

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/246727904>

Pyramid Methods in Image Processing

Article · November 1983

CITATIONS

942

READS

10,857

5 authors, including:



James R. Bergen
SRI International

83 PUBLICATIONS 15,230 CITATIONS

SEE PROFILE



Joan M Ogden
University of California, Davis

167 PUBLICATIONS 6,621 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Energy System Modelling [View project](#)



Low Carbon Fuels Policies [View project](#)

Pyramid methods in image processing

The image pyramid offers a flexible, convenient multiresolution format that mirrors the multiple scales of processing in the human visual system.

Digital image processing is being used in many domains today. In image enhancement, for example, a variety of methods now exist for removing image degradations and emphasizing important image information, and in computer graphics, digital images can be generated, modified, and combined for a wide variety of visual effects. In data compression, images may be efficiently stored and transmitted if translated into a compact digital code. In machine vision, automatic inspection systems and robots can make simple decisions based on the digitized input from a television camera.

But digital image processing is still in a developing state. In all of the areas just mentioned, many important problems remain to be solved. Perhaps this is most obvious in the case of machine vision: we still do not know how to build machines

that can perform most of the routine visual tasks that humans do effortlessly.

It is becoming increasingly clear that the format used to represent image data can be as critical in image processing as the algorithms applied to the data. A digital image is initially encoded as an array of pixel intensities, but this raw format is not suited to most tasks. Alternatively, an image may be represented by its Fourier transform, with operations applied to the transform coefficients rather than to the original pixel values. This is appropriate for some data compression and image enhancement tasks, but inappropriate for others. The transform representation is particularly unsuited for machine vision and computer graphics, where the spatial location of pattern elements is critical.

Recently there has been a great deal of interest in representations that retain spatial localization as well as localization in the spatial—frequency domain. This is achieved by decomposing the image into a set of spatial frequency bandpass component images. Individual samples of a component image represent image pattern information that is appropriately localized, while the bandpassed image as a whole represents information about a particular fineness of detail or scale. There is evidence that the human visual system uses such a representation,¹ and multiresolution schemes are becoming increasingly popular in machine vision and in image processing in general.

The importance of analyzing images at many scales arises from the nature of

images themselves. Scenes in the world contain objects of many sizes, and these objects contain features of many sizes. Moreover, objects can be at various distances from the viewer. As a result, any analysis procedure that is applied only at a single scale may miss information at other scales. The solution is to carry out analyses at all scales simultaneously.

Convolution is the basic operation of most image analysis systems, and convolution with large weighting functions is a notoriously expensive computation. In a multiresolution system one wishes to perform convolutions with kernels of many sizes, ranging from very small to very large, and the computational problems appear forbidding. Therefore one of the main problems in working with multiresolution representations is to develop fast and efficient techniques.

Members of the Advanced Image Processing Research Group have been actively involved in the development of multiresolution techniques for some time. Most of the work revolves around a representation known as a "pyramid," which is versatile, convenient, and efficient to use. We have applied pyramid-based methods to some fundamental problems in image analysis, data compression, and image manipulation.

Image pyramids

The task of detecting a target pattern that may appear at any scale can be approached in several ways. Two of these, which involve only simple convolutions, are illus-

Abstract: *The data structure used to represent image information can be critical to the successful completion of an image processing task. One structure that has attracted considerable attention is the image pyramid. This consists of a set of lowpass or bandpass copies of an image, each representing pattern information of a different scale. Here we describe a variety of pyramid methods that we have developed for image data compression, enhancement, analysis and graphics.*

©1984 RCA Corporation

Final manuscript received November 12, 1984

Reprint Re-29-6-5

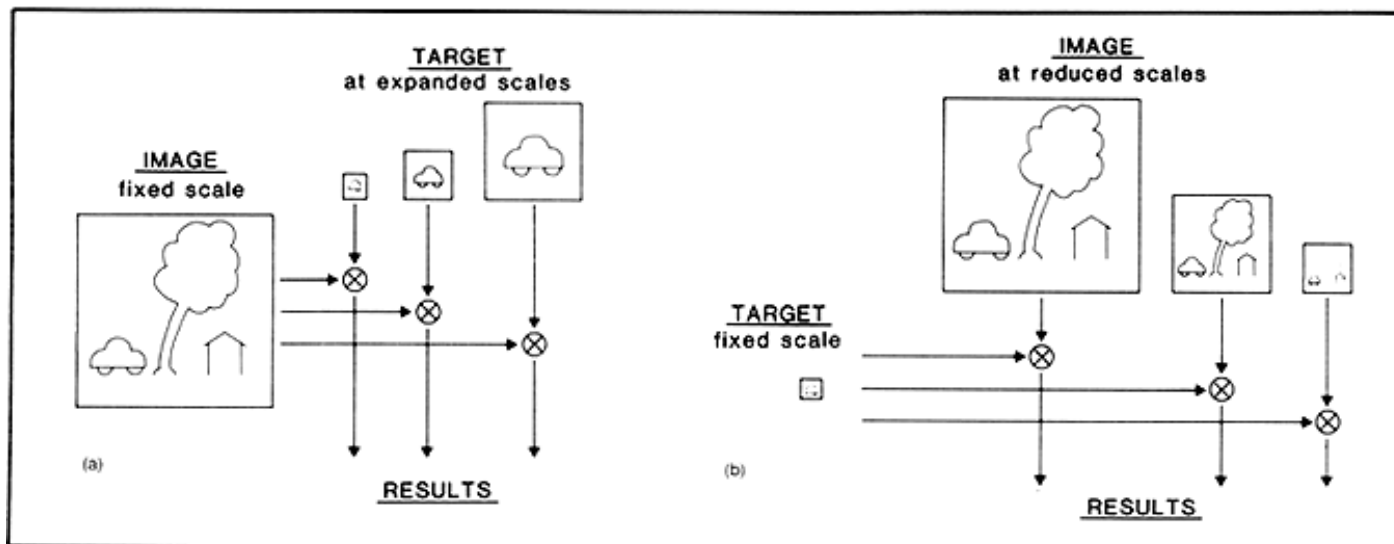


Fig. 1. Two methods of searching for a target pattern over many scales. In the first approach, (a), copies of the target pattern are constructed at several expanded scales, and each is convolved with the original image. In the second approach, (b), a single copy of the target is convolved with

copies of the image reduced in scale. The target should be just large enough to resolve critical details. The two approaches should give equivalent results, but the second is more efficient by the fourth power of the scale factor (image convolutions are represented by 'X').

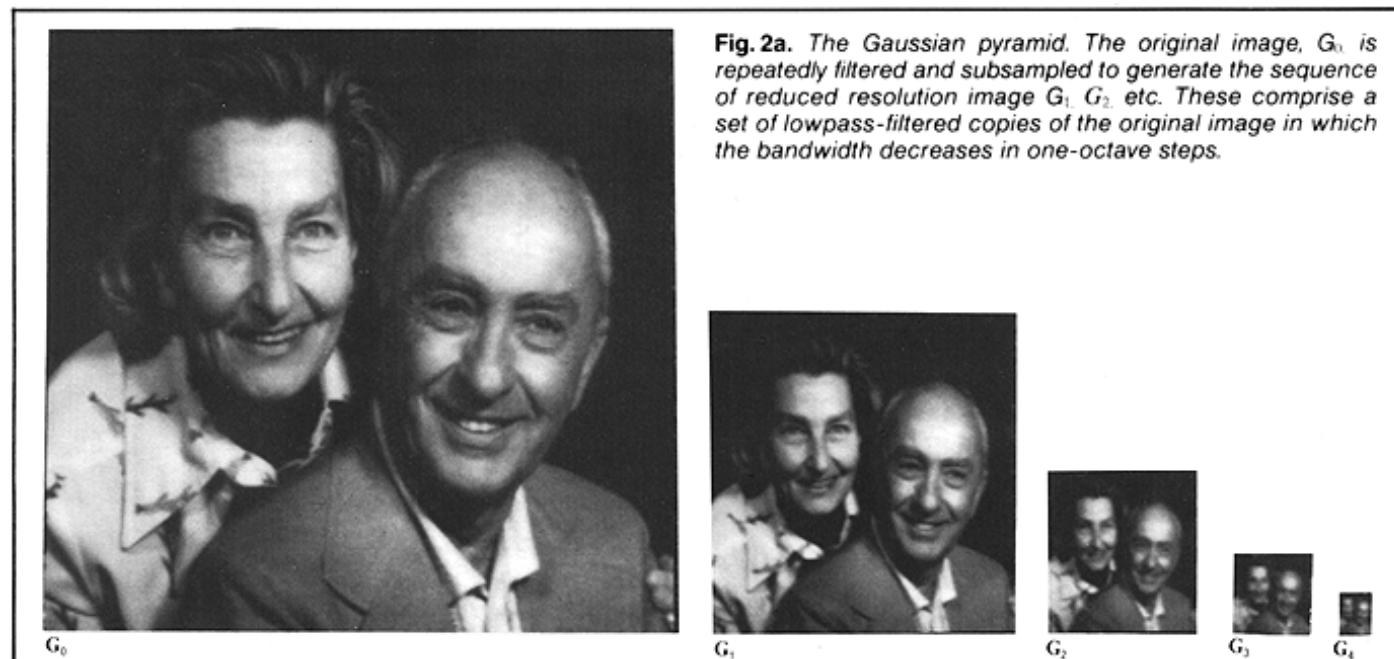


Fig. 2a. The Gaussian pyramid. The original image, G_0 , is repeatedly filtered and subsampled to generate the sequence of reduced resolution image G_1 , G_2 , etc. These comprise a set of lowpass-filtered copies of the original image in which the bandwidth decreases in one-octave steps.

trated in Fig. 1. Several copies of the pattern can be constructed at increasing scales, then each is convolved with the image. Alternatively, a pattern of fixed size can be convolved with several copies of the image represented at correspondingly reduced resolutions. The two approaches yield equivalent results, provided critical information in the target pattern is adequately represented. However, the second approach is much more efficient: a given convolution with the target pattern expanded in scale by a factor s will require s^4 more arithmetic operations than the corresponding

convolution with the image reduced in scale by a factor of s . This can be substantial for scale factors in the range 2 to 32, a commonly used range in image analysis.

The image pyramid is a data structure designed to support efficient scaled convolution through reduced image representation. It consists of a sequence of copies of an original image in which both sample density and resolution are decreased in regular steps. An example is shown in Fig. 2a. These reduced resolution levels of the pyramid are themselves obtained through a highly efficient iterative algorithm. The

bottom, or zero level of the pyramid, G_0 , is equal to the original image. This is low-pass-filtered and subsampled by a factor of two to obtain the next pyramid level, G_1 . G_1 is then filtered in the same way and subsampled to obtain G_2 . Further repetitions of the filter/subsample steps generate the remaining pyramid levels. To be precise, the levels of the pyramid are obtained iteratively as follows. For $0 < l < N$:

$$G_l(i,j) = \sum_m \sum_n w(m,n) G_{l-1}(2i+m, 2j+n) \quad (1)$$

However, it is convenient to refer to this

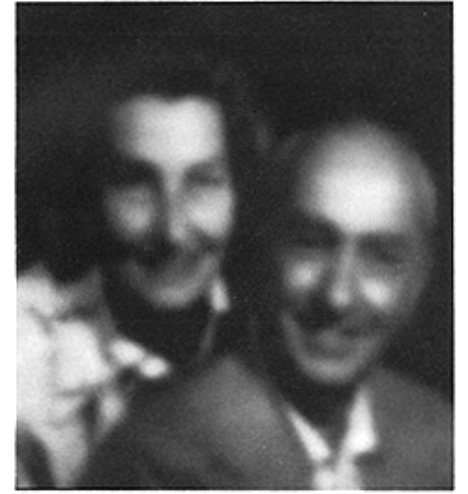


Fig. 2b. Levels of the Gaussian pyramid expanded to the size of the original image. The effects of lowpass filtering are now clearly apparent.

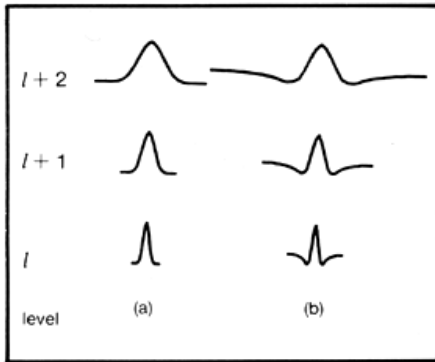


Fig. 3. Equivalent weighting functions. The process of constructing the Gaussian (lowpass) pyramid is equivalent to convolving the original image with a set of Gaussian-like weighting functions, then subsampling, as shown in (a). The weighting functions double in size with each increase in l . The corresponding functions for the Laplacian pyramid resemble the difference of two Gaussians, as shown in (b).

process as a standard REDUCE operation, and simply write

$$G_l = \text{REDUCE}[G_{l+1}].$$

We call the weighting function $w(m,n)$ the "generating kernel." For reasons of computational efficiency this should be small and separable. A five-tap filter was used to generate the pyramid in Fig. 2a.

Pyramid construction is equivalent to convolving the original image with a set of Gaussian-like weighting functions. These "equivalent weighting functions" for three successive pyramid levels are shown in Fig. 3a. Note that the functions double in width with each level. The convolution acts as a lowpass filter with the band limit

reduced correspondingly by one octave with each level. Because of this resemblance to the Gaussian density function we refer to the pyramid of lowpass images as the "Gaussian pyramid."

Bandpass, rather than lowpass, images are required for many purposes. These may be obtained by subtracting each Gaussian (lowpass) pyramid level from the next-lower level in the pyramid. Because these levels differ in their sample density it is necessary to interpolate new sample values between those in a given level before that level is subtracted from the next-lower level. Interpolation can be achieved by reversing the REDUCE process. We call this an EXPAND operation. Let G_{lk} be the image obtained by expanding G_l k times. Then $G_{lk} = \text{EXPAND}[G_{l,k-1}]$ or, to be precise, $G_{l,0} = G_l$, and for $k > 0$,

$$G_{lk}(i,j) = 4 \sum_m \sum_n G_{l,k-1} \left(\frac{2i+m}{2}, \frac{2j+n}{2} \right) \quad (2)$$

Here only terms for which $(2i+m)/2$ and $(2j+n)/2$ are integers contribute to the sum. The expand operation doubles the size of the image with each iteration, so that $G_{l,1}$ is the size of $G_{l,1}$, and $G_{l,1}$ is the same size as that of the original image. Examples of expanded Gaussian pyramid levels are shown in Fig. 2b.

The levels of the bandpass pyramid, L_0, L_1, \dots, L_N , may now be specified in terms of the lowpass pyramid levels as follows:

$$\begin{aligned} L_l &= G_l - \text{EXPAND}[G_{l+1}] \\ &= G_l - G_{l+1,1}. \end{aligned} \quad (3)$$

The first four levels are shown in Fig. 4a.

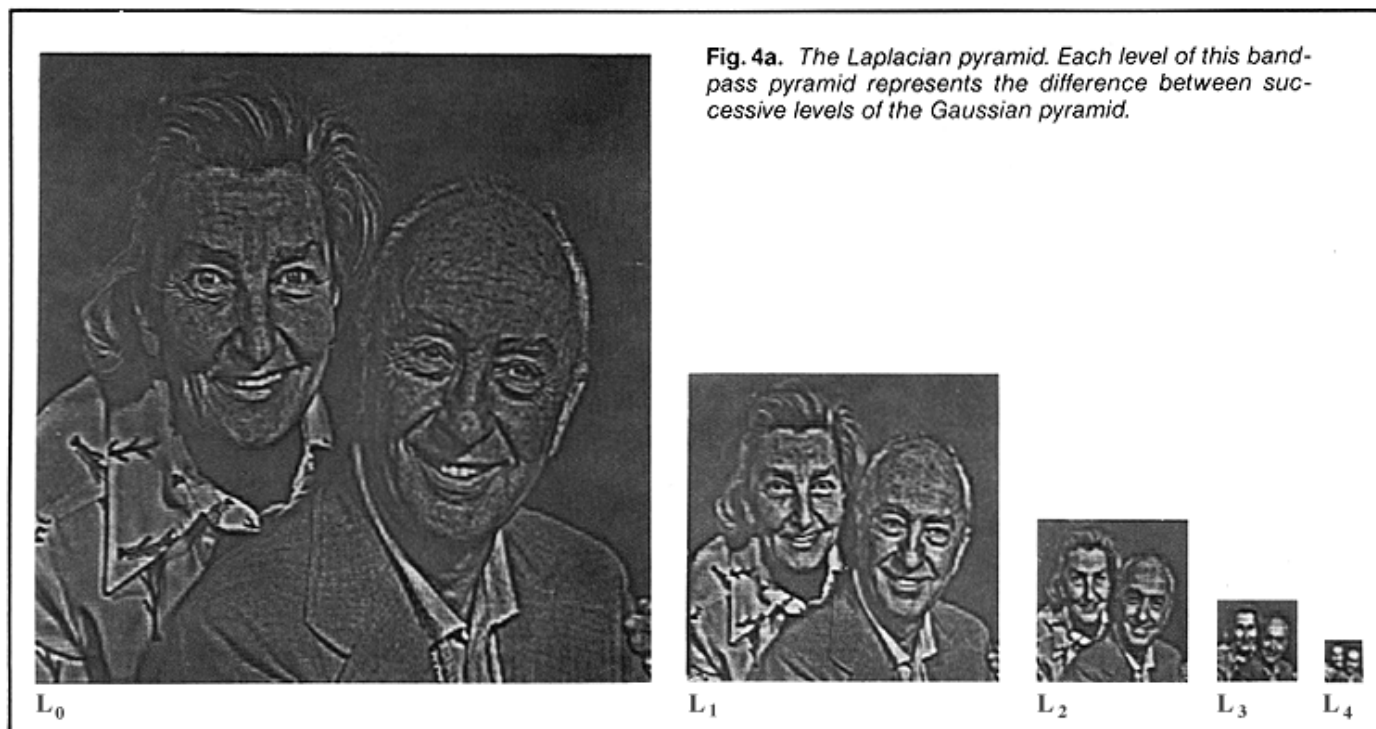
Just as the value of each node in the

Gaussian pyramid could have been obtained directly by convolving a Gaussian-like equivalent weighting function with the original image, each value of this bandpass pyramid could be obtained by convolving a difference of two Gaussians with the original image. These functions closely resemble the Laplacian operators commonly used in image processing (Fig. 3b). For this reason we refer to the bandpass pyramid as a "Laplacian pyramid."

An important property of the Laplacian pyramid is that it is a complete image representation: the steps used to construct the pyramid may be reversed to recover the original image exactly. The top pyramid level, L_N , is first expanded and added to L_{N-1} to form G_{N-1} , then this array is expanded and added to L_{N-2} to recover G_{N-2} , and so on. Alternatively, we may write

$$G_0 = \sum L_{li} \quad (4)$$

The pyramid has been introduced here as a data structure for supporting scaled image analysis. The same structure is well suited for a variety of other image processing tasks. Applications in data compression and graphics, as well as in image analysis, will be described in the following sections. It can be shown that the pyramid-building procedures described here have significant advantages over other approaches to scaled analysis in terms of both computation cost and complexity. The pyramid levels are obtained with fewer steps through repeated REDUCE and EXPAND operations than is possible with the standard FFT. Furthermore, direct convolution with large equivalent weighting functions requires 20- to 30-bit arithmetic to maintain the same ac-



curacy as the cascade of convolutions with the small generating kernel using just 8-bit arithmetic.

A compact code

The Laplacian pyramid has been described as a data structure composed of bandpass copies of an image that is well suited for scaled-image analysis. But the pyramid may also be viewed as an image transformation, or code. The pyramid nodes are then considered code elements, and the equivalent weighting functions are sampling functions that give node values when convolved with the image. Since the original

image can be exactly reconstructed from its pyramid representation (Eq. 4), the pyramid code is complete.

There are two reasons for transforming an image from one representation to another: the transformation may isolate critical components of the image pattern so they are more directly accessible to analysis, or the transformation may place the data in a more compact form so that they can be stored and transmitted more efficiently. The Laplacian pyramid serves both of these objectives. As a bandpass filter, pyramid construction tends to enhance image features, such as edges, which are important for interpretation. These features

are segregated by scale in the various pyramid levels, as shown in Fig. 4. As with the Fourier transform, pyramid code elements represent pattern components that are restricted in the spatial-frequency domain. But unlike the Fourier transform, pyramid code elements are also restricted to local regions in the spatial domain. Spatial as well as spatial-frequency localization can be critical in the analysis of images that contain multiple objects so that code elements will tend to represent characteristics of single objects rather than confound the characteristics of many objects.

The pyramid representation also permits data compression.³ Although it has one



Fig. 5. Pyramid data compression. The original image represented at 8 bits per pixel is shown in (a). The node values of the Laplacian pyramid representation of this image were quantized to obtain effective data rates of 1 b/p and 1/2 b/p. Reconstructed images (b) and (c) show relatively little degradation.

third more sample elements than the original image, the values of these samples tend to be near zero, and therefore can be represented with a small number of bits. Further data compression can be obtained through quantization: the number of distinct values taken by samples is reduced by binning the existing values. This results in some degradation when the image is reconstructed, but if the quantization bins are carefully chosen, the degradation will not be detectable by human observers and will not affect the performance of analysis algorithms.

Figure 5 illustrates an application of the pyramid to data compression for image transmission. The original image is shown in Fig. 5a. A Laplacian pyramid representation was constructed for this image, then the values were quantized to reduce the effective data rate to just one bit per pixel, then to one-half bit per pixel. Images reconstructed from the quantized data are shown in Figs. 5b and 5c. Humans tend to be more sensitive to errors in low-frequency image components than in high-frequency components. Thus in pyramid compression, nodes at level zero can be quantized more coarsely than those in higher levels. This is fortuitous for compression since three-quarters of the pyramid samples are in the zero level.

Data compression through quantization may also be important in image analysis to reduce the number of bits of precision carried in arithmetic operations. For example, in a study of pyramid-based image motion analysis it was found that data could be reduced to just three bits per sample without noticeably degrading the computed flow field.⁴

These examples suggest that the pyramid is a particularly effective way of representing image information both for transmission and analysis. Salient information is enhanced for analysis, and to the extent that quantization does not degrade analysis, the representation is both compact and robust.

Image analysis

Pyramid methods may be applied to analysis in several ways. Three of these will be outlined here. The first concerns pattern matching and has already been mentioned: to locate a particular target pattern that may occur at any scale within an image, the pattern is convolved with each level of the image pyramid. All levels of the pyramid combined contain just one third more nodes than there are pixels in the original image. Thus the cost of searching for a pattern at many scales is just one third more than that of searching the original image alone.

The complexity of the patterns that may be found in this way is limited by the fact that not all image scales are represented in the pyramid. As defined here, pyramid levels differ in scale by powers of two, or by octave steps in the frequency domain. Power-of-two steps are adequate when the patterns to be located are simple, but complex patterns require a closer match between the scale of the pattern as defined in the target array, and the scale of the pattern as it appears in the image. Variants on the pyramid can easily be defined with squareroot-of-two and smaller steps. However, these not only have more levels, but many more samples, and the computational

cost of image processing based on such pyramids is correspondingly increased.

A second class of operations concerns the estimation of integrated properties within local image regions. For example, a texture may often be characterized by local density or energy measures. Reliable estimates of image motion also require the integration of point estimates of displacement within regions of uniform motion. In such cases early analysis can often be formulated as a three-stage sequence of standard operations. First, an appropriate pattern is convolved with the image (or images, in the case of motion analysis). This selects a particular pattern attribute to be examined in the remaining two stages. Second, a nonlinear intensity transformation is performed on each sample value. Operations may include a simple threshold to detect the presence of the target pattern, a power function to be used in computing texture energy measures, or the product of corresponding samples in two images used in forming correlation measures for motion analysis. Finally the transformed sample values are integrated within local windows to obtain the desired local property measures.

Pattern scale is an important parameter of both the convolution and integration stages. Pyramid-based processing may be employed at each of these stages to facilitate scale selection and to support efficient computation. A flow diagram for this three-stage analysis is given in Fig. 6. Analysis begins with the construction of the pyramid representation of the image. A feature pattern is then convolved with each level of the pyramid (Stage 1), and the resulting correlation values may be passed through

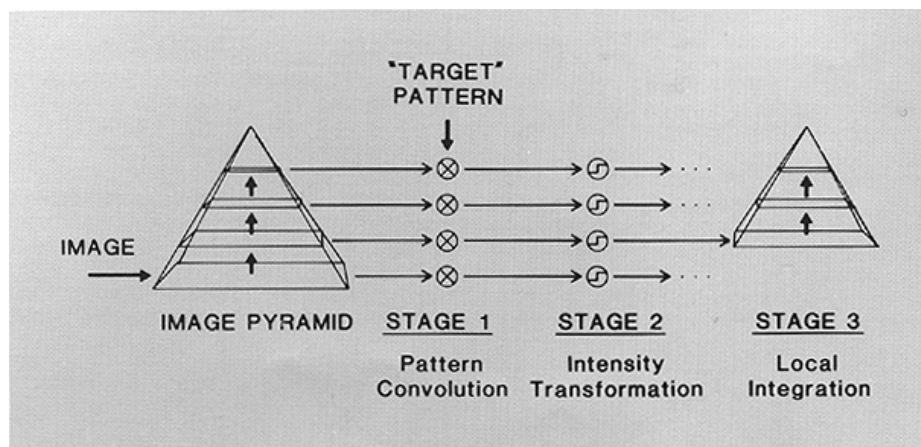


Fig.6. Efficient procedure for computing integrated image properties at many scales. Each level of the image pyramid is convolved with a pattern to enhance an elementary image characteristic, step 1. Sample values in the filtered image may then be passed through a nonlinear transformation, such as a threshold or power function, step 2. Finally, a new "integration" pyramid is built on each of the processed image pyramid levels, step 3. Node values then represent an average image characteristic integrated within a Gaussian-like window.

a nonlinear intensity transformation (Stage 2). Finally, each filtered and transformed image becomes the bottom level of a new Gaussian pyramid. Pyramid construction has the effect of integrating the input values within a set of Gaussian-like windows of many scales (Stage 3).

As an example, integrated property estimates have been used to locate the boundary between the two textured regions of Fig. 7a. The upper and lower halves of this image show two pieces of wood with differently oriented grain. The right half of the image is covered by a shadow. The boundary between the shaded and unshaded regions is the most prominent feature in the image, and its location can be detected quite easily as the maximum of the gradient of the image intensity (Fig. 7b). However, a simple edge-detecting operation such as this gradient-based procedure cannot be used to locate the boundary between the two pieces of wood. Instead it would isolate the line patterns that make up the wood grain.

The texture boundary can be found through the three-step process as follows: A Laplacian pyramid is constructed for the original texture. The vertical grain is then enhanced by convolving the image with a horizontal gradient operator (Stage 1). Each pyramid node value is then squared, (Stage 2) and a new integration pyramid is constructed for each level of the filtered image pyramid (Stage 3). In this way energy measures are obtained within windows of various sizes. Figure 7c shows level 2 of the integration pyramid for level L_0 of the filtered-image pyramid.

Note that texture differences in the original image have been converted into differences in gray level. Finally, a simple gradient-based edge-detection technique can be used to locate the boundary between image regions, Fig. 7d. (Pyramid levels have been expanded to the size of the original image to facilitate comparison.)

A third class of analysis operations concerns fast coarse-fine search techniques. Suppose we need to locate precisely a large complex pattern within an image. Rather than attempt to convolve the full pattern with the image, the search begins by convolving a reduced-resolution pattern with a reduced-resolution copy of the image. This serves to roughly locate possible occurrences of the target pattern with a minimum of computation. Next, higher-resolution copies of the pattern and image can be used to refine the position estimates through a second convolution. Computation is kept to a minimum by restricting the search to neighborhoods of the points identified at the coarser resolution. The search may proceed through several stages of increased resolution and position refinement. The savings in computation that may be obtained through coarse-fine search can be very substantial, particularly when size and orientation of the target pattern and its position are not known.

Image enhancement

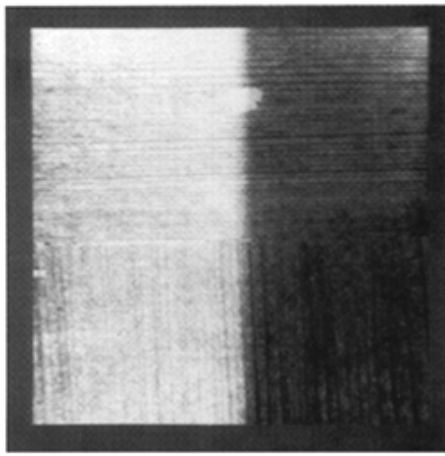
Thus far we have described how pyramid methods may be applied to data compression and image analysis. But there are other areas of image science where these

methods have proved be useful. For example, a method we call multi-resolution coring may be used to reduce random noise in an image while sharpening details of the image itself.⁵ The image is first decomposed into its Laplacian pyramid (bandpass) representation. The samples in each level are then passed through a coring function where small values (which include most of the noise) are set to zero, while larger values (which include prominent image features) are retained, or "peaked." The final enhanced image is then obtained by summing the levels of the processed pyramid. This technique is illustrated in Fig. 8. Figure 8a is the original image to which random noise has been added, and Fig. 8b shows the image enhanced through multiresolution coring.⁶

We have recently developed a pyramid-based method for creating photographic images with extended depth of field. We begin with two or more images focused at different distances and combine them in a way that retains the sharp regions of each. As an example, Figs. 9a and 9b show two pictures of a circuit board taken with the camera focused at two different depth-planes. We wish to construct a composite image in which all the components and the board surface are in focus. Let LA and LB be Laplacian pyramids for the two original images in our example. The low-frequency levels of these pyramids should be almost identical because the low spatial-frequency image components are only slightly affected by changes in focus. But changes in focus will affect node values in the pyramid levels where high-spatial-frequency information is encoded. However, corresponding nodes in the two pyramids will generally represent the same feature of the scene and will differ primarily in attenuation due to blur. The node with the largest amplitude will be in the image that is most nearly in focus. Thus, "in focus" image components can be selected node-by-node in the pyramid rather than region-by-region in the original images. A pyramid LC is constructed for the composite image by setting each node equal to the corresponding node in LA or LB that has the larger absolute value:

$$\begin{aligned} &\text{If } |LA_i(i,j)| > |LB_i(i,j)|, \\ &\text{then } LC_i(i,j) = LA_i(i,j) \\ &\text{otherwise, } LC_i(i,j) = LB_i(i,j) \end{aligned} \quad (7)$$

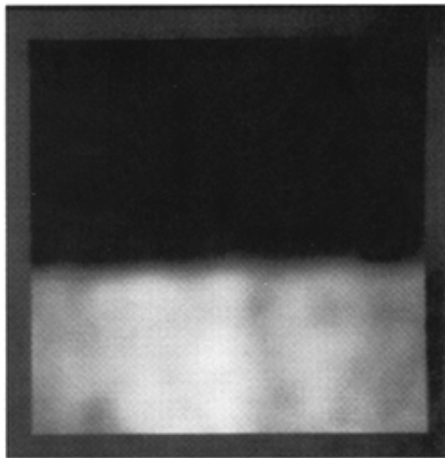
The composite image is then obtained simply by expanding and adding the levels of LC . Figure 9c shows an extended depth-of-field image obtained in this way.



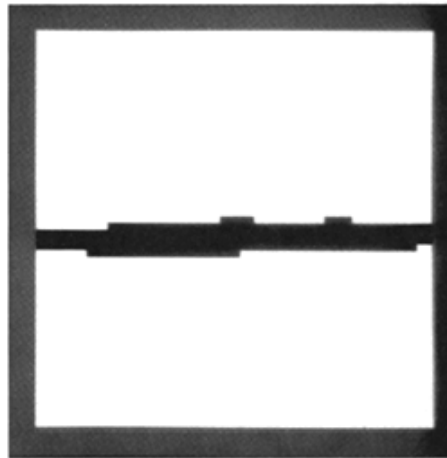
(a)



(b)



(c)



(d)

Fig. 7. Texture boundary detection using energy measures. The original image, (a), contains two pieces of wood with differently oriented grain separated by a horizontal boundary. The right half of this image is in a shadow, so an attempt to locate edges based on image intensity would isolate the boundary of the shadow region, (b). In order to detect the boundary between the pieces of wood in this image we first convolve each level of its Laplacian pyramid with a pattern that enhances vertical features. At level L_0 this matches the scale of the texture grain on the lower half of the image. The nodes at this level are squared and integrated (by constructing an additional pyramid) to give the energy image in (c). Finally, an intensity edge-detector applied to the energy image yields the desired texture boundary.

A related application of pyramids concerns the construction of image mosaics. This is a common task in certain scientific fields and in advertising. The objective is to join a number of images smoothly into a larger mosaic so that segment boundaries are not visible. As an example, suppose we wish to join the left half of Fig. 10a with the right half of Fig. 10b. The most direct method for combining the images is to catenate the left portion of Fig. 10a with the right portion of Fig. 10b. The result, shown in Fig. 10c, is a mosaic in which the boundary is clearly visible as a sharp (though generally low-contrast) step in gray level.

An alternative approach is to join image components smoothly by averaging pixel values within a transition zone centered on the join line. The width of the transition zone is then a critical parameter. If it is too narrow, the transition will still be visible as a somewhat blurred step. If it is too wide, features from both images will be visible within the transition zone as in a photographic double exposure. The blurred-edge effect is due to a mismatch of low frequencies along the mosaic boundary, while the double-exposure effect is due to a mismatch in high frequencies. In general, there is no choice of transition zone width that can avoid both defects.



(a)

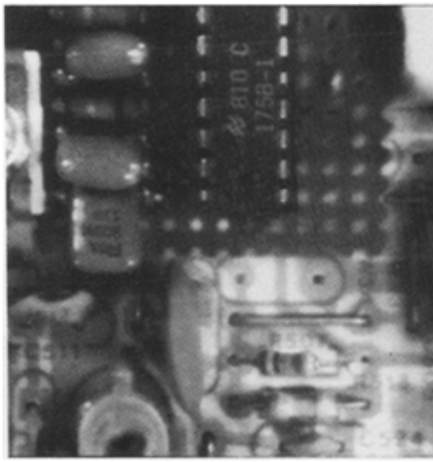


(b)

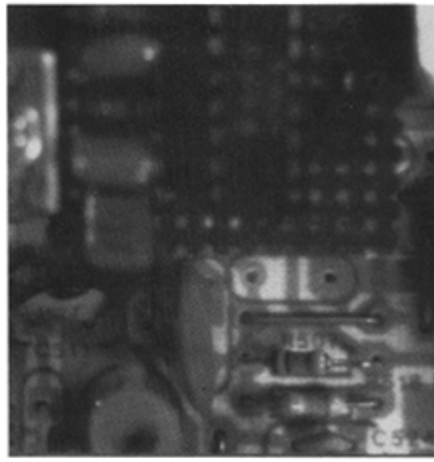
Fig. 8. Multiresolution coring. Part (a) shows an image to which noise has been added to simulate transmission degradation. The Laplacian pyramid was constructed for this noisy image, and node values at each level were "cored." As a result, much of the noise is removed while prominent features of the original image are retained in the reconstructed image, (b).

This dilemma can be resolved if each image is first decomposed into a set of spatial-frequency bands. Then a bandpass mosaic can be constructed in each band by use of a transition zone that is comparable in width to the wavelengths represented in the band. The final mosaic is then obtained by summing the component band-pass mosaics.

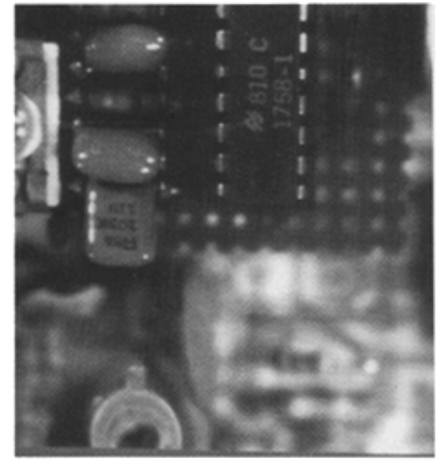
The computational steps in this "multiresolution splining" procedure are quite simple when pyramid methods are used.⁶ To begin, Laplacian pyramids LA and LB are constructed for the two original images. These decompose the images into the required spatial-frequency bands. Let P be the



(a)



(b)



(c)

Fig. 9. Multifocus composite image. The original images with limited depth of field are shown in (a) and (b). These are combined digitally to give the image with an extended depth of field in (c).

summed to yield the final mosaic, Fig. 10d. Note that it is not necessary to average node values within an extended transition zone since this blending occurs automatically as part of the reconstruction process.

Conclusions

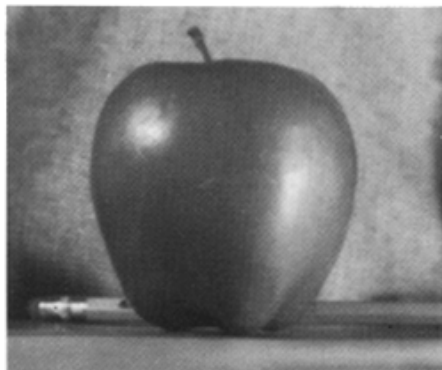
The pyramid offers a useful image representation for a number of tasks. It is efficient to compute: indeed pyramid filtering is faster than the equivalent filtering done with a fast Fourier transform. The information is also available in a format that is convenient to use, since the nodes in each level represent information that is localized in both space and spatial frequency.

We have discussed a number of examples in which the pyramid has proven to be valuable. Substantial data compression (similar to that obtainable with transform methods) can be achieved by pyramid encoding combined with quantization and entropy coding. Tasks such as texture analysis can be done rapidly and simultaneously at all scales. Several different images can be combined to form a seamless mosaic, or several images of the same scene with different planes of focus can be combined to form a single sharply focused image.

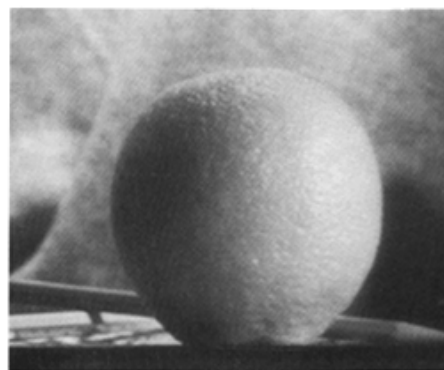
Because the pyramid is useful in so many tasks, we believe that it can bring some conceptual unification to the problems of representing and manipulating low-level visual information. It offers a flexible, convenient multiresolution format that matches the multiple scales found in the visual scenes and mirrors the multiple scales of processing in the human visual system.

References

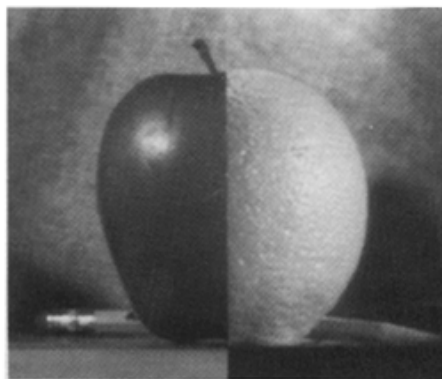
1. H. Wilson and J. Bergen' "A four mechanism model for threshold special vision", *Vision Research*. Vol. 19, pp. 19-31, 1979.



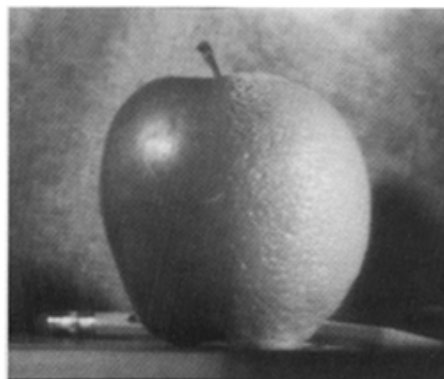
(a)



(b)



(c)



(d)

Fig. 10. Image mosaics. The left half of image (a) is catinated with the right half of image (b) to give the mosaic in (c). Note that the boundary between regions is clearly visible. The mosaic in (d) was obtained by combining images separately in each spatial frequency band of their pyramid representations then expanding and summing these bandpass mosaics.

locus of image points that fall on the boundary line, and let R be the region to the left of P that is to be taken from the left image. Then the pyramid LC for the composite image is defined as:

If the sample is in R , then

$$LC_i(i,j) = LA_i(i,j)$$

If the sample is in P , then

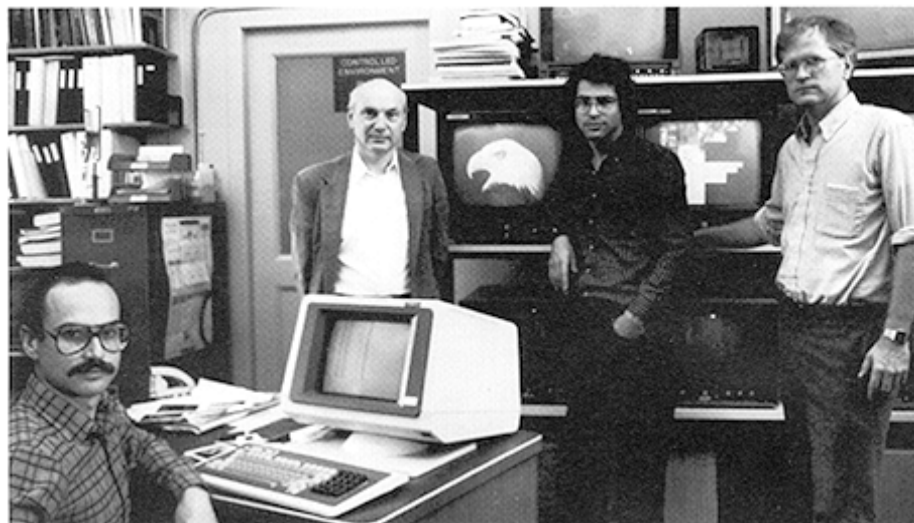
$$LC_i(i,j) = LB_i(i,j),$$

Otherwise,

$$LC_i = LC_i(i,j) \quad (8)$$

The levels of LC are then expanded and

2. C. Anderson, "An alternative to the Burt pyramid algorithm", memo in preparation.
3. P. Burt and E. Adelson, "The Laplacian Pyramid as a Compact Image Code," *IEEE Transactions on Communication*, COM-31 pp. 532-540, 1983a.
4. P. Burt, X. Xu and C. Yen, "Multi-Resolution Flow-Through Motion Analysis," RCA Technical Report, PRRL-84-TR-009, 1984.
5. J. Ogden and E. Adelson, "Computer Simulations of Oriented Multiple Spatial Frequency Band Coring," in preparation, 1984.
6. P. Burt and E. Adelson, "Multiresolution Spline with Application to Image Mosaics," *ACM Transactions on Graphics*, Vol. 2, pp. 217-236, 1983b.



Authors, left to right: Bergen, Anderson, Adelson, Burt.

Edward H. Adelson received a B.A. degree, summa cum laude, in Physics and Philosophy from Yale University in 1974, and a Ph.D. degree in Experimental Psychology from the University of Michigan in 1979. His dissertation dealt with temporal properties of the photoreceptors in the human eye. From 1978 to 1981 Dr. Adelson did research on human motion perception and on digital image processing as a Postdoctoral Fellow at New York University. Dr. Adelson joined RCA Laboratories in 1981 as a Member of the Technical Staff. As part of the Advanced Image Processing Research group in the Advanced Video Systems Research Laboratory, he has been involved in developing models of the human visual system, as well as image-processing algorithms for image enhancement and data compression.

Dr. Adelson has published a dozen papers on vision and image processing, and has made numerous conference presentations. His awards include the Optical Society of America's Adolph Lomb medal (1984), and an RCA Laboratories Outstanding Achievement Award (1983). He is a member of the Association for Research in Vision and Ophthalmology, the Optical Society of America, and Phi Beta Kappa.

Contact him at:
RCA Laboratories
Princeton, N.J.
Tacnet: 226-3036

Peter J. Burt received the B.A. degree in Physics from Harvard University in 1968, and the M.S. degree from the University of Massachusetts, Amherst, in 1974 and 1976, respectively. From 1968 to 1972 he conducted research in sonar, particularly in acoustic imaging devices, at the U.S. Navy Underwater Systems Center, New London, Conn. and in London, England. As a Postdoctoral Fellow, he has studied both natural vision and computer image under-

standing at New York University (1976-1978), Bell Laboratories (1978-1979), and the University of Maryland (1979-1980). He was a member of the engineering faculty at Rensselaer Polytechnic Institute from 1980 to 1983. In 1983 he joined RCA David Sarnoff Research Center as a Member of the Technical Staff, and in 1984 he became head of the Advanced Image Processing Group.

Contact him at:
RCA Laboratories
Princeton, N.J.
Tacnet: 226-2451

Charles H. Anderson received B.S. degree in Physics at the California Institute of Technology in 1957, and a Ph. D. from Harvard University in 1962. Dr. Anderson joined the staff of RCA Laboratories, Princeton, N.J. in 1963. His work has involved studies of the optical and microwave properties of rare-earth ions in solids. These studies have produced an optically-pumped microwave maser and a new spectrometer for acoustic radiation in the 10-to 300-GHz range. In 1971 he was awarded an RCA fellowship to do research at Oxford University for a year, and in 1972 became a Fellow of the American Physical Society. Upon returning to RCA, he became involved in new television displays. Between 1973 and 1978 he was a leader of a subgroup developing electron-beam guides for flat-panel television displays. In March 1977 he was appointed a Fellow of the Technical Staff of RCA Laboratories. From August 1978 through December 1982 he was head of the Applied Mathematical and Physical Sciences group. In January 1983 he returned full time to research as a member of the Vision Group, while maintaining a role as a task force leader in studies of the stylus/disc interface. In January 1984 he spent 5 weeks as a Regents Lecturer at the invitation of the Physics Department of UCLA. This was



Joan Ogden received a B.S. in Mathematics from the University of Illinois, Champaign-Urbana in 1970, and a Ph.D. in Physics from the University of Maryland in 1977. Coming to the Princeton Plasma Physics laboratory as a Post-Doctoral research associate, she continued her work in nuclear fusion, specializing in plasma theory and simulation. In 1980, she started her own consulting company, working on a variety of applied physics problems. In December 1982, she began working with the Advanced Image Processing Research Group, and has recently joined RCA as a part-time Member of the Technical Staff. Her research interests at RCA include applications of the pyramid algorithm to problems of noise reduction, data compression, and texture generation.

Contact her at:
RCA Laboratories
Princeton, N.J.

did research into and developed a model of the structure of the primate visual system from the retina to the striate cortex. Research was also done on the Hopfield model of associative memory.

Contact him at:
RCA Laboratories
Princeton, N.J.
Tacnet: 226-2901

James R. Bergen received the B.A. degree in Mathematics and Psychology from the University of California, Berkeley, in 1975, and the Ph.D. in Biophysics and Theoretical Biology from the University of Chicago in 1981. His work concerns the quantitative analysis of information processing in the human visual system. At the University of Chicago he was involved in the development of a model of the spatial and temporal processing that occurs in the early stages of the system. From 1981 to 1982 he was with Bell Laboratories, Murray Hill, N.J. His work concentrates on the effect of visual system structure on the extraction of information from a visual image. His current work includes basic studies of visual perception as well as perceptual considerations for design of imaging systems.

Contact him at:
RCA Laboratories
Princeton, N.J.
Tacnet: 226-3003