CHAPTER 2 - LITERATURE SURVEY

2.1 Introduction

As the theme of this thesis is vision based image processing, basics of Human Visual System (HVS), Boundary Contour System (BCS) and functions of retina are reviewed in this chapter. Further, motivation and methods behind the hexagonal image processing are discussed. The importance of the vision based image processing operations such as image enhancement and edge detection has been brought out from the study of biological visual system. Features and acquisition of Hexagonal Sampled Images and Hexagonal Structure Addressing are also reviewed in this chapter. Implementation of Boundary Contour System is included. Limitations of the existing work and need for further work is discussed.

2.2 Review of Biological Visual System

Vision is to be understood as the process of acquiring knowledge about environmental objects and events by extracting information from the light they emit or reflect [7]. Visual perception is the ability to interpret information and surroundings from visible light reaching the eye. The resulting perception is also known as eyesight, sight or vision. The various physiological components involved in vision are collectively referred as the visual system and are the focus of much research in psychology, cognitive science, neuroscience and molecular biology. The field of biological vision studies and models the physiological processes behind visual perception in humans and other animals. Computer vision on the other hand studies and describes the processes implemented in software and hardware behind artificial vision systems. Interdisciplinary exchange between biological and computer vision has proven fruitful for both fields.

Any understanding of the function of the human eye serves as an insight into how machine vision might be solved. Some of the early work by Hubel and Wiesel [6] on the receptive fields in the retina has led to the fundamental operation of spatial filtering that dominates initial stages of image processing. Ennio Mingolla, William Ross and Stephen Grossberg [4] introduced recent development of the Boundary Contour System (BCS) model which suggests
how the laminar, columnar, and map organization of the visual cortex accomplishes boundary segmentation. The BCS is a dynamic neural network model composed of shunting neurons designed to replicate the properties of illusory contour formation as observed in psychophysical studies. From an implementation point of view, the BCS is a serial interconnection of number of separate computational processes each with a large degree of parallelizable computation [4, 5].

2.2.1 Hierarchy of Visual Processes

The visual system is the most complex of all the various sensory systems in the human brain [10]. The visual system contains over 30 times the number of neurons associated with auditory processing. The brain contains hundreds of aggregations of cells which may take the form of balls or of stacks of layered plates [8, 60]. The balls or plates are often connected in serial order as pathways. A good example of such a serially connected system is the visual pathway. The retina of each eye consists of a plate having three layers of cells, one of which contains the light-sensitive receptor cells or rods and cones. Each eye contains over 125 million receptors. The two retinas send their output to the lateral geniculate bodies. These structures in turn connect their fibers to the visual part of the cerebral cortex either to the striate cortex or primary visual cortex. From there, after being passed from layer to layer through several sets of synaptically connected cells, the information is sent to several neighboring higher visual areas.

The lateral geniculate body receives its main input from the retina. Each cell in the geniculate body will receive connections from rods and cones by way of intermediate retinal cells. The population of rods and cones that feed into a given cell in the visual path-way are not scattered about all over the retina but are clustered into a small area. This area of the retina is called the receptive field of the cell. The inputs received by the two eyes from the adjacent regions of the field are mapped to LGN to preserve the spatial relationships. This nature of mapping continues to the first stage of the visual cortex (V1) which is about 2 mm thick and also consists of a six-layer structure. The M (Magno) and P (Parvo) pathways (Figure.2.1) project to distinct sub-layers within this layer.
The Magno cells have large receptive fields and they respond to large objects and follow rapid changes in stimulus. Parvo ganglion cells have smaller receptive fields and selectively respond to specific wavelengths. The main function of V1 is to decompose the results from the LGN into distinct features which can be used by other parts of the visual system. The LGN cells respond to spots (circles) of light whereas the simple cells in V1 respond to bars of light at specific orientations.

Figure 2.1 Information flow in the human visual system [10]

2.2.2 Organization of the Retina

Figure 2.2 describes a drawing of a section through the human eye with a schematic enlargement of the retina [9]. The retina is a multi-layered structure, containing 3 nuclear layers (of neurons) and two plexiform layers (for interconnections amongst the neurons). The photoreceptors are at the back, so light must first travel through all layers of the retina before being absorbed by the pigments in the rods and cones. There are basically two directions of signal flows in the retina: one longitudinal (photoreceptors → bipolar cells → ganglion cells) and the other lateral (via horizontal cells in the outer plexiform layer, and amacrine cells in the inner plexiform layer). The retina can be described as an “image capture” device like a camera, having analogue input photo transducers that convert photons into voltage changes, and discrete output devices that send pulses down the optic nerve. There are 120 million “input channels” (the photoreceptors, similar in a sense to pixels), but only 1 million...
“output channels” (the axons of the ganglion cells which constitute the optic nerve). Clearly the retina does a lot of processing on the image and sends its coded results to the brain.

2.2.3 Receptive Field Structure in the Retina

The centre-surround receptive field structure of retina which is essential for modeling of BCS is explained in Section 2.2.5. In both space and time, retinal neurons can be described as filters and they act as linear devices (having the properties of proportionality and superposition of responses to components of stimuli). An important aspect of retinal receptive fields (as distinct from those found in most neurons of the visual cortex) is that their spatial structure is isotropic or circularly symmetric, rather than oriented. Functions of each layer in the retina (Figure 2.2) can be summarized as follows [9, 60]:

Figure 2.2 The human visual pathway - cellular representation of the retina (top) and various parts of the brain associated with visual processing (bottom) [8]
An interesting aspect of the retina is its structure. The superficial layers, which are transparent, consist of neurons while the photoreceptors are found at the deepest layer. In the thinnest part of the retina, called the fovea, the neurons are moved aside to let light pass through directly to the photoreceptors. There is also a region of the retina known as the optic disc where there is an absence of photoreceptors to permit neural wiring to carry information out to the brain. This gives rise to a blind spot.

Horizontal cells pool together the responses from large numbers of photoreceptors within a local area. With these “surround” signals, they inhibit bipolar cells.

Bipolar cells are the first to have a “centre-surround” receptive field structure (Figure 2.4). Their response to light in a central disk is opposite from their response to light in the local surrounding area. Field boundaries are circular and roughly concentric (i.e. annular). Amacrine cells are “on - off” in temporal, as opposed to spatial terms.

Ganglion cells combine these spatial and temporal response properties and thus serve as integro - differential image operators with specific scales and time constants. Moreover they convert their responses to impulses in a spike frequency code, traveling down their axons which are the fibers of the optic nerve to the thalamus and hence on to the primary visual cortex in the brain.

Figure 2.3 The retinal signal processing using center-surround receptive fields [60].
Figure 2.4 Typical classical receptive fields of neurons in the visual pathway. Plus signs denote regions of the visual field where light causes excitation, minus signs denote regions where light causes inhibition. (a) Retinal ganglion and LGN neurons typically exhibit center-surround receptive fields organization, in one of two arrangements (ON-centre, OFF surround or OFF centre, ON surround) (b) The majority of simple cells in V1, on the other hand, have oriented receptive fields [8].

2.2.4 Arrangement of Rods and Cones in the Eye

The arrangement of the photoreceptors along the spherical retinal surface is illustrated in Figure 2.5. Here, the larger circles correspond to the rods and the smaller circles to the cones. A significant fact to notice is that the general topology in this diagram is roughly hexagonal. This is because, all naturally deformable circular structures pack best in two dimensions within a hexagonal layout such as found in honeycombs which inspires one to work on the image processing operations on hexagonal lattice.
2.2.5 Boundary Contour System (BCS) Model

The BCS provides a powerful technique to recognize patterns and restore image quality under excessive fixed pattern noise as in Synthetic Aperture Radar (SAR) images [4, 5]. BCS is used to model extraction of image boundary features in visual cortex, and encompasses visual processing at different levels, including several layers of cells interacting through shunting inhibition, long-range cooperative excitation and renormalization.

The BCS consists of a series of boundary detection, competition, and co-operation stages as shown in Figure 2.6. Stage 1 models the contrast enhancement resulting from on-center, off-surround (ON channel) and off-center, on-surround (OFF channel) interactions at the retina as described in Figure 2.4. These ON and OFF cells compensate for variable illumination by computing locally normalized contrast ratios throughout the image.

At stage 2, these ON and OFF cells generate half-wave rectified outputs which together drive the activation of oriented simple cells. Simple cells compute a measure of the local image gradient magnitude and orientation. Complex cells are insensitive to directions-of-contrast (dark-to-light vs light-to-dark).

---

**Figure 2.6 Schematic representations of BCS [4]**

BCS model proposed by Ennio Mingolla et.al [4] consists of nine layers, in which after local illumination normalization and contrast enhancement of an input image, performs local edge extraction for different spatial orientations and
scales, and then is able to identify consistent long range contours of the shapes in the input image through processing layers with feed forward and feedback connections. From Layer 1 to Layer 3 there are only feed forward filtering operations, while Layers 4 to 8 are connected in a feedback loop configuration, which means the system will reach a steady state after a certain number of iterations. Complex cell activations compete at stage 3 processing. Competition occurs through on-center off-surface processing across both image space (spatial competition) and across orientation space (orientational competition). Functionally, competition sharpens boundary localization and orientation tuning of individual complex cell filters. Since the modeling of these cells is quite complex, the real process is not understood.

A consequence of the finding about the receptive fields shown in Figure 2.4 is that the early visual processing can be modeled as a sequence of filter convolutions. The center-surround receptive fields can be modeled as the difference of two Gaussian kernels, a classical model of which is given by the Difference-of-Gaussian (DoG) filter. The orientation selective receptive fields can be modeled by Gabor filters which are products of sinusoidal gratings and Gaussian envelopes [61]. Gabor filters are closely related to the function of simple cells in the primary visual cortex of primates as explained in the Section 2.2.5. Since simple cells play a critical role in texture boundary detection, a computational model of this type of cell is essential. Previous studies have shown that Gabor filters closely resemble experimentally measured receptive fields in the visual cortex [62] and have been widely used to model the response of visual cortical neurons. Bank of oriented Gabor filters can be used to approximate the responses of simple cells in the primary visual cortex.

Therefore, the initial effort of this work had been to study and implement a model for Boundary Contour System. The preprocessing stages of the BCS has the properties of Dynamic Compression and Image Edge Enhancement [46] as shown in Figure.2.7.
2.2.6 Implementation Details of BCS

2.2.6.1 Stage 1: Contrast Enhancement (Center-ON OFF-Surround)

Let us assume

\[ I(p, q) = \begin{cases} \{ & p = 1, 2, \ldots, M \\ & q = 1, 2, \ldots, N \} \end{cases} \]

is an \( M \times N \) input image provided by a vision sensing front end. This input image is applied to a 2D filter where the kernel has radial symmetry. The pixels close to the center region of the kernel are going to contribute with positive weights to the convolution, while pixels further away will contribute negatively.

The mathematical expression for this kernel is

\[ S1(p, q) = A_1 e^{-p^2 + q^2 / \sigma g} - A_2 e^{-p^2 + q^2 / \alpha \sigma g} \quad (2.1) \]

where \( A_1 > A_2 \) are positive parameters (\( A_1 = 12 \) and \( A_2 = 0.5 \)), \( \alpha > 1 \) and \( \sigma g \) controls the spatial scale of the filtering. The result of convolution with the kernel of Equation (2.1) is performing contrast enhancement. ON channel shows enhanced response to image locations of high intensity relative to their surrounding image locations. OFF channel shows enhanced response to image locations of low intensity relative to their surrounding image locations. The
response shown was obtained using $\sigma = 3.7$ and $\alpha = 1.3$. Convolution with the generated kernel (Figure 2.8) shows image enhancement which can be observed from the results as shown in Figure 2.9.

![Figure 2.8 Centre on-off surround kernel of radial symmetry with $\sigma = 3.7$ and $\alpha = 1.3$](image)

Figure 2.8 Centre on-off surround kernel of radial symmetry with $\sigma = 3.7$ and $\alpha = 1.3$

(a) (b)

(c) (d)

Figure 2.9 (a), (c) Input Images (b), (d) Images obtained by convolution with center surround kernel which shows image enhancement
2.2.6.2 Stage 2: Boundary Detection

At the second stage, oriented contrast is detected by nodes with connectivity analogous to that of cortical simple cells. Both ON and OFF network activity is used for oriented contrast at each image location. An edge elicits a strong ON response next to a strong OFF response, an optimal coincidence for these boundary detectors. This stage of the BCS system applies an orientation specific convolution for detecting edges oriented within a narrow angle range. This is performed by convoluting the output of stage 1 with the kernel of this stage for different orientations. The kernel is described by the difference between displaced Gaussians as described in the Equation (2.2) [63].

\[
F_g(p_k, q_k) = \frac{1}{2\pi \sigma_{gh} \sigma_{gv}} e^{-\frac{1}{2} \left( \frac{p_k}{\sigma_{gh}} \right)^2} \left[ e^{-\frac{1}{2} \left( \frac{q_k}{\sigma_{gv}} + \frac{1}{2} \right)^2} - e^{-\frac{1}{2} \left( \frac{q_k}{\sigma_{gv}} - \frac{1}{2} \right)^2} \right] (2.2)
\]

where the coordinate system \((p_k, q_k)\) is rotated at certain angle with respect to the coordinate system of the input image \((p, q)\) and \(\sigma_{gh}, \sigma_{gv}\) represent deviation or simple cell width for three different scale. Generally \(\sigma_{v0} = 0.75, \sigma_{v1} = 1.5, \sigma_{v2} = 3.0\) for simple, medium and large simple cell width respectively [5] and \(\sigma_{gh} = 3\sigma_{v2}\). The response of stage 2 model is shown in Figure 2.10 which highlights the edges depending on the orientation K.

\(p_k\) and \(q_k\) are described as

\[
p_k = p \cos \frac{\pi k}{nr} - q \sin \frac{\pi k}{nr}
\]

and

\[
q_k = p \cos \frac{\pi k}{nr} + q \sin \frac{\pi k}{nr}
\]

(2.3)

where \(nr\) being the total number of orientations to be considered.

Stage 1 and Stage 2 kernels of BCS cell were modeled and operated on some images as an initial study. Stage 1 of BCS cell is performing image enhancement operation of the image and Stage 2 performs the edge enhancement operation of the image. Hence, these two operations are considered for the analysis of Hex-Gabor proposed in this work which is discussed in the subsequent chapters.
As the aim of this research work is vision based hexagonal image processing operations, basics of Human Visual System (HVS), Boundary Contour System (BCS) and functions of retina were reviewed in this section. Motivation behind the hexagonal image processing has been brought out. The importance of the vision based image processing operations such as image enhancement and edge detection has been brought out from the study of biological visual system. Features and sampling schemes of hexagonal image processing are explored in section 2.3.

2.3 Hexagonal Sampling Scheme and Features

In order to store, process, display and transfer images by digital devices the image plane must be quantized into spatial elements of finite dimension, which is generally referred to as pixels. Digitization, i.e. to convert real images into discrete sets of points is one of the earliest subjects of study for computer scientists involved in vision and graphics research. Each point which forms an image on the screen must be properly addressed in order to be indexed. The disposition of the points on the plane called digitization scheme can take different choices. Considering technical implementation, these points must be placed as regularly as possible on the plane so that the coverage of the plane is as efficient as possible.
2.3.1 Three Possible Regular Tessellation Schemes

There are three possible regular tessellation schemes to tile a plane without overlapping among the samples and gaps between them. They are tessellation with squares, hexagons and regular triangles (Figure 2.11). Any other types of spatial tessellation will result in either unequal distance between neighboring pixels or introduce gaps or overlaps among samples. A simple explanation is given below.

We use the symbol \{p, q\} to denote the tessellation of regular polygon search of which has pixels surrounding the vertex. It is easy to see that the \{p, q\} pairs are \{4, 4\}, \{6, 3\} and \{3, 6\} for the three tessellation schemes as illustrated in Figure 2.11, and in each case the polygon drawn in bold lines is the vertex. A tessellation is said to be regular if it has regular faces and a regular vertex. On the left is the square case \{4, 4\}, which is familiar and usual because it is aligned with the standard Cartesian axes which helps to make operations simple and intuitive. The middle illustrates the triangular case \{6, 3\}, which yields a denser packing than the square case. This means that more information is contained in the same area of the image. The tessellation in the far right hexagonal case \{3, 6\}, is often used for tiled floors and it can also be seen in any beehive. It is believed to be the most efficient tessellation scheme amongst all of them and its features are briefly discussed in the following sections.

![Figure 2.11 Tessellation Schemes (a) Square (b) Triangle (c) Hexagon](image)

2.3.2 Hexagonal sampling

Band-limited two-dimensional signals can be sampled and processed as array of numbers, which is fundamental and well known. With rectangular sampling a band-limited function of two independent variables are sampled at
evenly spaced values of each of those variables. Petersen and Middleton [64] showed that rectangular sampling as a special case in general sampling strategy by which a band-limited waveform is sampled on a skewed (i.e., non orthogonal) sampling raster. Hexagonal sampling is another special case of this general strategy. It is the optimal sampling scheme for signals which are band limited over a circular region of the Fourier plane, in the sense that exact reconstruction of the waveform requires a lower sampling density compared with alternative schemes. For such signals hexagonal sampling requires 13.4 percent fewer samples than rectangular sampling [17].

2.3.3 Rectangular and Hexagonal Sampling

As per M. Mersereau [17], if \( x_a(t_1, t_2) \) denotes an analog waveform (continuous independent variables), the operation of rectangular sampling can be described by

\[
x(n_1, n_2) = x_a(n_1 T_1, n_2 T_2)
\]

(2.3)

where \( T_1 \) and \( T_2 \) are the horizontal and vertical sampling intervals. The doubly indexed array of sample values \( x(n_1, n_2) \) will be referred to as a rectangular sequence or a rectangular array. Rectangular sampling thus corresponds to sampling a waveform at the sampling locations.

Let \( X_a(\Omega_1, \Omega_2) \), the Fourier transform of a waveform \( x_a(t_1, t_2) \). Then, \( x_a(t_1, t_2) \) will be said to be band limited with a band region \( R \) if \( X_a(\Omega_1, \Omega_2) \) satisfies the condition

\[
X_a(\Omega_1, \Omega_2) = 0, \quad (\Omega_1, \Omega_2) \notin R
\]

(2.4)

For \( x_a(t_1, t_2) \) to be exactly recoverable from the rectangular sequence defined by Equation (2.3), it must be band limited with the rectangular band region as shown in Figure 2.13 with

\[
T_1 < \frac{\pi}{w_1} \quad \text{and} \quad T_2 < \frac{\pi}{w_2}
\]

(2.5)

\( w_1 \) and \( w_2 \), the horizontal and vertical bandwidths, are measured in radians under the assumption that the independent variables \( t_1 \) and \( t_2 \) are dimensionless.
Hexagonal sampling is similar to rectangular sampling in that the values of the data sequence correspond to sample values of the analog waveform. However, the sample locations differ. Hexagonal sampling can be described by

\[ x(n_1, n_2) = x_a \left( \frac{2n_1 - n_2}{2} T_1, n_2 T_2 \right) \]  

(2.6)

This corresponds to sampling \( x_a(t_1, t_2) \) at the locations indicated in Figure 2.12(b). Alternate rows of the hexagonal sampling raster are identically positioned and the odd-numbered rows are shifted horizontally one-half sample interval with respect to the even-numbered rows. For \( x_a(t_1, t_2) \) to be exactly recoverable from the hexagonal sequence defined by Equation (2.5), it must be band limited with a hexagonal band region as shown in Figure 2.13(b) with

\[ T_1 < \frac{4\pi}{2w_1 + w_2}, \quad T_2 < \frac{\pi}{w_2} \]  

(2.7)

It should be noted that even when the band shape is a regular hexagon \((W_1 = W_3 = 2 W_2 / \sqrt{3})\) the horizontal and vertical sampling intervals will not be the same. Figure 2.14 shows a circular band region imbedded in a rectangular region and hexagonal band region.

Figure 2.12 (a) A rectangular sampling raster (b) A hexagonal sampling raster

[17]
Sampling theory [64] can be used to determine necessary conditions for reliably sampling a continuous signal $f(x)$. The conditions are based on the relationship between the sampled spectra and the original signal’s spectra. Given a continuous signal $f(x)$, its Fourier transform $F(\Omega)$ is defined as follows [10]:

\[
F(\Omega) = \int_{-\infty}^{\infty} f(x) e^{-j\Omega \cdot x} dx
\]

\[
f(x) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} F(\Omega) e^{j\Omega \cdot x} d\Omega
\]

Here, $\Omega \in \mathbb{R}^2$ is a frequency domain vector and $x \in \mathbb{R}^2$ is a vector in the spatial domain. Using the above relationship, it is possible to determine the effect of sampling the signal on its spectrum as follows:
Using a substitution \( \omega = V^T \Omega \) yields:

\[
  f_s(n) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} F((V^{-1})^T \omega) e^{j\omega^T n} \frac{d\omega}{|\det V|}
\]

(2.10)

where \( V = \{v_1,v_2\} \) is one of the valid basis of the rectangular sampling as shown in Figure 2.15. Instead of integrating this over the entire plane it is possible to split it into a series of smaller sub integrals with area \( 4\pi^2 \). If \( V \) is spatial domain lattice, then frequency domain lattice or reciprocal lattice \( \tilde{V} \) is obtained as \( \tilde{V} = 2\pi(V^{-1})^T \). Then the basis vectors in the two domains are seen to be mutually orthogonal, which is due to the fact that the Fourier transform is an orthogonal projection. As the effect of sampling a signal on a spatial domain lattice \( V \) is to replicate its spectrum on the frequency domain lattice sites \( 2\pi(V^{-1})^T \). Then,

\[
  f_s(n) = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \left[ \frac{1}{|\det V|} \sum_k F((V^{-1})^T (\omega - 2\pi k)) e^{j\omega^T n} e^{-2j\pi k^T n} \right] d\omega
\]

(2.11)

Now, \( e^{-2j\pi k^T n} \) always has unit value for all possible values of \( n \) and \( k \). Hence, by defining a new variable \( F_s(\omega) \), the sampled signal can be computed analogously to an inverse Fourier transform as follows:

\[
  F_s(\omega) = \frac{1}{|\det V|} \sum_k F((V^{-1})^T (\omega - 2\pi k))
\]

or,

\[
  F_s(V^T \Omega) = \frac{1}{|\det V|} \sum_k F(\Omega - Uk)
\]

(2.12)

where \( U \) is the reciprocal lattice, or frequency domain lattice. The spectra of the sampled signal \( F_s(V^T \Omega) \) is a periodic extension of the original continuous signal \( F(\Omega) \). The sampling lattice \( V \) and the reciprocal lattice are related as \( U^T V = 2\pi I \). For the sampling lattice given in Figure 2.15(a), the reciprocal lattice in
the frequency domain is illustrated in Figure 2.15(b). The gray circles illustrate the individual spectra of the sampled function.

![Figure 2.15](image)

Figure 2.15 (a) A possible uniform sampling on the Euclidean plane (b) A Frequency domain sampling lattice [10]

It is evident from the figure that, to avoid aliasing there should be no overlap between adjacent copies of the spectra. Hence, it is important that the original signal is band or wave-number (using the terminology of Petersen [64]) limited. This means that the spectrum of the original function $F(\Omega)$ is equal to zero outside some region of finite extent $B$, which is known as the baseband or region of support. There is no constraint on the shape of $B$, though certain shapes make the reconstruction simpler. It is possible to vary the spatial sampling matrix $V$ so that there is no overlap among the periodically repeated versions of $F(\Omega)$. Consequently, there is no aliasing.

### 2.3.4 Features of Hexagonal sampling

#### 2.3.4.1 More Efficient Sampling Schemes

An insufficient sampling rate can always introduce unwanted effects in the reconstructed signal, referred as aliasing in any sampling scheme. Peterson and Middleton [64] investigated sampling and reconstructing wave number-limited multi-dimensional functions and concluded that the hexagonal sampling scheme as most efficient. Hexagonal sampling scheme uses a minimum number of sampling points to achieve exact reproduction of a wave-number-limited function which is not the case in general rectangular sampling scheme. When a two dimensional isotropic function is considered, the optimum sampling lattice
is the 120° rhombic (hexagonal) with spacing of sample points equal to $1/(B*\sqrt{3})$ if the spectrum of a function is bounded by a circle of radius $2\pi B$ in the wave-number plane (Figure 2.16). The sampling efficiency is 90.8%, compared with 78.5% for the largest possible square lattice.

![Figure 2.16 Optimum sampling lattice for two dimensional isotropic functions](image)

A similar conclusion was obtained by Mersereau [17] who developed a hexagonal discrete Fourier transform and hexagonal finite impulse response filters. Mersereau showed that for signals which are band-limited over a circular region in Fourier space, 13.4% fewer sampling points are required with the hexagonal grid to maintain high frequency image information compared with the rectangular grid. Thus, less storage and less computation time are required. An example is that in image coding application, one may expect that the coding efficiency can be increased by using the hexagonal sampling scheme. This can be proved as follows: Let the sampling matrix of rectangular lattice be

$$V_{rect} = \begin{pmatrix} T_1 & 0 \\ 0 & T_2 \end{pmatrix}$$

which simplifies to the commonly used regular (Cartesian) sampling for $T_1 = T_2$.

Hexagonal sampling is most conveniently described by the matrix

$$V_{hex} = \begin{pmatrix} T_1 & \frac{1}{2}T_1 \\ 0 & T_2 \end{pmatrix}$$

for $T_2 = \frac{\sqrt{3}}{2}T_1$.

$$V_{hex} = \begin{pmatrix} T_1 & \frac{1}{2}T_1 \\ 0 & T_2 \end{pmatrix}$$

$$for \quad T_2 = \frac{\sqrt{3}}{2}T_1$$

(2.14)
When \( V \) describes the sampling in spatial domain, the matrix \( U \) describes its reciprocal lattice, satisfying
\[
U^T V = I
\]  
(2.15)

where \( U^T \) being the transpose of \( U \) and \( I \) being the identity matrix, describes the positions of the replicas in frequency domain. \( U \) is therefore called the periodicity matrix. Then, we obtain the periodicity matrix for rectangular sampling
\[
U_{\text{rect}} = \begin{pmatrix}
u_1 & 0 \\ 0 & u_2
\end{pmatrix}
\]
(2.16)

with \( u_1 = \frac{2\pi}{T_1} \) and \( u_2 = \frac{2\pi}{T_2} \). The periodicity matrix for hexagonal sampling is
\[
U_{\text{hex}} = \begin{pmatrix}
u_1 & 0 \\ -\frac{1}{2}u_2 & u_2
\end{pmatrix}
\]
for \( u_2 = \frac{2}{\sqrt{3}}u_1 \)
(2.17)

where again \( u_1 = \frac{2\pi}{T_1} \). The 2D Fourier Transform \( F \) of a circularly band-limited signal has the property
\[
F(\omega_1, \omega_2) = 0 \quad \text{for} \quad \omega_1^2 + \omega_2^2 \geq W^2
\]  
(2.18)

where \( W \) is the maximum frequency in the data set. This baseband can be inscribed, for example in a square with length \( l = 2W \) (corresponding to rectangular sampling). In other words, \( u_1 \) and \( u_2 \) in Equation (2.16) have to be equal to \( 2W \). On the other hand, the baseband can also be inscribed in a hexagon with side length \( l = \frac{2}{\sqrt{3}}W \) (corresponding to hexagonal sampling). This means that in Equation (2.17), \( u_2 \) must be equal to \( 2W \). Calculating the sampling matrices from these periodicity matrices,
\[ V_{\text{rect}} = \begin{pmatrix} \frac{\pi}{W} & 0 \\ 0 & \frac{\pi}{W} \end{pmatrix} \]  

(2.19)

Where

\[ |\det V_{\text{rect}}| = \frac{\pi^2}{W^2} \]  

(2.20)

and

\[ V_{\text{hex}} = \begin{pmatrix} \frac{2}{\sqrt{3}} \frac{\pi}{W} & \frac{1}{\sqrt{3}} \frac{\pi}{W} \\ \frac{1}{\sqrt{3}} \frac{\pi}{W} & \frac{\pi}{W} \end{pmatrix} \]  

(2.21)

where

\[ |\det V_{\text{hex}}| = \frac{\pi^2}{W^2} \frac{2}{\sqrt{3}} \]  

(2.22)

The sampling density is proportional to \( \frac{1}{|\det V|} \). By taking the ratio

\[ \frac{|\det V_{\text{rect}}|}{|\det V_{\text{hex}}|} = \frac{\sqrt{3}}{2} = 0.866 \]  

(2.23)

From Equation (2.23), it is clear that hexagonal sampling requires 13.4\% less samples than rectangular sampling.

Vitulli [18] investigated the efficiency of hexagonal sampling and concluded that about 13\% less number of pixels are needed to obtain the same performance as compared to square sampling with the same signal. These conclusions are briefly illustrated below. Figure 2.17 (a) is a generic hexagonal sampling lattice. Vitulli showed that the Fourier Transform (FT) of a hexagonal lattice is still a hexagonal lattice. Fourier transform of the hexagonally sampled image is also composed of infinite replicas of the spectrum \( G(\xi, \eta) \), the FT of the image \( g(x, y) \) to be sampled. These replicas are centered in the points of the hexagonal lattice which is the FT of the hexagonal sampling lattice (Figure
2.17(b)). As a result, using hexagonal grid wider spectra can be sampled without aliasing with the same number of pixels or less pixels than using square grid.

Figure 2.17 Hexagonal sampling lattice and its Fourier transform [18]

2.3.4.2 Smaller Quantization Error

As mentioned earlier, in order to process an image by a digital computer the continuous image in real world must be quantized into spatial elements of finite dimensions, generally referred as pixels. Due to the limited resolution capabilities of image sensors, this array is usually too small to adequately represent the scene in real world. Quantization error thus is inevitable. In computer vision, quantization error is a very important measurement to investigate the merits of different types of sensory configurations in order to find which spatial sampling would introduce less quantization error into computations. Kamgar-Parsi [13, 65] developed formal expressions for estimating quantization error in hexagonal spatial sampling and found that, for a given resolution capability of the sensor, hexagonal spatial sampling yields smaller quantization errors than square sampling. Figure 2.18 shows the spectral packaging for best rectangular and hexagonal sampling.

Figure 2.18 Spectral packing for (a) Rectangular sampling (b) Hexagonal sampling [13]
2.3.4.3 Consistent Connectivity Definition

Connectivity between pixels is a fundamental concept that simplifies the definition of numerous digital image concepts, such as regions and boundaries. To decide if two pixels are connected, it must be determined if they are neighbors and if they satisfy a specified criterion of similarity [47]. On a square grid, there are two possible ways to define neighbors of a pixel. We can either regard pixels as neighbors when they have a common edge or when they have at least one common corner, so that four and eight neighbors exist (referred as a 4-neighborhood and an 8-neighborhood). On a square grid, object connectivity can be defined as 4-way to any of the four nearest neighbors or 8-way if connectivity to diagonal neighbors is permitted. This is illustrated in Figure 2.19 (a) and (b).

Figure 2.19 Neighborhood relationships in Square grid and Hexagonal grid

Consider the pattern shown in Figure 2.20(a). Assuming 4-way connectivity for the object and the background, the number of vertices H in the
pattern is 16, the number of edges $E$ is 16, and the number of faces $F$ is 4. Application of the Euler formula $H-E+F$ to the pattern should give its genus. Thus, by the above formula the genus is $16-16+4 = 4$. The hexagonal grid, however offers no connectivity choice. We can only define a 6- neighborhood. Neighboring pixels have always one common edge and two common corners (Figure 2.19(c)). The absence of such choice in hexagonal grid results in easier and more efficient algorithms, such as thinning algorithm [67, 68, 69] since fewer connectivity situations have to be accounted for. Accordingly, connectivity in hexagonal objects is consistent as it is six-way to either of the nearest neighbors for both the object and the background image components [14, 15]. Assuming 6-neighbor connectivity, the number of vertices $H$ in Figure 2.20(b) is 24, the number of edges $E$ is 30 and the number of faces $F$ is 6. Using the Euler formula, the genus is equal to $24-30+6 = 0$.

2.3.4.4 Equidistance

With the introduction of neighborhood relation, distance function can be easily measured. In square grid we have two types of distances, where the distance between adjacent pixels in the diagonal direction is $\sqrt{2}$ times that of in the horizontal (or vertical) direction which is a unit distance (Figure 2.21(a)). While in hexagonal case, each hexagonal pixel has only six neighboring pixels and each pixel is equidistantly adjacent to their six neighbors along the six sides of the pixels. The centroid of the middle pixel is at the same distance from the centroids of the six adjacent pixels (Figure 2.21(b)).

![Figure 2.21 Distance in (a) square grid and (b) Hexagonal grid [66]](image-url)
2.3.4.5 Greater Angular Resolution

Image processing on a hexagonal lattice is also advantageous due to its greater angular resolution to represent curved objects. It has been noted that hexagons offer greater angular resolution as the nearest neighbors of the same type are separated by 60° instead of 90° [13]. An example showing a familiar curved figure and a representation on hexagonal and square lattices are shown in Figure 2.22.

Figure 2.22 Curved figure represented in (a) Hexagonal grid and (b) Square grid [66]

Notice that the hexagonal case shown in Figure 2.22(a) appears to have smoother curves than the square case. There are several reasons for this. The first is due to the consistent connectivity in the hexagonal lattice. This means that all neighbors are uniform distances away from each other and leads to the smoother curvature. Another reason is the oblique effect in human vision [70]. This means that we have a visual preference for lines at oblique angles. This also helps to make the hexagonal curves look smoother. The theory developed and the simulation done on a physical screen showed that hexagonal grids represent a reasonable alternative to conventional square grid display techniques not only for circle drawing which was somehow predictable, but also for straight lines. Digitization on hexagonal grid display yields a better connectivity and is perceived as being approximated by small polylines whereas on square grid, digitization is still perceived as being approximated by pixels.

2.3.4.6 Higher Symmetry

Serra [71] has developed many of morphological operators that were currently used for image processing. He prefers the hexagonal grid to the
rectangular because of the connectivity definition and the higher symmetry, which lead to simpler processing algorithms. It can be seen in Figure 2.20(b) that the cluster of hexagonal pixels possesses the same symmetry about the three different lines connecting pairs of two pixels and the central pixel. This symmetry degree is one higher than that of square grid and this feature makes image processing more accurate. For example, when an image on a hexagonal grid is rotated, more image information will be retained on it compared to the rotation on square grid.

2.3.4.7 Other Features of Hexagonal Grid

- **Isoperimetry**: As per the Isoperimetric theorem, a hexagon encloses more area than any other closed planar curve of equal perimeter except a circle [10]. This implies that the sampling density of a hexagonal lattice is higher than that of a square lattice.

- The image over the hexagonal grid has more clarity and visual appeal than that over the rectangular grid. Modeling of digital images over hexagonal pixel grid has a very positive effect on image clarity and appearance of features [20].

- The research done in the animal vision [6, 7, 62] clearly demonstrates that the arrangement of rods and cones in the fovea is more nearly approximate a hexagonal tessellation than a rectangular one (Figure 2.5).

- The improvements in CCD technology make hexagonal sampling feasible or practical applications and bring a new interest on this topic [18]. A customized fiber bundle comprising 19 100-micron hexagonal-packed fibers was designed with all but one fiber utilized for sky light collection for a 4STAR Spectrometer proposed by S. Dunagan et.al. [72]. A sun tracker based on CCD array technology is under consideration, that would better discriminate the solar disk from other glint or bright scattering sources that contaminate the quadrant detector signal. Additionally, the CCD radiometric measurements would be of value in spatially mapping cloud edge scattering.

- An integrated CMOS image acquisition system suitable for retinal information processing in bionic vision systems or product inspection has been developed and tested. The test chip incorporates 32 x 32
random addressable active sensor cells arranged on a hexagonal grid and on-chip control and readout electronics. A noteworthy example is the 4th generation super CCD of Fuji Film that uses pixel elements in a hexagonal-like arrangement [18].

![Figure 2.23 Super CCD by Fuji Co [18]](image)

- Besides this, the SPOT 5 satellite using a couple of identical linear CCD transmits on-ground two images that are quasi-identical but shifted by half a pixel. Then the two images are interwoven and interpolated. After further processing, the resulting image is similar to that would have been obtained using a hexagonal detector (Figure 2.23) [18].
- Thus by modeling and processing images on such a grid space, one can mimic the natural behavior to realize computer vision.

### 2.4 Acquisition of Hexagonal Sampled Images

The acquisition stage deals with generating image data from a real world source. This can be performed via a camera, a scanner or some more complex input device. The data may require additional processing before the acquisition is complete. Processing involves manipulation of the image data to yield meaningful information. This could be the application of a simple linear filtering algorithm using a convolution operator or something more complicated such as extracting a detailed description of the structures contained in the image. The visualization stage is very useful and often essential for human observers to make sense of the processed information. To achieve this, it is
necessary to define the lattice or a grid on which the visual information is processed. The lattice of interest here is the hexagonal lattice. There are two approaches for the acquisition of hexagonal lattice from the square lattice as (i) Hardware based approach known as complete Hexagonal Image Processing (HIP) system (Figure 2.24) and (ii) Software based approach known as Mixed system approach (Figure 2.25).

**Figure 2.24 Complete HIP system [10]**

**Figure 2.25 Mixed system approach [10]**

### 2.4.1 Hardware based Acquisition

A hexagonal image can be acquired in a cost-effective way by modifying an existing hardware system to perform hexagonal sampling. R.C.Staunton [21] designed pipeline architecture to take video images and by introducing a delay to alternate lines produced a hexagonal sampled image. For the industrial inspection application the hexagonal grid processing system was able to outline defects as reliably and accurately than the square processing system. There has also been much interest in building custom hardware for
acquiring hexagonal images. R.C.Staunton [19] notes that the technology to fabricate hexagonal grids does exist as it is widely used for large RAM devices. A pioneer in the hexagonal sensor field is Carver Mead [2] who has built sensors mimicking various biological sensors including the retina. The last fifteen years has witnessed increased activity in custom-built hexagonal sensors, many of which are CMOS based. These range from general purpose to application specific. Superior ability of hexagonal grids to represent curves has motivated a CMOS fingerprint sensing architecture to be developed based on the hexagonal lattice [73, 74]. With the ability to grow crystals in space, several projects have been performed to grow hexagonal sensing arrays for applications such as satellite sensors and replacing the human retina [75] after it has been damaged. Hexagonal sensors also find a place in remote sensing [76]. Though a lot of advantages are associated with the hexagonal acquisition and display systems, such devices are not readily available. Due to lack of hexagonal display systems image processing operations have been simulated using the existing rectangular lattice display devices (software based approach) which are discussed in the next section.

2.4.2 Software based Approach and Image Representation in Hexagonal Sampled Grid

Manipulating sampled data on one lattice to produce the data on a different lattice is termed resampling. In the current context, the original data is sampled on a square lattice while the desired image is to be sampled on a hexagonal lattice. Various software based approaches of generating hexagonal lattice from rectangular lattice are reviewed in this section.

2.4.2.1 Hexagonal Lattice and Triangular Pixels

The approach of Hartman [32] used a hexagonal lattice and triangular pixels. The construction process is illustrated in Figure 2.26 where black squares indicate square pixels. Two square pixels that are vertically adjacent are averaged to generate each individual triangular pixel. Hartman observed that the resulting pixel does not produce perfect equilateral triangles but instead a triangle with base angles of $63.4^0$ and a top angle of $53.2^0$. 

62
2.4.2.2 Hexagonal Orthogonal-Oriented Pyramid

With the goal of deriving an image transformation that was mathematically consistent with the primary visual cortex, Watson and Ahumada [22] proposed the hexagonal orthogonal-oriented pyramid. The square to hexagonal conversion process used the affine relationship between the square and hexagonal lattice points. This means that rectangular images are skewed to form hexagonal images. After the skewing process, a hexagon becomes elongated in an oblique direction which is illustrated in Figure 2.27. A consequence of this stretching is that the hexagonal shape no longer exhibits as many degrees of rotational symmetry. For the target application of image compression under consideration, this distortion was deemed unimportant.

2.4.2.3 Sampling by Half a Pixel Width

Horn [77] has described how a practical hexagonal data may be captured by delaying sampling by half a pixel width on alternate TV scan lines in
horizontal direction (Figure 2.28). In his scheme, the pixel shape is square. In other words, the sampling intervals in horizontal and vertical directions are identical. This scheme simplifies the hardware design by setting identical sampling intervals in both horizontal and vertical direction. However, the equidistance property of hexagonal pixels is not preserved. As shown in Figure 2.28, if we denote the distance between any two neighbors in horizontal and vertical direction as 1 unit, the distance between any two neighboring pixels in diagonal direction will be $\sqrt{3}/2$.

![Figure 2.28 Hexagonal scheme of Horn [77]](image)

2.4.2.4 Hexagonal Data Structure with a Rectangular Shape

Staunton [23] described a hexagonal data structure with a rectangular shape, where all the six neighbors of a pixel lie on a circle. The centre of the circle is the sampling point as illustrated in Figure 2.29. The major advantages with this structure are that, all sampling points are equidistant from their nearest neighbors, the angle subtended by two nearest neighboring points is $60^0$, and the horizontal sampling distance is $2/\sqrt{3}$. The pixel size is $\frac{1}{2\sqrt{3}}$ and thus for systems employing an equal number of pixels horizontally and vertically, the image aspect ratio would be $2/\sqrt{3}:1$. 

64
2.4.2.5 Pseudo Hexagonal Pixel

Wuthrich and Stucki [24] proposed a pseudo hexagonal pixel in order to evaluate the visual effect of hexagonal pixel and square pixel (Figure 2.30).

A comparative simulation of two screens based on the square and the hexagonal lattices has been made in using this approach. A hexagonal pixel, called a hyperpel, is simulated using a set of many square pixels and the simulated square grid had to be adapted in order to make its density comparable with the hexagonal grid. The resulting hyperpels, which are illustrated in Figure 2.30, were displayed at a resolution of 60 x 60 pixels, both for the hexagonal and for the square grids [24]. This results in a loss in the screen resolution [66].

2.4.2.6 Quincunx Sampling

Another approach to generate hexagonal sampled images is via the use of quincunx sampling. These samples are arranged as in a chessboard as illustrated in Figure 2.31. Laine et.al. [25] followed this approach and used
linear interpolation to double the image size in the horizontal direction and triple it in the vertical direction. The main purpose of the interpolation step is to scale the image in a way that emphasizes the hexagonal arrangement. The interpolated image was masked to produce the quincunx pattern, following which the remaining data was mapped on to a hexagonal grid.

![Image of linear interpolation and quincunx pattern]

Figure 2.31 Quincunx sampling [10]

2.4.2.7 Spiral Architecture

Unlike the square lattice, the points in a hexagonal lattice do not easily lend themselves to be addressed by integer Cartesian coordinates. This is because the points are not aligned in two orthogonal directions. Due to the nature of the hexagonal lattice an alternative choice for the coordinate axes would be the axes of symmetry of the hexagon [10, 66]. Labeling in rows and columns order is possible if the lattice is represented by true hexagon. Labeling in the mimicking hexagonal lattice are not representation of the spatial coordinates in the true sense. In order to properly address and store hexagonal images data Sheridan et.al. [28] proposed a one-dimensional addressing scheme for a hexagonal structure, with the definitions of two operations spiral addition and spiral multiplication. This hexagonal structure is called the Spiral Architecture (SA) (Figure 2.32). In the Spiral Architecture, spiral addition and spiral multiplication correspond to image translation and rotation respectively. The result of an addition or a multiplication is a spiral address, and it can be computed based on the spiral addition between the seven addresses from 0 to 6.
Middleton and Sivasamy [10, 26, 27] proposed sampling as a tiling in which a sampling process is performed by dividing the Euclidean plane into regular and reproducible regions and analyzing the signal in each region. Formally, a tiling, $T$, is a collection of closed sets, $T = \{T_1, T_2, \ldots, T_L\}$, which cover the Euclidean plane without gaps or overlaps, $T_i \cup T_j \neq \emptyset$ where $i \neq j$. Each $T_i$ is a series of points that completely define the tile and all such occurrences of the tile. The set of points within $T_i$ which completely describe a single tile is known as the prototile of the tiling $T$. The concatenation of the tiles is the entire Euclidean plane. A tiling in which all the individual tiles are of the same size and shape is called a monohedral tiling. Three such monohedral regular tilings are possible. These use triangles, squares, or hexagons as the prototile. The hexagon has rotational symmetry at multiples of $\pi/3$ radians and the square has rotational symmetry at multiples of $\pi/2$ radians. Thus, the hexagon exhibits 6-fold rotational symmetry and the square exhibits 4-fold rotational symmetry. As an example, Figure 2.15(a) illustrates a possible uniform sampling of the spatial domain on the Euclidean plane which is roughly hexagonal with the horizontal spacing being twice the vertical. Figure 2.15(b) shows the Frequency domain sampling lattice. Middleton and Jayanthi Sivaswamy [10] also proposed a one-dimensional addressing system addressing by modifying the Generalized Balanced Ternary system (GBT) which is elaborated in section 2.5.3.

2.4.2.8 Virtual Hexagonal Structure

On spiral architecture proposed by Philip Sheridan [30], each pixel is identified by a designated positive number. The numbered hexagons form the
size $7^n$, where $n$ is the layer number. The hexagons tile the plane in a recursive modular manner along the spiral direction. Qiang Wu et.al. [31] proposed virtual spiral architecture which exists during the procedure of image processing. In order to make such mimicking methods reliable, virtual spiral architecture retains the resolution of the image on virtual spiral space.

![Figure 2.33 A square grid and hexagonal grid which has the same area](image)

Figure 2.33 A square grid and hexagonal grid which has the same area

For a given picture represented on rectangular architecture, if it is represented on spiral architecture on which each hexagonal grid has the same area as square grid on rectangular architecture, the image resolution is retained which is shown in the Figure 2.33.

![Figure 2.34(a) One virtual hexagonal grid and its four connected square pixels (b) One square grid and its three connected virtual hexagonal grids [31]](image)

Figure 2.34(a) One virtual hexagonal grid and its four connected square pixels (b) One square grid and its three connected virtual hexagonal grids [31]

In Virtual Spiral Architecture and rectangular architecture, each grid is considered as a set which is composed of many small points [31]. The accuracy of approximation will be improved when the number of small points in a grid is
increased. The gray value of hexagonal grid is calculated as the weighted average of the grey values of the connected square grids in this method (Figure 2.34(a)). The reverse operation is performed in order to map the images from virtual spiral architecture to rectangular architecture after image processing on spiral architecture (Figure 2.34(b)). It was proved by Qiang Wu et.al. [31] that compared with common mimic Spiral Architecture, performance of Virtual Spiral Architecture is better and it retained the image resolution during mimicking procedure. Such virtual spiral architecture only exists during the procedure of image processing. It builds up a virtual hexagonal grid system on memory space on computer. Then, processing algorithms can be implemented on such virtual spiral space. Finally, resulting data can be mapped back to rectangular architecture for display (Figure 2.35).

This mimicking operation significantly does not introduce distortion or reduce image resolution, which is the most remarkable advantage over other mimicking methods, while keeping the isotropic property of the hexagonal architecture. The virtual hexagonal structure is very much useful for hardware implementation.
2.4.2.9 Alternatively Suppressing Rows and Columns of the Existing Rectangular Grid

E.G. Rajan [20] shows that a hexagonal lattice can be simulated on the rectangular grid as shown in Figure 2.36(b). Each pixel of this new grid is surrounded by six nearest neighbor pixels in a hexagonal structure. We can obtain the hexagonal grid by alternatively suppressing rows and columns of the existing rectangular grid. The sub sampling procedure is given by the rule as follows:

\[
pixel_{val\_hex\_hex\_i, j} = \begin{cases} 
\pixel_{val\_hex\_2*i, 2*j} & \text{if } i \text{ is even} \\
\pixel_{val\_hex\_2*i, 2*j+1} & \text{if } i \text{ is odd}
\end{cases}
\]

(2.24)

All the pixels in the rectangular grid that do not have any correspondence to those in their hexagonal counterparts are suppressed. Suppressing is assigning the value of these pixels to zero. While processing this sub sampled image these suppressed pixels are not considered in computation. However, the newly constructed hex lattice has only one-fourth of the number of pixels, as that of the rectangular grid. So, the modeling and processing of images on the hexagonal lattice is not comparable to that over the given regular rectangular lattice. For such comparison purposes we need to obtain a rectangular lattice that has the same resolution as that of the modeled hexagonal lattice which is the limitation of this method.

Figure 2.36(a) Rectangular Pixel Grid (b) Simulated Hexagonal Grid [20]
2.4.2.10 Half Pixel Shift

Senthil Periasamy [33] proposed half pixel shift method to obtain hexagonal lattice from the conventional rectangular lattice which is described as follows: For each odd line, find the midpoint between two adjacent pixels by simple linear interpolation \( \text{mid} = (\text{left} + \text{right}) / 2 \). Discard the left and right, keeping only the mid values.

\[
P_{\text{new}}^{\text{old}}(x, 2y) = P_{\text{old}}^{\text{old}}(x, 2y)
\]

\[
P_{\text{new}}^{\text{new}}(x, 2y + 1) = \left( P_{\text{old}}^{\text{old}}(x, 2y + 1) + P_{\text{old}}^{\text{old}}(x + 1, 2y + 1) \right) / 2
\]

This gives us a hexagonal mapping from a regular square or rectangular grid. It was proved by Senthil Periasamy [33], that the half pixel shift approach was computationally more effective due to three axis symmetry of hexagonal pixels. This approach was used for designing hexagonal wavelet and enhancement was performed on each directional channel, i.e., a particular direction can be separately enhanced. Using the hexagonal wavelet, it is possible to obtain certain features that exist in particular scales can be selectively in one direction, and to suppress features in orthogonal direction. Hexagonal wavelet using half pixel shift method was used for detection of microcalcification of mammogram images. It has been determined experimentally by using this approach results in 100% True Positive identification. Control of the various parameters during the detection phase is extremely simple (basically a threshold factor and a calcification amplification factor) and is easily understood which makes the process intuitive. Hence, the half pixel shift method is computationally more effective for the wavelet based applications if the mimicking hexagons are used.

Therefore, initially the half pixel shift method was used for the applications such as edge detection, image interpolation and enhancement. However, the results in Hex-Gabor interpolation (chapter 3) disputes the fact that such methods require radial distance corrections and they defeat edge enhancement as edges get averaged out. In the case of three direction Gabor processing as the edges are enhanced in one direction, image smoothening
occurs in orthogonal direction. This method is used to enhance the intrapixel information automatically, specifically when \( \sigma = 2/\pi \) which is elaborated in chapter 3.

![Conversion of a regular square grid to a hexagonal grid](image)

Figure 2.37 Conversion of a regular square grid to a hexagonal grid [33]

### 2.5 Hexagonal Structure Addressing

Generally, addressing and storage are important issues when it comes to image processing. The specific storage mechanism can often drastically affect the performance of a system. In the human visual system (HVS) discussed in the section 2.2, the Lateral Geniculate Nucleus (LGN) serves to organize the information from the retina for use in the rest of the visual system. There are two sorts of arrangements that are apparent in the LGN. Firstly, is spatiotopic where information from neighboring sensors is stored near each other. And, secondly is hierarchical organization which selectively pools information from many sensors producing a variety of resolution-specific arrangements. The visual system manages to achieve these arrangements via the way in which it is organized. Such features are also desirable in computer vision systems, since neighboring pixel information plays an important role in image manipulation and easy access to such information can impact system performance. The existing addressing schemes suitable for hexagonal structure are reviewed in this chapter. There is no specific addressing scheme is described yet. A new addressing scheme is proposed (Chapter 4) which is suitable for hardware implementation of image processing operations.
2.5.1 Two-Axes Oblique Coordinate Addressing Scheme

We can address the hexagonal images using Cartesian co-ordinates, needing only two filters when the filter is isotropic. When diffusion along different directions is to be different and nonlinear, four orientations have to be used in Cartesian coordinates. Additional benefits of orthogonal directions to manipulate an intra pixel are difficult in terms of edges. As a result of this, weighted four orientation masks are employed using some soft computing technique for interpolation [55, 120, 130, 131].

Two oblique axes (Figure 2.38) to address hexagonal structure was suggested by Luczak et.al. [78], where two basis vectors are not orthogonal. Due to the nature of the hexagonal lattice, an alternative choice for the coordinate axes would be the three axes of symmetry of the hexagon. This is convenient as it will provide purely integer coordinates for every point in the lattice. Since there are more than two axes of symmetry, many schemes have been developed for addressing points on a hexagonal lattice. The simplest way to address points on a hexagonal lattice is to use a pair of skewed axes which are aligned along axes of rotational symmetry of the hexagon [10]. This will yield integer coordinates and is efficient as two coordinates are sufficient to represent a point on a plane. There are two distinct possibilities for the skewed axes as illustrated in Figures 2.38(a) - (d) where each axis has a 60 degree and 120 degree separation.
2.5.2 Three-Coordinate Symmetrical Coordinate Frame

Her [29] developed a symmetrical hexagonal coordinate frame, denoted as $\mathbb{R}^3$, for hexagonal grid. This uses three coordinates $x$, $y$, $z$ instead of two to represent each pixel on the grid plane, as shown in Figure 2.39. The three coordinates at any pixel has a relationship among them as $x + y + z = 0$. Here the distance between two neighboring grid points is defined as one unit. The major advantage of this coordinate system is a one-to-one mapping between $\mathbb{R}^3$ and the 3-dimensional Cartesian frame $\mathbb{R}^3$, as illustrated in Figure 2.40, where $x$, $y$ and $z$ are the three orthogonal axes of $\mathbb{R}^3$. As a result of this many geometrical properties of $\mathbb{R}^3$ can be readily transferable to $\mathbb{R}^3$. Since the $x$ and $y$ coordinates of a point of this symmetrical hexagonal coordinate frame $\mathbb{R}^3$ are actually the two coordinates used in the oblique coordinate frame. Theories and equations previously developed for the oblique coordinate frame can directly be used in $\mathbb{R}^3$. Moreover, the use of this symmetrical hexagonal coordinate frame is demonstrated to derive various affine transformations by Her [29].

Due to the physical relationships between the symmetrical hexagonal coordinate frame and the 3-dimensional Cartesian frame $\mathbb{R}^3$, geometric transformations on the hexagonal grid are conveniently simplified and the symmetrical property of the hexagonal grid is successfully preserved. This three-axis coordinate system is used for mathematically handling the hexagonal structure, i.e., numerically calculating the distance between two objects. This coordinate system reflects the geometrical symmetry of the grid. This three -
coordinate symmetrical frame is mainly used in geometric transformations such as projection, rotation, scaling and reflection.

Figure 2.39. Symmetrical hexagonal frame $*R^3$ [29]

Figure 2.40 Relation between frames $*R^3$ and $R^3$ [29]

2.5.3 Single Indexing System

Sheridan [30] proposed a one-dimensional addressing system as well as two operations based on this addressing system for hexagonal structure (Figure 2.41). Q. Wu, X. He and T. Hintz [31] developed an algorithm to realize image rotation without scaling which improves the spiral architecture’s usage in image processing. Middleton and Jayanthi Sivaswamy [10] also proposed a one-dimensional addressing system addressing by modifying the Generalized Balanced Ternary system (GBT) which was shown in Figure 2.32.
This addressing scheme takes inspiration from the Human Visual System (HVS) which has some or all of the following features: (i) Spatiotopic arrangement (ii) Hierarchical (iii) Computational efficiency. This address grows from the centre of image in powers of seven along a spiral like curve [10]. This addressing scheme is combined with two mathematic operations, spiral addition and spiral multiplication which correspond to image translation and image rotation respectively. The use of the balanced ternary indexing scheme is central to the Hexagonal Image Processing (HIP) framework. The indexing scheme can be constructed intuitively by tiling the image. Consider a single hexagon to be a tile at layer 0. This can then be surrounded moving anti clockwise by a further 6 hexagons (Figure 2.32). This new structure forms a tile at layer 1, with the new tile being the super-tile of layer 0. Similarly, this process can be extended to any number of layers. The number of hexagons that are contained in a given layer $\lambda$ are $7^\lambda$. Each hexagon can then be numbered uniquely as a sequence of numbers where every digit gives its position in a given tile. The HIP addressing scheme that has been described is a radix 7 positional number system.
A square sampled image of size $M \times N$ pixels with a quantization depth of 24 bits will have a size of $3MN$ bytes [10]. Similarly for a $\lambda$-level image the required storage space is $3 \times 7^\lambda$. Given a square sampled image, the number of layers for an equivalent hexagonal sampled image can be found by equating the numbers of points:

$$\lambda = \frac{\log M + \log N}{\log 7}$$

Due to the nature of the addressing scheme, the value of $\lambda$ is also the number of digits in the corresponding addresses for the points in the image. For example if a square image of size $128 \times 128$ is to be examined, the formula gives $\lambda \approx 5$ as the number of layers in the equivalent hexagonal image. This structure contains 16807 points compared with a square array which contains 16384 points. In the spiral addressing scheme, increasing in number of layers leads to more computational complexity.

In this section, the addressing schemes such as two - axes oblique coordinate addressing scheme, three-coordinate symmetrical frame and single addressing scheme proposed by various authors were discussed. The advantages and limitations of the existing addressing scheme were also addressed. In this research work, a new addressing scheme suitable for hardware implementation is proposed and the functionality of this addressing scheme is explained in chapter 4.

2.6 Limitations of the existing work and need for further work

- About 97% receptive field of the neurons is very closely described as 2D Gabor wavelet and it is mostly suitable for vision system modeling. Immense work is available on texture information, especially for rectangular structures. However, there is a little work in recognizing minute details in an image on hexagonal structure by either interpolation or enhancement. In this work, the two important operations of biological visual system such as enhancement and interpolation are considered and performed using the Hex-Gabor process which are elaborated in chapter 3..

- In the past years, there had been many attempts in representing hexagons in the regular square lattice or in a Spiral Addressing Scheme (SAS) [11-32]. Regular
geometry is only kept in the case of SAS, however processing such pseudo lattices gave rise to better results compared with square lattices. Even though consistent gradient operators were designed [79] for the edge detection on hexagonal lattice, it has not proved its efficiency on the mimicking scheme other than spiral addressing scheme [113]. The reasons for the former are (i) The intra-pixel distances in the two orientations differ and (ii) Pixels do not physically exist in the processed square lattice space. Hence, two solutions are proposed for the edge detection such as (i) CLAP algorithm based edge detection suitable for hardware implementation (Chapter 4) (ii) Hex- Gabor based edge detection (Chapter 3).

- A new addressing scheme is proposed which is suitable for CLAP algorithm based edge detection. The operation of new addressing scheme is discussed briefly in section 4.5.6.1. The features of new addressing scheme can be summarized as (i) It is possible to use the proposed addressing scheme for both rectangular and hexagonal lattice structures by suitable read/write operation. (ii) The architecture designed using new addressing scheme is capable of producing one edge-pixel every clock cycle. (iii) As CLAP algorithm is used for edge detection operation using new addressing scheme, multipliers are not used which leads to less hardware requirements.

- Human mind with the help of eyes, has a capability to see objects under occlusion. Thus, this aspect of building image information at least partially from vision is another area for research. The author has attempted to thin and reconstruct images using the Hilditch algorithm in hexagonal domain which is addressed in Chapter 5.