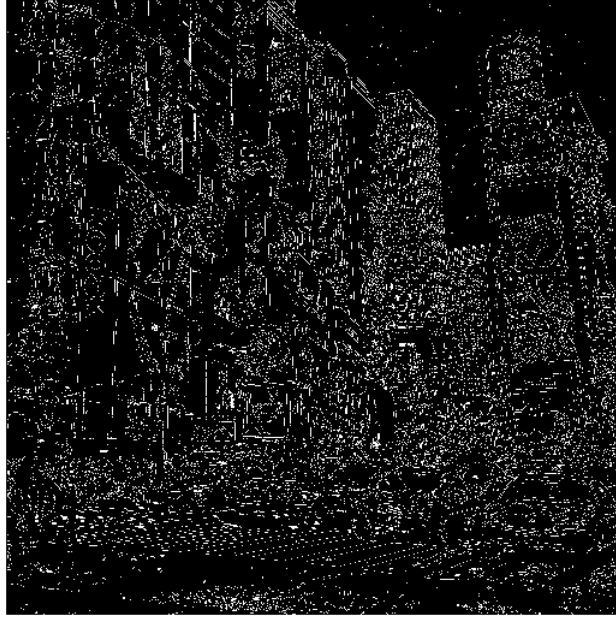
Figure 5.25: Noised Times square edge image using Canny edge detector

Figure 5.25 shows the edges of the Times Square noised image using Canny edge detector.

It is clear that our algorithm is not immune to noise but still it is better than Canny edge detector with respect to noise.

5.6 Comparison of Haar with db2 and db3 in edge detection

In this section, we will apply our algorithm using Haar, db2 and db3 wavelets to the Pepsi can image and compare the edge images in all cases. This section, will

show the distinction of the Haar wavelet in edge detection as we proposed in the beginning of this work. This distinction is due to the better correlation of the Haar wavelet with the characteristic edge jump discontinuity. Figure 5.26 shows the Pepsi can image.
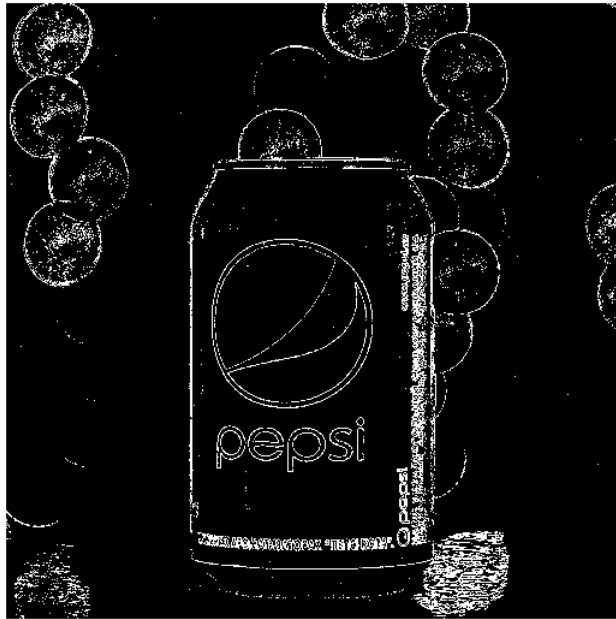

Figure 5.26: Pepsi can image

Figure 5.27: Pepsi can edge image using dwt2 with Haar

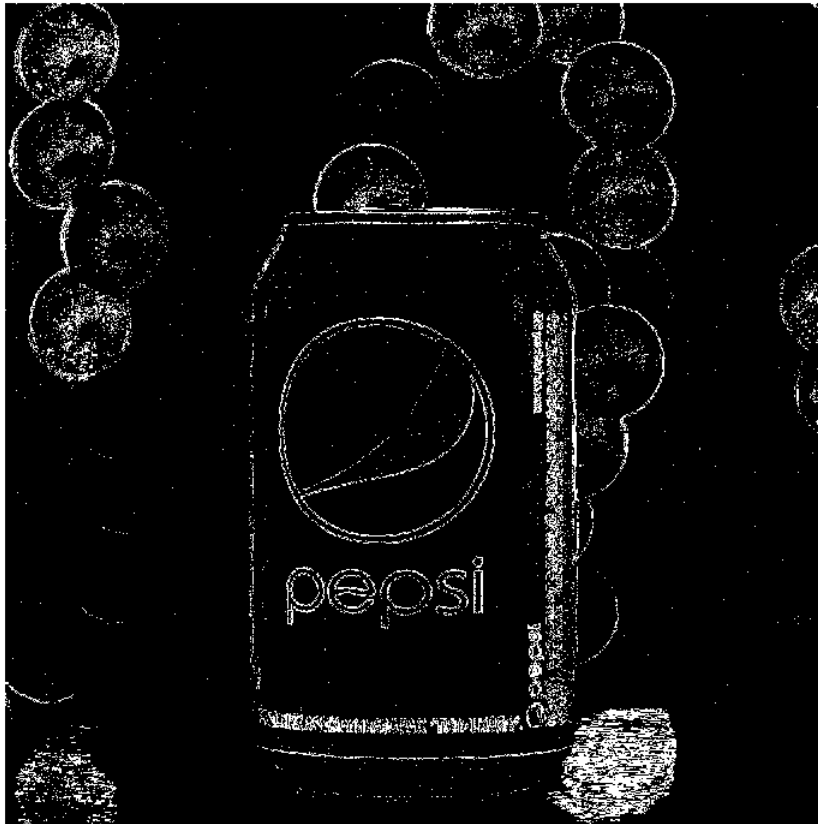Figure 5.27 shows the edges of the Pepsi can image using Haar wavelet.

Figure 5.28: Pepsi can edge image using dwt2 with db2

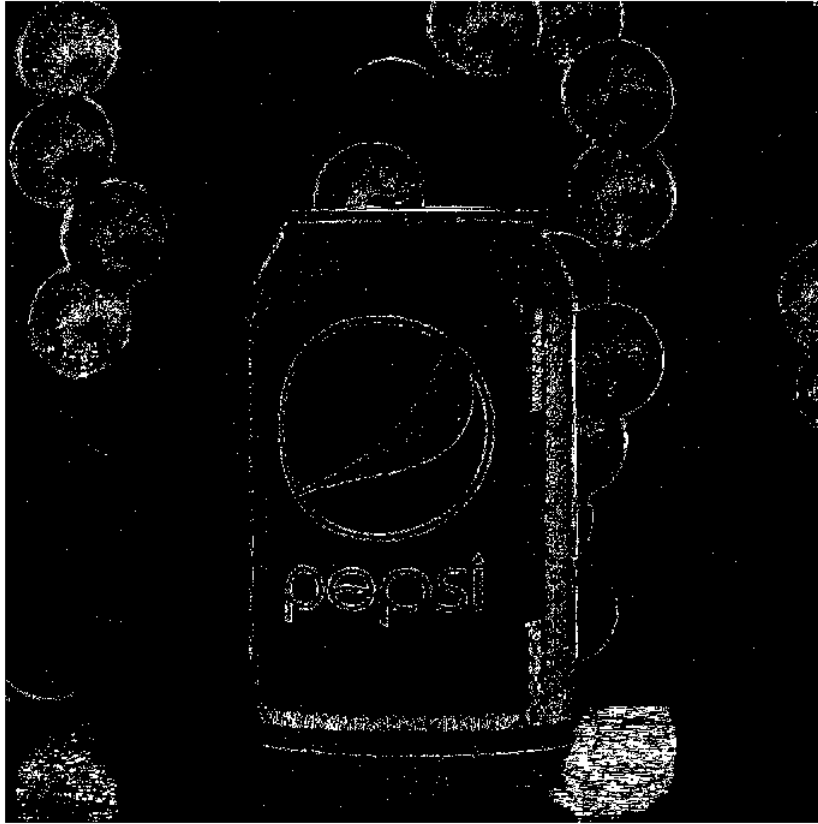Figure 5.28 shows the edges of the Pepsi can image using db2 wavelet.

Figure 5.29: Pepsi can edge image using dwt2 with db3

Figure 5.29 shows the edges of the Pepsi can image using db3 wavelet. It clear that the Haar wavelet is distinct in detecting edges. Moreover, the edge image loses sharpness of edges as we go from db1(Haar) to db2 to db3 an so on.

5.7 Edge Detection In Colored Images

In this section, we will directly apply our algorithm to colored images instead of transforming first to a black-white image. The algorithm will be applied to

the colored Pepsi can image (Figure 5.30) and to the toys room image (Figure 5.33). Figures 5.31 and 5.32 show the edges of the colored Pepsi can image using Haar wavelet and Canny edge detector, respectively. Figures 5.34 and 5.35 show the edges of the toys room image using Haar wavelet and Canny edge detector, respectively. Recall that a colored image can be thought of as a function

$$u : [0, 1] \times [0, 1] \longrightarrow \mathbb{R}^3$$

where

$$u(x, y) = \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$

and r, g, b are the red, green and blue color intensities at the point (x, y). The idea is to apply the edge detection algorithm to each of the rgb channels representing the color intensities at the image pixels.
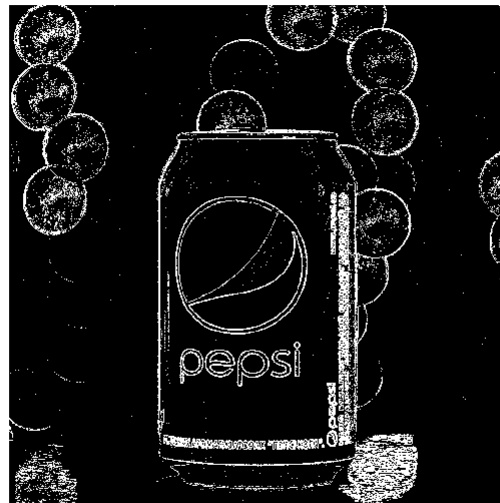
Figure 5.30: Colored Pepsi can image

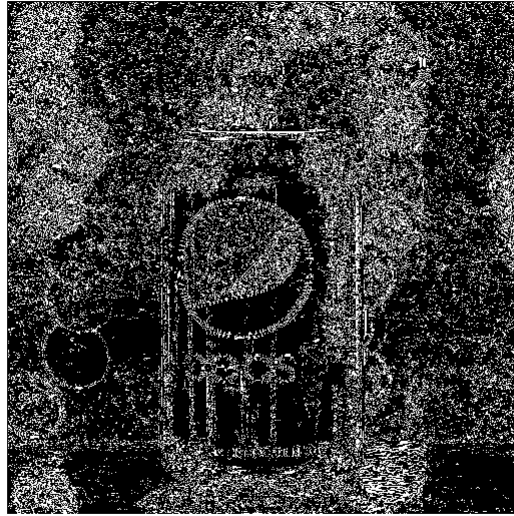Figure 5.31: Colored Pepsi can edge image using dwt2 with Haar

Figure 5.32: Colored Pepsi can edge image using Canny edge detector

Figure 5.33: Colored toys room image

Figure 5.34: Colored toys room edge image using dwt2 with Haar

Figure 5.35: Colored toys room edge image using Canny edge detector

Clearly, from the above images you can see that our algorithm outperforms

the Canny edge detector when colored images are processed instead of black-white images.

Concluding Remarks 5.6

At the end of this work, we can summarize our findings as follows:

- Our algorithm outperforms Canny edge detector when processing high-definition images.

- Our algorithm is much more immunized than Canny edge detector when processing noised high-definition images.

- Our algorithm is performing the best when using the Haar wavelet.

- Our algorithm outperforms Canny edge detector when colored images are processed.

Conclusion 5.7 *fsd*

REFERENCES

[1] Rashmi, M. Kumar, and S. Rohini, "Algorithm and Technique on Various Edge Detection : A Survey," *Signal & Image Processing : An International Journal*, vol. 4, no. 3, pp. 65–75, 2013.

[2] S. Bhardwaj and A. Mittal, "A Survey on Various Edge Detector Techniques," *Procedia Technology*, vol. 4, pp. 220–226, 2012.

[3] M. Ali and D. Clausi, "Using the Canny edge detector for feature extraction and enhancement of remote sensing images," in *IEEE International Geoscience and Remote Sensing Symposium*, vol. 5. IEEE, 2001, pp. 2298–2300.

[4] Sivakumar P and Meenakshi S, "A REVIEW ON IMAGE SEGMENTATION TECHNIQUES," *International Journal of Advanced Research in Computer Engineering & Technology*, vol. 5, no. 3, 2016.

[5] P. A. Hajare and P. A. Tijare, "Edge Detection Techniques for Image Segmentation," *International Journal Of Computer Science And Applications*, vol. 4, no. 1.

[6] R. Maini and H. Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques," *Himanshu Aggarwal International Journal of Image Processing*, vol. 3, no. 1, 2009.

[7] M. Gudmundsson, E. El-Kwae, and M. Kabuka, "Edge detection in medical images using a genetic algorithm," *IEEE Transactions on Medical Imaging*, vol. 17, no. 3, pp. 469–474, jun 1998.

[8] Z. Yu-qian, G. Wei-hua, C. Chen Zhen-cheng, T. Jing-tian, and L. Ling-yun, "Medical Images Edge Detection Based on Mathematical Morphology," in *IEEE Engineering in Medicine and Biology 27th Annual Conference*, vol. 6. IEEE, 2005, pp. 6492–6495.

[9] S. Mallat, *A Wavelet Tour of Signal Processing*, 2nd ed. Academic Press, 1999.

[10] L. S. Davis, "A Survey of Edge Detection Techniques," *Computer Graphics and Image Processing*, vol. 4, no. 3, pp. 248–270, 1975.

[11] L. Roberts, "Machine Perception of Three-Dimensional Solids," Ph.D. dissertation, MIT, 1963.

[12] I. Sobel and G. Feldman, "A 3x3 isotropic gradient operator for image processing," *a talk at the Stanford Artificial Intelligence Project*, 1968.

[13] J. Prewitt, "Object Enhancement and Extraction," in *Picture Processing and Psychopictorics*, B. Lipkin and A. Rosenfeld, Eds. Academic Press, 1970, ch. 1, pp. 75–151.

[14] D. Marr and E. Hildreth, "Theory of edge detection." *Proceedings of the Royal Society of London. Series B, Biological sciences*, vol. 207, no. 1167, pp. 187–217, feb 1980.

[15] J. Canny, "A Computational Approach to Edge Detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, nov 1986.

[16] F. Bergholm, "Edge Focusing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 6, pp. 726–741, nov 1987.

[17] V. Chaganti, "Edge Detection of Noisy Images Using 2-D Discrete Wavelet Transform," Master's thesis, Florida State University, 2005.

[18] D. Mumford and J. Shah, "Optimal approximations by piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, jul 1989.

[19] J.-M. Morel and S. Solimini, "Segmentation of images by variational methods: a constructive approach," 1988.

[20] S. Solimini and J. M. Morel, "Variational methods in image segmentation," *Progress in Nonlinear Differential Equations and Their Applications*, 1995.

[21] U. Massari and I. Tamanini, "Regularity properties of optimal segmentations," *Journal fur die reine und angewandte Mathematik*, vol. 420, pp. 61–84, 1991.

[22] S. Mallat and S. Zhong, "Characterization of Signals from Multiscale Edges," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 7, pp. 710–732, jul 1992.

[23] R. Monika, "Image Edge Detection using Discrete Wavelet Transform," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 4, no. 4, pp. 169–175, 2016.

[24] H. L. Royden and P. M. Fitzpatrick, *Real Analysis*, 4th ed. Pearson, 2010.

[25] D. F. Walnut, *An introduction to wavelet analysis*, 2nd ed. Birkhauser, 2004.

[26] A. Haar, "Zur theorie der orthogonalen funktionensysteme," *Mathematische Annalen*, vol. 69, no. 3, pp. 331–371, 1910.

[27] R. Islam, "Approximation By Wavelets," Master's thesis, KFUPM, 2005.

[28] K. Urban, *Wavelet methods for elliptic partial differential equations.* Oxford University Press, 2009.

[29] A. Chambolle, R. A. De Vore, N.-Y. Lee, and B. J. Lucier, "Nonlinear wavelet image processing: variational problems, compression, and noise removal through wavelet shrinkage," *IEEE Transactions on Image Processing*, vol. 7, no. 3, pp. 319–335, 1998.

Vitae

- Name: Azzam Abdulaziz Alfarraj

- Nationality: Saudi

- Date of Birth: 21, Feb 1992

- Email: *azzamfr@gmail.com*

- Permenant Address: P.O. Box 450, KFUPM, Dhahran 31261, Saudi Arabia