2

Current approaches to vision

Many advances in the physical sciences have come from examination of the world around us. By performing experiments our understanding of the governing laws of the universe expands and thus we are able to build systems to take advantage of our newly acquired knowledge. The early understanding of the nature of light came from understanding and trying to mimic our eyes. The models were incorrect but they provided the foundation for all future research which truly was performed standing upon giants' shoulders. So it was in the early days of the science of image processing. The primary motivation was to recreate the abilities of the visual system in modern computing devices.

In the years since its inception the science of image processing has forked many times, each time with a resulting name change. Many of these forks disregarded the influence of the visual system when devising image processing algorithms. However, due to the rapid rise in computational power in recent times, it is possible to accurately model portions of the brain. This has led to a resurgence in research into image processing using the results from biological visual system.

It is with these ideas in mind that this chapter provides an overview of both biological and computer vision. The study of the visual system will lead to an architecture in which the brain performs its visual processing. This generic architecture will be then be applied to the study of conventional vision. Central to this thrust is the specific hexagonal arrangement implicit in the visual system's sensor array. Logically, this arrangement affects all other aspects of the visual system. In line with the historical perspective, the biological system will be discussed first followed by current computer vision approaches.

2.1 Biological vision

The complexities of the brain and all its subsystems have fascinated mankind for an extremely long time. The first recorded references to brain dissection and study date to Galen in the 2nd century AD though there is evidence, in papyri, that the Egyptians were also interested in brain function [3]. The awareness of a distinct subsystem associated with vision dates from the 18th century [4].

This section will provide a brief overview of some key aspects of the human visual system (HVS) which occupies two thirds of the human brain's volume. The first of part of the HVS is the eye. The eye performs a similar function to a smart sensor array in a modern camera. The key feature of the visual system is that it performs the processing of the information using a hierarchy of cooperative processes.

2.1.1 The human sensor array

The visual system, according to Kepler who founded the modern study of the eye in 1604 [5], begins when"the image of the external world is projected onto the pink superficial layer of the retina". Later, Descartes [6] studied the optics of the eye and Helmholtz [7] studied the retina. This early work promoted the view that the eye performed in the same way as a pinhole camera (or camera obscura). Advances made since then however, have led to the view held today that the eye is more sophisticated and functions like a mini-brain. We will now briefly explain the reasons for this view.

The eye is roughly spherical with a slightly protruding part that is exposed while the remaining part sits in the eye socket. The light enters the eye through the pupil behind the cornea and is projected by a lens onto the inner spherical surface at the rear part of the eye. Here, the light is converted into electrical signals in an array of interconnected nerve cells known as the retina. 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.

There are two distinct sorts of photoreceptors, namely, rods and cones, with their nomenclature stemming from their shapes. Their functions are mutually complementary as summarised in Table 2.1.

A remarkable feature of the photoreceptive layer of the retina is that the rods and cones are distributed non-uniformly as illustrated in Figure 2.1. There is a radial distribution of these receptors: cones are concentrated in the central foveal region and, as one moves away from the centre, rods are found in abundance but gradually diminish in number. The foveal region, rich in cones, specialises in high resolution, colour vision under bright illumination such as during the day. This region however, is very small in extent. The region outside the fovea is rod-rich and hence contributes towards vision under low levels of illumination such as during night time. The field of view afforded by

Table 2.1. Differences between rods and cones.

Rods	Cones
high sensitivity	low sensitivity
more photopigment	less photopigment
high amplification	lower amplification
slow response	fast response
low resolution	high resolution
achromatic	chromatic (red, green, blue)
night vision	day vision

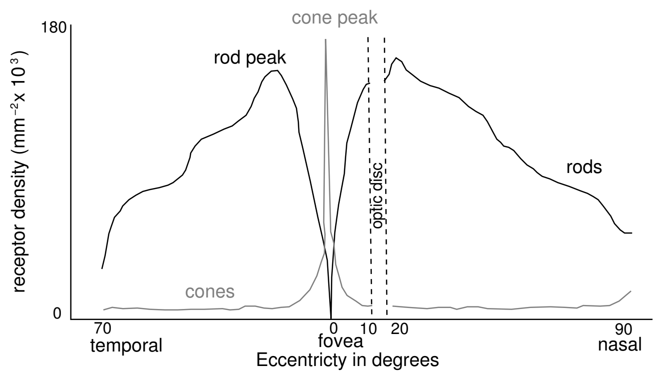
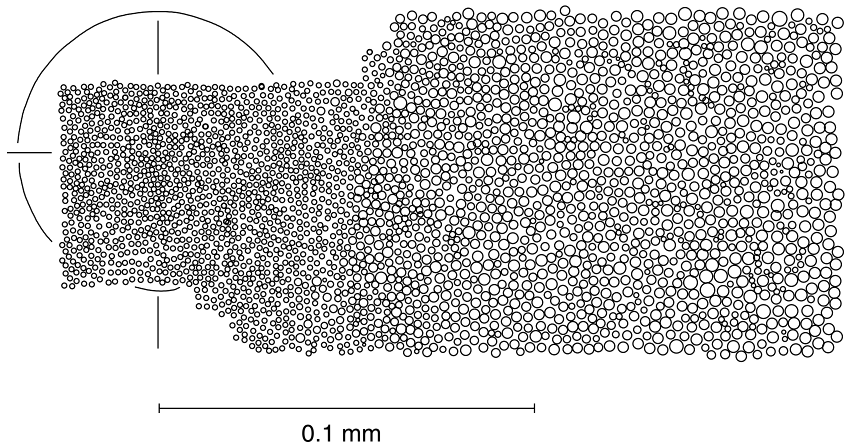


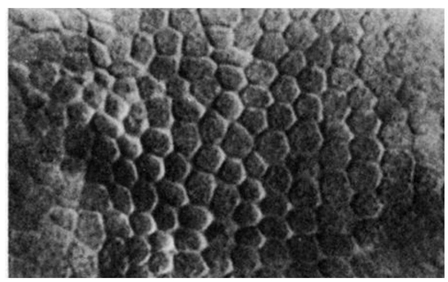
Fig. 2.1. Distribution of rods and cones in the retina (redrawn from Osterberg [8]).

high resolution and colour vision sensors is complemented by a combination of eye and head movements. The arrangement of the photoreceptors along the spherical retinal surface, is illustrated in Figure 2.2(a). 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, as we shall see later, all naturally deformable circular structures pack best in two dimensions within a hexagonal layout such as found in honeycombs. An example of an enlarged portion of the foveal region of the retina, showing this behaviour, is given in Figure 2.2(b).

The signals from the photoreceptors are preprocessed by a neuronal assembly made of four major types of neurons: bipolar, horizontal, amacrine, and ganglion. Of these cells, the horizontal and amacrine are purely used as lateral connections joining remote regions. The lateral connections enable receptors to influence each other and help in contrast correction and adaptation to sudden changes in ambient illumination. The ganglion cells are specialised for processing different aspects of the visual image such as movement, fine spatial detail, and colour. Two of the widely studied types of ganglion cell are the magno and parvo cells. Functionally speaking, these two types of cells give rise to the formation of two distinct pathways (called the M and P pathways)



(a)



(b)

Fig. 2.2. (a) Arrangement of rods and cones in eye adapted from Pirenne 1967 [9] (b) A close up of the foveal region (reprinted from Curcio et al. [10]. Copyright 1987 AAAS).

through which visual information is passed to the brain and processed. The magno cells have large receptive fields due to their large dendritic arbours, and respond relatively transiently to sustained illumination. Thus, they respond to large objects and follow rapid changes in stimulus. For this reason it is believed that magno cells are concerned with the gross features of the stimulus and its movement. On the other hand, the more numerous parvo ganglion cells have smaller receptive fields and selectively respond to specific wavelengths. They are involved with the perception of form and colour and are considered responsible for the analysis of fine detail in an image. The ganglion cells are collected together in a mylenated sheath at the optic disk to pass the visual information to the next stage in the visual system in the brain.

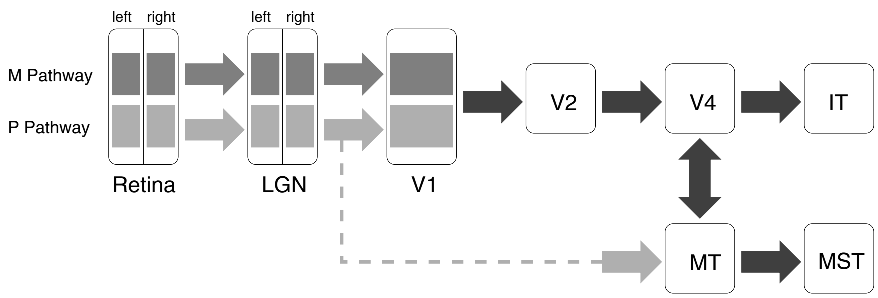


Fig. 2.3. Information flow from the retina to different cortical areas.

2.1.2 Hierarchy of visual processes

The visual system is the most complex of all the various sensory systems in the human brain. For instance, the visual system contains over 30 times the number of neurons associated with auditory processing. The sampled and pre-processed visual input is passed on to a mid-brain structure called the lateral geniculate nucleus (LGN) and then to the visual cortex. A diagram illustrating the information flow in the human visual system is shown in Figure 2.3.

The LGN is a six-layered structure that receives inputs from both the left and right visual fields via a crossover mechanism. The mapping of information from different eyes to the LGN is retinotopic which means cells in adjacent retinal regions project to cells in adjacent regions in the LGN. Additionally, inputs received by the two eyes from adjacent regions of the visual field are mapped to LGN to preserve the spatial relationships. Hence, the information mapping to LGN is spatiotopic as well. This nature of mapping continues to the first stage of the visual cortex (V1) which is about 2mm thick and also consists of a six-layer structure. The M and P pathways project to distinct sub-layers within this layer. The main function of V1 is to decompose the results from the LGN into distinct features which can used by other parts of the visual system.

Studies of the receptive fields of cells in V1 have found the cells to be considerably more complex than the cells in the retina and the LGN [11]. For instance, the LGN cells respond to spots (circles) of light whereas the simple cells in V1 respond to bars of light at specific orientations. The complex cells in V1 appear to pool outputs of several simple cells as they respond to orientation but not to the position of the stimulus. The cells in V1 are further organised into large functional structures. These include orientation-specific columns, ocular dominance columns, and colour-sensitive blobs. Neurons with similar responses but in different vertically oriented systems are linked by long range horizontal connections. Information thus flows both between the layers and between the columns, while remaining within the layer. This pattern of interconnections seems to link columnar systems together. For instance, such linkage might produce an effect of grouping all orientations in a specific region

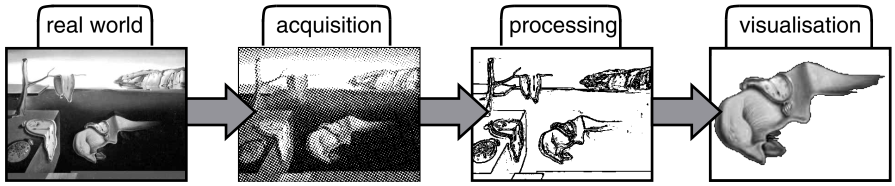


Fig. 2.4. The image processing pipeline.

of the visual field. The linked columns serve the purpose of an elementary computational module. They receive varied inputs, transform them, and send the results to a number of different regions in the HVS.

As we move further up the visual pathway, there is an increased pooling of information from lower levels and more complex forms of specialised processing carried out in different regions.

2.2 Hexagonal image processing in computer vision

Image processing in computer vision systems consists of three important components: acquisition, processing and visualisation. A simple image processing system with the information flowing from one end to another is depicted in Figure 2.4.

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. In terms of the discussion of the visual system in Section 2.1.2 the acquisition stage is equivalent to the eye(s). 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 visualisation stage is very useful and often essential for human observers to make sense of the processed information. Underpinning these three components is the use of a lattice or a grid on which the visual information is defined. The acquisition stage uses this to capture the image of the real world while the processing stage uses it to define appropriate data structures to represent and manipulate the image. The lattice of interest here is the hexagonal lattice, hence we will restrict the literature survey to hexagonal image processing.

The beginning of digital image processing is generally traced to the early 1960s, at which time it was spurred by the need to enhance images transmitted by the Ranger 7 [12]. These images were sampled on a square lattice. Interest in images defined on hexagonal lattices can also be traced to the 1960s. McCormick, reporting in 1963, considered a *rhombic* array, which is a hexagonal lattice, in addition to a rectangular array for a thinning algorithm to process digital images of bubble chamber negatives [13]. The work was part of the design of the ILLIAC III, a parallel computer developed exclusively for pattern recognition. Another theoretical work on 2-D signals from this period is that of Petersen [14] who found the hexagonal lattice to be the optimal arrangement for sampling of 2-D bandlimited signals. However, as we shall show next, the work on hexagonal image processing has not been sustained or intense compared to square image processing, the reasons for which are open to speculation.

The overview of the work that has been carried out by researchers in the last 40 years on hexagonal image processing is organised for convenience, along the lines of the information flow shown in Figure 2.4.

2.2.1 Acquisition of hexagonally sampled images

There are two main approaches to acquiring hexagonally sampled images. Since conventional acquisition devices acquire square sampled images, the first approach is to manipulate the square sampled image, via software, to produce a hexagonally sampled image. The second approach is to use dedicated hardware to acquire the image. We will discuss each of these in turn.

Software-based acquisition

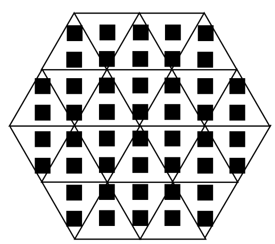
Manipulating data sampled on one lattice to produce data sampled on a different lattice is termed resampling. In the current context, the original data is sampled on a square lattice while the desired image is to be sampled on a hexagonal lattice.

The approach of Hartman [15] used a hexagonal lattice and triangular pixels. The construction process is illustrated in Figure 2.5(a), 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 $63.4°$ and a top angle of $53.2°$.

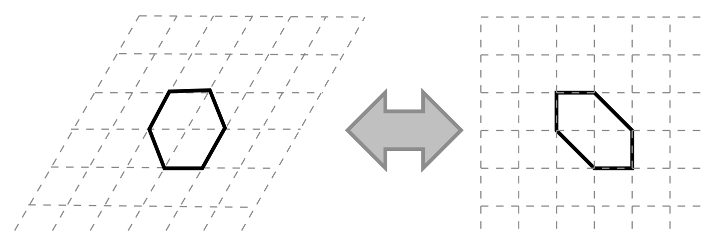
With the goal of deriving an image transform that was mathematically consistent with the primary visual cortex, Watson [16] proposed the hexagonal orthogonal-oriented pyramid. The square to hexagonal conversion process used the affine relationship between the square and hexagonal lattice points. The idea is illustrated in Figure 2.5(b). This means that rectangular images are skewed to form hexagonal images. After the skewing process, a hexagon becomes elongated in an oblique direction. 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 in the work, this distortion was deemed unimportant.

An approximation to the hexagonal lattice that is simple and easy to generate is a brick wall. Here, the pixels in alternate rows are shifted by half a pixel to simulate the hexagonal lattice. In Fitz and Green [17] this approach is taken. First they perform weighted averaging and subsampling on an image to halve the resolution in both directions and then they perform the pixel shift in alternate rows. The resulting image is like a brick wall made of square pixels. A different implementation of the brick wall is found in Overington's work [18]. It was noted that a hexagonal lattice is achievable with an array of square pixels where the horizontal separation is 8 pixels and the vertical separation is $5\sqrt{2}$, with alternate rows being staggered by 4 pixels. This can be approximated by a brick wall of rectangles having an 8×7 aspect ratio. Overington observes that there are no measurable errors in this approach, even though the shapes are incorrect. The process to generate hexagonal images in this case then involves first generating 8 rows of data from each 7 rows in the original image. Alternate rows are computed using the mean of adjacent pairs of data. These two steps are illustrated in Figure 2.6(a) and 2.6(b). The first step is a simple interpolation with a linear kernel and the second step is a nearest neighbour interpolation.

Another approach to generating hexagonally sampled images is via the use of quincunx sampling. These samples are arranged as in a chessboard as illustrated in Figure 2.7. Laine [19] 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 emphasises the hexagonal arrangement. The interpolated image was masked to produce the quincunx pattern, following which the remaining data was mapped onto a hexagonal grid.

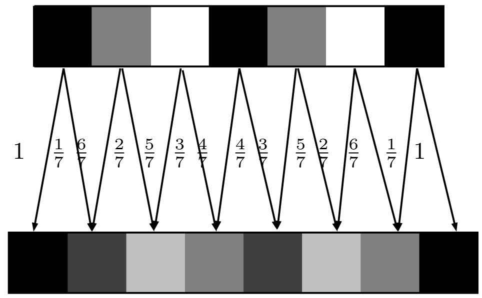


(a)

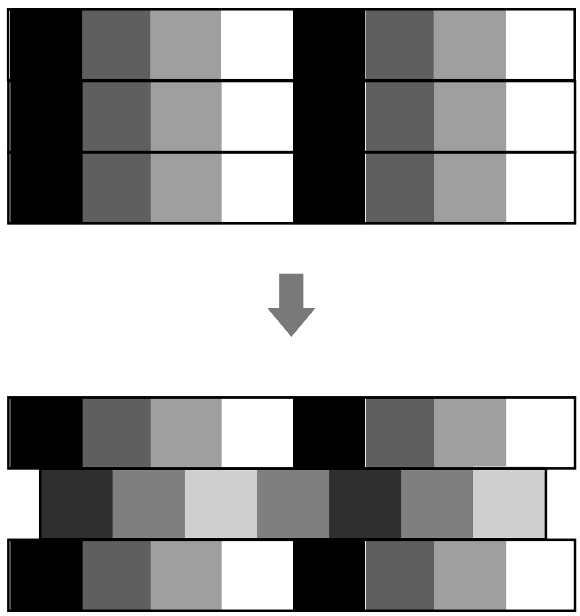


(b)

Fig. 2.5. Hexagonal resampling schemes of (a) Hartman and Tanimoto (1984) (b)Watson and Ahumada (1989).



(a)



(b)

Fig. 2.6. Hexagonal resampling scheme of Overington (1992) (a) making 8 rows from 7 (b) combining alternate rows.

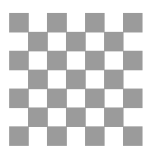
The resampling approach of Her [20] is also an approximate one. It is similar in idea to the work of Hartman. Here, interpolation is performed on the original image to halve the vertical resolution. After comparing a variety of different interpolation functions, Her found that bi-cubic interpolation performed the best, although the computational speed was low. Hence, bi-linear sampling was advocated as a viable alternative.

More recently, least squares approximation of splines has been proposed for resampling square images onto a hexagonal lattice [21–23]. This is an exact method and is computationally intensive. It has been used for suppressing aliasing artifacts in high quality colour printing.

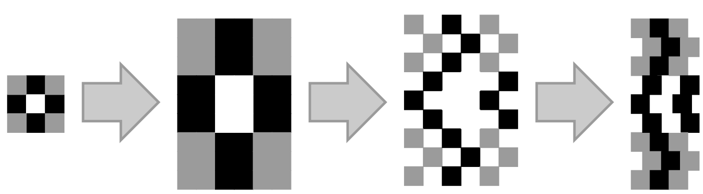
Hardware based acquisition

A hexagonal image can be acquired in a cost-effective way by modifying an existing hardware system to perform hexagonal sampling. Staunton [24] designed a pipeline architecture to take video images and, by introducing a delay to alternate lines, produced a hexagonal sampled image. There has also been much interest in building custom hardware for acquiring hexagonal images. There are two good reviews of the field of hardware sampling and sensors [25, 26]. Staunton [25] 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 Mead [27] who has built sensors mimicking various biological sensors including the retina. In 1982, Gibson and Lucas [28] referred to a specialised hexagonal scanner. The last ten years has witnessed increased activity in custom-built hexagonal sensors, many of which are CMOS based. These range from general purpose to application specific. Examples of general purpose sensors are found in [29–32]. The superior ability of hexagonal grids to represent curves has motivated a CMOS fingerprint sensing architecture to be developed based on the hexagonal lattice [33]. With the ability to grow crystals in space, several projects have been performed to grow hexagonal sensing arrays for applications such as satellite sensors [34] and replacing the human retina [35] after it has been damaged. Hexagonal sensors have also been developed for high speed colour and position sensing [36]. Interestingly, hexagonal sensors (be it solid state or photomultiplier based) also find a place in medical imaging [37] and remote sensing [38].



(a)



(b)

Fig. 2.7. Quincunx sampling (a) arrangement (b) hexagonal image construction.

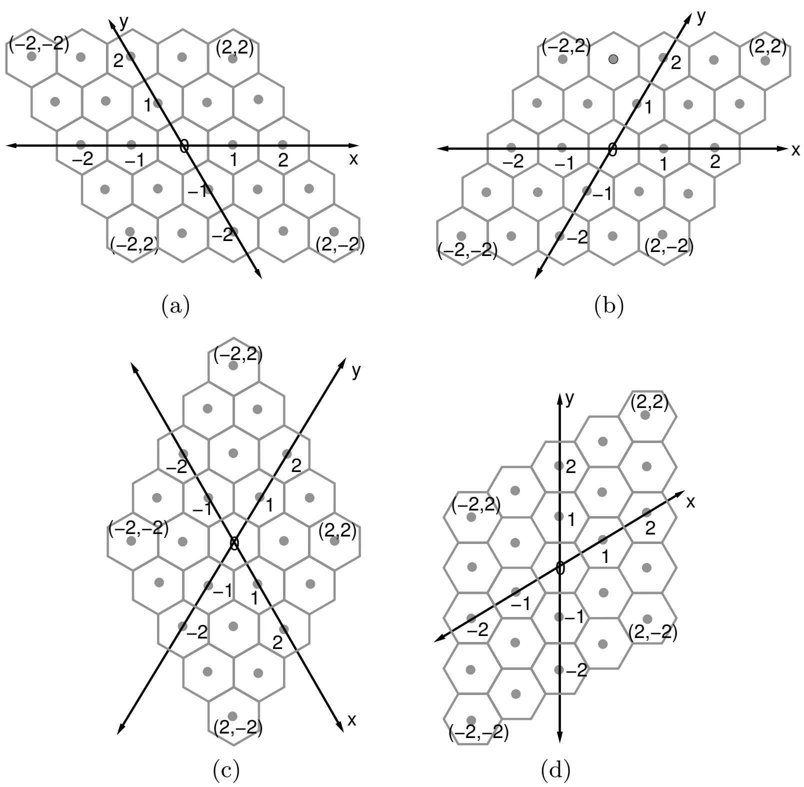


Fig. 2.8. Addressing points on a hexagonal lattice using two skewed axes.

Non-square grids, such as quincunx grids have also been found to be advantageous for high-speed imaging due to the reduction in aliasing and improved resolution after interpolation [39].

2.2.2 Addressing on hexagonal lattices

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. 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 in which 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. This will yield integer coordinates and is efficient as

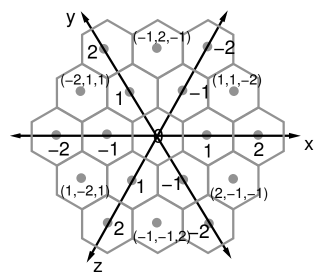


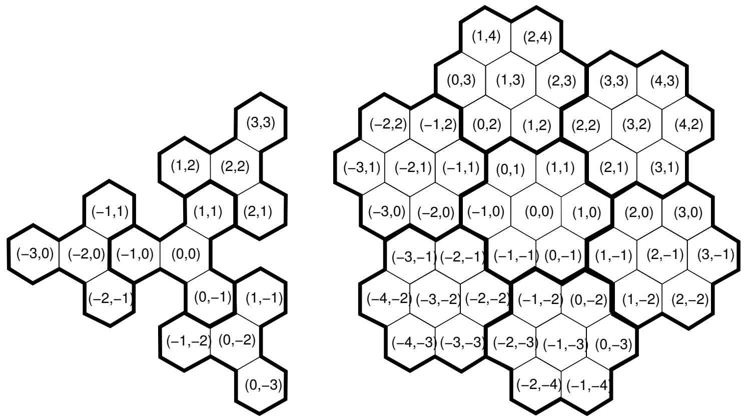
Fig. 2.9. Addressing points on a hexagonal lattice using three skewed axes.

two coordinates are sufficient to represent a point on a plane. There are two distinct possibilities for the skewed axes as illustrated in Figures 2.8(a) and 2.8(b), where the axes are either 120° apart or 60° apart. Further variations are possible by rotating the axes such that one them is vertical, if desired. Nevertheless, the coordinate system remains the same. Several examples of using these skewed axes can be found in the literature. Those using the axes in Figure 2.8(a) are found in [40–42] while examples of the axes in Figure 2.8(b) can be found in [16, 26, 43–45]. Combinations of these are also possible such as that of Rosenfeld [46, 47] and Serra [48]. An example of usage of the rotated axes in Figure 2.8(c) is seen in [49] while that of the axes in Figure 2.8(d) can be seen in [50].

Another approach to addressing a hexagonal lattice is that taken by Her [20, 51] using the three axes of symmetry of the hexagon instead of two axes. The third axis will be a linear combination of the other two axes and hence this coordinate system suffers from redundancy. However, this symmetric hexagonal coordinate frame has advantages when it comes to operations which involve a large degree of symmetry such as rotation and distance measures. The corresponding coordinate scheme uses a tuple of coordinates (l, m, n) which correspond to the distance from the lines $x = 0$, $y = 0$, and $z = 0$ respectively, and they obey the following rule:

$$l + m + n = 0$$

It is clear that the distance between any two neighbouring points in this scheme is 1. Additionally, as this scheme uses all three axes of symmetry, it is possible to reduce the coordinates to any of the skewed axis coordinate systems. Thus, any theories or equations derived for the two skewed axes schemes can then be applied to Her's tuple. An illustration of the coordinate system is given in Figure 2.9. The coordinate scheme can also be considered to be the projection of a 3-dimensional Cartesian coordinate scheme, $\mathbb{R}^3$, onto an oblique plane through the origin, with equation $x+y+z = 0$. Hence, many of the geometrical properties of $\mathbb{R}^3$ can be readily exploited in this coordinate scheme. However, according to Her [20], this coordinate scheme leads to more



(a) (b)

Fig. 2.10. The hierarchical addressing scheme proposed by Burt [52].

cumbersome data structures than using skewed axes and can lead to increased processing time, especially for non-symmetric operations.

The approach taken by Overington [18] is a little different from the approaches outlined previously. The entire hexagonal array is treated as if it is a rectangular array and Cartesian coordinates are directly employed to address all points. However, for practical use, it must be remembered that either odd or even rows should correspond to a half pixel shift. He notes that it is possible to approximately predetermine local operators with this shift taken into account.

There are other methods for addressing a hexagonal lattice and the motivation behind these lie in the exploitation of hierarchical hexagonal structures and their properties. The earliest of these was the work of Burt [52] who examined the construction of trees and pyramids for data sampled on hexagonal lattices. He distinguished four distinct constructions as being compact and symmetric. The coordinate scheme used a tuple, (i, j, k), and is called a pyramidal address. The last coordinate, k, is the level of the pyramid and the remaining two coordinates (i, j) were found using a skewed axis such as the one illustrated in Figure 2.8(b). The coordinate scheme is relative to each layer of the pyramid. Figure 2.10 illustrates two examples of this scheme for hexagonal lattices employing skewed axes (120°). The first two layers of the pyramid are shown highlighted with thick lines. The pyramid address of these pixels has a third coordinate of 0 or 1 indicating the first or second layer. The second of these examples which is hexagonal, Figure 2.10(b), was called a sept-tree. It was considered to be particularly promising as it was the largest compact (there are no nodes outside the pattern which are closer to the centroid than those inside the pattern) tile in either square or hexagonal lattices.

Gibson [28, 53–55] also looked at an addressing regime based on pyramidal decomposition of hexagonal lattices. The structure used was extensible to arbitrary dimensions. In two dimensions, the structure was identical to Burt's sept-tree in Figure 2.10(b), but the addressing scheme was different. The motivation for this work was the conversion of map data to other forms for graphical information systems and used a data structure known as the generalised balanced ternary (GBT) [56, 57]. The GBT consists of hierarchical aggregation of cells, with the cells at each level constructed from those at a previous level using an aggregation rule. For the 2-D case, the aggregation rule is the same as that of Burt's sept-tree. Each level consists of clusters of seven cells, and any GBT-based structure can be represented by unique, base-seven indices. The work of Sheridan [58] studies the effect of modulo operations on these addresses.

Independent of the work on GBT-based data structures, Hartman and Tanimoto [15] also investigated generating hexagonal pyramid structures, for the purpose of modelling the behaviour of the orientation- and location-specific cells in the primary visual cortex. The structure they devised was based on aggregations of triangular pixels which resulted in an image which was hexagonal in shape. Each pixel in the image was labelled uniquely by a three-tuple of coordinates: the first being the level of the pyramid and the other two being the position within the level. The position was measured on a coordinate system with the x-axis at $0°$ and the y-axis at $-60°$.

2.2.3 Processing of hexagonally sampled images

Contributions to the body of knowledge about processing aspects of hexagonally sampled images can be found at the level of theory as well as applications.

At the theoretical level, the earliest work is by Petersen [14], who considered it as a possible alternative sampling regime for a 2-D Euclidean space. Petersen concluded that the most efficient sampling schemes were not based on square lattices. The contribution of Mersereau [41, 59] for hexagonally processed signals is extensive. Based on earlier studies of general sampling [14], he proposed a hexagonal sampling theorem and used it to study properties of linear shift invariant systems for processing hexagonally sampled signals. The hexagonal discrete Fourier transform (HDFT) was also formulated by Mersereau along with a fast algorithm for its computation. The Fourier transform kernel is non-separable on a hexagonal lattice, hence an entirely different approach based on the work of Rivard [60] was employed to speed up the computations. Despite the non-separability problem, the fast algorithm was found to be 25% faster than the equivalent fast Fourier transforms for the square grid. Other contributions by Mersereau include recursive systems and of the design of FIR filters for hexagonal sampled signals.

Overington's [18] work is focused upon modelling the human visual system. Motivated by a keen interest in the limits of the visual system and its implications, hexagonal lattice was chosen as it models the retinal arrangement.

The problems that he examined are varied. He studied such areas as image preprocessing, optical flow analysis, stereo vision, colour processing, texture analysis, and multi-resolution image analysis. A good review of his work can be found in [61].

Work on morphological processing of hexagonally sampled images dates back to Golay [62] who proposed a parallel computer based on hexagonal modules which could be connected to perform different morphological operations. It was also shown that it required fewer interconnections compared to a similar square based architecture [63]. A good review of this work is found in Preston [64]. Serra has also contributed in a major way to the knowledge of mathematical morphology on a hexagonal lattice. He showed a distinct preference for hexagonal images by applying all the algorithms to hexagonal lattices prior to square lattices in his book [65,66]. The uniform connectivity and other topological properties of the hexagonal lattice are cited as reasons to make it an ideal candidate for morphological operations. Staunton [24, 25, 67–70] is another researcher who has worked extensively in both the hardware and software aspects of hexagonal image processing. As mentioned earlier, he has designed specialised hardware to generate hexagonal images and worked on a variety of applications, such as edge operators and thinning, that can help make a hexagonal system a replacement for automated inspection systems. For a good summary of the wide-ranging work see [25].

Distance transforms have been studied by several researchers starting from 1968. Rosenfeld [46] studied distance functions on digital pictures using a variety of coordinate systems. Simple morphological operators were also evaluated in these coordinate systems via some simple applications. Distance transforms on hexagonal grids were studied extensively later as well, with integer and non-integer arithmetic [71–73] and extended for 3-D images [74]. A parallel algorithm for deriving a convex covering for non-convex shapes, defined on a hexagonal grid, proposes to use the distance transform for shape analysis [75].

Thinning algorithms for hexagonal and square grids were first considered by McCormick [13] in the context of parallel implementations of thinning. The hexagonal grid was found to offer a specific edge for parallelisation over a square grid since there are only six neighbours compared to eight. Deutsch [76] also implemented a thinning algorithm on hexagonal, square and triangular lattices. The results on the hexagonal lattice were found to be better in terms of the number of edge points and resistance to noise. These results were confirmed later by Staunton [67,68] and shown to be quite efficient under parallel implementation.

Frequency domain processing of hexagonally sampled images has motivated several researchers to investigate various transforms. The work on HDFT by Mersereau was revisited in 1989 by Nel [45] who also introduced a fast Walsh transform. The fast Walsh transform was derived using a similar formulation to that for the fast Fourier transform. An analytical derivation of the DCT for hexagonal grids is available in [77]. An unusual work in the area of fast algorithms for DFT takes hexagonal images as input but computes the frequency domain samples on a square grid [78]. The formulation allows the 2-D DFT computation by two 1-D DFTs which is, as noted before, not possible on a hexagonal grid. Fitz and Green [17] have used this algorithm for fingerprint classification and report that the algorithm is not as efficient as the square FFT, but is more efficient in memory usage since fewer data points are required for the same resolution. Another work which has a square lattice formulation to compute the hexagonal DFT is given in [79]. Hexagonal aggregates identical to those defined by Gibson [28] have been used in [80] to develop a fast algorithm for computing the DFT. The radix-7 algorithm is very efficient with computational complexity of $Nlog_7N$ for an image containing a total of N pixels.

Multiresolution image processing on hexagonal images is an area which is seeing a slow pace of activity although the theoretical foundations have been laid for over a decade via filter bank design [81, 82] and wavelet bases [83, 84]. It has been pointed out that the advantage of using a hexagonal grid for multiresolution decomposition is the possibility of having *oriented* subbands which is not possible in a square lattice. Simoncelli's work [82] on perfect reconstruction filter banks has been used as a basis in [40] for sub-band coding and in [19, 85] for tumour detection in mammograms. Both these works use a pyramidal decomposition of the image with three sub-band orientations namely at $30°$, $90°$ and $120°$.

The hexagonal, orthogonal, oriented pyramid of Watson [16] was used for an image coding scheme. The main motivation was to be comparable to the receptive fields of the human visual system. It was built upon an earlier work called the cortex transform. The levels of compression achieved in this scheme were greater than the equivalent schemes for square images.

The orientation selectivity afforded by hexagonal lattice and multiscale representations have been used a bit differently to define a set of ranklets for feature detection [86]. Such ranklets on square lattices resemble the Haar wavelets. Results of applying them to face detection are said to be good and consistent with those using square ranklets.

The work of Gibson and Lucas [28, 53–56] was primarily in the addressing of hexagonal lattices. However, several applications were performed upon their addressing scheme. The first was the rasterisation of digitised maps for a geographical information system. Another application was automated target recognition where the aim was to recognise tanks from a cluttered background for the United States military. Snyder [26] revisited addressing of hexagonal images and defined his own addressing scheme. Within the context of this addressing scheme, neighbourhoods were defined and associated operators such as convolution and gradient were developed.

The contribution of Her [20, 51], like Gibson and Lucas, is primarily in the addressing scheme. His 3-coordinate scheme exploits the implicit symmetry of the hexagonal lattice. Her also introduced a series of geometric transformations such as scaling, rotation and shearing on the hexagonal 3-coordinate system. Operators were also devised for efficient rounding operations to find nearest integer grid points in a hexagonal lattice, which are useful in geometric transformations.

Kimuro [87] used a spherical hexagonal pyramid to analyse the output from an omni-directional camera. The results were promising and were better than such systems designed using square images. Surface area estimates of scanned images, have been computed using hexagonal and other tilings by Miller [44]. Hexagonal tiles were found to perform with less error than the other tiles. Sheridan [58] revisited the work of Gibson and Lucas and labelled his version of the generalised balanced ternary as the Spiral Honeycomb Mosaic and studied modulo arithmetic upon this structure. Texture characterisation using co-occurrence matrices have also been studied on the hexagonal lattice [88]. It was found that there were benefits for textures which have information at multiples of 60°. Additionally, texture synthesis using Markov random fields has been proposed which permit classes of texture which are not simple to generate using square images [89]. A family of compactly supported hexagonal splines are proposed in [23]. It has an increased number of symmetry compared to the B-splines used on square lattices. Among the suggested applications are zooming using a hexagonal spline transform and resampling.

Quantitative studies comparing image quality on square and hexagonal lattices have also been carried out in [90–93]. Quantisation error measures and profiles have been developed and used to compare the square and hexagonal sampling in [92, 93]. The results reported for line representation show the hexagonal grid somewhat favourably. However, it is cautioned that the specific algorithm to be used should be taken into consideration for an accurate evaluation. [91] presents an interesting comparison of the square and hexagonal lattices with respect to image quality measurements. They propose a metric in terms of the area enclosed within the Fourier transform of the Wigner-Seitz cell. Effective resolution is a term that has been often overused in digital camera literature. Almansa [90] presents a quantitative means to measure the effective resolution of image acquisition systems which can be used as a basis of comparison of square and hexagonal sampling as well as for improving image resolution. He also gives a good overview of hexagonal image processing in his doctoral dissertation [94].

2.2.4 Visualisation of hexagonally sampled images

The final part of the review we will present is on the visualisation aspect of hexagonal image processing. Visualisation is strongly tied to advances in display technology. There have been many approaches over the years to this problem and some of these will be now presented. The presentation will be roughly chronological.

The earliest approach to displaying hexagonal images dates back to the 1960s and Rosenfeld [46]. His intuitive approach, illustrated in Figure 2.11, was to use an evenly spaced grid with alternate rows offset by one point. The images were purely monochrome, so an *on* pixel was illustrated with a '.'

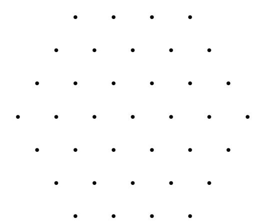
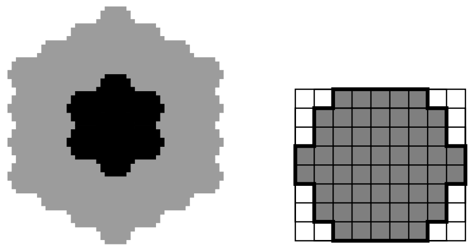


Fig. 2.11. Rosenfeld's method for displaying a hexagonally sampled image [46].

and an *off* pixel was illustrated with a ' '. Intensities and other information, however, could be conveyed by employing different characters. The results were generated using a dot matrix printer. As the horizontal spacing is fixed, the sole manipulation that could be performed was the vertical spacing. This approach generated hexagonal lattices that appeared very regular.

In the 1980s there were several significant advances in technology which directly impacted on the visualisation of hexagonal images. The first of these was in the printing industry and the second was in the visual display industry. In the print industry, devices were developed which afforded a much higher resolution and more control of the printed output than previously available in dot matrix devices. These had far reaching effects on print media, specifically newspapers, and led to research into how to display continuous tone photographs on the printed page. This area was known as half-toning. Stevenson [42] proposed the use of a hexagonal layout of dots. The dots, being squirts of ink, are circular and hence a hexagonal lattice was more robust to errors and compensated the characteristics of the non-ideal printing device. His method was analogous to that of Rosenfeld except that individual dots were used to indicate the hexagonal pixels instead of entire characters. These dots had a controllable radius and were slightly bigger than an individual print element. In this way, alternate rows could be offset to produce the required hexagonal lattice arrangement. This allowed hexagonal images to be printed with much denser arrangements of hexagonal pixels. A variation of this method of display is used in newspapers even today. Recently, resampling square to hexagonal images using splines has also been used to reduce alias artifacts in colour printing [22].

The visual display industry also experienced significant advances in the 1980s with the arrival of high resolution visual display units which were capable of displaying millions of colours. This had an effect on the display of hexagonal images. For instance, the work of Hartman and Tanimoto [15] made up a hexagonal image using accumulations of screen pixels which approximated an equilateral triangle. These were coloured with the desired intensity of the hexagonal pixel. Wüthrich [50] used an accumulation of screen pixels which, when taken together, looked hexagonal in shape. This is illustrated in Figures 2.12(a) and 2.12(b). The logical consequence of this approach is that the screen resolution is significantly reduced when compared with square



(a) (b)

Fig. 2.12. Wüthrich's method for displaying a hexagonally sampled image [50] (a) an image (b) an individual picture element.

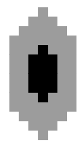


Fig. 2.13. Her's method for displaying a hexagonally sampled image [20].

images. The design of these hexagonally shaped aggregations of pixels is also problematical.

Her [20] used two pixels to approximate the hexagonal pixels in a display. Using the assumption that the individual screen pixels are oblate (slightly elongated in the horizontal direction), an approximation to a hexagonal pixel could be made with just two pixels. This gave the individual hexagonal pixels an aspect ratio of 1:1.64. For an equilateral triangle drawn through the centres of three adjacent hexagons, this gave base angles of $51°$ and a top angle of $78°$. A one-pixel shift in alternate rows is all that is required here to generate the image. A hexagonal image using these pseudo-hexagonal pixels is illustrated in Figure 2.13. This method has the advantage that it is easy to compare results with square sampled images as the only modification required is to remove the offset on alternate rows.

The work of Staunton [49] uses a consistent display scheme throughout opting for the brick wall approximation. Unlike Her however, larger pixels which have a width and height tuned to give the same aspect ratio as a real hexagonal pixel are employed. For ideal pixels which have a vertical separation of 1 the corresponding horizontal separation is $\frac{2}{\sqrt{3}}$. Thus, it is possible to plot a rectangular accumulation of pixels that has an aspect ratio of 1.15:1. Overington [18] uses different approaches to display hexagonal images depending on the application. For coarse images, a methodology similar to Rosenfeld is utilised. However, in the case of high resolution images, a methodology similar to Staunton is preferred. Overington remarks that the brick wall display process has little error and can be used rather than going to the trouble of using a real hexagonal grid.

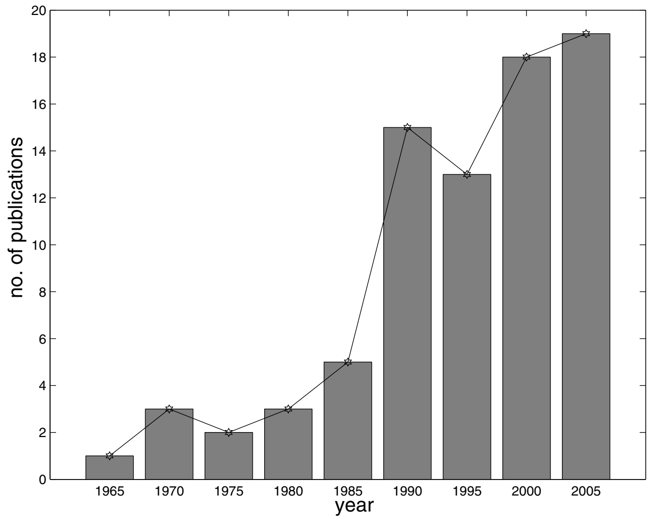


Fig. 2.14. Publications about hexagonal image processing over the last 40 years.

The work of Gray et al. [95] was in the perception of differently shaped pixels. They experimented with a variety of different pixel shapes and arrangements for displaying images. Experiments included a conventional square image and a hexagonal image, both displayed using a brick wall approach. For pixel shapes, real hexagons, using the method of Wüthrich, and many more complex shapes were tested. The results demonstrated that the human visual system has a preference for a pixel shape that has dominant diagonals in it. This includes diamonds and hexagons.

2.3 Concluding Remarks

The aim of this chapter has been to provide an overview of the visual systems. The two systems that were discussed were the human visual system (HVS) and computer vision systems.

The brief discussion of the HVS revealed it to be a complex one. In the language of modern computer vision, the HVS appears to employ *smart* sensing in two ways: (i) it is adaptable to ambient illumination due to the use of two types of receptors and long range connections and (ii) it is capable of a wide field of high resolution imaging, achieved by optimising the sensing methodology via a combination of movable sensory system and non-uniform sampling. Other remarkable aspects of the HVS are the hexagonal arrangement of photoreceptors (especially in the fovea) and the hierarchical nature of representation of visual information and its processing.

The review of computer vision was restricted to hexagonal image processing. A chronological picture of the research activity in this field is shown in Figure 2.14. From the review and this figure we can see that the field of hexagonal image processing is as old as image processing itself. Researchers in this field have had disparate motivations for considering hexagonal sampling, ranging from modelling HVS to parallel implementation of thinning and fast algorithm for Fourier Transform to efficient map data handling for geographical information systems. Almost all traditional areas of image processing have been investigated by researchers in this field.

The length of time for which hexagonal sampling has held interest among researchers and the wide range of areas covered, we believe, points to the fact that there is not only significant interest but also merit in processing hexagonally sampled images. However, the second point to note from the figure is that the level of activity is low in general, notwithstanding the upward trend that is apparent in the last 15 years. This is true both in terms of the number of researchers and in the number of publications that has come out in this area over 40 odd years. The reason for this can only be described as the weight of *tradition.* Hardware for image acquisition and visualisation have traditionally been based on a square lattice. The reason for the former is ascribed to the complications in designing a hexagonal arrangement of sensors, notwithstanding the availability of technology for hexagonal layouts in large RAMs.

The scenario in computer vision as a whole is witnessing a change. There appears to be renewed interest (both at research and commercial levels) in fovea-type image sensors and non-uniform sampling, features which we have noted to be hallmarks of HVS. These developments are motivated by the search for new solutions to many difficult problems in vision as well as the need to innovate. In keeping with this trend, we take a fresh look at hexagonal image processing as a whole in the remaining parts of this monograph and propose a complete framework. This idea was initially introduced in [96–101] and is being expounded further in this monograph. The framework is intended to serve as a test bed for studying the impact of changing the sampling lattice on image processing.

http://www.springer.com/978-1-85233-914-2

Hexagonal Image Processing
A Practical Approach
Middleton, L.; Sivaswamy, J.
2005, XIII, 254 p. With CD-ROM., Hardcover
ISBN: 978-1-85233-914-2