IMAGE PROCESSING

by

ANDY BING-BILL CHAN

B.Sc., The University of British Columbia, 1993

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF

THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

THE FACULTY OF GRADUATE STUDIES

Department of Statistics

We accept this thesis as conforming

to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA

April 1995

Department of ___STATISTICS___

The University of British Columbia
Vancouver, Canada

Date ___21 APRIL 1995___

DE-6 (2/88)

ABSTRACT

In this thesis, we consider the restoration of multiple grey levels image. The problem is to clean up or restore the dirty picture, that is, to construct an estimate of the true image from a noisy picture of that true image. Following a method proposed by Meloche and Zamar (1994), we estimate the colour at each site by a function of the data available in a neighbourhood of that site. In this approach, the local characteristics of that image, that is, the frequency with which each pattern appears in the true unobserved image are particularly important. We will propose a family of unbiased estimates of the pattern distribution and the noise level which are used in the restoration process. We will use our estimates of the pattern distribution in an attempt to select the best neighbourhood shape for the restoration process.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

1 Introduction.

In this thesis, we consider multiple grey levels image restoration. One kind of grey levels is grey levels from white to black. Some other kinds are grey levels are from white to red, white to green, white to blue, etc. If we combine the three latter kinds, we will get a coloured image. The problem is to *clean up* or *restore* the dirty picture, that is to construct an estimate of the true image from a noisy picture of that true image. The true image is not observed, while the noisy picture is. Figure 2.1 is the true image while Figure 2.2 is a degraded version of true image, the noisy picture. We want to recover the true image (Figure 2.1) from the noisy picture (Figure 2.2) which is the observed data.

The estimate of the colour at each site proposed by Meloche and Zamar (1994) is a function of the data available in a neighbourhood of that site. The local characteristics of the underlying image is central to their approach. The notion of local characteristics is formalized in what follows.

We assume that there are n sites or pixels on a plane (or a line), numbered 1 to n. The vector $\{z_1, z_2, \ldots, z_n\}$ represents the true image, and $z_i \in C$ where $C = \{c_1, c_2, \ldots, c_k\}$ is a set of k grey levels. The observations $Y_1, Y_2, \ldots, Y_n$ (called the records) form the degraded image and are modelled as

$$Y_i = z_i + \sigma e_i \tag{1}$$

where $e_1, \ldots, e_n$ i.i.d. $N(0,1)$, $Y_1, \ldots, Y_n$ are the records, and σe_i's are the noise. The objective of image restoration is to work out estimates $\hat{z}_1, \ldots, \hat{z}_n$ of the true image $z_1, \ldots, z_n$ from the records $Y_1, \ldots, Y_n$.

Following Meloche and Zamar (1994), we define a system of neighbourhoods, $N_1, \ldots, N_n$.

The neighbourhood N_i, centred at i, consists of r sites called the *neighbours* of site i and includes the site i itself. In general, the neighbourhood N_i is the set of all sites j at a small distance from site i. The neighbourhoods $N_1, \ldots, N_n$ have the same shape but different centres. For examples, if $N_i = \{1, \ldots, n\}$ for all i, then r equals n and we have just one neighbourhood, the whole image. On the other hand, if $N_i = \{i\}$ for all i, then r equals 1 and then we have n single-site neighbourhoods. Neighbourhoods can be 1-dimensional (linear), 2-dimensional (planar), 3-dimensional, or even higher than 3-dimensional. A *pattern* is a particular colouring of the sites (or pixels) in a neighbourhood. Some examples of neighbourhoods are graphically shown in the following examples.

Example 1.1 : If the true image consists of $0 - 1$ bits sent over a line, then the following is a typical linear 3-site neighbourhood:
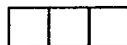
The true image is one-dimensional, and the colour set C is $\{0, 1\}$, where $0 = $ black, and $1 = $ white. In this example of 3-site neighbourhood, there will be $2^3 = 8$ possible patterns:

Mathematically, a pattern here is a vector of 3 coordinates of black or white that represent the colour of neighbourhood sites in some fixed order.●

As we will notice from the next example, the size of the pattern space increases rapidly with the size of the neighbourhood.

Example 1.2 : Suppose we have a colour set $C = \{0, 1, 2\}$ ($0 = $ black, $1 = $ middle grey,

2 = white), and a cross-shaped neighbourhood
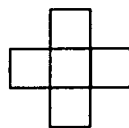
There will then be $3^5 = 243$ possible patterns:
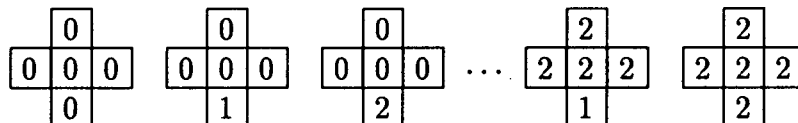
Again, mathematically, a pattern here is a vector of 5 coordinates of black, middle grey or white that represents the colour of neighbourhood sites in some fixed order.●

There is a need for a systematic convention for the ordering of the components of a neighbourhood. Firstly, we denote a pattern by δ. We use the bold face $\boldsymbol{\delta}$ as a vector of δ's in a neighbourhood. In Example 1.1, $\boldsymbol{\delta}$ is a vector of 3 coordinates, that is, $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)$, and in Example 1.2, $\boldsymbol{\delta}$ is a vector of 5 coordinates, that is, $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3, \delta_4, \delta_5)$. In the case of 3-site neighbourhoods, we use $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3)$ in the following order:
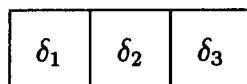
And then in the case of cross-shaped neighbourhoods, we use $\boldsymbol{\delta} = (\delta_1, \delta_2, \delta_3, \delta_4, \delta_5)$ in the following order:
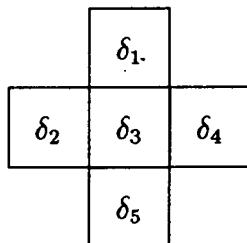
Bold face notation is used to represent a vector of variables in a neighbourhood. The coordinates are ordered in the same way as the coordinates of the patterns are. For

3

example, for 3-site neighbourhoods, $\mathbf{Y}_i = (Y_{i,1}, Y_{i,2}, Y_{i,3})$ is the vector of records, Y_j, for j in the neighbourhood N_i of site i. $Y_{i,j}$ is the record at the jth neighbour of site i. By equation (1),

$$\boxed{Y_{i,1}}\ \boxed{Y_{i,2}}\ \boxed{Y_{i,3}} \quad = \quad \boxed{z_{i,1}}\ \boxed{z_{i,2}}\ \boxed{z_{i,3}} \quad + \quad \boxed{\sigma e_{i,1}}\ \boxed{\sigma e_{i,2}}\ \boxed{\sigma e_{i,3}}$$

so that the distribution of $\mathbf{Y}_i$ is multivariate normal with mean $\mathbf{z}_i$ and covariance matrix $\sigma^2 I_3$.

Throughout this thesis, we look at the distribution of all the possible patterns in the true image. The true distribution of the patterns is usually unknown, but it can be estimated from the records (the noisy picture) in many different ways. As stated before, the pattern distribution plays a useful role in the process of image restoration. The estimates $\hat{z}_1, \ldots, \hat{z}_n$ of the true image proposed by Meloche and Zamar (1994) are based on the knowledge of the pattern distribution and σ.

Example 1.3 : In this example, we use the Figure 2.1 with 3 grey levels as the true image. Let r denote the neighbourhood size, and $C = \{0, 1, 2\}$. When $r = 1$, there are only three possible patterns, namely, black, middle grey, and white. Table 1.1 below shows how frequently each pattern appears in the true image.

Table 1.1: Frequency distribution ($r = 1$)

pattern	frequency
black	0.48340
grey	0.42505
white	0.09155
Total	1.00000

This frequency distribution reflects the fact that 48% of all pixels are black, 43% of them are middle grey, and 9% are white. When $r = 2$, a horizontal 2-site neighbourhood, there is a total of $3^2 = 9$ possible patterns, namely, 00, 01, 02, 10, 11, 12, 20, 21, and 22. Table 1.2 below shows the frequency with which each pattern appears in the true image.

Table 1.2: Frequency distribution ($r = 2$)

pattern	frequency
0 0	0.47192
0 1	0.00842
0 2	0.00305
1 0	0.00842
1 1	0.41541
1 2	0.00122
2 0	0.00305
2 1	0.00122
2 2	0.08728
Total	1.00000

The high frequency of the patterns 00, 11, and 22 in this frequency distribution reflects the fact that neighbouring pixels typically have the same colour.

When $r = 5$, we consider the cross-shaped neighbourhood, there are $3^5 = 243$ possible patterns. We use $\delta = (\delta_1, \delta_2, \delta_3, \delta_4, \delta_5)$ in the following order:

	δ_1	
δ_2	δ_3	δ_4
	δ_5	
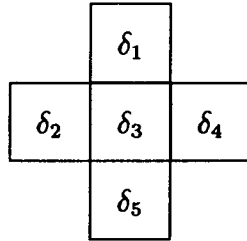
Table 1.3 below is the corresponding frequency table.

Table 1.3: Frequency distribution ($r = 3$)

pattern	frequency	pattern	frequency	pattern	frequency
0 0 0 0 0	0.43811	1 0 0 0 0	0.00824	2 0 0 0 0	0.00293
0 0 0 0 1	0.00824	1 0 0 1 0	0.00006	2 0 0 1 0	0.00012
0 0 0 0 2	0.00293	1 0 0 2 0	0.00012	2 0 2 2 0	0.00018
0 0 0 1 0	0.00824	1 0 1 1 0	0.00024	2 0 2 2 2	0.00281
0 0 0 2 0	0.00293	1 0 1 1 1	0.00806	2 1 1 1 1	0.00110
0 0 1 1 1	0.00012	1 1 1 0 0	0.00012	2 1 2 2 2	0.00110
0 0 2 2 2	0.00006	1 1 1 0 1	0.00793	2 2 1 1 1	0.00012
0 1 0 0 0	0.00824	1 1 1 0 2	0.00012	2 2 2 0 0	0.00006
0 1 0 0 1	0.00006	1 1 1 1 0	0.00793	2 2 2 0 1	0.00012
0 1 0 0 2	0.00012	1 1 1 1 1	0.38757	2 2 2 0 2	0.00269
0 1 1 0 1	0.00024	1 1 1 1 2	0.00110	2 2 2 1 0	0.00012
0 1 1 1 1	0.00806	1 1 1 2 0	0.00012	2 2 2 1 2	0.00110
0 2 0 0 0	0.00293	1 1 1 2 1	0.00110	2 2 2 2 0	0.00269
0 2 0 0 1	0.00012	1 1 2 2 2	0.00012	2 2 2 2 1	0.00110
0 2 0 0 2	0.00018	1 2 1 1 1	0.00110	2 2 2 2 2	0.07532
0 2 2 2 2	0.00281	1 2 2 2 2	0.00110	Total	1.00000

Note that Table 1.3 has fewer than 243 lines. It is because some patterns do not appear in the image. The high frequency of the patterns 00000, 11111, and 22222 reflects the fact that cross-shaped neighbourhoods of 5 pixels typically have the same colour.

In general, the pattern distribution is unknown, and Meloche and Zamar (1994) propose to estimate it directly from the records $Y_1, \ldots, Y_n$, for binary (black/white) images. In this thesis, we extend this to the case of images with an arbitrary, but finite, number of grey levels, and we develop a family of estimates of the pattern distribution.

In the case of single-site neighbourhoods, the above frequency distribution can be written as

$$Q_1(\delta) = \frac{1}{n} \sum_{i=1}^{n} 1\{z_i = \delta\}, \tag{2}$$

where $1\{z_i = \delta\} = 1$ when the true colour at site i is equal to δ, otherwise, it is equal to zero. The subscript 1 is used to indicate that we are dealing with single-site neighbourhoods. More generally, for an arbitrary neighbourhood size r, the above frequency distribution can be written as

$$Q_r(\delta) = \frac{1}{n} \sum_{i=1}^{n} 1\{z_{i,1} = \delta_1, \ldots, z_{i,r} = \delta_r\}, \tag{3}$$

and the δ in equation (3) is r-dimensional. The subscript r in the above equation is used to indicate that we are dealing with r-site neighbourhoods.

For binary images, Meloche and Zamar (1994) propose to use

$$\hat{z}_i = \frac{\sum_{\delta} \delta_c \phi_\sigma(\mathbf{Y}_i - \delta) Q_r(\delta)}{\sum_{\delta} \phi_\sigma(\mathbf{Y}_i - \delta) Q_r(\delta)} \tag{4}$$

where Q_r defined by equation (3) depends on the $z_1, z_2, \ldots, z_n$. Note that $\hat{z}_i$ depends

on the typically unknown parameters Q_r and σ. For this reason, Meloche and Zamar (1994) propose methods to estimate Q_r and σ in the case of $C = \{0,1\}$, that is, binary images. While equation (4) defines an estimate no matter what the colour set C is, the estimates of Q_r and σ proposed by Meloche and Zamar (1994) are only valid in the case of $C = \{0,1\}$. In this thesis, we propose a family of simple estimates for Q_r and σ in the case of $C = \{c_1, c_2, \ldots, c_k\}$, that is, images with an arbitrary, but finite, number of grey levels. Chapter 2 will focus on the estimation of Q_r, assuming σ is known, while Chapter 3 will deal with the joint estimation of Q_r and σ.

2 Estimations of Q.

2.1 Introduction

Firstly, let $\mathbf{Q}_r$, which will be used frequently, be the $|\mathbf{C}|^r$ vector of all possible $Q_r(\delta)$'s, where $|\mathbf{C}|^r$ is the total number of patterns. Meloche and Zamar (1994) propose a biased estimate of $\mathbf{Q}_r$ based on indicator functions for binary images, and a general formula for obtaining an unbiased estimate from the biased one. In this chapter, we extend their idea to images with an arbitrary, but finite, number of grey levels. And we also discover a simple representation which makes the computations of the estimates easier. In this chapter we assume that σ is known. The joint estimation of $\mathbf{Q}_r$ and σ will be discussed in Chapter 3.

2.2 Unbiased Estimations of $\mathbf{Q}_r$

We start by reviewing the biased and unbiased estimates of $\mathbf{Q}_r$ based on indicator functions for binary images proposed by Meloche and Zamar (1994). Later on, we will extend to image with more than two grey levels.

Consider a case of 2 grey levels (eg. black/white), that is, $\mathbf{C} = \{0, 1\}$, with single-site neighbourhoods $(r = 1)$. Let

$$\varphi(0, y) = 1\{y < \frac{1}{2}\}$$
$$\varphi(1, y) = 1\{y \geq \frac{1}{2}\}$$

Now, define

$$\tilde{Q}_1(0) = \frac{1}{n}\sum_{i=1}^{n}\varphi(0, Y_i) = \frac{1}{n}\sum_{i=1}^{n}1\{Y_i < \frac{1}{2}\}$$
$$\tilde{Q}_1(1) = \frac{1}{n}\sum_{i=1}^{n}\varphi(1, Y_i) = \frac{1}{n}\sum_{i=1}^{n}1\{Y_i \geq \frac{1}{2}\}$$

and $\tilde{\mathbf{Q}}_1$ to be the vector $(\tilde{Q}_1(0), \tilde{Q}_1(1))^T$. This $\tilde{\mathbf{Q}}_1$ is a biased estimate of $\mathbf{Q}_1$ because, for example,

$$
\begin{aligned}
E\tilde{Q}_1(0) &= E\frac{1}{n}\Sigma 1\{Y_i < \frac{1}{2}\} \\
&= \frac{1}{n}\Sigma P(Y_i < \frac{1}{2}) \\
&= \frac{1}{n}\Sigma P(z_i + \sigma e_i < \frac{1}{2}) \\
&= \frac{1}{n}\Sigma\{P(e_i < \frac{1}{2\sigma})1\{z_i = 0\} + P(e_i < \frac{-1}{2\sigma})1\{z_i = 1\}\} \\
&= \frac{1}{n}\Sigma 1\{z_i = 0\}\Phi(\frac{1}{2\sigma}) + \frac{1}{n}\Sigma 1\{z_i = 1\}\Phi(\frac{-1}{2\sigma}) \\
&= \Phi(\frac{1}{2\sigma})Q_1(0) + \Phi(\frac{-1}{2\sigma})Q_1(1)
\end{aligned}
$$

Similarly, $E\tilde{Q}_1(1) = \Phi(\frac{-1}{2\sigma})Q_1(0) + \Phi(\frac{1}{2\sigma})Q_1(1)$, where $\Phi(\cdot)$ is the cumulative distribution function of a N(0,1) random variable. Now we let $a = \Phi(\frac{1}{2\sigma}), b = \Phi(\frac{-1}{2\sigma})$, then we can write the above in matrix notation.

$$
E\tilde{\mathbf{Q}}_1 = \mathbf{A}_1\mathbf{Q}_1
$$

where

$$
\mathbf{A}_1 = \begin{pmatrix} a & b \\ b & a \end{pmatrix}.
$$

So, an unbiased estimate can be obtained by inverting $\mathbf{A}_1$ and defining $\hat{\mathbf{Q}}_1 = \mathbf{A}_1^{-1}\tilde{\mathbf{Q}}_1$, so that $E\hat{\mathbf{Q}}_1 = \mathbf{A}_1^{-1}E\tilde{\mathbf{Q}}_1 = \mathbf{A}_1^{-1}\mathbf{A}_1\tilde{\mathbf{Q}}_1 = \mathbf{Q}_1$. That is,

$$
\begin{aligned}
\hat{Q}_1(0) &= \frac{1}{a^2 - b^2}[a\tilde{Q}_1(0) - b\tilde{Q}_1(1)] \\
\hat{Q}_1(1) &= \frac{1}{a^2 - b^2}[-b\tilde{Q}_1(0) + a\tilde{Q}_1(1)]
\end{aligned}
$$

and $\hat{\mathbf{Q}}_1$ is unbiased for $\mathbf{Q}_1$.

Using the same colour set C, we now consider linear 3-site neighbourhoods, that is, 3 pixels in a row, either vertically or horizontally, with C = $\{0,1\}$. As we have just mentioned in Example 1.1, there are $2^3 = 8$ possible patterns, namely 000, 001, ..., 111. Recall from equations (1) and (2) in Chapter 1 that

$$Q_3(\delta_1, \delta_2, \delta_3) = \frac{1}{n} \sum_{i=1}^{n} 1\{z_{i,1} = \delta_1, z_{i,2} = \delta_2, z_{i,3} = \delta_3\},$$

where $\delta_i = 0, 1, i = 1, 2, 3$. Following the same idea as above, define

$$\tilde{Q}_3(000) = \frac{1}{n} \sum 1\{Y_{i,1} < \frac{1}{2}\} 1\{Y_{i,2} < \frac{1}{2}\} 1\{Y_{i,3} < \frac{1}{2}\}$$

$$\tilde{Q}_3(001) = \frac{1}{n} \sum 1\{Y_{i,1} < \frac{1}{2}\} 1\{Y_{i,2} < \frac{1}{2}\} 1\{Y_{i,3} \geq \frac{1}{2}\}$$

$$\vdots$$

$$\tilde{Q}_3(111) = \frac{1}{n} \sum 1\{Y_{i,1} \geq \frac{1}{2}\} 1\{Y_{i,2} \geq \frac{1}{2}\} 1\{Y_{i,3} \geq \frac{1}{2}\}.$$

Again, we write $\tilde{Q}_3$ for the vector $(\tilde{Q}_3(000), \tilde{Q}_3(001), \tilde{Q}_3(010), \ldots, \tilde{Q}_3(111))^T$. In the case of 2 grey levels, we use two indicator functions, $1\{Y_i < \frac{1}{2}\}$ for black pixel, and $1\{Y_i \geq \frac{1}{2}\}$ for white pixel (for single-site neighbourhoods). That is, if $Y_i < \frac{1}{2}$, then we say that the colour at site i is black. And if $Y_i \geq \frac{1}{2}$ then we say that the colour at site i is white. Observe that there is a relationship between $\tilde{Q}_1$ and $\tilde{Q}_3$. In the case of linear 3-site neighbourhoods, we use products of 3 indicators chosen from $1\{Y_i < \frac{1}{2}\}, 1\{Y_i \geq \frac{1}{2}\}$. So $\tilde{Q}_1$ and $\tilde{Q}_3$ have similar structure. Since $\tilde{Q}_1$ is biased for Q_1, we would expect that $\tilde{Q}_3$ is also biased for Q_3.

Now, we can compute the expectations of $\tilde{Q}_3(\delta_1, \delta_2, \delta_3)'s$. For example,

$$E\tilde{Q}_3(000) = E\frac{1}{n}\Sigma 1\{Y_{i,1} < \frac{1}{2}\} 1\{Y_{i,2} < \frac{1}{2}\} 1\{Y_{i,3} < \frac{1}{2}\}$$

$$= \frac{1}{n}\Sigma P(Y_{i,1} < \frac{1}{2})P(Y_{i,2} < \frac{1}{2})P(Y_{i,3} < \frac{1}{2})$$

$$= \frac{1}{n}\Sigma\{[P(e_{i,1} < \frac{1}{2\sigma})1\{z_{i,1} = 0\} + P(e_{i,1} < \frac{-1}{2\sigma})1\{z_{i,1} = 1\}] \times$$

$$[P(e_{i,2} < \frac{1}{2\sigma})1\{z_{i,2} = 0\} + P(e_{i,2} < \frac{-1}{2\sigma})1\{z_{i,2} = 1\}] \times$$

$$[P(e_{i,3} < \frac{1}{2\sigma})1\{z_{i,3} = 0\} + P(e_{i,3} < \frac{-1}{2\sigma})1\{z_{i,3} = 1\}]\}.$$

Performing similar calculations for the other seven expectations, and letting a be $\Phi(\frac{1}{2\sigma})$ and b be $\Phi(\frac{-1}{2\sigma})$, we get

$$E\begin{pmatrix} \tilde{Q}_3(000) \\ \tilde{Q}_3(001) \\ \tilde{Q}_3(010) \\ \tilde{Q}_3(011) \\ \tilde{Q}_3(100) \\ \tilde{Q}_3(101) \\ \tilde{Q}_3(110) \\ \tilde{Q}_3(111) \end{pmatrix} = \begin{pmatrix} a^3 & a^2b & a^2b & ab^2 & a^2b & ab^2 & ab^2 & b^3 \\ a^2b & a^3 & ab^2 & a^2b & ab^2 & a^2b & b^3 & ab^2 \\ a^2b & ab^2 & a^3 & a^2b & ab^2 & b^3 & a^2b & ab^2 \\ ab^2 & a^2b & a^2b & a^3 & b^3 & ab^2 & ab^2 & a^2b \\ a^2b & ab^2 & ab^2 & b^3 & a^3 & a^2b & a^2b & ab^2 \\ ab^2 & a^2b & b^3 & ab^2 & a^2b & a^3 & ab^2 & a^2b \\ ab^2 & b^3 & a^2b & ab^2 & a^2b & ab^2 & a^3 & a^2b \\ b^3 & ab^2 & ab^2 & a^2b & ab^2 & a^2b & a^2b & a^3 \end{pmatrix}\begin{pmatrix} Q_3(000) \\ Q_3(001) \\ Q_3(010) \\ Q_3(011) \\ Q_3(100) \\ Q_3(101) \\ Q_3(110) \\ Q_3(111) \end{pmatrix}.$$

In matrix notation, we can write the above as

$$E\tilde{\mathbf{Q}}_3 = \mathbf{A}_3\mathbf{Q}_3$$

As before, we obtain an unbiased estimate of $\mathbf{Q}_3$ by defining $\hat{\mathbf{Q}}_3 = \mathbf{A}_3^{-1}\tilde{\mathbf{Q}}_3$.

As shown in Meloche and Zamar (1994), the above matrices $\mathbf{A_1}, \mathbf{A_3}$ (as well as the corresponding ones for all r) can be analytically inverted. Meloche and Zamar (1994) show that, in general, $E\tilde{\mathbf{Q}}_r = \mathbf{A}_r\mathbf{Q}_r$ with

$$\mathbf{A}_r(\boldsymbol{\delta}, \boldsymbol{\gamma}) = (a)^{\#(\boldsymbol{\delta}, \boldsymbol{\gamma})}(b)^{r-\#(\boldsymbol{\delta}, \boldsymbol{\gamma})}, \tag{5}$$

13

where $\mathbf{A}_r(\boldsymbol{\delta}, \boldsymbol{\gamma})$ is the $(\boldsymbol{\delta}, \boldsymbol{\gamma})$ entry of $\mathbf{A}_r$, and $\#(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \#(k \ni \delta_k = \gamma_k)$ is the number of k such that $\delta_k = \gamma_k$, for $k = 1, 2, \ldots, r$, and that

$$\mathbf{A}_r^{-1}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = (-1)^{\#(\boldsymbol{\delta}, \boldsymbol{\gamma})} \frac{(a)^{\#(\boldsymbol{\delta}, \boldsymbol{\gamma})}(b)^{r-\#(\boldsymbol{\delta}, \boldsymbol{\gamma})}}{(b^2 - a^2)^r} \tag{6}$$

where $a = \Phi(\frac{1}{2\sigma})$, $b = \Phi(\frac{-1}{2\sigma})$. Equation (6) is valid for all r-site neighbourhoods, where r is arbitrary and finite, with $\mathbf{C} = \{0, 1\}$.

As an application of equations (5) and (6) in the case of 2-site neighbourhoods, an unbiased estimate of $\mathbf{Q}_2$ is

$$\begin{pmatrix} \hat{Q}_2(00) \\ \hat{Q}_2(01) \\ \hat{Q}_2(10) \\ \hat{Q}_2(11) \end{pmatrix} = \begin{pmatrix} aa & ab & ba & bb \\ ab & aa & bb & ba \\ ba & bb & aa & ab \\ bb & ba & ab & aa \end{pmatrix}^{-1} \begin{pmatrix} \tilde{Q}_2(00) \\ \tilde{Q}_2(01) \\ \tilde{Q}_2(10) \\ \tilde{Q}_2(11) \end{pmatrix}$$

where

$$\begin{aligned} \tilde{Q}_2(00) &= \frac{1}{n}\Sigma 1\{Y_{i,1} < \frac{1}{2}\} 1\{Y_{i,2} < \frac{1}{2}\} \\ \tilde{Q}_2(01) &= \frac{1}{n}\Sigma 1\{Y_{i,1} < \frac{1}{2}\} 1\{Y_{i,2} \geq \frac{1}{2}\} \\ &\vdots \end{aligned}$$

We now extend this idea for estimating $\mathbf{Q}$ based on indicator functions to a colour set with more than 2 grey levels. We will notice from the following examples that the bias matrix $\mathbf{A}_r$ that needs to be inverted may not have a simple structure.

<u>Example 2.1</u> : Suppose $\mathbf{C} = \{0, 1, 2\}$ and $r = 1$. Now let

$$\varphi(0, y) = 1\{y < \frac{1}{2}\}$$

14

$$\varphi(1, y) = 1\{\frac{1}{2} \le y < \frac{3}{2}\}$$

$$\varphi(2, y) = 1\{y \ge \frac{3}{2}\}$$

and define

$$\tilde{Q}_1(0) = \frac{1}{n}\sum_{i=1}^{n}\varphi(0, Y_i) = \frac{1}{n}\sum_{i=1}^{n}1\{Y_i < \frac{1}{2}\}$$

$$\tilde{Q}_1(1) = \frac{1}{n}\sum_{i=1}^{n}\varphi(1, Y_i) = \frac{1}{n}\sum_{i=1}^{n}1\{\frac{1}{2} \le Y_i < \frac{3}{2}\}$$

$$\tilde{Q}_1(2) = \frac{1}{n}\sum_{i=1}^{n}\varphi(2, Y_i) = \frac{1}{n}\sum_{i=1}^{n}1\{Y_i \ge \frac{3}{2}\}$$

After carrying out some calculations, and letting $a = \Phi(\frac{1}{2\sigma}), b = \Phi(\frac{-1}{2\sigma}), c = \Phi(\frac{-3}{2\sigma}), d = \Phi(\frac{3}{2\sigma}) - \Phi(\frac{1}{2\sigma}), e = \Phi(\frac{1}{2\sigma}) - \Phi(\frac{-1}{2\sigma}), f = \Phi(\frac{-1}{2\sigma}) - \Phi(\frac{-3}{2\sigma}), g = 1 - \Phi(\frac{3}{2\sigma}), h = 1 - \Phi(\frac{1}{2\sigma}), i = 1 - \Phi(\frac{-1}{2\sigma})$, we obtain

$$E\begin{pmatrix} \tilde{Q}_1(0) \\ \tilde{Q}_1(1) \\ \tilde{Q}_1(2) \end{pmatrix} = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}\begin{pmatrix} Q_1(0) \\ Q_1(1) \\ Q_1(2) \end{pmatrix},$$

or, in matrix notation, $E\tilde{\mathbf{Q}}_1 = \mathbf{A}_1\mathbf{Q}_1$. When $r = 2$ (linear 2-site neighbourhoods), there are $3^2 = 9$ possible patterns, namely, 00, 01, 02, 10, 11, 12, 20, 21, and 22. And again by carrying out some calculations, we obtain

$$
E \begin{pmatrix} \tilde{Q}_2(00) \\ \tilde{Q}_2(01) \\ \tilde{Q}_2(02) \\ \tilde{Q}_2(10) \\ \tilde{Q}_2(11) \\ \tilde{Q}_2(12) \\ \tilde{Q}_2(20) \\ \tilde{Q}_2(21) \\ \tilde{Q}_2(22) \end{pmatrix} = \begin{pmatrix} aa & ab & ac & ba & bb & bc & ca & cb & cc \\ ad & ae & af & bd & be & bf & cd & ce & cf \\ ag & ah & ai & bg & bh & bi & cg & ch & ci \\ da & db & dc & ea & eb & ec & fa & fb & fc \\ dd & de & df & ed & ee & ef & fd & fe & ff \\ dg & dh & di & eg & eh & ei & fg & fh & fi \\ ga & gb & gc & ha & hb & hc & ia & ib & ic \\ gd & ge & gf & hd & he & hf & id & ie & if \\ gg & gh & gi & hg & hh & hi & ig & ih & ii \end{pmatrix} \begin{pmatrix} Q_2(00) \\ Q_2(01) \\ Q_2(02) \\ Q_2(10) \\ Q_2(11) \\ Q_2(12) \\ Q_2(20) \\ Q_2(21) \\ Q_2(22) \end{pmatrix},
$$

or $E\tilde{\mathbf{Q}}_2 = \mathbf{A}_2\mathbf{Q}_2$ in matrix notation. But now, $\mathbf{A}_2$ does not seem to have a simple structure.●

In the case of binary images, observe that the two bias matrices, $\mathbf{A}_1, \mathbf{A}_2$ are closely related. In fact $\mathbf{A}_2$ is a **2-fold** Kronecker product of $\mathbf{A}_1$.

<u>**Definition 2.1**</u> (Kronecker product of matrices) :

Let $\mathbf{A} = (a_{ij})$ be a $p \times m$ martix and $\mathbf{B} = (b_{\alpha\beta})$ be a $q \times n$ matrix. The $pq \times mn$ matrix with $a_{ij}b_{\alpha\beta}$ as the element in the i,αth row and the j,βth column is called the *Kronecker or direct product* of $\mathbf{A}$ and $\mathbf{B}$ and is denoted by $\mathbf{A} \otimes \mathbf{B}$; that is,

$$
\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \ldots & a_{1m}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \ldots & a_{2m}\mathbf{B} \\ \vdots & \vdots & & \vdots \\ a_{p1}\mathbf{B} & a_{p2}\mathbf{B} & \ldots & a_{pm}\mathbf{B} \end{pmatrix}.
$$

16

According to Anderson (1984),

$$(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}, \tag{7}$$

that is, the inverse of a Kronecker product of 2 matrices is the Kronecker product of the inverses of the 2 matrices. In what follows, $\mathbf{A}^{(r)}$ denotes the **r-fold** Kronecker product of $\mathbf{A}$, that is

$$\mathbf{A}^{(r)} = \mathbf{A} \otimes \mathbf{A} \otimes \cdots \otimes \mathbf{A}. \tag{8}$$

Since $\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are r-dimensional, we can write the $(\boldsymbol{\delta}, \boldsymbol{\gamma})$ entry of $\mathbf{A}^{(r)}$ as

$$A^{(r)}(\boldsymbol{\delta}, \boldsymbol{\gamma}) = \prod_{i=1}^{r} A(\delta_i, \gamma_i) \bullet \tag{9}$$

In case of binary images, we denote the matrix (bias part) of the 1-site case by $\mathbf{A}_1$, the matrix of the 2-site case by $\mathbf{A}_2$, and the matrix of the 3-site case by $\mathbf{A}_3$. Take a closer look at $\mathbf{A}_1$, and $\mathbf{A}_2$.

$$\mathbf{A}_1 = \begin{pmatrix} a & b \\ b & a \end{pmatrix},$$

$$\mathbf{A}_2 = \begin{pmatrix} aa & ab & ba & bb \\ ab & aa & bb & ba \\ ba & bb & aa & ab \\ bb & ba & ab & aa \end{pmatrix} = \begin{pmatrix} a\mathbf{A}_1 & b\mathbf{A}_1 \\ b\mathbf{A}_1 & a\mathbf{A}_1 \end{pmatrix}.$$

By Definition 2.1, $\mathbf{A}_2$ is a **2-fold** Kronecker product of $\mathbf{A}_1$. Similarly, $\mathbf{A}_3$ is a **3-fold** Kronecker product of $\mathbf{A}_1$.

Note that in Example 2.4, for the colour set of 3 grey levels, $\mathbf{A}_2$ is a **2-fold** Kronecker

product of $\mathbf{A}_1$. Also note that $\mathbf{A}_r^{-1}(\delta, \gamma)$ in equation (6) is the (δ, γ) entry of the r-fold Kronecker product of $\mathbf{A}_1^{-1}$ for the case of binary images.

Given a function $\varphi : \mathbf{C} \times \mathbf{R} \longrightarrow \mathbf{R}$, define

$$\tilde{Q}_1(\delta) = \frac{1}{n} \sum_{i=1}^{n} \varphi(\delta, Y_i) \tag{10}$$

$$\tilde{Q}_r(\delta) = \frac{1}{n} \sum_{i=1}^{n} \varphi(\delta_1, Y_{i,1}) \varphi(\delta_2, Y_{i,2}) \cdots \varphi(\delta_r, Y_{i,r}) \tag{11}$$

In the examples so far, we have been using indicator function as the choice of $\varphi(\delta, y)$.

Proposition 2.1 :

If

$$\tilde{Q}_r(\delta) = \frac{1}{n} \sum_{i=1}^{n} \varphi(\delta_1, Y_{i,1}) \varphi(\delta_2, Y_{i,2}) \cdots \varphi(\delta_r, Y_{i,r})$$

then

$$E\tilde{\mathbf{Q}}_r = \mathbf{A}_1^{(r)} \mathbf{Q}_r \tag{12}$$

where $A_1(\delta, \gamma) = E\varphi(\delta, \gamma + \sigma e)$, and $\mathbf{A}_1^{(r)}$ is the r-fold Kronecker product of $\mathbf{A}_1$. $\mathbf{A}_1^{(r)}$ is a $|\mathbf{C}|^r \times |\mathbf{C}|^r$ matrix, where $|\mathbf{C}|$ is the size of the colour set. Furthermore, if $|\mathbf{A}_1| > 0$, then $\hat{\mathbf{Q}}_r = (\mathbf{A}_1^{-1})^{(r)} \tilde{\mathbf{Q}}_r$ is an unbiased estimate of $\mathbf{Q}_r$. $\bullet$

Proof : We start the proof by expressing the equation $E\tilde{\mathbf{Q}}_r = \mathbf{A}_r \mathbf{Q}_r$. We need to show

$$E\tilde{Q}_r(\delta) = \sum_{\gamma} A_r(\delta, \gamma) Q_r(\gamma) \tag{13}$$

for all δ. Note that

$$
\begin{aligned}
E\tilde{Q}_r(\delta) &= E\frac{1}{n}\sum_{i=1}^{n}\varphi(\delta_1, Y_{i,1})\cdots\varphi(\delta_r, Y_{i,r}) \\
&= \frac{1}{n}\sum_{i=1}^{n}E\varphi(\delta_1, Y_{i,1})\sum_{\gamma_1}1\{z_{i,1}=\gamma_1\}\cdots E\varphi(\delta_r, Y_{i,r})\sum_{\gamma_r}1\{z_{i,r}=\gamma_r\} \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{\gamma_1}1\{z_{i,1}=\gamma_1\}E\varphi(\delta_1, \gamma_1+\sigma e_1)\cdots\sum_{\gamma_r}1\{z_{i,r}=\gamma_r\}E\varphi(\delta_r, \gamma_r+\sigma e_r) \\
&= \frac{1}{n}\sum_{i=1}^{n}\sum_{\gamma_1}1\{z_{i,1}=\gamma_1\}\cdots\sum_{\gamma_r}1\{z_{i,r}=\gamma_r\}E\varphi(\delta_1, \gamma_1+\sigma e_1)\cdots E\varphi(\delta_r, \gamma_r+\sigma e_r) \\
&= \sum_{\gamma}Q_r(\gamma)A_1(\delta_1, \gamma_1)\cdots A_1(\delta_r, \gamma_r) \\
&= \sum_{\gamma}\Pi_{i=1}^{r}A_1(\delta_i, \gamma_i)Q_r(\gamma) \\
&= \sum_{\gamma}A_r(\delta, \gamma)Q_r(\gamma) \\
&= \sum_{\gamma}A_1^{(r)}(\delta, \gamma)Q_r(\gamma)
\end{aligned}
$$

which is just equation (12). Therefore,

$$
E(\mathbf{A}_1^{(r)})^{-1}\tilde{\mathbf{Q}}_r = E(\mathbf{A}_1^{-1})^{(r)}\tilde{\mathbf{Q}}_r = (\mathbf{A}_1^{(r)})^{-1}\mathbf{A}_1^{(r)}\mathbf{Q}_r = \mathbf{Q}_r,
$$

and $\hat{\mathbf{Q}}_r = (\mathbf{A}_1^{-1})^{(r)}\tilde{\mathbf{Q}}_r$ is unbiased for $\mathbf{Q}_r$. $\bullet$

Proposition 2.1 shows how to obtain an unbiased and consistent estimate of $\mathbf{Q}_r$ for any function $\varphi(\delta, y)$ provided $|\mathbf{A}_1| \neq 0$. We therefore have a family of unbiased and consistent estimates. In order to avoid the problem of inverting a large matrix, we can invert $\mathbf{A}_1$ first for the case of single-site neighbourhood, then we can use Definition 2.1 to build the $\mathbf{A}_r^{-1} = (\mathbf{A}^{-1})^{(r)}$.

Example 2.2 : Suppose $C = \{0, 1, 2\}$, $r = 3$, and define

$$
\begin{aligned}
\varphi(0, y) &= 1 \\
\varphi(1, y) &= y \\
\varphi(2, y) &= y^2
\end{aligned}
$$

and as usual

$$\tilde{Q}_3(\delta) = \frac{1}{n}\sum_{i=1}^{n}\varphi(\delta_1, Y_{i,1})\varphi(\delta_2, Y_{i,2})\varphi(\delta_3, Y_{i,3})$$

For instance,

$$\tilde{Q}_3(0,0,0) = \frac{1}{n}\sum_{i=1}^{n} 1\cdot 1\cdot 1 = 1$$

$$\tilde{Q}_3(1,1,1) = \frac{1}{n}\sum_{i=1}^{n} Y_{i,1}\cdot Y_{i,2}\cdot Y_{i,3}$$

$$\tilde{Q}_3(2,2,2) = \frac{1}{n}\sum_{i=1}^{n} 1\cdot Y_{i,2}^2\cdot Y_{i,3}^2$$

Obviously, $\tilde{Q}_3$ is biased for Q_3. Just notice, for instance, that $\tilde{Q}_3(0,0,0) = 1$ independently from the records. By definition, $A_1(\delta, \gamma) = E\varphi(\delta, \gamma + \sigma e)$, we have

$$A_1(0, \gamma) = E1 = 1$$

$$A_1(1, \gamma) = E(\gamma + \sigma e) = \gamma$$

$$A_1(2, \gamma) = E(\gamma + \sigma e)^2 = \gamma^2 + \sigma^2$$

By Proposition 2.1, $\hat{Q}_3 = (A_{-1})^{(3)}\tilde{Q}_3$ is unbiased for Q_3, where

$$
\mathbf{A}_1 = \begin{pmatrix}
A_1(0,0) & A_1(0,1) & A_1(0,2) \\
A_1(1,0) & A_1(1,1) & A_1(1,2) \\
A_1(2,0) & A_1(2,1) & A_1(2,2)
\end{pmatrix} = \begin{pmatrix}
1 & 1 & 1 \\
0 & 1 & 2 \\
\sigma^2 & 1+\sigma^2 & 4+\sigma^2
\end{pmatrix},
$$

and

$$
\mathbf{A}_1^{-1} = \begin{pmatrix}
1 - \frac{\sigma^2}{2} & \frac{-3}{2} & \frac{1}{2} \\
\sigma^2 & 2 & -1 \\
\frac{-\sigma^2}{2} & \frac{-1}{2} & \frac{1}{2}
\end{pmatrix}.
$$

Now, we express $\hat{\mathbf{Q}}_3 = (\mathbf{A}_{-1})^{(3)}\tilde{\mathbf{Q}}_3$ in matrix form.

$$
\begin{pmatrix} \hat{Q}_3(0,0,0) \\ \hat{Q}_3(0,0,1) \\ \hat{Q}_3(0,0,2) \\ \vdots \end{pmatrix} = \begin{pmatrix} (1-\frac{\sigma^2}{2})^3 & (1-\frac{\sigma^2}{2})^2(\frac{-3}{2}) & \cdots & \cdots \\ (1-\frac{\sigma^2}{2})^2(\sigma^2) & (1-\frac{\sigma^2}{2})^2(2) & \cdots & \cdots \\ \vdots & \vdots & \ddots & \\ \vdots & \vdots & & \ddots \end{pmatrix} \begin{pmatrix} \tilde{Q}_3(0,0,0) \\ \tilde{Q}_3(0,0,1) \\ \tilde{Q}_3(0,0,2) \\ \vdots \end{pmatrix} \bullet
$$

2.3 A Simple Representation for $\hat{\mathbf{Q}}_r$

Proposition 2.1 provides a class of unbiased estimates $\hat{\mathbf{Q}}_r$ of $\mathbf{Q}_r$. The proposed estimates have the form $\hat{\mathbf{Q}}_r = (\mathbf{A}^{-1})^{(r)}\tilde{\mathbf{Q}}_r$, where

$$
\tilde{Q}_r(\boldsymbol{\delta}) = \frac{1}{n}\sum_{i=1}^{n}\varphi(\delta_i, Y_{i,1})\varphi(\delta_2, Y_{i,2})\cdots\varphi(\delta_r, Y_{i,r}).
$$

The matrix $(\mathbf{A}^{-1})^{(r)}$ is the r-fold Kronecker product of $\mathbf{A}_1$, and when we have a large colour set, $\mathbf{C}$, and a big r, $(\mathbf{A}^{-1})^{(r)}$ is a large matrix.

In this section, we show that $\hat{\mathbf{Q}}_r$ can be expressed as

$$
\hat{Q}_r(\boldsymbol{\delta}) = \frac{1}{n}\sum_{i=1}^{n}\Psi(\delta_i, Y_{i,1})\Psi(\delta_2, Y_{i,2})\cdots\Psi(\delta_r, Y_{i,r}) \tag{14}
$$

for some $\Psi(\delta, y)$'s.

<u>Proposition 2.2</u> : If we define

$$
\tilde{Q}_r(\boldsymbol{\delta}) = \frac{1}{n}\sum_{i=1}^{n}\varphi(\delta_1, Y_{i,1})\varphi(\delta_2, Y_{i,2})\cdots\varphi(\delta_r, Y_{i,r}),
$$

then $E\tilde{\mathbf{Q}}_r = \mathbf{A}^{(r)}\mathbf{Q}_r$, where $\mathbf{A}^{(r)}$ is the r-fold Kronecker product of the bias matrix $\mathbf{A}_1$, and by definition, $A_1(\delta, \gamma) = E\varphi(\delta, \gamma + \sigma e)$ ($\mathbf{A}_1^{(r)}$ is $|\mathbf{C}|^r \times |\mathbf{C}|^r$). If $|\mathbf{A}_1| > 0$, then

21

$\hat{\mathbf{Q}}_r = (\mathbf{A}^{-1})^{(r)}\tilde{\mathbf{Q}}_r$ is unbiased and consistent for $\mathbf{Q}_r$, and

$$\hat{Q}_r(\boldsymbol{\delta}) = \sum_{\gamma} A_{\boldsymbol{\delta},\gamma}^{-(r)} \tilde{Q}_r(\gamma) \qquad (15)$$

can be expressed as

$$\hat{Q}_r(\boldsymbol{\delta}) = \frac{1}{n} \sum_{i=1}^{n} \Psi(\delta_1, Y_{i,1}) \cdots \Psi(\delta_r, Y_{i,r}) \qquad (16)$$

where

$$\Psi(\delta, y) = \sum_{\gamma} A_{\delta\gamma}^{-1} \varphi(\gamma, y). \bullet \qquad (17)$$

Proof:

$$
\begin{aligned}
\hat{Q}_r(\boldsymbol{\delta}) &= \sum_{\gamma} A_{\boldsymbol{\delta},\gamma}^{-(r)} \tilde{Q}_r(\gamma) \\
&= \sum_{\gamma_1} \cdots \sum_{\gamma_r} [A_{\delta_1,\gamma_1}^{-1} \cdots A_{\delta_r,\gamma_r}^{-1} \tilde{Q}_r(\gamma_1,\ldots,\gamma_r)] \\
&= \sum_{\gamma_1} \cdots \sum_{\gamma_r} [A_{\delta_1,\gamma_1}^{-1} \cdots A_{\delta_r,\gamma_r}^{-1} \frac{1}{n} \sum_{i=1}^{n} \varphi(\gamma_1, Y_{i,1}) \cdots \varphi(\gamma_r, Y_{i,r})] \\
&= \frac{1}{n} \sum_{i=1}^{n} \{ \sum_{\gamma_1} \cdots \sum_{\gamma_r} [A_{\delta_1,\gamma_1}^{-1} \varphi(\gamma_1, Y_{i,1}) \cdots A_{\delta_r,\gamma_r}^{-1} \varphi(\gamma_r, Y_{i,r})] \} \\
&= \frac{1}{n} \sum_{i=1}^{n} \{ (\sum_{\gamma_1} A_{\delta_1,\gamma_1}^{-1} \varphi(\gamma_1, Y_{i,1})) \cdots (\sum_{\gamma_r} A_{\delta_r,\gamma_r}^{-1} \varphi(\gamma_r, Y_{i,r})) \} \\
&= \frac{1}{n} \sum_{i=1}^{n} [\Psi(\delta_1, Y_{i,1}) \cdots \Psi(\delta_r, Y_{i,r})].
\end{aligned}
$$

Note that with this representation, $\hat{\mathbf{Q}}_r$ is easily seen to be unbiased. Indeed,

$$
\begin{aligned}
E\Psi(\delta, \gamma + \sigma e) &= E \sum_{\tau} A_{\delta\tau}^{-1} \varphi(\tau, \gamma) \\
&= \sum_{\tau} A_{\delta\tau}^{-1} A_{\tau,\gamma} \\
&= 1\{\delta = \gamma\}.
\end{aligned}
$$

And so

$$E\hat{Q}_r(\delta) = \frac{1}{n}\sum_{i=1}^{n} E\Psi(\delta_1, Y_{i,1})\Psi(\delta_2, Y_{i,2})\cdots\Psi(\delta_r, Y_{i,r})$$

$$= \frac{1}{n}\sum_{i=1}^{n} 1\{z_{i,1} = \delta_1\}1\{z_{i,2} = \delta_2\}\cdots 1\{z_{i,r} = \delta_r\}$$

$$= \frac{1}{n}\sum_{i=1}^{n} 1\{z_{i,1} = \delta_1, \ldots, z_{i,r} = \delta_r\}$$

$$= Q_r(\delta).$$

Note that since $\hat{Q}_r(\delta)$ is an average of an r-dependent sequence, it converges to its mean. Therefore, $\hat{Q}_r(\delta)$ is an unbiased and consistent estimate of $Q_r(\delta)$.●

Example 2.2 (Continued) : Recall that $\varphi(0, y) = 1, \varphi(1, y) = y, \varphi(2, y) = y^2$. By equation (17),

$$\begin{pmatrix} \Psi(0,y) \\ \Psi(1,y) \\ \Psi(2,y) \end{pmatrix} = \begin{pmatrix} 1 - \frac{\sigma^2}{2} & \frac{-3}{2} & \frac{1}{2} \\ \sigma^2 & 2 & -1 \\ \frac{-\sigma^2}{2} & \frac{-1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} \varphi(0,y) \\ \varphi(1,y) \\ \varphi(2,y) \end{pmatrix}$$

Therefore, we can write

$$\Psi(0,y) = (1 - \frac{\sigma^2}{2})1 + (\frac{-3}{2})y + (\frac{1}{2})y^2,$$

$$\Psi(0,y) = (\sigma^2)1 + (2)y + (-1)y^2,$$

$$\Psi(1,y) = (\frac{-\sigma^2}{2})1 + (\frac{-1}{2})y + (\frac{1}{2})y^2,$$

and so

$$\hat{Q}_3(0,0,0) = \frac{1}{n}\sum_{i=1}^{n} \Psi(0, Y_{i,1})\Psi(0, Y_{i,2})\Psi(0, Y_{i,3}),$$

$$\hat{Q}_3(0,0,1) = \frac{1}{n}\sum_{i=1}^{n} \Psi(0, Y_{i,1})\Psi(0, Y_{i,2})\Psi(1, Y_{i,3}),$$

$$\vdots$$

23

$$\hat{Q}_3(2,2,2) \;\; = \;\; \frac{1}{n}\sum_{i=1}^{n}\Psi(2,Y_{i,1})\Psi(2,Y_{i,2})\Psi(2,Y_{i,3}).$$

As stated in Proposition 2.2, such a new representation can be obtain generally.●

In summary, we have the following general procedure for obtaining unbiased and consistent estimates for $\mathbf{Q}_r$: Let $\mathbf{C} = \{c_1, c_2, \ldots, c_k\}$, and $r > 0$. Take any $\varphi : \mathbf{C} \times \mathbf{R} \longrightarrow \mathbf{R}$, and then define $\mathbf{A}$ as the $|\mathbf{C}| \times |\mathbf{C}|(= k \times k)$ bias matrix with elements $A_{\delta\gamma} = E\varphi(\delta, \gamma + \sigma e)$. If $|\mathbf{A}| > 0$, then define

$$\hat{Q}_r(\boldsymbol{\delta}) \;\; = \;\; \frac{1}{n}\sum_{i=1}^{n}\Psi(\delta_1, Y_{i,1})\Psi(\delta_2, Y_{i,2})\cdots\Psi(\delta_r, Y_{i,r})$$

where

$$\Psi(\delta, Y) = \sum_{\gamma} A_{\delta\gamma}^{-1}\varphi(\delta, Y).$$

Then $\hat{Q}_r(\boldsymbol{\delta})$ is an unbiased and consistent estimate of $Q_r(\boldsymbol{\delta})$.

<u>Example 2.3</u> : We use Figure 2.1 as the true and unobserved image, and Figure 2.2 as the noisy and observed image. The colour set $\mathbf{C}$ is $\{0, 1, 2\}$, r is the neighbourhood size, and $\sigma = 0.50$. The following tables shows how far away the biased $\tilde{\mathbf{Q}}_r$ is from the true $\mathbf{Q}_r$, and how good and close the unbiased $\hat{\mathbf{Q}}_r$ is to the true $\mathbf{Q}_r$.

Table 2.1: Pattern distribution with $r = 1$, $\varphi(\delta, y) =$ Indicator functions

δ	$Q_1(\delta)$	$\tilde{Q}_1(\delta)$	$\hat{Q}_1(\delta)$
0	0.48340	0.47498	0.48505
1	0.42505	0.37836	0.42074
2	0.09155	0.14667	0.09421
Total	1.00000	1.00000	1.00000

Table 2.2: Pattern distribution with $r = 1$, $\varphi(\delta, y) =$ Power functions

δ	$Q_1(\delta)$	$\tilde{Q}_1(\delta)$	$\hat{Q}_1(\delta)$
0	0.48340	0.37644	0.48306
1	0.42505	0.22966	0.42380
2	0.09155	0.39390	0.09314
Total	1.00000	1.00000	1.00000

Table 2.3: Pattern distribution with $r = 2$, $\varphi(\delta, y) =$Indicator functions

δ	$Q_2(\delta)$	$\tilde{Q}_2(\delta)$	$\hat{Q}_2(\delta)$
0 0	0.47192	0.34601	0.46966
0 1	0.00842	0.11395	0.01315
0 2	0.00305	0.01501	0.00224
1 0	0.00842	0.11377	0.01274
1 1	0.41541	0.20496	0.40092
1 2	0.00122	0.05963	0.00709
2 0	0.00305	0.01520	0.00266
2 1	0.00122	0.05945	0.00667
2 2	0.08728	0.07202	0.08488
Total	1.00000	1.00000	1.00000

Table 2.4: Pattern distribution with $r = 2$, $\varphi(\delta, y) =$ Power functions

δ	$Q_2(\delta)$	$\tilde{Q}_2(\delta)$	$\hat{Q}_2(\delta)$
0 0	0.47192	0.09996	0.46610
0 1	0.00842	0.06099	0.00452
0 2	0.00305	0.10460	0.00598
1 0	0.00842	0.06099	0.00480
1 1	0.41541	0.07745	0.42015
1 2	0.00122	0.12890	0.00000
2 0	0.00305	0.10460	0.00570
2 1	0.00122	0.12896	0.00000
2 2	0.08728	0.23355	0.09275
Total	1.00000	1.00000	1.00000

For $r = 3$ (linear 3-site neighbourhood), the columns $Q_3(\delta), \tilde{Q}_3(\delta)$, and $\hat{Q}_3(\delta)$ are displayed graphically as barplots in Figures 2.3a, 2.3b, 2.3c, 2.3d, 2.3e, 2.3f. Finally, for $r = 5$ (cross-shaped neighbourhood), the columns $Q_5(\delta), \tilde{Q}_5(\delta)$, and $\hat{Q}_5(\delta)$ are also displayed graphically as barplots in Figures 2.4a, 2.4b, 2.4c, 2.4d, 2.4e, 2.4f.

We have seen that different neighbourhood shapes and sizes give different pattern distributions. They also give different estimates $\{\hat{z}_1, \hat{z}_2, \ldots, \hat{z}_n\}$, because as mentioned before, the estimate of the colour at each site is a function of the data, Y_i, available in a neighbourhood of that site. The estimate $\hat{z}_i$ proposed by Meloche and Zamar (1994) requires estimates of pattern distributions, $Q_r(\delta)$, and σ. Recall from equation (4) that

$$\hat{z}_i = \frac{\sum_\delta \delta_c \phi_\sigma(\mathbf{Y}_i - \delta) Q_r(\delta)}{\sum_\delta \phi_\sigma(\mathbf{Y}_i - \delta) Q_r(\delta)},$$

where ϕ_σ denotes the r-dimensional normal density with mean zero and covariance $\sigma^2 I_r$, and δ_c denotes the colour/pattern at the center of the neighbourhood. Figures 2.5a to 2.5h are the restored images based on different neighbourhood shapes and sizes, and different estimations of the pattern distributions.●

3 Joint Estimations of Q and σ.

3.1 Introduction

Meloche and Zamar (1994) propose estimating z_i by

$$\hat{z}_i = \frac{\sum_{\delta} \delta_c \phi_{\sigma}(\mathbf{Y}_i - \delta) \hat{Q}_r(\delta)}{\sum_{\delta} \phi_{\sigma}(\mathbf{Y}_i - \delta) \hat{Q}_r(\delta)},$$

where δ_c denotes the colour/pattern at the center of the neighbourhood, and ϕ_{σ} is the normal density with mean zero and variance σ^2. Furthermore, our proposed estimation of $\mathbf{Q}_r$ involves σ, and we have been assuming that σ is known. Equation (16) states that

$$\hat{Q}_r(\delta) = \frac{1}{n} \sum_{i=1}^{n} \prod_{j=1}^{r} \Psi_{\sigma}(\delta_j, Y_{i,j}).$$

The subscript σ indicates that $\hat{Q}_r(\delta)$ involves σ. But σ is likely to be unknown and must be estimated from the noisy image. Ideally, we want to estimate both $\mathbf{Q}_r$ and σ, but we will focus on the estimations of $\mathbf{Q}_1$ and σ. Once σ is estimated, $\mathbf{Q}_r$ can also be estimated. For the rest of this thesis, we write $\mathbf{Q}$ instead of $\mathbf{Q}_1$. In this chapter, we propose some estimating equations for σ.

3.2 Estimation of σ using Estimating Equation derived from $E(\frac{1}{n}\sum_{i=1}^{n} Y_i^2)$

Note that

$$
\begin{aligned}
E(\frac{1}{n}\sum_{i=1}^{n} Y_i^2) &= \frac{1}{n}\sum_{i=1}^{n} E(Y_i^2) \\
&= \frac{1}{n}\sum_{i=1}^{n} E(z_i + \sigma e)^2 \\
&= \frac{1}{n}\sum_{i=1}^{n} E(z_i^2 + 2z_i\sigma e + \sigma^2 e^2)
\end{aligned}
$$

$$\begin{aligned}
&= \frac{1}{n}\sum_{i=1}^{n}(z_i^2 + \sigma^2) \\
&= \sigma^2 + \frac{1}{n}\sum_{i=1}^{n} z_i^2 \\
&= \sigma^2 + \frac{1}{n}\sum_{i=1}^{n}\sum_{\delta} 1\{z_i = \delta\}\delta^2 \\
&= \sigma^2 + \sum_{\delta}\frac{1}{n}\sum_{i=1}^{n} 1\{z_i = \delta\}\delta^2 \\
&= \sigma^2 + \sum_{\delta}\delta^2 Q(\delta).
\end{aligned}$$

Thus, if $\hat{\mathbf{Q}}_\sigma$ is any unbiased estimate of $\mathbf{Q}$,

$$E\left(\sigma^2 - \frac{1}{n}\sum_{i=1}^{n} Y_i^2 + \sum_{\delta}\delta^2\hat{Q}_\sigma(\delta)\right) = 0. \tag{18}$$

The subscript σ is used to indicate that $\hat{\mathbf{Q}}_\sigma$ involves σ. Now define

$$\lambda_n(a) = a^2 - \frac{1}{n}\sum_{i=1}^{n} Y_i^2 + \sum_{\delta}\delta^2\hat{Q}_a(\delta) \tag{19}$$

and taking expection of equation (19),

$$\lambda(a) = E\lambda_n(a) = a^2 - \left(\sigma^2 + \sum_{\delta}\delta^2 Q(\delta)\right) + \sum_{\delta}\delta^2 E\hat{Q}_a(\delta). \tag{20}$$

Note that according to equation (18), σ is one of the root of $\lambda(a)$. Empirical evidence suggests that when $\hat{\mathbf{Q}}$ is derived from indicator functions, $\lambda(a)$ has 2 roots, the smaller of which is σ (irrespective of the colour set $\mathbf{C}$). Theoretical results to that effect have not been reached yet.

When $\hat{\mathbf{Q}}_\sigma$ is derived from power functions, and when $\mathbf{C} = \{c_1, c_2\}$, equation (19) yields a closed-form estimate of σ^2 instead of σ.

$$\lambda_n(a) = a^2 - \frac{1}{n}\sum_{i=1}^{n} Y_i^2 + \frac{1}{n}\sum_{i=1}^{n} \frac{c_1^2(c_2 - Y_i) + c_2^2(-c_1 + Y_i)}{c_2 - c_1},$$

$$\lambda(a) = a^2 - (\sigma^2 + c_1^2 Q(c_1) + c_2^2 Q(c_2)) + c_1^2 Q(c_1) + c_2^2 Q(c_2).$$

By solving $\lambda_n(a) = 0$, σ^2 can be estimated as

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{1}{n}\sum_{i=1}^{n} \frac{c_1^2(c_2 - Y_i) + c_2^2(-c_1 + Y_i)}{c_2 - c_1}.$$

This is just the estimate of σ^2 proposed by Meloche and Zamar (1994) for the colour set $\{0, 1\}$. It is unbiased and consistent.

When $|C| = 3$, it can be shown that $\lambda(a) \equiv 0$ if $\hat{Q}_\sigma$ is derived from power functions. If $C = \{c_1, c_2, c_3\}$, and $\varphi(c_1, y) = 1, \varphi(c_2, y) = y, \varphi(c_3, y) = y^2$, then by equation (9),

$$\tilde{Q}(c_1) = \frac{1}{n}\sum_{i=1}^{n} 1,$$

$$\tilde{Q}(c_2) = \frac{1}{n}\sum_{i=1}^{n} Y_i,$$

$$\tilde{Q}(c_3) = \frac{1}{n}\sum_{i=1}^{n} Y_i^2,$$

then $\tilde{\mathbf{Q}} = \mathbf{A}_a \mathbf{Q}$, where

$$A_a = \begin{pmatrix} 1 & 1 & 1 \\ c_1 & c_2 & c_3 \\ c_1^2 + a^2 & c_2^2 + a^2 & c_3^2 + a^2 \end{pmatrix},$$

and

$$A_a^{-1} = \frac{1}{|\mathbf{A}_a|} \begin{pmatrix} (c_2 - c_3)(a^2 - c_2 c_3) & (c_2 - c_3)(c_2 + c_3) & -(c_2 - c_3) \\ -(c_1 - c_3)(a^2 - c_1 c_3) & -(c_1 - c_3)(c_1 + c_3) & (c_1 - c_3) \\ (c_1 - c_2)(a^2 - c_1 c_2) & (c_1 - c_2)(c_1 + c_2) & -(c_1 - c_2) \end{pmatrix},$$

31

where $|\mathbf{A}| = -(c_1 - c_2)(c_1 - c_3)(c_2 - c_3)$. We can write equation (20) in matrix form. Let $\Delta = (c_1^2, c_2^2, c_3^2)^T$,

$$
\begin{aligned}
\lambda(a) &= a^2 - \left(\sigma^2 + \sum_\delta \delta^2 Q(\delta)\right) + \sum_\delta \delta^2 E \hat{Q}_a(\delta) \\
&= a^2 - \sigma^2 + \sum_\delta \delta^2 E \hat{Q}_a(\delta) - \sum_\delta \delta^2 Q(\delta).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\lambda(a) &= a^2 - \sigma^2 + \Delta^T \mathbf{A}_a^{-1} \mathbf{A}_\sigma \mathbf{Q} + \Delta^T \mathbf{Q} \\
&= a^2 - \sigma^2 + \Delta^T (\mathbf{A}_a^{-1} \mathbf{A}_\sigma - \mathbf{I_3}) \mathbf{Q},
\end{aligned}
$$

where

$$
\mathbf{A}_a^{-1}\mathbf{A}_\sigma - \mathbf{I} =
\begin{pmatrix}
-\left(\frac{a^2-\sigma^2}{c_2 c_3 - c_1 c_2 + c_1^2 - c_1 c_3}\right) & -\left(\frac{a^2-\sigma^2}{c_2 c_3 - c_1 c_2 + c_1^2 - c_1 c_3}\right) & -\left(\frac{a^2-\sigma^2}{c_2 c_3 - c_1 c_2 + c_1^2 - c_1 c_3}\right) \\
\frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_2 c_3 - c_2^2} & \frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_2 c_3 - c_2^2} & \frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_2 c_3 - c_2^2} \\
-\left(\frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_3^2 - c_2 c_3}\right) & -\left(\frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_3^2 - c_2 c_3}\right) & -\left(\frac{a^2-\sigma^2}{-c_1 c_3 + c_1 c_2 + c_3^2 - c_2 c_3}\right)
\end{pmatrix},
$$

and

$$
\begin{aligned}
\Delta^T (\mathbf{A}_a^{-1} \mathbf{A}_\sigma - \mathbf{I_3}) \mathbf{Q} &= \frac{-c_1^2(a^2 - \sigma^2)}{c_2 c_3 - c_1 c_2 + c_1^2 - c_1 c_3} + \frac{c_2^2(a^2 - \sigma^2)}{-c_1 c_3 + c_1 c_2 + c_2 c_3 - c_2^2} \\
&\quad + \frac{-c_3^2(a^2 - \sigma^2)}{-c_1 c_3 + c_1 c_2 + c_3^2 - c_2 c_3} \\
&= -(a^2 - \sigma^2).
\end{aligned}
$$

As a result,

$$
\lambda(a) = a^2 - \sigma^2 - (a^2 - \sigma^2) = 0.
$$

Thus, when $|\mathbf{C}| = 3$, $\lambda(a)$ defined by equation (20) is identically zero. To get an estimate of σ^2 when $|\mathbf{C}| = 3$, one possibility is to derive a different estimating equation starting from higher moments of Y_i. For example, we start from Y_i^4.

32

3.3 Estimation of σ using Estimating Equation derived from $E(\frac{1}{n}\sum_{i=1}^{n} Y_i^4)$

By simple computation,

$$E(\frac{1}{n}\Sigma Y_i^4) = 3\sigma^4 + 6\sigma^2 \Sigma_\delta \delta^2 Q(\delta) + \Sigma_\delta \delta^4 Q(\delta). \tag{21}$$

Thus, if $\hat{Q}_\sigma$ is any unbiased estimate Q,

$$E(3\sigma^4 + 6\sigma^2 \sum_\delta \delta^2 \hat{Q}_\sigma(\delta) + \sum_\delta \delta^4 \hat{Q}_\sigma(\delta) - \frac{1}{n}\sum_{i=1}^{n} Y_i^4) = 0. \tag{22}$$

Define

$$\lambda_n(a) = 3a^4 + 6a^2 \sum_\delta \delta^2 \hat{Q}_a(\delta) + \sum_\delta \delta^4 \hat{Q}_a(\delta) - \frac{1}{n}\sum_{i=1}^{n} Y_i^4 \tag{23}$$

and taking expectation of equation (23),

$$\lambda(a) = 3a^4 + 6a^2 \sum_\delta \delta^2 E\hat{Q}_a(\delta) + \sum_\delta \delta^4 E\hat{Q}_a(\delta) - E(\frac{1}{n}\sum_{i=1}^{n} Y_i^4). \tag{24}$$

According to equation (22), σ is one of the root of $\lambda(a)$. By equation (14),

$$\hat{Q}_a(c_1) = \frac{1}{|\mathbf{A}|}\{\frac{1}{n}\sum_{i=1}^{n}[(c_2 - c_3)(a^2 - c_2 c_3) + (c_2^2 - c_3^2)Y_i - (c_2 - c_3)Y_i^2]\},$$

$$\hat{Q}_a(c_2) = \frac{1}{|\mathbf{A}|}\{\frac{1}{n}\sum_{i=1}^{n}[-(c_1 - c_3)(a^2 - c_1 c_3) - (c_1^2 - c_3^2)Y_i + (c_1 - c_3)Y_i^2]\},$$

$$\hat{Q}_a(c_3) = \frac{1}{|\mathbf{A}|}\{\frac{1}{n}\sum_{i=1}^{n}[(c_1 - c_2)(a^2 - c_1 c_2) + (c_1^2 - c_2^2)Y_i - (c_1 - c_2)Y_i^2]\}.$$

Then note that $\lambda_n(a)$ and $\lambda(a)$ defined by equations (23) and (24) are second degree polynomials in a^2. As a result, $\lambda(a)$ has 2 roots. The smaller of which is σ^2. In general, σ^2 can be estimated as the smallest root of $\lambda_n(a)$. In particular, when $\mathbf{C} = \{0, 1, 2\}$,

$$\hat{Q}_a(0) = \frac{1}{n}\sum_{i=1}^{n}[(1 - \frac{a^2}{2})1 + (\frac{-3}{2})Y_i + (\frac{1}{2})Y_i^2]$$

$$\hat{Q}_a(1) = \frac{1}{n}\sum_{i=1}^{n}[(a^2)1 + (2)Y_i + (-1)Y_i^2]$$

$$\hat{Q}_a(2) = \frac{1}{n}\sum_{i=1}^{n}[(\frac{-a^2}{2})1 + (\frac{-1}{2})Y_i + (\frac{1}{2})Y_i^2]$$

By solving $\lambda_n(a) = 0$, we have

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{7}{6} \pm \sqrt{(\frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{7}{6})^2 - \frac{1}{3n}\sum_{i=1}^{n}(-6Y_i + 7Y_i^2 - Y_i^4)},$$

$$\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2 + Q(1) + 4Q(2) - \frac{7}{6} \pm \sqrt{(Q(1) + 4Q(2) - \frac{7}{6})^2}.$$

We can conclude that for the colour set $\{0, 1, 2\}$, σ^2 can be estimated as

$$\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{7}{6} - \sqrt{(\frac{1}{n}\sum_{i=1}^{n} Y_i^2 - \frac{7}{6})^2 - \frac{1}{3n}\sum_{i=1}^{n}(-8Y_i + 8Y_i^2 - Y_i^4)},$$

Although $\hat{\sigma}_n^2$ may not be unbiased, it is consistent. The estimate of σ can be obtained by taking the square root of $\hat{\sigma}_n^2$. In general, we can derive an estimating equation for σ^2 with an arbitrary colour set in a similar fashion when $\mathbf{Q}_r$ is derived from power functions.

4 Neighbourhood Shapes.

We have seen in the previous examples that the bigger the neighbourhood size, the better the restoration performance. But this will no longer hold if the neighbourhood size is big while the image is small. Sometimes a good and small neighbourhood may result in better performance in restoration than a big and bad one. The performance of the estimates $\hat{z}_1, \ldots, \hat{z}_n$ can be measured by the average expected square error:

$$AMSE = \frac{1}{n} \sum_{i=1}^{n} E(\hat{z}_i - z_i)^2$$

According to the theorem by Chan and Meloche (1995),

$$AMSE = \frac{1}{n} \sum_{i=1}^{n} E(\hat{z}_i - z_i)^2 = \sigma^2 (1 - \sigma^2 I_0(\eta_\sigma * Q)). \tag{25}$$

$\sigma^2(1 - \sigma^2 I_0(\eta_\sigma * Q))$ can be obtained by numerical integration or Montecarlo. $I_0(\eta_\sigma * Q)$ is the middle element on the diagonal of the Fisher information matrix of $\eta_\sigma * Q$ (recall that our vectors are indexed from 1 to r, so the middle element on the diagonal element is the $(\frac{1+r}{2})^{th}$ element. η_σ is the normal density with mean zero and variance σ^2. Figure 3.1 is a 129×129 true image with alternating horizontal strips of 0, 1, and 2, and Figure 3.2 is the degraded version of the true image with $\sigma = 0.50$. Table 3.1 provides the approximate $AMSE$ (obtained by Montecarlo) for various neighbourhood sizes and shapes. In the table,

$$
\begin{aligned}
AMSE_1 &= \sigma^2(1 - \sigma^2 I_0(\eta_\sigma * Q)), \\
AMSE_2 &= \sigma^2(1 - \sigma^2 I_0(\eta_\sigma * \hat{Q}_{indicator})), \\
AMSE_3 &= \sigma^2(1 - \sigma^2 I_0(\eta_\sigma * \hat{Q}_{power})),
\end{aligned}
$$

where Q is the true pattern distribution, $\hat{Q}_{indicator}$ is the estimated pattern distribution based on the indicator functions, and $\hat{Q}_{power}$ is the estimated pattern distribution based on the power functions.

Table 4.1: $AMSE$ for various neighbourhood sizes and shapes

neighbourhood	$AMSE_1$	$AMSE_2$	$AMSE_3$
linear 3-site nbhd (horizontal)	0.042	0.050	0.054
linear 3-site nbhd (vertical)	0.021	0.032	0.031
linear 5-site nbhd (horizontal)	0.013	0.042	0.080
linear 5-site nbhd (vertical)	0.0053	0.043	0.056
5-site nbhd (cross-shaped)	0.0025	0.045	0.043
linear 7-site nbhd (horizontal)	0.0042	0.092	0.125

Note that when the $Q_r(\delta)$'s are known, the cross-shaped neighbourhood results in a better performance with the lowest $AMSE_1$ than the other neighbourhoods used here. In particular, it performs better than a larger neighbourhood which is the linear and horizontal 7-site neighbourhood. In pratice, we substitute $\hat{Q}_r(\delta)$'s for $Q_r(\delta)$'s in equation (25) when we only have a noisy image. We choose the neighbourhood shape which gives the smallest $AMSE$.

When we use the estimates $\hat{Q}_r(\delta)$'s in equation (25) for obtaining the $AMSE_2$ and $AMSE_3$, the linear and vertical 3-site neighbourhood seems to be the best choice because its corresponding $\hat{Q}_3(\delta)$'s are relatively more accurate than the $\hat{Q}_5(\delta)$'s and $\hat{Q}_7(\delta)$'s. When the neighbourhood size gets larger, the estimates $\hat{Q}_r$'s become less accurate. Therefore, larger neighbourhood results in a worse restoration performance for our particular noisy image in Figure 3.2. If we have a larger image, the larger neighbourhoods may perform better than the smaller ones because the estimates $\hat{Q}_r$ for the larger neighbourhoods become more accurate.

5 Conclusions.

By extending the idea of estimating $\mathbf{Q}_r$ based on indicator functions by Meloche and Zamar (1994), we have developed a family of estimates of $\mathbf{Q}_r$ which can be based on any arbitrary choice of a set of $\varphi(\delta, y)$'s. We start with any set of $\varphi(\delta, y)$'s which seems to have a simple structure. Then we define

$$\tilde{Q}_r(\boldsymbol{\delta}) = \frac{1}{n} \sum_{i=1}^{n} \varphi(\delta_1, Y_{i,1}) \varphi(\delta_2, Y_{i,2}) \cdots \varphi(\delta_r, Y_{i,r})$$

which is biased for $\hat{Q}_r(\boldsymbol{\delta})$. By applying the propositions stated before, we obtain a new set of $\Psi(\delta, y)$'s, where

$$\Psi(\delta, y) = \sum_{\gamma} A_{\delta,\gamma}^{-1} \varphi(\gamma, y),$$

such that

$$\hat{Q}_r(\boldsymbol{\delta}) = \frac{1}{n} \sum_{i=1}^{n} \Psi(\delta_1, Y_{i,1}) \Psi(\delta_2, Y_{i,2}) \cdots \Psi(\delta_r, Y_{i,r})$$

is an unbiased and consistent estimate of $Q_r(\boldsymbol{\delta})$. At the moment, we lack the theoretical results on judging which set of $\Psi(\delta, y)$'s give the best and the most accurate estimate of $\mathbf{Q}_r$.

We have addressed the problem of estimating σ by proposing some estimating equations for σ. We have derived an estimating equation from $E(\frac{1}{n} \sum_{i=1}^{n} Y_i^2)$ for a colour set with $|\mathbf{C}| = 2$. By solving $\lambda_n(a) = 0$, we have obtained an estimate of σ^2 for any colour set with $|\mathbf{C}| = 2$.

But the estimating equation derived from $E(\frac{1}{n} \sum_{i=1}^{n} Y_i^2)$ does not work for a bigger colour set when $\mathbf{Q}_r$ is derived fron power functions. So we have derived another estimating equation from $E(\frac{1}{n} \sum_{i=1}^{n} Y_i^4)$ for a colour set with $|\mathbf{C}| = 3$. Again, by solving $\lambda_n(a) = 0$,

we have obtained an estimate of σ^2 for any colour set with $|C| = 3$. To estimate σ^2 when $|C| \geq 3$, we can derive a different equation starting from a higher moment of Y_i. In general, we estimate σ^2 by deriving an estimating equation for σ^2 in this fashion. The esimate may not be unbiased, but it is consistent. By taking the square root of σ_n^2, we obtain the estimate of σ.

When Q_r is derived from indicator functions, the empirical evidence suggests that $\lambda(a)$ has 2 solutions for a, and the smaller of which is σ. Irrespective of the colour set C, estimating equation (18) always works. Therefore, we do not have to consider higher moments of Y_i. But theoretical results have not been reached yet.

REFERENCES

Anderson, T.W. (1984) *An Introduction to Multivariate Statistical Analysis.* Wiley.

Chan, A., Meloche, J. (1995) Estimation of a Gaussian Mean and Image Restoration. *To be submitted for publication.*

Meloche, J., Zamar, R. (1994) Black and White Image Restoration. *The Canadian Journal of Statistics.* **22, 3**, 335-355.

Figure 2.1: True Image (Unobserved)
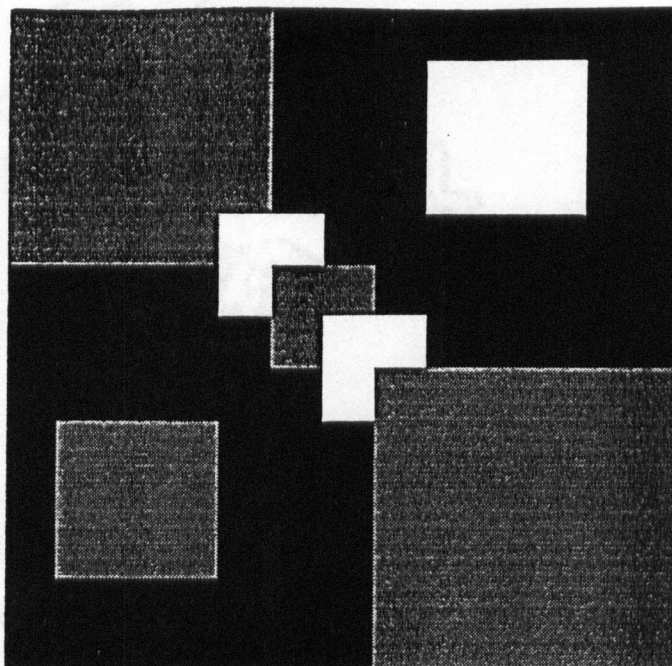
Figure 2.2: Noisy Image (Observed)

41
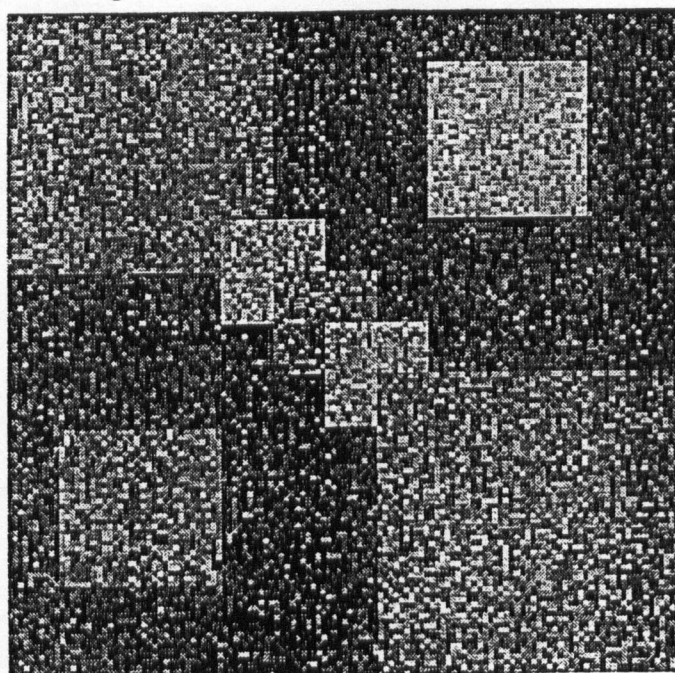
Figure 2.3: Pattern distribution with $r = 3$, and (i) $\varphi(\delta, y) =$ Indicator functions
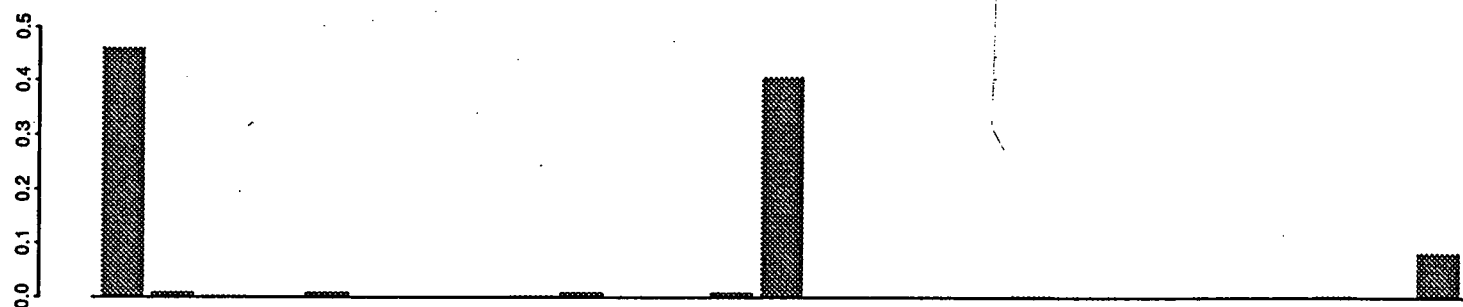
(a) $Q_3(\delta)$ vs δ
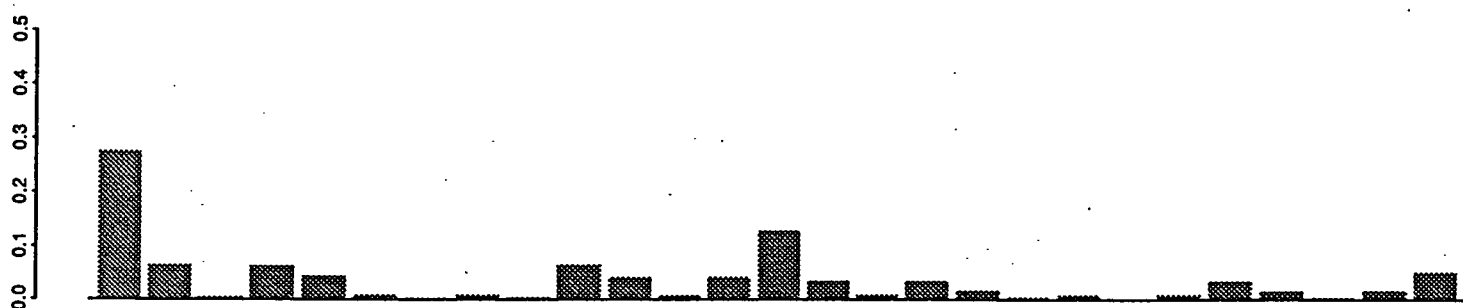
(b) $\tilde{Q}_3(\delta)$ vs δ
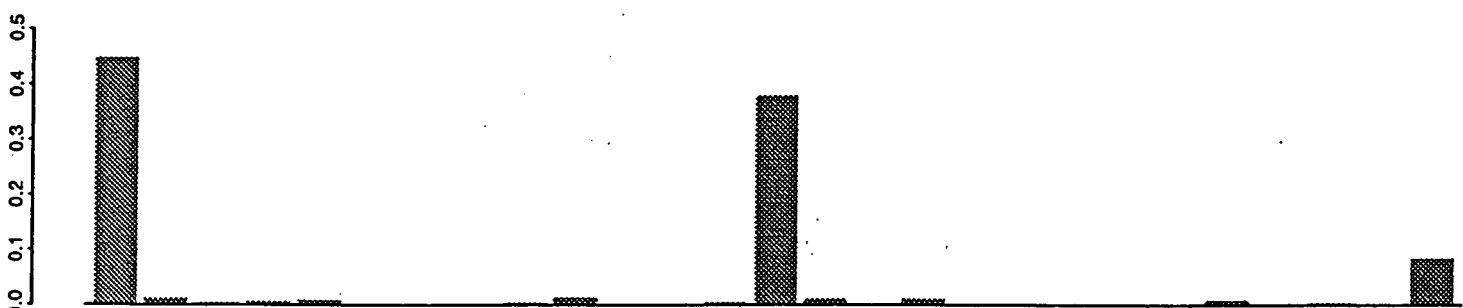
(c) $\hat{Q}_3(\delta)$ vs δ

Figure 2.3: Pattern distribution with $r = 3$, and (ii) $\varphi(\delta, y) =$ Power functions

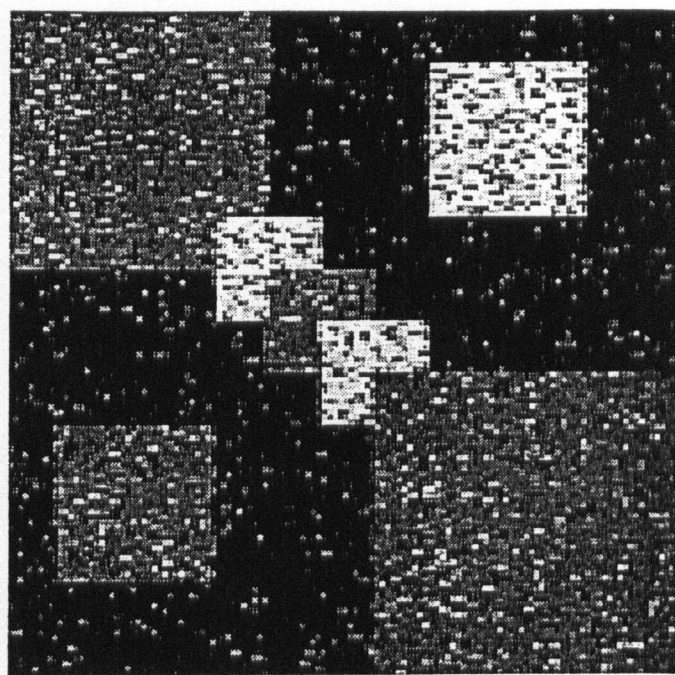(d) $Q_3(\delta)$ vs δ

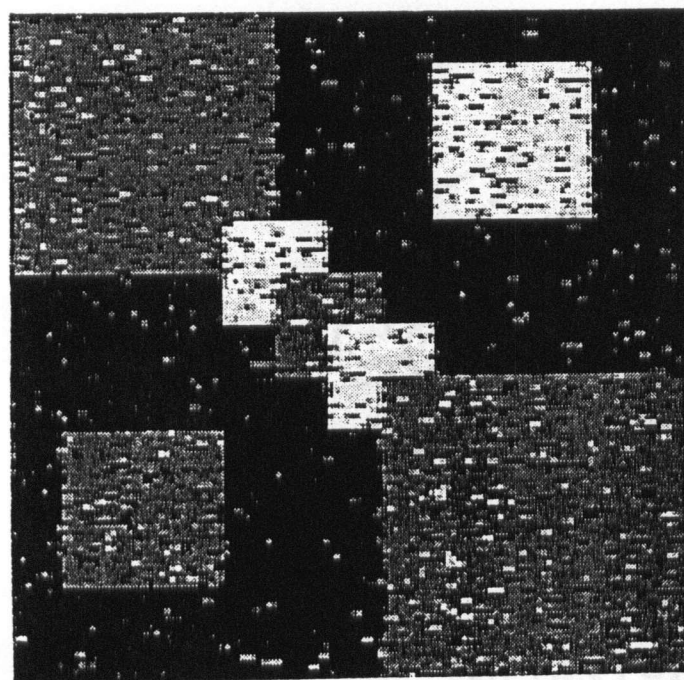(e) $\tilde{Q}_3(\delta)$ vs δ

(f) $\hat{Q}_3(\delta)$ vs δ
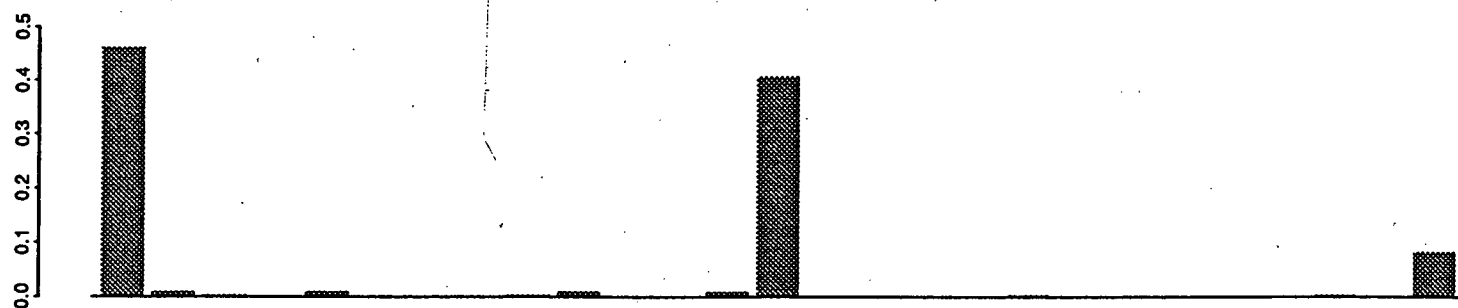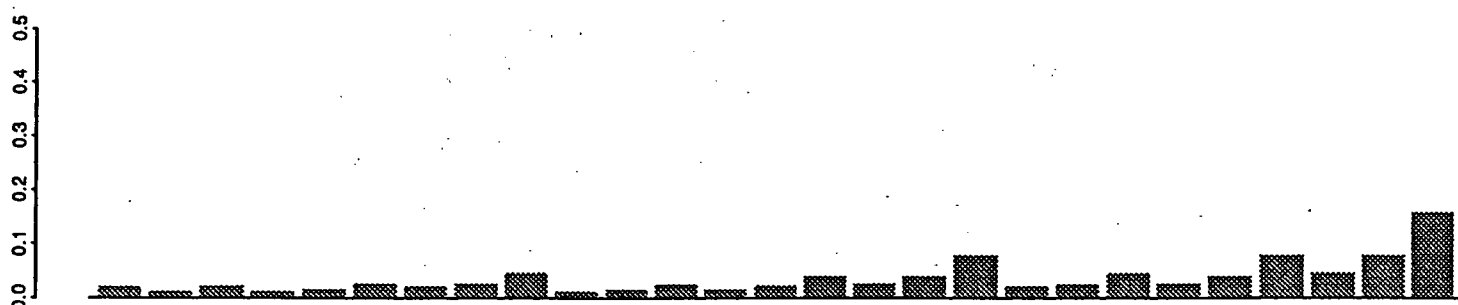
Figure 2.4: Pattern distribution with $r = 5$, and (i) $\varphi(\delta, y) = $ Indicator functions
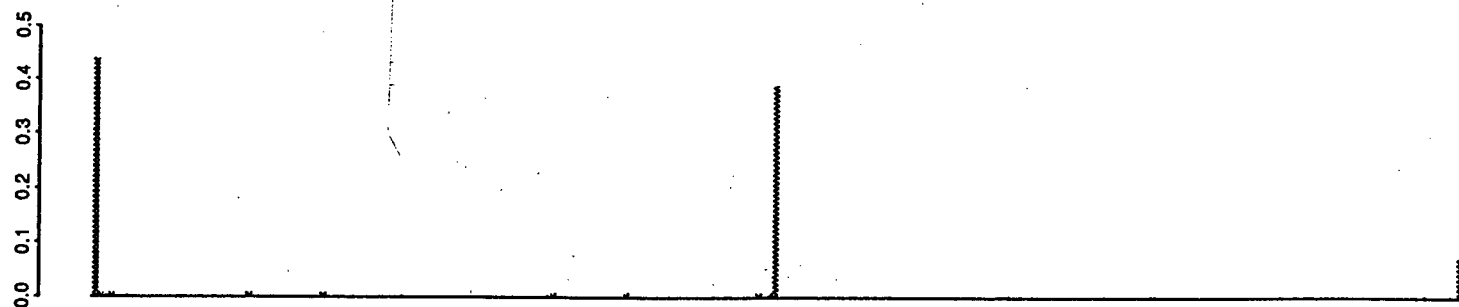
(a) $Q_5(\delta)$ vs δ
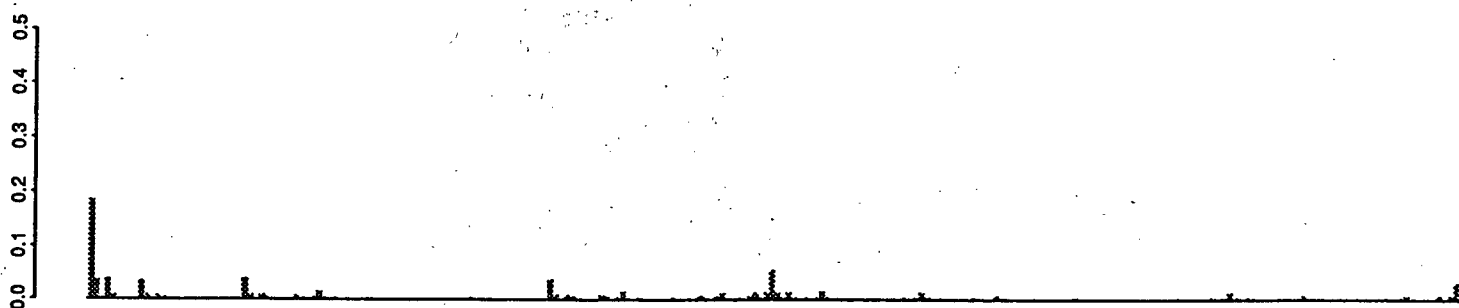
(b) $\tilde{Q}_5(\delta)$ vs δ
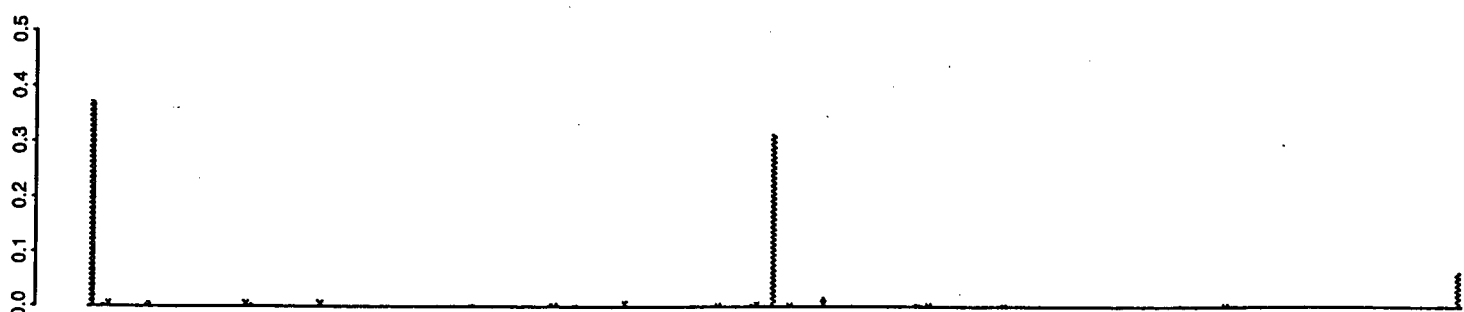
(c) $\hat{Q}_5(\delta)$ vs δ

Figure 2.4: Pattern distribution with $r = 5$, and (ii) $\varphi(\delta, y) =$ Power functions
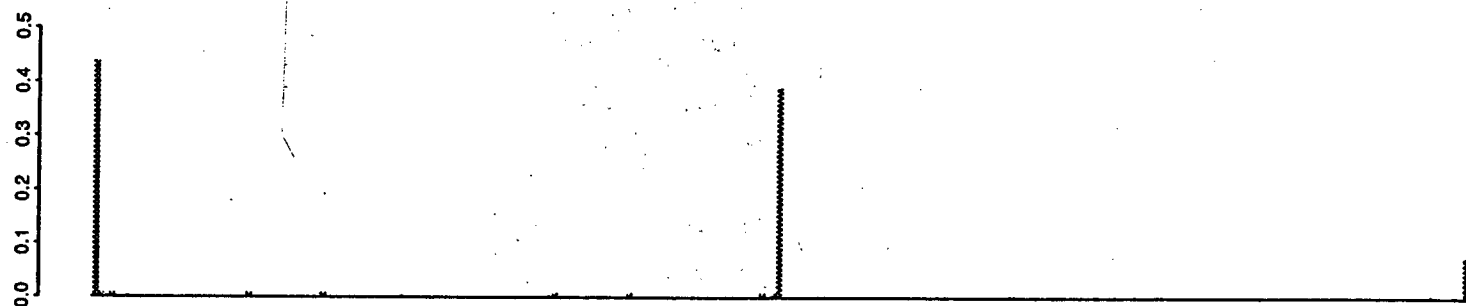
(d) $Q_5(\delta)$ vs δ
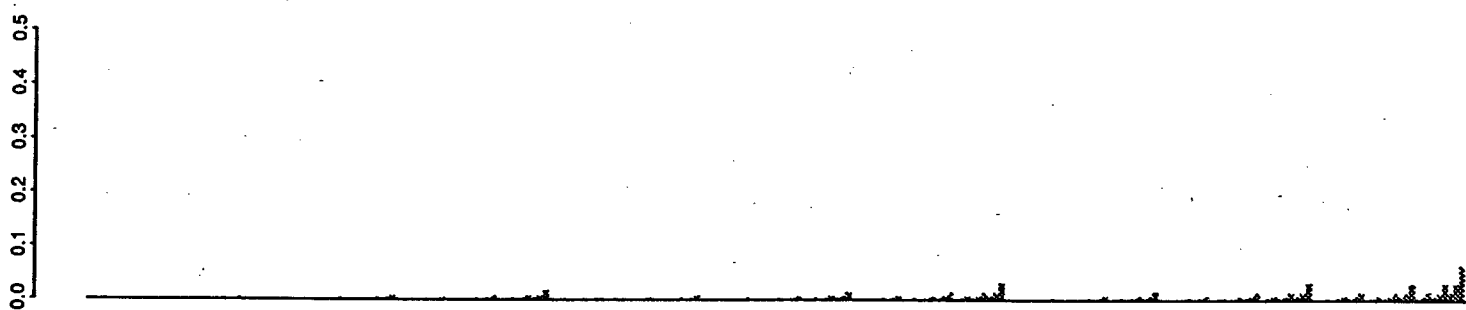
(e) $\tilde{Q}_5(\delta)$ vs δ
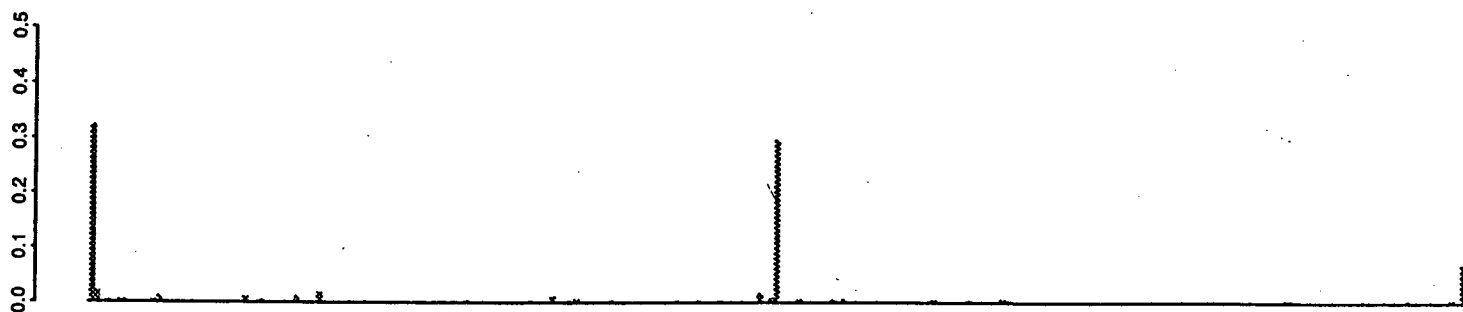
(f) $\hat{Q}_5(\delta)$ vs δ

Figure 2.5: Restored Image, $\varphi(\delta, y) =$ Indicator
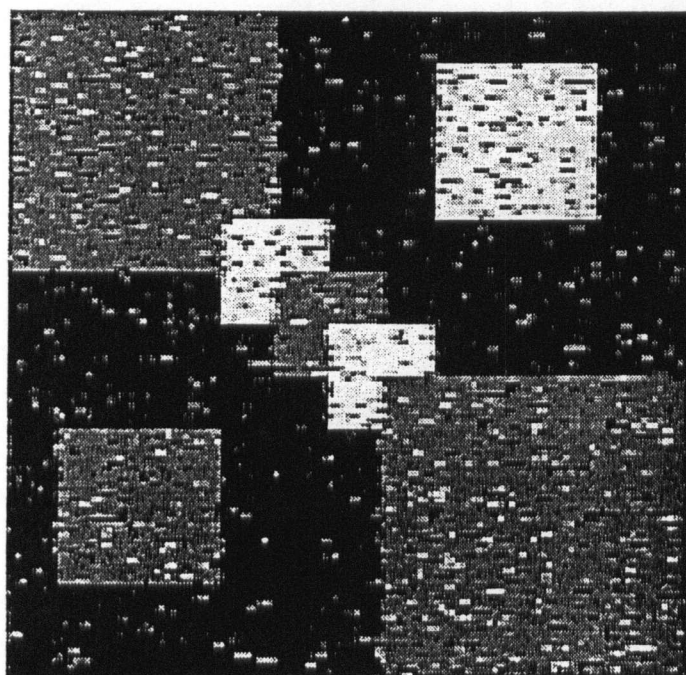

(a) $r = 1$

(b) $r = 2$

(c) $r = 3$

46
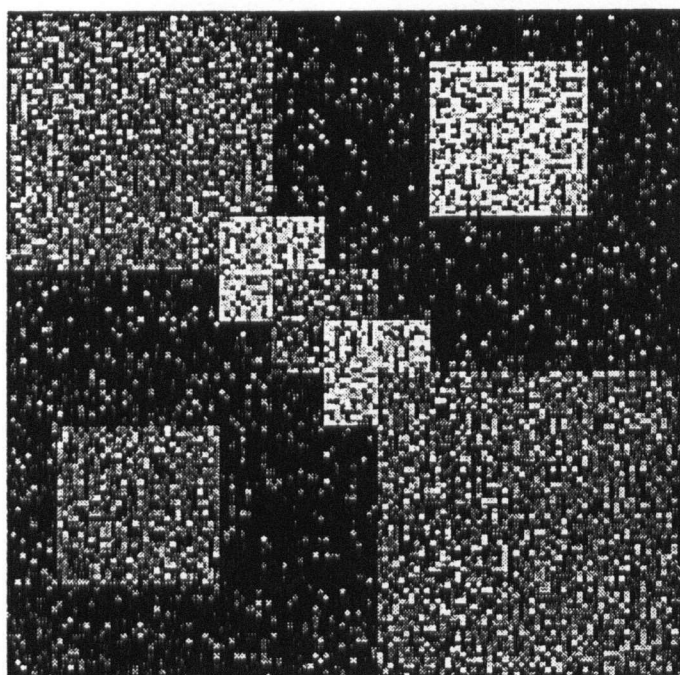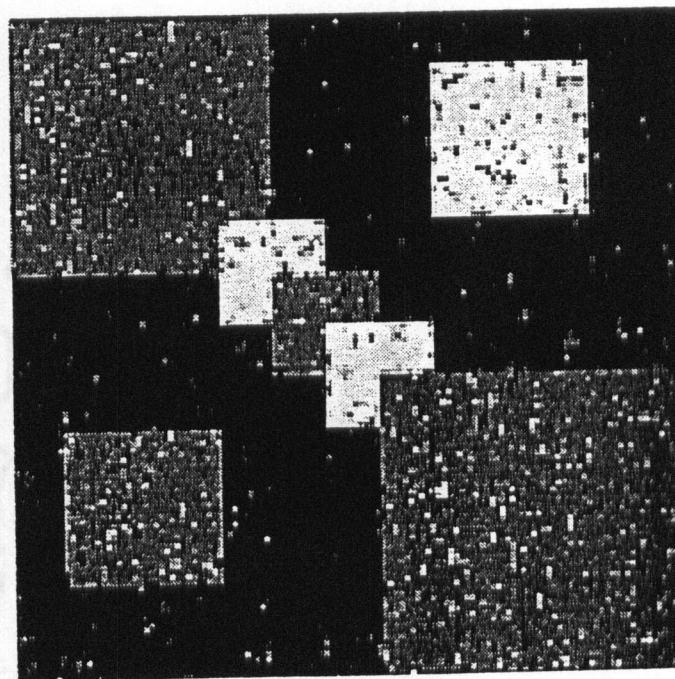
(d) $r = 5$

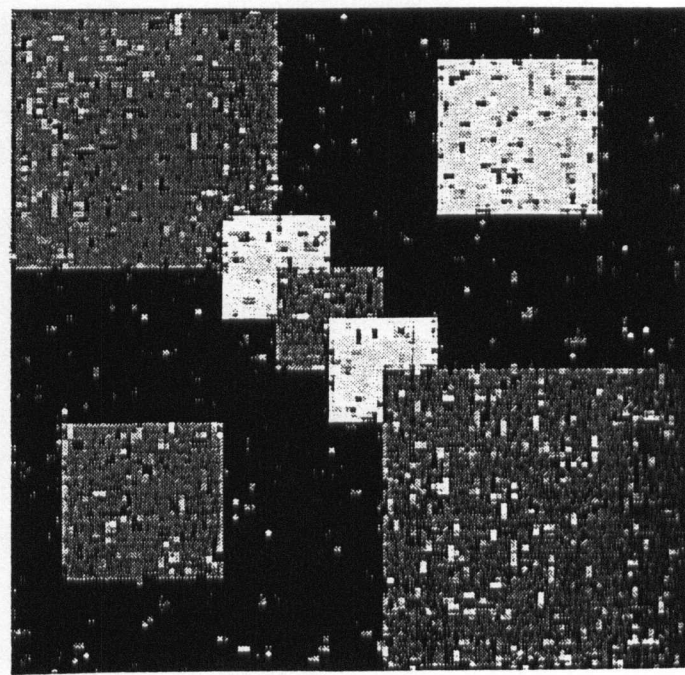Figure 2.5: Restored Image, $\varphi(\delta, y) =$ Power

(e) $r = 1$

(f) $r = 2$
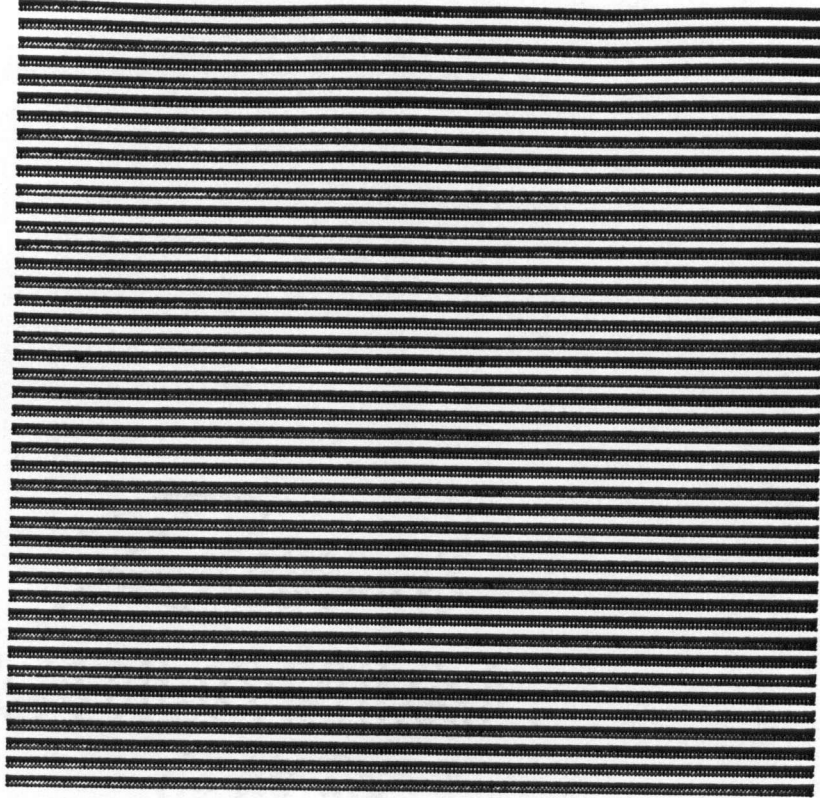
(g) $r = 3$

47
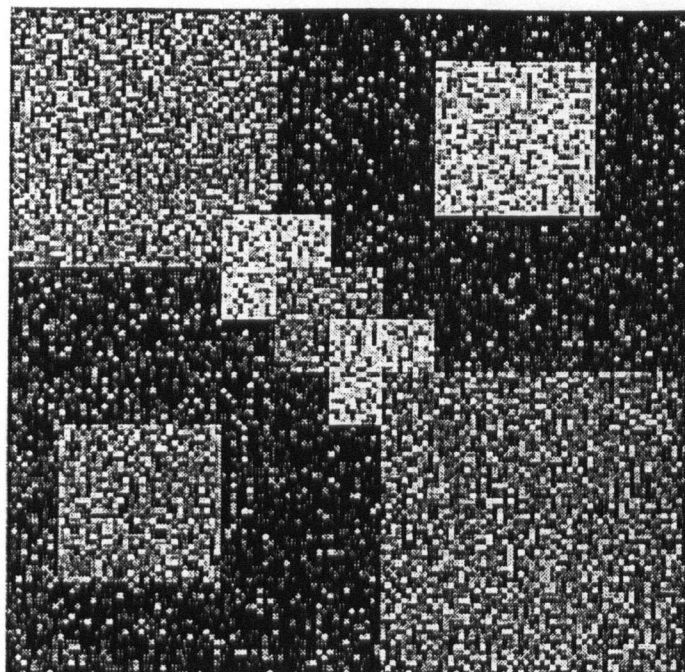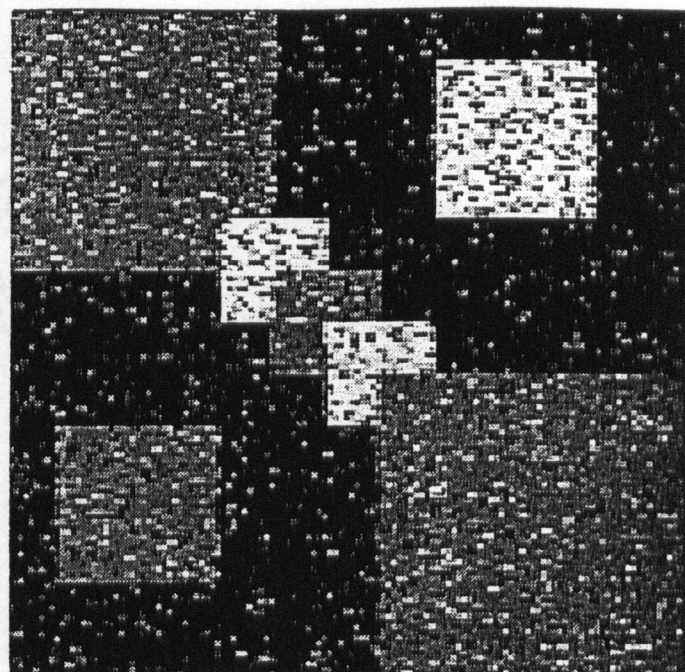
(h) $r = 5$

Figure 3.1: True Image (0-1-2 strips)

Figure 3.2: Noisy Image (0-1-2 strips)