4.0 Blocking Artifact

4.1 Boundary Characteristic

Transform coding, one of the most popular image compression techniques, generally divides the original image into subimages called blocks. Each block is transformed and the selected large coefficients are quantized and then stored/transmitted. One of these techniques' drawbacks is that the discontinuities at the block boundaries are visible. Such a discontinuity implies the appearance of high frequency components which can be eliminated by using lowpass filters. The advantage of using a lowpass filtering technique is that no additional information is needed, and, as a result, the bit rate is not increased. However, it results in blurred images.

Let us investigate the pixel intensity at the boundary of a transform coded image. An 512×512 original image is divided into 8×8 subimages before performing MBR coding with 16% and 24% MSE. Figure 4.1 shows sample boundary pixels of the original image and of its compressed representatives at different degrees of compression. The figure illustrates that the slopes formed by the pixels at the boundaries are higher than those of the neighboring pixels. The higher the compression ratio (more mean square error), the higher the slopes at the boundaries.

Most local constraints discussed in Chapter 3 perform well in the sense of MSE reduction; however, those constraints enforce nothing between blocks in order to remove the block discontinuity. Thus, information on individual block boundaries is necessary to smooth the entire image. So far, we know that the slopes at the boundaries in the compressed representations are normally higher than those in the original. Therefore, inter-block constraints should be used to remove the artifacts at the boundaries.

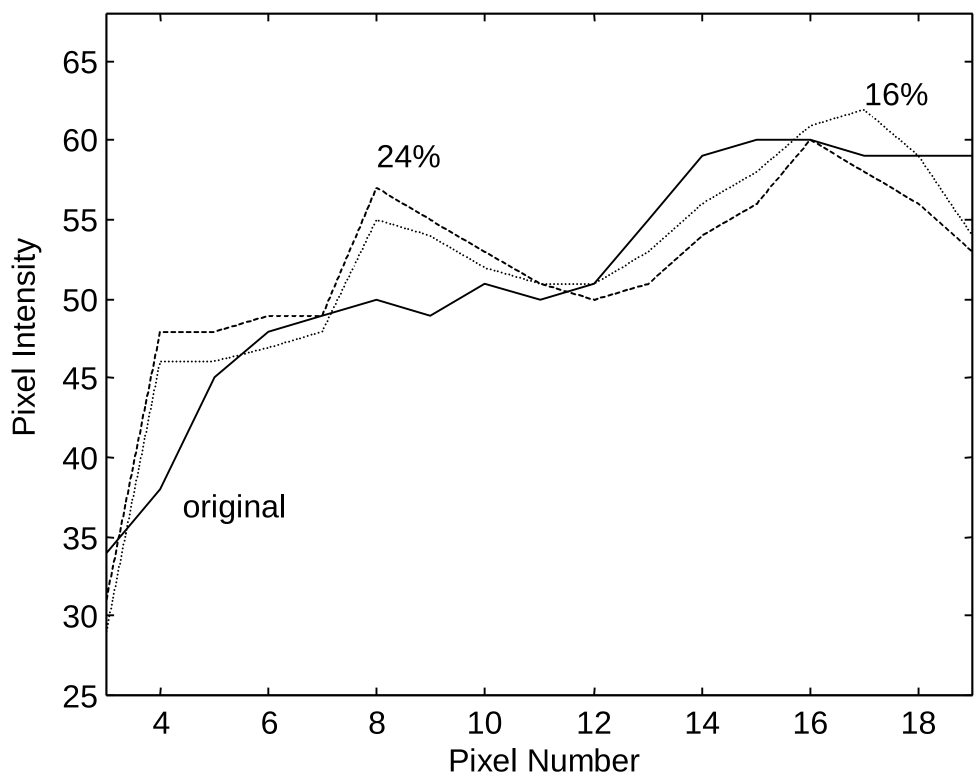

Figure 4.1 Pixel intensity at boundaries.

[7,8] and [15,16] form the boundaries.

4.2 Slope Constraint

Let $\mathbf{f} = [f_0, f_1, \ldots, f_{N-1}]^T$ and $\mathbf{g} = [g_0, g_1, \ldots, g_{N-1}]^T$ be N-dimensional vectors. The differential vector, $\mathbf{d}_{\mathrm{fg}} = [d_{fg,0}, d_{fg,1}, \ldots, dfg_{,N-1}]^T$ is defined by

$$\mathbf{d}_{\mathrm{fg}} = \mathbf{f} - \mathbf{g} \tag{4.1}$$

Assuming **f** and **g** represent adjacent column vectors of an image, d_{fg} can be viewed as a slope vector. The Absolute Slope constraint is now defined as

$$C_{\mathrm{slope}} = \{(\mathbf{f},\mathbf{g}): |f_i - g_i| \le s \text{ , for } i = 0,1,\ldots,N-1\} \tag{4.2}$$

where *s* is the maximum slope allowed by the constraint.

Let N-dimensional vectors x and y be in a set S which all absolute value of their elements $|x_i|$ and $|y_i|$ are less than or equal to s. For $0\le\mu\le1$, $\mu|x_i|+(1-\mu)|y_i| \le \mu s+(1-\mu)s = s$. Thus, a vector formed by $\mu x+(1-\mu)y$ is still in the set. Therefore, this set is convex. Consider x_p the point where all absolute values of its elements are s. x_p is a limit point of S since neighborhood of x_p contains a point of S distinct from x_p [Valentine 1964]. Because S contains x_p, its limit point, S is a closed set.

Equation (4.2) illustrates that C_{slope} constrains the absolute pixel value of the differential vector formed by **f** and **g**. Therefore, the element-wise slope projection applied to **f** and **g** is defined by

$$
P_{slope}(f_i, g_i) = \begin{cases} (f_i, g_i), & |f_i - g_i| < s \\ \left(\frac{f_i + g_i}{2} + \frac{s}{2}, \frac{f_i + g_i}{2} - \frac{s}{2}\right), & |f_i - g_i| \geq s, \quad and \quad f_i < g_i \\ \left(\frac{f_i + g_i}{2} - \frac{s}{2}, \frac{f_i + g_i}{2} + \frac{s}{2}\right), & |f_i - g_i| \geq s, \quad and \quad f_i \geq g_i \end{cases} \tag{4.3}
$$

The projection does nothing to f_i and g_i if the absolute value of $d_{fg,i}$ is less than the specific threshold s. If the absolute value of $d_{fg,i}$ is greater than s, the projection forces the slope to s without changing the sign of the $d_{fg,i}$.

The threshold value s might be transmitted as side information for each boundary. In the extreme case, s is set to be zero. The projection operator defined by (4.3) becomes

$$
P_{slope2}(f_i, g_i) = \left(\frac{f_i + g_i}{2}, \frac{f_i + g_i}{2}\right) \tag{4.4}
$$

The operator just averages the boundary columns vectors. Consequently, both boundary column vectors are the same.

4.3 Norm-of-Slope Constraint

Let **f** and **g** be vectors at the boundaries of an original image and **f**¢and **g**¢be the vectors at the same location of a transform coded version of the image. Due to the block by block transformation, the slopes at the boundaries of the transform coded image increase.

Therefore, the norm of the different vector of the transformed image, $\|\mathbf{f}' - \mathbf{g}'\|$, is also larger than the norm of the original, $\|\mathbf{f} - \mathbf{g}\|$. As a consequence, a constraint named the Norm-of-Slope constraint is introduced.

4.3.1 Definition

Let $\mathbf{d}_{fg}$ be the difference between boundary vectors **f** and **g**. The norm of $\mathbf{d}_{fg}$, $\|\mathbf{d}_{fg}\|$, is evaluated by

$$\|\mathbf{d}_{fg}\| \quad = \left[\sum_{i=0}^{N-1} d_{fg,i}{}^{2} \right]^{\frac{1}{2}} \qquad (4.5)$$

The norm-of-slope constraint is defined by

$$C_{NS} \quad = \{(\mathbf{f},\mathbf{g}) : \|\mathbf{d}_{fg}\| \leq E\} \qquad (4.6)$$

where E is a predetermined scalar, for example based on the original image.

Consider an N×N image, **q**. The image is divided into M×M subimages (blocks). N is, in general, a multiple of M. Each block is then transformed. The reconstructed images based on the partial transform coefficient information result in blocking artifacts. These blocking artifacts are noticeable at the boundaries of the blocks. Let us index row and column vector by 0 to N-1. Rows pM-1 and pM, and columns kM-1 and kM, where p,k=1,2,...,N/M-1 , form boundaries where the blocking artifacts show up.

To utilize the norm-of-slope constraint, let us define **f** and **g**. Based on the location where the blocking artifact occurs, we have some freedom to define the vectors **f** and **g**. For example, if the constraint is applied locally to each block, the M-dimensional vector **f** might be the last column of the block and **g** is then the first column of the adjacent block.

In stead on constraining one block at a time, we can constrain two or more adjacent blocks simultaneously. The length of vector **f** and **g** is then a multiple of the block size. For instance, if two blocks are to be constrained simultaneously, i.e., 2M-dimensional vectors are to be constrained, **f** is from merging two vertically adjacent blocks and **g** is then the next column with the same size. Figure 4.2 depicts the various choices to define the vectors **f** and **g**. The M-dimensional vector $\mathbf{f_M}$ in Figure 4.2 is defined to constrain block by block and the 2M-dimensional vector $\mathbf{f_{2M}}$ is for constraining two adjacent blocks simultaneously.

For N-dimensional vectors **f** and **g**, we can define the entire row or column vectors that form the "inside-image" blocking boundaries as the vectors to be constrained. Let q_i where i = 0,1,...,N-1 be the column vector of the N×N image.

$$\mathbf{q} = \{q_0, q_1, q_2, ..., q_{N-1}\} \tag{4.7}$$

Thus, column q_{kM-1} and q_{kM}, where k = 1,2,...,N/M-1, form the inside-image blocking boundary. There are then (N/M-1) column pairs and also (N/M-1) row pairs that form the boundaries. In this particular case where $f = q_{kM-1}$ and $g = q_M$, the constraint defined by (4.6) becomes

$$C_{NS} \quad = \{(q_{kM-1},q_{kM}): \| q_{kM-1}-q_{kM}\| \le E\} \tag{4.8}$$

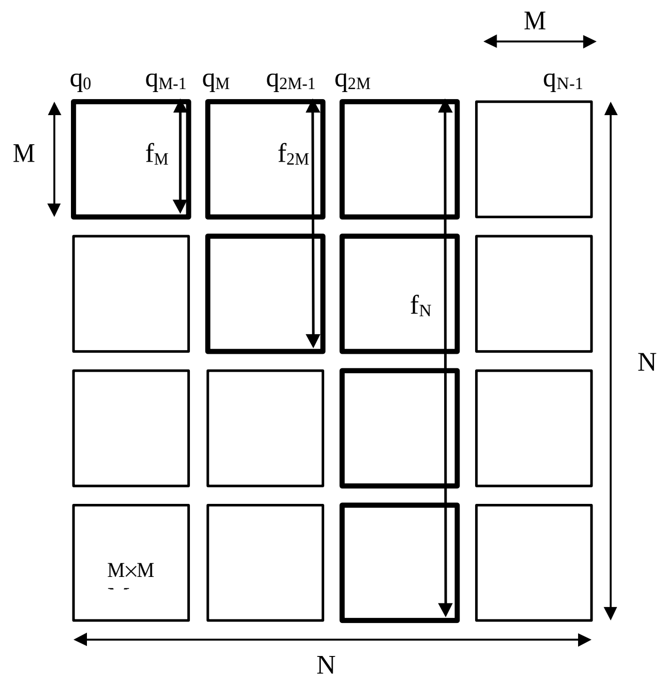

Figure 4.2 Choices of vectors for C_{NS}.

Yang [Yang 1993] defines a vector Q, which is equivalent to the vector $\mathbf{d}_{fg}$, as the difference between the columns at the internal block boundaries.

$$Q = \begin{bmatrix} q_{M\text{-}1} - q_M \\ q_{2M\text{-}1} - q_{2M} \\ q_{3M\text{-}1} - q_{3M} \\ \vdots \\ q_{N\text{-}M\text{-}1} - q_{N\text{-}M} \end{bmatrix} \tag{4.9}$$

Q is an N×(N/M-1)-dimensional vector for which the norm can be computed by

$$\|Q\| = \left[\sum_{i=1}^{\frac{N}{M}-1} \|q_{iM\text{-}1} - q_{iM}\|^2 \right]^{\frac{1}{2}} \tag{4.10}$$

Let us take an example in which **q** is a 512×512 image for which the subimage block size is 16×16. The total number of columns that form internal boundaries is 31. The vector Q can be illustrated by

$$Q = \begin{bmatrix} q_{15} - q_{16} \\ q_{31} - q_{32} \\ q_{47} - q_{48} \\ \vdots \\ q_{495} - q_{496} \end{bmatrix} \tag{4.11}$$

and the norm is then

$$\|Q\| = \left[\sum_{i=1}^{31} \|q_{16i\text{-}1} - q_{16i}\|^2 \right]^{\frac{1}{2}} \tag{4.12}$$

The constraint on the column block boundaries is defined to force $||Q||$ to be less than the predetermined scalar value E. In much the same way this constraint can be applied to row boundaries. The variations between the rows at block boundaries are then captured and used to smooth those boundaries in reconstruction.

As illustrated, there are many ways to define the **f** and **g**. If the exact value of E needs to be stored or transmitted, the greater the length of vectors **f** and **g,** the less additional information required. However, the greater length might effect the performance of the constraint. The comparison will be illustrated in Chapter 6.

4.3.2 Projection

Define a projection operator corresponding to the norm-of-slope constraint, $P_{NS}(f,g) = (\tilde{f}, \tilde{g})$. The projection is implemented [Yang 1993] as follows

$$\tilde{f} = \alpha \cdot f + (1-\alpha) \cdot g \tag{4.13a}$$

$$\tilde{g} = (1-\alpha) \cdot f + \alpha \cdot g \tag{4.13b}$$

where $\alpha = \frac{1}{2}\left[\frac{E}{||f - g||} + 1\right]$.

The exact value of E can be easily provided at the transmitter end. The exact E is the best for the constraint but will cost some extra information. It is ideal if a good-enough estimated value of E can be determined from the distorted signal. That means there is no extra information needed for this constraint.

4.4 Estimate of the Scalar E

Since image signals have high correlation among the neighboring signals, any slopes formed by the difference of two adjacent pixels in the neighborhood are not much different. The slopes at the boundaries of the distorted image are generally greater than of the original. However, the slopes in the distorted boundary neighborhood do not much differ from the original neighboring slopes as much as at the boundaries. Consequently, the scalar E could be reasonably estimated from the slopes of the boundary neighborhood of the distorted signal.

Consider column vector q_i. The blocking artifact boundaries are at q_{kM-1} and q_{kM} where k = 1,2,...,N/M-1. $\{q_{kM-M/2}, q_{kM-M/2+1}, \ldots, q_{kM-2}\}$ and $\{q_{kM+1}, q_{kM+2}, \ldots, q_{kM+M/2-1}\}$ are the neighboring column vectors of the boundary column vectors q_{kM-1} and q_{kM}. An estimate of real E, E′, can be written as

$$E'_{kM} = \frac{1}{M-2} \sum_{\substack{i=kM-\frac{M}{2}+1 \\ i \neq kM}}^{kM+\frac{M}{2}-1} \| q_{i-1} - q_i \| \tag{4.14}$$

Let us take an example for M = 8 and k = 1. E'_8 is determined by

$$E'_8 = \frac{1}{6} \sum_{\substack{i=5 \\ i \neq 8}}^{11} \| q_{i-1} - q_i \| \tag{4.15}$$

E'_8 is a measure of the average of the norm of difference vectors from column 4 to column 11 with omitting $\|q_7-q_8\|$ which is the boundary.