Improvements to NTSC by Multidimensional Filtering

By Eric Dubois and William F. Schreiber

This article presents, in a tutorial fashion, the basic elements required for understanding compatible improvement techniques for NTSC color multiplexing using multidimensional filtering at the transmitter and receiver. It discusses the multidimensional spectrum of moving pictures and relates the form of this spectrum to different types of image material. It also presents the basics of multidimensional filtering that are key to the techniques we discuss for improving NTSC. Finally, systems for NTSC encoding and decoding using multidimensional processing are described.

When color television was introduced in the U.S. in the early 1950s, a fateful decision was taken that was to have an enduring influence on the development of this medium. The Federal Communication Commission decreed that the new system was to be compatible with the then-existing monochrome system and was to fit into the same channel. New and old receivers were both to function adequately with old monochrome as well as new color signals. The wisdom of that decision is open to question, but it is a part of our technological history and we must deal with it.

Color signals clearly contain more information than monochrome signals; the extra information had to be somehow squeezed into the same channel and, moreover, the extra signals were to be essentially invisible on existing receivers. To meet these requirements, the designers of the color system had to look for "holes" in the spectrum that had gone mostly unused, and, it turned out, their search was rather successful. The backward compatibility requirement was met easily by using the monochrome signal as one of the color components.

The commercial and technological success of the NTSC[a] system is undeniable. Yet, as equipment has improved and viewers, both professional and lay, have become more critical, its defects[b] have become more apparent. These are, principally, cross color (false colors produced by the luminance signal in certain detailed areas), cross luminance (artifacts of luminance produced by the color signal, principally moving edge distortion at boundaries of brightly colored areas), and a reduction in horizontal resolution. To some extent, the color subcarrier is also visible.

It is fair to say that the theoretical basis for these problems is now understood much better than it was 35 years ago. In addition, modern signal-processing methods, which depend at least as much on better circuit components as on superior understanding, make it possible to achieve, almost perfectly, what the ingenious NTSC pioneers tried to do at that time. There is, of course, some cost associated with these improvements. However, if we are willing to pay the price, we can have back nearly the full resolution[c] of the monochrome system as it existed before color was added, and we can completely eliminate cross color and cross luminance. We find it ironic that most contemporary proposals for improving television include the abandonment of the NTSC band-sharing principle just at the time we have learned how to make it work!

The NTSC System

Owing to a fortunate quirk of human vision, excellent color reproduction is possible by superimposing three images, preferably very pure red, green, and blue, each derived from a different portion of the visible spectrum using three color filters of somewhat overlapping transmission characteristics, in front of three cameras operating in register. For transmission, the three signals are subject to a linear transformation, resulting in a luminance signal much like the signal of black-and-white TV, plus two color-difference signals. The latter may be transmitted with less resolution, thus simplifying the problem of squeezing them into the monochrome channel. At the receiver, the linear transformation is inverted, producing the required signals for display.

Mertz and Gray[1] showed that the spectrum[d] of a still monochrome TV signal consists of harmonics of the horizontal scanning frequency, each surrounded by clusters of harmonics of the vertical scanning frequency. Naturally, a still picture scanned in a repetitive pattern must produce a signal that is precisely periodic, the spectrum thus consisting of infinitesimal lines, all harmonics of the frame rate. If successive frames are slightly different, the spectral lines spread out. If the motion is sufficient to cause these broadened lines to overlap, motion rendition starts to deteriorate, with the appearance of stroboscopic effects, now called temporal aliasing.

A remarkable feature of the video spectrum is that it is mostly empty — there is very little energy between the spectral lines *most of the time*. This makes it possible for luminance and chrominance signals to share the same band, a suggestion first made by Frank Gray in 1929.[2] A modulated signal whose carrier and associated harmonics fall halfway between these lines can, in principle, be multiplexed with the original signal, and the two can later be separated by a suitable "comb" filter.[e] This relationship is achieved with any signal whose frequency is an odd multiple of one-half the horizontal scanning frequency. Thus a color subcarrier was chosen with frequency

$$f_{sc} = (455/2)f_h$$

(about 3.58 MHz) near the upper end of the video band, and the color difference signals (called "I" and "Q") are

---

A contribution received from Eric Dubois, Ph.D., INRS-Télécommunications, Verdun, Quebec, Canada H3E 1H6 (who was on sabbatical leave at MIT); and William F. Schreiber, Ph.D., MIT, Cambridge, MA 02139.

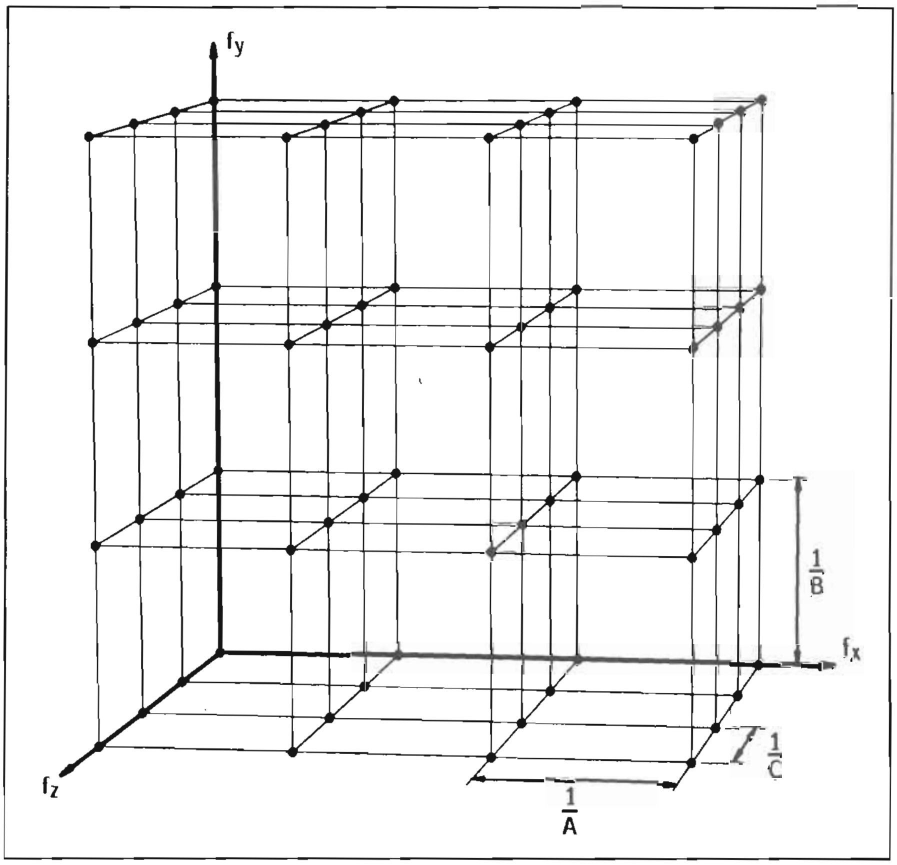

Figure 1. Discrete spatiotemporal frequency components in the Fourier series representation of a time-varying image of dimension A by B over a time interval C.

quadrature modulated onto it, forming the "composite" signal. I and Q are band-limited to about 1.3 and 0.5 MHz, respectively, before the subcarrier is modulated. After the modulated chrominance is added to the luminance signal, the composite signal is low-pass filtered to 4.2 MHz, so that the I and Q sidebands fit within the 4.2-MHz video bandwidth (although part of the upper I sideband is removed).

Comb filters of the type indicated here, nowadays called "line combs," were not used until recently. (They are now common in professional equipment and are beginning to be found in higher-priced consumer equipment as well.) Rather, the nonvisibility of the subcarrier and its sidebands was made to depend primarily on the fact that, for still images, the subcarrier phase reverses on every frame at each point in the image.[f] Since the eye does not integrate for exactly $1/15$ sec and because the picture tube transfer characteristic is nonlinear, this hoped-for cancellation is not perfect, even for still pictures, and is less so for those that move. This is the source of cross luminance. On the older monochrome receivers that had the full 4.2-MHz bandwidth, this appeared as a fine cross-hatch pattern in all colored areas. On conventional color receivers using notch filters, cross luminance produces a strong moving pattern around the edges of brightly colored objects.

Cross color, an equally disturbing artifact in which spurious colors appear around sharp edges or fine detail, is due to the interpretation of high-frequency luminance information as chrominance. This often occurs even with comb filters at the receiver, since the space between the interleaved horizontal harmonics of the two signals is not empty — it is filled with additional harmonics, spaced apart by the vertical scanning frequency, whose amplitude depends on the amount of vertical detail.

Spectral analysis is the most powerful mathematical tool to deal with these problems. We start by considering the light intensity at the focal plane of a camera, pointed at a scene, as a continuous function of x, y, and t. Such a function can be represented by a three-dimensional (3-D) Fourier series. The components of the 3-D spectrum are spoken of as "spatiotemporal frequencies." In the complex exponential form that is more convenient to use than sines and cosines, these have the form:

$$\exp[2\pi j(f_x x + f_y y + f_z z)]$$

Considering the video information to be confined to a rectangular box of dimensions A, B, and C in x, y, z space (z represents time), f_x is a harmonic of one cycle per picture width, f_y is a harmonic of one cycle per picture height, and f_z is a harmonic of one cycle per chosen time interval. The harmonics are thus located on a rectangular grid in 3-D frequency space, with the separation between adjacent lines of components being the fundamental frequency of one cycle per box dimension in the corresponding direction (Fig. 1).

By choosing to represent the signal within this box, the spectrum is given by the coefficients of a Fourier series. The latter is periodic in all three dimensions, the box size being the period. Therefore, outside the box, the series produces an endless repetition of the picture information in all three directions. This makes it very easy to derive the temporal spectrum produced by periodic scanning, since unidirectional scanning through the triple infinity of repeated boxes gives the same result as periodic scanning within the box. If f_h, f_v, and f_t are the horizontal, vertical, and temporal scanning frequencies, respectively, then the path of the scanning spot in x, y, z space is given by

$$x = Af_h t; \quad y = -Bf_v t; \quad z = Cf_t t = t$$

Since scanning is normally from top to bottom, a negative sign is required in the equation for y. Note that f_h, the number of lines per sec (15734.26), and f_v, the number of fields per sec (59.94), are meaningful. Since the focal plane image is really not confined to the time duration C, the "fundamental" frequency of one cycle per C sec is not of particular significance. It is used here merely to simplify the derivation. A more precise derivation is given later.

Substituting these expressions into the exponential, we convert it into a temporal component of the type

$$\exp[2\pi j(nf_h - mf_v + lf_t)t],$$

where n, m, and l are the orders of the respective harmonics. Thus each spatiotemporal component of the focal plane image corresponds to one temporal component of the video signal that results from scanning. The converse is not true, however, and we now have a problem of uniqueness.

Because of the vertical-temporal sampling that has occurred, each temporal frequency in the video signal corresponds to an infinite number of spatiotemporal components. If more than one of these spatiotemporal components (usually the one of "lowest" frequency) is nonzero, confusion occurs, since we cannot decide which of the possible spatiotemporal components correspond to the given temporal component. This confusion is referred to as aliasing. A set of spatiotemporal components that can be represented without aliasing by the scanned video signal is shown in Fig. 2. The diamond shape of the band in the vertical-temporal frequency plane is due to the line-interlaced scanning used in NTSC. The horizontal frequency limits of this region correspond to the 4.2-MHz bandwidth of the video signal,[g] while the vertical frequency limit is based on an active picture height of 484 lines.

By considering the frequency of the color subcarrier as the sum of multiples of the horizontal, vertical, and temporal fundamental frequencies, we can make a correspondence between it and a set of spatiotemporal frequency components in the basic cell. Although the latter have no "real" existence, since they do not exist in the focal plane image, they are real enough to subject to 3-D filtering. The subcarrier is located in four of the eight corners of 3-D frequency space (Fig. 2). The chrominance spectrum is centered around the subcarrier in the same way the luminance is centered at the origin (Fig. 3). The proportion of luminance information found in those corners is small, and in addition, visual sensitivity to such components, which are simultaneously high-frequency in all directions, is low. There is thus little loss in quality by eliminating them from the luminance signal. By this means, a space can be made for the chrominance components that is completely separate from that reserved for luminance (Fig. 4).

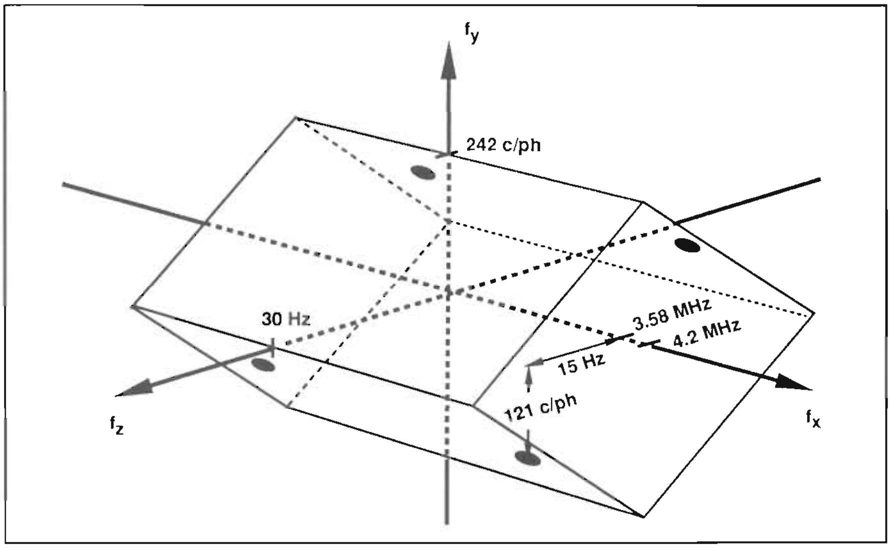

Figure 2. Spatiotemporal frequencies which can be represented by the scanned signal. The equivalent spatiotemporal subcarrier is located in four of the eight octants.

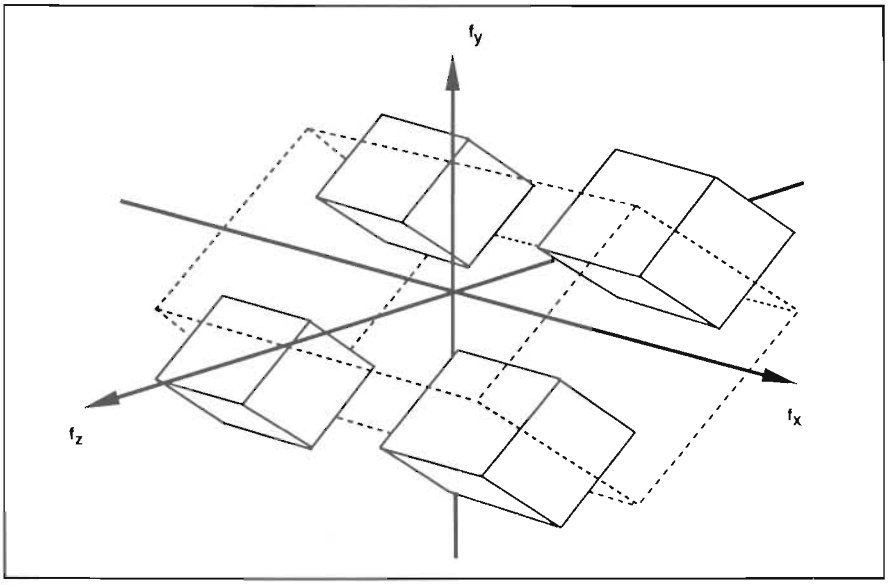

Figure 3. The chrominance spectrum is centered around the subcarrier in the same way the luminance spectrum is centered at the origin. The figure shows chrominance spectra with half the vertical and half the temporal bandwidth of the luminance signal.

Achieving the complete separation of luminance and chrominance in 3-D frequency space requires that all components be appropriately filtered *before* they are added together to form the composite signal. In particular, this means that the separation cannot be done perfectly at the receiver unless the appropriate filtering has been done at the transmitter. The way the prefiltering is best done depends on a prior decision on how to partition the space between the two signals. If one insists on doing the separation with simple temporal (1-D) filters, then both luminance and chrominance bandwidth must be reduced, their sum being no more than 4.2 MHz. If 2-D spatial filtering is to be done, the bands allocated to luminance and chrominance are cylinders perpendicular to the f_xf_y plane. In the spatial projection of the NTSC spectrum, the subcarrier lies in each of the four quadrants. Thus, the corners must be notched out of the 2-D luminance spectrum to allow room for chrominance (Fig. 5). Systems that incorporate line-comb filters are an example of 2-D filtering. Both components would then have the full temporal bandwidth. With the arrangement of Fig. 5, diagonal luminance resolution must be restricted, but the luminance can have full horizontal resolution at the lowest vertical frequencies. Also, vertical chrominance resolution must be restricted.

If the decision is made to use 3-D filters, then a corner of the 3-D luminance spectrum is notched out and assigned to chrominance (Fig. 4). Systems that use field- or frame-comb filters are examples of 3-D filters. Now the luminance can have nearly the entire space, except for those components that are simulta-

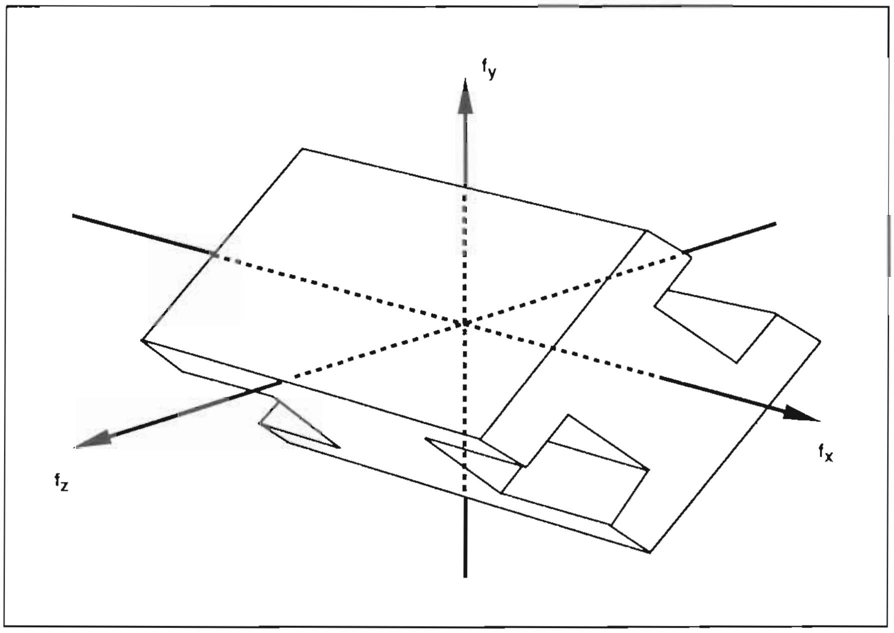

Figure 4. Luminance and chrominance spectra can be made disjoint in the composite signal by spatiotemporally prefiltering the luminance to carve out a notch into which the chrominance may be placed. The chrominance must also be spatiotemporally prefiltered to fit into the allotted space. One possible division of the available frequency space between luminance and chrominance is shown.

neously high in all three directions; that is, luminance temporal resolution must be restricted at the highest spatial frequencies.

Some Historical Notes

The spectrum of video signals was not generally understood during the TV experimentation of the 1920s. It was first described accurately by Mertz and Gray in 1934,[1] and the first suggestion of frequency-interleaving was made by Gray in 1929.[2] Comb filters for other purposes were known at an early date and were widely used for noise reduction of periodic signals.[3] A full description was given by Flesher in 1961.[4] Remarkably enough, they do not seem to have been specifically proposed for NTSC decoding until 1967 by Parker.[5]

The theory of filtering in one, two, and three dimensions to avoid cross effects in the PAL system was fully developed by Drewery in 1975,[6] including the description of the 3-D spectrum of the PAL signal. The application of some of these ideas using line-comb filters was given by Auty et al.[7] A treatment of the 3-D spectrum of composite NTSC signals was given by Dubois et al.[8]

The use of prefiltering of luminance and chrominance using line-comb filters to reduce cross effects in the NTSC system was proposed by Turner in 1977.[9] The extension to higher-order comb filters to obtain better suppression of cross effects has been suggested by Tsinberg and Fisch,[10] and a technique using field-comb filters has been presented by Strolle.[11] A report of experiments with high-order nonseparable 2-D digital filters is given by Dubois and Faubert.[12]

Adaptive comb filters for chroma/luma separation were proposed by Rossi in 1976[13] and Kaiser in 1977,[14] while a series of increasingly more sophisticated adaptive filters were described by Faroudja.[15] A recent treatment of adaptive filtering techniques is by Teichner.[16]

Three-Dimensional Spectrum of Images

A more detailed discussion of the form of the three-dimensional spatiotemporal spectrum of a time-varying television image follows. In particular, the general characteristics of the three-dimensional spectrum for several classes of typical image types are described. The effect of motion on the temporal component of the spectrum is also discussed. A good discussion of the spectral representation of various types of images is given by Kretz and Sabatier.[17]

Spatiotemporal Spectrum

As stated, any moving image of dimension A by B over a time interval C can be represented by a Fourier series. This representation has the form

$$u(x, y, z) = \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} U \times \left(\frac{n}{A}, \frac{m}{B}, \frac{l}{C}\right) \exp\left[j2\pi\left(\frac{nx}{A} + \frac{my}{B} + \frac{lz}{C}\right)\right]$$

The coefficients $U(n/A, m/B, l/C)$ define the spatiotemporal spectrum of the image, giving the relative weight of the different frequency compo-

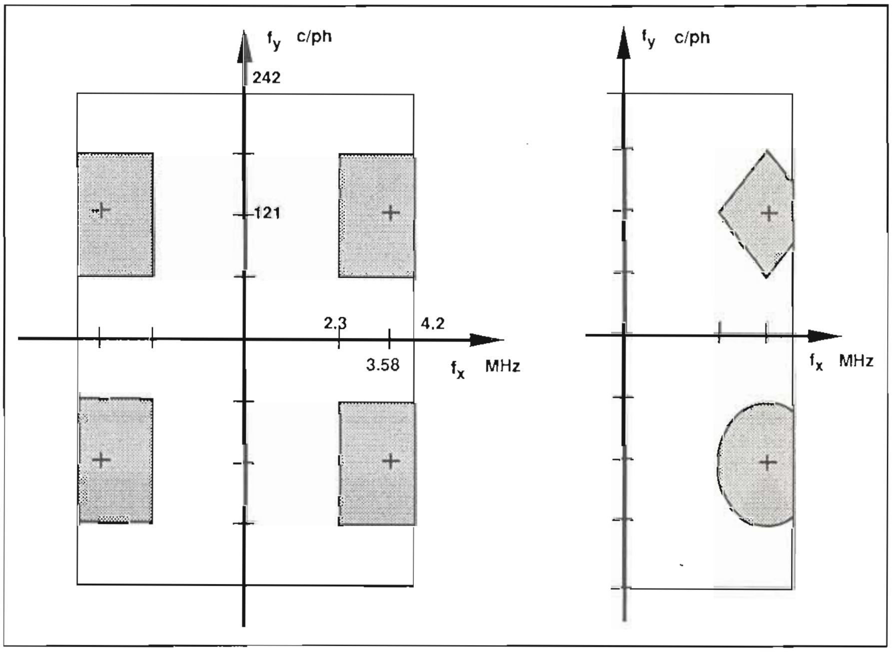

Figure 5. Luminance and chrominance can be made disjoint in two dimensions by spatial filtering only. Three possibilities for allocating the available frequency space to luminance and chrominance are shown. The spatiotemporal frequency bands are cylinders with the three different cross sections shown.

nents in the image. They are complex numbers that specify the magnitude and phase of the components. Since the time interval C can be arbitrarily long, the temporal frequency components can be spaced arbitrarily closely, and are usually considered to be continuous. However, the horizontal and vertical frequency components are constrained to be multiples of one cycle per picture width and one cycle per picture height respectively.

In television, where the absolute picture size is variable, the active picture height is often chosen as the basic unit of spatial distance, and spatial frequencies are expressed in units of cycles per picture height (c/ph). If we assume an active picture height of 484 lines and a total picture height of 525 lines, a vertical spatial frequency m/B is equivalent to $484m/525 \approx 0.922m$ c/ph. With an active picture aspect ratio of 4:3 and a horizontal active line period equal to 0.827 of the total line period, a horizontal spatial frequency n/A is equivalent to $0.827n/1.33 \approx 0.620n$ c/ph. Spatial frequencies in TV lines are double the corresponding figure in c/ph. Temporal frequencies are of course expressed in Hz. The coefficients of the Fourier series can be obtained from the image by the Fourier transform operation

$$U\left(\frac{n}{A}, \frac{m}{B}, \frac{l}{C}\right) = \frac{1}{ABC} \int_0^C \int_0^B \int_0^A$$

$$\times\, u(x, y, z) \exp\left[-j2\pi\left(\frac{nx}{A} + \frac{my}{B} + \frac{lz}{C}\right)\right]$$

$$\times\, dx\, dy\, dz$$

A fixed frequency component of spatial frequency (f_x, f_y) is of the form shown in Fig. 6. When f_x is large and f_y is small, we refer to the component as a high horizontal frequency. Similarly, when f_y is large and f_x is small, the component is said to be a high vertical frequency. Finally, if both f_x and f_y are large, we call the component a high diagonal frequency. If f_z is nonzero, the component will be seen to move in a direction perpendicular to the lines of constant intensity.

Relationship of the Spatiotemporal Spectrum to Spatial Content and Motion

Consider a fixed spatial pattern denoted $u(x, y)$ that has a frequency domain representation $U(f_x, f_y)$. If we now suppose that this pattern is moving in the image plane with a fixed velocity $\mathbf{v} = (v_x, v_y)$, we can show that the resulting three-dimensional spectrum is concentrated in a plane in the three-dimensional frequency space.[18] This plane is given by

$$v_x f_x + v_y f_y + f_z = 0$$

and the frequency spectrum is given by

$$U(f_x, f_y, f_z) = U(f_x, f_y)\delta(v_x f_x + v_y f_y + f_z)$$

where $\delta(\cdot)$ is a Dirac delta function. For an arbitrary image, there are many different velocities, and so the frequency spectrum does not lie in a single plane, but is dispersed throughout the frequency space. However, if a large area of the plane is translating with a certain velocity, as in a camera pan or large-area motion, there is a concentration of energy in the corresponding plane in the frequency domain.

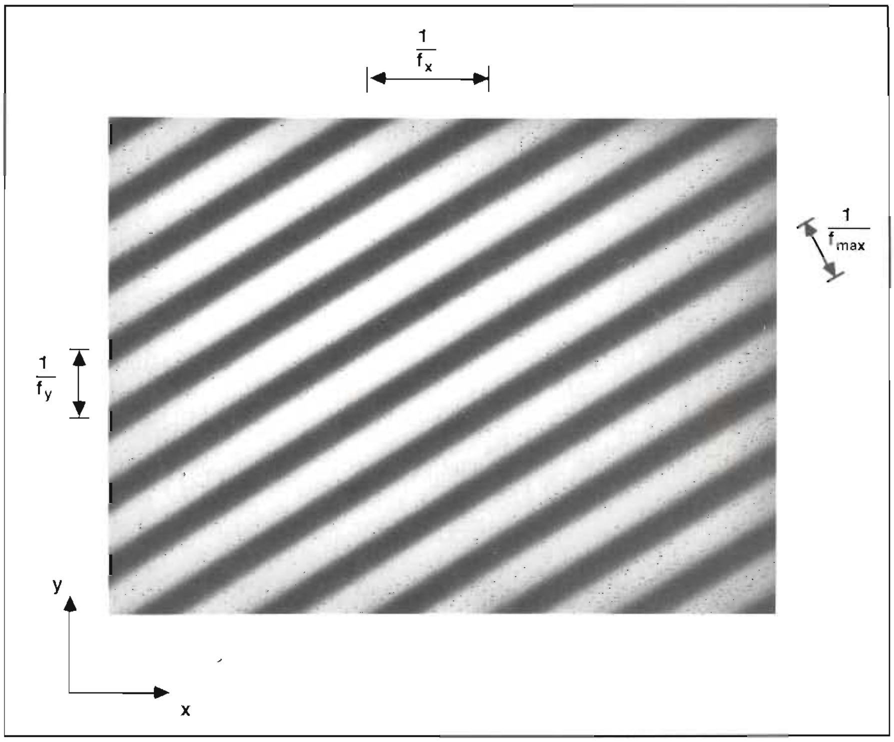

Figure 6. Illustration of a two-dimensional sinusoid. f_x is the horizontal frequency, f_y is the vertical frequency, and f_{max} is the maximum frequency.

Spectrum of Spatial Patterns

The following briefly describes the two-dimensional spectrum of a number of specific patterns that commonly occur in television pictures, including edges, lines, textures, and periodic structures.

Straight Lines and Edges

Straight lines and edges (or simply contours) are one-dimensional functions oriented in a particular direction. Suppose the contour passes through the origin and is at an angle α with the horizontal. There is no spatial variation in the direction parallel to the edge contour. Thus, the spectrum is concentrated on a line in the two-dimensional frequency domain. The value of the spectrum along this line is given by the one-dimensional Fourier transform of the profile of the contour. For a step edge, this has the form $1/f$, while for a line, considered as a narrow pulse, it has the form of a sinc function (Fig. 7).

The spectrum of a picture that contains many lines and edges oriented in various directions is seen to have many needle-like components passing through the origin in many directions. These directions are, of course, related to the orientations of the lines and edges in the picture, being perpendicular to the corresponding image elements if the appropriate scaling of the axes is used.

Textures and Periodic Patterns

Pictures often contain spatial patterns that can be described as either pseudorandom "textures" or as periodic patterns. If a pattern has a random appearance, its spectrum will tend to be of a continuous hill-like form. If the texture has a certain directionality, the spectrum will also be oriented in the corresponding (perpendicular) direction in the frequency domain.

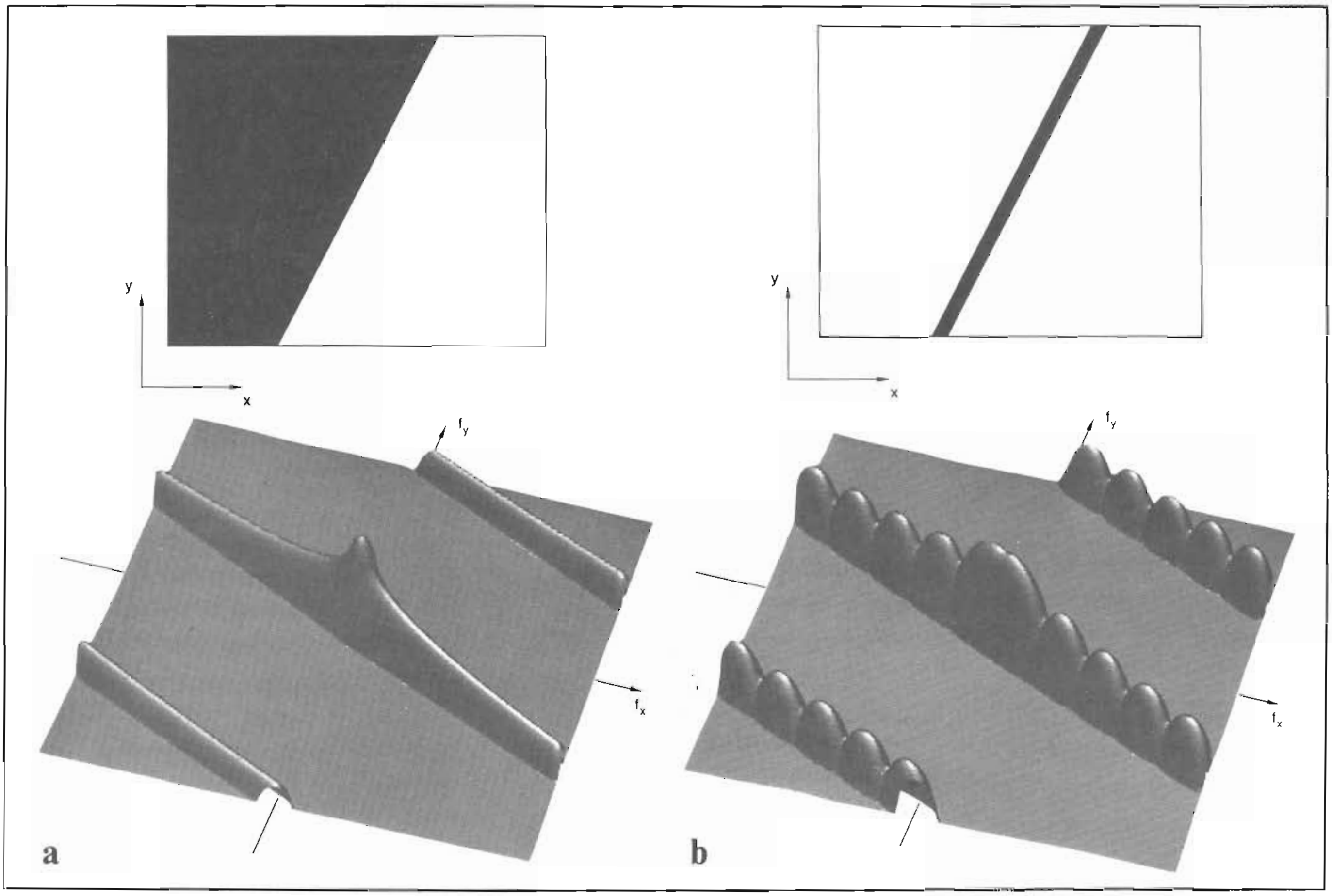

Figure 7. The spectra of straight edges and contours are one-dimensional functions in the frequency domain: (a) step edge with 1/f spectrum; (b) line with sinc(f) spectrum.

On the other hand, if the texture has a strong periodicity, the spectrum will have energy located in clumps at the harmonics of the fundamental frequency of the pattern. If the pattern is basically one-dimensional, the spectrum will be concentrated at regular intervals along a line in the frequency domain. If it is basically two-dimensional, it will be concentrated on a two-dimensional grid. Figure 8 shows an example of a quasi-periodic and a random texture, with their spectra.

Spectrum of the Scanned Signal

The scanning operation is equivalent to a vertical-temporal sampling of the image, followed by the concatenation of the scanned lines to form a one-dimensional temporal signal. The sampling has the effect of replicating the basic spectrum of the optical image in the vertical and temporal frequency dimensions. The concatenation of the lines is equivalent to a scanning of the spatiotemporal spectrum of the sampled image. The two operations can be analyzed in one step as we have shown. The result is that a spatiotemporal frequency $(f_x, f_y, f_z) = (n/A, m/B, f_z)$ corresponds to the video frequency

$$nf_h - mf_v + f_z = (525n - 2m)\frac{f_v}{2} + f_z \qquad (1)$$

where m and n are integers. Conversely, for a video frequency of f Hz, the corresponding horizontal spatial frequency is approximately $(f/f_h)/B$ or $0.62 f/f_h$ c/ph. In particular, 1 MHz corresponds to 39.4 c/ph and 4.2 MHz corresponds to 165.5 c/ph. An infinite number of spatiotemporal frequency components are mapped to the same video frequency and are thus indistinguishable.

From Eq. 1, any spatiotemporal frequency of the form

$$(n/A, m/B, (2m - 525n)f_v/2) \qquad (2)$$

is mapped to zero video frequency. Thus, any two spatiotemporal frequencies that differ by an amount of this form map to the same video frequency; we say that these spatiotemporal components are *confused*, or *aliased*. From this it is clear that temporal frequencies can only be aliased if they are separated by a multiple of $f_v/2$.

Figure 9 shows the first few spatiotemporal frequencies that are aliased with zero frequency. The effect of the sampling is to replicate the spectrum of the optical image at the points shown in Fig. 9, resulting in a periodic spectrum. We can be assured that no aliasing will occur if all spatiotemporal components present in the image are closer to the origin than to any of these points. Thus, the original image spectrum before scanning should be confined to the region shown in Fig. 2.[h] If this is so, it should be possible to reconstruct the continuous image at the receiver by an appropriate vertical-temporal interpolation, the characteristics of which depend, to some extent, on consumer preference and details of the receiver design.

Three-Dimensional Spectrum of the NTSC Signal

Again we recall that a spatiotemporal frequency $(f_x, f_y, f_z) = (n/A, m/B, f_z)$ corresponds to the video frequency $nf_h - mf_v + f_z$, where m and n are

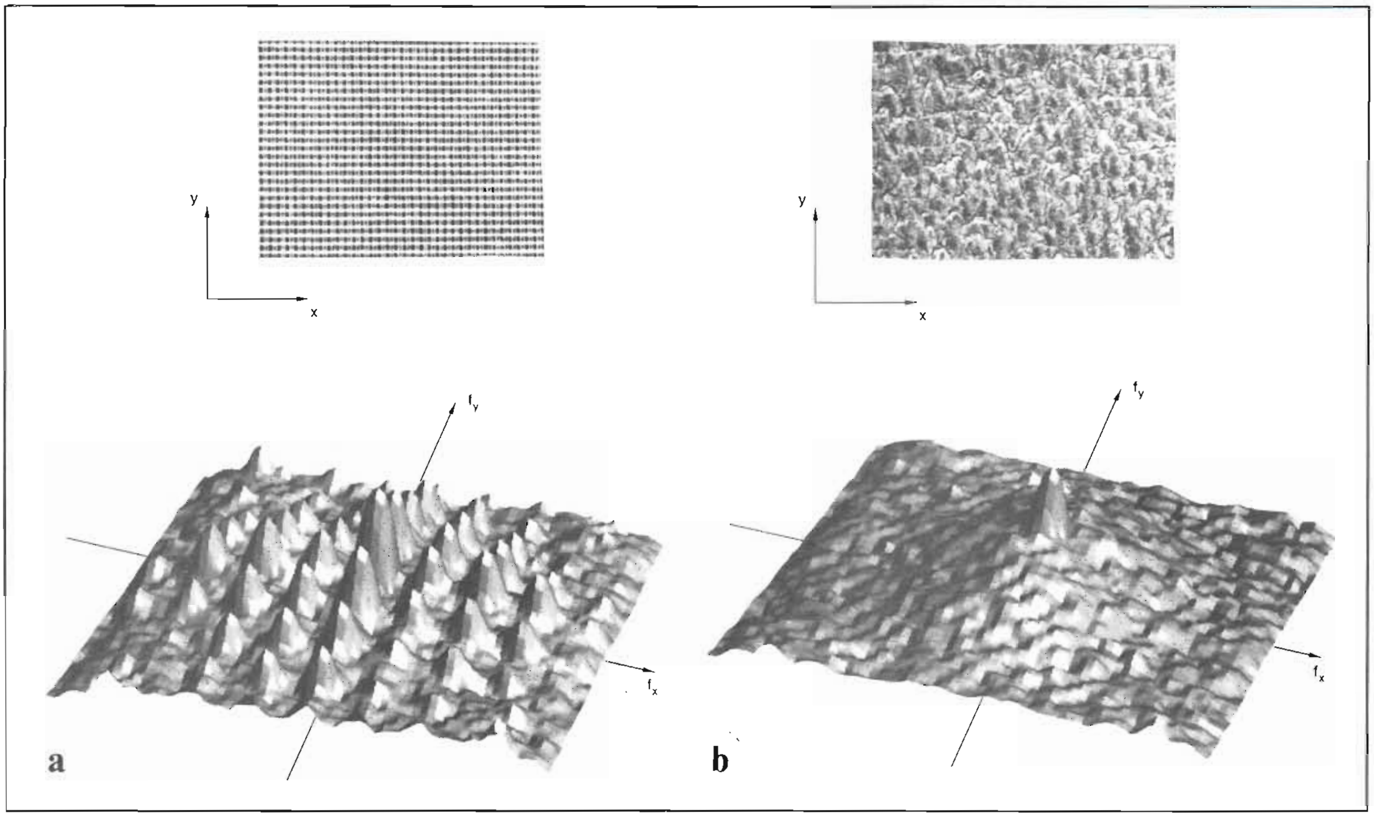

Figure 8. The spectrum of textured images: (a) a quasi-periodic two-dimensional texture whose spectrum has significant energy at harmonics of the fundamental frequency; (b) a "random" texture with a continuous spectrum.

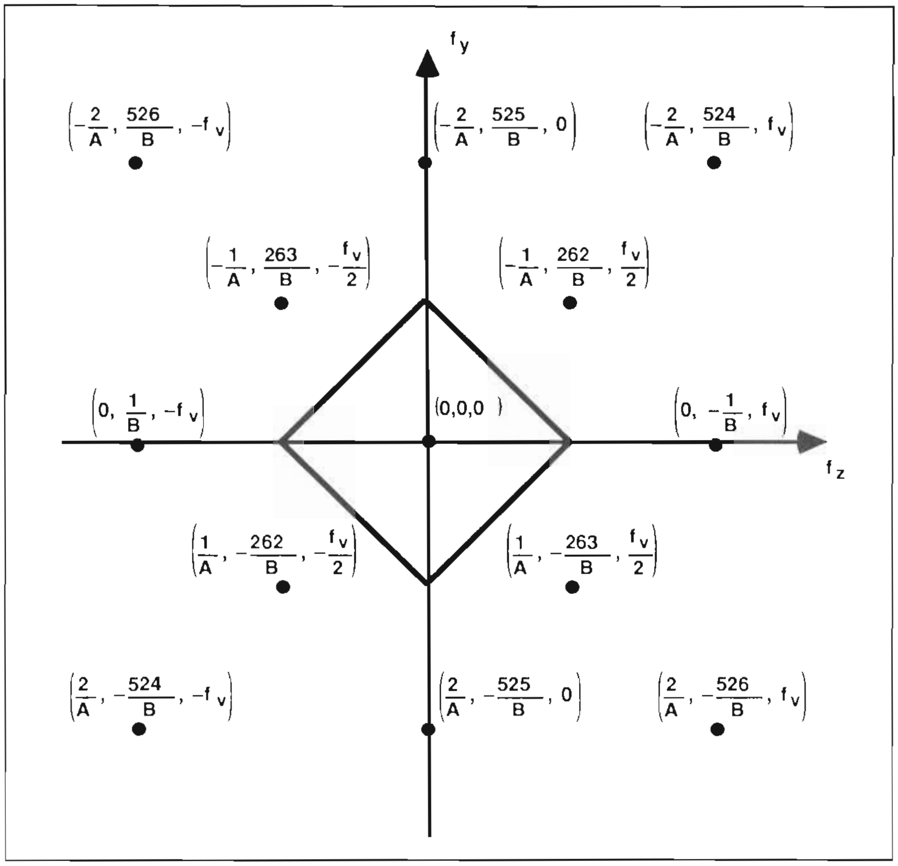

Figure 9. Projection in vertical-temporal plane of spatiotemporal components which are aliased with zero frequency because of the vertical-temporal sampling of the scanning process. The effect of the scanning is to replicate the baseband spectrum of the optical image on each of these points; aliasing occurs if they overlap. To avoid aliasing, the vertical-temporal spectrum should be confined to the indicated region.

integers. Now the subcarrier frequency is given by

$$f_{sc} = \left(\frac{455}{2}\right) f_h$$

$$= \left(\frac{455}{2}\right)\left(\frac{525}{2}\right) f_v.$$

This can be written in the form above as

$$f_{sc} = 227 f_h + 131 f_v + f_v/4$$

or

$$f_{sc} = 228 f_h - 131 f_v - f_v/4$$

so that f_{sc} corresponds to the spatiotemporal frequencies

$$\left(\frac{227}{A}, \frac{-131}{B}, \frac{f_v}{4}\right) \text{ and }$$

$$\left(\frac{228}{A}, \frac{131}{B}, \frac{-f_v}{4}\right).$$

The components corresponding to $-f_{sc}$ are at

$$\left(\frac{-227}{A}, \frac{131}{B}, \frac{-f_v}{4}\right) \text{ and }$$

$$\left(\frac{-228}{A}, \frac{-131}{B}, \frac{f_v}{4}\right).$$

In units of c/ph and Hz, these frequencies are (±140.74 c/ph, ∓120.78

c/ph, ±14.985 Hz) and (±141.36 c/ph, ±120.78 c/ph, ∓14.985 Hz). Because of the scanning, the subcarrier frequency also corresponds to an infinite number of spatiotemporal components displaced by frequencies of the form in Eq. 2 such as are shown in Fig. 9.

The process of modulating the color subcarrier by the chrominance components has the effect of shifting the 3-D spectrum of the chrominance to these spatiotemporal frequencies. Adding them to the luminance results in a spatiotemporal spectrum for the composite signal of the form shown in Fig. 3. An alternate derivation of the form of the spatiotemporal spectrum of the NTSC signal using direct 3-D analysis of the modulation process is given by Dubois et al.[8,19] Figure 10 shows the spatial and the vertical-temporal projection of the three-dimensional spectrum of a typical NTSC image sequence, of which one field is shown. Of course, these spectra are replicated in the vertical-temporal frequency dimensions, centered on the points shown in Fig. 9, because of the vertical-temporal sampling.

To perform digital processing of the NTSC signal, it is necessary to sample the signal along the scan lines, that is, in the horizontal dimension. The horizontal frequency and the spatiotemporal sampling structure must be specified in order to do this. Although many sampling structures are possible,[20,21] the simplest for subsequent processing is the aligned sampling structure, where all samples in all fields line up vertically. More complicated sampling structures, such as the "field quincunx,"[i] may be appropriate when bit-rate reduction is an issue.

The effect of aligned sampling is simply to replicate the spatiotemporal spectrum in the horizontal frequency dimension at multiples of the horizontal sampling frequency. To avoid aliasing, the sampling frequency must be greater than twice the horizontal bandwidth of the signal. Thus, sampling frequencies greater than 8.4 MHz are required. For convenience, the sampling frequency should be a multiple of the subcarrier frequency or of the line frequency. Sampling frequencies commonly used are $3f_{sc}$, $4f_{sc}$, and 13.5 MHz, the international studio standard for component signals.

Multidimensional Filtering

Many techniques for improving the performance of the NTSC system,

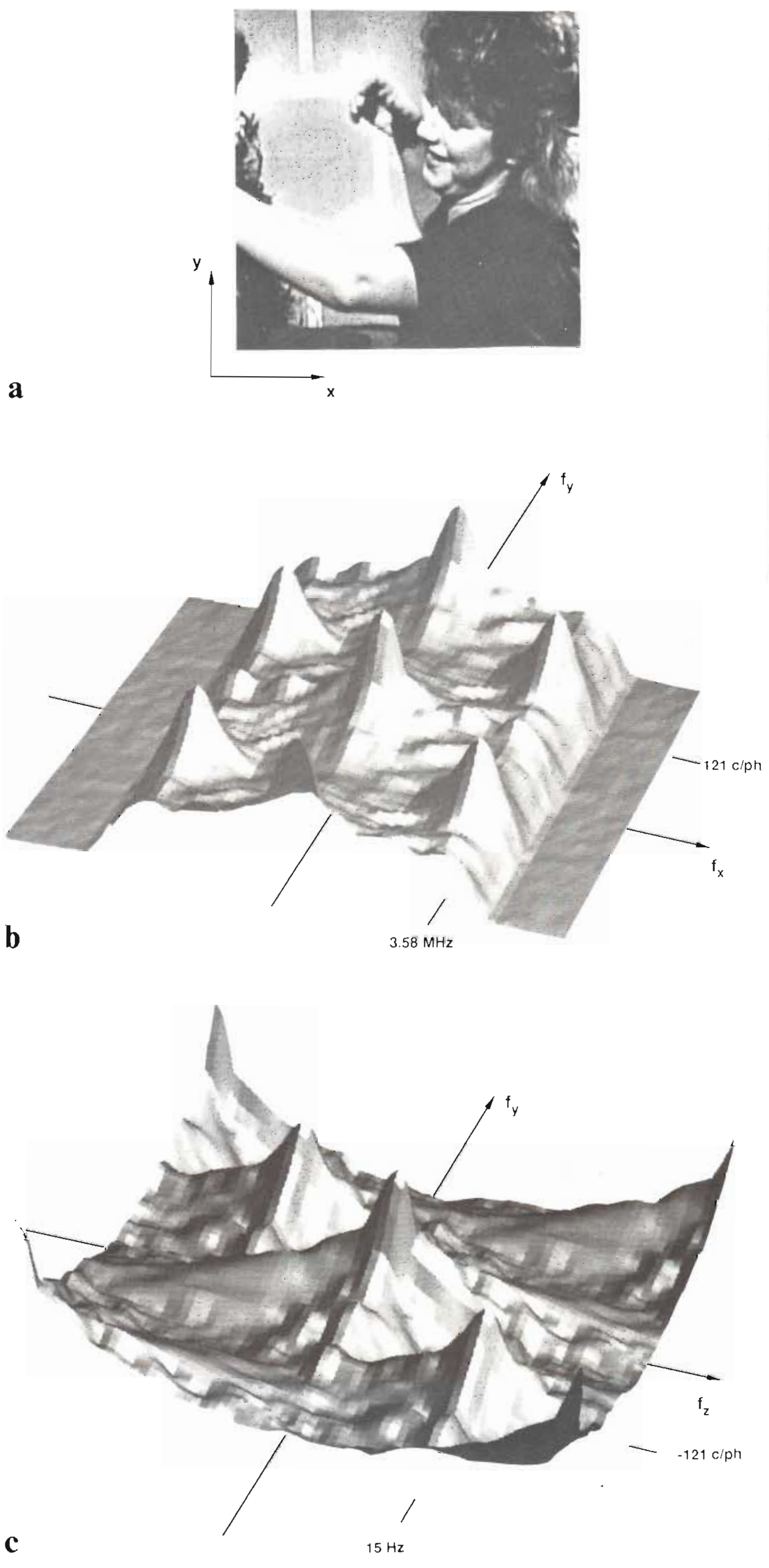

Figure 10. Projections of spectrum of NTSC signal: (a) one field on an NTSC image sequence; (b) spatial projection of the 3-D spectrum of the sequence; (c) vertical-temporal projection of the 3-D spectrum of the sequence.

particularly those described in this article, are based on two- and three-dimensional filtering. The main purpose of a multidimensional filter is to shape the spectrum of the incoming signal. For our purposes, the goal of the filter is usually to limit the spectrum of the signal to some specific subarea of the frequency domain, primarily to control aliasing or crosstalk. This subarea is called the passband. However, it is neither possible nor desirable to have an "ideal" filter whose transmission is unity in the passband and zero elsewhere. Such a filter is not realizable, and in any case, a filter with a sharp transition from passband to stopband would introduce unacceptable artifacts in the form of ringing. Thus, the problem at hand is to design filters that perform the desired spectral shaping of the signal without introducing undue degradations.

Basic Notions of Multidimensional Filters

In reviewing the basic ideas of multidimensional filters, we are primarily interested in two- and three-dimensional filters for spatial and spatiotemporal processing, respectively. Since the ideas are the same for both, we will discuss two-dimensional filters; the extension to three-dimensional filters is straightforward.

The input to a two-dimensional digital filter is a two-dimensional sampled signal $u(mX, nY)$, where X is the horizontal sample spacing and Y is the vertical sample spacing. The filter produces an output signal $z(mX, nY)$. We write $z = T[u]$ where T denotes the operation of the filter. The filter is said to be linear if the response to a weighted sum of two signals is equal to the same weighted sum of the individual responses to the two signals:

$$T[a_1u_1 + a_2u_2] = a_1T[u_1] + a_2T[u_2]$$

The filter is said to be shift invariant if a given shift in the input signal causes an equal shift in the output signal:

$$T[u(mX - m_0X, nY - n_0Y)] = z(mX - m_0X, nY - n_0Y)$$

We will deal only with linear shift-invariant filters in this article, denoting them LSI filters.

An important property of an LSI filter is that the response of the filter to a sinuosidal signal is a sinusoid of the same frequency, but with amplitude and phase modified. In terms of complex exponentials, this is given by

$$T[\exp(j2\pi(mXf_x + nYf_y))] = H(f_x,f_y)\exp(j2\pi(mXf_x + nYf_y))$$

The constant of proportionality $H(f_x, f_y)$, a function of f_x and f_y and generally a complex number, is called the frequency response of the filter and incorporates both phase and amplitude variation. It can be shown to be the Fourier transform of the impulse response of the filter. This result extends to the Fourier representation of an arbitrary signal, which is a linear combination of sinusoids. Thus if $U(f_x, f_y)$ is the Fourier transform of the input to the filter, the Fourier transform of the output is given by

$$Z(f_x,f_y) = H(f_x,f_y)U(f_x,f_y)$$

The spectral shaping operation of the filter is clear from this formula. Just as the spectra of the sampled signals are periodic, the frequency response of the two-dimensional filter is periodic:

$$H\left(f_x + \frac{k}{X}, f_y + \frac{l}{Y}\right) = H(f_x,f_y)$$

for any integers k and l.

The impulse response of a two-dimensional filter is the response to a two-dimensional unit sample: $h(mX, nY) = T[\delta(mX, nY)]$, where

$$\delta = \begin{cases} 1 & \text{if } m = 0 \text{ and } n = 0 \\ 0 & \text{otherwise} \end{cases}$$

The response of a filter to an input signal u can be expressed as a linear combination of the input signal samples, weighted by the values of the impulse response, in a form known as a convolution sum.

$$z(mX, nY) = \sum_{k=-\infty}^{\infty}\sum_{l=-\infty}^{\infty} h(kX, lY) \times u(mX - kX, nY - lY) \quad (3)$$

These concepts can be easily extended to three-dimensional filters sampled on an orthogonal structure simply by writing all the expressions with three indices instead of two. When the three-dimensional sampling structure is not orthogonal, as in the case of interlaced signals, the expressions are similar, but the indices must be restricted so that the samples lie on the desired structure. In this case, the filter impulse response is also defined on the same sampling structure as the image.[20]

Filter Types

There are several important classes of filters that should be distinguished. The main subdivision is finite impulse response (FIR) and infinite impulse response (IIR) filters. The impulse response of an FIR filter is only nonzero for a finite range of values of m and n. Then, the limits in the summation in Eq. 3 are finite, and the filter can actually be implemented by this equation. Since the output is a linear combination of a finite number of input samples only, this type of realization is often called a nonrecursive filter. IIR filters, on the other hand, result from a recursive structure, where a given filter output sample is computed as a linear combination of both input samples and previously computed output samples. Recursive filters may be unstable if the filter coefficients do not satisfy certain conditions. For two- and three-dimensional filters, these conditions are somewhat complex and complicate the filter design process.

Linear-phase filters are another important class. These have the property that all frequency components undergo the same group delay, so that phase distortion does not occur. This is particularly troublesome in image processing from a perceptual standpoint, so that exact or approximate linear phase is very desirable for image filters.

Zero-phase filters are a particular class of linear phase filters where the frequency response is purely real and there is no shift. The impulse response of such a filter must satisfy the constraint $h(mX, nY) = h(-mX, -nY)$. It is very easy to design zero-phase FIR filters by simply incorporating this constraint in the filter design process. In general, stable IIR filters do not have exactly zero phase because of their recursive nature, and the incorporation of additional constraints for approximate phase linearity is required in the design process. For these reasons, FIR filters are used almost exclusively in spatial image processing applications. However, IIR filters have the potential of offering better amplitude selectivity for a given filter order, resulting in fewer field delays for spatiotemporal filters. The use of IIR filters may thus be justified if adequate phase response can be obtained.

A final filter property of note is separability. A two-dimensional filter is said to be separable if it can be expressed as the cascade of a one-dimensional horizontal filter and a

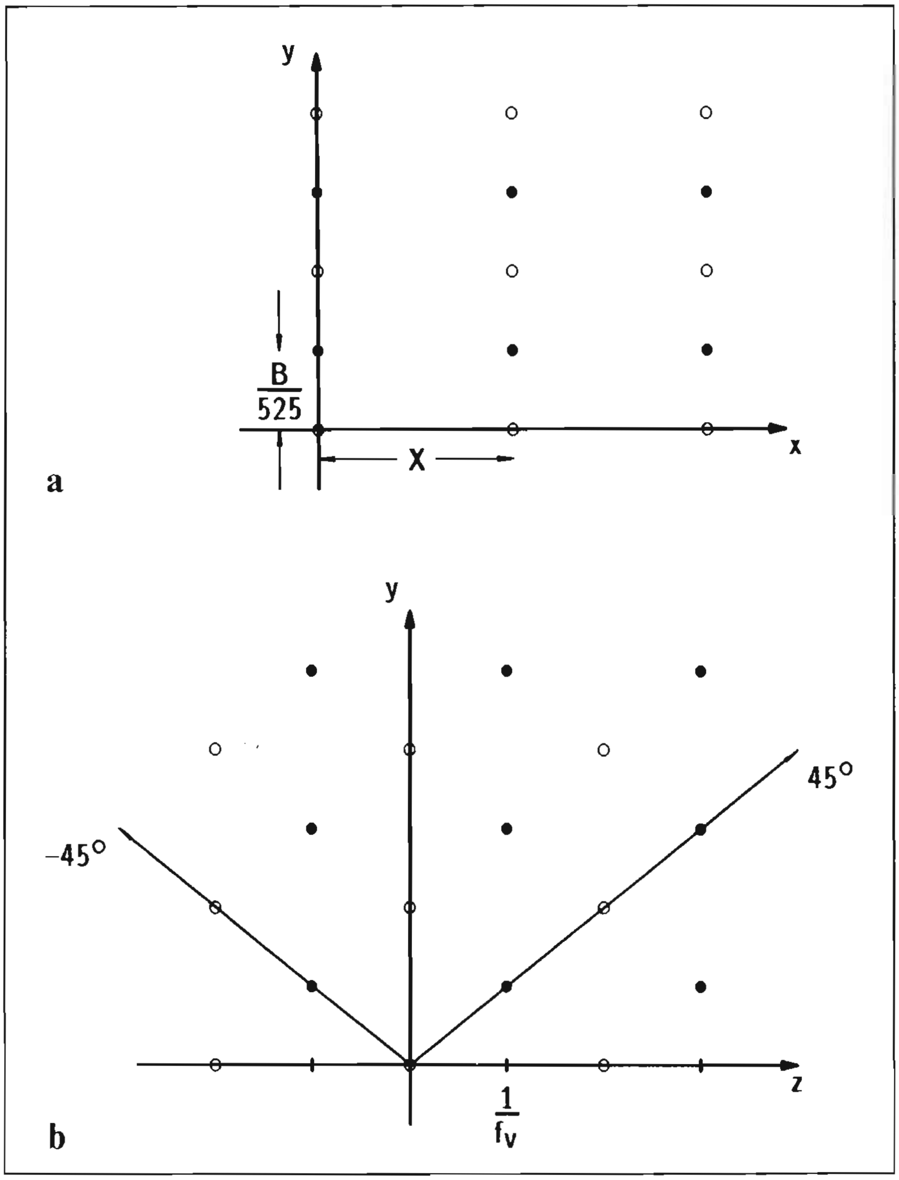

Figure 11. Aligned sampling structure for interlaced video signals: (a) spatial projection; (b) vertical-temporal projection.

one-dimensional vertical filter. Separability reduces complexity in both the design phase, where only one-dimensional filters need to be designed, and in the implementation. In general, fewer operations are required to realize a separable two-dimensional filter than an arbitrary two-dimensional filter of the same order. However, the price to be paid is a potential loss in performance because of constraints on the type of response that can be achieved. In particular, if filters with rapid cutoffs are desired, the passband of separable filters will have a rectangular shape, which may be inadequate if circular or diamond-shaped passbands are desired.

These ideas can be extended to three-dimensional filters in several ways. A three-dimensional filter can be fully separable, as a product of three one-dimensional filters in the horizontal, vertical, and temporal directions respectively. The filter can also be partially separable, as a product of a one-dimensional filter and a two-dimensional nonseparable filter. This can be done in three ways, corresponding to horizontal-vertical, horizontal-temporal, and vertical-temporal nonseparable filters.

Two- and Three-Dimensional Filtering of the NTSC Signal

The pattern for aligned sampling of the NTSC signal is shown in Fig. 11 in spatial and vertical-temporal projection. The horizontal spacing X depends on the chosen sampling frequency.

For spatial (intrafield) filtering, the vertical spacing between coefficients is $B/262.5$, so the filter has a vertical periodicity of $0.922(262.5) = 242$ c/ph. Thus, the vertical response of the filter can be specified only up to 121 c/ph, which is only half of the vertical bandwidth available for still pictures (see Fig. 2). This can be a serious limitation of intrafield filters when one desires to modify the vertical frequency components above 121 c/ph independently of those components below 121 c/ph. Vertical-temporal filters are required for this, but of course their response in the presence of motion may be a problem.

Any vertical-temporal filter has the same periodicity as the spectrum of the signal. Thus, a single period of the frequency response is the region shown in Fig. 9, and this basic period is replicated on the points shown in the figure. Similarly, any three-dimensional filter has a basic period which is the region shown in Fig. 2.

Any three-dimensional filter can be considered a one-dimensional filter operating on the temporal video signal. In this case, a vertical delay of one (intrafield) sample is equivalent to a temporal delay of one line period in the video signal. A purely vertical filter is equivalent to a one-dimensional filter with coefficients separated by one line period. The frequency response of such a filter is periodic, the period being f_h. A filter whose coefficients are separated by a constant delay has a periodic frequency response, resembling the shape of a comb, and is sometimes referred to as a comb filter. The one described above is called a line-comb filter.

Other comb filters are also useful. The next order of complexity is the field-comb, of which there are two types. They involve choosing the line above or the line below in the previous field, corresponding to a delay of 262 or 263 lines respectively. These filters are equivalent to vertical-temporal filtering in the −45° and +45° directions indicated in Fig. 11b. Finally, we have frame-comb filters, using a delay of one frame or 525 lines.

Three-dimensional filters can be constructed by interconnecting various one-dimensional filters corresponding to comb filters. Most proposed approaches to NTSC filtering are of this type. However, this imposes limitations on the obtainable shape of the pass region, as mentioned.

Filter Design

Filter design involves the determination of the impulse response (often called the filter coefficients) required to achieve the desired level of performance. Among the parameters that may determine the performance are passband ripple, stopband transmission, overshoot in the step response, etc. If precise specifications can be placed on these quantities (including the definition of the passband and the stopband), then a numerical design technique can be employed to determine a filter of minimum order meeting the specifications.

The step-response overshoot (which is manifested in pictures as ringing about sharp edges) is normally related to the width of the transition band between the passband and the stopband. The step-response overshoot increases as the width of the transition band is decreased. Thus ringing can usually be controlled by allowing a sufficiently wide transition band. An example of a typical pass-stop design specification is shown in Fig. 12. The desired filter response is 1 in the passband and 0 in the stopband. The value in the transition band is unconstrained, although it should fall smoothly and monotonically from the passband to the stopband.

Many design techniques for determining a filter approximating the desired frequency response have been published. A survey of such techniques is given in the book by Dudgeon and Mersereau.[22] The minimax design criterion is usually used, whereby the maximum deviation from the desired response in the passband and the stopband is minimized. If desired, different weights can be placed on passband and stopband ripple.

Effects of Filtering on Spatiotemporal Patterns and Images

The parameters given for filter design, while mathematically sufficient for the purpose, can themselves be derived only by considering their relationship to image quality. There are two ways in which filtering affects image quality. One has to do with the vertical-temporal sampling inherent in the interlaced scanning structure. The other involves separation of luminance and chrominance. In both, a particular bandwidth is assigned to each component. Filters are required that confine the signal, more or less, to the assigned bandwidth in order to control aliasing or crosstalk. In addition, the sharpness of the image rendered by each component must be as high as possible without introducing other artifacts, such as overshoot or increased visibility of the sampling structure.

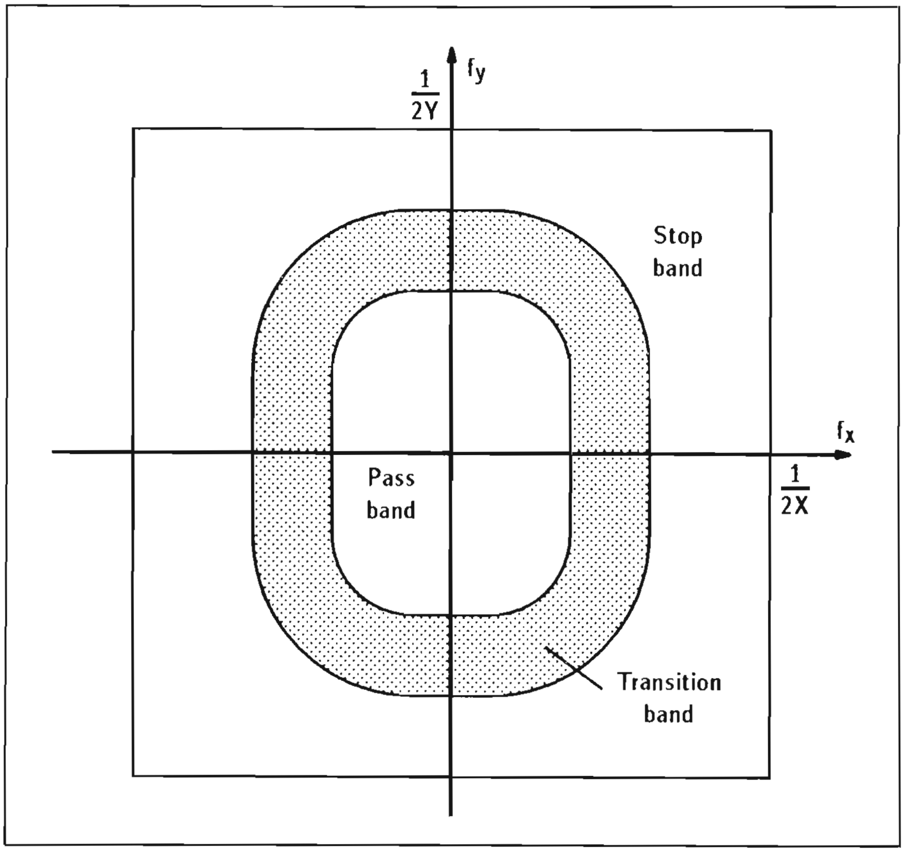

Figure 12. Typical specification for a two-dimensional filter design.

It is important to understand the effect that multidimensional filtering has on real images. A few examples of filters applied to real pictures and the associated effects are presented here.

Figure 13a shows a test picture to which a number of filters have been applied. On the left is a zone plate; the response of a filter to this pattern gives an illustration of the filter's spatial frequency response. On the right is a natural scene. Figures 13b to 13d show the effect of different circularly symmetric filters on this image. All the filters have approximately the same 3-dB contour, but with different transition bands. Figure 13b shows the response to a Gaussian filter; the image shows marked blurring. The filters used in Figs. 13c and 13d were obtained using the McClellan transformation technique. The filter in 13c has a maximally flat radial profile, while the filter in 13d has a sharper, equiripple profile. The image in 13c is less blurred than that in 13b, and does not suffer from ringing. The sharper filter shown in 13d gives the least blurring, but the image suffers from significant ringing. These examples highlight the trade-off between resolution loss and ringing in two-dimensional filter design.

Figures 13e and 13f show examples of purely horizontal and vertical filters respectively, with their effect on horizontal and vertical resolution. Finally, Fig. 13g shows an example of a filter with a rectangular stopband at high horizontal and vertical frequency.

Multidimensional Filters for NTSC Encoding and Decoding

The NTSC encoding process, as shown, is a three-dimensional frequency multiplexing operation for luminance and chrominance. At the present time, contrary to conventional practice in frequency multiplexing (at least in one-dimensional systems), steps are not normally taken to prevent overlap in the multiplexed signals. This results in significant crosstalk between luminance and chrominance, as described. This crosstalk can be effectively eliminated, or at least greatly reduced, by performing suitable prefiltering of luminance and chrominance before the multiplexing operation.

Generalized NTSC Encoders and Decoders

Prefiltering is discussed here in the context of a generalized NTSC

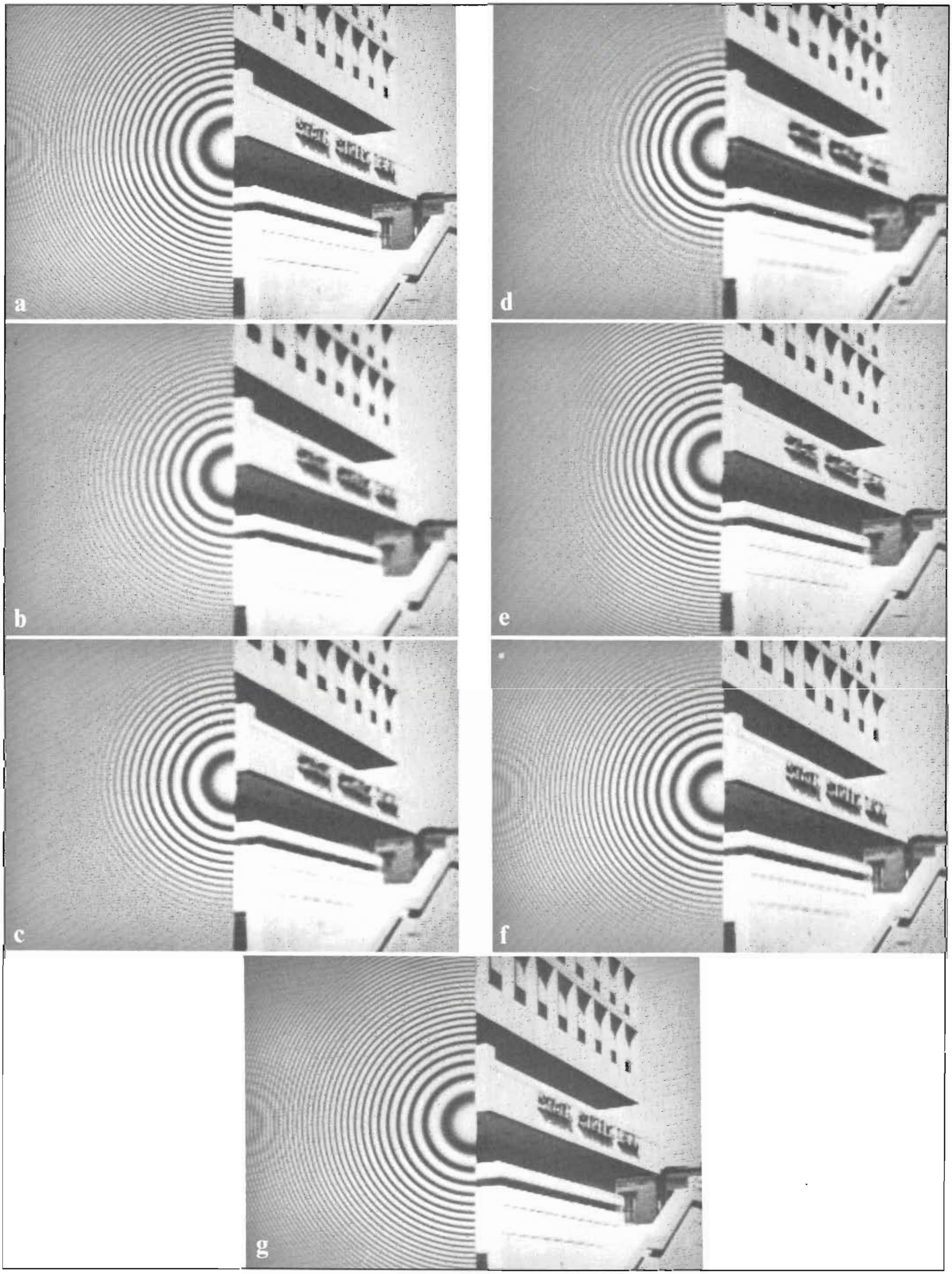

Figure 13. Effects of different filters on a still picture: (a) original; (b) Gaussian low-pass with circular symmetry; (c) maximally flat low-pass with circular symmetry; (d) sharp cutoff FIR filter with circular symmetry; (e) maximally flat horizontal low-pass; (f) maximally flat vertical low-pass; (g) maximally flat high-frequency rejection filter.

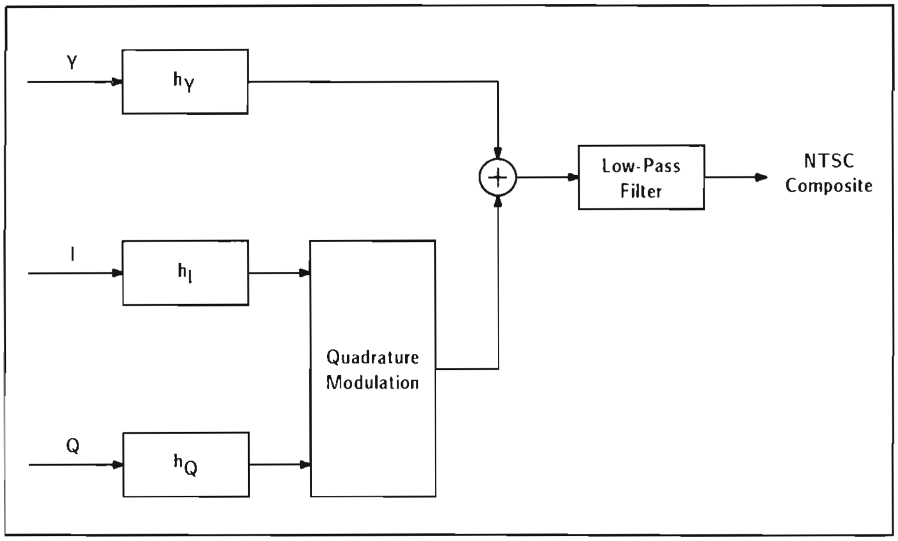

Figure 14. Block diagram of generalized NTSC encoder using multidimensional filters.

encoder (Fig. 14), of which the conventional encoder with one-dimensional filters is a special case. In the general case, the *Y*, *I*, and *Q* signals are each subjected to a separate multidimensional filtering operation before modulation and multiplexing. The purpose of the filtering, which can be one-, two-, or three-dimensional, is to limit the crosstalk that would otherwise occur between luminance and chrominance components. After multiplexing, the composite signal is filtered with a one-dimensional 4.2-MHz low-pass filter for further summation with the audio, RF modulation, and transmission. We will not be concerned with these subsequent operations and will refer to the 4.2-MHz vision signal as the NTSC signal.^{k}

The goal in designing the filters for the generalized encoder is to maximize the quality obtained after decoding the composite signal, subject to constraints such as compatibility with existing standards. (Of course, we are also concerned with the cost and complexity of the filter.) This optimization involves a trade-off between such effects as sharpness, crosstalk, and ringing. Furthermore, the optimization may depend on the type of decoder used. Currently, one-dimensional notch filter decoders and two-dimensional comb-filter decoders are most common. However, new decoders could be proposed with improved performance.

In principle, since spatiotemporal frequency multiplexing is being performed, three-dimensional filters should be used in the decoder, even if only one-dimensional prefilters are used at the encoder. Thus, we can also consider a generalized NTSC decoder, as shown in Fig. 15. In this system, the received composite signal is passed through a multidimensional stopband filter to remove the chrominance component and thus derive the luminance component. It is also passed through a multidimensional bandpass filter to extract the chrominance component for subsequent demodulation into *I* and *Q*. Again, these filters can be one-, two-, or three-dimensional, so that the commonly used decoders fit within this structure. Although the structures are general, specific implementations may have different forms in order to reduce complexity due to redundant or duplicated operations. Thus luminance and chrominance filters may share common elements, part of the chrominance prefiltering may take place after quadrature modulation, and so forth.

Filters for the Generalized Encoder

The purpose of the filters in the generalized NTSC encoder is to reduce crosstalk between luminance and chrominance components introduced by the frequency mulitplexing, without unduly sacrificing resolution. In fact, the optimal compromise between these two sources of degradation is sought. The best solution will depend on the dimensionality of the filters used in the encoder and decoder.

In the design of the encoder, compatibility is a major concern. The generalized encoder must respect any standards in force for conventional encoders. Furthermore, the use of a new encoder should not greatly reduce the quality obtained with a conventional decoder, while, of course, significant quality improvement should be expected with the use of a matched decoder.

One-Dimensional Prefilters

In one-dimensional encoders, which are the standard encoders now in use, the *Y* signal is not prefiltered at all, while the *I* and *Q* signals are filtered to approximately 1.3 MHz and 0.5 MHz respectively. The FCC has specified some constraints on these filters, shown in Table 1. Any two- or three-dimensional filters for *I* and *Q* should also respect the specifications of Table 1 in regard to their horizontal frequency response. Of course, it would be possible to filter the luminance component so that it does not contain significant energy in the band occupied by the chrominance components. However, this would entail low-pass filtering of the luminance to about 2.3 MHz, resulting in severe loss of resolution. In fact, the resolution loss is judged more severe than the elimination of crosstalk impairments, conflicting with our desire for an optimal trade-off. Therefore, in practice, the luminance component is not filtered at all.

Two-Dimensional Prefilters

Two-dimensional (intrafield) fil-

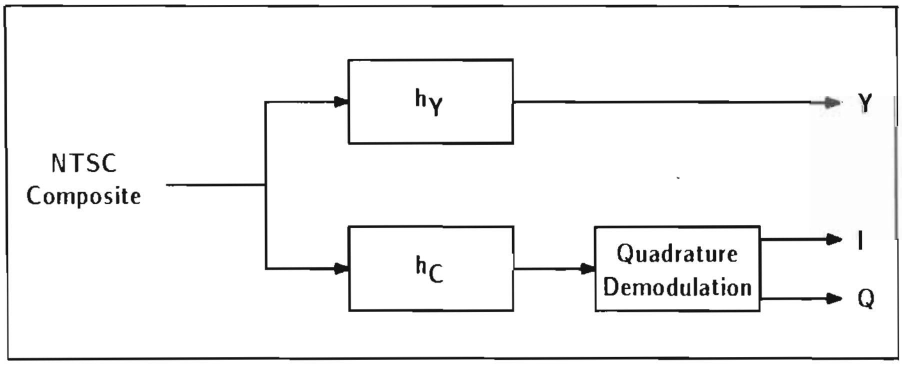

Figure 15. Block diagram of generalized NTSC decoder using multidimensional filters.

Table 1 — Specifications for the Horizontal Response of Filters for *I* and *Q*

Component	Frequency	Attenuation
I	1.3 MHz	< 2 dB
	3.6 MHz	> 20 dB
Q	0.4 MHz	< 2 dB
	0.5 MHz	< 6 dB
	0.6 MHz	> 6 dB

ters can be used to prefilter luminance and chrominance to separate bands. In three-dimensional frequency space, these bands are four cylinders perpendicular to the spatial frequency plane. The vertical chrominance resolution can be reduced, allocating the recovered spectral space to the luminance. A reasonable choice is to allocate the chrominance components a vertical spatial frequency bandwidth comparable to the horizontal spatial frequency bandwidth of the *I* signal. The main issue is the choice of the shapes of the passbands (that is, the cross section of the cylinders) for the three components; the only constraints relate to the horizontal component of the *I* and *Q* signals.

Figure 5 shows some possible allocations of the available frequency space to luminance and chrominance. These correspond to rectangular, circular, and diamond-shaped passbands for the chrominance components. Assuming that the horizontal and vertical extents of the chrominance passband correspond to the specifications of Table 1, the diamond shape is the one that allocates the greater part of the available spectral space to the luminance component. Experiments carried out at INRS-Télécommunications indicate that this gives the most favorable tradeoff, the additional luminance resolution being quite noticeable in some cases.[12]

Given the choice of frequency bands, the next issue is that of filter design. The simplest approach is to use a combination of horizontal and line-comb filters, as proposed by Turner[9] and Faroudja.[23] The chrominance components are prefiltered by a separable low-pass filter, whereas the luminance is prefiltered by the complement of a separable bandpass filter. The structure of this encoder is shown in Fig. 16. A vertical (line-comb) low-pass filter for each chrominance component can be replaced

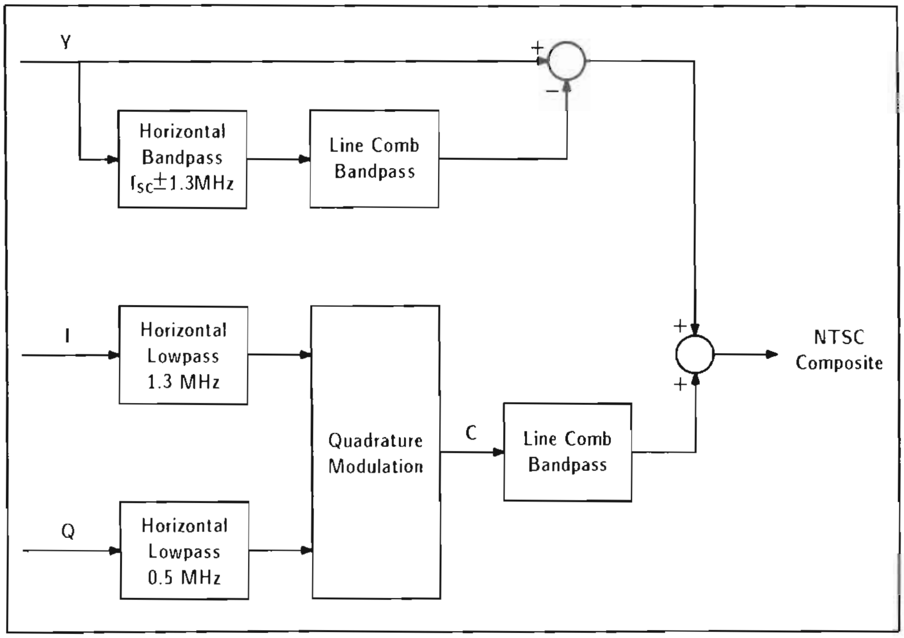

Figure 16. NTSC encoder using horizontal filters and line-comb filters.

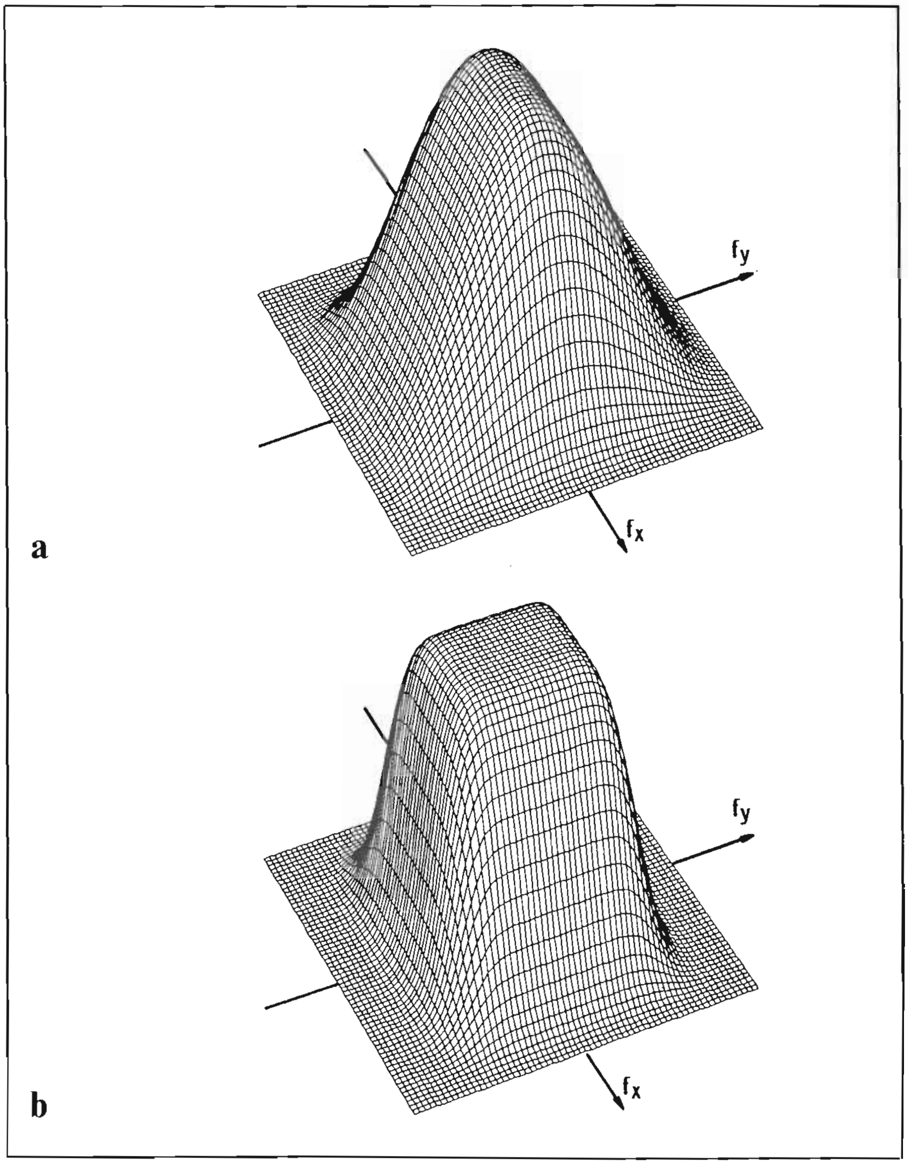

Figure 17. Response of the I signal prefilter using the encoder of Fig. 16: (a) comb filter with 2 line delays; (b) comb filter with 20 line delays.

by a single line-comb bandpass filter on the quadrature-modulated chrominance signal. Figure 17a shows a perspective view of the frequency response of the *I* chrominance prefilter using two line delays proposed by Faroudja.

Such a low order for the vertical filter is not sufficient to eliminate crosstalk between the components, nor does it give very high vertical sharpness. The vertical response can be made sharper by using more lines in the comb filter.[10] An example using a 21st-order maximally flat filter in the vertical direction is shown in Fig. 17b. However, the passband of the chrominance remains rectangular, and luminance diagonal resolution is significantly reduced.

It is not possible to achieve the diamond-shaped chrominance band with separable filters; nonseparable filters are required. Such filters have been obtained using an approximate equiripple design method by Dubois and Faubert.[12] The design of these filters involved an attempt to find a good trade-off between ringing, cross effects, and resolution. Perspective views of the filter responses obtained for the *Y*, *I*, and *Q* prefilters are shown in Fig. 18.

Three-Dimensional Prefilters

The goal of three-dimensional filtering is to limit the chrominance spectrum in three dimensions and to allocate space in four of the eight octants in which to place it. One possible approach, proposed by Faroudja[23] and further discussed by Strolle,[11] is the use of line- and frame-comb filters. This is simply achieved by plac-

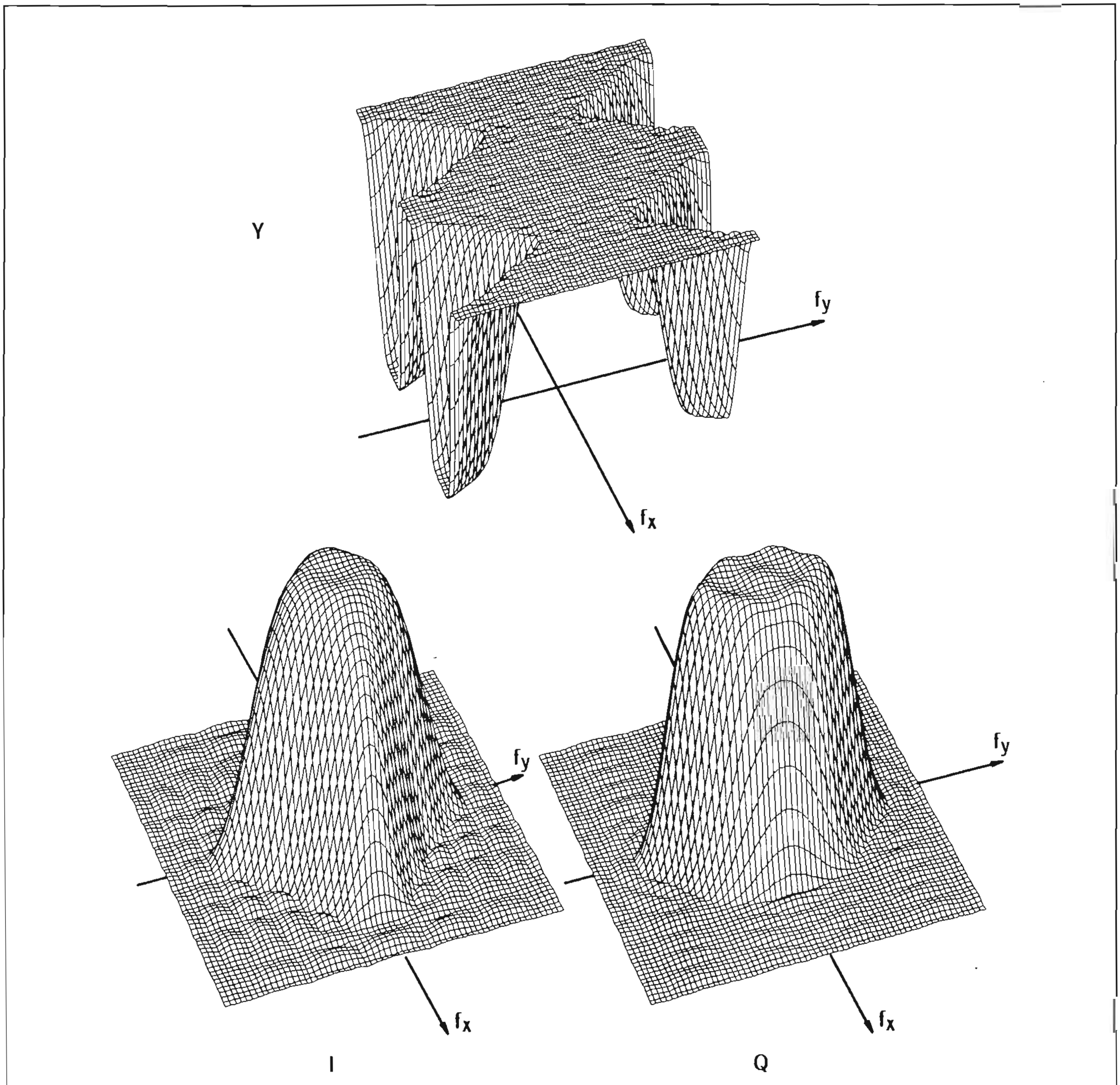

Figure 18. Response of the Y, I, and Q prefilters for a system using a diamond-shaped chrominance band. The filters are nonseparable FIR filters with a vertical extent of 11 lines. The I and Q filters have a horizontal extent of 15 samples, and the Y filter has a horizontal extent of 31 samples.

ing a frame-comb bandpass filter after each of the line-comb filters shown in Fig. 16. Such a filter limits the chrominance to box-shaped regions, aligned with the main frequency axes, in all eight octants of the frequency space. The four octants opposite those containing the subcarrier hold chrominance information of high vertical and high temporal resolution.

Another approach is to perform separable filtering in the ±45° vertical-temporal directions (Fig. 11b) instead of line- and frame-comb filtering.[11] This is a cascade of the two types of field-comb filters discussed. This approach can be used to confine chrominance to the four desired octants, with pass regions of the form shown in Fig. 4. This is probably a good choice. Many field delays may be required to get adequate selectivity in the field-comb filters using FIR techniques. The filters reported by Strolle use six field delays in each of the field-comb filters, for a total memory requirement of twelve fields. However, Strolle reports that the use of two IIR field-comb filters, each using two field delays, gives equivalent performance, without perceptible phase distortion.

Filters for the Generalized Decoder

The decoder also can use one-, two-, or three-dimensional filters, independent of the dimensionality of the filters used at the encoder. However, the advantage of multidimensional filters at the decoder is reduced if prefiltering has not been performed, since crosstalk still exists and cannot be removed by any linear filtering operation. The generalized decoder structure described applies to one-, two-, and three-dimensional decoders.

One-Dimensional Filters

One-dimensional decoders were used in virtually all television receivers until recently, when comb-filter decoders began to appear in professional and some consumer television sets. The one-dimensional decoder uses a bandpass filter to extract the chrominance from the composite signal, from which the I and Q signals are decoded. To obtain the luminance, a notch filter is used to suppress the frequency band about the subcarrier. The notch filter is narrower than the chrominance-band filter, again since the additional luminance resolution is judged more important than impairment caused by a small amount of high-frequency modulated chrominance left in the signal.

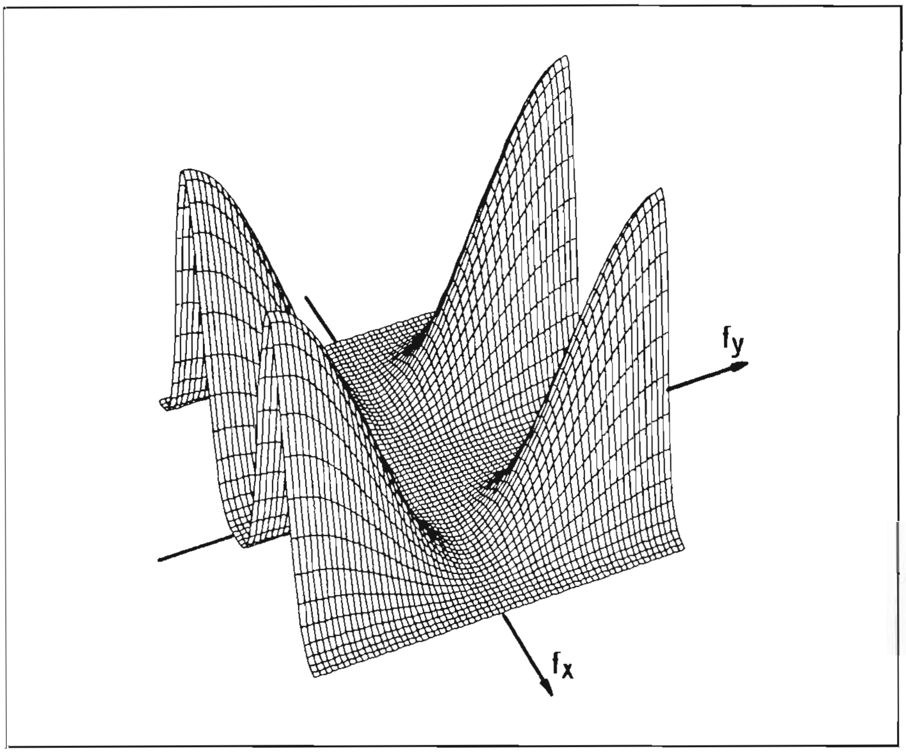

Figure 19. Chrominance filter in a line-comb decoder using two line delays.

Two-Dimensional Filters

The comb-filter decoders that have recently come into use are a form of two-dimensional filter. The chrominance bandpass filter is separable in the horizontal and vertical dimensions. The vertical order is normally 2 or 3 (that is, using one or two line delays). If no prefiltering is done, there is no reason to use higher-order filters, since crosstalk that cannot be removed is the main impairment. The response of the two-dimensional bandpass filter is shown in Fig. 19.

If an effective prefiltering scheme, such as that using the filters in Fig. 18, has been employed, then significant gain can be obtained by using higher-order decoder filters with a diamond-shaped passband. A chrominance filter designed in the same way as the prefilters of Fig. 18 is shown in Fig. 20. The luminance can be obtained by subtracting the output of the chrominance filter from the composite signal. Thus only one of the two filters in Fig. 15 is required, the function of the other filter being obtained by simple subtraction. A compensating delay may be needed.

Three-Dimensional Filters

The filters for three-dimensional decoding are very similar to those discussed for the encoder. The same principle of separable filtering in the diagonal directions in the vertical-temporal dimensions is a good approach.

Performance of One- and Two-Dimensional Encoders and Decoders

A number of simulations and demonstrations have clearly shown the considerable improvement in image quality that can be obtained with some of the multidimensional prefiltering and postfiltering techniques described here. The prefiltering has been shown to be a crucial element in eliminating the cross effects that are present in most television pictures today.

A study comparing the performance of various one- and two-dimensional NTSC encoders and decoders has recently been carried out at INRS-Télécommunications.[12] A subjective experiment to evaluate the picture quality obtained with different combinations of prefilters and decoders was performed. The prefilters used were the standard one-dimensional encoder and the diamond-shaped filters of Fig. 18. Three decoders were compared: the standard one-dimensional notch filter, a three-line comb-filter decoder, and the diamond-shaped two-dimensional filter decoder of Fig. 20. The different combinations that were tested are shown in Table 2. Also included was a component benchmark, in which the I and

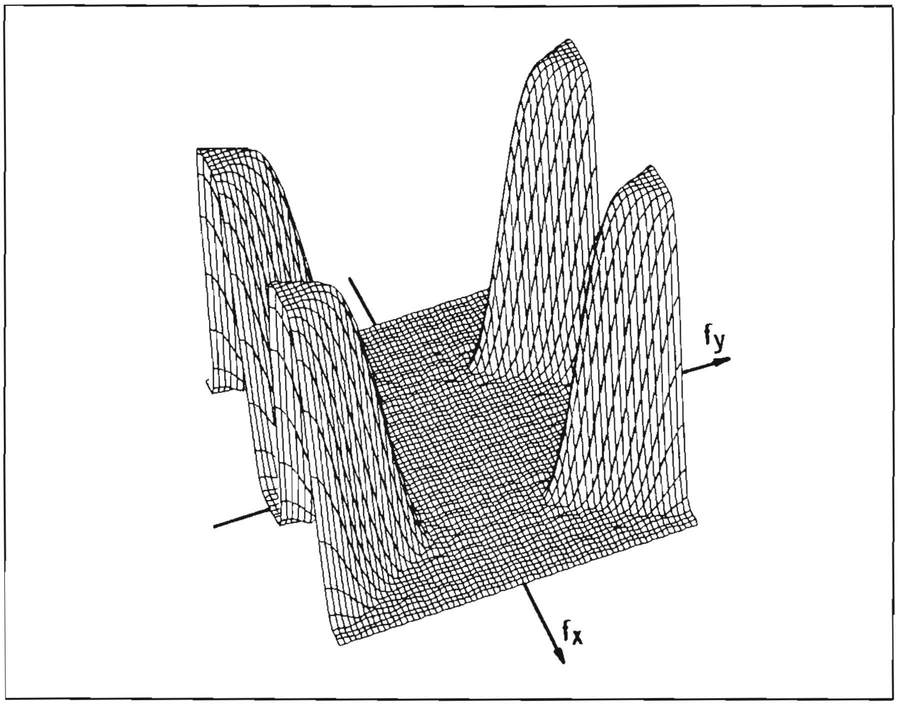

Figure 20. Chrominance filter for a diamond-shaped chrominance band. The filter is a nonseparable FIR filter with a vertical extent of 11 lines and a horizontal extent of 31 samples.

Q components were horizontally filtered according to the specifications of Table 1, but no encoding or decoding was done.

Three critical sequences containing slow motion, detailed areas, and sharp color transitions, and in general displaying the defects in the standard NTSC encoding/decoding system, were processed with the different encoder/decoder combinations described in Table 2. These sequences were rated on the CCIR 5-point quality scale in a subjective test. Fourteen observers participated, each taking the test twice. The results are shown in Table 2.

The following observations can be made concerning these results:

1. The notch-filter decoder gave the lowest quality for these sequences. The quality was essentially the same for the conventional encoder and for the two-dimensional diamond-filter encoder. Sophisticated encoding does little for conventional receivers.

Table 2 — Subjective Rating of Encoder/Decoder Combinations

Encoder	Decoder	Rating
1-D	Notch	2.1
Diamond	Notch	2.2
1-D	Comb	3.5
1-D	Diamond	3.3
Diamond	Comb	4.5
Diamond	Diamond	4.6
Benchmark		4.8

2. When the conventional encoder was used, the comb-filter decoder and the two-dimensional diamond-filter decoder gave about the same quality. With a rating of 3.3 to 3.5, this was more than a full point above the rating obtained with the notch-filter decoder. This is the quality that would be obtained with current high-quality television receivers equipped with comb filters. Such filters are a worthwhile investment.
3. When the two-dimensional diamond-filter encoder was used, the quality was improved significantly, giving a rating of 4.5 for the comb-filter decoder and 4.6 for the diamond-filter decoder. This indicates that prefiltering can give a significant improvement in quality, even if today's "conventional" comb-filter decoders are used.
4. The rating obtained for the benchmark component system, which is the highest quality possible when the specifications of Table 1 are met, was 4.8. The best composite system, using both two-dimensional diamond encoders and decoders, was within 0.2 of the benchmark performance. The small difference is probably due to some loss of diagonal luminance resolution and to chrominance ringing introduced by the filters.

The results of these experiments show that excellent results, essentially equal to the performance of a very good component system, can be obtained with a composite system using two-dimensional processing at the encoder and receiver. However, these results were obtained on only three sequences. These sequences were produced by standard current TV equipment and probably did not have the full spatiotemporal spectrum inherent in the NTSC specifications. In addition, the results undoubtedly depend on the particular viewing conditions. It is to be expected that the higher the quality of the display and the more carefully viewed, the greater the disparity among the various systems. For example, when the processed sequences were viewed on a large-screen projection display, the superiority of the two-dimensional diamond decoder over the standard comb-filter decoder, when prefiltering had been applied, became much more apparent. However, the difference between the diamond encoder/decoder combination and the benchmark remained small.

Adaptive Two-Dimensional Filters

As reported, the performance of fixed high-order two-dimensional filters is excellent. While such filters are clearly feasible at the encoder, they would significantly raise the cost of receivers. A number of systems have been proposed for obtaining very good performance with simpler filters by making them adaptive.

A series of improved decoders was developed by CBS in the early 1970s for use in frame-recursive noise reducers and in the Electronic Still Store. In general, they all used three-line combs (two line delays), but were adaptive in various ways. In Kaiser's scheme,[14,24] separate 3-line combs are used to find luminance and chrominance. The difference between the first and last lines at the subcarrier frequency is used to find vertical chrominance detail. If such detail is present, it is assumed that cross color would result, so the luminance signal is low-pass filtered.

Rossi's scheme was intended to preserve vertical resolution of chrominance by using the comb filter only in areas lacking vertical detail.[13,25] In this system, comparisons were made between bandpass signals of the upper two and lower two scan lines to locate vertical chrominance detail, if any. Depending on the location and contrast of the detail, combing was performed on the upper two lines, the lower two lines, all three lines, or not at all. Luminance was found by subtracting chrominance from the composite signal.

Faroudja's system[26] is an improvement on the CBS systems in that switching between comb filters and notch filters depends on two detail measurements — high vertical frequency/low horizontal frequency luminance and simultaneous chrominance and high-frequency luminance. It is claimed that this technique permits the use of two-line (one delay) comb filters.

Teichner's method[16] entails the use of 3-point line and frame combs. The latter achieves excellent separation and spatial resolution in stationary areas, while the former does well with moving images having low vertical resolution. A linear combination of the two kinds of processing is used, with the weighting controlled by a movement detector.

Conclusion

The potential application of multidimensional filtering techniques to improve NTSC has been discussed. These improvements are mainly related to the reduction of cross color and cross luminance and to the recovery of some of the spatiotemporal resolution that was lost when color was added to the existing monochrome system. From the discussion of the spectrum of composite signals, it becomes apparent that the three color components can be separated completely given two conditions: first, that 3-D frequency space must be divided, once and for all, among the components; and second, that appropriate filters must be used at both encoder and decoder. Under these conditions, the received image quality is negligibly lower than that of the component signals, equally bandlimited.

This conclusion is not, however, the end of the matter. If cameras and displays ever have resolution equivalent to the original monochrome standard, which is substantially higher than what is now in use, the reduction in resolution required to accommodate color may be noticeable. In addition, the filters required at the receiver are likely to be fairly expensive; one or two line delays are not enough. Finally, filtering at the encoder, although not expensive by studio standards, is impractical in composite studios.

A short review of work in adaptive comb filters for receivers indicates that some of the techniques may produce rather good pictures with less expensive filters than are required for the nonadaptive procedures.

Acknowledgment

This study was supported in part by contracts from the Center for Advanced Television Studies. The work on design of nonseparable two-dimensional filters for NTSC encoding and decoding was carried out at INRS-Télécommunications under a research contract from the Canadian Broadcasting Corporation. Louis-Georges Lefebvre, of INRS-Télécommunications, produced most of the figures and images in this paper.

Notes

a. Most of the remarks here hold for PAL as well.

b. In this article, we deal only with the NTSC problems associated with the addition of color. Motion defects and artifacts associated with interlace, although important, are not covered.

c. If the "Fukinuki hole" is used, even higher horizontal resolution may be achievable.

d. The method of analysis of Mertz and Gray is used throughout this article. We assume here a general familiarity with basic electrical engineering and television concepts.

e. This clearly requires a filter that passes one set of harmonics and rejects the other, thus having the shape of the teeth of a comb: a "comb" filter. For an early suggestion for using comb filters for this purpose, see Ref. 3.

f. Thus the fundamental frame rate of a color signal is 15 Hz, not 30 Hz, a fact that causes some complications when single frames stored in composite form are edited or continuously replayed. Note that the out-of-phase condition at corresponding points in successive frames, and at vertically adjacent points in successive lines in each field, occurs only if the image information is identical at the two associated points. It is more precise to say that the *reference* phase at these points is reversed. Note also that the harmonics of a stationary color image are separated by 15 Hz, not 30 Hz, as in a stationary monochrome image.

g. In usual practice, the spatiotemporal spectrum is not limited to a region of the type shown, and aliasing does occur. This can be avoided by initially scanning the signal at a higher density (for example, with twice the number of scanning lines per picture height) and performing vertical-temporal prefiltering to limit the spectrum to the indicated region. The signal can then be rescanned at the desired density without aliasing. This additional technique for improving NTSC by multidimensional filtering is not further addressed in this article.

h. In 2- or 3-D sampling patterns, there is no unique shape of the alias-free baseband, although there is generally, as in this case, a most "natural" one.

i. With the field quincunx structure, samples within a field are vertically aligned, while samples in successive fields are offset by half the horizontal sampling period.

j. It is possible to realize an FIR filter with a recursive structure.

k. For some applications which do not involve normal RF transmission, the NTSC signal bandwidth may be greater than 4.2 MHz. This has little effect on the discussion here.

References

1. P. Mertz and F. Gray, "A Theory of Scanning and Its Relation to the Characteristics of the Transmitted Signal in Telephotography and Television," *Bell Syst. Tech. J.*, *13*:464–515, 1934.
2. F. Gray, "Electro-optical Transmission System," U.S. Patent No. 1,769,920, 1929.
3. C. R. Trimble, "What Is Signal Averaging?" *Hewlett-Packard J.*, 2–13, Apr. 1968.
4. G. T. Flesher et al., "The General Theory of Comb Filters," in *Proc. NEC*, 282–295, 1958.
5. N. W. Parker, "Color TV Signal Separation System," U.S. Patent No. 3,542,945, Nov. 1970.
6. J. O. Drewery, "The Filtering of Luminance and Chrominance Signals to Avoid Cross-Colour in a PAL Colour System," *BBC Eng.*, *104*:8–39, Sept. 1976.
7. S. J. Auty, D. C. Read, and G. D. Roe, "PAL Color Picture Improvement Using Simple Analog Comb Filters," *SMPTE J.*, *87*:677–681, Oct. 1978.
8. E. Dubois, M. S. Sabri, and J.-Y. Ouellet, "Three-Dimensional Spectrum and Processing of Digital NTSC Color Signals," *SMPTE J.*, *91*:372–378, Apr. 1982.
9. R. Turner, "Some Thoughts on Using Comb Filters in the Broadcast Television Transmitter and at the Receiver," *IEEE Trans. Consum. Electron.*, *CE-23*:248–256, Aug. 1977.
10. M. Tsinberg and E. Fisch, "A System for Artifact-Free NTSC Encoding and Decoding," *IEEE Trans. Consum. Electron.*, *CE-32*:228–236, Aug. 1986.
11. C. H. Strolle, "Cooperative Processing for Improved NTSC Chrominance/Luminance Separation," *SMPTE J.*, *95*:782–789, Aug. 1986.
12. E. Dubois and P. Faubert, "Two-Dimensional Filters for NTSC Color Encoding and Decoding," in *Proc. Int. Broadcast. Conv.*, Brighton, U.K. 252–255, 1986.
13. J. Rossi, "Comb Filter for Television Signals Having Adaptive Features," U.S. Patent No. 4,050,084, Sept. 1977.
14. A. Kaiser, "Comb Filter Improvement with Spurious Chroma Deletion," *SMPTE J.*, *86*:1–5, Jan. 1977.
15. Y. Faroudja, "Method and Apparatus for Separation of Chrominance and Luminance with Adaptive Comb Filtering in a Quadrature-Modulated Color Television System," U.S. Patent No. 4,179,705, Dec. 1979.
16. D. Teichner, "Adaptive Filter Techniques for Separation of Luminance and Chrominance in PAL TV Signals," *IEEE Trans. Consum. Electron.*, *CE-32*:241–250, Aug. 1986.
17. F. Kretz and J. Sabatier, "Échantillonnage des images de télévision: analyse dans le domaine spatio-temporel et dans le domaine de Fourier," *Ann. Télécommunications*, *36*:231–273, 1981.
18. U. Boes, "Linear Filtering in Image Sequences," in *Image Sequence Processing and Dynamic Scene Analysis*, T. S. Huang, ed., 437–447, Springer-Verlag: Berlin, 1983.
19. E. Dubois, "Multidimensional Spectra of Sampled NTSC Colour Signals with Application to Coding," in *Proc. Picture Coding Symposium*, 8.3.1–8.3.2, 1979.
20. E. Dubois, "The Sampling and Reconstruction of Time-Varying Imagery with Application in Video Systems," *Proc. IEEE*, *73*:502–522, Apr. 1985.
21. J.-Y. Ouellet and E. Dubois, "Sampling and Reconstruction of NTSC Video Signals at Twice the Color Subcarrier Frequency," *IEEE Trans. Commun.*, *COM-29*:1823–1832, Dec. 1981.
22. D. E. Dudgeon and R. M. Mersereau, *Multidimensional Digital Signal Processing*, Prentice-Hall: Englewood Cliffs, N.J., 1984.
23. Y. Faroudja, "Optimizing NTSC to *RGB* Performance," presented at the 127th SMPTE Technical Conference, Los Angeles, Oct. 1985.
24. A. Kaiser, "Chrominance-Luminance Separator," U.S. Patent No. 4,072,984, Feb. 1978.
25. J. Rossi, "Digital TV Comb Filter with Adaptive Features," in *Proc. IERE Conf. on Video and Data Recording*, 267–281, 1976.
26. Y. Faroudja, "Method and Apparatus for Separation of Chrominance and Luminance with Adaptive Comb Filtering in a Quadrature-Modulated Color Television System," U.S. Patent No. 4,240,105, Dec. 1980.