# Chapter 2

# Mathematical Background

This chapter deals with the necessary mathematics. It introduces the Markov Random Field (MRF) and the Wavelets. First part is devoted to the MRF, and the second part deals with wavelets.

## 2.1 Markov Random Field

The MRF branch of probability theory provides a foundation for the characterization of contextual constraints and the densities of the probability distributions of interaction features in images. Along with the methods from decision and estimation theory, MRF theory provides a systematic approach for deriving optimality criteria such as those based on the *maximum a posteriori* (MAP) concept. In recent years, the MRF-MAP framework has enabled a systematic development of algorithms for a variety of vision tasks using rational principles rather than an *ad-hoc* heuristic. We give a brief outline of MRF in this chapter. More information can be found in Li [78].

The pioneering work of extending the one-dimensional MRF to higher dimensions is due to Dorbushin [24], Woods [151, 152] and Wong [29]. In recent years, statistical methods for image processing is drawing attention of researchers. Many of these studies involve the use of MRF models and processing techniques based on these models [17, 44, 40, 22, 68, 46, 25]. Chellappa and Kashyap [19] have extended Markovian property to 2-D and include the simultaneous auto-regressive (SAR) models and the conditional Markov models.

MRF as a spatial interaction model was not exploited completely till the discovery of the equivalence between MRF and the Gibbs distribution (GD). With the GD, the difficulties involved in dealing with conditional distributions in MRF is eliminated as

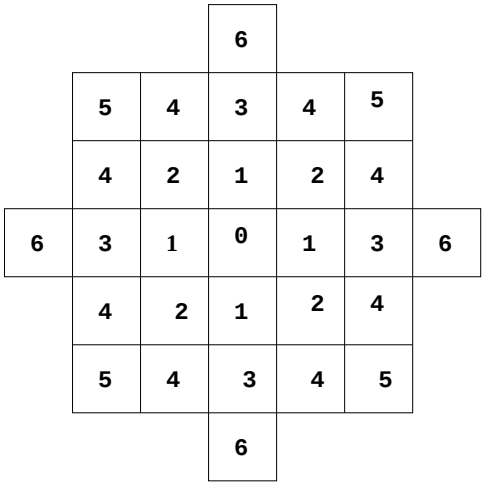


Figure 2.1: A neighborhood system

the joint distribution is readily available with the GD, and the local characteristics are easily obtained from the joint distribution. Besag [9] laid the groundwork for the GD characterization of MRFs.

### 2.1.1 Basic Framework

Let $L = \{(i,j); 1 \leq i \leq M; 1 \leq j \leq M\}$ denote a square lattice of size $M \times M$. The ordered pair $(i,j)$ is called the $ij^{th}$ site of this lattice.

**Definition 2.1:** *Neighborhood system:*
A collection of subsets of $L$, $\eta = \{\eta_{i,j}; i,j \in L, \eta_{i,j} \subseteq L\}$ is a neighborhood system on $L$ if and only if $(i,j) \notin \eta_{i,j}$ and if $(i,j) \in \eta_{k,l}$ then $(k,l) \in \eta_{i,j}$ for any $(i,j) \in L$

Here, $\eta_{i,j} \subseteq L$ denotes the set of all sites which are neighbors of the site $(i,j)$. A hierarchically ordered sequence of a neighborhood system that is commonly used in image processing applications is $\eta_1\eta_2\eta_3\eta_4 \ldots$. For example, $\eta_1$ is the neighborhood system that is formed using the four closest neighbors of each pixel, that is $\eta_{i,j} = (i,j-1),(i,j+1),(i-1,j),(i+1,j)$. This is known as the nearest neighbor, or more commonly, the $1^{st}$ order model. This may have to be adjusted appropriately at the boundaries. Fig. 2.1 shows the neighborhoods. The numbers indicate the order of the neighborhood, including and up to that order. In general, a neighborhood system denoted by $\eta_k$ is called the $k^{th}$ *order neighborhood*. A neighborhood structure is said to be symmetric if $\forall i,j,(i,j) \in S,(i,j) \in \eta_{k,l}$ if and only if $(k,l) \in \eta_{i,j}$.

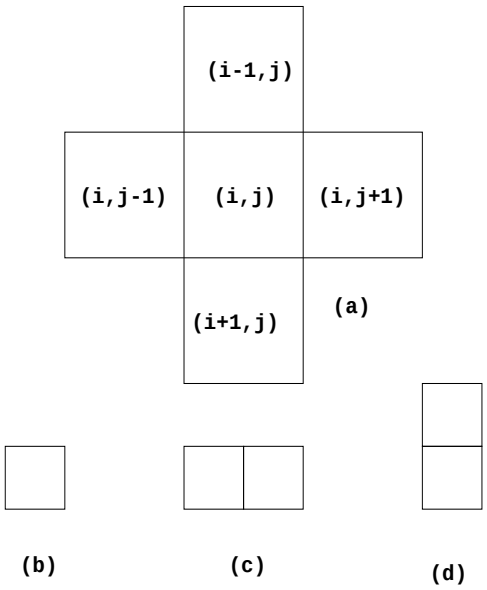


Figure 2.2: First order cliques

### 2.1.2 Clique

**Definition 2.2** *Clique:*
A clique is defined as a set of sites which either consists of a single site or multiple sites such that every site is a neighbor of every other site in the set

Putting it in another way, a clique is a subset $S$ such that

1. $c$ consists of a single pixel, or,

2. for every $(i,j) \neq (k,l)$ if $(i,j) \in \mathcal{C}$ then $(i,j) \in \eta_{i,j}$

where $\mathcal{C}$ denotes set of all cliques. Thus, for the nearest neighbor structure or the $1^{st}$ order system, cliques are of the form $\{(i,j)\}, \{(i,j-1),(i,j)\}$ and $\{(i-1,j),(i,j)\}$. Cliques corresponding to the first order neighborhood are depicted in Fig. 2.2. The cliques corresponding to second order neighborhood structure are shown in Fig. 2.3.

Now, let $X = \{X_{i,j}, i,j \in S\}$ denote a family of random variables indexed by $i,j$, where each $X_{i,j} \in G$ and $G = \{0,1,\ldots,n-1\}$ and $S$ representing a subset. Let the sample space $\kappa$ denote all possible realizations, and $x \in \kappa$ denote one such realization i.e., $x = \{X_{i,j} = x_{i,j}; (i,j) \in (M \times M)\}$. Then, $X$ is an MRF with the neighborhood system $\eta$, if the following conditions are satisfied [9, 36]:

1. *Positivity:* $P(X = x) > 0 \;\; \forall x \in \kappa$

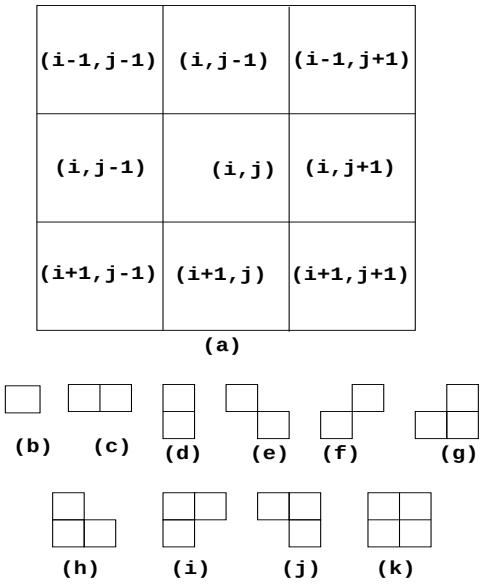


Figure 2.3: Second order cliques

2. *Local Characteristics:* $P(X_{i,j} = x_{i,j}|X_{k,l} = x_{k,l}), (k,l) \neq (i,j) = P(X_{i,j} = x_{i,j}|X_{k,l} = x_{k,l})$ , $k, l \in \eta_{i,j}$

The joint probability structure $P(X = x)$ of any process that satisfies the positivity condition is uniquely determined by the conditional probability distribution. With the neighborhood system $\eta$, the joint distribution of $X$, which is an MRF, is a Gibbs distribution with the following representation:

$$P(X = x) = \frac{1}{Z} exp^{-U(x)/T} \tag{2.1}$$

where $Z$ is the normalizing constant, also known as the *partition function* in the physics literature, $T$ is the temperature of the field. The exponent $U(x)$ in 2.1 is called the energy function. Energy function is of the form

$$U(x) = \sum_{c \in \mathcal{C}} V_c(x) \tag{2.2}$$

where $\mathcal{C}$ denotes the set of cliques and $V_c(x)$ is the clique potential.

A clique is a subset $\mathcal{C} \in S$ such that every pair of distinct sites in $\mathcal{C}$ are neighbors. This definition is best illustrated by a few examples. For a first order neighborhood ( Fig. 2.2) corresponding to the site $(i,j)$ the clique set is

$$\mathcal{C} = \{[(i,j)], [(i,j),(i,j+1)], [(i,j),(i+1,j)][(i,j),(i,j-1)], [(i,j),(i,j-1)]\}$$

Similarly, one can define second order neighborhood cliques. This will have 25 terms (See Fig. 2.3). It can be noticed that the number of cliques quickly blows up by increasing the neighborhood order. Thus, in most applications, one considers a first or a second order neighborhood system, because computational complexity is proportional to the number of cliques.

MRF theory tells us how to model the *a priori* probability of contextually dependent patterns, such as a class of textures and an arrangement of object features [78]. For the purpose of illustration, we cite the popular Ising model [53]. Here $F$ is a binary image with $f_{i,j} = 1(\mathcal{L} = \{+\infty, -\infty\})$; $+1$ representing an up-spin and $-1$ representing a down spin. Ising model assumes a first order neighborhood, and the function is given by

$$U(f) = -\sum[\alpha f_{i,j} + \beta\{f_{i,j}f_{i+1,j} + f_{i,j}f_{i,j+1} + f_{i,j}f_{i-1,j} + f_{i,j}f_{i,j-1}\}]$$

Such models were studied in the context of ferro magnetism. $\beta$ depends on the property of the material; $\beta > 0$ represents the attractive case, while $\beta < 0$ represents the repulsive case. The parameter $\alpha$ represents the effect of the external force field. A typical problem would be to find that configuration of up and down spins $f$ which maximizes $P[X = x]$. We see that for $\beta > 0$ this will mount to aligning all the spins in the same direction as the external force field.

A general form of the energy function is given by

$$U(f) = \sum_{i,j} f_{ij}G_{ij}f_{ij} + \sum_{ij,lm} f_{ij}f_{lm}G_{ij,lm}(f_{ij}f_{lm}) + \sum_{ijlmnk} f_{ij}f_{lm}f_{nk}G_{ij,lm,kn}(f_{ij}f_{lm}f_{kn}) + \ldots$$

The above expansion can be viewed as a Reed-Müller expansion of $U(f)$, where, $f$ is a vector, each element of which can take $p$ values (i.e., a $p$-ary variable, as opposed to a binary variable).

### 2.1.3 Incorporating Discontinuities

Within the MRF-Gauss Distribution framework, Geman and Geman [36] introduced an important concept of line fields located on the dual lattice to detect discontinuities in the random field. This dual lattice has two sites corresponding to the vertical and horizontal line fields, $v$ and $h$ respectively. Here, $h, v \in \{0, 1\}$, and the corresponding line field vectors are $H$ and $V$. One can then define *a priori* Gibbs distribution that combines $X, H$ and $V$ as

$$P[X = x, H = h, V = v] = \frac{1}{Z}\ exp[-U(x, h, v)].$$

where $Z = \sum_{x,h,v} \exp(-U(x,h,v))$. For a typical first-order weak membrane model [14]

$$\begin{aligned} U(x) &= \sum_{c \in \mathcal{C}} [\mu(x_{i,j} - x_{i,j-1})^2(1 - v_{i,j}) + (x_{i,j} - x_{i,j+1})^2(1 - v_{i,j+1}) + \\ &\quad (x_{i,j} - x_{i-1,j})^2(1 - h_{i,j}) + (x_{i,j} - x_{i+1,j})^2(1 - h_{i+1,j})] + \\ &\quad \gamma[v_{i,j} + h_{i,j} + v_{i,j+1} + h_{i+1,j}] \end{aligned} \tag{2.3}$$

The terms in the first bracket signify the interaction between the neighboring pixels; if the gradient is high, then the corresponding line field gets activated to indicate a discontinuity. For example, $v_{i,j} = 1$ if $|x_{i,j} - x_{i,j-1}| > \theta$ ( $\theta$ is threshold, and is constant), else $v_{i,j} = 0$. The term in the second bracket (weighted by a factor $\gamma$) imposes a penalty on every discontinuity created so as to prevent spurious discontinuities.

Energy functions of the form of Eqn. (2.3) are used for a variety of applications. Bhat [11, 12] used it for robust image restoration. Rajagopalan [106, 107] uses a similar model for depth from defocus. A super-resolution image model based on the above equation is used by Deepu [110, 111]. However, the parameters for Equation (2.3) are usually selected by trial and error. Use of iterative algorithms will be a bottleneck for such a trial and error method. Methods to estimate the parameters from image features are reported by Sridhar [76], and Nanda [92, 93]

### 2.1.4 Simulated Annealing

When the energy function to be minimized is non-convex, normal gradient descent type of algorithms fail. They may end up settling at a *local minima* . It can be observed that a solution procedure that aims for the global minima should have three capabilities: (i) it should be able to span whole of the solution space, (ii) it should be able to converge to the global solution when it is in the neighborhood of the solution in the solution space, and, (iii) some amount of randomness should be incorporated to come out of local minima. Simulated annealing (SA) and genetic algorithms are two well known algorithms which fulfill these criteria. We use simulated annealing algorithm, which guarantees convergence to global minima, in probability. A heuristic description of SA is given below:

SA algorithm allows us to go *uphill* (in the energy function) occasionally. This allows the energy function to climb out of a local minima. The algorithm also allows random jumps from one position to another, with probability. These two movements are controlled by *temperature T,* which decreases with time. This is called *simulated annealing*, because it simulates the annealing (gradual cooling) process used in metals. During the annealing process, free energy of a solid is minimized. This is possible if the metal is heated to a high temperature and then cooled very slowly.

A pseudo code for the SA algorithm is given below. For more details, one can refer Aarts [3]

```
BEGIN
    Initialize I-start,Tk and L
    k=0;i=I-start;
    REPEAT
        FOR  l = 1 to L DO
        BEGIN
            generate j ;
            IF f(j) <= f(i) THEN i = j
            else
                IF exp((f(i)-f(j))/Tk) > random [0,1) THEN i = j ;
        ENDFOR
    k = k +1 ;
    Tk = Tk/(log l) ;
    UNTIL stop criterion
    ENDREPEAT
END.
```

## 2.2 Hotelling (K-L) Transform

Let $\{x(n), 1 \leq n \leq M\}$ be a complex random sequence whose auto-correlation matrix is $\mathbf{R}$. Let $\Phi$ be an $M \times M$ unitary matrix, which reduces $\mathbf{R}$ to its diagonal form $\mathbf{\Lambda}$. Then transformed vector

$$\mathbf{y} = \mathbf{\Phi}^{*\mathbf{T}}\mathbf{x} \tag{2.4}$$

is called Karhunen-Loéve (KL) transform (or Hotelling Transform) of $\mathbf{x}$. It satisfies the property

$$E[yy^{*T}] = \mathbf{\Phi}^{*T}\{E[xx^{*T}]\}\mathbf{\Phi} = \mathbf{\Phi}^{*\mathbf{T}}\mathbf{R}\mathbf{\Phi} = \mathbf{\Lambda}E[y(k)y^{*}(l)] = \lambda_k\delta(k-l) \tag{2.5}$$

That is, elements of the transformed sequence $y(k)$ are orthogonal. If $\mathbf{R}$ represents the covariance matrix rather than the autocorrelation matrix of $x$, then the sequence $y(k)$ is uncorrelated. The unitary matrix $\mathbf{\Phi}^{*\mathbf{T}}$ is called the KL transform matrix. Its rows are the conjugate eigenvectors of $\mathbf{R}$, that is, it is the conjugate transpose of the eigenmatrix of $\mathbf{R}$.

## 2.3 Wavelets

Wavelet representation provides a multiresolution/multifrequency expression of a signal with localization in both time and frequency. Moreover, it decomposes a given signal into a set of coefficients associated with multiscaled wavelets. Discrete wavelet theory is found to be closely related to the framework of multiresolution analysis (MRA) and sub-band decomposition [85, 149], which have successfully been used in image processing applications [144]. In MRA, an image is represented as a limit of successive approximations, each of which is a smoothed and sub-sampled version of the image at the given resolution. In sub-band coding, entire frequency band of an image is decomposed into a number of sub-bands by a bank of band-pass filters. Each sub-band is then encoded separately. For reconstruction, the sub-band signals are decoded and then summed up to give a close replica or an estimate of the original image. Sub-band coding approach yields a signal-to-noise ratio comparable to the transform coding approach and yields a superior subjective perception due to the lack of the *blocking effect* [149]. Because of multiresolution analysis and sub-band decomposition, coding schemes and parameters can be adapted to the statistical properties of the wavelet coefficients at each scale of multiresolution decomposition. This property is very useful for image compression and the proposed zooming applications.

Multiresolution representation and sub-band coding approach have recently been integrated into the framework of the wavelet theory. Wavelet theory provides a systematic way to construct a set of filter banks with regularity constraint and compact support. In a wavelet representation, overall number of image samples is conserved after the decomposition, due to the orthogonality of wavelet basis at different scales. Wavelet representation is more flexible and can easily be adapted to the nature of the human visual system (HVS). It is also free from blocking artifacts, commonly encountered in block-transform based coding approaches for very low bit rate.

### 2.3.1 Continuous Wavelet Transform

Standard Fourier analysis decomposes a signal into frequency components and determines the relative strengths of each component. It does not give any information as to when the signal has exhibited a particular frequency component. That is time information is lost in the frequency domain and vice-versa. In signal analysis, it may be desirable to have a transformation that represents the signal in time and frequency domains simultaneously. For such applications, Gabor suggested a "windowed" Fourier transform - the short time

Fourier transform (STFT). Here, a fixed duration window is moved over time and extracts the frequency content within that interval of time. This fixed duration window is accompanied by a fixed frequency resolution, thereby, allowing only fixed time-frequency resolutions. This is a limitation of the STFT. The need is to have a variable window length, which is satisfied by the wavelet transform.

Wavelet transforms have basis functions formed by dilation and translation of a prototype function $\psi(t)$. These functions are of short duration - high frequency and long duration - low frequency. They are better suited for representing short bursts of high frequency signals and slowly varying signals of long duration.

The continuous wavelet transform (CWT) is defined in terms of dilation and translations of a prototype *mother wavelet* function $\psi(t)$, defined as:

$$\psi_{a,b}(t) = a^{-\frac{1}{2}}\psi\left(\frac{t-b}{a}\right). \tag{2.6}$$

Then, the CWT maps a square integrable function ($f(t) \in L^2(I\!R)$) in time scale to

$$W_f(a,b) = \int_{-\infty}^{\infty} f(t)\psi^*_{ab}(t)\ \ dt. \tag{2.7}$$

For a large $a$, the basis function becomes a stretched version of the prototype wavelet, that is, a low-frequency function; while for a, small $a$ this basis function is a contracted version of the mother wavelet function, which is short-time duration high-frequency function. Depending on the scaling parameter $a$, the wavelet function $\psi(t)$ dilates or contracts in time, causing corresponding contraction or dilation in the frequency domain. Thus the wavelet transform provides the needed time-frequency localization and resolution.

For a function to be a (mother) wavelet, it must satisfy admissiblilty criterion. A square integrable function $\psi \in L^2(I\!R)$ is admissible if

$$C_\psi = \int_{-\infty}^{\infty} \frac{|\Psi(\omega)^2|}{|\omega|}\, d\omega \ \ < \ \infty \tag{2.8}$$

where, $\Psi(\omega)$ is the Fourier transform of $\psi(t)$. The constant $C_\psi$ is the admissibility constant of the function $\psi(t)$, and the requirement is that it is finite, allows inversion of the wavelet transform. For a given function $\psi(t)$, the condition $C_\psi < \infty$ holds only if $\Psi(0) = 0$, or equivalently if

$$\int_{-\infty}^{\infty} \psi(t)\ dt = 0.$$

If the admissibility condition is satisfied, then inversion of $f(t)$ from $W_f(a,b)$ is given by

$$f(t) = \frac{1}{C_\psi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_f(a,b)\psi_{ab}(t)\frac{da\ db}{a^2}. \tag{2.9}$$

### 2.3.2 Discrete Wavelet Transform

The CWT is highly redundant for $a$ and $b$ varying continuously over the entire range. Therefore, the transform is usually evaluated on a discrete grid on the time-scale plane, corresponding to a discrete set of continuous basis function as:

$$\psi_{mn}(t) = a_0^{-m/2}\psi(a_0^{-m}t - nb_0) \tag{2.10}$$

where, $m, n \in \mathbf{Z}$ ($\mathbf{Z}$ is a set of integers), $a_0 > 1$ and $b_0 \neq 0$. This corresponds to $a = a_0^m$ and $b = na_0^m b_0$.

We observe that the translation steps depend on the dilation, since, long wavelets have to be advanced by large steps and short ones by small steps. On such a grid, the discrete wavelet transform is given by:

$$W_f(m,n) = a_0^{\frac{-m}{2}} \int_{-\infty}^{\infty} f(t)\psi^*(a_0^{-m}t - nb_0)\, dt \tag{2.11}$$

A special case, which corresponds to $a_0 = 2$ and $b_0 = 1$, called dyadic sampling, has been extensively studied in literature. Discussion here is limited to dyadic wavelets. A special case of M-adic wavelets are reported by Subrahmanya and Ramakrishna [134]. For dyadic case, it is possible to construct function $\psi(t)$ such that the set

$$\psi_{mn}(t) = 2^{-m/2}\psi(2^{-m}t - n) \tag{2.12}$$

is orthonormal for both $m, n \in \mathbf{Z}$. Such wavelets corresponds to mathematical tool proposed by Mallat[83], called Multiresolution analysis (MRA). MRA enables the characterization of a class of functions $\psi(t) \in L^2(I\!R)$ that generate an orthogonal bases in terms of discrete-time filter coefficients. The idea of MRA is to write an $L^2(I\!R)$ function $f$ as a limit of successive approximations, each of which is a smoothed version of $f$. These successive approximations correspond to different resolutions, much like a pyramid. This smoothing is accomplished by convolution with a low pass kernel called the *scaling function*, $\phi(t)$.

An MRA consists of sequence of closed subspaces $V_m | m \in \mathbf{Z}$, of $L^2(I\!R)$, that have the following properties:

- These spaces describe successive approximation spaces - (Containment)

$$\ldots V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \ldots$$

- Completeness:

$$\bigcap_{m \in \mathbf{Z}} V_m = \{0\}$$

and

$$\bigcup_{m \in \mathbf{Z}} V_m = L^2(I\!R)$$

- Scaling:

$$f(x) \in V_m \Leftrightarrow f(2x) \in V_{m-1}$$

for any function $f \in L^2(I\!R)$.

- There exists a scaling function $\phi(t) \in V_0$ such that

$$\phi_{mn}(t) = 2^{-\frac{m}{2}} \phi(2^{-m}t - n)$$

is orthonormal for a fixed $m$. The space spanned by $\phi_{mn}$ for a fixed $m$ is denoted by $V_m$. For each $m$, let $\psi_{mn}$ span a space $W_m$, which is exactly the orthogonal complement of $V_m$ in $V_{m-1}$. That is,

$$V_{m-1} = V_m \oplus W_m$$

where, $\oplus$ represents the direct sum. Furthermore, the direct sum of all the $W_m$ spans $L^2(I\!R)$:

$$L^2(I\!R) = \ldots W_{-2} \oplus W_{-1} \oplus W_0 \oplus W_1 \oplus W_2 \oplus \ldots$$

Suppose we start with a scaling function $\phi(t)$, such that its translation $\{\phi(t-n)\}$ spans $V_0$. Then $V_{-1}$ is spanned by $\{\phi(2t-n)\}$, dilates of the function $\phi(t)$ in $V_0$. Thus, $V_{-1}$ is spanned by the integer translation of two functions, and $\phi(t)$ can be expressed as a linear combination of translates of $\phi(2t)$ as:

$$\phi(t) = 2 \sum_n h_n \phi(2t - n) \tag{2.13}$$

This equation is a consequence of the containment property. The coefficient set $\{h_n\}$ are the *inter-scale basis coefficients*. Similarly, the wavelet function $\psi(t)$ can be expressed as a linear combination of translates of $\phi(2t)$ as:

$$\psi(t) = 2 \sum_n g_n \phi(2t - n). \tag{2.14}$$

These two-scale difference equations are fundamental in the generation of orthonormal, compact support, discrete wavelets. They show how the digital filters $\{h_n\}$ and $\{g_n\}$ wholly determine the scaling function $\phi(t)$ and $\psi(t)$.

Taking discrete time Fourier transform of the two equations (2.13) and (2.14) yields

$$\Phi(\omega) = H\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right)$$

and

$$\Psi(\omega) = G\left(\frac{\omega}{2}\right)\Phi\left(\frac{\omega}{2}\right)$$

where,

$$\begin{array}{l} H(\omega) = \sum_n h_n e^{-j\omega n} \\ G(\omega) = \sum_n g_n e^{-j\omega n} \\ \Psi(\omega) = \int_{-\infty}^{\infty} \phi(t) e^{-j\omega n} \end{array} \tag{2.15}$$

Integrating results in

$$\Phi(\omega) = \Psi(0) H\left(\frac{\omega}{2^k}\right)$$

and

$$\Psi(\omega) = G(0)\Psi(0) H\left(\frac{\omega}{2^k}\right)$$

Normalizing the scaling function by $\int \phi(t)dt = 1$ leads to $\Phi(0) = 1$ and hence $H(0) = 1$. The wavelet must be a band-pass function, satisfying the admissibility condition

$$\Psi(0) = \int_{-\infty}^{\infty} \psi(t)dt = 0$$

and hence $G(0) = 0$, or equivalently $\sum_n g_n = 0$.

The above analysis shows that the Fourier transform of the continuous time scaling function is obtained by the infinite product of the discrete time Fourier transform of the inter-scale coefficients $\{h_n\}$.

If the function $f$ is given in a sampled form, then one can take these samples for the highest order resolution approximation coefficients $a_{0n}(f)$ and describe a sub-band coding algorithm on these sampled values, with low-pass filter $g$ and high-pass filter $h$ as

$$\begin{array}{l} a_{m,n}(f) = \sum_k h_{2n-k} a_{m-1,k}(f) \\ c_{m,n}(f) = \sum_k g_{2n-k} a_{m-1,k}(f) \end{array} \tag{2.16}$$

This filter bank approach is pictorially depicted in Fig.2.4

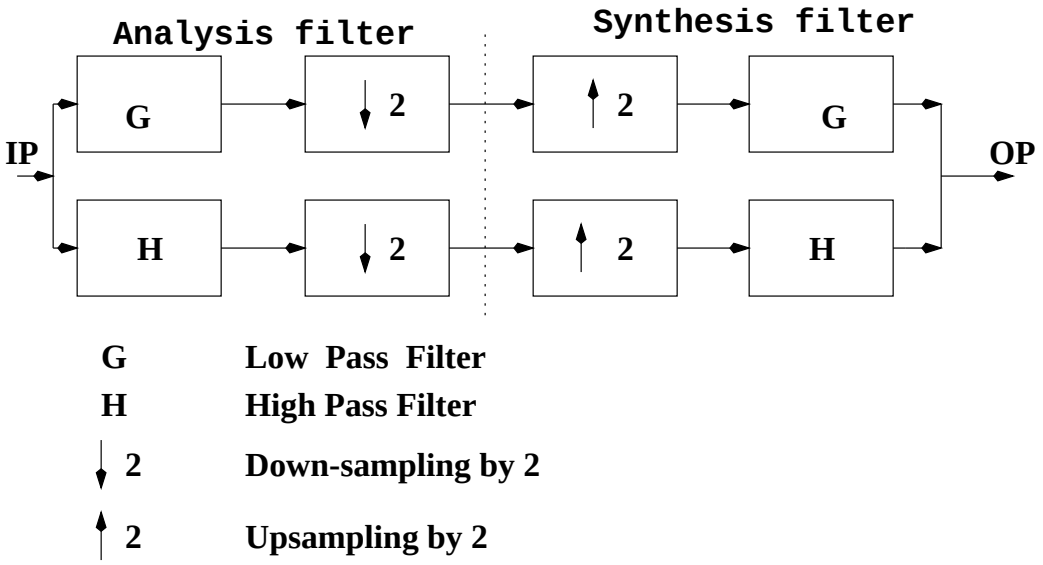


Figure 2.4: Wavelet Filter bank

Because of their association with orthonormal wavelet bases, these filters give a prefect reconstruction. That is:

$$a_{m-1,l}(f) = \sum_n [h_{2n-l} a_{m,n}(f) + g_{2n-l} c_{m,n}(f)]$$

Moreover, these orthogonal wavelet bases are closely linked with the unitary 2-band perfect reconstruction quadrature mirror filter (PR-QMF)

Most of the orthonormal wavelet bases have infinitely supported $\phi$, corresponding to filters $h$ and $g$ with infinitely many taps. The construction in [21] gives $\psi$ with finite support, and therefore corresponds to Finite impulse Response (FIR) filters. It follows that the orthonormal bases of Daubechies [21] correspond to a sub-band coding scheme with perfect reconstruction property, using the same FIR filter for reconstruction, as for the decomposition.

One can use biorthogonal wavelets [1], instead of orthogonal ones. Here, the wavelet used for the analysis is different from the one used at the synthesis stage. The added flexibility in the biorthogonal case permits use of linear phase and unequal filters. Perfect reconstruction conditions are satisfied by imposing orthogonality across the analysis and synthesis stages. Biorthogonality provides additional degrees of freedom so that both perfect reconstruction and linear phase filters can be realized simultaneously.

Consider $\{h_n, g_n\}$ constituting the analysis and synthesis filter banks respectively. Then, for orthogonal wavelets, the analysis and synthesis filters are time reversal of each

other. Equivalently, $h(n) = h(-n)$, $g(n) = g(-n)$ and

$$\begin{array}{l} g(n) = (-1)^n h(1-n) \\ \sum_n h(n)h(n+2k) = \delta(k) \end{array} \tag{2.17}$$

That is, $h_n$ is orthogonal to translations of itself. For biorthogonal wavelets, these restrictions are relaxed. However, in order to perfectly reconstruct the input, these four filters have to be related as [1]:

$$\begin{array}{l} g(n) = (-1)^n h(1-n) \\ g(n) = (-1)^n h(1-n) \end{array} \tag{2.18}$$

and

$$\sum_n h(n)h(n+2k) = \delta(k)$$

Thus, they are cross related by time reversal and flipping signs of every other element. Here, $h(n)$ is orthogonal to $h_n$, and thus, the name *biorthogonal*.

### 2.3.3 2-Dimensional DWT

One dimensional wavelet case of Mallat [83] can be extended to two dimensional case. For two-dimensional wavelet analysis, a scaling function $\phi(x, y)$, similar to the one-dimensional case, is defined as,

$$\phi(x, y) = \phi(x)\phi(y) \tag{2.19}$$

where, $\phi(.)$ is a one-dimensional scaling function. The separability of $\phi(x, y)$ is normally assumed, but not essential. Let $\psi(x)$ be the one-dimensional wavelet associated with the scaling function $\phi(x)$. Then, the three two-dimensional wavelets are defined as:

$$\psi^H(x, y) = \phi(x)\psi(y)\psi^V(x, y) = \psi(x)\phi(y)\psi^D(x, y) = \psi(x)\psi(y) \tag{2.20}$$

where, the superscripts $H, V$ and $D$ stand for *horizontal*, *vertical* and *diagonal* (parts) respectively. With the compact support assumption, filtering can be done on "rows" followed by "columns" in the two dimensional array, as shown in Fig.2.5

Wavelet coefficients of the image can be computed using the sub-band coding algorithm, similar to the one-dimensional case. Filters $H$ and $G$ are one-dimensional filters. This decomposition provides sub-images corresponding to different resolution levels and orientations. For example, $LH$ are the horizontal edges, $HL$ the vertical and $HH$ the diagonal edge. Hence, the superscripts $H, V$ and $D$. Note that if the original image consists of an $N \times N$ array, output image consists of four arrays, each of $N/2 \times N/2$. Output image

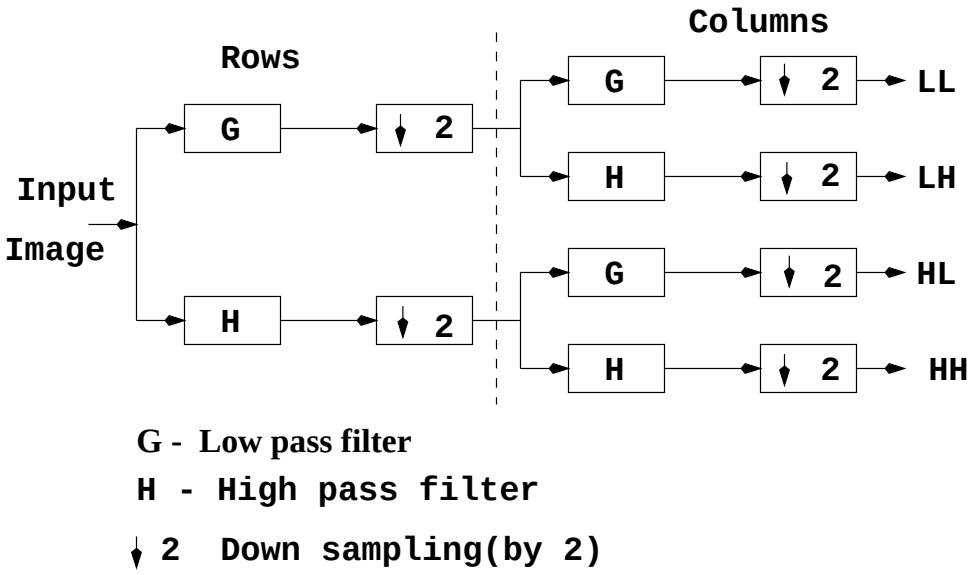


Figure 2.5: One stage of wavelet decomposition with 2-D separable filter

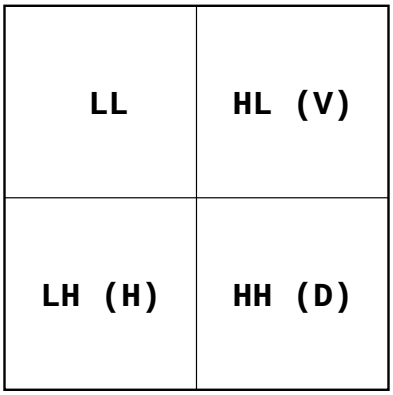


Figure 2.6: Schematic representation of a 2-D wavelet transform

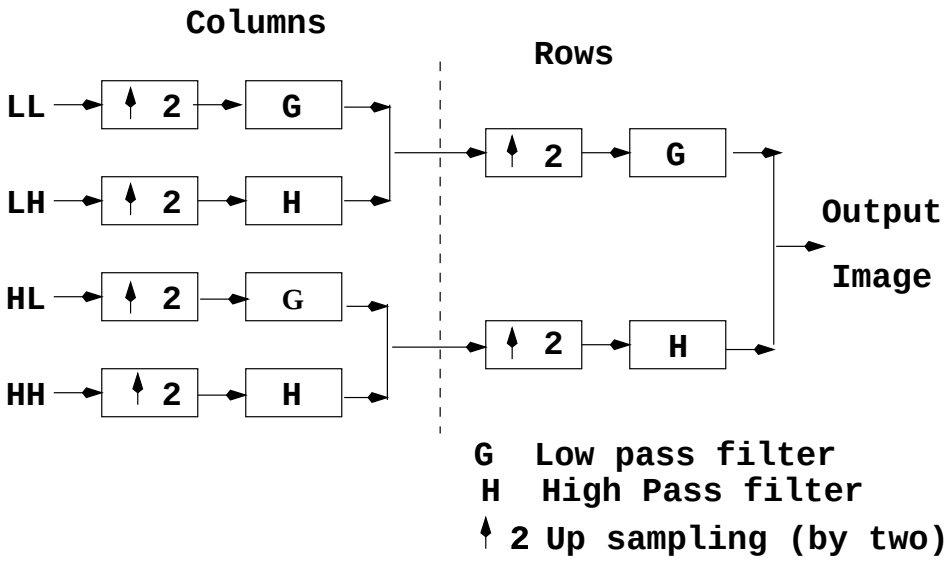


Figure 2.7: 2-D Wavelet reconstruction (synthesis) filter

can be represented by an image of the same size as the original image. These relations are shown in Fig.2.6. One can decompose $LL$ further if more multiresolution layers are required. Reconstruction (inverse) for 2-D is the exact reverse of the forward operation, as shown in Fig.2.7

If *all four* sub-bands are subjected to another set of filtering and down-sampling, it leads to *uniform decomposition*, or a balanced tree. If *only* the low band $LL$ is further decomposed, we get an *octave band* or a *logarithmic tree*. Decomposition between these two extremities that is, decomposition of a few selected bands leads to *wave packets*.

## 2.4 Summary

This chapter provided necessary mathematical background for the subsequent chapters. Following things were discussed. A *neighborhood system* which relates nearby elements in a lattice and *clique* which implies how neighboring elements interact were defined. Normal image restoration imposes smoothness criterion, which smooths sharp edges in an image. To retain sharp edges, or to retain discontinuities, *line fields* are introduced. An image can be modeled as MRF using neighborhood system, cliques and line field. Such a model gives rise to an energy function, which is to be minimized. *Simulated annealing*, which guarantees convergence to global minima is discussed. We introduced *wavelets* and showed how wavelets (filter banks) can generate a *multi-resolution* representation of a signal. One dimensional wavelet was extended to two dimensional wavelets.