THE MOMENT-PRESERVING PYRAMID FOR PROGRESSIVE IMAGE TRANSMISSION

Jeng-Jong Pan
Research and Data Systems Corporation
Greenbelt, Maryland
U.S.A.

Sheng-Tun Li
Dept. of Computer Science
University of Houston
Houston, Texas
U.S.A.

Commission II

ABSTRACT

The pyramid structure used in image encoding provides an efficient approach for progressive image transmission. The pyramid consists of a sequence of reduced-resolution images generated from an image. The original image is processed and decimated to a one-quarter of its size, the reduced image is expanded and subtracted from the original image, the difference is then encoded. Repeat the same procedure, we can generate a very samll image at the top level of the pyramid and save a set of difference images which resulted from the difference between two adjacent levels. The top level image can be transmitted rapidly due to its samll size, it is then expanded and displayed at the receiving station. The difference images are also transmitted sequentially to improve the image quality. The viewer can terminate the transmission quickly if the image is not desired. Most approximations used to compute the reduced-resolution images are based on different filters. In this paper, a moment-presevering method in a 2 x 2 moving window is proposed to generate the pyramid which can reduce the entropy and variance of the difference image at each level. Compared with Gaussian-Laplacian pyramid, it can achieve a higher compression ratio.

KEY WORDS: Pyramid, Moment, Encoding, Transmission.

INTRODUCTION

The multiresolution pyramid for progressive image transmission has been widely discussed (e.g., Anderson et al., 1984, Chin, et al., 1989, Rioul and Vetterli, 1991, Goldberg and Wang, 1991). The pyramid structure used in the image encoding provides an efficient approach for progressive image transmission. The pyramid consists of a sequence of reduced-resolution images which resulted from an image. The original image is processed (e.g., filtering) and decimated to one-quarter of its original size, the reduced image is expanded and subtracted from the original image, the difference is then coded. Repeat the same procedure, we can generate a very samll image at the top level of the pyramid and save a set of difference images which resulted from the difference between two adjacent levels. In the progressive image transmission, the top level image can be transmitted rapidly due to its samll size, it is then expanded and displayed at the receiving station. The difference images are also transmitted to improve the image quality successively. The viewer can terminate the transmission quickly if the image is not desired.

There are two concerns in the progressive image transmission using the pyramid structure. The first one is that we hope that the expanded image at the receiving station still looks like its original image, so the viewer can make a decision earlier to cease the image transmission or not. One way to evaluate the performance objectively is besed on the variance of the difference image, which provides a quantitative measurement of the approximation. The second concern is how to minimize the entropy of each difference image. Each difference image can be compressed for efficient transmission. Because the entropy of an image is the minimum bits required for lossless encoding, it is a goal for us to develop a method which can reduce the entropy of the difference image at each level.

Most approximations used in the computation of reduced-resolution images are based on different low-pass or band-pass filters (e.g., Anderson et al., 1984). In this paper, we first review the Gaussian pyramid and the corresponding Laplacian pyramid. Then a new moment-presevering approach is discussed. Several test results indicate that the new approach, compared with the Gaussian-Laplacian pyramid, can improve the two requirements mentioned above.

GAUSSIAN-LAPLACIAN PYRAMID

Let us assume that an image at the level k in the pyramid structure is $G_k(i,j)$, the reduced-resulution at the next level $k+1$ is $G_{k+1}(i,j)$, and usually the original image is referred as k=0. Burt and Adelson (1983) proposed a 5 x 5 Gaussian-like weighting function, which is equivalent to a low-pass filter, to remove the pixep-to-pixel correlations when compute the $G_{k+1}(i,j)$ from $G_k(i,j)$.

The level-to-level reduction is performed as

$$G_k(i,j) = \sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n)\ G_{k-1}(m+2i,n+2j) \quad (1)$$

where $w(m,n)$ is a two-dimensional Gaussian-like weighting function which can be generated by $w(m,n)=f(m)f(n)$, $f(m)$ is a normalized and symmetric function defined as

$$\begin{gathered}\sum_{m=-2}^{2} f(m) = 1 \\ f(m) = f(-m) \quad m=0,1,2.\end{gathered} \quad (2)$$

One of such functions which satisfies the above conditions is (Burt, 1983)

$$\begin{gathered}f(o) = c \\ f(-1) = f(1) = \frac{1}{4} \\ f(-2) = f(2) = \frac{1}{4} - \frac{c}{2}.\end{gathered} \quad (3)$$

where c is a constant and can be adjusted to minimize the entropy and variance of the Laplacian images which will be discussed later.

If the dimension of $G_{k-1}(i,j)$ is $2^M+1 \times 2^N+1$, then the dimension of $G_k(i,j)$ is reduced to $2^{M-1}+1 \times 2^{N-1}+1$, i.e., the resolution at level k is always reduced from the level k-1 by 2 in each dimension. The sequential images computed from this approach is organized as a pyramid, which is called Gaussian Pyramid, and the lower level represents a higher resolution. The $G_k(i,j)$ can be expanded to the same dimension as $G_{k-1}(i,j)$ through an appropriate interpolation. Let us assume that the expanded image from $G_{k+1}(i,j)$ is $G_k'(i,j)$,

$$G'_k(i,j) = 4\sum_{m=-2}^{2}\sum_{n=-2}^{2} w(m,n)\; G_{k+1}([\frac{i-m}{2}],[\frac{j-n}{2}]) \quad \textbf{(4)}$$

where only the pixels which make $[x]$ to be an integer are computed in the sum.

The difference image, which is called the Laplacian image if the Gaussian-like weighting function is used in eqs. 1 and 4, $L_k(i,j)$ at the level k is defined as

$$L_k(i,j) = G_k(i,j) - G'_k(i,j) \quad \textbf{(5)}$$

where $L_k(i,j)$ is also recognized as the prediction error, which usually has small dynamic range, and the entropy of $L_k(i,j)$ could be less than that of $G_k(i,j)$. Therefore, the encoding of the image $G_0(i,j)$ is replaced by encoding $L_k(i,j)$, $k = 0, 1, 2, .., L-1$ and $G_L(i,j)$. The set of Laplacian images are organized as the Laplacian pyramid. One potential problem in the Laplacian pyramid generation is how to select the optimal coefficient c in eq. 3, so that the entropy and variance of each Laplacian image have the minimal values. Burt and Adelson (1983) found that c=0.6 was an optimal value in their test. The same value of c is also used in the following images test.

In the next section, a moment-preserving processing is used to replace the Gaussian-like weighting function in the Gaussian-Laplacian pyramid generation. Rather than smoothing the pixels through a 5 x 5 window, the first and second moments in a 2 x 2 window at level k are still preserved maximally in the corresponding pixel at level $k+1$.

MOMENT-PRESERVING PYRAMID

Moment-preserving method has been implemented recently in various image processing, such as image segmentation (Tsai, 1985), optical character recognition (Cash and Hatamian, 1987), image enhancement (Chen and Tsai, 1988), and quantization (Delp and Mitchell, 1991). The rth moment m_r of pixels $p(i,j)$ in a 2 x 2 window is defined as

$$m_r = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} p^r(i,j) \qquad r=1,2,3,... \quad \textbf{(6)}$$

We use these four pixels at level k to determine a new pixel q at level k+1. Let us consider $m_{1,k}$ and $m_{2,k}$ at level k,

$$m_{1,k} = \frac{1}{4}(p(1,1)+p(1,2)+p(2,1)+p(2,2))$$
$$m_{2,k} = \frac{1}{4}\sum_{i=1}^{2}\sum_{j=1}^{2} p^2(i,j) \quad \textbf{(7)}$$

and $m_{1,k+1}$ and $m_{2,k+1}$ at level k+1,

$$m_{1,k+1} = q$$
$$m_{2,k+1} = q^2 \quad \textbf{(8)}$$

The new pixel q is determined by minimizing an objective function, which is defined as

$$minimize \quad a(m_{1,k} - m_{1,k+1})^2 + b(m_{2,k} - m_{2,k+1})^2 \quad \textbf{(9)}$$

where a and b are two weighting coefficients which can be used to emphasize the first or the second moment in the pyramid generation. Here we assume that both a and b are 1. In this case, the second moment could have more impact on the minimization. If b=0, then the moment-preserving method will be equivalent to the mean pyramid (Goldberg and Wang, 1991).

The solution for eq. 9 can be simplified through the least-squares approximation. It is equivalent to find the root of the following equation,

$$2q^3 + (1-2m_{2,k})q - m_{1,k} = 0 \quad \textbf{(10)}$$

There should have three roots in eq. 10. If there is only one real-value root, then it will replace q, otherwise, the root with real value which is closest to the $m_{1,k}$ is selected. The value is truncated to the nearnest integer before replace pixel q. The first and the second moments of the reduced-resolution image is optimally preserved at each level in the pyramid. The image at level k can be expanded either using the nearest neighbor interpolation, that repeats the same gray value in the expanded 2 x 2 window, or the Gaussian-like interpolation function (eq. 10). The difference image at each level is obtained through the same approach as defined in eq. 5.

In order to evaluate the image transmission efficieny and quality, the impact on the entropy and variance of difference images using these two different interpolation methods is discussed in the next section. However, the interpolation method used at the receiving station could be irrelative to that used in the compression procedure.

ENTROPY AND VARIANCE COMPARISONS

In order to compare the performances of progressive image transmission using the Gaussian-Laplacian pyramid and the moment-preserving pyramid, three images appeared in most image analysis, 'Lena', 'Cameraman', and 'City', are used here for the experimental evaluation of entropy and variance at each level.

Figures 1 and 2 show these three original images and the corresponding histograms, respectively. The size of each original image is 256 x 256 pixels, but it is expanded to 257 x 257 by repeating the rightmost column and the lowest row in the Gaussian-Laplacian pyramid computation. The entropy of image is 7.58 (Lena), 7.01 (Cameraman), and 7.34 (City). The histograms indicate that the dynamic range of pixels is wide enough to segment several objects.

Figures 3 and 4 show the Gaussian-Laplacian pryamid and the moment-presevering pyramid of image 'Lena', respectively. Visually, the reduced-resolution images generated from both methods still preserved most significant features, which include the tone and shapes of major objects. The entropy and variance of each image at each level using these two methods are compared (Tables 1 to 3). The two different expansion methods, nearest neighbor and Gaussian-like expansions, are also included in the comparsions.

Tables 1 to 3 list the entropy of each test images at the first five levels, respectively. The values in the last column in tables 1 to 3 corresponds to the entropy of the reduced-resoultion image, and values in orther four columns are entropy of difference images at the first four levels. Apparently the entropy of the difference image obtained from the moment-preserving method with the nearnest neighbor expansion has the lowest entropy, and which is around 15% to 25% less than that from the Gaussian-Laplacian pyramid method.

Tables 4 to 6 list the variance of each difference image at the first five levels, respectively. Again, the variance of the difference image computed for the moment-preserving method with the nearest neighbor expansion has the minimal value, and which is around 30% to 60% less than that from the Gaussian-Lapalcian method. The lower variance could reflect both the higher image quality and the lower tone variation.

The comparsions indicate that the moment-preserving pyramid with nearest neighbor expansion can speed the progressive image transmission by reducing both the entropy and variance of difference images. The transimitted images from level by level can be used to reconstruct its origianl image without any error if no quantization is implemented on the difference image and no other thansmission noise is invloved. Entropy can be further reduced if the pixels of the difference image at each level are quantized, however, at the expense of sacrifice image quality.

CONCLUSIONS

A moment-preserving pyramid is proposed to speed the progressive image transimission. Its performance, based on the entropy and variance criteria, is better than that of the Gaussian-Laplacian pyramid. Because the entropy and variance are reduced around 20% and 45%, respectively, the moment-preserving pyramid requires less bits in encoding of the difference images, but still preserves higher image quality at each level. The moment-presevering pyramid works on both lossless encoding and lossy encoding if quantization is involved.

REFERENCES

Anderson, E. H., Anderson, C. H., Bergen, J. R., Burt, P. J., and Ogden, J. M., 1984. Pyramid methods in imagfe processing. RCA Engineer, 29(6):33-41.

Burt, P. J. and Adelson, E. H., 1983. The Laplacian pyramid as compact image code. IEEE Trans. on Communications, 31(4):532-540.

Cash, G. and Hatamian, M., 1987. Optical character recognition by the method of moments. Computer Vision, Graphics, and Image Processing, 39:291-310.

Chen, L. H. and Tsai, W. H., 1988. Moment-preserving sharpening - A new approach to digital picture deblurring. Computer Vision, Graphics, and Image Processing, 42:1-13.

Chin, F., Choi, A., and Luo, Y., 1989. An optimal generating kernel for image pyramids by linear fitting. International Symposium on Computer Architecture and Digital Signal Processing, Hong Kong, 612-617.

Delp, E. J. and Mitchell, O. R., 1991. Moment preserving quantization. IEEE Trans. on Communications, 39(11):1549-1991.

Goldberg, M. and Wang, L., 1991. Comparative performance of pyramid data structures for progressive image transmission. IEEE Trans. on Communications, 39(4):540-548.

Rioul, O. and Vetterli, M., 1991. Wavelets and signal processing. IEEE Signal Processing Magazine, 8(4):14-38.

Tsai, W. H., 1985. Moment-preserving thresholding: A new approach. Computer Vision, Graphics, and Image Processing, 29:377-393.

(a)

(b)

(c)

Figure 1. Three test images. (a) Lena, (b) Cameraman, and (c) City.

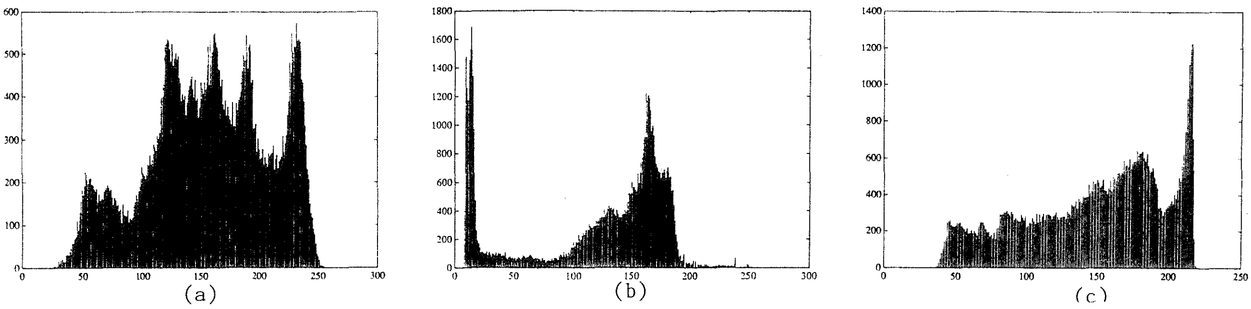

Figure 2. The corresponding histograms of Fig. 1. (a) Lena, (b) Cameraman, and (c) City.

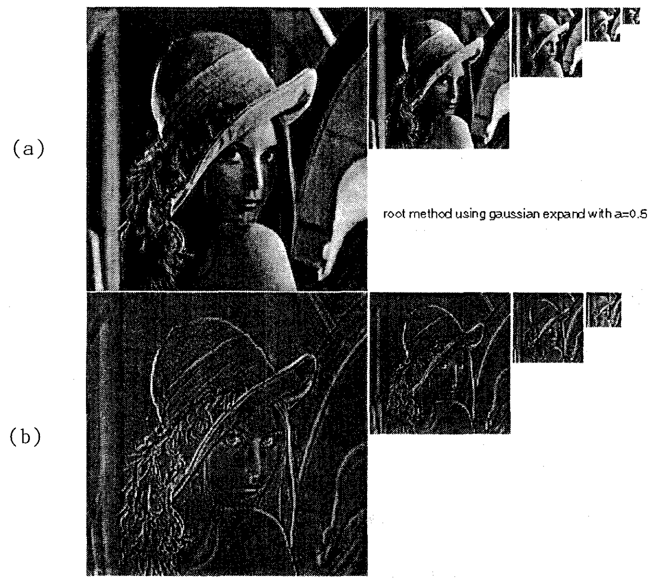

Figure 3. The first five levels of (a) the Gaussian pyramid and (b) the Laplacian pyramid of the image 'Lena'.

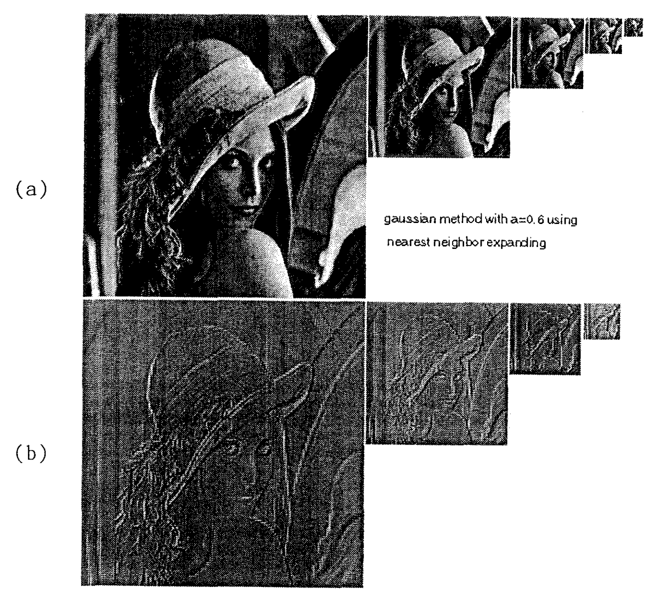

Figure 4. The first five levels of (a) the moment-preserving pyramid and (b) the corresponding difference pyramid of the image 'Lena'.

Table 1. Entropy comparison of the image 'Lena'

	256x256	128x128	64x64	32x32	16x16
Laplacian	5.62	6.14	6.70	7.03	6.87
Moment 1	5.05	5.54	6.14	6.57	6.57
Moment 2	4.71	5.21	5.76	6.06	6.77

* Moment 1 is based on Gaussian-like expansion with c=0.6,
Moment 2 is based on nearest neighbor expansion,
The size of Laplacian pyramid is the number plus 1, i.e.
(257x257, 129x129, ...)

Table 2. Entropy comparison of the image 'Cameraman'

	256x256	128x128	64x64	32x32	16x16
Laplacian	5.27	5.42	5.70	5.89	6.57
Moment 1	4.90	5.02	5.29	5.66	6.49
Moment 2	4.58	4.65	4.83	5.08	6.49

Table 3. Entropy comparison of the image 'City'

	256x256	128x128	64x64	32x32	16x16
Laplacian	6.01	6.03	6.32	6.62	6.74
Moment 1	4.90	5.02	5.29	5.66	6.49
Moment 2	4.58	4.65	4.83	5.08	6.49

Table 4. Variance comparison of the image 'Lena'

	256x256	128x128	64x64	32x32	16x16
Laplacian	342	608	1008	1382	-
Moment 1	144	255	443	672	-
Moment 2	101	186	302	380	-

Table 5. Variance comparison of the image 'Cameraman'

	256x256	128x128	64x64	32x32	16x16
Laplacian	494	636	804	1047	-
Moment 1	252	302	370	459	-
Moment 2	190	247	291	343	-

Table 6. Variance comparison of the image 'City'

	256x256	128x128	64x64	32x32	16x16
Laplacian	675	600	754	878	-
Moment 1	430	349	392	364	-
Moment 2	358	239	278	274	-