COMMITTEE T1

CONTRIBUTION

Document Number: T1A1.5/95-104

****

STANDARDS PROJECT:	All Digital Video Performance Standards Projects
TITLE:	Objective Measures for Detecting Digital Tiling
ISSUE ADDRESSED:	Objective Performance Measurements for Digital Video Systems
SOURCE:	NTIA/ITS - Dwight Melcher, Stephen Wolf
DATE:	January 9, 1995
DISTRIBUTION TO:	T1A1.5
KEYWORDS:	Video Performance Testing, Objective Video Quality

****

Objective Measures for Detecting Digital Tiling

Dwight Melcher, Stephen Wolf

1.0 Introduction

This contribution describes an extension of the objective measurements of video quality (presented in [1]) that specifically detects the "tiling" or "blocking" artifacts produced by many digital coding systems.

1.1 General Observations

The development of extended measures to specifically detect blocking is motivated by the following observations:

- Current spatial measures detect both blocking and blurring combined. However, from informal analysis of subjective data, it is clear that viewers subjectively rate blocking and blurring differently. Anecdotal evidence suggests viewers find tiling more objectionable than blurring. A measure that separates blocking and blurring into two components is an advantage, since models using separate components could weight the two differently, resulting in higher correlation between the objective and subjective measures.
- The spatial information measurements currently used are based on the gradient image produced by Sobel filtering each frame of video [1]. The standard deviation of the magnitude of the pixels in the gradient image is used as a feature. Features are then compared in the source and degraded images to form a parameter. The parameter is then used in a model in order to quantify the perceived degradation between the two input images.
- Current measures compress each frame in the video stream into two numbers. This is a highly desirable property of the feature extraction process, since the feature data can be easily transmitted between test instruments over a low bandwidth serial link. The down side of the high level of compression is that too much information is being carried by these features, which makes it difficult to separate components of the feature that might be more useful if weighted differently.

1.2 Extensions to Spatial Features

This section describes additional low-bandwidth features that can be used for measuring tiling artifacts. These features are an extension of the basic spatial information (SI) features described in [1] and [4].

In order to separate out information that is relevant to measuring tiling effects, several attributes of the Sobel gradient operators can be exploited. From vector analysis, it is known that the gradient vector points in the direction of maximum rate of change of a function *f* at (*h*,*v*). The gradient is the vector:

$$\overline{\nabla}f = \begin{bmatrix} SI_h \\ SI_v \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial h} \\ \frac{\partial f}{\partial v} \end{bmatrix}$$

Where *h* and *v* are the horizontal and vertical directions, respectively. Also, SI_h and SI_v are the partial derivatives of *f* with respect to *h* and *v*, respectively. If *f* is an image, SI_h and SI_v can be computed using Sobel filtering as described in [1]. The current spatial features use an approximation of the magnitude of the gradient vector to characterize the edge content of an image:

$$\nabla f \approx |SI_h| + |SI_v|$$

However, the direction of the gradient vector is also an important quantity. The direction of the gradient vector of *f* at (*h*,*v*) is given by:

$$\theta(h, v) = \operatorname{atan}\frac{SI_v}{SI_h}$$

In summary, the gradient image contains both edge magnitude and angle information, which is derived from the SI_h and SI_v components of the Sobel filter. The edge magnitude information is used in the current set of spatial features to characterize the edge content of a video image. The remainder of this contribution is devoted to an extension of the current features that takes advantage of the angle information present in the gradient image in order to detect specific artifacts such as digital tiling.

2.0 Gradient Magnitude and Angle Histogram

In order to make explanation of the new features more clear, this section presents a method of creating a two dimensional histogram that makes it possible to visualize both magnitude and angle information at once.

While computing the gradient image, one can accumulate a 2-dimensional histogram. Each bin in the histogram is identified by an *h* and *v* coordinate. For a pixel in the image, the two components of the gradient are computed, yielding SI_h and SI_v at that point. SI_h and SI_v are then used as the coordinates of a bin in the histogram. The bin at coordinate (SI_h, SI_v) is incremented. This process is repeated for all pixels in the image. The result is a histogram that shows both the distributions of *intensity* and *direction* of the edges in the image.

The histogram can be displayed as a normal X-Y plot. By assigning different values of color or intensity to the values of the bins, one can see how the distribution of angles in the gradient image varies. Figure 1 shows an example image and its gradient magnitude and angle histogram. Lighter pixels in the plot indicate bins with higher values.

One can view the histogram as an extension of the SI feature [1] into a 2-dimensional function that is computed from an image. Since it is convenient to talk about the histogram in terms of angles and radii, later discussions are done in terms of polar coordinates. We define a function SIH(r, θ) where r is the magnitude of the gradient, and θ is the angle of the gradient and SIH(r,θ) is the number of pixels in the gradient image whose gradient radius and angle is r and θ, respectively. In all of the equations below, θ is assumed to be between 0 and 2π.

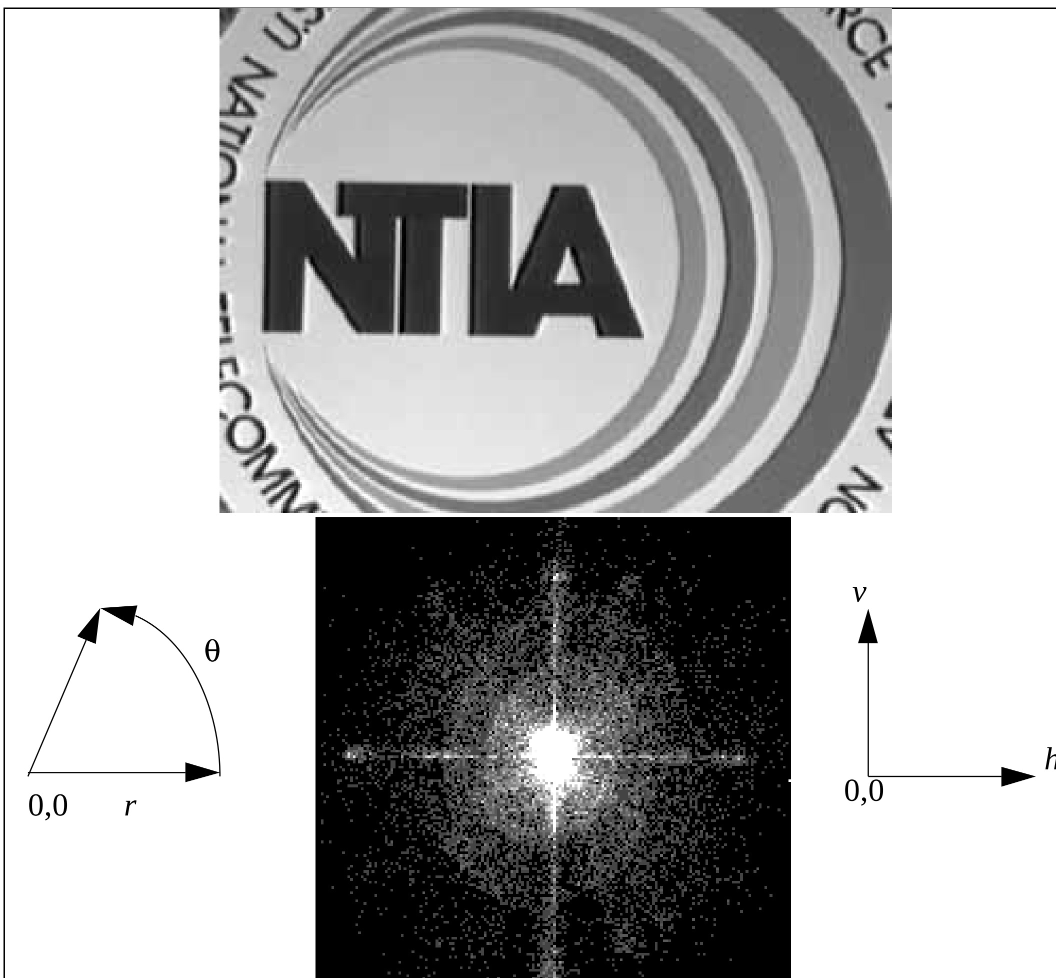

FIGURE 1. An Image and its Gradient Magnitude and Angle Histogram, SIH(r,θ)

2.1 Features extracted from SIH(r,θ)

This section describes features that can be computed from SIH(r, θ) that are useful for detecting digital tiling.

The basic method for extracting the features is to select various ranges for the values of r and θ, then compute some simple statistics from SIH(r, θ) based on these ranges. An area of particular interest is that containing the principal axes of the histogram, where horizontal and vertical edges produce a response.[1]

Tiling artifacts appear as faint horizontal and vertical edges added to an image. Accordingly, tiling artifacts tend to increase the value of SIH(r, θ) where θ lies along one of the principal axes (e.g. where $\theta = k\pi / 2$, $k = 0, 1, 2, 3$). Figure 2 and Figure 3 show how SIH(r, θ) changes for an image exhibiting tiling. The image in Figure 3 has been chosen by the T1A1.5 committee as an example of the “blocking” or “tiling” impairment [3]. The important thing to note about the two histograms is that the histogram for the tiling-impaired image has a distinctly more “plus-like” shape, indicating that there is a clustering of points around the horizontal and vertical axes. This observation is the basis for the formulation of the features given below.

A feature whose value increases as the number or sharpness of horizontal and vertical edges increase is given as:

$$g_{hv} = \frac{1}{p} \cdot \sum_{r} \sum_{\theta} SIH(r, \theta) \cdot r$$

$$0 < c_a \leq r \leq c_b \qquad \theta = \frac{k\pi}{2} \ (k = 0, 1, 2, 3)$$

Where r and θ are as defined above and c_a and c_b are clipping limits, and p is the number of pixels in the image (*not* the number of pixels in the summation). This feature is a weighted sum of the bins that accumulate horizontal and vertical edges in the gradient image. The lower clipping limit is used to restrict the computation to a region where the approximations of angle are more accurate (note that since the Sobel gradient filter produces only integer approximations of the gradient, as r becomes small the angle calculation becomes very coarse, making it difficult to use these low level values to accurately characterize the line orientations of the image.) The r clipping limits can also be used to “focus” the measurement to include only those edges of a desired intensity. Empirical evidence indicates the value for c_a should be from 5 to 20, depending on the source material

1. There are many other parts of the histogram that are of potential interest. One such part is the main diagonals (45, 135, 225, 315 degrees), where digital impulse noise is accumulated. Such noise is produced by pixel dropouts or pixel errors. This paper is concerned mostly with horizontal and vertical areas of the histogram and doesn’t explore these other ideas.

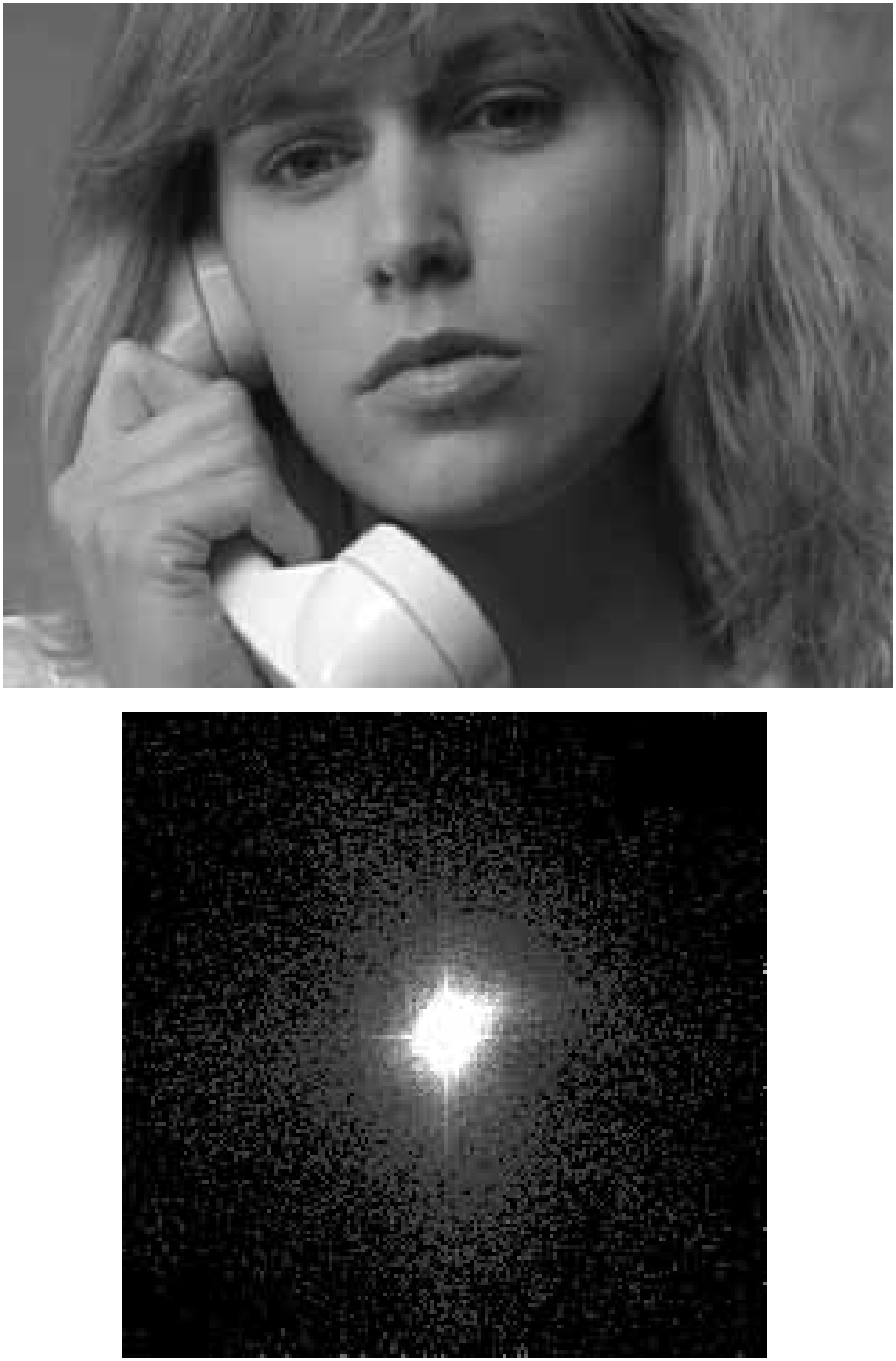

FIGURE 2. Source Image (NULL impairment) and Gradient Angle Histogram. The image is taken from the draft ANSI standard test scenes [2].

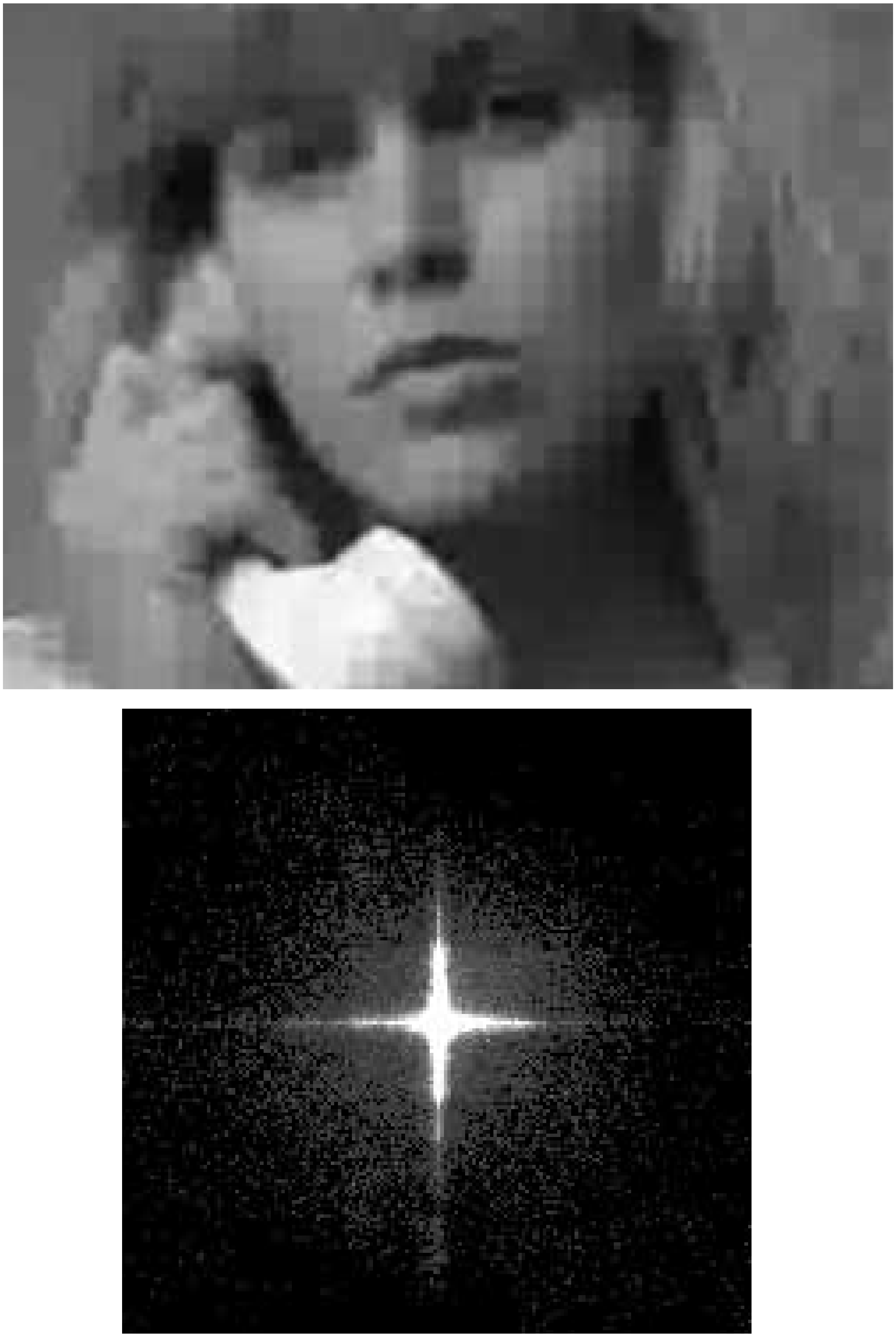

FIGURE 3. The image from Figure 2 after passing though a VTC system. Note that the image exhibits tiling and blurring (this image was selected by Committee T1A1.5 as an example of the tiling impairment [3]) The gradient angle histogram shows a distinctive “plus” shape compared to the histogram in Figure 2.

being processed. In practice, it is useful to include a small wedge of a few degrees around horizontal and vertical axis. This allows for slight variations in the angle calculations that are caused by noise.

It should be noted that the g_{hv} feature characterizes *all* of the horizontal and vertical edges in an image, natural as well as those produced by tiling. This feature also tends to decrease in value as the image becomes blurry and increase in value as the image suffers tiling (just as the original SI measurement [1]).

In order to produce a parameter that more clearly separates blurring from tiling, a feature that characterizes the edge content of the image *without* the inclusion of horizontal and vertical edges is used:

$$g_{hv'} = \frac{1}{p} \cdot \sum_{r} \sum_{\theta} SIH(r, \theta) \cdot r$$

$$0 < c_a \le r \le c_b \qquad \theta \ne \frac{k\pi}{4},\ k = 0 \ldots 7$$

The $g_{hv'}$ feature is computed over the entire histogram, excluding the horizontal and vertical angles. It also excludes the exact diagonal angles, since these angles tend to accumulate low level impulse noise as described above.

2.2 A Composite Ratio Feature

A composite feature can be formulated that uses the g_{hv} and $g_{hv'}$ features in order to enhance information related to tiling in the image. Throughout the discussion below, HV refers to “horizontal and vertical”. One possible feature, denoted by *R*, is given by:

$$R = \frac{g_{hv} + \varepsilon}{g_{hv'} + \varepsilon}$$

Where ε is a small constant (usually on the order of 1) that is used to stabilize the value of *R* as $g_{hv'}$ becomes small. The *R* feature is the ratio of the HV edge content of the image and the non-HV edge content. As the number of HV edges becomes high in relation to the non-HV edges in the image, the value of *R* increases (visually, an increase in R increases the “plus-like” appearance of the histogram, as in Figure 3.)

2.3 Combining the Features into Parameters

Three basic features have been described that can be used to characterize the horizontal and vertical edge content of an image (g_{hv}), the non-vertical and non-horizontal edges ($g_{hv'}$) and the relative amounts of HV vs. non-HV edges in an image (*R*).

These features alone cannot be used to measure the tiling content of an image. The features must be computed on the source and degraded images and compared in order to discover if tiling has been added to the degraded images.

The three basic features described in the previous section are computed frame-by-frame on the source and degraded video images. This produces a time history of the features for both the source and degraded sequences. However, for the purposes of the following discussion, only one frame of video each of source and degraded is considered. The frames are assumed to be corresponding frames from each stream (i.e. the frames are assumed to have been selected from time aligned video sequences.)

In order to compare the source and degraded video frame, the features are combined into a *parameter* using the form:

$$P = \frac{s - d}{s}$$

Where s denotes the feature computed from the source video frame, and d is the feature computed from the degraded video frame. The resulting value is called a parameter. In this case, the parameter measures the fractional difference between the source and degraded features. So, for example, if there is no change in the feature between the source and degraded images, the value of P would be 0, indicating no change. This form of parameter (the error divided by the source reference) has proven to be one of the most useful for quantifying video distortions [1].

One parameter that is useful for detecting tiling is as follows:

$$P_{hv1} = \frac{R_s - R_d}{R_s}$$

The P_{hv1} parameter becomes negative when the proportion of HV edges in the degraded image is greater than the proportion of HV edges in the source image. Note that in the case of digital tiling artifacts, faint HV edges are added to the degraded image, resulting in a more negative value for P_{hv1}.

Two more parameters of the same form can be computed from the given features:

$$P_{hv2} = \frac{g_{hvs} - g_{hvd}}{g_{hvs}} \qquad P_{hv2'} = \frac{g_{hv's} - g_{hv'd}}{g_{hv's}}$$

In this case, the P_{hv2} and $P_{hv2'}$ parameters measure separately the fractional change in the HV edges and the non-HV edges, respectively. This allows one to examine these changes between source and degraded images in isolation. These parameters exhibit responses to blurring that are similar to the response of the SI parameter presented in [1]. If a degraded image is blurred (and no extra edges are added to it) the values of these parameters will increase by roughly the same fraction, indicating a global reduction in edge energy in the degraded image. However, suppose the degraded image is blurred, but also has extra HV energy added in the form of digital tiling. Since $P_{hv2'}$ is not affected by HV edges, $P_{hv2'}$ retains the same value as in the first case. However, conceptually g_{hvd} will decrease by the same amount as $g_{hv'd}$, and then g_{hvd} will be increased by some amount that is proportional to the amount of edge energy added by the tiling. These observations are the motivation behind the next parameter, which is the difference between the $P_{hv2'}$ and P_{hv2} parameters :

$$P_{hv4} = P_{hv2'} - P_{hv2}$$

The P_{hv4} parameter detects the relative change in HV edges *vs.* the relative change in non-HV edges. The idea is that if both the HV and non-HV edge features change by the same fraction, some global degradation has occurred to the image (such as uniform blurring to all edges in the image, as mentioned above.) In this case, the P_{hv2} and $P_{hv2'}$ parameters will track, since they are presumably affected by the degradation in the same manner. This results in P_{hv4} taking on a value close to zero.

However, if the non-HV edge parameter ($P_{hv2'}$) stays constant, and the HV edge parameter (P_{hv2}) decreases (possibly becoming negative), the indication is that some extra HV edges are being added to the degraded output.[1] These extra HV edges are likely the result of tiling being added to the degraded image. Hence, P_{hv4} tends to increase as the amount of tiling increases in the degraded image.

Figure 4 through Figure 7 present a series of degraded images with varying degrees of tiling and blurring. Table 1, "Feature and Parameter Values for Images," on page 13 shows the computed values for the images in this paper. Note that for the images which exhibit tiling, the P_{hv1} and P_{hv4} parameters show a marked response (P_{hv1} strongly negative, and P_{hv4} strongly positive) compared to the images with little or no tiling. In all cases, the source image is the NULL impairment from Figure 2.

2.4 Conclusions

This contribution presented several objective measurements that can be used to detect digital tiling. A two-dimensional histogram of the gradient angle and magnitudes of an image can be created and used as the basis for extracting new features. Two basic features can be extracted from the histogram and these features can be combined in a variety of ways and compared in source and degraded video streams to characterize the tiling content of the degraded video. An example was presented that shows that the presented parameter values respond in the presence of tiling in the degraded video.

2.5 Acknowledgments

The authors wish to thank Ismail Dalgic at Stanford University for informative correspondence leading to the development of these measures. The authors would also like to thank

1. Note that P_{hv2} and $P_{hv2'}$ are *positive* when the degraded image has *less* edge energy (for example, when blurring is present). P_{hv2} and $P_{hv2'}$ are *negative* when there is *added* edge energy in the degraded image.

Margaret Pinson for her effort in implementing and testing these measures on the NTIA real-time video quality measurement system.

2.6 References

1. S. Wolf, M. Pinson, C. Jones, and A. Webster. “A Summary of Methods of Measurement for Objective Video Quality Parameters Based on the Sobel Filtered Image and the Motion Difference Image”. Committee T1 Contribution T1A1.5/93-152, November 1993.

2. Standards Committee T1A1 Letter Ballot. “Draft Proposed American National Standard - Digital Transport of Video Teleconferencing/Video Telephony Signals - Video Test Scenes for Subjective and Objective Performance Assessment.”, T1A1/LB94-19, December 1994.

3. Standards Committee T1A1, “Proposal for an ANSI Standard Specification of Video Performance Terms and Definitions.”, T1A1.5/95-108, January 1995.

4. Standards Committee T1A1, “Draft Proposed American National Standard - Digital Transport of Video Teleconferencing/Video Telephony Signals - Parameters for Objective Performance Assessment.”, T1A1.5/95-107, January 1995.

5. A. K. Jain, *Fundamentals of Image Processing*, Prentice-Hall, 1989.

6. I. Dalgic, personal correspondence, 1994.

7. R. C. Gonzalez, R. E. Woods, *Digital Image Processing*, Addison-Wesley, 1992.

FIGURE 4. Image compressed using JPEG with a quality level of 25 (CJPEG -Q 25) Notice that some minor blocking and contouring is visible.

FIGURE 5. Image compressed using JPEG with a quality of 5 (CJPEG -Q 5). Note that tiling artifacts are evident.

FIGURE 6. Image compressed using MPEG1. Compression parameters were set to give approximately 400Kbps on a 256x243 grayscale video stream with no interframe coding.

FIGURE 7. Image blurred by several passes of a low-pass filter.

TABLE 1. Feature and Parameter Values for Images

Degradation	g_{hv}	$g_{hv'}$	R	P_{hv1}	P_{hv2}	$P_{hv2'}$	P_{hv4}
NULL	5.5	30.6	0.19	n/a	n/a	n/a	n/a
T1A1	8.2	16.3	0.52	-1.7	-0.50	0.47	0.95
CJPEG -Q 25	7.8	28.5	0.29	-0.48	-0.41	0.07	0.48
CJPEG -Q 5	12.4	22.8	0.56	-1.9	-1.3	0.25	1.5
MPEG1-400kbps	10.0	16.4	0.62	-2.2	-0.82	0.46	1.3
Blurry	3.21	19.6	0.18	0.04	0.42	0.36	-0.06

Notes: Calculations were performed using the following values. All angles were folded into the 0° to 45° area of the SIH histogram. The g_{hv} and $g_{hv'}$ features are scaled by the number of pixels in the image (which is identical for all images.)

Edge Filter	Sobel
c_a	10
c_b	max
HV $\Delta\theta$	0° to 5° (for computation of g_{hv})
non-HV $\Delta\theta$	6° to 40° (for computation of $g_{hv'}$)
ε	0.5