

# HAAR WAVELET AND ITS APPLICATION IN IMAGE COMPRESSION

JASON NGO

## 1. INTRODUCTION

Since the start of the 20th century, we have seen rapid development in the theory and applications of wavelets. As a mathematical tool, wavelets can be used to extract information from different kinds of data such as audio signals and images. As an attempt to explore wavelets at an introductory level, this paper will examine the first wavelet developed, called the Haar Wavelet and discuss its applications in image compression.

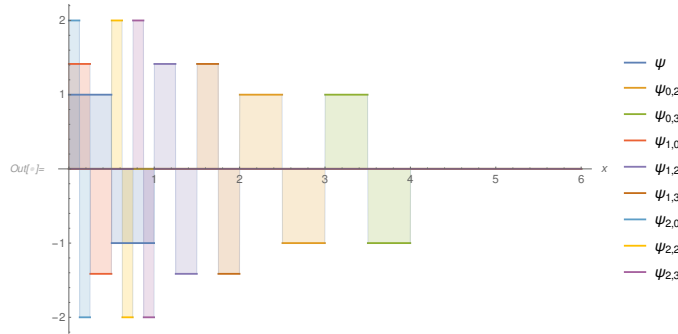


FIGURE 1. Elements of the Haar system

Wavelets were introduced relatively recently, in the beginning of the 1980s.

**1.1. Motivation for Wavelets.** “Develop an important variation on Fourier series, replacing the sine and cosine functions with new families of functions, called wavelets”. Want: good properties of trig functions (orthogonal basis, invariant under translation, dilation (??)).

Fourier coefficients, and hence the Fourier series approximation, depend on all values of the function. For example, if you change a function  $f$  a small amount on the interval  $[0; 0.01]$ , it is possible that every Fourier coefficient changes. This will then have an effect on the partial sums  $S_n f(\theta)$  for all

Date: 2019-05-03.

fix introduction, maybe mention where fourier series is failing and then the motivation for our wavelets.

maybe include a little bit of history of the Haar wavelet

values of  $\theta$ . Although these changes may be small, there are many subtleties in analyzing Fourier series approximations, as we have seen.

This suggests looking for a series expansion with better local properties, meaning that coefficients reflect the local behaviour of the function and a small change on one interval affects only a few of the series coefficients and leaves unchanged the partial sums elsewhere in the domain.

## 2. DEFINITION

To get a better sense of our goals, recall that  $L^2(\mathbb{R})$  is a vector space of square integrable functions, taken with the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}} f(x)g(x)dx.$$

**Definition 2.1** ([2, p. 515]). The *Haar function* is the function  $\Psi = \chi_{[0,0.5)} - \chi_{[0.5,1)}$ . The *Haar system* is the family

$$\{\Psi_{j,k}(x) = 2^{j/2}\Psi(2^j x - k) \mid j, k \in \mathbb{Z}\}.$$

Note that each term in the Haar system is constructed by translating and/or dilating the original  $\Psi$  (see Figure 1). By this construction, since  $\int_{\mathbb{R}} \Psi(x)dx = 0$ , it follows that  $\int_{\mathbb{R}} \Psi_{j,k}(x)dx = 0$  for  $j, k \in \mathbb{Z}$ . Furthermore, it is important to calculate the width and height for each  $\Psi_{j,k}$ . We have  $\Psi(2^j - k) = 1$  when  $0 \leq 2^j - k < \frac{1}{2}$ , or equivalently,  $\frac{k}{2^j} \leq x < \frac{k+\frac{1}{2}}{2^j}$ . Therefore, in this interval,  $\Psi_{j,k} = 2^j$ . Similar calculations yield:

$$(2.1) \quad \Psi_{j,k} = \begin{cases} 2^{j/2} & \text{if } \frac{k}{2^j} \leq x < \frac{k+\frac{1}{2}}{2^j}, \\ -2^{j/2} & \text{if } \frac{k+\frac{1}{2}}{2^j} \leq x < \frac{k+1}{2^j}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, we will prove that the Haar function  $\Psi$  satisfies Definition 2.2 of an *orthonormal wavelet*.

**Definition 2.2** (Orthonormal Wavelet, [4, p. 303]). If  $\Psi \in L^2(\mathbb{R})$ ,  $\Psi_{j,k}(x) = 2^{j/2}\Psi(2^j x - k)$ , and the set  $\{\Psi_{j,k} : j, k \in \mathbb{Z}\}$  is an orthonormal basis for  $L^2(\mathbb{R})$ , then  $\Psi$  is called an *orthonormal wavelet*.

To prove the Haar system is an *orthonormal basis*, we have to show:

- (1) The Haar system  $\{\Psi_{j,k}\}$  is an orthonormal set, i.e.  $\|\Psi_{j,k}\| = 1$  for all  $j, k \in \mathbb{Z}$  and  $\langle \Psi_{j,k}, \Psi_{j',k'} \rangle = 0$  for all  $(j, k) \neq (j', k')$ .
- (2) The span of Haar system, denoted  $\text{span}\{\Psi_{j,k}\}$ , is dense in  $L^2(\mathbb{R})$ .

In Section 3, we will first prove that the Haar system is orthonormal. Section 4 will show that it spans a dense subspace of  $L^2(\mathbb{R})$ . Finally, in Section ??, we will discuss the Haar wavelet's local property and its application in image compression.

### 3. ORTHONORMALITY

To get started, building upon the terse proof in [2], this section proves that the Haar system is orthonormal, followed by a short discussion of the normalization factor.

**Lemma 3.1** ([2, p. 409]). *The Haar system is an orthonormal set in  $L^2(\mathbb{R})$ .*

*Proof.* First, it is easy to see that  $\|\Psi\| = 1$ . Let us compute:

$$\begin{aligned}\|\Psi_{j,k}\|^2 &= \int_{\mathbb{R}} \left| 2^{j/2} \Psi(2^j x - k) \right|^2 dx = \int_{\mathbb{R}} 2^j |\Psi(2^j x - k)|^2 dx \\ &= \int_{\mathbb{R}} |\Psi(y)|^2 dy = \|\Psi\|^2,\end{aligned}$$

by change of variable  $y = 2^j x - k$  in the integral. Therefore, all the elements in the Haar system has norm 1.

Now, we want to show the orthogonality. Consider  $\Psi_{j,k}$  and  $\Psi_{j',k'}$ , for  $k \neq k'$ . Since these two elements have disjoint support<sup>1</sup>, their inner product is 0. If  $j < j'$ , then either  $\Psi_{j,k}$  and  $\Psi_{j',k'}$  have disjoint supports (when  $k \neq k'$ ), or  $\text{supp } \Psi_{j',k'}$  is contained in an interval on which  $\Psi_{j,k}$  is constant to  $2^{j/2}$  (when  $k = k'$ ). For the former case, their inner product is obviously 0; for the latter case, we have:

$$\langle \Psi_{j,k}, \Psi_{j',k'} \rangle = \int_{\mathbb{R}} 2^{j/2} \cdot \Psi_{j',k'} dx = 2^{j/2} \int_{\mathbb{R}} \Psi_{j',k'} dx = 0.$$

Hence, the set is orthonormal.  $\square$

*Remark 3.2.* The Haar function is a dyadic function, meaning that the dilations are taken to be powers of 2. Here, the factor  $2^{j/2}$  is called the normalization factor, which is there so that the dilated and translated Haar function has norm 1. Although powers of 2 are a common choice, they are certainly not the most general one<sup>2</sup>.

### 4. HAAR WAVELET BASIS FOR $L^2(\mathbb{R})$

Now that we have shown the Haar system is orthonormal, this section will finish showing  $\Psi$  satisfies Definition 2.2 by proving the following theorem:

**Theorem 4.1** ([2, p. 411]). *The Haar system  $\{\Psi_{j,k}\}$  spans all of  $L^2(\mathbb{R})$ .*

To prove this theorem, we will start out with  $C_c(\mathbb{R})$ , the set of continuous functions with compact support. Since  $C_c(\mathbb{R})$  is dense in  $L^2(\mathbb{R})$ <sup>3</sup>, if we can

<sup>1</sup>The *support* of a function, denoted  $\text{supp}$ , is the subset of the domain containing those elements which are not mapped to zero.

<sup>2</sup>See examples of non-dyadic wavelets and a discussion why they are more appropriate for statistical data analysis in [5].

<sup>3</sup>The proof that  $C_c(\mathbb{R})$  is dense in  $L^2(\mathbb{R})$  can be found in [3, p. 326].

show that any function  $f \in C_c(\mathbb{R})$  is in the closure of the span of  $\{\Psi_{j,k}\}$ , then it follows that  $\{\Psi_{j,k}\}$  spans  $L^2(\mathbb{R})$ .

Motivated by [1, p. 3] and [2, p. 516], we will set up the integral transform  $P_n f$  in Section 4.1, and then proceed to define the Haar inner product expansion in Section 4.2, which indeed can approximate any function in  $C_c(\mathbb{R})$ .

**4.1. Integral Transform.** Set  $\phi(x) = \chi_{[0,1)}$ . For  $n \in \mathbb{Z}$ , we define the integral kernel

$$K_n(x, y) = 2^n \sum_{k \in \mathbb{Z}} \phi(2^n x - k) \phi(2^n y - k).$$

Note that  $K_n$  is either equal to  $2^n$  (when there is some  $k \in \mathbb{Z}$  s.t.  $2^n x - k$  and  $2^n y - k$  lie in  $[0, 1)$ ) or equal to 0 (otherwise). In other words, if  $K_n = 2^n$ , then there is some  $k$  s.t.

$$(4.1) \quad x, y \in \left[ \frac{k}{2^n}, \frac{k+1}{2^n} \right) = I_{n,k}.$$

Here, note that each  $I_{n,k}$  has length  $2^{-n}$  and that the sets  $\{I_{n,k}\}$  are disjoint.

We define the integral transform

$$P_n f(x) = \int_{\mathbb{R}} K_n(x, y) f(y) dy.$$

If  $x \in \mathbb{R}$  then there is a unique  $k_x \in \mathbb{Z}$  with  $x \in I_{n,k_x}$  and

$$(4.2) \quad P_n f(x) = 2^n \int_{I_{n,k_x}} f(y) dy.$$

Since  $I_{n,k_x}$  has length  $2^{-n}$ , we can say that  $P_n f$  is the average value of a function  $f$  on the dyadic interval containing  $x$ .

Notice that the integral kernel  $K_n$  is dependent on  $\phi$ . In order to write any function in terms of the Haar system  $\{\Psi_{j,k}\}$ , we must eliminate  $\phi$  somehow from our calculations. The following Lemma 4.2 will write the integral kernels in terms of  $\{\Psi_{j,k}\}$ .

**Lemma 4.2** ([4, p. 293]). *If  $n \in \mathbb{Z}$ , then*

$$K_{n+1} - K_n = \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x) \Psi_{n,k}(y), \quad x, y \in \mathbb{R}.$$

*Proof.* First, let us calculate the right-hand side. By Equation (2.1),  $\Psi_{n,k}(x) = 2^{n/2}$  when  $x \in I_{n+1,2k} = \left[ \frac{2k}{2^{n+1}}, \frac{2k+1}{2^{n+1}} \right)$ , and similarly,  $\Psi_{n,k}(x) = -2^{n/2}$  when

$x \in I_{n+1,2k+1}$ . Therefore, it follows that:

$$\Psi_{n,k}(x)\Psi_{n,k}(y) = \begin{cases} 2^n & \text{if } x, y \in I_{n+1,2k} \text{ or } x, y \in I_{n+1,2k+1}, \\ -2^n & \text{if one is in } I_{n+1,2k} \text{ and the other is in } I_{n+1,2k+1}, \\ 0 & \text{otherwise.} \end{cases}$$

Now, before calculating the left-hand side  $K_{n+1} - K_n$ , note that:

$$I_{n+1,2k} \cup I_{n+1,2k+1} = \left[ \frac{2k}{2^{n+1}}, \frac{2k+2}{2^{n+1}} \right) = I_{n,k}.$$

Therefore, if both  $x, y$  are in  $I_{n+1,2k}$  or  $I_{n+1,2k+1}$ , then we have (1)  $K_{n+1} = 2^{n+1}$  (by Equation (4.1)) and (2)  $K_n = 2^n$  (because  $x, y \in I_{n,k}$ ), Hence,

$$K_{n+1}(x, y) - K_n(x, y) = 2^{n+1} - 2^n = 2^n.$$

By similar argument, if either  $x$  or  $y$  is in  $I_{n+1,2k}$  and the other is in  $I_{n+1,2k+1}$ , then  $K_{n+1}(x, y) - K_n(x, y) = 0 - 2^n = -2^n$ . Otherwise, both  $K_{n+1}$  and  $K_n$  are 0.

Thus, we conclude  $K_{n+1}(x, y) - K_n(x, y) = \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x)\Psi_{n,k}(y)$ .  $\square$

*Remark 4.3.* From this Lemma, the difference between two integral transforms can be written as the *inner product expansion* of the function  $f$ :

$$\begin{aligned} (P_{n+1} - P_n)f(x) &= \int_{\mathbb{R}} K_{n+1}(x, y)f(y)dy - \int_{\mathbb{R}} K_n(x, y)f(y)dy \\ &= \int_{\mathbb{R}} \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x)\Psi_{n,k}(y)f(y)dy \\ &= \sum_{k \in \mathbb{Z}} \Psi_{n,k}(x) \int_{\mathbb{R}} \Psi_{n,k}(y)f(y)dy \\ (4.3) \quad &= \sum_{k \in \mathbb{Z}} \langle f, \Psi_{n,k} \rangle \Psi_{n,k}(x). \end{aligned}$$

The terms  $\langle f, \Psi_{j,k} \rangle$  are called the the Haar coefficients. One can easily notice that this inner product expansion is only one-sided, meaning we keep the dilation factor  $j = n$  fixed and vary the translation factor  $k$ . In the next section, we will abstract this expansion to a more general setting, varying both the translation and dilation factors.

**4.2. Two-sided Haar Representation.** From Equation 4.3, we can write

$$(4.4) \quad P_{n+1}f(x) - P_{-m}f(x) = \sum_{j=-m}^n \sum_{k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi_{j,k}(x).$$

To prove that any function  $f \in C_c(\mathbb{R})$  is in the closure of the span of  $\{\Psi_{j,k}\}$ , we want to show that the Haar inner product expansion on the right-hand

side converges to  $f$ . Fortunately, we can do so by showing that  $P_{-m}f \rightarrow 0$  and  $P_{n+1}f \rightarrow f$  with respect to the  $L^2$  norm as  $m, n \rightarrow \infty$ . The two claims are presented as Lemmas below.

First, Lemma 4.4 tells us that the larger the intervals on which we average a continuous function, the smaller the supremum of the averaged function.

**Lemma 4.4** (Average Function over Dyadic Intervals, [4, p. 295]). *If  $f \in C_c(\mathbb{R})$ , then  $\|P_{-m}f\|_2 \rightarrow 0$  when  $m \rightarrow \infty$ .*

*Proof.* Suppose  $f \in C_c(\mathbb{R})$  and  $\text{supp}(f) \subseteq [-K, K]$ . There exists  $M > 0$  s.t.  $\text{supp}(f) \subset [-2^M, 2^M]$ . Now, if  $m > M$ , then  $\text{supp}(f) \subseteq I_{-m,-1} \cup I_{-m,0}$ ; Equation (4.2) therefore holds. We compute:

$$\begin{aligned} \|P_{-m}f\|_2^2 &= \int_{\mathbb{R}} |P_{-m}f(x)|^2 dx = \sum_{k \in \mathbb{Z}} \int_{I_{-m,k}} |P_{-m}f(x)|^2 dx \\ &= \sum_{k \in \mathbb{Z}} \int_{I_{-m,k}} \left| 2^{-m} \int_{I_{-m,k}} f(y) dy \right|^2 dx = 2^m 2^{-2m} \sum_{k \in \mathbb{Z}} \left| \int_{I_{-m,k}} f(y) dy \right|^2 dx. \end{aligned}$$

Since  $f$  is only supported on a subset of  $I_{-m,-1}$  and  $I_{-m,0}$ , we have:

$$\begin{aligned} \|P_{-m}f\|_2^2 &= 2^{-m} \left( \left| \int_{I_{-m,-1}} f(y) dy \right|^2 dx + \left| \int_{I_{-m,0}} f(y) dy \right|^2 dx \right) \\ &= 2^{-m} \left( \left| \int_{-K}^0 f(y) dy \right|^2 dx + \left| \int_0^K f(y) dy \right|^2 dx \right). \end{aligned}$$

Note that  $\int_{-K}^0 f(y) dy = \langle f(y), \chi_{[-K,0]} \rangle$ ; by Cauchy-Schwarz inequality,

$$\|P_{-m}f\|_2^2 \leq 2^{-m} (K\|f\|_2^2 + K\|f\|_2^2) = 2^{-m} \cdot 2K\|f\|_2^2.$$

As  $m \rightarrow \infty$ , the right-hand side converges to 0. Therefore, by Squeeze Theorem, it follows that  $\|P_{-m}f\|_2 \rightarrow 0$ .  $\square$

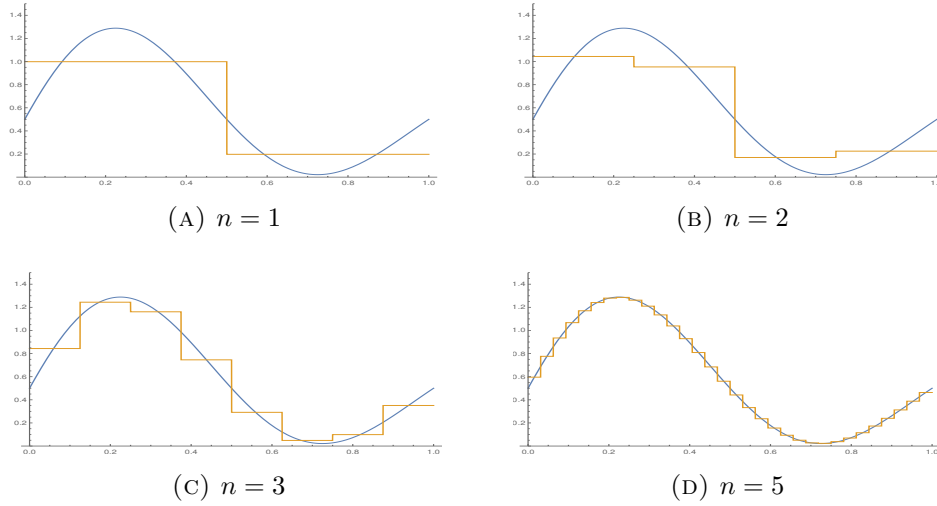
Now, before proving our next Lemma, recall that  $P_n f(x)$  is equal to the average value of a function  $f$  on a small dyadic interval of length  $2^{-n}$ . Therefore, if we increase  $n$ , then the intervals will get smaller, and the average value  $P_n f$  will get closer to  $f$  (see Figure 2 for the sequence  $\{P_n f\}$  approaching  $f(x) = e^{-x} \sin 2\pi x$  on the unit interval).

**Lemma 4.5** (Convergence of Integral Transform, [2, p. 517]). *If  $f \in C_c(\mathbb{R})$ , then  $P_n f \rightarrow f$  in the uniform norm as well as in the  $L^2$  norm.*

*Proof.* Suppose  $f \in C_c(\mathbb{R})$  and  $\text{supp}(f) \subset [-2^M, 2^M]$  for  $M \geq 0$ . Since  $f$  is continuous on a compact set, it is uniformly continuous.

Given  $\epsilon > 0$ . By uniform continuity, there exists  $\delta > 0$  such that if  $|x - y| < \delta$ , then  $|f(x) - f(y)| < \epsilon$ .

This Lemma is particularly useful in signal denoising: given a time series, we

FIGURE 2. Approximation of  $f(x) = e^{-x} \sin 2\pi x$  by  $P_n f(x)$ 

Now, choose  $N$  such that  $2^{-N} < \delta$ ; if  $n > N$ , then we have:

$$\begin{aligned}
 |P_n f(x) - f(x)| &= \left| 2^n \int_{I_{n,k_x}} f(y) dy - f(x) \right| \\
 &= \left| 2^n \int_{I_{n,k_x}} f(y) dy - 2^n \int_{I_{n,k_x}} f(x) dy \right| \\
 &\leq 2^n \int_{I_{n,k_x}} |f(y) - f(x)| dy < 2^n \int_{I_{n,k_x}} \epsilon dy = \epsilon.
 \end{aligned}$$

Therefore,  $P_n f$  converges to  $f$  uniformly. Since uniform convergence implies  $L^2$  convergence, it follows that  $P_n f \rightarrow f$  with respect to the  $L^2$  norm.  $\square$

Now, by Lemma 4.4 and Lemma 4.5, it follows that for any  $f \in C_c(\mathbb{R})$ ,

$$f = \sum_{j \in \mathbb{Z}} \sum_{k \in \mathbb{Z}} \langle f, \Psi_{j,k} \rangle \Psi_{j,k}.$$

Since  $C_c(\mathbb{R})$  is a dense subspace of  $L^2(\mathbb{R})$ , we conclude that the Haar system spans  $L^2(\mathbb{R})$ . The Haar function  $\Psi$  is therefore an orthonormal wavelet.  $\square$

## 5. TITLE

**GRAPH THE COEFFICIENTS** A good feature of the Haar wavelets can be noticed. By increasing the wavelet number, the wavelet coefficients rapidly decrease and higher coefficients are practically zero.  $\therefore$  can confine to a small number of terms in the wavelet series.

Tie back to the motivation at the beginning of the paper

maybe i can compare and contrast haar wavelet basis versus fourier series. draw some more graphs and that's it. to see how much more useful haar wavelet basis is.

pls fix

## REFERENCES

- [1] Jordan Bell. “Haar wavelets and multiresolution analysis”. 2014.
- [2] Kenneth R. Davidson and Allan P. Donsig. *Real Analysis with Real Applications*. Prentice Hall, 2002. ISBN: 0-13-041647-9.
- [3] RH Farrell. “Dense Algebras of Functions in  $L^p$ ”. In: *Proceedings of the American Mathematical Society* 13.2 (1962), pp. 324–328.
- [4] Mark A Pinsky. *Introduction to Fourier analysis and wavelets*. Vol. 102. American Mathematical Soc., 2008.
- [5] Stephen Pollock and Iolanda Lo Cascio. “Non-dyadic wavelet analysis”. In: *Optimisation, Econometric and Financial Analysis*. Springer, 2007, pp. 167–203.