THE DISCRETE REPRESENTATION OF SPATIALLY CONTINUOUS IMAGES

by

John Newland Ratzel

S.B., Massachusetts Institute of Technology

1975

S.M., E.E., Massachusetts Institute of Technology

1977

Submitted in Partial Fulfillment

of the Requirements for the

Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

August, 1980

Signature of Author

Department of Electrical Engineering and Computer Science

Certified by ______ Thesis Supervisor

Accepted by ______ Chairman, Department Committee

ARCHIVES
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY
MAR 23 1981
LIBRARIES - 1 -


THE DISCRETE REPRESENTATION OF SPATIALLY CONTINUOUS IMAGES

by

John Newland Ratzel

Submitted to the Department of
Electrical Engineering and Computer Science
in August, 1980 in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy.

ABSTRACT

Recent years have witnessed a vast amount of work in image processing, much of it with very impressive results. However, the majority of coding, enhancement, and analysis algorithms which have been developed assume digitized input and output, without much regard for the conversions to and from the spatially continuous domain. To a large degree, the successful operation of these procedures depends on how well these conversions have been made or appear to have been made to a human observer. It is this problem, primarily a psychovisual one, which has been addressed in this thesis.

Any notion of image quality can only be discussed within a very limited context. That which has been used herein is pulse code modulation, PCM, in which the digitized image representation is obtained by prefiltering the spatially continuous input, then subsampling it on a constant raster. A spatially continuous output is obtained by interpolation via linear filtering. The choice of PCM was due to its simplicity and utility, as evidenced by its widespread use.

Many tradeoffs arise during the three phases of PCM, for example, blur versus aliasing or smoothness versus sharpness. Several of these were investigated in a series of three experiments. The first experiment examined the tradeoff between space-width and frequency-width which arises in both the PCM prefiltering and postfiltering operations. Gaussian low pass filters were used for bandlimiting and interpolation, and a metric which characterized their frequency-width was taken as the percentage of filter energy within the passband implied by the subsampling frequency. Subjective picture quality was measured as a function of this metric, and was maximized at values of 60% and 75% for the bandlimiting and interpolating filters,
respectively. In the space domain, these numbers correspond to prefilter and postfilter widths of .30 and .375 subsample spacings. The second experiment examined the effects of a variety of interpolation filters, each of which traded off smoothness, sharpness, and isotropy to different degrees. The single lobe sharpened gaussian was found to maximize output quality given a prefiltered and subsampled input. Such a filter is a member of a family of filters which can be constructed by presharpening an all-positive basis interpolating function -- a 2N+1 point presharpening filter resulting in a net impulse response with N lobes. Any resulting single lobe function has the property that it is similar to the types of responses found naturally in the human visual system. The final experiment investigated the subjective effects resulting from the choice of sampling raster, hexagonal versus rectangular. When the interpolation scheme did not have a zero at the subsampling frequency, hexagonal subsampling performed significantly better than rectangular. However, this was not necessarily the case if the interpolation scheme did have a zero at the subsampling frequency, particularly for very structured images, for example, text or test patterns.

Two basic pictures were used in the experiments. The first was the well-known cameraman, representative of the general class of continuous tone images. The second test picture was a combination resolution wedge and text segment, representative of line art images. All of the work was performed on the MIT Cognitive Information Processing Group's Image Processing System (IPS), a dedicated, small-scale computer based picture processing facility.

Name and Title of Thesis Supervisor: Donald E. Troxel
Associate Professor Electrical Engineering
ACKNOWLEDGMENTS

I would like to acknowledge Professors Troxel, Schreiber, and McClellan for their suggestions and criticism throughout this work. As advisors, mentors, and friends, they have made my experience as a graduate student in the lab rewarding in both an academic and real-world sense.

Then there are the CIPG Imps. You couldn't hope to find a nicer group of people, and I owe something to all of them. Let me mention in particular Charlie Lynn (CLYNN), who first indoctrinated me into the wonderful world of IPS way back in the beginning, Mike McIlrath (MBM), who has done much the same with our unix system, and Sam Goldwasser (SG) and Len Picard (LEN), principal co-conspirators in our adventures of the past five years.

Finally, this might never have been completed without the tremendous support and encouragement of my family, for whom I am the most grateful. Marsha, thank you. hhh!
# TABLE OF CONTENTS

I. Introduction .......................................................... 8
II. A Model for Brightness Perception .............................. 14
III. A Characterization of the Spatial Coding Process
    A. The Digital Coding Problem .................................. 26
    B. PCM Picture Coding ........................................... 31
    C. Tradeoffs in PCM Picture Coding .......................... 36
IV. Experimental Investigation of
    The Space-width/Frequency-width Tradeoff .................. 47
V. Experimental Investigation of
    Interpolation Tradeoffs ......................................... 65
VI. Experimental Investigation of
    Hexagonal vs. Rectangular Sampling .......................... 99
VII. Discussion and Conclusions ..................................... 114
VIII. References ......................................................... 126
Appendix A. Tone Scale Calibration Procedure .................. 129
Appendix B. Non-Standard IPS and Unix Files ................... 134
<table>
<thead>
<tr>
<th>Figure</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>General Structure of Image Processing System</td>
<td>9</td>
</tr>
<tr>
<td>2.1</td>
<td>Simultaneous Contrast Phenomenon</td>
<td>15</td>
</tr>
<tr>
<td>2.2</td>
<td>Mach Band Phenomenon</td>
<td>17</td>
</tr>
<tr>
<td>2.3</td>
<td>A Three Stage Model for the Visual System</td>
<td>19</td>
</tr>
<tr>
<td>2.4</td>
<td>Mach Bands Explained By Spatial Frequency Response</td>
<td>21</td>
</tr>
<tr>
<td>2.5</td>
<td>Simultaneous Contrast Explained By Spatial Frequency Response</td>
<td>23</td>
</tr>
<tr>
<td>2.6</td>
<td>Normalized Frequency Response of Visual System</td>
<td>25</td>
</tr>
<tr>
<td>3.1</td>
<td>The Sampling Theorem in the Frequency Domain</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>The Effects of Discrete Subsampling</td>
<td>33</td>
</tr>
<tr>
<td>3.3</td>
<td>Varying the Amount of Aliasing Using Spatial Gaussian Filters</td>
<td>40</td>
</tr>
<tr>
<td>3.4</td>
<td>Intersymbol Interference in Interpolation</td>
<td>43</td>
</tr>
<tr>
<td>4.1</td>
<td>Original CMAN Picture</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Experimental Gaussian Filter Values</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>Two Dimensional Picture Space</td>
<td>52</td>
</tr>
<tr>
<td>4.4</td>
<td>Typical Slice of Processed Pictures</td>
<td>55</td>
</tr>
<tr>
<td>4.5</td>
<td>Relative Quality of Processed Pictures</td>
<td>60</td>
</tr>
<tr>
<td>4.6</td>
<td>Relative Noise Powers of Processed Pictures</td>
<td>64</td>
</tr>
</tbody>
</table>
I. Introduction

In the broadest sense, image processing can be construed as any manipulation of information which is inherently two-dimensional in nature. More specifically, however, the techniques which have been investigated fall into three general areas: image enhancement or restoration, pictorial pattern recognition or scene analysis, and picture coding. As the name implies, image enhancement is concerned with the improvement of pictures which have been degraded (i.e. blurred or made noisy) by known or unknown processes. Scene analysis attempts to segment a picture into a map of the real-world objects within the scene, often for purposes of identification or classification of those objects. Picture coding deals with the conversion of real spatially continuous images to discrete form, and the compression of these numbers in order to save storage space or transmission time.

A general block diagram which encompasses all three of these types of image processors is shown in figure 1.1.[1] An examination of its structure reveals that, although at first glance enhancement, analysis, and coding may seem rather mutually exclusive, they share common features and problems. Image formation begins when a real-world object reflects incident radiation (e.g. visible light or x-rays), or a self-luminous ob-
Compression System:

Enhancement/Recognition System:
ject emits it. The light is then collected and focused into an image plane, and usually converted into a form which is more easily dealt with, typically video electric signals. At this point a fundamental issue arises: whether or not to convert the continuous analog signal to a discrete digital one, usually for the purpose of using a computer. More and more often this is done, and for good reason: the range of sophisticated digital signal processing techniques which can be applied to the resulting numbers is much greater than that of analog circuitry processing techniques. If the choice is made not to do the conversion, the system is primarily limited to an image transmission system; the great majority of segmentation, coding, and enhancement algorithms assume digitization. At the present time the great majority of systems in the public domain are transmission systems and are all analog, for example commercial television, printing presses, and facsimile networks. Current research, however, indicates that digital systems in these areas may not be too far away.

Given the conversion from video to an array of numbers, subsequent processing depends on the type of system, as outlined in the figure. In the case of a coding system, the numbers are operated upon by psychovisual and/or statistical coders, and then stored or transmitted through a communications channel. A psychovisual coder is one which reduces the amount of information in the image by throwing away what is unimportant as far as the human visual system is concerned. A statistical coder is
one which reduces the amount of data which needs to be transmitted or stored by recognizing the statistical redundancies in the image and removing them. The communications channel or storage device may itself have processors at each end; operations such as n-bit error detection and correction or modulation and demodulation are common in modern equipment. At the receiving end of the channel, the signal passes through decoding stages which un-do the effects of the coders on the transmitting end, resulting in a faithful replication of the original array of numbers. This replication is then passed on to an output stage.

In systems other than coders, the processing is done on the original digitized image by any number of general processing sections, and the result passed to an output stage. These processing sections are digital hardware or software, most typically computer programs.

In general the desired output of the system will be an image, requiring first the conversion back from discrete numbers to continuous video, followed by a general reversal of the process of the input device to produce either a self-luminous image -- for example the face of a TV picture tube -- or an object which reflects light, like a photograph. From here, the light goes on to the ultimate receiver in the system, the human eye. In the case of the scene recognition system, the output may not be an image, but rather a typed list in English of the objects in the scene which the system recognizes. We shall not consider these further.
Within the general framework of the figure, image processing systems can further be classified by two criteria: the quality of the image that is being dealt with, and complexity and/or expense of the processing which is performed on it. Because these two factors vary over such a tremendous range, it is clear that no single system can be optimum in any sense for all cases. Therefore, what follows is limited to a consideration of what can be accomplished by a small to medium scale computer capability on graphic-arts-quality images. The decision on computation capability results from the fact that small scale computing power is very inexpensive, and getting cheaper all the time. As a result, the decision whether or not to include computers in any image processing system depends only on what they can do, not what they cost. The limitation to graphic-arts-quality images results simply because they comprise the vast majority of hard-copy pictures which are produced every day. "Graphic-arts-quality" means the quality typically found in books, i.e. much better than a television image, somewhat better than a pocket Instamatic picture, not as good as a large format camera picture.

A vast amount of work on image coding, enhancement, and analysis procedures has gone into image processors of the general type outlined in figure 1.1.[2],[3] However, they almost exclusively assume digital input and output without much regard for the conversions to and from the analog domain. To a large degree, the successful operation of these procedures, as well as
the quality of their results, depends on how well these conver-
sions have been made, or appear to have been made to a human ob-
server. Therefore, it is this specific problem which this
thesis addresses -- in simplest terms, "what is the best, or at
least a good, way to represent a continuous picture in discrete
terms, given a human observer?"

Because the ultimate receiver in the canonical system is
a person, with any notion of quality which is highly variable,
the approach which has been taken to solve the problem is better
categorized by analysis of subjective experimental results
than by theoretical calculation. The preparation of test pic-
tures, experimental procedures, and analysis methods which were
involved in this are the subjects of sections four, five, and
six of this document. Sections two and three provide some
relevant background information; the former on the human visual
system, the latter, on a mathematical characterization of the
coding process. Section seven consists of a discussion of the
experimental results, including conclusions and suggestions for
further research. It is followed by references, and appendices
which describe experimental details.
II. A Model for Brightness Perception

Visual perception is clearly an incredibly complex process. The visual system is non-linear, inhomogeneous (i.e., space variant), time variant, and anisotropic -- about as unwieldy as can be imagined. However, even a very basic model can lend considerable insight into some of the features which make an image "good" or "bad". That which follows is a combination of both physiological and psychological evidence.[4] The units involved in this description are light intensity for input, and "apparent brightness" for output. The input intensity is a physically measurable parameter in terms of, for example, light quanta per second per square millimeter. The perceived brightness, on the other hand, is a psychological, descriptive term, which introduces large subjective variances into human judgments of image quality.

A good way to motivate a model for the visual system is by the phenomena which it must explain. Three of the most important are those of simultaneous contrast, Mach bands, and brightness constancy. Simultaneous contrast is illustrated in the familiar squares of figure 2.1. Briefly, simultaneous contrast is the phenomenon in which the apparent brightness of an object of constant known intensity varies strongly as the local background intensity changes. In the squares, the object inten-
figure 2.1. Simultaneous Contrast Phenomenon
sity (little square) is 128 (in the range \(0 - 255\)); the background intensities, 16, 85, 170, 240. The darker the background, the lighter the object looks; the lighter the background, the darker the object looks. Mach bands, named after Ernst Mach, who first described them in 1855, are anomalous bands of brighter-than-bright and darker-than-dark brightness which border high contrast, dark-bright edges in pictures. An example is shown in figure 2.2. Also plotted is a graph of the light intensity which is reflected off of the page and a graph of the perceived brightness as you view it. The phenomenon of brightness constancy is best described by a two-step experiment:

1. A subject is given an enclosed box containing comparison shades of gray. He views them through a window while they are lighted by a constant illuminance. In front of him hangs a larger test shade of gray which is to be matched by one in the box. Initially, with the room lights on, it is not surprising that the subject picks out the comparison shade which is physically the same as the test shade.

2. Now the room lights are dimmed, and the subject is asked to make the match again. Surprisingly, he chooses the same comparison shade, as if to say that the brightness of the target hanging in front of him is independent of the illumination falling on it.

This then is what is meant by "brightness constancy": over a wide range of intensities, the apparent brightness of an object
figure 2.2. Mach Band Phenomenon
can remain constant if it maintains a constant contrast ratio with its background. The subject acts as though he is responding to a property of the object itself, its reflectance, rather than the amount of light which is entering his eye (the only source of information which he has).

A complete model which explains these phenomena is shown in figure 2.3. Its first stage is a familiar optical low pass filter which represents the effects of the eye's lens and cornea. Any describing function for the visual perceptual system must be bandlimited. This stage of the model incorporates that limit in a way that makes physical sense in terms of the optical modulation transfer function (MTF). Such low pass functions are usually obtained using sophisticated opthalmoscopes which measure the actual light intensity on the retinas of living eyes subject to incident light. [5],[6] The cutoff frequency for the low pass characteristic is between .1 and 1.0 cycles per minute of arc. The second stage is a memory-less transformation representing the processes which occur between the absorption of light and the generation of nerve impulse signals. Simple electro-chemical models predict a form which is gently non-linear:

\[ A = \frac{k_1 * I}{k_2 + I} \]

where
A is output
I is input intensity
k<sub>1</sub>, k<sub>2</sub> are constants
optical MTF

nonlinear transformation

recurrent lateral inhibition

figure 2.3  A Three Stage Model for the Visual System

- 19 -
The characteristic plotted in the figure is of a modified log form:

\[ A = \frac{I_{\text{max}} \log(1+ai)}{\log(1+ai_{\text{max}})} , \ a=.1. \]

It was derived by Hashizume using quantization noise sensitivity measurements.[7] Despite functional differences, almost all representations of this nonlinear relationship have very nearly the same shape. The final stage of the model is a filter with sensitivity (apparent brightness) as a function of spatial frequency. Quite unexpectedly, this characteristic resembles a spatial differentiator. The curve shown was determined by Davidson using a technique of contrast matching two sequentially flashed sine wave images of different spatial frequencies.[8]

With this model it isn't too difficult to explain the phenomena of Mach bands, simultaneous contrast, and brightness constancy. The low pass and differentiator stages of figure 2.3 result in something resembling a net bandpass response. With this in mind, it is not surprising that standard Fourier analysis predicts the overshoots exhibited in the perceived brightness of Mach bands. A simpler way to obtain this result is to consider the frequency effects of the model as simply suppressing some amount of the low frequencies, and all of the very high frequencies, of the input intensity. This approach is outlined in figure 2.4. The first plot is of the original intensity distribution. The second plot is of the low frequencies
figure 2.4  Mach Bands Explained by Spatial Frequency Response
of the input, multiplied by a fraction less than one. The third plot is what remains after removing (subtracting) these lows from the original and rounding off the sharp corners (in order to represent the removal of the very high frequencies). The similarity to figure 2.2 is obvious. The simultaneous contrast effect of figure 2.1 results from the same cause. Figure 2.5 is a cross-section view of the intensities of figure 2.1. Applying Fourier transforms gives the output which is expected -- namely, that the apparent brightness of the little squares decreases as the background intensity increases.

An explanation of brightness constancy requires both the logarithmic-like transformation and the differentiator (high pass) filter. Given an object and its background, the intensity of the scene is the product of the reflectance and the illumination. If the intensity goes through a logarithmic stage early in the visual system, the output of that stage is a sum of two components, one due to reflectance and one due to illuminance. Because the illumination is generally slowly varying over the scene, i.e. composed of low frequencies, it disappears going through the high pass stage. Therefore, the apparent brightness is relatively independent of the illumination -- brightness constancy.

Although it is very general, the three stage model is rather cumbersome to use. However, by linearizing the middle stage about an operating point, the sequence of the stages can be commuted, and they can be compressed into a single frequency
figure 2.5 Simultaneous Contrast Explained by Spatial Frequency Response
response for small signals. Mannos and Sakrison have done this, with the result:[9]

$$H(f) = 2.6 \cdot (0.0192 + 0.114f) \cdot \exp(-0.114f^{1.1})$$

Here $f$ is analog spatial frequency, in units of cycles per degree of arc. With a given viewing distance and sample resolution, $f$ can be transformed to discrete normalized frequency, valid on (or, approximately, near) the axis of viewing. The result plotted in figure 2.6 is for 155 points/inch at 30 inches.

Both the three stage model and its simplified version explain real-world phenomena very well. However, they have their limitations. For one, they assume isotropy, which is clearly not the case (we see better horizontally and vertically than diagonally). Both also ignore the remarkable adaptation which the eye exhibits in varying brightness environments. Such considerations should be second order, however.
Figure 2.6 Normalized Frequency Response of Visual System

Sampling rate: 155 pels/inch

Viewing distance: 30 inches
III. A Characterization of the Spatial Coding Process

A. The Digital Coding Problem

Conversion from a continuous spatial domain to a discrete one, and back again, is generally considered an analog problem. However, there are advantages to casting it into digital form, among them: 1) existing scanning equipment can be used without modification of analog circuitry, and 2) computers can be utilized to do the signal processing digitally. The preceding section makes it clear that since the human visual system is inherently band-limited, the images that are dealt with might as well be also. With reference to figure 2.3, visual response starts to fall off at .25 cycles/minute. At the closest possible viewing distance at which the visual system can focus, about six inches, this corresponds to about 150 cycles/inch or a Nyquist rate of 300 points/inch. By spatially sampling above this rate, a conversion from analog to very high resolution digital and back to analog should be effected with absolutely no loss in apparent quality. Accordingly, the "input sub-system" of figure 1.1 can be split into two stages: a very high resolution, fixed-rate analog to digital conversion, followed by a purely digital processor which extracts the best lower resolution representation of the image. Conversely, the
output sub-system can be split up into the reverse of these two stages.

The mathematical justification for this conversion rests in the multi-dimensional sampling theorem. First described by Peterson and Middleton in 1962[10], the theorem serves to relate the frequency spectrum of an original continuous signal to that of a discrete signal obtained by sampling it. The case of interest is for sampling on a square grid of period $T$ -- the usual mode of operation of the scanner performing the conversion. If $X(jU, jV)$ is the spectrum of the continuous input and $Y(e^{ju}, e^{jv})$ is the spectrum of the discrete sequence obtained by sampling, then:

$$Y(e^{ju}, e^{jv}) = \frac{1}{T^2} \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} X\left(\frac{ju}{T} + \frac{j2\pi p}{T}, \frac{jv}{T} + \frac{j2\pi q}{T}\right),$$

where $u = U \times T$ and $v = V \times T$.

The general effect is to scale the continuous spectrum in frequency by a factor of $T$, and then replicate it throughout the frequency plane at all integral multiples of $2\pi$. This is illustrated in figure 3.1. For $T$ very small, i.e. a very high resolution conversion, the spectra are clustered closely around the multiples of $2\pi$. If the input bandwidth is limited, e.g. by the scanner optics, then little overlap of the replications occurs, and the single replication at the origin can be obtained with a non-critical continuous low-passing operation, resulting in almost the original spectrum.
original continuous input spectrum

two-dimensional bandwidth (region of support)

U (analog domain)

discrete spectrum after very high resolution sampling of period T

V

2π 2π

-2π -2π

u (discrete domain)

output spectrum after continuous low pass filter

V

a

U (analog domain)

-28-

figure 3.1 The Sampling Theorem in the Frequency Domain
Besides the spatial geometrical quantization which occurs in the conversion from analog to very high resolution digital, there remains the matter of amplitude quantization. The fact that the d.c. response of the eye is roughly logarithmic means that amplitude quantization noise visibility in any part of an image is proportional to the brightness there. Therefore, uniform quantizers result in the dark parts of a picture looking noisier than the lighter parts. Tapered quantizers, on the other hand, which have their quantization steps closer together at small levels, make the errors equally visible over the entire dynamic range of the input. Such quantizers are usually implemented as a cascade of a nonlinear transformation and a uniform quantizer. Because any digital processing adds noise, it is generally advisable to use this nonlinear domain (the so-called "lightness" domain) throughout the course of the processing to minimize the effects.[11],[12] In this case, eight bit samples are more than adequate to represent the amplitude; few imaging systems have signal-to-noise ratios which permit an observer to distinguish between more than 256 shades of gray. Because this number is also conveniently an integral fraction of typical computer word lengths, it is often used without question.

Given a very high resolution digitized picture, there are two general classes of coders which can be used to represent the information: transform coders and direct coders.[13] As the name implies, transform coders transform the input signal into a formulation whereby it is expressed as a linear summation of
orthonormal two-dimensional basis functions. The extraction of a low 
resolution "digitized image" then consists of selecting the coefficients in this sum which are the "most important" in terms of the amount of information they contain (usually these correspond to low frequencies). The post-processing operation consists of inverse transforming these selected samples to obtain an approximation to the input signal. Direct coders, on the other hand, are the class for which the "digitized image" samples are picture brightness values. In this case the pre-processing operation is a filter and the extraction is done in the spatial domain on a non-random two-dimensional grid. Post-processing then consists of interpolating between these points to obtain the output approximation. Both of these classes have their advantages and disadvantages. Given a statistical characterization of the ensemble of pictures to be coded, in the form of a cross-correlation function, the Karhunen-Loeve transformation is the optimal coder in the minimum mean square error sense. However, much computation is involved. Less, but still quite involved, work is required for the more well known Fourier and Hadamard transforms, in which the basis functions are complex sinusoids and square waves respectively. The major disadvantage with these and all other transform techniques is that the coded representation of the image doesn't make sense to a human observer or the majority of coding/enhancement/recognition algorithms -- they require a direct rather than transformed representation. Direct coding schemes, while generally not as
efficient as transform coding ones, produce digitized images which look like the original and are amenable to simple fundamental picture processing operations like domain (tone scale) transformation and geometrical processes (e.g. scaling, rotation and cropping.) Another significant advantage is that while transform coders require the entire image to be buffered as it is coded, most direct schemes require only a small portion of it. Transform coders sometimes get around this restriction by arbitrarily blocking the picture into square sub-pictures and coding these independently. The fact that this blocking is arbitrary can only hurt, and usually does.

B. PCM Picture Coding

Pulse Code Modulation, PCM, is the direct coding scheme which uses a regular, space-invariant two-dimensional sampling grid. It is the most common method of representing a continuous signal in terms of discrete samples -- essentially the "intuitively obvious" way to do it. In the extraction procedure, the input is prefiltered and then subsampled to obtain the digitized image. In reconstruction, the digitized image samples are first expanded, so that each occupies its original spatial location, and then interpolated between.

The key problems involved in PCM are most easily motivated by a consideration of the subsampling stage. Recall that the input is a very high resolution discrete version of the
original continuous image, with a spectrum which is periodic in 2\pi in both spatial frequency variables (figure 3.1b.) Any arbitrary two-dimensional subsampling grid can be characterized by the three parameters (integers) \(a, b,\) and \(c\) shown in figure 3.2a. Functionally this grid can be represented by

\[
g(\vec{r}) = \sum_{\text{all } \vec{k}} \delta(\vec{r} - B \cdot \vec{k})
\]

where \(\vec{r} = [m \ n]^T\)

\[
B = \begin{bmatrix}
a & b \\
0 & c
\end{bmatrix}.
\]

The vector \(\vec{r}\) is the location in the space plane, the matrix \(B\) has columns which are the two, two-dimensional basis vectors of the pattern, and \(\vec{k}\) is a vector which spans over all possible values. Because this pattern is periodic, it has a Fourier series representation. In vector notation it turns out to be:

\[
g(\vec{r}) = \frac{1}{ac} \sum \bar{I} e^{(j2\pi \bar{I} \cdot B^{-1} \cdot \vec{r})}
\]

where \(B^{-1} = \begin{bmatrix}
1/a & -b/ac \\
0 & 1/c
\end{bmatrix}\)

and \(\bar{I}\) is an index vector that takes on values from \([0 \ 0]^T\) to \([a-1 \ c-1]^T\). The subsampling operation consists of two steps: multiplying the input sequence by the sampling grid and then extracting the non-zero values. [14] If the multiplication operation is expressed as

\[
m(\vec{r}) = x(\vec{r}) \ast g(\vec{r})
\]
Figure 3.2 The Effects of Discrete Subsampling

a. subsampling grid

b. spectrum after multiplying by subsampling grid.

c. spectrum after extracting nonzero subsamples (arrows show frequency scaling).
then its transform can be developed as follows:

\[ M(e^{j\tilde{w}^T}) = \sum_{\tilde{r}} m(\tilde{r}) * e^{(-j\tilde{w}^T \cdot \tilde{r})} \text{, with } \tilde{w}^T = [w_x \ w_y] \]

\[ = \sum_{\tilde{r}} x(\tilde{r}) * \left( \frac{1}{ac} \sum_{I} e^{(j2\pi I^T \cdot B^{-1} \cdot \tilde{r})} \right) * e^{(-j\tilde{w}^T \cdot \tilde{r})} \]

\[ = \frac{1}{ac} \sum_{I} \sum_{\tilde{r}} x(\tilde{r}) * e^{-j(\tilde{w}^T - 2\pi I^T \cdot B^{-1}) \cdot \tilde{r}} \]

\[ = \frac{1}{ac} \sum_{I} X(e^{j(\tilde{w}^T - 2\pi I^T \cdot B^{-1})}) . \]

This result is plotted in figure 3.2b. It is similar to that for the continuous case in that the input spectrum, unscaled in frequency, is replicated at a finite number of points in the frequency plane. The replication points are a function of the set of indices \( I \) and the inverse of the basis matrix, \( B^{-1} \). Given \( m(\tilde{r}) \), the digitized image is comprised of its non-zero values as defined by the equation

\[ y(\tilde{r}) = m(B^* \tilde{r}). \]

In the frequency domain this is equivalent to

\[ Y(e^{j\tilde{w}^T}) = M(e^{j\tilde{w}^T \cdot B^{-1}}) \]

meaning that a plot for \( Y \) is obtained by warping the plot for \( M \) (figure 3.2b) in such a way as to make the replications periodic in 2\( \pi \). This is simply a two-dimensional frequency scaling, as demonstrated in figure 3.2c. The figure also demonstrates the relationship between the output and the input, namely:
\[ Y(e^{j \tilde{w}^T}) = \frac{1}{ac} \sum I X(e^{j(\tilde{w}^T \cdot B^{-1} - 2\pi I^T \cdot B^{-1})}) = X(e^{j(\tilde{w}^T \cdot B^{-1})}). \]

This equation says, to plot the output spectrum, take every point in the frequency plane for the input and transform it by B. Clearly, if B is not a square transformation, the spectrum of the digitized image has a shape very different from that of the input. Any signal processing operations subsequently performed on the digitized image must take this into account. Essentially, in this case the image is being represented in an oblique coordinate system -- a non-trivial complication for most image processing systems.

When the digitized image is converted back to the very high resolution representation, the spatial warping demonstrated in figure 3.2c is of no further consequence. This is because the first step of the reconstruction, the expansion, can be expressed as

\[ e(B \cdot \tilde{r}) = y(\tilde{r}) \]

or

\[ E(e^{j \tilde{w}^T}) = Y(e^{j \tilde{w}^T \cdot B}). \]

The inverse transformation \( B^{-1} \) maps the plot of figure 3.2c directly back into that of figure 3.2b, restoring the original input shape. It is the function of the interpolater to then ob-
tain the form of the original as plotted in figure 3.1b.

With this examination of the subsampling process, the functions of the prefilter and postfilter become clearer. The purpose of the prefilter is to bandlimit the input. In any real world system this is impossible. What is of interest is the best way to bandlimit the input approximately, and the best way to subsample the result, keeping in mind that its quality will be judged subjectively. The purpose of the postfilter is to extract the single replication at the origin (and multiples of 2πi) in the frequency plane, in order to obtain the original spectrum. In the space plane, this amounts to optimally interpolating between the subsampled points after they have been expanded. Again, this requires a band-limiting low pass filter, which must be approximated in some good way in a real system.

C. Tradeoffs in PCM Picture Coding

The classical approach to bandlimiting and interpolation is a frequency-domain approach -- based on the development of the previous section -- which is theoretically optimal in the sense of minimum mean square error between input and output. Given a desired two-dimensional sampling raster (which usually reflects a desired amount of scaling up or down), the procedure is to employ ideal low pass filters for both bandlimiting and interpolating. Such filters are constant over the region of support in the frequency plane and zero elsewhere. In using
them, there is no aliasing, and as large a bandwidth as possible is maintained. From the frequency domain standpoint this technique seems ideal. However, there are some disadvantages.

1. It doesn't take into account the properties of the eventual receiver, the human visual system.

2. The spatial domain counterparts of ideal filters are sinc functions. Because they are infinite in extent, implementation of the filters by direct convolution is impossible. Because the Fourier series converges relatively slowly and nonuniformly, truncation of the sinc functions results in ringing at sharp edges -- the Gibbs phenomenon. This is not necessarily a priori bad. However, such artifacts are extremely visible, and usually annoying to an observer because the differentiator-like characteristic of his visual transfer function (fig 2.6) accentuates the ringing relative to the average pel value in the local area.

3. Optimality in the minimum mean square error sense is not a particularly good choice from a psychovisual point of view. Consider a picture of medium resolution composed of four bit pels. A substantial error results from adding one to every pel, however, the effect is not too disturbing. Exactly the same mean square error results from adding one to every even numbered pel in even lines and odd numbered pel in
odd lines, and subtracting one from the rest (i.e., something like a checkerboard). This resulting picture looks much worse than the previous one, despite the same error. Clearly while mean square error is certainly correlated with subjective picture quality, it cannot serve alone as the definitive criterion.

An alternative to the ideal low pass filter is suggested by noting that the requirements imposed by the subsampling process are at extreme odds. The filters involved should be narrow in the space domain in order not to blur out the image if it can be helped. They should also be narrow in the spatial frequency domain in order to minimize aliasing. It is a general consequence of linear system theory that both can not be possible at the same time. If the width of a filter in either domain is taken as its second central moment (i.e. moment of inertia about the center of gravity), then

\[ W_s \times W_f \geq \frac{1}{4\pi} \]

where \( W_s \) is the width in space and \( W_f \) is the width in frequency.\[15\] As the equation suggests, any individual filter has to trade blur for aliasing. The question is how to do it and get the best looking picture (the object of the experiment described in section four).

Equality in the above equation is obtained only for
gaussian filters. Besides this minimum space-width*frequency-width property, these filters also have the advantages of circular symmetry and separability in two orthogonal dimensions, all of which suggests their applicability to the problem at hand. The space/spatial frequency gaussian pair is:

\[ g(r) = \frac{1}{2\pi \sigma^2} e^{-r^2/2\sigma^2} \]

\[ G(e^{j\omega r}) = e^{-\omega^2 \sigma^2/2} \quad \text{around the origin for } \sigma \geq 1. \]

With reference to figure 3.2b, the replication of the input spectrum in the frequency plane defines a cutoff frequency, \( w_{co} \), equal to half the distance between nearest neighbor replication points. One way to quantify the effect of the space-width/frequency-width tradeoff for any filter is to consider how much of the filter energy is within such a given radial cutoff frequency. The two equations above make it clear that for a gaussian, the tradeoff is completely characterized by a single parameter, the standard deviation sigma, and as a result, the percentage of energy is a simple function of sigma alone:

\[ \text{Energy (percent)} = 100(1 - e^{-\sigma^2 w_{co}^2}). \]

For more complex filter functions (still low pass in nature), this metric is still obtainable, but perhaps only by numerical means. Figures 3.3a and b demonstrate the space-width/frequency-width tradeoff for a gaussian filter for two different percentages, 75% and 95%.
a. effects of Gaussian filters in the frequency domain.

b. corresponding spatial filters.

Figure 3.3 Varying the Amount of Aliasing Using Spatial Gaussian Filters
The space-width/frequency-width tradeoff is the fundamental one involved in the PCM bandlimiting operation. Although other issues arise (e.g. the degree of circular symmetry of the filters used), their effects are greatly diminished by the resolution reduction inherent to the subsampling operation itself. In the interpolating operation, on the other hand, other factors must be considered (e.g. filter isotropy or sample visibility) because they are visible in the higher resolution output. The functions of bandlimiting and interpolating are distinct, so there is little reason to assume that identical filter types should be used for both. It should also be noted that the reconstruction phase of PCM is an important problem in its own right -- that of the best way to enlarge a given input to result in a higher sample density output. For these reasons it is worthwhile considering the interpolation process on its own, given a bandlimited, sampled input.

Any interpolation scheme which uses a space-invariant interpolating function can be expressed easily in terms of linear system equations. This class includes most popular techniques, but excludes polynomial-fitting approaches, for example LaGrange interpolation. Within this context there are two basic goals:

1. obtain a higher resolution output than input, and
2. insure that the output is exactly equal to the input at the points where the input is given.

With a low-pass shaped interpolating function, the first cri-
terion is met more or less automatically. In the case where the interpolating function has a span which is less than or equal to the input point separation, e.g. linear interpolation, the second criterion is also met. However, in the general case, the interpolating function has a span which is several times the input point separation, and each output point is a linear combination of several input points. This situation, non-zero intersymbol interference, is demonstrated in figure 3.4a. Considering the input values at the set of points where the input is non-zero (the sample "knots") as a sequence \( x(n) \), and the output values \( y(n) \), the effect of a linear filtering with the interpolating function can be expressed (in one dimension and with only nearest neighbor interaction, for simplicity) by the difference equation:

\[
y(n) = x(n) + a x(n-1) + a x(n+1) .
\]

Clearly for non-zero \( a \), the outputs at the knots are not equal to the inputs. If this filtering can be preceded by another, however, such that the net system function is unity, zero intersymbol interference will result. The pre-filter will in general have a sharpening effect. Taking the reciprocal of the system function resulting from the difference equation above defines how this filter must behave:

\[
H(z) = \frac{1}{1 + a(z^{-1} + z^{+1})} .
\]
y(n) = x(n) \cdot a \cdot x(n-1) + a \cdot x(n+1)

a. intersymbol interference in the space domain

\[ F(e^{j\omega}) = \frac{1}{1 + 2a \cos(\omega)} \]
\[ a = 0.25 \]

b. necessary sharpening filter

Figure 3.4 Intersymbol Interference in Interpolation
This filter is both noncausal, and has an impulse response which is infinitely long. Because of the latter fact, stability must be considered. The poles of the z-transform are located at:

\[
\frac{-1}{2a} \pm \frac{1}{\sqrt{(2a)^2 - 1}}.
\]

For the system to be stable, \( a \) must be less than \( 1/2 \). This isn't too serious a restriction in view of typical interpolating functions. The actual implementation of this processing may consist of either recursive or non-recursive filters. For the recursive case, note that the z transform can be factored to give:

\[
H(z) = 1 - \frac{a}{a + z + az^2} - \frac{a}{a + z^{-1} + az^{-2}}.
\]

The output is then a sum of three components: the original, a recursively filtered component proceeding forward, and a recursively filtered component proceeding backward. This approach is optimal, but requires the entire input to be buffered. A FIR filter design can be used to approximate the recursive filter, and requires only a small amount of data at a time. The Fourier transform of such a \( 2N+1 \) point filter is best approximated, in the MMSE sense, by the \( 2N+1 \) samples of the z transform uniformly spaced around the unit circle.
These are given by:

\[ H(k) = \frac{1}{1 + ae^{j2\pi k/2N+1} + e^{-j2\pi k/2N+1}}, \quad -N \leq k \leq N. \]

The spatial domain filter which corresponds to these samples is then given by the inverse Discrete Fourier Transform, as follows:

\[ h(n) = \frac{1}{2N+1} \sum_{k=-N}^{+N} H(k) e^{j2\pi nk/2N+1}, \quad -N \leq n \leq N. \]

\[ = \frac{1}{2N+1} \left[ \frac{1}{1 + 2a} + \sum_{k=1}^{N} \frac{2\cos(2\pi nk/2N+1)}{1 + 2a \cos(2\pi k/2N+1)} \right]. \]

For example, for the simple case of a 3 point filter, the values are

\[ h(0) = \frac{1+a}{(1+2a)(1-a)} \]

\[ h(1) = h(-1) = \frac{-a}{(1+2a)(1-a)} . \]

Algebraic interpolation, then, can be represented by the cascade of two operations. The first is a filtering to remove the intersymbol interference. The second is a simple convolution with the basis interpolating functions to do the actual interpolation. The effect of the second phase is clearly a low-passing operation since the interpolating functions are all positive. The effect of the first phase is a sharpening opera-
tion. This can be seen from the Fourier transform, evaluating 
$H(z)$ for $z = e^{jw}$:

$$
H(e^{jw}) = \frac{1}{1 + 2a \cos w}, \quad a < .5
$$

Figure 3.4b shows the result for $a = 1/4$.

Within the confines discussed above, any interpolating function can be made to have zero intersymbol interference. The choice of good interpolating functions therefore must depend on other criteria (the determination of which is the object of the experiment described in section five). The cubic b-spline has the property that it maintains continuous values, slopes, and curvatures for any input. Hou and Andrews have demonstrated its use in image magnification.[16] The use of the gaussian, which is not too different in shape from a cubic b-spline, has been demonstrated by Schreiber, Hoover, and Grass.[17],[18] A great many other interpolating filters suggest themselves, particularly the good low pass window type of filters, e.g. Hamming, Hanning, and Blackman functions, numerically derived functions[19], and the set of two dimensional filters derived from transformed single dimensional designs, e.g. McClellan transform filters.[20],[21]
IV. Experimental Investigation of
The Space-width/Frequency-width Tradeoff

The first of the experiments described in this thesis was undertaken in order to investigate the tradeoff between blur and aliasing as it affects subjective image quality. Like the subsequent experiments, it was performed on the MIT CIPG Image Processing System (IPS). This system is a dedicated, small-scale computer based picture processing facility, developed to conveniently manipulate graphic arts quality images.[22]

In general, images fall into one of two broad categories, continuous tone and line art. The CIPG "CMAN", a continuous tone image, was chosen as the test picture for the experiment because: 1) it offered a wide range of difficult-to-deal-with image characteristics, such as highlight and shadow detail and texture, slowly varying low spatial frequency regions, and high contrast, diagonally oriented edges, and 2) it was typical in content of real world images. Because the experiment required the production of many (25) filtered versions, only a single picture was used, and the class of line art pictures was not represented.

CMAN was scanned on an ECRM Autokon electronic process camera at an effective resolution of 380 points per inch, and immediately transformed to the lightness domain. The very high
resolution was achieved as a combination of optical photographic enlarging (by approximately a factor of four), and the use of the Autokon magnification controls. The tone scale transformation function necessary to put the scanned picture into the lightness domain resulted from a careful calibration scheme, which is described in appendix A. The limitations of the IPS separable filtering processes, used later, were established by the transpose task, which was only capable of processing a picture 1024 pels square or smaller. For this reason, a 1024 pel square segment of CMAN was cropped and used as the "original" very high resolution picture; it is shown in figure 4.1. Hard copy outputs of all pictures were obtained on an AP Laserphoto© facsimile receiver. Like the Autokon, its output tone scale transformation was carefully measured, and compensated for when the pictures were printed. It should be noted that the Laserphoto output pictures which were used were continuous tone images; the copies which appear here are halftone screened versions of those images. (Because all the pictures referenced in this thesis are archived on magnetic tape (standard IPS TAPe format), additional continuous tone copies can be made.)

Gaussian filters were used in the prefiltering (bandlimiting) and postfiltering (interpolating) operations because they could be characterized by a single parameter. The metric which was chosen was the filter's percentage of energy within the passband implied by the subsampling frequency, as discussed earlier. For the purposes of subjective evaluation, five distinct
**figure 4.2** Experimental Gaussian Filter Values

<table>
<thead>
<tr>
<th>Sub-sample spacing:</th>
<th>5.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Energy percentage:</td>
<td>50.</td>
</tr>
<tr>
<td>Sigma (pels):</td>
<td>1.325</td>
</tr>
<tr>
<td>D.C. gain:</td>
<td>66.4</td>
</tr>
<tr>
<td>IPS file name:</td>
<td>GSQ9.TBL GSQ11.TBL GSQ13.TBL GSQ15.TBL GSQ17.TBL</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Coefficients, h(0):</th>
<th>100.</th>
<th>100.</th>
<th>100.</th>
<th>100.</th>
<th>100.</th>
</tr>
</thead>
<tbody>
<tr>
<td>h(1),h(-1):</td>
<td>75.</td>
<td>85.</td>
<td>88.</td>
<td>92.</td>
<td>94.</td>
</tr>
<tr>
<td>h(2),h(-2):</td>
<td>32.</td>
<td>52.</td>
<td>61.</td>
<td>71.</td>
<td>77.</td>
</tr>
<tr>
<td>.</td>
<td>8.</td>
<td>23.</td>
<td>33.</td>
<td>46.</td>
<td>55.</td>
</tr>
<tr>
<td>.</td>
<td>1.</td>
<td>7.</td>
<td>14.</td>
<td>25.</td>
<td>35.</td>
</tr>
<tr>
<td>0...</td>
<td>2.</td>
<td>5.</td>
<td>12.</td>
<td>19.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>0...</td>
<td>1.</td>
<td>5.</td>
<td>9.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td>0...</td>
<td>0...</td>
<td>1.</td>
<td>4.</td>
<td></td>
</tr>
<tr>
<td>.</td>
<td></td>
<td>0...</td>
<td>1.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
values for this metric were chosen: 95%, 90%, 80%, 70%, and 50%. A subsampling factor of 5 was chosen because it was large enough to represent a substantial decrease in resolution, yet small enough to permit the use of reasonable width spatial filters. (The implications of this subsampling factor in the frequency domain are discussed in section seven). For this subsampling factor, the corresponding standard deviations of the gaussian filters, and the actual filter values, are given in figure 4.2. The IPS FILter and transpose processes (TRaNsPose, INSERT, and eXTRACT) were used to implement the filters as separable convolutions in the space domain. The IPS ARithmetic process was used to do the subsampling.

The bandlimiting, subsampling, and interpolation processing resulted in a set of 25 pictures (5 choices of prefilter times 5 choices of postfilter). As illustrated in figure 4.3a, this set was organized along two axes: "increasing prefilter width" and "increasing postfilter width". Units along these axes were considered simply as picture indices (0,1,2,3,4) or the units of the metric itself (percentages, 0-100). In either case, the set of pictures determined a set of points over the range, and their qualities, a contour over the range.

Subjective picture quality measures typically take one of two forms: specification of absolute goodness (e.g. Q=10 for a very good picture, Q=0 for an extremely bad one), or scaled rank ordering within a group (e.g. Q=10 for best picture in group, Q=0 for worst picture in group).[23] The former technique
figure 4.3 Two Dimensional Picture Space

<table>
<thead>
<tr>
<th>Energy % ( \sigma )</th>
<th>50 ( 1.3 )</th>
<th>70 ( 1.7 )</th>
<th>80 ( 2.0 )</th>
<th>90 ( 2.4 )</th>
<th>95 ( 2.7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 ( 1.3 )</td>
<td>CM99</td>
<td>CM0911</td>
<td>CM0913</td>
<td>CM0915</td>
<td>CM0917</td>
</tr>
<tr>
<td>70 ( 1.7 )</td>
<td>CM1109</td>
<td>CM1111</td>
<td>CM1113</td>
<td>CM1115</td>
<td>CM1117</td>
</tr>
<tr>
<td>80 ( 2.0 )</td>
<td>CM1309</td>
<td>CM1311</td>
<td>CM1313</td>
<td>CM1315</td>
<td>CM1317</td>
</tr>
<tr>
<td>90 ( 2.4 )</td>
<td>CM1509</td>
<td>CM1511</td>
<td>CM1513</td>
<td>CM1515</td>
<td>CM1517</td>
</tr>
<tr>
<td>95 ( 2.7 )</td>
<td>CM1709</td>
<td>CM1711</td>
<td>CM1713</td>
<td>CM1715</td>
<td>CM1717</td>
</tr>
</tbody>
</table>

increasing postfilter width

a. axes and picture names

b. slices

slice shown in figure 4.4
has the advantage that pictures to be judged may be considered individually, but the disadvantage that the subject's concept of goodness and badness may vary widely over the course of the experiment. The latter technique overcomes this disadvantage, but puts an upper limit on the number of pictures which can be judged effectively, typically about nine. The method which was used was a compromise between these two. Rather than compare the entire set of pictures at once, subsets were created by extracting pictures which fell along slices taken through the input range (see figure 4.3b). All possible horizontal, vertical, and 45 degree diagonal slices were used, so each picture was included with four other pictures four times. Each subset was then considered independently, and the pictures within each were rank ordered. In this comparison, a picture received four points for being the best in the group of five; zero points for being the worst. After all the comparisons had been made, the relative quality of a picture was taken as the sum of the points it had accumulated.

The experimental conditions under which subjects made quality judgments were carefully controlled. The pictures to be compared were presented on a neutral background, illuminated by diffuse fluorescent light. Subjects were seated directly in front of the pictures in a chair with an adjustable headrest which served to maintain a constant distance between the subject's eye and the presented pictures. This distance, 30 inches, was sufficient to insure that the Laserphoto output
resolution, 155 points/inch, was well beyond the Nyquist rate corresponding to the visual transfer function cutoff frequency (i.e. so limitations in the Laserphoto output process would have no discernible effect). Subjects were given as much time as desired to make their comparisons. After the first comparison was completed, they were told how many more remained in order to permit them to establish their own rate. One subject took as little as 15 minutes to make the 20 comparisons; one, as long as 45 minutes. A total of ten subjects participated in this experiment. Most were affiliated with CIPG, and were considered experienced, if not expert, observers.

A typical subset of pictures is shown in figure 4.4a-e (this is the diagonal slice, noted in figure 4.3b, ordered from top left to bottom right). They should be viewed at arms length to simulate the experimental conditions.

The combined raw experimental results for all ten subjects are plotted below in figure 4.5a. In the figure, the contour represents the mean value, while the vertical bars represent ±1 standard deviation of experimental error. The data is noisy, as in most psychovisual studies, but does exhibit some very definite trends. These trends can be made quantitative by the application of a regression analysis to smooth the data. The simple linear regression plotted in figure 4.4b exhibits a generally negative slope along the prefilter axis, and a slight positive one along the postfilter axis, but clearly cannot exhibit a peak at the optimum values. The higher order re-
figure 4.4a. Typical Slice of Processed Pictures
Subsample spacing: 5
Energy percentage: 50
Standard deviation of gaussian filters: 1.325
figure 4.4b. Typical Slice of Processed Pictures
Subsample spacing: 5
Energy percentage: 70
Standard deviation of gaussian filters: 1.746
figure 4.4c. Typical Slice of Processed Pictures
Subsample spacing: 5
Energy percentage: 80
Standard deviation of gaussian filters: 2.019
figure 4.4d. Typical Slice of Processed Pictures
Subsample spacing: 5
Energy percentage: 90
Standard deviation of gaussian filters: 2.415
figure 4.4e. Typical Slice of Processed Pictures
Subsample spacing: 5
Energy percentage: 95
Standard deviation of gaussian filters: 2.755
figure 4.5 Relative Quality of Processed Pictures

a. raw data

b. 1st order regression
figure 4.5, continued

c. 2nd order regression

relative quality

increasing postfilter width

increasing prefilter width
gression plotted in figure 4.4c does exhibit this peak, and serves to smooth the data out very well. Mathematically the contour can be expressed as[24]

$$\log Q = a + bx + cy + dxy + ex^2 + fy^2$$

where: $Q$ is subjective quality ($Q > 0$)
$x$ is prefilter energy percentage ($100 > x > 0$)
$y$ is postfilter energy percentage ($100 > y > 0$)
$a, b, c, d, e, f$ are constants determined by the regression:
\[a = -6.85 \quad b = .095 \quad c = .139 \quad d = .00001 \quad e = -.00079 \quad f = -.00093 \quad R^2 = .85\]

The $R^2$ statistic is a measure of how well the curve fits the data. A value of one is a perfect fit; a value approaching zero is a very poor fit.[25] The form of the equation is a standard second order linear regression, except that the logarithm of the quality is used to take advantage of the fact that it must be a positive quantity. Because it is second order, derivatives can be taken to find the optimum values of $x$ and $y$ which maximize $Q$:

$$\begin{bmatrix} x_{opt} \\ y_{opt} \end{bmatrix} = \frac{1}{4ef-d^2} \begin{bmatrix} 2f-d & -b \\ -d & 2e \end{bmatrix} \begin{bmatrix} -c \end{bmatrix} = \begin{bmatrix} 60.4 \\ 75.2 \end{bmatrix}.$$  

For a gaussian filter, 60% and 75% are equivalent to standard deviations of .30 and .375 subsample spacings, or 1.5 and 1.875 pels for a subsample spacing of 5.

An important unsolved problem in picture coding is the determination of objective quality measures which predict results similar to subjective techniques, such as the one described above. As discussed earlier, mean square error does not work well in this capacity. However, several investigators,
e.g. Stockham in [12], have suggested the use of a squared error criterion weighted by the types of processes found in the visual system. For example, given an original input picture $P_i(x,y)$ and a processed output picture $P_o(x,y)$, rather than using

$$\sum_{x,y} [P_o(x,y) - P_i(x,y)]^2$$

and its various statistics, they suggest using

$$\sum_{x,y} [F(x,y) \triangleq (L(P_o(x,y)) - L(P_i(x,y)))]^2$$

where: $L()$ is the visual nonlinear transformation $F(x,y)$ is the visual impulse response.

Figure 4.6a is a plot of the unweighted noise powers of each of the set of 25 processed pictures. The noise power of each output picture was obtained by subtracting it from the original CMAN, resulting in a signed 7-bit error picture, and computing the variance. Figure 4.6b is a plot of the weighted noise powers, obtained by: 1) transforming the input and output to the lightness domain, 2) subtracting to form the error picture, 3) filtering the result with the frequency response of figure 2.6, and 4) calculating the variance. The shapes of the weighted and unweighted curves are similar, but very much different from that of the subjective contour. Clearly, the weighted square error criterion offers no improvement in predicting quality; its failure only serves to accentuate the extreme difficulty in finding objective measures (discussed further in section seven).
figure 4.6 Relative Noise Powers of Processed Pictures

a. unweighted noise powers

b. weighted noise powers
V. Experimental Investigation of Interpolation Tradeoffs

While the first experiment was designed to acquire an understanding of the single fundamental tradeoff involved in the bandlimiting operation, the second experiment was undertaken in order to investigate the wider variety of issues involved in the reconstruction phase of PCM.

Because the number of filtered pictures involved in this experiment was smaller than in the first one, a second test picture, representative of the class of line art images, was utilized in addition to CMAN. This picture, "RWEDGE", was a combination resolution wedge and text segment. The resolution wedge was obtained from a copy of the Japanese facsimile test chart number 1R; the text, from a Monotype sample book of Times New Roman 24 point type. Both segments were scanned in on the Autokon at a resolution of 250 points per inch, then cropped, concatenated, and transformed into the lightness domain. The resulting original is shown below in figure 5.1. It is clearly quite different looking than CMAN, but allows the same kind of determinations about the filters to be made. (For example, RWEDGE's type and CMAN's tripod are essentially high contrast, diagonal edges -- they both test the isotropy of the coding procedure.)

The CMAN and RWEDGE originals were first bandlimited
In the vocation of thy gained by the quiet, I have equalized my joy ABCDEFGHIJKLMNOP

figure 5.1. Original RWEDGE Picture
with a gaussian filter of standard deviation 1.7 pels (this filter was the one in the set of filters used in the first experiment closest to the optimum value determined there). The two pictures were then subsampled by a factor of five, and subjected to a variety of interpolation filters. As in the first experiment, the IPS ARI process was used to do the subsampling, and the IPS FIL, TRN, INSERT, and XTRACT processes were used to implement the filters as separable spatial convolutions.

Several of the criteria discussed in section III were considered in designing the filters, among them:

smoothness - no ringing or contours, minimum visibility of sampling structure

sharpness - no intersymbol interference, maximum frequency bandwidth

isotropy - no stair-step rendition of diagonal edges, minimum spatial quantizing effects

A total of eight filter functions were used: sample-and-hold, bilinear, truncated gaussian, cubic b-spline, truncated sinc, truncated sharpened gaussian, and two varieties of sharpened cubic b-spline. The detailed values are given in figure 5.2, as well as an indication of how well each meets the above criteria.

In general, the filters can best be characterized by their frequency responses and impulse responses. These are shown in figures 5.3a-h, along with contour maps of the impulse responses -- useful in describing spatial characteristics. In these figures, medium gray represents a value of zero, lighter
**figure 5.2. Interpolation Filter Values**

All filters are zero phase. Values are given only for n >= 0.

<table>
<thead>
<tr>
<th>Name</th>
<th>Sample and hold</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPS file name</td>
<td>FWSAH.TBL</td>
</tr>
<tr>
<td>Functional form</td>
<td>h(n) = 1 , n &lt;&lt; 2</td>
</tr>
<tr>
<td>Normalizer</td>
<td>1</td>
</tr>
<tr>
<td>Coefficients</td>
<td>1, 1, 1</td>
</tr>
<tr>
<td>Smoothness</td>
<td>poor</td>
</tr>
<tr>
<td>Sharpness</td>
<td>poor</td>
</tr>
<tr>
<td>Isotropy</td>
<td>poor</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Bilinear</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPS file name</td>
<td>FWBLIN.TBL</td>
</tr>
<tr>
<td>Functional form</td>
<td>h(n) = 1 - (n/5) , n &lt;&lt; 5</td>
</tr>
<tr>
<td>Normalizer</td>
<td>10</td>
</tr>
<tr>
<td>Coefficients</td>
<td>10, 8, 6, 4, 2</td>
</tr>
<tr>
<td>Smoothness</td>
<td>good</td>
</tr>
<tr>
<td>Sharpness</td>
<td>poor</td>
</tr>
<tr>
<td>Isotropy</td>
<td>fair</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPS file name</td>
<td>FWGS13.TBL</td>
</tr>
<tr>
<td>Functional form</td>
<td>[ h(n) = \frac{1}{\sqrt{2\pi}2.6} e^{-0.5n^2/2.6^2}, \ n &lt;&lt; 6 ]</td>
</tr>
<tr>
<td>Normalizer</td>
<td>27</td>
</tr>
<tr>
<td>Coefficients</td>
<td>21, 20, 16, 11, 6, 3, 1</td>
</tr>
<tr>
<td>Smoothness</td>
<td>average</td>
</tr>
<tr>
<td>Sharpness</td>
<td>poor</td>
</tr>
<tr>
<td>Isotropy</td>
<td>good</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name</th>
<th>Cubic b-spline</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPS file name</td>
<td>FW3SPL.TBL</td>
</tr>
</tbody>
</table>
| Functional form | \[ h(n) = \frac{2}{3} - (n/5)^2 + 0.5(n/5)^3, \ n << 5 \]  
\[ = 4/3 - 2(n/5) + (n/5)^2 - (n/5)^3/6, \ 5 < n < 10 \] |
| Normalizer    | 10000                           |
| Coefficients  | 667, 631, 539, 415, 283, 167, 85, 36, 11, 1 |
| Smoothness    | good                            |
| Sharpness     | poor                            |
| Isotropy      | average                         |
figure 5.2, continued

Name: Multiple lobe sharpened cubic b-spline
IPS file name: FWSHCB.TBL (preceding FW3SPL.TBL)
Functional form: (see section IIIA.)
Normalizer: 1000
Coefficients: 1733,0,0,0,0,-464,0,0,0,125,0,0,0,0,
-33,0,0,0,0,9,0,0,0,0,-3
Smoothness: average
Sharpness: good
Isotropy: average

Name: Truncated sinc
IPS file name: FWSINC.TBL
Functional form: h(n) = (5/πn) sin(πn/5) , n =< 50
Normalizer: 1000
Coefficients: 1000,935,757,505,234,0,-156,-216,-189,
-104,0,85,126,116,57,0,-58,-89,-84,
-49,0,45,69,66,39,0,-36,-56,-54,-32,0,
30,47,46,28,0,-25,-41,-40,-24,0,23,36,
35,21,0,-20,-32,-32,-19
Smoothness: poor
Sharpness: good
Isotropy: average

Name: Single lobe sharpened cubic b-spline
(TRW cubic spline)
IPS file name: FWTRWS.TBL
Functional form: h(n) = 1 - 2(n/5)^2 + (n/5)^3 , n =< 5

= 4 - 3(n/5) + 5(n/5)^2 - (n/5)^3 , 5 < n < 10
Normalizer: 1000
Coefficients: 1000,928,744,496,232,0,-128,-144,-96,-32
Smoothness: good
Sharpness: good
Isotropy: average

Name: Single lobe sharpened gaussian
IPS file name: FWSHG.TBL (preceding FWGSL3.TBL)
Functional form: (see section IIIA.)
Normalizer: 10
Coefficients: 14,0,0,0,0,-2
Smoothness: average
Sharpness: good
Isotropy: good
figure 5.3a Sample and hold frequency and impulse responses

vertical range is 20.00000
horizontal range is 3.14159
figure 5.3b Bilinear frequency and impulse responses
figure 5.3c Gaussian frequency and impulse responses
Figure 5.3d: Cubic b-spline frequency and impulse responses.
figure 5.3e Multiple lobe sharpened b-spline frequency and impulse responses.
figure 5.3f Truncated sinc
frequency and impulse responses
figure 5.3g Single lobe sharpened b-spline frequency and impulse responses.
figure 5.3h Single lobe sharpened gaussian frequency and impulse responses
indicates positive values, and darker represents negative values. Sample-and-hold and bilinear interpolation are the two commonest techniques, in which the interpolating functions are, respectively, square blocks and tent-like pyramids. These two are very simple to implement computationally, but their performance in the spatial and frequency domains is relatively poor. The frequency responses make clear that although both have a zero at the replication frequency (meaning a d.c. picture in gives a d.c. picture out), the responses up to the cutoff frequency don't remain near the d.c. value, and the responses past the cutoff frequency are not consistently close to zero. The contours of the impulse responses point out the directionality in the spatial domain. The gaussian and cubic b-spline functions have much better spatial characteristics. The former is absolutely circularly symmetric and relatively smooth, while the latter is absolutely smooth (i.e. continuous) in value, first derivative, and second derivative, and relatively circularly symmetric. Although mathematically distinct, the functions look very similar. In the frequency domain, the responses are much better at the cutoff frequency and beyond. However, in the pass band, the responses still fall off from the d.c. value rather quickly. Two alternatives are found in the frequency responses for the sharpened cubic b-spline filter and the truncated sinc filter. The sharpened cubic b-spline filter was obtained by removing the intersymbol interference of the cubic b-spline function with an 11 point sharpening filter via the method developed
in section III. The truncated sinc filter was obtained by windowing the ideal LPF impulse response with a 101 point rectangle function. Because the response in the passband for these filters is close to the d.c. value, the resulting sharpness is relatively good. However, this is obtained at the expense of ringing in the impulse response -- a particularly annoying phenomenon. In the case of an 11 point presharpening filter, the net impulse response has 5 lobes. However, not all lobes are undesirable. As noted in section II, a single lobe overshoot or undershoot is a very natural visual phenomenon (recall the simultaneous contrast phenomenon, figure 2.5.) The final two filters, the sharpened gaussian and 3-point sharpened cubic b-spline (the "TRW cubic spline"[26]), take advantage of this fact. The sharpening effect of their frequency responses is faintly reminiscent of that of the visual system (figure 2.6). Both filters combine a reasonable low pass frequency characteristic with the desired single lobe impulse response.

Figures 5.4a-h and 5.5a-h show the results of processing the two test pictures with the set of eight filters. The varying edge effects are artifacts of the IPS filtering process, and reflect the different spatial widths of the impulse responses (small for the sample-and-hold filter, large for the truncated sinc filter.) The single letter labels on each picture were randomly chosen and used to identify the pictures during the subsequent experimentation.

Subjective quality judgments were undertaken in a manner
figure 5.4a. Interpolated CMAN
Subsample spacing: 5
Filter: sample and hold
figure 5.4b. Interpolated CMAN
Subsample spacing: 5
Filter: bilinear
figure 5.4c. Interpolated CMAN
Subsample spacing: 5
Filter: gaussian
figure 5.4d. Interpolated CMAN
Subsample spacing: 5
Filter: cubic b-spline
figure 5.4e. Interpolated CMAN
Subsample spacing: 5
Filter: multiple lobe sharpened b-spline
figure 5.4f. Interpolated CMAN
Subsample spacing: 5
Filter: truncated sinc
figure 5.4g. Interpolated CMAN
Subsample spacing: 5
Filter: single lobe sharpened b-spline
figure 5.4h. Interpolated CMAN
Subsample spacing: 5
Filter: single lobe sharpened gaussian
figure 5.5a. Interpolated RWEDGE
Subsample spacing: 5
Filter: sample and hold
figure 5.5b. Interpolated RWEDGE
Subsample spacing: 5
Filter: bilinear
figure 5.5c. Interpolated RWEDGE
Subsample spacing: 5
Filter: gaussian
In the vacation of I
armed by the quiet.
I have equalized my
ABCDEFGHJKLM

figure 5.5d. Interpolated RWEDGE
Subsample spacing: 5
Filter: cubic b-spline
figure 5.5e. Interpolated RWEDGE
Subsample spacing: 5
Filter: multiple lobe sharpened b-spline
In the vocation of the gained by the quiet.
I have equalized my jo

figure 5.5f. Interpolated RWEDGE
Subsample spacing: 5
Filter: truncated sinc
figure 5.5b. Interpolated RWEDGE
Subsample spacing: 5
Filter: single lobe sharpened gaussian
similar to that of the first experiment. In this case, however, an entire set of 8 pictures was presented to be rank ordered, rather than a subset of 5 pictures at a time. Again, an individual picture received 0 points for being the worst in the group, and the maximum, 7, for being the best. Because fewer picture comparisons were involved, this experiment generally proceeded more quickly than the first, and more subjects, a total of 12, participated.

The combined results for the two sets of pictures are plotted in figure 5.6. In the figure, the solid bars represent the main values, while the vertical ranges represent ±1 standard deviation of error. For both CMAN and RWEDGE, the single lobe sharpened gaussian was the preferred method of interpolation. The single lobe sharpened spline, similar in shape, also scored above average. In the case of CMAN, the multiple lobe sharpened cubic b-spline worked fairly well, probably because of its good sharpness. However, in the much more structured RWEDGE image, its ringing produced apparently objectionable artifacts. In the case of RWEDGE, the ordinary gaussian gave good results, probably due to its excellent isotropy. In the CMAN image, however, isotropy is less necessary than sharpness, and the gaussian did not perform as well. The sinc function worked very much like the multiple lobe cubic b-spline: its sharpness was used to some advantage in CMAN, but its ringing hurt in RWEDGE. Of the remaining filters, the sample-and-hold technique scored particularly poorly. Otherwise, the filter effects were gen-
Figure 5.6
Relative Quality of Interpolated Pictures
erally regarded as indistinguishable, as evidenced by their approximately equal means and wide variances.
VI. Experimental Investigation of

Hexagonal vs. Rectangular Sampling

The third and final experiment was undertaken to determine any subjective advantage gained by the use of a hexagonal rather than standard rectangular sampling raster.

A good theoretical argument can be made for the hexagonal case. [27] Given an input with a finite two-dimensional bandwidth, or "region of support", it is desirable to pack the replications of figure 3.2b as closely together as possible in the frequency plane without overlapping. In doing this, the parameters a, b, and c are maximized, which reduces the number of subsamples needed for a given size original. Viewed another way, given a constant number of subsamples, it is desirable to arrange the replications in the frequency plane so that the largest possible input region of support can be handled. The amount of "uncovered" area left in the frequency plane can be used as a measure of the subsampling strategy's efficiency. The strangely shaped region of support of figure 3.2 is rather artificial. Consider an input with an isotropic, or circularly symmetric, region of support. When subsampling is done in the most common way, on a Cartesian grid, the replications are arranged in orthogonally stacked rows and columns. In this case 78.5 percent of the frequency plane area is covered. If the
circular replications are packed as tightly as possible in the frequency plane, they form a hexagonal pattern, which corresponds (by duality) to a hexagonal subsampling raster. In this case a maximum amount of the frequency plane is covered by the circles, 90.7 percent, a 12.2 percent improvement over the cartesian case.

One immediately obvious problem involved in hexagonal subsampling is that given a very high resolution rectangularly sampled image, it is difficult to extract a hexagonal pattern for small reduction factors. For example, with reference to figure 3.2a, if the desired reduction along the m (horizontal) direction is 4, then the proper hexagonal parameters are a=4, b=2, and c=3.464. The best approximation for c, 3, results in a 13% error. For other small reduction factors:

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>ideal c</th>
<th>c</th>
<th>% error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>1</td>
<td>1.732</td>
<td>2</td>
<td>15.5</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>3.464</td>
<td>3</td>
<td>13.4</td>
</tr>
<tr>
<td>6</td>
<td>3</td>
<td>5.196</td>
<td>5</td>
<td>3.8</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>6.928</td>
<td>7</td>
<td>1.</td>
</tr>
<tr>
<td>10</td>
<td>5</td>
<td>8.66</td>
<td>9</td>
<td>3.9</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>10.39</td>
<td>10</td>
<td>3.75</td>
</tr>
<tr>
<td>14</td>
<td>7</td>
<td>12.12</td>
<td>12</td>
<td>0.99</td>
</tr>
<tr>
<td>16</td>
<td>8</td>
<td>13.85</td>
<td>14</td>
<td>1.1</td>
</tr>
<tr>
<td>18</td>
<td>9</td>
<td>15.58</td>
<td>16</td>
<td>2.6</td>
</tr>
</tbody>
</table>

The raster with a=8 was chosen for the experiment because: 1) it was very close to the ideal hexagonal raster, 2) it represented a reduction factor similar to that used in the previous experiments, and 3) the corresponding rectangular raster with an identical number of points per unit area was not too anamorphic. The detailed structures of these two rasters are diagrammed in
figure 6.1. Both are composed of rows of samples 7 lines apart. In the case of the rectangular raster, these rows are repeated identically, whereas for the hexagonal raster, every other row is offset.

Both RWEDGE and CMAN were used as test pictures for this experiment. They were first lowpassed with a gaussian filter with optimum standard deviation as defined by the results of the first experiment (.30 times the sample spacing of 8). After subsampling on the hexagonal and rectangular grids, two different filters were used for interpolation. The first of these was a gaussian, with a standard deviation equal to the optimum result of the first experiment. This filter was designed to be slightly elliptical in shape to account for the anamorphism of the rectangular grid -- the detailed values are in figure 6.2. As expected, the technique resulted in some ripple in the d.c. areas of the pictures. The second interpolation filter which was utilized was the preferred method of the second experiment, the sharpened gaussian. The output of this technique was ripple free because it was also made slightly elliptical to accommodate the anamorphism (see figure 6.2). Figure 6.3a-d shows the CMAN picture processed by the two rasters and the two interpolation filters; figure 6.4a-d does the same for RWEDGE.

The subjective evaluation involved was the simplest of the three experiments. Subjects were shown the four pairs of pictures, each member of the pair having been processed identically except for the sampling raster, and asked to select the
figure 6.1 Hexagonal and Rectangular Subsampling Rasters

[Diagram showing hexagonal and rectangular subsampling patterns]

figure 6.2 Experimental Interpolation Filter Values

<table>
<thead>
<tr>
<th>Name</th>
<th>75% Gaussian</th>
<th>Sharpened Gaussian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sub-sample spacing:</td>
<td>8.</td>
<td>7.</td>
</tr>
<tr>
<td>D.C. gain:</td>
<td>113.25</td>
<td>113.7</td>
</tr>
<tr>
<td>IPS file name:</td>
<td>G23.TBL</td>
<td>G21.TBL</td>
</tr>
<tr>
<td>Coefficients, h(0):</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>h(1), h(-1):</td>
<td>95</td>
<td>95</td>
</tr>
<tr>
<td>h(2), h(-2):</td>
<td>86</td>
<td>82</td>
</tr>
<tr>
<td></td>
<td>71</td>
<td>64</td>
</tr>
<tr>
<td></td>
<td>54</td>
<td>45</td>
</tr>
<tr>
<td></td>
<td>39</td>
<td>29</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>17</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>4</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td></td>
<td>0...</td>
<td></td>
</tr>
</tbody>
</table>

- 102 -
figure 6.3a. Processed CMAN
Subsample raster: hexagonal 8 x 7
Filter: 75% gaussian
figure 6.3b. Processed CMAN
Subsample raster: rectangular 8 x 7
Filter: 75% gaussian
figure 6.3c. Processed CMAN
Subsample raster: hexagonal 8 x 7
Filter: sharpened gaussian
figure 6.3d. Processed CMAN
Subsample raster: rectangular 8 x 7
Filter: sharpened gaussian
In the vacation of 1
correlations the quest.

A B C D E F G H I J K L M

figure 6.4b. Processed RWEDGE
Subsample raster: rectangular 8 x 7
Filter: 75% gaussian
figure 6.4c. Processed RWEDGE
Subsample raster: hexagonal 8 x 7
Filter: sharpened gaussian
In the location of u
the grafted by the quiet.
I have equalized our of
ABCD EFGHIJKL

figure 6.4d. Processed RWEDGE
Subsample raster: rectangular 8 x 7
Filter: sharpened gaussian
The results of this procedure were surprising in their unanimity (predicted in part by the very quick rate at which most of the subjects made their decisions). In the case of the gaussian filter, which produced output ripples in d.c. areas, the hexagonal raster was preferred by every subject. This is most probably because the hexagonal raster minimized both the maximum sample separation and ripple amplitude. The effect is illustrated in figure 6.5, which lists the output values for the gaussian filtering given a d.c. input of 1.0. In the case of the sharpened gaussian filter, which produced no ripples in d.c. areas, the choice of raster was very much correlated with picture subject matter. For RWEDGE, the rectangular raster was preferred, probably because it resulted in relatively good edges along the many horizontal and vertical bar shapes in the text segment. While the rectangular raster permits sharp edges along horizontal and vertical directions which are orthogonal, the hexagonal raster permits them along the horizontal and two vertical directions which are skewed by 30 degrees from orthogonal. Typically, text segments contain more features comprised of orthogonal bars than those with this 30 degree slant, and the rectangular raster results in better reproduction. In the frequency domain, if a picture has many horizontal and vertical edges, its region of support is rectangular, and a rectangular stacking in the frequency plane works better than a hexagonal one (figure 3.2). In the case of CMAN, the hexagonal raster was
figure 6.5  Processed Output Values for D.C. Input

Input = 1.0
Sample points indicated by circles

Rectangular raster:

<table>
<thead>
<tr>
<th>1.08</th>
<th>1.07</th>
<th>1.04</th>
<th>1.01</th>
<th>.99</th>
<th>1.01</th>
<th>1.04</th>
<th>1.07</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.06</td>
<td>1.05</td>
<td>1.02</td>
<td>1.00</td>
<td>.98</td>
<td>1.00</td>
<td>1.02</td>
<td>1.05</td>
</tr>
<tr>
<td>1.03</td>
<td>1.02</td>
<td>.99</td>
<td>.97</td>
<td>.95</td>
<td>.97</td>
<td>.99</td>
<td>1.02</td>
</tr>
<tr>
<td>1.00</td>
<td>.99</td>
<td>.96</td>
<td>.95</td>
<td>.93</td>
<td>.95</td>
<td>.96</td>
<td>.99</td>
</tr>
<tr>
<td>1.00</td>
<td>.99</td>
<td>.96</td>
<td>.95</td>
<td>.93</td>
<td>.95</td>
<td>.96</td>
<td>.99</td>
</tr>
<tr>
<td>1.03</td>
<td>1.02</td>
<td>.99</td>
<td>.97</td>
<td>.95</td>
<td>.97</td>
<td>.99</td>
<td>1.02</td>
</tr>
<tr>
<td>1.06</td>
<td>1.05</td>
<td>1.02</td>
<td>1.00</td>
<td>.98</td>
<td>1.00</td>
<td>1.02</td>
<td>1.05</td>
</tr>
</tbody>
</table>

Hexagonal raster:

<table>
<thead>
<tr>
<th>1.08</th>
<th>1.07</th>
<th>1.04</th>
<th>1.02</th>
<th>.99</th>
<th>1.02</th>
<th>1.04</th>
<th>1.07</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.06</td>
<td>1.05</td>
<td>1.02</td>
<td>1.00</td>
<td>.98</td>
<td>1.00</td>
<td>1.02</td>
<td>1.05</td>
</tr>
<tr>
<td>1.02</td>
<td>1.01</td>
<td>.99</td>
<td>.98</td>
<td>.96</td>
<td>.98</td>
<td>.99</td>
<td>1.01</td>
</tr>
<tr>
<td>.97</td>
<td>.98</td>
<td>.96</td>
<td>.96</td>
<td>.96</td>
<td>.95</td>
<td>.96</td>
<td>.98</td>
</tr>
<tr>
<td>.96</td>
<td>.96</td>
<td>.96</td>
<td>.98</td>
<td>.97</td>
<td>.98</td>
<td>.96</td>
<td>.96</td>
</tr>
<tr>
<td>.96</td>
<td>.98</td>
<td>.99</td>
<td>1.01</td>
<td>1.02</td>
<td>1.01</td>
<td>.99</td>
<td>.98</td>
</tr>
<tr>
<td>.98</td>
<td>1.00</td>
<td>1.02</td>
<td>1.05</td>
<td>1.06</td>
<td>1.05</td>
<td>1.02</td>
<td>1.00</td>
</tr>
</tbody>
</table>
preferred by almost all subjects. Again, this is a manifesta-
tion of matching the sampling raster to the frequency content of
the picture. CMAN has a less directional content than RWEDGE,
i.e. its region of support is more circular, so the hexagonal
raster can be used to advantage.
VII. Discussion and Conclusions

The three experiments described in the previous sections have investigated a number of the issues involved in PCM picture coding. Operations in the continuous spatial domain were simulated by sampling the test pictures involved on a very high resolution discrete grid, then processing them digitally. In general this processing consisted of three phases: bandlimiting the very high resolution input, subsampling it to obtain the desired resolution digitized image, and interpolating this back into a very high resolution approximation to the original.

The first experiment examined the tradeoff between space-width and frequency-width which arises in both the PCM prefiltering and postfiltering operations. Gaussian low pass filters were used for bandlimiting and interpolation, and a metric which characterized their frequency-width was taken as the percentage of filter energy within the passband implied by the subsampling frequency. For a subsampling frequency of 5, the value which was used, the input spectrum was replicated in the frequency plane at multiples of 2\pi/5, and the resulting passband was the range 0 to \pi/5 (similar to the case of figure 3.2). Each member of the family of gaussian filters which was used was accordingly characterized by its percentage of energy within this band (illustrated in figure 3.3). Subjective quali-
ty was measured experimentally as a function of the metric, and
regression techniques were used to fit a smooth curve to the
resulting data (figure 4.5). Most psychophysical experiments,
by their nature, tend to have results which are noisy. With
this in mind, the standard errors for this experiment, denoted
by black vertical bars in the raw data, are not unreasonable --
in general the standard deviations are below 25% of the mean
values. Moreover, if two surfaces are imagined which connect,
respectively, the upper and lower endpoints of each error bar,
they both exhibit the same general curvatures of the sample mean
curve. The second order regression fit, which fortifies the
trends of the raw data, is a good one ($R^2 = .85$ in the range 0-1),
however the exact optimum metric values which it provides, 60%
and 75% for the bandlimiting and interpolating filters, respec-
tively, should not be taken with too many significant digits.
In the the space domain, these numbers correspond to prefilter
and postfilter widths of .30 and .375 subsample spacings.
Graphically, the resulting filters are shown in figure 7.1. In
the physical sense, this plot might represent, for example, the
intensity cross-section of the beam in a laser scanner/recorder
relative to the scan line separation.

Always significant is the question regarding the gen-
erality of the experimental results. The exact numbers as
determined here are theoretically valid only for the single con-
trolled viewing distance and test picture which were used. An
obvious suggestion for further research is to repeat the experi-
figure 7.1 Optimum Gaussian Pre- and Post-filters
ment for a variety of different viewing distances, pictures, and subsampling frequencies. However, the work involved in producing a full new set of 25 carefully controlled test pictures, and the subsequent data logging at different viewing distance, is prohibitive for more than a very few cases. Of the two filters involved, the result for the bandlimiting filter is more likely to vary with the specific picture content than that for the interpolating filter. This is because the degree of aliasing resulting from the prefiltering operation depends not only on the prefilter and subsampling frequency, but on the picture frequency content as well. The exact amount of aliased energy results from the product of the picture's Fourier transform and the bandlimiting filter transfer function. When coding a specific picture, its transform could be taken in order to optimize the bandlimiting filter. However, this is out of the range of the limitations discussed in the first section. Over the ensemble of continuous tone images, the autocorrelation function is typically approximated with a decaying exponential, resulting in a simple single pole power spectrum. Assuming that this spectrum remains somewhat flat out to the edge of the passband, to a first approximation the aliased energy depends mostly on the bandlimiting filter. The numbers along the "increasing prefilter width" axis in figure 4.5 suggest that this is the case for the CMAN image used in the experiment. If the energy in the picture were concentrated in frequencies much lower than the edge of the passband, the various frequency-
widths of the filters would have little effect. Most real world continuous tone images are similar to CMAN in not having large periodic components at high spatial frequencies, and would probably perform in much the same way. For line art images, however, the higher frequencies make aliasing more of an issue, and the results might be different. This could form the basis of a further investigation.

As far as the ordinary gaussian interpolating filter is concerned, the conclusion to be made from the first experiment, simply stated, is that people prefer sharp-but-slightly-ripply pictures to smooth-but-slightly-blurred ones. The result is consistent with experiments which have traded off sample resolution with #bits/pel, in which subjects preferred noisy high resolution pictures to clean low resolution ones.[28]

The second phase of the first experiment involved a comparison of the determined subjective quality measure with an objective one which has been suggested in the literature. As the plots of figure 4.6 indicate, the weighted and unweighted square error criteria have similar shapes, but very different from that of the subjective measure. In both the unweighted and weighted case, the noise power is much more influenced by the postfilter than the prefilter. This is because the effects of the noise-inducing prefiltering are greatly masked by the subsampling operation: except when a subsample point falls directly on a high frequency area which has been blurred by the prefiltering, the subsampled value is essentially the original value at that
point. Although it doesn't show up well in the figure, the slope along the prefilter axis is small but negative. Along the postfilter axis, the weighted and unweighted square error are correlated with the amplitude of the ripple resulting from the filtering operation. This is as it should be from a computational point of view. The general effect of the weighting function (the bandpass visual transfer function, figure 2.6), is to magnify the fourier components of the ripple waveform. This occurs to a greater degree when the waveform has a large amplitude than a small one, most likely because at the viewing distance at which the experiment was performed, the fundamental component is located close to the maximum in the visual transfer function (meaning the amount of weighting depends mostly on the amplitude of the fundamental).

It is clear that the weighted square error criterion fails as an objective predictor of picture quality because it is too simple, but few alternatives (none that work well) have been suggested in the literature. Some notable attempts are contained in references [29]-[32]. However, current psychological and neurobiological research is beginning to understand how the later stages of the visual system work. Only after these processes are better understood will better models become possible.

The second experiment examined the effects of a variety of interpolation filters, each of which traded off smoothness, sharpness, and isotropy to different degrees. The single lobe
sharpened gaussian was found to maximize output quality given a prefiltered and subsampled input. Significantly, it was the preferred technique for both the continuous tone test picture and the line art test picture which were used, which suggests the generality of the result. Besides the clear superiority of the sharpened gaussian, the statistics shown in figure 5.5 also make the point that many of the other techniques which were tried were often considered indistinguishable by observers (i.e. large variances in quality with respect to the mean). In particular, the truncated sinc function, which required a large amount of computation, did not perform significantly better than the simple bilinear filter. The simple-minded sample-and-hold technique also performed much worse than any other filter, and can't accordingly be considered a reasonable method for interpolating pictures. As far as computation time is concerned, the sharpened gaussian also has the advantage that it is the convolution of two separate, separable functions, a presharpen followed by an all-positive interpolated. When filtering, the presharpen can be applied to the lower sample density image, before it is expanded in the space plane for interpolation, rather than after. On IPS, this accounted for time savings of about 30% relative to filters with comparable sized impulse responses which cannot be split into these two stages.

The important result of this experiment is not as much the exact functional form of the sharpened gaussian as much as the general shape of its impulse and frequency responses. Such
a filter is a member of a family of filters which can be con-structed by presharpening an all-positive basis interpolation function -- a 2N+1 point presharpening filter resulting in a net impulse response with N lobes. An interesting investigation might be the systematic evaluation of the first n members of such a family, starting with a gaussian or bilinear function. Any resulting single lobe function has the property that it is similar to the types of responses found naturally in the human visual system.

All of the interpolating filters used in the experiment were primarily designed with spatial characteristics in mind. However, the utility of the specialized gaussian is also suggested from the spatial frequency standpoint. Given a gaussian-like bandlimiting filter, the mid-frequency boost of the sharpened gaussian's frequency response (figure 5.3h) tends to make it look like the inverse filter to the gaussian over a frequency range close to the origin. The exact net frequency response, gaussian followed by sharpened gaussian, is plotted in figure 7.2. Also shown are the individual responses for each filter. As the figure indicates, for the net transfer function to be flat, either the interpolating filter or bandlimiting filter needs slightly boosted mid-frequencies. If this were done to the interpolating filter, turning it into the inverse gaussian filter, its nice spatial properties would probably be distorted. However, the bandlimiting filter could be slightly sharpened to achieve a flatter net response. Although it might be worth in-
vestigating whether or not this actually results in a better looking output, the gains will probably be negligible. This is due to the phenomenon mentioned above, whereby the detailed effects of the bandlimiting operation are made smaller by the resolution reduction inherent to the subsampling operation.

The final experiment investigated the subjective effects resulting from the choice of sampling raster, hexagonal versus rectangular, given identical pre- and post-filtering. A gaussian filter was used as the bandlimiting filter for all cases, while gaussian and sharpened gaussian filters were used as two distinct interpolating filters. The experimental results were nearly unanimous among the subjects, and strongly correlated with the picture content and interpolation scheme. They also vividly illustrated the effects of discrete subsampling on a given raster shown in figure 3.2. In the case of the gaussian interpolator, its frequency response did not have a zero in the frequency plane at the subsampling frequency for either the rectangular or hexagonal raster, with the result that a d.c. area in the input picture contained some ripple in the output picture. However, the ripple amplitude was less in the hexagonal case than in the rectangular case because the input spectrum replications in the frequency plane for the former are spread out as far as possible, thereby minimizing the effect. The basic result is that, if the interpolating scheme produces ripples, a hexagonal raster works better mathematically and looks better visually. This was true for both the continuous tone and line
art test pictures.

The sharpened gaussian interpolator did have a zero at the subsampling frequency for both rasters, so d.c. areas in the input picture were reproduced as d.c. areas in the output picture. In this case quality judgments were based more on aliasing effects than interpolation artifacts, and the hexagonal raster was judged to work better for CMAN, the continuous tone test picture, while the rectangular raster was judged to work better for RWEDGE, the line art test picture. This result is a manifestation of the frequency content of the two pictures, and how well the rasters can "stack them up" in the frequency plane. CMAN is relatively non-directional, like most continuous tone pictures, and his approximately circular spectrum is closest packed in the frequency plane in a hexagonal pattern. RWEDGE, on the other hand, has larger horizontal and vertical frequency components, resulting in a squarer spectrum, and so is better packed in a rectangular pattern.

It must be noted that the advantages gained by using a hexagonal raster are not free. In the case of a computer based system with human interaction, the hexagonal representation is obtained basically at the expense of a more complicated internal representation of pels (e.g. non square aspect ratio.) It is also a fundamentally unnatural way for people to think about images.

One final way to evaluate the effectiveness of the filtering techniques investigated in this thesis is to determine
the coding efficiency of the best looking output pictures relative to a yardstick of various resolution sample-and-hold originals. Versions of CMAN and RWEDGE have been produced at 1/1, 1/2, 1/3, 1/4, and 1/5 of original resolution for this purpose, and are available for general use. Comparisons indicate that for a subsampling frequency of 5, the sharpened gaussian CMAN is approximately equivalent to the 1/4 resolution original, while this version of RWEDGE is approximately equivalent to its 1/3 resolution original. This results in savings of 1 - 16/25 for CMAN and 1 - 9/25 for RWEDGE, or 36% and 64% respectively.

All of the filters in this thesis were implemented as separable convolutions, with the exception of those involved in the visual system weighted noise powers of the first experiment (figure 4.6). These filtering operations, which simulated the effect of a circularly symmetric visual transfer function (shown radially in figure 2.6) were implemented via a McClellan transform algorithm. This RADFIL process is interesting and bug-free, but runs slower than necessary because it maintains internal 32 bit precision. An interesting project would be to determine how many bits are necessary and what the resulting subjective effects on pictures are in the IPS environment.
VIII. References


Appendix A. Tone Scale Calibration Procedure

The tone scale calibration procedure was carried out in two phases. The first of these was the determination of the transformation from the input photograph, scanned via the Autokon, to the resulting computer number; the second, the determination of the transformation from computer number to the output, a Laserphoto facsimile. For the reasons discussed previously, it was desired that the computer number representation of the picture be in the lightness domain, so filtering operations could take place in that domain. Therefore, on input, the general approach was to measure the mapping from input lightness to computer number, and then apply the inverse of that mapping as a tone scale correction immediately after scanning. Similarly, on output, the functional mapping from computer number to output lightness was determined, and the inverse of that function was applied immediately before sending a picture out.

The lightnesses of the input and output pictures were evaluated using a virtual "lightness meter". Such a device didn't exist in the CIPG lab -- it was actually a combination of a real densitometer and several computer programs. Measured picture densities were converted to intensities by a relationship of the form:

\[ I = k \cdot 10^{-D} \]
Here "k" was a scaling factor used to obtain a maximum luminance value of 255 (its value was approximately 400). The luminance values obtained in this way were further converted to lightness values by the relationship discussed earlier and repeated here:

\[ A = \frac{I_{\text{max}} \cdot \log(1+aI)}{\log(1+aI_{\text{max}})} \, , \, a=0.1, \, \, I_{\text{max}}=255. \]

This table is shown below in figure A1.1.

The gray scale steps of the IEEE facsimile test chart served as the input calibrating picture. This test chart provided 16 values of input lightness, with corresponding densities between 0.13 and 1.84 on an absolute scale. These values were used to determine 16 points on the input mapping function; the remainder were obtained by linear interpolation. The chart was scanned several times, using different Autokon tone control settings each time. After low pass filtering for noise reduction, cross sections were taken through the scanned pictures to determine the resulting computer numbers. The Autokon setting which gave the best results in terms of a clean, stair-step cross section was: \( \text{dmin}=0, \, \text{dmax}=200, \, \text{mid}=+4, \, \text{sharp}=+0. \) The corresponding mapping from input lightness to output computer number is plotted below in figure A1.2. The inverse of this table, used immediately after scanning to transform the computer representation of the picture to the lightness domain, is shown in figure A1.3.
figure A1. Tone Scale Calibration Tables

fig a1.1  
input: luminance  
output: lightness

fig a1.2  
input: measured lightness  
output: computer number

fig a1.3  
input: computer number  
output: lightness

fig a1.4  
input: computer number  
output: measured lightness

fig a1.5  
input: lightness  
output: computer number
Output calibration proceeded in a similar way, utilizing a set of computer generated gray steps of value 0, 16, 32, ... as the test picture. This picture was transmitted out facsimile channel 2 and received on a Laserphoto using heavyweight paper. No tone scale adjustments were made in either the transmitting or receiving hardware. The measured resulting lightness values are plotted in figure A1.4. Its inverse, the pre-output correction table, is plotted in figure A1.5.

This procedure worked very well to match the tone scale characteristics of an original input picture to those of the final output. It was well defined, but not as automated as it could have been. Given a standard input picture with pre-measured densities, input calibration could proceed automatically using little more than current IPS processes. Output calibration requires either hand measurements of the output densities and a new IPS process, or faith in the input tone scale and computer measurement of the scanned output densities. Since only two original pictures were scanned in over the course of the investigation, the stability of the input tone scale transformation was not tested. However, a large volume of output pictures was produced over several months, and the output tone scale was consistently correct. The following two points regarding future calibration should be noted. First, it is doubtful that the two Autokons behave similarly. The extent to which either is factory calibrated is unknown. Second, the tables which were used as the post-input and pre-output
transformations are named INPUT.TBL and FAXOUT.TBL. Until any recalibration is done, these remain valid.
Appendix B. Non-Standard IPS and Unix Files

The vast majority of computer programs and files used in this investigation were standard system components. Those which were newly created or modified are listed below, along with a brief description. More thorough explanations can be obtained from IPS online help files and the program listings. At this time there are no known bugs in any program. Some of the programs which follow are more useful than others; anything that might be of possible utility to future users is included. IPS names are in upper case, Unix names in lower case.

BEEP.MAK - makes a loud noise on a selected fax machine when invoked. Useful for signalling the completion of a very long IPS process (it can be heard all over the lab).

CONTUR.MAK - allows an operator, under joystick control, to select a band of tone scales to be displayed on the tv. When an impulse response is being displayed on the tv, sweeping the joystick displays the contour levels from 0 to 255. One pot selects the center of the band being displayed (0-255), another pot selects the width of that band (1-32). To exit, all switches must be asserted.

CSECT.MAK - creates a table from the first 256 pels of the first line of a specified picture. Useful for obtaining edge cross-sections, impulse response profiles, and the
like. The table editor can be used to examine the values of the cross section in detail. The "dual" of this process is TEST, which can make a picture from a table with a given cross section.

FAXOT1.SCF - segment control file which takes a 512x512 picture and prepares it to be sent out the fax. The name of the picture must be FAXOUT.PIC, the name of the caption FAXCUT.TXT, the name of the output correction table FAXOUT.TBL. In the process the picture is toned by the correction table and flipped horizontally, the caption text is inserted at the bottom, a thin black border is put around the picture, and the whole thing is given a one inch white border.

FAXOT3.SCF - same as FAXOT1.SCF except for a 1024x1024 picture.

FAXOT5.SCF - same as FAXOT1.SCF except for a 770x1024 picture.

FAXOUT.TBL - pre-output correction table to be used when sending a computer picture in the lightness domain out fax channel 2. See appendix A for details.

FI2.TSK - a modified version of the filter task which does signed 7 bit arithmetic rather than unsigned 8 bit arithmetic. All IPS pictures have their pels in the range 0 to 255. After reading a picture line, this task converts each pel to 16 bit accuracy and subtracts 128 before further processing; when all processing is complet-
ed, it adds 128 to each pel before writing it out. The output pels are clamped at 0 and 255. Several small bugs involved in arithmetic overflow in both the standard filter task and this one were fixed.

FIL.MAK - modified to allow the operator to specify a normalizing coefficient, so the dc gain of the filtering operation could be other than one. Like all of the coefficients, the normalizer is a decimal integer, and is specified after the name of the coefficient table.

FW3SPL.MAK - a generator file for cubic b-spline filter coefficients.

FWBLIN.MAK - a generator file for bilinear filter coefficients.

FWGS13.MAK - a generator file for gaussian filter coefficients.

FWSAH.MAK - a generator file for sample-and-hold filter coefficients.

FWSHCB.MAK - a generator file for cubic b-spline presharpening filter.

FWSHG.MAK - a generator file for gaussian presharpening filter.

FWSINC.MAK - a generator file for sinc function filter coefficients.

FWTRWS.MAK - a generator file for TRW cubic spline filter coefficients.

GD.TSK - a new task which calculates the vector gradient of
a specified input picture. Originally intended but never used, it is included in this list only for completeness.

GRAPH.SCF - a segment control file which inserts labels into plots produced by PLOT.MAK. The caption should be in GRAPH.TXT, the label for the origin in 0.TXT, the label for the axes in 255.TXT. The labels are inserted directly into tv core, before the plot, by using TVPGS. After plotting as many functions as desired, the contents of the tvcore can be made into a picture using ZTV, and the result can be output on any IPS hardcopy device.

GRATE.MAK - a generator file for a resolution grating. The resulting table can be used with TEST to make useful test patterns.

GRDINT.MAK - process definition for GD.TSK. Calculates vector gradient (x component, y component, magnitude) for a specified input picture.

INPUT.TBL - post-input correction table used to put a picture scanned in on the Autokon into the lightness domain. Since the Autokon tone scale controls work for pictures coming in, this table requires a standard setting. See appendix A for details.

INSERT.MAK - inserts a specified picture into a tvu-format scratchpad. Tvu-format scratchpads are allocated when the disk is initialized. The basic effect is to segment
the picture into vertical strips 128 pels wide. These strips are then easily transposed or used to quickly update the tv display.

LIGHT.TBL - transformation table from input luminance to output lightness. Its form is given in section two.

MES.MAK - modified to allow the creation and execution of command files, useful for queuing up long routes of filtering operations.

PLOT.MAK - plots a table, with axes and graticules, on new PCTV. Other IPS processes can be used to make a hard copy. Besides the name of the table to plot, it allows the operator to specify background and video values and masks, so plotting can proceed on a bit plane basis.

plot3d - the standard Unix three dimensional plot routine, written in fortran. Several bugs in it were fixed, and code was added to plot the x and y coordinate axes.

pprogØ - a simple single dimensional plotting program written in fortran. When invoked, it takes as an argument the name of the plot file to be produced. The function to be plotted should be programmed in, as well as its input x range. The output y range is automatically scaled to fit on the paper. Another variable which can be changed is the number of graticule subdivisions which are plotted. The function currently programmed in is normalized frequency response, which asks for filter coefficient values.
pprog4 - a plotting program written in fortran which plots a perspective view of a three dimensional contour. The first argument is the name of a file with the contour values at sampled points. These are read into an array, and the contour is linearly interpolated between them. The second argument is the name of the plot file to produce. The x and y input range of the contour are programmed in. The program prompts the user for the vantage point from which the perspective display is to be viewed.

pprog5 - similar to pprog4, but it asks the user for the 6 second order coefficients and plots the function as $Ax + By + Cxy + \ldots$, rather than from sample values.

prog0 - given standard deviation sigma as input, this fortran program prints out scaled gaussian filter coefficients.

prog13 - a fortran program which generates the luminance to lightness tone scale transformation table.

prog14 - given an energy percentage and a sample spacing, this program gives the correct standard deviation for a gaussian filter.

prog15 - given a programmed-in normalized frequency response, this fortran program does a frequency-sampling type design, resulting in a zero phase FIR filter with a frequency response which approximates the desired one. There are better methods, particularly when the desired
frequency response is piecewise discontinuous, but few
which are simpler. The frequency response currently
programmed is that of the human visual system.
progl7 - a utility program which takes input densities and
types out the corresponding luminance with a desired
scale factor. Typically, the input numbers are read off
the densitometer.
progl8 - a fortran program which simulates radial transform
filtering on a small internal impulse picture. Used to
debug RADFIL, it is also useful to obtain small impulse
responses with 64 bit floating point precision.
prog20 - a fortran program which performs a simple linear
regression, first order. The first argument is the ma-
trix of independent variables, the second is the obser-
vation vector. On output, the regression coefficients
are printed out, as well as the r-squared parameter.
prog203 - similar to prog20, except it performs a second
order fit (six coefficients rather than three). It also
assumes positive observations, and takes their log be-
fore performing the fit.

RADFIL.MAK - performs a circularly symmetric filtering
operation on a specified input picture using the McClell-
lan transform algorithm.[18] Basically this algorithm
performs a two dimensional filtering given a prototype
zero phase one dimensional filter. The input parameters
are the prototype single dimensional filter coeffi-
cients, h(0), 2h(1), 2h(2), 2h(3), ..., and a normalizer. For a dc gain of one, the normalizer should be the sum h(0) + 2h(1) + 2h(2) + ....

RF.TSK - radial filter task for RAFPIL.MAK. Maintains internal 32 bit precision. Whether or not this is necessary might make an interesting investigation. The arithmetic in this task could be further optimized.

SIGN.TBL - transformation table from unsigned 8 bit pels to signed 7 bit pels.

STATZ.MAK - computes and prints out statistics (e.g. mean, variance) for a specified IPS histogram. The operator has the choice of including the endpoint values at 0 and 255 or not. If the file is a table, rather than a histogram, without a scale factor, the last number will be incorrect.

TRN.MAK - transposes a tvu-format scratchpad. Such a scratchpad can contain a picture of up to 1024x1024 pels. The basic algorithm is to transpose 128 pel square blocks within the scratchpad, while transposing the individual pels within each block (i.e. a recursive procedure).

XTRACT.MAK - extracts a picture from a tvu-format scratchpad. This is the opposite of INSERT.