### 4.5.1 Noise in the Histograms

The histograms obtained from real image data always have a certain amount of noise, noise which cannot always be ignored. There are, in fact, several stages where noise can be injected into the color image data. The response of the camera, may not be the same for all the three channels (r, g, b). Further, when a combined 3-dimensional histogram is formed, the noise from the individual channels are added. Also, when the histograms are formed, the last two bits of each of the three channels are ignored as explained in Section 3.1.1. This could be another source of noise.

One way to study the noise content of the data is to trace the variations of the histogram data along 1-dimensional curve, i.e., along a line in 3-dimensional space passing through the peak of the histogram. This concept is shown on a 2-dimensional graph in Figure 4.10. Consider a 2-dimensional histogram which has a peak at cell *M*. A line *l* is drawn from the origin through the cell *M* in the 2-dimensional space. The shaded cells are those cells through which the line passes through, and the values at these locations are plotted along the x-axis. Figure 4.11a shows the variations in the 3-dimensional histogram seen along a line in 3-dimensional space that passes through the highest point in the histogram. The line in 3-dimensional space is shown in Figure 4.11(b).

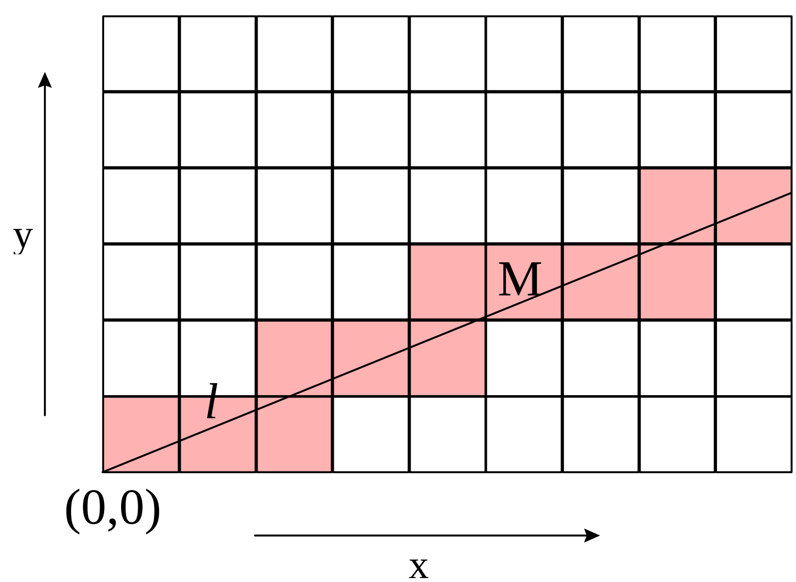
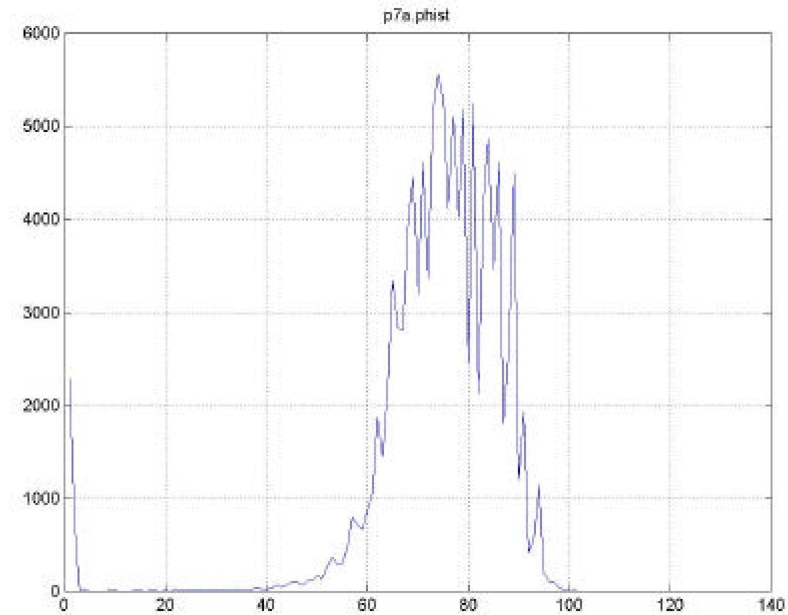



Figure 4.10: Concept of a line through a 2-dimensional histogram

The noise in the image data causes of false peaks to occur in the histogram. Figure 4.11(a) shows that there are a lot of spikes and abrupt changes in the 3-dimensional histogram. If the threshold $T_i$ is selected to be very low, the entire histogram falls under one cluster. On the other hand if it is set to be very high, the main cluster containing the clear wood will be broken up into a number of clusters. This problem is compounded by the fact that when the elements in set *S2* (Section 2.7.1) are assigned to the nearest cluster, those elements that actually belong to the clearwood cluster could end up in the defect cluster.

A 2-dimensional histogram of the pine board shown in Figure 4.2(a) are shown in Figures 4.12, each illustrated using two representations. To illustrate the effect of irregularities in the histogram, consider the r-b histogram shown in Figure 4.12(a). There appears to be two main peaks in this histogram. But, in between the two peaks there are several spikes where each spike appears as a smaller peak. Now all these histograms are in two dimensions. The effect will be more pronounced in three dimensions. The reason for this is as follows. The value at each element $(i,j)$ in the 2-dimensional histogram in is the sum of all the elements of the third channel which has co-ordinates $(i,j)$ in the 3-dimensional histogram. The formation of 2-dimensional histograms from 3-dimensional histograms is explained in Section 3.1.2. This has an averaging effect, which reduces the appearance of irregularities. The contour plots also indicate the presence of several peaks. The white regions in the center of the contour plots represent the peaks in the histogram. All these features contribute to the observed segmentation results of Section 4.2



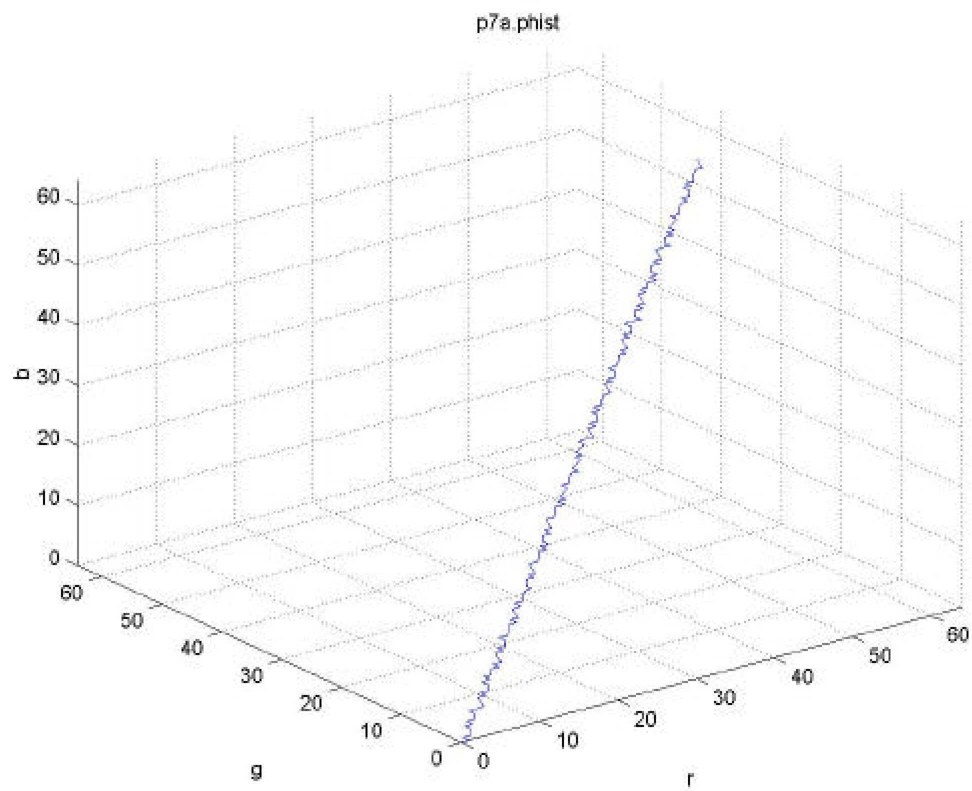




Figure 4.11: (a) variation in the 3-dimensional histogram along a line in 3-dimensional space, (b) line in 3-dimensional Space