

Copyright © 1990, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

**OPTIMAL ARCHITECTURES FOR AN
INTEGRATED NTSC DECODER**

by

Ken A. Nishimura

Memorandum No. UCB/ERL M90/63

24 June 1990

COVER PAGE

99

**OPTIMAL ARCHITECTURES FOR AN
INTEGRATED NTSC DECODER**

by

Ken A. Nishimura

Memorandum No. UCB/ERL M90/63

24 June 1990

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

TITLE PAGE

3.7 Chrominance Demultiplexing	39
4.0 Existing Circuit Embodiments of NTSC Decoders	41
4.1 Overall Performance of Various Circuit Embodiments	41
4.2 Analog vs. Digital Architectures	43
4.3 Discrete and Mixed Circuit Embodiments	43
4.4 Integrated Circuit Embodiments	44
4.5 Comparisons with the Proposed Chip	46
5.0 Architecture and Circuit Implementation of the Proposed Chip	48
5.1 Overall Chip Architecture	48
5.2 DC Restoration and Input Buffer	49
5.2.1 CMOS Operational Amplifiers for Resistive Loads	52
5.2.2 Operation of the Clamping Circuit	53
5.3 Sync Processing Circuits	54
5.4 Phase Locked Loop and Clock Generators	55
5.4.1 Choice of Sampling Rate	55
5.4.2 Anti-Aliasing and Oversampling	57
5.4.3 Monolithic MOS Crystal Voltage Controlled Oscillator	58
5.5 Oversampling S/H Stage with Integral FIR Decimation Filter	62
5.5.1 First Stage Filter Design	62
5.5.2 Circuit Implementation of the First Stage Filter	63
5.5.3 Effects of Circuit Non-Idealities on Filter Performance	66
5.6 Second Stage Decimator and Filter	69
5.7 Luminance/Chrominance Separator	72
5.7.1 Development of an Analog RAM Delay Line	73
5.7.2 Write/Read Circuit for Analog RAM Cells	74
5.7.3 Implementation of the 2H Comb Filter	75
5.7.4 Effects of Circuit Non-Idealities	76
5.8 Chrominance Signal Demodulator	78

Table of Contents

1.0 Introduction	1
2.0 An Overview of the NTSC Signal	4
2.1 The Raster Scanning Process	4
2.1.1 Spectral Analysis of the Raster Scanned Image	8
2.2 Principles of Color Vision	9
2.2.1 Additive Color Theory	9
2.2.2 Color Responsivity of the Human Eye	11
2.3 Color Video Signals	12
2.3.1 Methods of Transmitting Color Information	12
2.3.2 The NTSC Color Encoding System	13
2.4 The PAL and SECAM Color Encoding Systems	16
3.0 The NTSC Decoding Process	19
3.1 Decoding of NTSC Signals	19
3.2 DC Restoration	19
3.3 Synchronization Signal Recovery and Processing	20
3.4 Filtering of the Composite Signal	22
3.5 Luminance Chrominance (Y/C) Separation	23
3.5.1 Bandpass Filtering	25
3.5.2 Comb Filters	26
3.5.2.1 Delay Elements for Video Signals	26
3.5.2.2 Simple (1H) Comb Filters	30
3.5.2.3 Three Line (2H) Comb Filters	32
3.5.2.4 Adaptive Comb Filters	33
3.6 Chrominance Signal Demodulation and Filtering	36

3.7 Chrominance Demultiplexing	39
4.0 Existing Circuit Embodiments of NTSC Decoders	41
4.1 Overall Performance of Various Circuit Embodiments	41
4.2 Analog vs. Digital Architectures	43
4.3 Discrete and Mixed Circuit Embodiments	43
4.4 Integrated Circuit Embodiments	44
4.5 Comparisons with the Proposed Chip	46
5.0 Architecture and Circuit Implementation of the Proposed Chip	48
5.1 Overall Chip Architecture	48
5.2 DC Restoration and Input Buffer	49
5.2.1 CMOS Operational Amplifiers for Resistive Loads	52
5.2.2 Operation of the Clamping Circuit	53
5.3 Sync Processing Circuits	54
5.4 Phase Locked Loop and Clock Generators	55
5.4.1 Choice of Sampling Rate	55
5.4.2 Anti-Aliasing and Oversampling	57
5.4.3 Monolithic MOS Crystal Voltage Controlled Oscillator	58
5.5 Oversampling S/H Stage with Integral FIR Decimation Filter	62
5.5.1 First Stage Filter Design	62
5.5.2 Circuit Implementation of the First Stage Filter	63
5.5.3 Effects of Circuit Non-Idealities on Filter Performance	66
5.6 Second Stage Decimator and Filter	69
5.7 Luminance/Chrominance Separator	72
5.7.1 Development of an Analog RAM Delay Line	73
5.7.2 Write/Read Circuit for Analog RAM Cells	74
5.7.3 Implementation of the 2H Comb Filter	75
5.7.4 Effects of Circuit Non-Idealities	76
5.8 Chrominance Signal Demodulator	78

5.8.1 Multiplication Stage	78
5.8.2 Sampling Phase Adjustment	79
5.8.3 Chrominance Signal Filtering	80
5.9 Projected Performance of the Chip	81
6.0 Conclusions	84
References	87

Optimal Architectures for an Integrated NTSC Decoder

ABSTRACT

The drive to integrate entire circuit systems on a single VLSI chip stems from the cost reductions and performance gains associated with integrated circuit technology. In this paper, a description of an optimal architecture for a NTSC decoder utilizing analog sampled data techniques is presented. This architecture makes use of recent developments in the area of multi-dimensional (comb) filtering used for luminance/chrominance (Y/C) separation of composite NTSC video signals. Issues of anti-alias filtering of sampled video signals are addressed through the use of oversampling and decimation. Moreover, circuit techniques which make use of the special characteristics of the video signal are discussed; topics related to integration of various circuit blocks into a system level circuit are addressed.

Specifically, this work presents two new analog sampled data techniques. The first is an sample and hold with a built-in decimation filter for oversampling applications. The second is the development of an analog RAM structure capable of random storage and access of an analog data sample. In addition, this paper introduces basic video and NTSC color encoding concepts with special emphasis on Y/C separation technologies. Care has been taken in the design to conform to the full NTSC standards — this will allow use of the chip in performance critical areas such as professional and broadcast arenas.

Acknowledgements

I would like to extend my sincere appreciation to Professor Paul R. Gray for his continuing support and guidance in my work. His insight into circuit techniques and broad knowledge in many areas have helped make this work possible. In addition, many thanks go to Professor Meyer and Professor Messerschmitt for their help in answering the countless questions I have posed to them in the course of this work.

Special thanks goes to Ampex Corporation and my friends there, especially Rylan Luke and Steve Wagner. It is there where I first became interested in video signal processing. Their willingness to answer at times very strange questions has contributed greatly to my understanding of video.

My fellow students at the EECS department at U.C. Berkeley have been a source of knowledge and companionship through my lengthy stay here. Special commendations go to my colleagues in the Analog Circuit Design group for their knowledge and willingness to entertain me on those long days. Kudos to Gregory Uehara and Cormac Conroy for their efforts in reviewing the manuscript.

Finally, a special note of appreciation goes to my parents and family for their constant support, patience, and encouragement throughout all these years.

This work was supported through a fellowship from the Fannie and John Hertz Foundation and through a National Science Foundation Research Grant.

Optimal Architectures for an Integrated NTSC Decoder

Chapter 1 Introduction

The rapid pace of development of VLSI circuits is making possible the implementation of circuit functions unheard of only a few years ago. Integration of circuits carries with it the advantage of smaller size, lower cost, reduced power consumption, higher reliability, and in general, a higher level of performance over non-integrated implementations of the circuit. Originally designed for high-performance computing applications, VLSI circuits have been introduced into almost all applications of electronic circuitry. Electronic equipment intended for household use, consumer electronics, is no exception. The tremendous growth in this field can be attributed in part to the introduction of VLSI circuits in order to perform the complex tasks associated with contemporary equipment, such as DSP and data conversion in compact disc players, video signal processing circuitry found in video cassette recorders and televisions, and the intelligent microcontrollers found in today's stereo equipment. The rapid growth in the consumer electronics industry has placed an upward pressure to improve quality while at the same time reducing the costs to implement the increasingly complex functions demanded by today's equipment.

One area of intense scrutiny in the consumer electronics field is that of video signal processing. The introduction of improved format video recorders for consumer use and the advent of HDTV has raised consumer awareness in the area of image quality and overall performance of consumer video equipment. Therefore, a need exists to integrate the functions found in video equipment to improve reliability and reduce costs at the same time.

A function common to most all consumer video equipment is that of processing a composite NTSC video signal to yield the luminance and two color difference channels used to reconstruct a color video image, often referred to as a NTSC decoder. Although NTSC decoding is not a new or innovative function and many discrete circuit designs exist — one exists in each color television set — of the few integrated circuit designs that have been made to perform this function to date, none have attempted to provide near broadcast quality performance coupled with advanced analog signal processing to reduce the overall size of the chip. Moreover, most integrated solutions

utilize a set of chips rather than integrating the function onto a single chip, often utilizing different technologies for each chip. Of the single chip solutions that have been reported, many have substantial flaws in performance that exclude their use in professional and higher-grade consumer equipment. As the demand for higher quality in consumer equipment grows, the gap between the traditional mediocre "consumer grade" and "broadcast grade" will continue to narrow.

1.1 Definition of the Work

This report describes the design methodology used to develop a new architecture to implement the NTSC decoding function in a single monolithic integrated circuit (chip). The proposed chip will take continuous time NTSC video as an input and will perform the necessary steps to yield the luminance and two color difference channels used to recreate the image on a CRT. Care will be taken in the design of the circuits and overall chip architecture to preserve the highest level of image quality. The design goal is such that the chip will produce signals of quality commensurate with that found in professional or studio equipment. Although the definition of "professional grade" is somewhat nebulous, it can be quantified in general terms, and usually implies the following: [1, 2]

- (1) The luminance signal bandwidth is substantially flat from DC to 4.2 Mhz. The passband ripple should be less than 1dB, preferably less than 0.5 dB.
- (2) The chrominance signal bandwidth should be on the order of 1 Mhz, with demodulating of the color subcarrier signal along the I/Q axes, as opposed to the Pr/Pb axes.
- (3) The non-linearity and step response as measured by the "K-Factor" is less than 2 percent.
- (4) The differential gain and differential phase is less than 2 percent and 2 degrees respectively.
- (5) The overall signal to noise ratio is greater than 54 dB (p-p vs. rms) in the luminance channel, measured at 50 IRE APL.
- (6) The luminance chrominance (Y/C) separation should employ some method of two dimensional (comb) filtering rather than a single dimensional (bandpass) filtering scheme. If at all possible, some method of adaptively controlling the Y/C separator should be employed to remove image artifacts that may arise through the separation process.

- (7) The field rate and line rate droop should be less than one percent.

1.2 Overall Design Methodology

The advent of scaled analog CMOS technologies ($L_d \approx 1\mu\text{m}$) makes possible the creation of analog signal processing circuits that operate at video rates. The underlying premise of this work is that a judicious use of analog signal processing techniques can result in the design of a NTSC decoder circuit with performance equivalent to or exceeding that of a DSP based circuit while consuming less than half the silicon area. An investigation made into the performance limiting areas of NTSC decoders singled out the luminance/chrominance (Y/C) separation stage as the area that caused the most degradation in image quality [3, 4]. Therefore, an attempt was made to include advanced processing techniques, usually performed by complex digital signal processors, within the chip for purposes of improving the Y/C separation process. An area often overlooked in sampled data systems is the need for anti-alias filtering. For video applications, this filter can be a source of considerable cost and complexity. As such, the issue of anti-alias filtering was addressed through the use of oversampling and decimation filtering on-chip.

1.3 Organization of the Report

This report is organized into six chapters. Chapter Two gives a brief overview of video concepts and the NTSC color television standard. It forms the basis for the remaining chapters as it covers the various aspects of the NTSC video signal. Chapter Three describes the technology associated with the design of a NTSC decoder. A typical NTSC decoder architecture is presented and various methods of implementation are discussed. Emphasis is placed on the luminance/chrominance separation technology and current work in that area. Chapter Four discusses prior work in the area of integrated NTSC decoder circuits. Comparisons of performance and silicon area are made between the proposed chip and prior work. Chapter Five is a detailed description outlining the various components of the proposed chip. Two new analog circuit techniques, the oversampling front-end with integral decimation filter and the analog RAM cell will be introduced as a means of providing functionality previously reserved for DSP techniques. Chapter Six, the last chapter, recapitulates the major points of the report and provides a feel for future work in the area.

Chapter 2

An Overview of the NTSC Signal

Prior to a detailed explanation of the decoder circuit architecture, it will prove helpful to discuss the properties of the video signal, and in particular the encoding scheme that is used to transmit color video signals in the United States. To this end, this chapter will attempt to explain the basic concept of video, and describe the salient points of the NTSC standard and the resulting video signal. Emphasis will be placed on topics that are of direct interest to the proposed IC. There are many excellent references on video, NTSC, and related topics for readers wishing a more thorough explanation [5 - 7].

2.1 The Raster Scanning Process

A video system is designed to transmit images, which are inherently two dimensional. If continually changing images are to be transmitted, the signal will consist of a time dimension in addition to the two spatial dimensions. As transmitting a three dimensional signal is rather complex, it is desirable to utilize a process which converts the two spatial dimensions of an image to a single dimension versus time.

For simplicity, consider a stationary, rectangular, monochrome image. Assign cartesian coordinates with the origin at the upper left corner of the image with h denoting the horizontal coordinate and v denoting the vertical coordinate (Fig 2.1). Associated with each point of the image is a quantity representing the image at the point. In the case of a monochrome image, this quantity in question is the instantaneous brightness or luminance at the point. Obtaining this quantity at each point of the image allows reconstruction of the image via an inverse process.

However, this method of converting an image requires an infinite amount of work, as there are an infinite number of distinct points within an image. Instead, a method of sampling is used which allows the image to be represented by a finite quantity of information. The method most often used in video is known as raster scanning. This is a mapping method where a locus is traced by moving a point at a constant rate in both horizontal and vertical directions. It is usual to make the rate of horizontal motion much greater than the vertical rate. If the horizontal rate is an integral multiple of the vertical rate, then a repeating scanning pattern results (Fig 2.2a). The locus begins at the upper left corner (A), and quickly moves to the right where it reaches the right

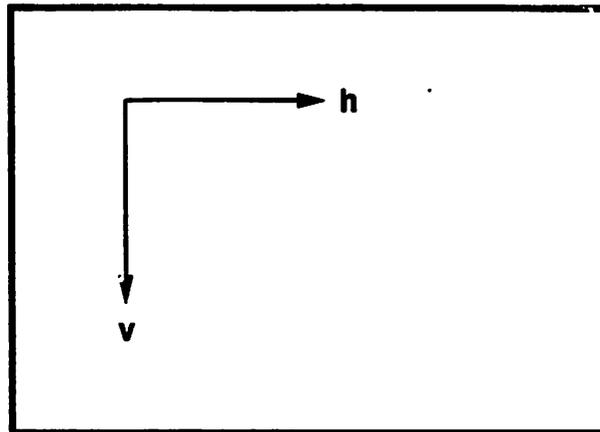


Figure 2.1: Coordinate System for Video Signals

edge of the image (B). During a period known as the horizontal blanking, an operation known as the horizontal retrace is performed, moving the locus back to the left edge (C). This process is repeated line by line until the lower right corner is reached (D). The image so scanned is called a frame. In the literature, the period of time it takes to traverse the image horizontally is denoted H , and the period of a frame is commonly denoted as V .

Video utilizes the same principle as motion pictures of rapidly transmitting a sequence of slowly changing images in order to create the illusion of motion. Thus, during an interval known as vertical blanking, the scanning locus moves back to (A) to begin processing the next image (frame). The number of scanning lines per frame and the rate of image repetition (frame rate) directly affect the quality of the transmitted image. The frame rate must be high enough to ensure that the illusion of motion is maintained. The minimum number of scan lines is limited by the required vertical resolution of the image as it is limited by the Nyquist theorem to half the number of scan lines per frame. However, the bandwidth necessary to transmit this data is proportional to the product of the frame rate and the number of scan lines. The video standard prevalent in the United States (System M) provides for 525 scan lines per frame with a frame rate 29.97 per second. This results in a line period, H , of 63.556 μs .

In reality, the frame rate specified, while high enough to provide an illusion of motion is insufficient to prevent objectionable flicker.[†] Thus, a technique known as interlaced scanning is used to double the effective frame rate without increasing the bandwidth necessary to transmit the data (Fig. 2.2b). In interlaced scanning, the frame is broken into two halves called fields, each containing half the number of scan lines by separating the odd and even numbered scan lines.

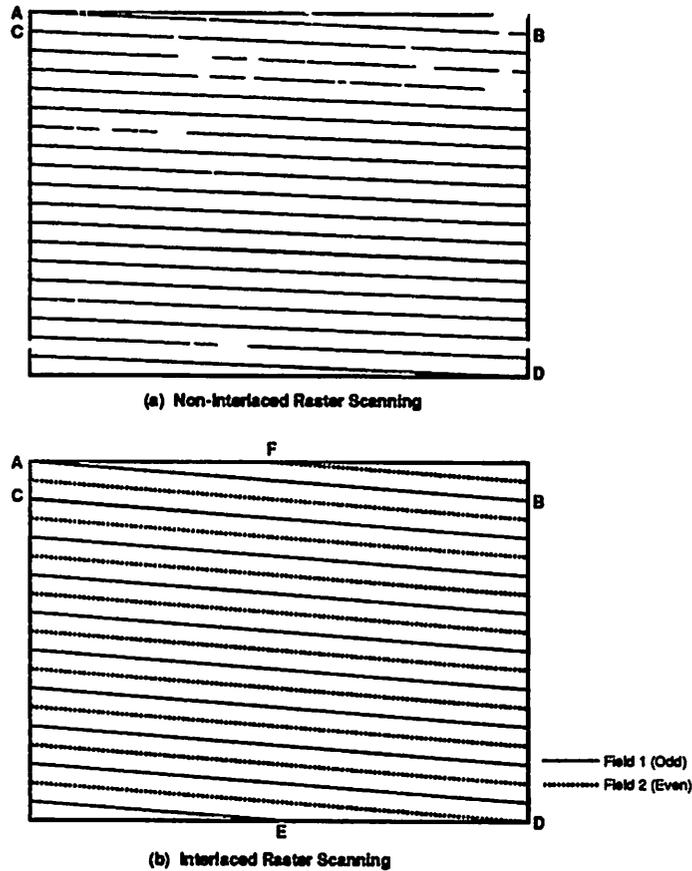


Figure 2.2: Raster Scanning, (a) Non-Interlaced, (b) Interlaced

Thus, the field containing lines 1, 3, 5, ... is known as the odd field and the field containing lines 2, 4, 6, ... is called the even field. The image scanning process begins at the upper left corner as before, but progresses vertically down the image at twice the previous rate. As the frame contains an odd number of lines (as in NTSC), the last line will be incomplete, ending at point E. The even field begins after the vertical blanking interval at point F, halfway across the image. The scanning process then ends at the lower right corner D. The process then repeats for the next image. Thus, the vertical period, V is halved for interlaced scanning.

In the case of a moving image, the sum of the odd and even fields is not exactly that which would be obtained by using a non-interlaced scan, as the image would move slightly between the odd and even fields. While this is not a concern with respect to this report, it is of great concern

† Although motion pictures are produced at 24 frames/sec, to avoid the same type of flicker, each image is shown twice in the theater, thus making the effective image update rate 48 frames/sec.

Chapter 2 An Overview of the NTSC Signal

when dealing with processes that convert between non-interlaced and interlaced images, such as may be encountered in image compression algorithms.

The scanned image is converted into an electrical signal by representing the image quantity as a voltage. In the case of monochrome images, the image quantity is simply the brightness. Dark areas result in a video signal of low amplitude, while light areas result in a large amplitude signal. The resulting signal is known as a video luminance signal, as it conveys the brightness of an image.

One requirement of a scanned image system such as this is that there must be perfect synchronization between the transmitting and receiving scan processes to prevent break up of the image. To facilitate this synchronization, signals collectively known as sync signals are inserted into the video signal which indicate the beginning of each scan line (horizontal sync) and the beginning of each field (vertical sync). The horizontal sync is inserted during the horizontal blanking period and instructs the electron beam to return to the left edge of the image. Similarly, the vertical sync is inserted during the vertical blanking period. During the blanking intervals, the electron beam in the receiver CRT is cut off, thus the sync signals themselves are not seen. The portion of the video signal that is actually seen, that is not blanked, is called active video. In the NTSC system, 11.1 μ s of the each line is reserved for the blanking interval, leaving about 52.4 μ s of active video per line. Likewise, 42 lines of each frame (21 lines per field), are used for the vertical blanking interval, leaving 483 lines (525-42) of active video information.

The horizontal sync consists simply of a negative going pulse at the start of each line period. The base to peak amplitude of the pulse is nominally 286 mV; particularly important is the fact that the peak of the sync called the sync tip, is the most negative voltage in the video signal. The vertical sync is contained within the vertical blanking interval and consists of a large series of pulses, most of which are not of direct consequence to the design of the proposed IC. However, owing to the long length of the vertical blanking interval, and the fact that information transmitted during this interval is not seen, this period is often used to transmit ancillary information ‡. As such, it is preferable to preserve the video signal during the vertical sync interval in signal processing circuits.

‡ Two common uses for the vertical sync interval are the VIR receiver calibration signal, and Closed Captioning (CC), which provides teletext style information to hearing impaired viewers.

2.1.1 Spectral Analysis of the Raster Scanned Image

It will be useful to perform a Fourier analysis of the video signal discussed so far, as it will give insight into the color video system in use today. Again, consider a stationary image for simplicity. If H and V denote the horizontal and vertical scan rates, performing a two-dimensional Fourier transform on the image results in [8],

$$x_{mn} = \frac{1}{HV} \int_0^H \int_0^V F(u,v) \exp \left[-j2\pi \left(\frac{mh}{H} + \frac{nv}{V} \right) \right] du dv$$

$$F(h,v) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x_{mn} \exp \left[j2\pi \left(\frac{mh}{H} + \frac{nv}{V} \right) \right]$$

where $F(h,v)$ is the image quantity of interest (intensity) and x_{mn} is the Fourier component at spatial frequency (m,n) . Thus, the video signal can be represented by

$$y(t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x_{mn} e^{j2\pi(mf_h + nf_v)t}$$

where f_h and f_v are the horizontal and vertical scanning rates respectively. A key property of this signal is that it is doubly periodic in f_h and f_v . The quantity x_{mn} is typically a monotonically decreasing function of m and n as most images contain less energy corresponding to high spatial frequencies.

In the case of a moving image, the periodicity between frames is lost, the result being the distinct lines at intervals of f_v blend together to form a continuous spectrum (Fig. 2.3b). The "width" of each clump of energy is dependent on the spatial frequency of the image. Images with high spatial frequencies in the vertical dimension (poor line to line correlation) will tend to spread out the clumps, while high frequencies in the horizontal dimension will tend to extend the series of clumps into higher frequencies, thus increasing the overall signal bandwidth.

The total bandwidth occupied by the video signal is determined by the horizontal resolution required in the image. A sinusoidal video signal will result in alternating light and dark zones. The number of these alternating zones that can be packed into one line is determined strictly by the number of cycles of the sinusoid that can be transmitted in one line period H , minus the horizontal blanking time. Additional corrections described below are made for the 4:3 aspect ratio used in video — the end result being that approximately 1 Mhz of bandwidth is required for each 80 lines of horizontal resolution. Thus, there exists a one to one relationship between perceived image quality and the bandwidth used to transmit the image. The vertical resolution is limited by the

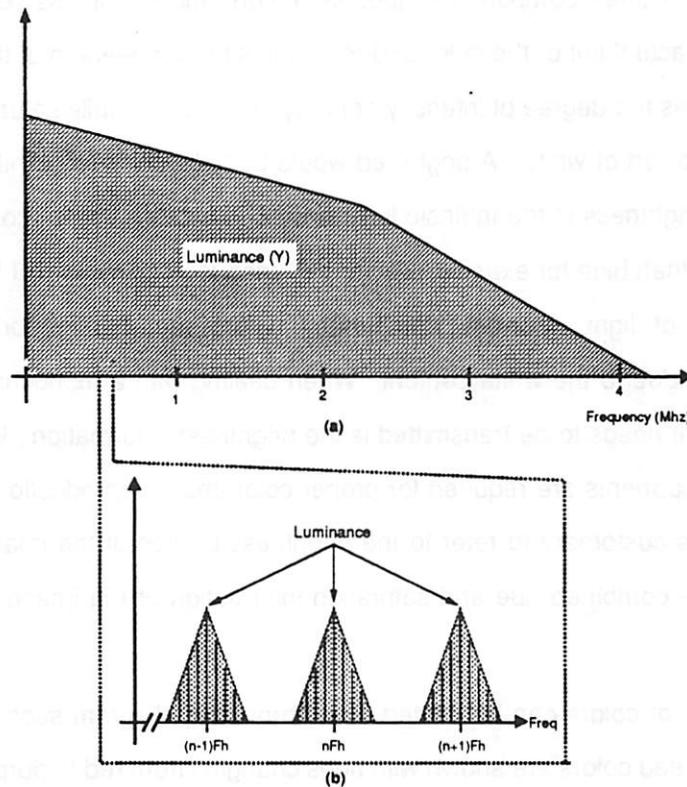


Figure 2.3: (a) Overall NTSC Luminance Spectral Density
(b) Magnified Luminance Spectral Density

number of scan lines present, less those used in vertical blanking, or 483 lines. In reality, interlacing reduces the perceived vertical resolution to about 70 percent of this value. Furthermore, it is desirable to maintain approximately the same horizontal and vertical resolution when transmitting an image. Thus, approximately 340 lines of horizontal resolution are needed. The NTSC luminance bandwidth of 4.2 Mhz reflects this requirement.

2.2 Principles of Color Vision

2.2.1 Additive Color Theory

Before generalizing this discussion to color video systems, a brief review of color theory is helpful. Video images unlike those on printed paper, consist of light generated by the screen, rather than being reflected from a light source. Thus, additive color theory is applicable as opposed to the more familiar subtractive color theory.

Optimal Architectures for an Integrated NTSC Decoder

A color consists of three components, hue, saturation, and brightness commonly abbreviated HSB. Hue is the actual tint of the color and represents the wavelength of the photons of the light. Saturation denotes the degree of intensity or purity of the tint. A fully saturated color can be desaturated by the addition of white. A bright red would be fully saturated, while pink is a desaturated red. Finally, brightness is the intrinsic luminosity of the color. Some colors, such as yellow are much brighter than blue for example, due to the unequal responsivity of the human eye to different wavelengths of light. Likewise, desaturated colors are usually brighter than their saturated counterparts due to the white content. When dealing with a monochrome system, the only image quantity that needs to be transmitted is the brightness information. In a color system, however, all three components are required for proper color image reproduction. When describing video systems, it is customary to refer to the brightness portion of the image as luminance, abbreviated Y, and the combined hue and saturation information of the image as chrominance, abbreviated C.

The visible range of colors can be plotted on a chromacity diagram such as that shown in Figure 2.4. Fully saturated colors are shown with hues changing from red to purple while traveling counterclockwise on the outer edge. Progressively desaturated colors are represented by the inside of the diagram with pure white near the center, as white is the sum of all colors. Analogous to the familiar subtractive color theory, new colors can be synthesized by forming linear combinations of given colors. The colors obtained by adding two colors are defined by the line connecting the two colors on the diagram. If three colors are used, then the colors produced by linear combinations of these three colors is the region enclosed by the triangle whose vertices represent the original three colors. The set of all possible colors obtained from a given trio of colors is known as the color gamut. Note that a judicious choice of these three colors results in a gamut that encompasses nearly all of the visible colors. This trio of colors is known as an additive primary set. The choice of a primary set for color video involves trying to encompass as large an area as possible in the chromacity diagram, at the same time selecting colors that are easily generated by phosphors on a CRT. The generally accepted standard for video are red, green, and blue, commonly abbreviated RGB.

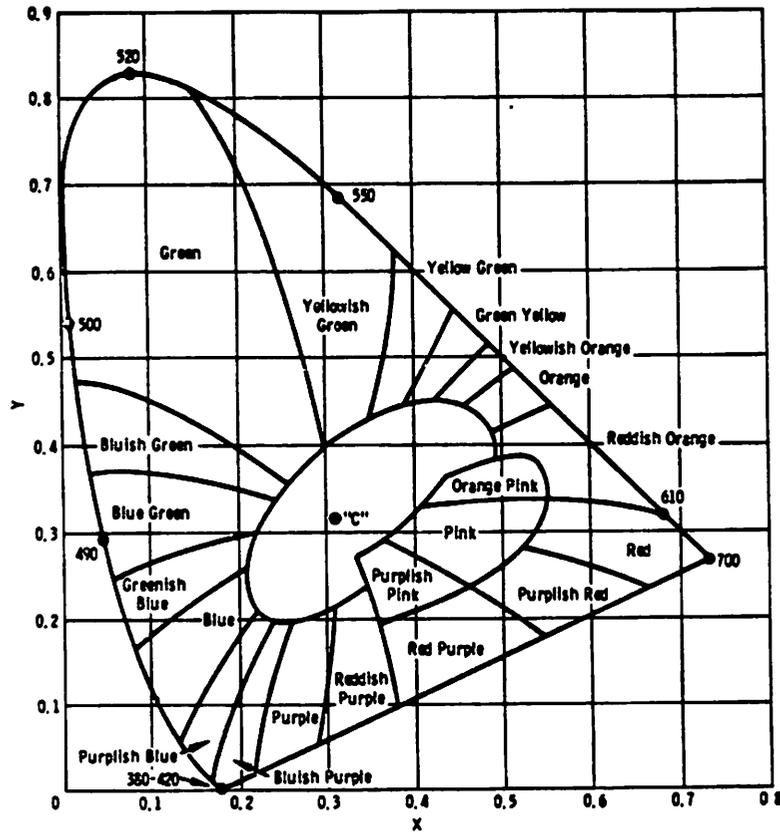


Figure 2.4: CIE diagram with color perceptions of the "standard" eye.
 (Numbers correspond to wavelength in nm.) (From [2])

2.2.2 Color Responsivity of the Human Eye

While it is important to provide a precise of a transmit/receive system as possible, equally important is matching the system's characteristics to the final receiver of the image, the human eye [9]. The eye utilizes two separate light sensitive structures in order to render images. The more plentiful, rods, are sensitive only to the presence or absence of photons, thus, they contribute to monochrome vision. Cones, on the other hand, are sensitive to the wavelength of the photon, thereby contributing to color vision. However, there are far fewer cones than rods, resulting in a significantly lower spatial resolution for the color portion of an image. Sharp edges and other high spatial frequency components of a color image are distinguished by changes in intensity or brightness rather than changes in color. Thus, a color transition with a relatively long transition time will appear to be sharp if it is accompanied with a sharp edge in luminosity. Therefore, a deficiency in color spatial resolution can be masked by preserving a good spatial frequency response in the luminance component of the image. This property of human vision allows a

reduction in the amount of chrominance information necessary for image reproduction.

Moreover, the resolution of the human eye is not constant for all colors. The eye has the highest chroma resolution in the orange and cyan hues, while it has a fairly poor color resolution in the magenta and green hues. This information is used to further reduce the bandwidth necessary to transmit the chrominance information.

2.3 Color Video Signals

2.3.1 Methods of Transmitting Color Information

The most straightforward method of transmitting a color image would be to record the red, green, and blue components of each image point in lieu of the intensity or brightness. This is known as RGB transmission and is widely used in computer displays. However, this system requires three times the bandwidth of a monochrome system. In addition, the resulting signal would be incompatible with the previously set standard for monochrome video. During the introduction of color television broadcasts, the FCC mandated that any color video standard must be compatible with the existing monochrome standard. That is, a pre-existing monochrome receiver must be able to reconstruct a satisfactory monochrome image from the color signal. This requirement ruled out RGB transmission, and also indirectly required that the color signal occupy roughly the same bandwidth as the monochrome signal, as broadcast channel allocations had already been made utilizing the 6 Mhz monochrome standard.

In order to maintain compatibility with the monochrome standard, the luminance information of the image must be transmitted in roughly the same method independent of the presence of chrominance information. Thus, a method of relating the luminance of a color with the RGB values is used. As white light is a particular linear combination of red, green and blue light, the luminance of a color can be determined by observing the amount of red, green, and blue light contained within that color. In the case of color video, the following equation is used to define luminance as a function of the RGB values.

$$Y=0.30R+0.59G+0.11B$$

Furthermore, color theory follows the rules of a linear space. Thus, any color that can be represented by the original primary set can be represented by another primary set, whose colors are independent linear combinations of the original primary set (RGB). (Mathematicians will

equate primary sets to a set of basis vectors. The operation just described can be viewed as a basis vector transformation.) Therefore, a new color space is created, translating RGB space into a space defined by a vector representing luminance (Y) and two independent linear differences of Y and the RGB, (R-Y) and (B-Y), commonly abbreviated *Pr* and *Pb*. Red and blue were chosen, being that the largest component of Y is G, making G-Y smaller in magnitude compared to *Pr* and *Pb*, resulting in a higher susceptibility to noise. The transformation, being linear, can be represented as a matrix multiplication:

$$\begin{bmatrix} Y \\ Pr \\ Pb \end{bmatrix} = \begin{bmatrix} 0.30 & 0.59 & 0.11 \\ 0.70 & -0.59 & -0.11 \\ -0.30 & -0.59 & 0.89 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The Y channel is therefore equivalent to the signal that would be obtained by scanning the image via a monochrome system. Thus, a method of encoding the *Pr* and *Pb* channels while providing minimal interference to the Y channel is needed. As stated earlier, high resolution is required for the Y channel to preserve image quality necessitating the full 4.2 Mhz bandwidth previously allocated to monochrome signals. However, the *Pr* and *Pb* channels contain information used strictly for color reproduction, and as a result, a much narrower bandwidth corresponding to the lower resolution required is allowable. Experimentation has shown that a bandwidth of approximately 1 Mhz is sufficient for color reproduction.

2.3.2 The NTSC Color Encoding System

A method which allows simultaneous transmission of the original monochrome channel and the *Pr* and *Pb* channels in the same bandwidth was developed by the National Television Standards Committee (NTSC) in 1953 [10]. The salient feature of the system is the frequency division multiplexing of a subcarrier that has been quadrature amplitude modulated by the *Pr* *Pb* channels. Recall that for most images, the power density spectrum of the existing signal is discontinuous, with energy concentrated in clumps centered at multiples of the line frequency H. Thus, if a method that inserts the color information in between the clumps is used, the color information can be transmitted without disturbing the existing luminance information. As the color information is obtained by the same scanning process as the luminance information, the power density spectrum of the *Pr* and *Pb* components will also be discontinuous with clumps centered at multiples of H. Amplitude modulation of the signal with a subcarrier will allow shifting of the spectrum by an amount equal to the subcarrier frequency. Therefore, if the subcarrier frequency is chosen to

reside at $\frac{1}{2}(2n+1)H$, the resulting spectrum of the P_r and P_b signals will interleave with the existing luminance spectrum (Figure 2.5).

Spectrum in vicinity of Color Subcarrier

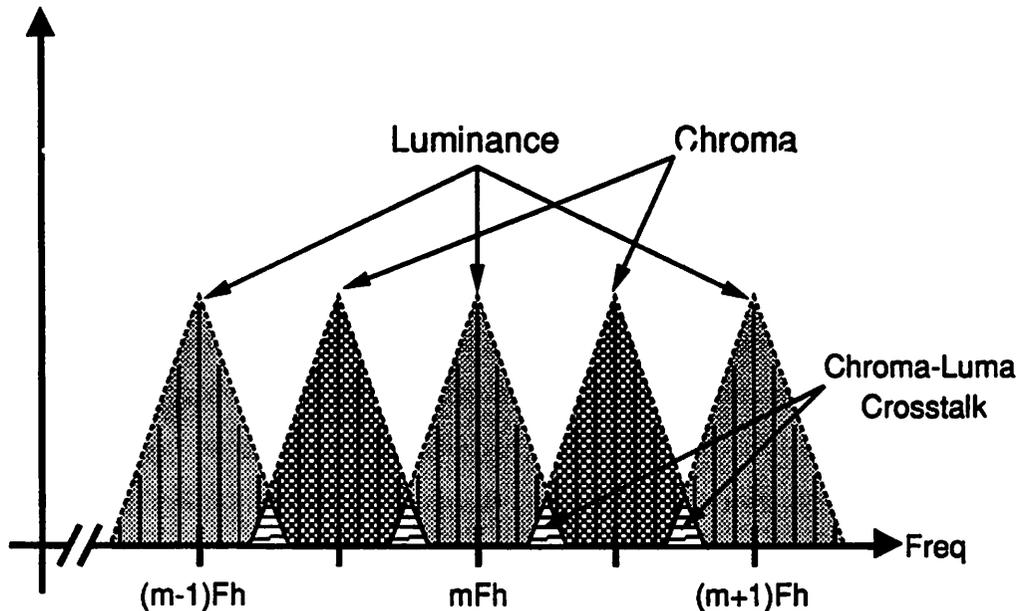


Figure 2.5: Frequency interleaving of Luminance and Chrominance Signals

By using QAM, two independent signals, P_r and P_b , can be transmitted in the same frequency spectrum. The NTSC standard sets the subcarrier frequency, f_{sc} at $455/2 H$ or 3.57954525 Mhz. Prior to modulation, the P_r and P_b signals are bandlimited to about 1 Mhz. The resulting QAM chrominance (C) signal contains all the information necessary to reproduce the color portion of the image.

A small problem with the signal described above is that the highest frequency present would be approximately 4.58 Mhz. This is higher than allowed by the existing monochrome standard and is especially troublesome for broadcast, as the audio channel is transmitted via FM on a subcarrier at 4.5 Mhz. Thus, a slight modification is made using the information known about human color vision. By changing the color space from $YPrPb$ to another space known as YIQ , it was found that one of the color components would represent colors that require relatively high resolution while the other component would represent colors that require only minimal resolution. These components are known as I and Q respectively for "In-Phase" and "Quadrature". The Q signal is

bandlimited to 600 KHz prior to modulation and thus does not cause a violation of the 4.2 Mhz limit. The I signal, however, is allowed a bandwidth of 1.3 Mhz; to prevent interference to the audio carrier, the modulated signal is low pass filtered at 4.2 Mhz resulting in a vestigial upper sideband. The unequal response about the subcarrier results in crosstalk between the I and Q channels which requires a more complex filter in the receiver. Thus, the complete frequency map of the NTSC signal is shown in Figure 2.6.

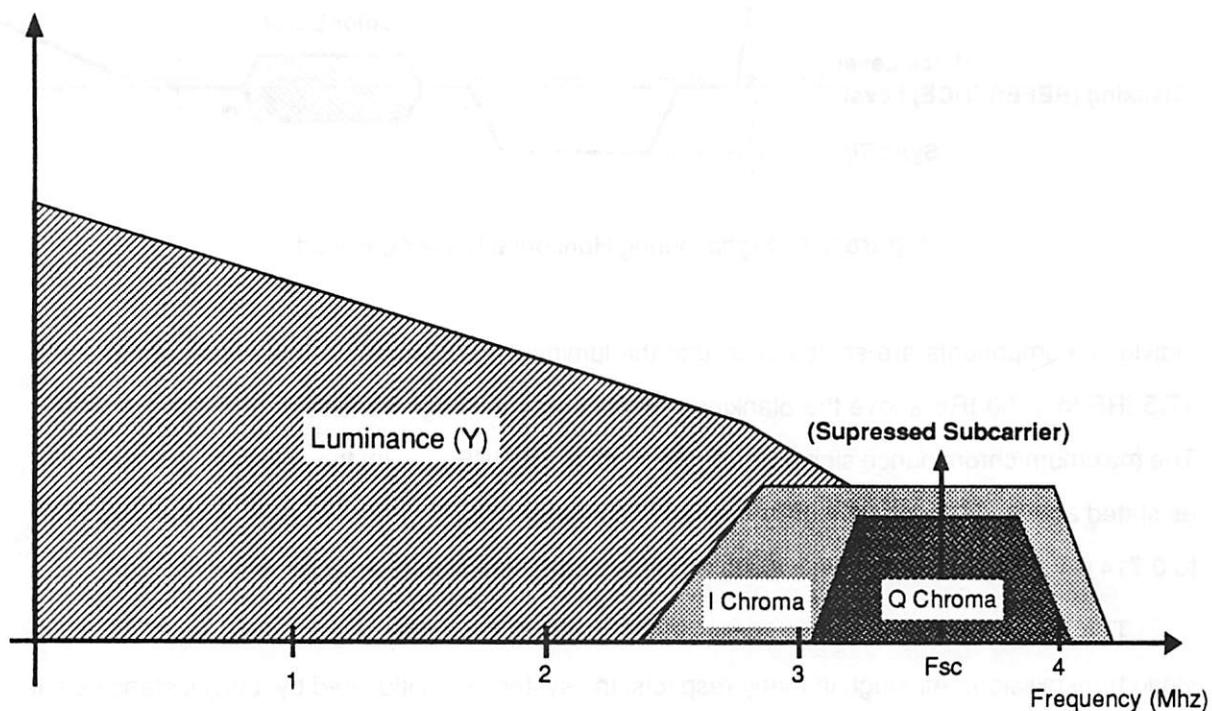


Figure 2.6: Overall NTSC Frequency Spectrum

Successful demodulation of the QAM chrominance signal requires that the receiver have a local oscillator that is phase locked to the oscillator used for modulation. To facilitate this, a signal known as the colorburst is added to the video signal, once per line right after the sync signal. This burst signal consists of eight or nine cycles of the subcarrier at a phase corresponding to a negative Pb signal, with an amplitude of 286 mV_{p-p}. This signal then serves as a phase and amplitude reference for the local demodulating oscillator. Thus, the video signal during horizontal blanking consists of the horizontal sync and colorburst. A diagram of the horizontal blanking interval is shown in Figure 2.7.

The final step involves adding together the luminance and chrominance signals along with the synchronization signals to form the final, composite NTSC video signal. The amplitudes of the

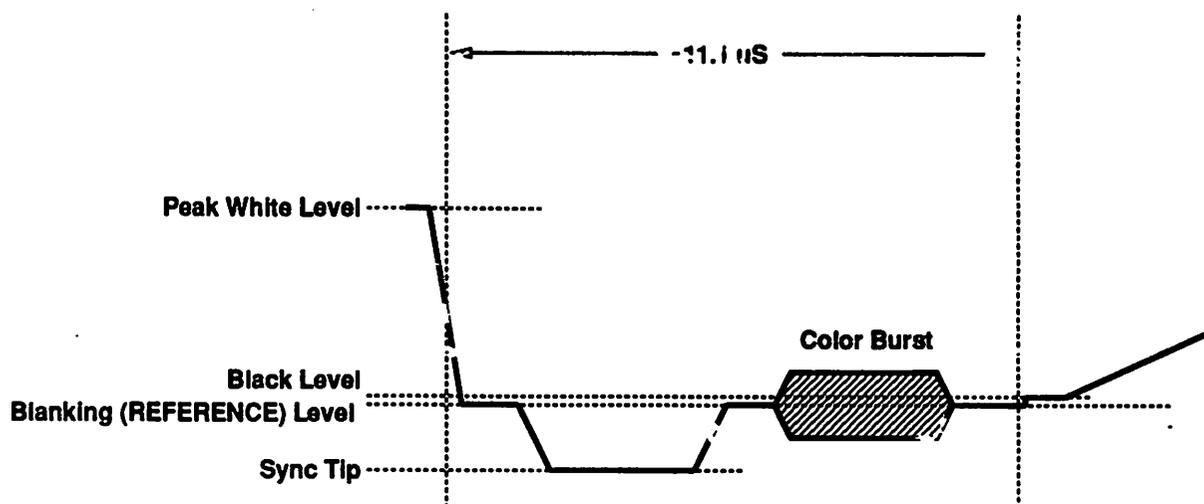


Figure 2.7: Signal during Horizontal Blanking Period

individual components are scaled such that the luminance component of the signal swings from +7.5 IRE to +100 IRE above the blanking reference as the image transitions from black to white. The maximum chrominance signal amplitude is set to 126 IRE_{p-p} , with the colorburst set at 40 IRE as stated above. (The IRE is a unit of voltage commonly used in video literature, and corresponds to 0.714 mV.) Figure 2.8 shows a video waveform representation of an albeit artificial image.

The signal described above corresponds to the NTSC video signal used today for color video transmission. Although in many respects, the system isn't antiquated by today's standards, it is in use by over 500 million television receivers worldwide. The performance of the system is being continually improved through more advanced signal processing techniques. The next chapter will describe the topic of the proposed chip, the NTSC decoding process — converting the NTSC signal back into the original R, G, and B signals that describe the image.

2.4 The PAL and SECAM Color Encoding Systems

The NTSC system is used in the United States, Canada, Mexico, Japan, and a few other countries. In Europe, two other systems, PAL (Phase Alternation by Line), and SECAM (Séquential Couleur à Mémoire) are used. Both these systems are similar to NTSC as they follow the principle of a luminance component and two derived color components. However, the particular method used to encode the chrominance information differs among the three systems.

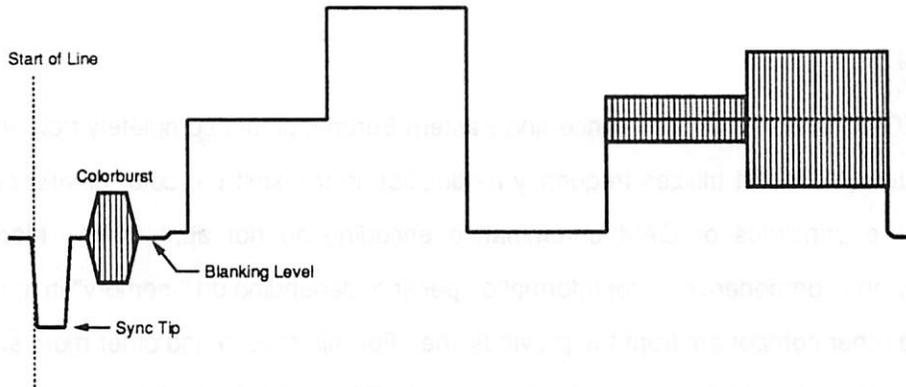
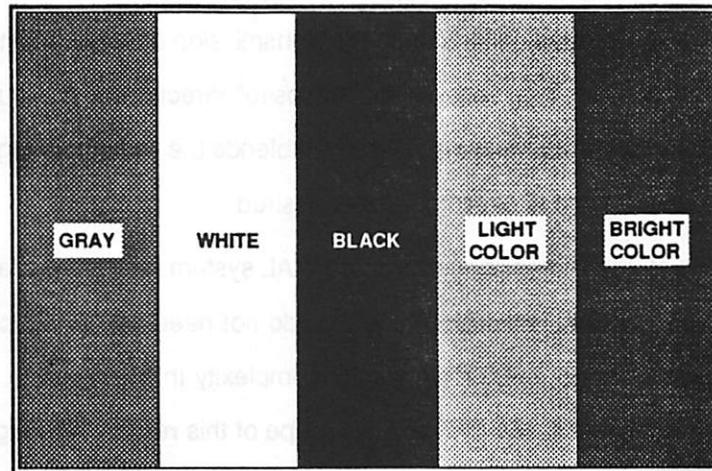


Figure 2.8: Video waveform resulting from image above.

2.4.1 PAL

The PAL system is used almost exclusively with the 625 line 50 field/sec systems that are prevalent in Europe. Like NTSC, it makes use of QAM to encode the two color signals. Unlike NTSC, modulation is performed on equal bandwidth Pr and Pb signals (renamed U and V in PAL) omitting the I/Q modification used in NTSC. Furthermore, as the 625/50 system has a higher bandwidth than the 525/60 systems, the subcarrier is moved higher in frequency to 4.43361875 Mhz, a frequency which allows frequency multiplexing of the chroma signals as in NTSC.

However, the biggest difference between PAL and NTSC lies in the line by line alternation of the reference phase used to modulate the Pb (V) signal [11]. The motivation for this is that transmission of a QAM signal through a channel with phase distortion normally results in errors in the recovered baseband signals. Such errors manifest themselves as unwanted hue changes in

the picture. By alternating the reference phase every line in the V component, and compensating for this in the receiver, a phase distortion in the transmission channel affects the reproduced color in one direction on a given line, and in the "opposite" direction on the subsequent line. As the lines are very close together in the image, the eye blends the two colors together and perceives the "average" of the colors, which is exactly what is desired.

This method of alternating lines makes the PAL system more robust and controls hue disturbances to the point that PAL television receivers do not need hue or tint controls. However, this improvement is made at the cost of significant complexity in the receiver. The specific changes necessary to demodulate PAL are beyond the scope of this report. Although the proposed chip is designed for NTSC operation, with the proper modifications, it is feasible to provide PAL operation, albeit at the cost of added silicon area and power for the more complex decoding circuits.

2.4.2 SECAM

The SECAM system used in France and Eastern Europe, differs completely from the NTSC and PAL systems in that it utilizes frequency modulation to transmit the color difference signals [12]. Thus, the principles of QAM chrominance encoding do not apply here. Moreover, it transmits only one component of color information per line, depending on "memory" in the receiver to provide the other component from the previous line. For this reason and other more subtle factors, the SECAM system is considered to be inferior to PAL and NTSC. As a result, no work is planned to modify the proposed chip to operate with the SECAM standard.

Chapter 3

The NTSC Decoding Process

3.1 Decoding of NTSC Signals

The NTSC signal described in the previous chapter contains all the information necessary to reproduce the original color image within the limits of the system. In order to accomplish this task, the video signal must be processed in order to recover the R, G, and B values that comprise the image, and to make available the sync signals so that the electron beam in the CRT can be synchronized to the scanning pattern originally used to transmit the image.

This chapter will focus mainly on the existing techniques and available technology to perform this task. Methods that are being considered for use in the proposed IC will be explained in full detail in Chapter 4. The task of decoding NTSC signals can be broken into several major sub-divisions; these are listed below, and explained in the subsequent sections.

- (1) DC Restoration
- (2) Sync Stripping and Timing Recovery
- (3) Filtering of the Composite Signal
- (4) Luminance Chrominance (Y/C) Separation
- (5) Chrominance Signal Demodulation and Filtering
- (6) Chrominance Signal Dematrixing

3.2 DC Restoration

The signal path between the circuit generating the NTSC signal and that which is decoding the signal is arbitrary and is not guaranteed to have benign characteristics. Moreover, in many cases, there is not a continuous electrical path between the two circuits, as in the case of a magnetic recording device (e.g. video tape recorder), where the signal could have been encoded years before it is subsequently decoded upon playback. Furthermore in many cases, transmitting the encoded video signal is much easier if the signal path is AC coupled. Therefore, by convention, video signals are intended to be AC coupled from one signal processing unit to another. This

practice has been carried to the point that some pieces of equipment will place a DC bias on the video signal for various reasons; thus, it is imperative that the input to any video signal processing circuit be capacitively coupled to prevent any spurious DC offsets from upsetting the circuitry. However, by capacitively coupling the circuit, the DC reference level is lost. Hence, it is necessary to somehow restore the video signal to a known reference by clamping a known value of the video signal to some reference voltage within the circuit. This task is known as DC restoration.

During a small time interval immediately after the horizontal sync signal, the video signal is at a known level, that corresponding to the blanking level. As it always is present and since its value corresponds to a brightness just "blacker than black," it is commonly used as a reference in video signals. Thus, the DC restore circuit attempts to clamp the blanking level to a reference voltage, usually ground. Most restoration circuits operate by finding and locking to the sync signal, and waiting a preset period of time until the blanking interval is present, approximately 3 μ s after the rising edge of sync. Although, the colorburst signal is present during this interval, a low pass filter is used to remove the burst signal since it is at a rather high frequency. The resultant filtered signal is applied to a feedback loop during the blanking interval to adjust the bias on the input coupling capacitor until the blanking signal is at ground. At the end of the blanking interval, the bias on the capacitor is held until the next line period.

The sync detection circuitry for this function does not need to perform to very high standards, as it is not used to determine system timing. Therefore, it is common to use a peak detection circuit which looks for a negative pulse every line period. On the other hand, a critical section of this circuit is the portion which holds the value of bias on the coupling capacitor until the next line period. The clamping value needs to be held to within 5 mV of the proper value during a line period. Larger drifts lead to a condition known as tilt, where the brightness of the picture changes from one edge of the picture to the other due to the changing DC reference level. Therefore, care must be taken in the design of the sample and hold, with emphasis on droop and sample to hold offset.

3.3 Synchronization Signal Recovery and Processing

As was stated in Chapter 2, the raster scanning process depends heavily on the ability of the receiver to identically reproduce the raster pattern used to create the original signal. To do this, synchronization signals have been added to the video signal at the start of each line and the start of each field. Related to this, and just as important is the issue of finding a reference for system

timing. As the most critical timing issue in NTSC is the frequency of subcarrier, it is used as the timing reference. The length of H and V and all other related timing issues are generated as a fraction or multiple of the subcarrier period. Therefore, the sync recovery and processing, commonly referred to as sync stripping, is vital to proper operation of the circuit.

Precise recovery of the horizontal sync requires that the point corresponding to the 50 percent point on the falling edge of the sync pulse be found. The most common method of accomplishing this involves the use of a peak detector and a comparator. As the sync tip is guaranteed to be the most negative voltage in the video signal, use of a peak detector to find and hold the most negative voltage will result in a signal at the sync tip voltage. By averaging this voltage with that of blanking, a voltage equal to the 50 percent point of sync can be generated. This voltage is then used as a comparator trip point to locate the midpoint of the falling edge of sync.

The method described above, known as a fine sync strip, requires that precautions be taken to prevent high amplitude chroma, whose peaks can fall below the 50 percent value of sync, from triggering the comparator. The simplest way of doing this is to lowpass filter the signal fed to the comparator to attenuate the chroma signal amplitude. A slight compensation must be made for the group delay of the lowpass filter in order to insure that the sync signal derived from the comparator accurately represents the sync signal of the video signal. As the coarse sync strip used for this function is so similar to that for the DC restoration, it is very common to combine these two circuits to reduce complexity and cost.

Once the sync edge has been found, the colorburst can be located by waiting an appropriate amount of time, $5.0 \mu\text{s}$. Locking of a local oscillator to this colorburst can be done in many ways, with phase lock techniques becoming more popular as costs of doing so keep falling. The actual oscillator structure in this case is nearly always uses a quartz crystal for stability. However, methods such as injection locking of a high Q resonant tank are still used. The frequency of colorburst is specified to be very close to the nominal frequency, typically within 25 Hz. From any one given video source, however, the drift over time is much less than that, typically on the order of one hertz. As such, a very stable local oscillator that accurately reflects the chroma modulation oscillator can be made to function even though a comparison to a reference is made only for a few microseconds every line period.

The recovery of both subcarrier and H signals from the sync can pose a system level problem. In theory, a video signal will have a perfect relationship between H and f_{sc} . Thus, generating

H from f_{sc} should result in the same signal as that derived from the sync signal. However, many types of equipment, especially those used in consumer electronics, do not conform to the RS-170A timing specification which binds together H and f_{sc} . Thus, generating one signal from the other may result in a conflict. As subcarrier recovery is more critical, in many cases, the H period is determined by dividing down the recovered f_{sc} , using the recovered H only to locate the beginning of the line.

3.4 Filtering of the Composite Signal

In theory, the video signal should be bandlimited to 4.2 Mhz at the source. However, this is not always the case, and along with noise during transmission, some signal energy exists above the standard bandwidth of 4.2 Mhz that needs to be filtered prior to subsequent processing. This is even more critical if sampled data techniques are utilized in the signal path. It is typical to specify a low pass cutoff of 5 Mhz, to prevent bandedge response degradation when multiple passes are made through the filter. The lowpass filter used for this function must have the following characteristics: (1) Stable, flat passband, with a minimum of ripple, preferably less than 0.3 dB, (2) Smooth transition band with an adequate attenuation above a critical frequency, usually the nyquist rate for sampled data architectures, and (3) Linear phase, or constant group delay within the passband. Of these three requirements, the last, constant group delay is probably the hardest to achieve. Unlike audio signals, preservation of group delay over frequency is an important characteristic in video. While the ear is relatively insensitive to phase distortions, such distortions in image signals manifest themselves as visible ringing, especially after an abrupt step in the image.

Symmetric Finite Impulse Response (FIR) filters are often touted for the linear phase characteristics. This fact is true, and is heavily exploited in the design of the proposed chip. However, FIR filters exist only in the sampled data domain — prior to sampling, the signal must be bandlimited to prevent aliasing distortion. In the continuous time domain, filters with complex poles at 30 degree angles from the real axis in the s-plane (Bessel Response), are often used as approximations to a linear phase response. Such filters can be made to have excellent group delay characteristics; the penalty being very poor transition band characteristics. As the poles are relatively near the real axis, the Q of the filter is low, thus requiring a high order filter for a given response [13]. Other filter topologies, such as Butterworth and Chebychev give better transition band response, but with correspondingly worse group delay characteristics. Causer or elliptic

filters give the sharpest rolloff through the use of transmission zeros. Not surprisingly, they also have the worst delay distortion. However, the increased transition band performance allows the use of much fewer filter sections for a given response. The savings in complexity can be used to construct a set of allpass filters to correct the delay distortion. Simulations show that a fifth order Cauer filter followed by two second order allpass sections results in a magnitude and phase response adequate for NTSC video.

The primary disadvantage of such a filtering scheme is the cost and complexity of such a filter, which usually employs several hand adjusted coils and precision capacitors. Such a filter also does not allow integration onto a chip. Recently, two alternative methods have been proposed. The first, is the incorporation of a continuous time filter on chip; elliptic filters for video applications have been demonstrated [14]. The delay distortion due to the filter however, must be corrected using DSP techniques after A/D conversion. Thus, the use of such a filter for signal paths not employing A/D converters is dubious. Moreover, these continuous time filters tend to inject spurious tones into the signal path, thus somewhat degrading the overall performance of the system. The other, involves oversampling the video signal, and performing the critical filtering using a FIR type of filter. Although a continuous time anti-aliasing filter is still required, the requirements of such a filter will be relaxed considerably. The primary disadvantage of this method is the need for a very high speed data converter if the FIR filtering is done in the digital domain. Moreover, such filters tend to be very large in silicon area and consume a considerable amount of power.

The proposed integrated circuit will take the latter approach and implement an oversampling front end with sampled data processing to perform decimation and filtering. In order to mitigate the need for a high speed data converter, the chip is designed to perform these tasks in the analog domain. The structure used will be described in full detail in chapter five.

3.5 Luminance Chrominance (Y/C) Separation

Subsequent to filtering, the NTSC decoder must first separate the luminance and chrominance portions of the composite signal before further processing. Although this seems innocuous at first, it is perhaps the most difficult task to perform in the entire decoder. Many of the annoying picture artifacts and "defects" of the NTSC system can be attributed to poor luminance chrominance separation techniques [15 - 18]. Thus, this is an area of intense scrutiny, where large gains in picture quality can be made.

Two picture artifacts are especially annoying and are the direct result of poor separation of the luminance and chrominance signals. These are known as cross color and cross luminance (luma). While there are many specific reasons for insufficient separation, the end results are the same. Cross color is the misinterpretation of luminance as chrominance information. This occurs when there is a large amount of luminance information near the subcarrier frequency, along with a low line to line correlation. This occurs with images that have fine diagonal stripes for example. The chrominance demodulator tries to decode a color from the misinterpreted luminance information. This results in areas of the image which contain this high frequency information having random colors superimposed on them.*

The other artifact, cross luma, is the opposite of cross color and results when chrominance information is misinterpreted as luminance information. As the amplitude of the luminance signal is translated directly into brightness, the chrominance signal, which is a high frequency sinusoid will translate into zones of alternating light and dark areas or dots. Since the frequencies of the subcarrier and the line period are related, under many conditions, the pattern of dots will appear to be fixed, thus making the artifact more noticeable. Different methods of luminance chrominance separation result in varying degrees of these artifacts, with the more complex separators generally providing a higher level of performance.

The Luminance Chrominance (Y/C) Separator processes the composite NTSC signal to yield a baseband luminance signal and the quadrature amplitude modulated chrominance signal. Due to the choice of subcarrier frequency and the QAM modulation process, the Y and C signals are interleaved in the frequency domain. Moreover, due to the nature of most images, the bulk of the luminance energy is concentrated in the lower frequencies (below 2 Mhz), while the chrominance signal energy is concentrated near the subcarrier (≈ 3.58 Mhz). The task of separating the luminance and chrominance can be approached in several ways, a few of which are outlined below.

* Perhaps the best example of this phenomena is Johnny Carson and his pinstriped shirts. Most television sets will show some degree of color fringing on the lapel of the shirt.

3.5.1 Bandpass Filtering

The simplest and until recently the most widely used method of Y/C separation is the bandpass chroma filter. This method depends solely on the fact that the energy of the chrominance signal occupies a fairly narrow band (approx. 700 KHz) centered about the subcarrier frequency. Outside of this band, the spectral density falls off rapidly, as there are few high frequency chrominance signals in most images. Moreover, the energy density of the luminance signal in the 3.5 Mhz range is relatively low compared to the chrominance signal as those signals would correspond to the fine textures in the images which are usually of small amplitude. Thus, by applying a bandpass filter with a bandwidth of about ± 600 KHz centered about the subcarrier frequency, the resulting signal will consist mainly of the chrominance signal. The luminance signal can be generated by delaying the composite signal to compensate for the group delay of the bandpass filter and then subtracting the chrominance signal.

While this method is simple and inexpensive, especially in continuous time systems, it suffers from many drawbacks:

(1) Within the passband of the filter, the system makes no differentiation between the chrominance and luminance components. Thus, high frequency luminance components will be interpreted as chrominance. This is the origin of "cross color" and will result in a shimmering rainbow effect surrounding fine patterns in the luminance. Moreover, as the luminance signal is generated by subtracting the output of the filter from the composite signal, these high frequency luminance components will be subtracted from the composite signal, resulting in a poor response of the luminance channel near f_{sc} . The result is that the luminance horizontal resolution is limited to approximately 240 TV lines.

(2) While most images contain little energy in the chrominance component outside of the filter passband, sharp horizontal chroma transitions will cause a significant amount of chrominance energy to fall outside of the filter passband. Thus, these components will be interpreted as luminance signals. As this type of transition usually occurs with highly saturated colors, the amplitude of the chroma signal that is misinterpreted will tend to be quite large. Demodulation of the subcarrier as a luminance signal will result in the "cross luma" phenomenon. This manifests itself visually as a line of dots, commonly known as "hanging dots", corresponding to a luminance pattern at the subcarrier frequency. These "dots" are especially visible since the frequency of the subcarrier under these conditions is generally

not one half of an odd multiple of the line frequency.

(3) This type of Y/C separator is most often implemented using discrete components in the continuous time domain. As such, the group delay characteristics of the bandpass filter are often not well controlled and frequently is a strong function of the frequency. Similarly, the delay equalizer, often implemented as a LC delay element, suffers from similar drawbacks. The result is that the luminance and chrominance signals often get out of step with each other, resulting in images where the colors are offset from the monochrome image.

3.5.2 Comb Filters

In order to overcome the drawbacks associated with a simple bandpass filter, structures known as comb filters have been employed in various forms. Comb filters perform the Y/C separation by making use of the frequency interleaving between the signals rather than their gross placement in the spectrum. These filters are a specialized form of FIR filters that have multiple zeros at evenly spaced intervals in frequency [19]. The resulting frequency response looks like the teeth of a comb, hence the name.

FIR filters make use of delay elements to perform their task; comb filters make use of long delays to obtain the multiple zeros in the frequency response. The quality of the delay elements is crucial to the overall performance as they lie in the signal path. Prior to describing the various kinds of comb filters, a brief discussion of delay elements appropriate to video comb filters is presented.

3.5.2.1 Delay Elements for Video Signals

As mentioned earlier, the efficient implementation of comb filters is critically dependent upon the development and use of a high-performance delay element, especially when multiple line comb filters are to be designed. These delay elements are used to create the 1H delays required in comb filter construction. The quality of the resulting video signal is a function of many parameters of the delay line; however, two are especially critical. The first, applicable to analog delay lines, is the noise introduced by the delay element. Depending on the type of delay element used, the noise introduced into the signal can be of sufficient magnitude to substantially impair the performance of the entire system. Second, is the accuracy of the delay itself. The frequency response of the resulting comb is determined by the amplitude matching of the delayed and

undelayed signal and the amount of delay. Comb filters for chrominance separation are used near the subcarrier frequency which is $227.5 f_H$. Complete loss of performance will result if the zeros of the filter move by one-half f_H at the frequency of operation, with marked degradation at much smaller errors. Thus, the amount of delay provided by the element must be controlled to within small fractions of a percent.

There are four predominant types of delay elements currently used for video signals. Each type has its merits and specific niche of application. The four in use now are the: (1) Glass Delay Line, (2) Modulated Bulk or Surface Wave Delay Lines, (3) Digital Memory Delay Lines, and (4) Analog Charge Coupled Device (CCD) Delay Lines. Each will be described below, with relevant comments pertaining to their use in a video comb filter. A fifth type of delay element, based on switched capacitor technology, has been designed for use in the proposed chip. This will be explained in the next chapter.

(1) Glass Delay Line - This type of delay element is the most popular for low cost, medium performance systems. The principle involved in this device is piezoelectricity and the fixed velocity of bulk waves through a solid. These delay lines consist of a piezoelectric element fixed to opposite ends of a slab of glass. By applying an electrical signal to one element, the piezoelectric characteristic of the element converts the electrical signal into mechanical waves. These waves then travel to the opposite end of the glass slab, where they are reconverted back to electrical signals by an inverse process. The amount of delay is fixed by the length of the glass slab, as the bulk velocity of what are essentially ultrasonic sound waves is a material constant.

As glass delay lines are used in large quantities in the consumer TV market, the cost of such delay lines is quite low, in the region of five dollars. However, their performance is quite limited. First, the bulk velocity of glass is a frequency dependent quantity. Thus, signals of different frequencies will be delayed by differing amounts — this is analogous to a non-linear phase characteristic in a filter. In addition, the process of converting an electrical signal to an acoustic wave and back imparts considerable loss to the signal, in the neighborhood of 12 dB. Moreover, this loss is not constant across the frequency band. Manufacturers have designed these elements to exhibit reasonably constant group delay and attenuation characteristics over a narrow range of frequencies about the subcarrier frequency. Outside of this range, the anomalies imparted by the delay line make it unsuitable for video applications [20]. Thus, circuits making use of this type of delay line always place a bandpass filter in

front of the delay element to prevent unwanted signals from entering. Therefore, these elements are suitable only for chrominance separation. Furthermore, there are non-linearities that are introduced mainly by the piezoelectric conversion process. This limits the dynamic range of these devices to roughly 45 dB. Delay of the entire composite signal for overall comb filter group delay matching is not possible. Finally, the glass delay line is not readily integratable on the traditional silicon chip; hence, its use is limited to discrete circuitry or as an external component to an integrated circuit.

(2) Modulated Bulk or Surface Acoustic Wave Delay Elements use roughly the same principle as the simple glass delay line with one major exception. Prior to conversion into an acoustic wave, the electrical signal is used to modulate a carrier signal, at a frequency usually much higher than the signal frequency. The principle is to constrain the resultant bandpass signal to a narrow enough range compared to the carrier frequency that the non-linear delay and attenuation characteristics of the delay element itself is no longer a factor. After reconversion into an electrical signal, demodulation is performed to recover the original baseband signal [21]. The higher frequencies used makes the use of surface acoustic waves possible; this usually results in smaller and more stable devices.

The chief drawback of this system is the added complexity and cost of the modulation and demodulation process. Commercially available devices cost on the order of two to three hundred dollars. However, the performance of these devices allows full bandwidth delays to be implemented — this allows group delay matching of the entire comb filter system, a significant advantage in implementing multiple line filters. In addition, as circuitry is already required to perform modulation and demodulation, much of the insertion loss can be compensated for by the electronics. The noise and distortion performance of these devices tend to be better, especially if frequency modulation is utilized. As with glass delay lines, the absolute delay of the element is fixed by the physical dimensions of the delay element.

(3) Digital Delay Elements - This class of delay elements depends on an analog to digital conversion prior to the delay operation. The structures involved usually are of a shift register type, although dual port RAM structures have also been used. As the video signal is now represented in discrete time as binary numbers, delaying the signal amounts to storing the numbers and waiting a fixed number of clock cycles prior to reading the numbers back.

The chief disadvantage of this element is the analog to digital conversion that has to be performed prior to delay, though this is less of a concern when the comb filter is to be followed by more complex DSP operations. However, since the signal is discrete and quantized, issues of noise introduced by the delay elements and delay errors are nonexistent. The discrete time nature of the signal, however, restricts the delay to a integral number of clock cycles. This restricts the sampling rate of the overall system to be an integral multiple of the line frequency f_H , in order to realize a $1H$ delay. Finally, these devices are easily integratable on an IC, along with a monolithic A/D.

(4) CCD Delay Elements - These devices are the "analog" equivalent of the digital delay elements. Rather than storing a digital number in a cell and waiting a fixed number of clock cycles to recover the number, CCD elements work with charge stored in a special cell similar to a MOS capacitor with switches. The delay function is accomplished by connecting a large number of these cells serially in a row. Charge stored in one cell can be transferred from the first to the second cell and so on, down the line at given clock intervals. Thus, the structure mimicks a shift register. The discrete time signal is converted to a charge and injected into the first cell; with subsequent clocks, the charge in each cell is transferred down the line with the newest signal being injected into the first cell. N clock cycles later, (N = number of cells in the line), the charge is reconverted back into a voltage; this results in a N clock cycle delay.

CCDs require no analog to digital conversion, and are especially well suited for applications where no subsequent DSP is desired. Integration into a standard IC is possible, although a special step or two are usually required to form the CCD cells. As with the digital equivalent, the discrete time nature of the device provides a delay accurate to the system clock. The main difference of this device, is that in handling an analog signal, it remains susceptible to noise and other imperfections not seen in an all digital implementation [22]. The two main sources of noise in CCDs are the non-unity charge transfer efficiency and dark current. The amount of charge transferred from one cell to another is known as the charge transfer efficiency. This is a random variable whose mean is very close to unity, typically 0.99995. However, the random nature of this value imparts an uncertainty in the resultant charge whenever a transfer operation is performed. This is equivalent to the addition of a small amount of noise each time a transfer is performed. Dark current is a leakage current associated with the charge storage cell, and again is a random quantity that depends strongly on

temperature and processing imperfections [23]. Both these effects accumulate over multiple shifts, thus becoming worse with long delay lengths. Moreover, long CCD delay lines tend to wander in the DC level of the signal from one end to another. The dynamic range of the CCD delay line is dependent on how well the random effects described above are controlled. Proper layout and circuit design has resulted in CCD structures with dynamic ranges above 56 dB [24].

3.5.2.2 Simple (1H) Comb Filters

The simplest comb filter utilizes a single delay element — by delaying the composite signal by a period equal to exactly one line, H, and subtracting it from itself, the resulting signal will be largely chrominance. (Fig. 3.1) This follows from the transfer function of this structure,

$$H(z) = K[1 - z^{-H}]$$

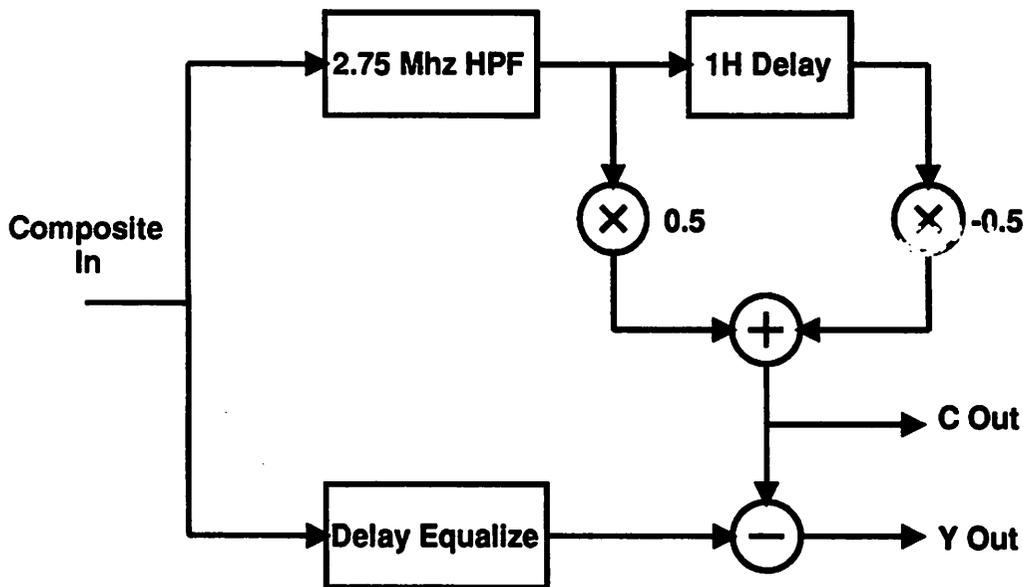


Figure 3.1: Simple 1H Comb Filter

The constant K is used to set the gain of the structure, and is usually set to be one half. The frequency response of this structure has zeros at DC and multiples of (1/H) or 1H. (Fig. 3.2) As the luminance energy is concentrated at multiples of fH, and the chrominance energy is concentrated in between, this filter effectively separates the chrominance from the luminance. As above, the luminance signal is obtained by subtracting the resulting chrominance signal from the composite signal. Most structures place a band pass filter in front of the delay element for the following

reasons: (1) The delay element, especially if implemented using a glass (piezoelectric) element delay line, may have poor group delay characteristics outside of a narrow operating frequency band. Thus, removing signal components outside of this band will prevent distorted signals from entering the signal path. (2) Most of the luminance energy is contained in the low frequencies; by removing these signals, the chance of spurious leakage into the chrominance path is minimized. (3) Moreover, by removing the large amplitude low frequency luminance signals, the signal passing through the delay element can be scaled appropriately, yielding a better signal-to-noise ratio.

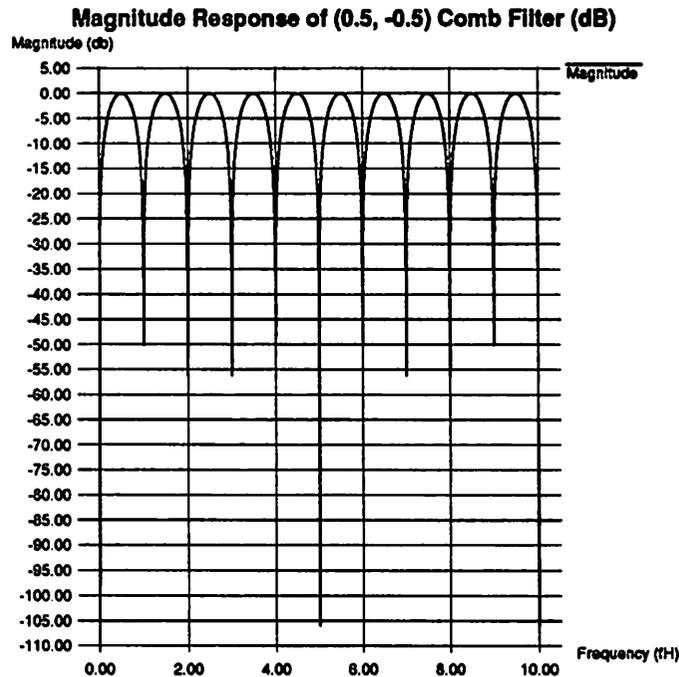


Figure 3.2: Frequency Response of Simple (1H) Comb Filter

This approach for Y/C separation is becoming more commonplace, especially with the proliferation of large screen monitors. The loss of luminance bandwidth with the first method described becomes intolerable as screen dimensions increase. Moreover, the cost for implementing this structure is reasonable, with glass delay lines priced in the approximately in the 5 dollar range. However, the architecture described above suffers from a group delay that is a non-integral multiple of a line period, rather for this structure, it is one-half of a line period. This will result in the luminance and chrominance being out of step in the vertical dimension. This is usually not noticeable when the filter is used once; however, if the signal is passed through this filter multiple times, the offset between the chrominance signal and luminance will cause the color

portion of the image to slip down from the monochrome image. This can be fixed by attempting to delay the luminance image appropriately. However, this requires that a delay element capable of handling the wide bandwidth luminance signal be employed. Glass delay elements are inappropriate as they do not have the bandwidth necessary. Thus, other more expensive means such as digital memories, modulated delay lines, or CCDs must be used. As low cost is a primary advantage of this type of comb filter, such methods are rarely used.

3.5.2.3 Three Line (2H) Comb Filters

If the group delay characteristics of the second method are unacceptable, a structure utilizing two line delays and having a filter length of three lines is used. Here, the equivalent transfer function is:

$$H(z) = Kz^{-H}[z^H + 2 + z^{-H}]$$

Again, K is a gain scaling factor, and is usually one-quarter for this structure. The z^{-H} factor is used to translate this non-causal filter structure into a causal structure. The frequency response of this filter is similar to the comb filter above, except that each null is a double zero. (Fig 3.3) Thus, the result of this filter, the chrominance component, will have a smaller magnitude of spurious luminance signals. As the group delay of the comb structure is one line, the composite signal must be delayed by one line prior to the subtraction operation. There are several different architectures for obtaining this response. The main tradeoff lies in the number of delay lines vs. the number of bandpass or highpass filters required. If the technology used places a high cost on the number of delay lines, then the structure in Figure 3.4 is appropriate. However, this architecture requires that at least the first delay line be a full bandwidth comb filter for luminance signal delay matching, and is also not well suited for adaptive comb filters that are described below for reasons that are explained later. A more elegant method of performing the three line comb is shown in Figure 3.5. This structure places a bandpass filter before the combing operation. This requires a separate full bandwidth delay element; thus for cost reasons, this structure is used only for applications requiring the highest performance.

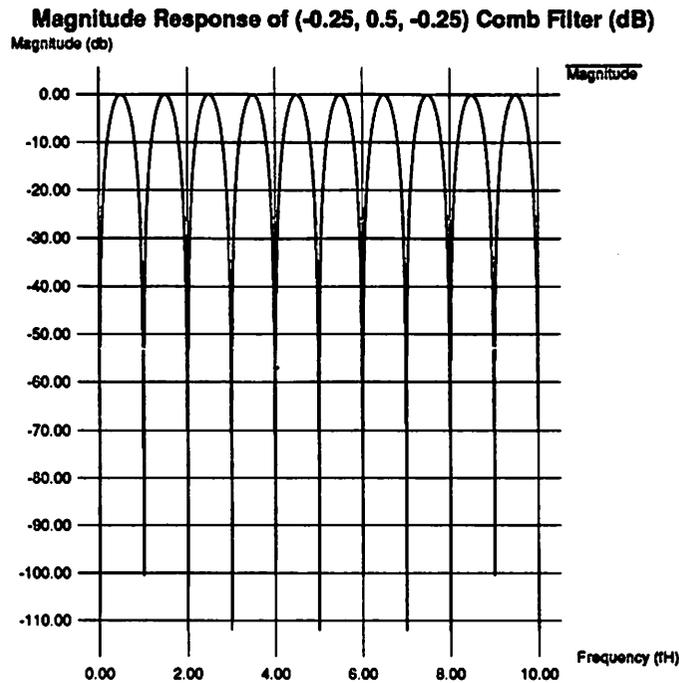


Figure 3.3: Frequency Response of the 3 Line (2H) Comb Filter

3.5.2.4 Adaptive Comb Filters

While comb filters do an excellent job of Y/C separation under most conditions, they suffer from a drawback which unfortunately, creates the exact condition the filters are designed to remove. Under certain conditions, comb filters promote the occurrence of the cross luminance condition, producing a very objectionable visual artifact. Comb filters operate under the premise that the chrominance is properly interleaved with the luminance structure. However, when there is a high chrominance vertical spatial frequency in the image, this will no longer be true, as the "wedges" in the spectrum will tend to spread out. This condition occurs most often with a horizontal color transition where the colors between successive scan lines are markedly different. The result is that a significant portion of the chrominance energy will fall outside the passband of the comb structure, and thus will not be properly subtracted from the composite signal. The residual chrominance in the luminance path will result in cross luminance which manifests itself as a row of "hanging dots."

Several solutions have been proposed for this defect, with varying degrees of success. The simplest is to perform both a bandpass and comb filter separation and average the two. Although this results in visual artifacts along both horizontal and vertical color transitions, the magnitude of

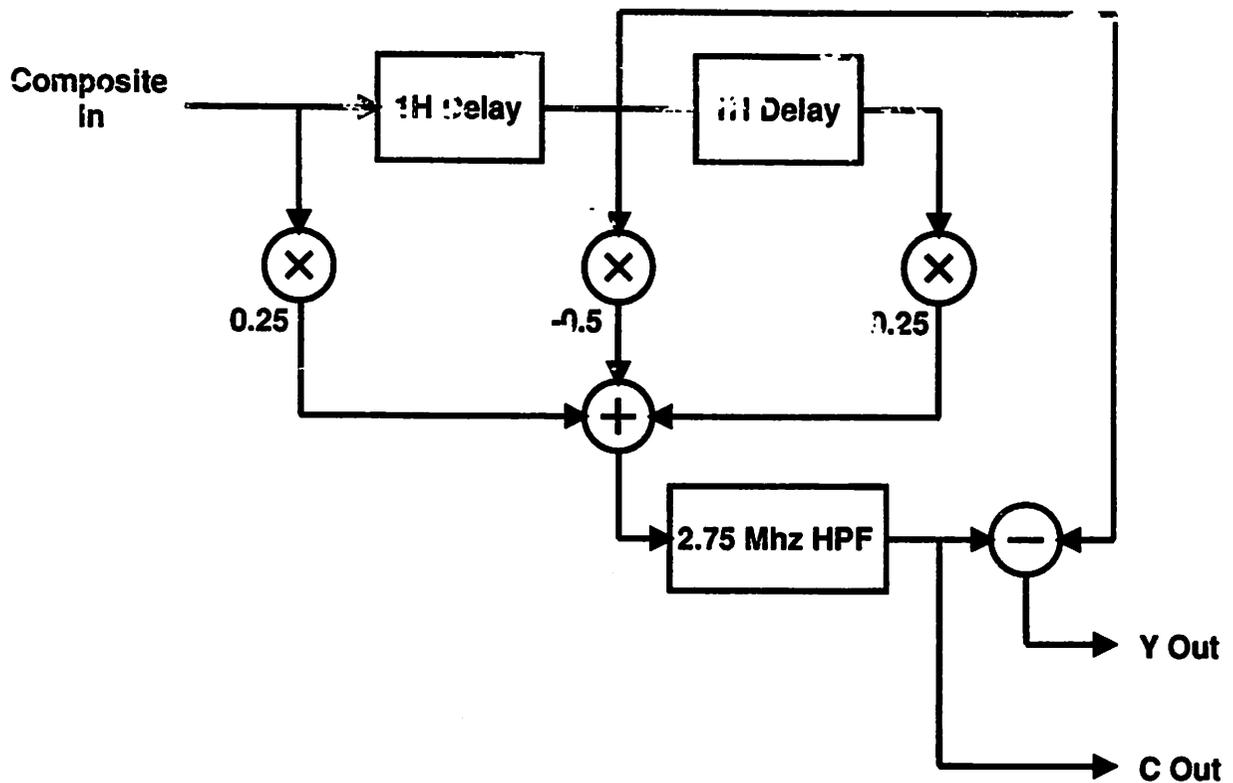


Figure 3.4: Minimum Delay Line 2H Comb Filter

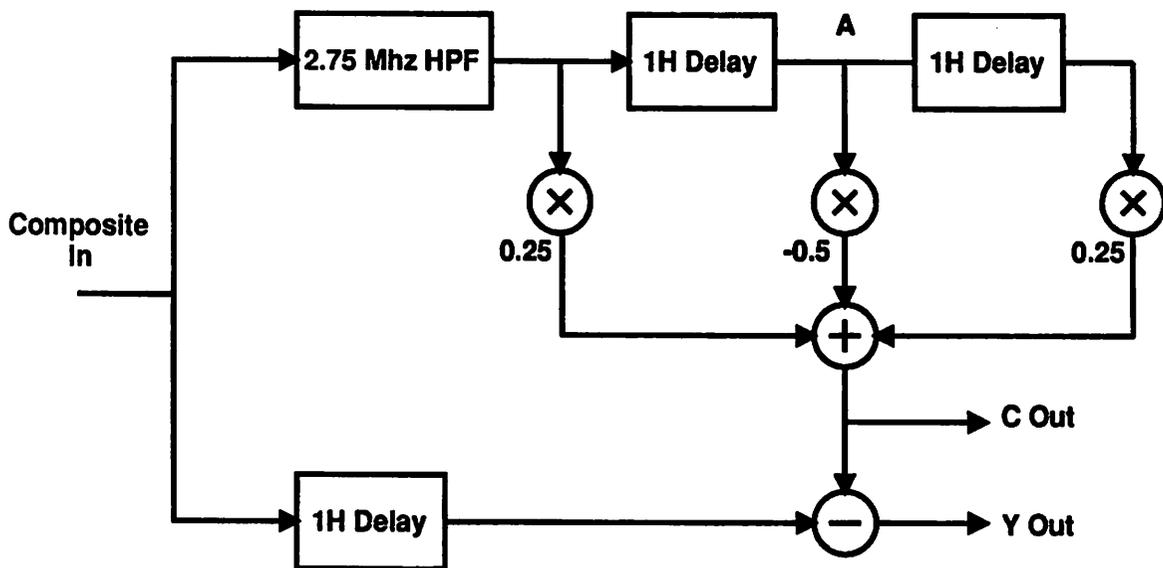


Figure 3.5: 3 Delay Line 2H Comb Filter

each is reduced and thus less noticeable. However, the most successful solutions to the problem involve adaptive methods [25 - 27]. This method is suited best for the three line comb structure, and will be discussed with respect to that system. The cross luminance problem in comb filters occurs when the chrominance correlation between lines in an image becomes non-existent as in the case of a horizontal color transition. Thus, by detecting this condition and switching the Y/C separation process from a comb to a bandpass system, the defect will no longer be visible. As the condition of non-correlation usually occurs over only one or two lines, the effects of switching back to the bandpass with its defects are well concealed. The drawbacks to this method are (1) the complexity of detecting the non-correlation and subsequent algorithms to make the decision when to switch between comb and bandpass, and (2) the added overhead of performing a bandpass separation in addition to a comb separation.

Fortunately, the signal at Point A in three line comb filter structure of Figure 3.5 is already a bandpassed chroma signal. Thus, by subtracting that signal from the composite, a Y/C separation function corresponding to the bandpass method is obtained. The topic of correlation detection and decision control has been investigated by several groups and a number of algorithms have been proposed. The general approach however, is common to all of these systems. First, a value corresponding to the interline chroma correlation is computed. The simplest method is to subtract pixel by pixel, the bandpassed chroma signals from the first and third lines in the comb filter structure. The resulting control signal is then averaged over a few pixels, and is representative of the average non-correlation of the chroma between lines. This value is then compared against a two level reference. Below the lower value, corresponding to high correlation, the system will use the comb filter. In between, it will choose a half comb, half bandpass approach. When the correlation falls to a very low value, the Y/C separation will be accomplished using a straight bandpass. The half bandpass half comb transition is necessary to prevent large transients from entering the signal path, as may be the case when switching directly from comb to bandpass [28].

This method of adaptive switching tends to be a bit conservative in the use of the comb filter. As the control signal is derived from a bandpass filtered composite signal, high frequency luminance information will adversely affect the system. In particular, the presence of high frequency diagonal luminance information will fool the adaptive circuit into believing that there is a high degree of non-correlation. This is particularly troublesome, as high frequency diagonal luminance information causes a large amount of cross color with bandpass Y/C separators. Thus, better solution would be to calculate the luminance correlation in addition to the chrominance correlation.

Such methods have been devised, but are of such complexity that inclusion in an analog signal processing circuit is not practical.

A diagram showing the adaptive three line comb filter is shown in Figure 3.6. The correlator implements the function described above to generate a decision signal to operate the switch. The availability of the bandpass signal in the structure of Figure 3.5 gives it a substantial advantage over that of Figure 3.4. In order to provide the same functionality, two additional bandpass filters are required to prevent DC transients through the switch. As a result, this structure is rarely used although it uses one less delay line.

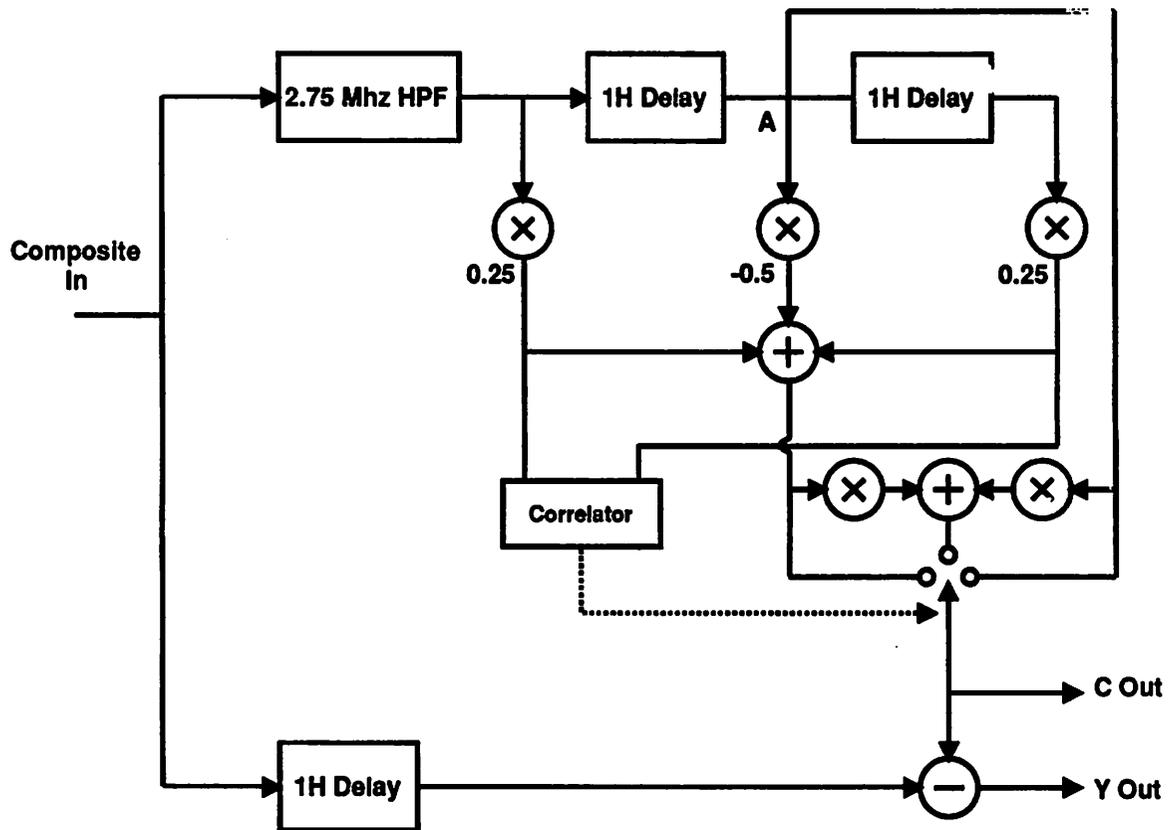


Figure 3.6: Adaptive 2H Comb Filter with 3 Delay Lines

3.6 Chrominance Signal Demodulation and Filtering

The outputs of the Y/C separator are the luminance and chrominance components of the composite video signal. The luminance signal is already at baseband, having never been modulated. However, the chrominance signal is a QAM signal which has two baseband components. Thus, a demodulation operation is necessary to recover the original two chrominance signals, Pr

and Pb.

The chrominance demodulation process consists of two parts. The first is an automatic gain control circuit which insures that the chrominance signals are at the proper amplitude for the dematrixing operation that converts YPrPb or YIQ to RGB. This is done by observing the amplitude of the colorburst signal, and applying a variable gain to the chrominance signal path to bring it to the proper value. The second circuit is the actual demodulation circuit. QAM demodulation involves the inverse process of modulation; the incoming chrominance signal is multiplied by a sinusoid and a cosinusoid of proper phase to recover the color components.

Although the original color equations were written in terms of Y, Pr, and Pb, recall that the NTSC standard calls for the use of I and Q channels to maximize the color bandwidth where needed, and minimize the bandwidth in other areas. The subsequent vestigial sideband modulation of the I channel produces an equivalent baseband frequency response of the I channel that is non-symmetric about DC. This results in the production of complex components in the signal at frequencies beyond the edge of the upper (vestigial) sideband. In the NTSC QAM system, this will result in crosstalk between the I and Q components above 620 KHz. Thus, demodulation of the full I and Q components requires a more complex filtering process. Moreover, the colorburst signal used to lock the local demodulating oscillator has a phase which corresponds to -Pb. In order to demodulate along the I/Q axes, the local oscillator must incorporate a 33 degree phase shift network. To remove the crosstalk of the I channel into the Q channel, the two signals are subjected to filters of different bandwidths, 600 KHz for Q channel, and 1.3 Mhz for the I channel.

The complications associated with I/Q demodulation can be avoided by using what is sometimes called a "narrowband color" demodulator. This method demodulates the chrominance along the Pr and Pb axes to recover Pr and Pb directly. To prevent crosstalk caused by the vestigial sideband, both channels are bandlimited to 600 KHz. The resulting loss of chrominance information in certain colors results in a less than optimal image; however, the degradation is minimal and is usually worth the cost and complexity savings.

Actual circuit implementations of the chroma demodulators fall into two broad classes, those employing continuous time techniques, and those employing sampled data techniques. Continuous time demodulators almost invariably use a form of the four quadrant analog multiplier. Two oscillators locked to the colorburst signal are used. Due to the invariance of the subcarrier frequency, phase shift networks are used to generate the quadrature signal, and also to provide the

33 degree phase shift when doing I/Q demodulation. Care must be taken in constructing the multiplier to insure that it is properly balanced to prevent subcarrier leakage into the signal path. This implementation is very popular, especially in circuits utilizing discrete components as both bipolar multiplier chips are readily available and inexpensive.

Sampled data demodulation has been done exclusively in the digital domain. Continuous time oscillators are replaced by counters and look up ROMs to generate the necessary sampled sinusoids. The clock feeding the counters is locked to an appropriate multiple of the colorburst frequency. Quadrature is guaranteed by the nature of the ROMs, and any arbitrary phase shift can be pre-programmed into the look up tables. The chief disadvantage of digital demodulation is the need for a high speed digital multiplier. Video is frequently quantized to eight bits, thus requiring eight by eight multiplies to be carried out at rates of approximately 15 Ms/s. The sampling rate of systems employing digital demodulation is restricted by the finite size of the look up ROMs. An irrational relationship between the sampling rate and the subcarrier frequency would require an infinite number of entries in the look up table. In practice, the sampling rate is usually a multiple of the subcarrier frequency, although the new CCIR 601 standard calls for a sampling rate of 13.5 Ms/s independent of whether NTSC or PAL encoding is used [29].

The special case of sampling at four times the subcarrier frequency, $4f_{sc}$, allows considerable simplification in the digital multiplier section if certain constraints are met. This modification also allows the possibility of performing sampled data demodulation in the analog domain, the technique of choice in the proposed chip. This method of demodulation will be discussed later in this report.

Regardless of the method used to demodulate the signal to baseband, the resulting signals contain double frequency terms, and in the case of I/Q decoding, crosstalk components. Thus, lowpass filtering of both signals is required. In the case of "narrowband" color demodulation, all that is required are two low pass filters that attenuate all signals above 600 KHz. Note that as the filters will have some amount of group delay, the luminance path must contain an equivalent compensating delay to keep the signals together in time. As with the composite filter, linear phase characteristics are a requirement. Systems utilizing I/Q demodulation require filters with unequal bandwidths for the I and Q channels as described above. Such filters, especially those in the continuous time domain, will have unequal group delays, thus necessitating a compensating delay network in both the I and luminance channels. A diagram of a typical full bandwidth chroma demodulator is shown in Figure 3.7.

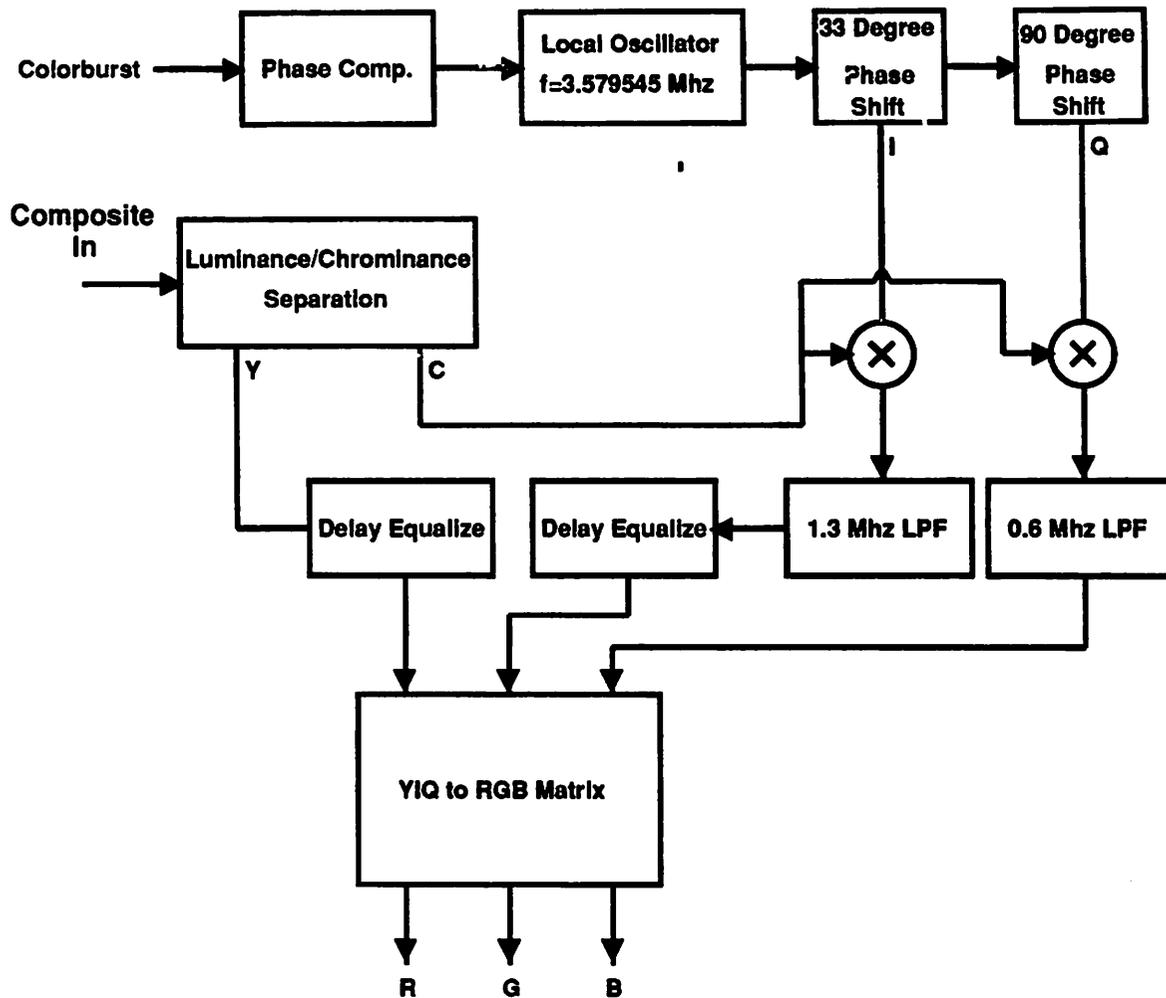


Figure 3.7: Typical Wide-Bandwidth NTSC Chroma Demodulator

3.7 Chrominance Demultiplexing

The final step in NTSC processing is the generation of RGB values from the recovered Y and IQ or PrPb depending on the type of demodulation performed. In the case of I/Q demodulation, the I and Q signals are first dematrixed to form the Pr and Pb signals. The Pr and Pb signals are then matrixed with the Y signal to form the RGB values that originally comprised the image [30]. Mathematically, the operation can be written as:

Optimal Architectures for an Integrated NTSC Decoder

$$\begin{bmatrix} Y \\ Pr \\ Pb \end{bmatrix} = \begin{bmatrix} 1.00 & 0.00 & 0.00 \\ 0.00 & 0.948 & -1.105 \\ 0.00 & 0.621 & 1.73 \end{bmatrix} \begin{bmatrix} Y \\ I \\ Q \end{bmatrix}$$

and

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.00 & 1.00 & 1.00 \\ 1.00 & -0.5085 & 0.00 \\ 1.00 & -0.1864 & 1.00 \end{bmatrix} \begin{bmatrix} Y \\ Pr \\ Pb \end{bmatrix}$$

The resulting RGB values are then applied to the image reconstruction device (e.g. CRT) to reproduce the original image.

For some applications, especially those incorporating digital special effects, the dematrixing operation is sometimes omitted deliberately. First, signals in the RGB domain require three high-bandwidth channels, while those in the YPrPb domain can utilize reduced bandwidth for the color components. This is especially useful in digital systems where bit rate reduction is a key concern. In addition, color processing is easier when working in the YPrPb domain, as the luminance and chrominance are separated. Thus, processing can be done on one component without disturbing the other. In these cases, dematrixing is done only to recover the Pr/Pb components from the I/Q components from wideband demodulation.

Chapter 4

Existing Circuit Embodiments of NTSC Decoders

In this chapter, a discussion and comparison of the various methods of performing the NTSC decoding task will be presented—the results from this chapter will be used to gauge the relative merits of the technique for the proposed chip discussed in chapter five. Unlike A/D converters or S/H circuits where, for a given application, it is relatively easy to compare one embodiment to another each embodiment of a NTSC decoder is different; hence, it is nearly impossible to perform a one to one comparison of different approaches. Therefore, extrapolation techniques will be made in order to facilitate comparisons when needed. Cost figures are based on rough estimates and therefore maybe inappropriate in some circumstances.

4.1 Overall Performance of Various Circuit Embodiments

Regardless of the actual technology or circuit techniques used to realize the NTSC decoding function, the overall circuit performance can be measured using industry standard techniques, some of which will be discussed here. Unlike audio, where the current circuit challenges are to maintain and extend the dynamic range to 100 dB and beyond, the challenges in video are to increase the speed (frequency response) and the complexity of signal processing to achieve the highest image quality possible. As with any practical implementations, the constraint is usually cost, with higher performance circuits demanding more expensive components and higher assembly costs. The NTSC decoder is no exception. Medium performance circuits, acceptable in items such as consumer television sets demand low cost, while other situations such as studio or broadcast equipment demand the highest performance possible.

One of the best "yardsticks" of circuit performance is the frequency response of the luminance channel. The standard NTSC luminance bandwidth is 4.2 Mhz. Many low-cost implementations suffer substantial loss in response above the subcarrier frequency, independent of the use of a notch filter for color information removal. This is a byproduct of the use of less complex filters in order to simplify the circuit, as maintaining a flat response to 4.2 Mhz usually involves a filter with a very steep rolloff above 5 Mhz. Of course, maintaining a flat response also necessitates use of a comb filter for Y/C separation which increases the complexity of the circuit substantially. Poorly designed filters exhibit ripple in the passband response, or worse yet, induce ringing in the

time domain, resulting in ghosting of step transitions in the image. A small degree of ringing always exists in circuits and is tolerable — however, a poorly aligned filter will cause visible ringing from the sync pulse which is very objectionable.

Another indication of the performance of the circuit is the type of Y/C separation performed. As mentioned in chapter three, Y/C separation can be done with and without comb filters. Due to the superior performance of comb filters, nearly all professional equipment utilize comb filters of some sort, with some higher grade consumer television sets also incorporating comb filters, especially in the larger screen sizes. The best performance is obtained with the three line (2H) comb filters which, due to their complexity and cost of implementation, are found only in studio grade equipment. Single line chroma only comb filters utilizing glass delay lines are increasingly popular, with attendant cost reductions. However, the lowest cost method of Y/C separation is still the single LC tuned notch filter, and is still commonly used on low cost television receivers.

The remaining methods of judging the performance of a NTSC decoder include:

- (1) Signal to Noise - Video signal to noise is measured using a p-p signal to rms noise ratio and should measure at least 54 dB in the luminance channel in a 100 Khz to 4.2 Mhz bandwidth.
- (2) Differential Gain and Phase - As the color information is encoded as the amplitude and phase of a subcarrier, circuits that exhibit changes in their transfer function with the average DC level of the signal will adversely affect the chroma signal. A simple example is a circuit that has a large parasitic junction capacitance. The value of the capacitance varies with the DC level of the signal and will cause a phase shift at the subcarrier frequency as a function of the luminance of the image. Differential gain (DG) is defined as the percentage change in gain of a network at f_{sc} as the DC level is varied from 0 to 100 IRE (black to peak white). Differential phase (DP) is defined as the phase change of a network at f_{sc} as the DC level is varied from 0 to 100 IRE. These figures should be below two percent and two degrees.
- (3) Non-Linearity - A specialized test signal known as a 2T test pulse is used to check the circuit for non-linearities and ringing in the filters. A metric known as the K-Factor is used to quantify the non-idealities. This measure should be less than two percent.

4.2 Analog vs. Digital Architectures

Nearly all of the architectures for NTSC decoding that have been reported to date neatly fall into one of two categories, those utilizing analog signal processing, and those using digital signal processing with the aid of an A/D converter at the start of the signal path. Of the two, the DSP approach is becoming more popular as the cost of high-speed A/D converters and DSP hardware falls. Moreover, digital processing lends itself well to integrated design, as the time to production is shortened with the use of standard cells and sea-of-gates approaches. Use of silicon compilers to generate the FIR filter sections have been reported with resultant shortening of the design cycle. However, the high data rate of digital video still results in large areas of silicon dedicated to the DSP core. A very rough estimate of the area consumed by a DSP equivalent to the front-end filter and decimator for the proposed chip is 4.9 mm² or five times the projected area for the equivalent function on the proposed chip. DSP does have its advantages in its ability to carry out complex signal processing functions — a particularly good application for DSP is in the area of adaptive comb filtering. However, in general, all-digital solutions carry the penalty of larger silicon area. Moreover, they require an A/D in front, and if oversampling is employed, the conversion rates can quickly reach the limits of monolithic CMOS converters.

As an alternative, use of analog signal processing as is being proposed in this chip, has the advantage of reducing silicon area by performing some functions in the analog sampled data domain. Analog sampled data does not require an A/D converter at the head of the signal path, and is not limited by the somewhat restrictive bandwidth of data converters. Functional blocks such as filters can be designed in a fraction of the area of an equivalent digital filter. The precision of analog signals is not limited by the number of bits used as is in digital systems. The somewhat limited functionality of analog sampled data circuits can be overcome through the use of novel circuit techniques, such as the transversal FIR filter and the Analog RAM cell proposed in this chip, to provide signal processing capabilities normally reserved for DSP architectures.

4.3 Discrete and Mixed Circuit Embodiments

Discrete components are most often used to implement the NTSC decoder in low-cost, mass-produced television receivers as currently available integrated circuits still cost more than the discrete components needed to perform the same task. The performance of these circuits is only fair, as the drive to decrease cost is stronger than that to increase performance.

At the other end of the spectrum, discrete components are also used in the highest performance arena, as the integrated circuit embodiments do not perform to the level required. These circuits take up thousands of square centimeters (hundreds of square inches) of printed circuit board space, and cost thousands of dollars each. These circuits usually implement both analog and digital signal processing, usually performing the filtering in the analog domain with LC filters that have been carefully hand tuned for amplitude and phase response. LSI A/D converters are used to digitize the signal for Y/C separation and demodulation, followed by D/A converters to reconstruct the analog signal. Extensive use is made of SSI and MSI circuits to reduce the overall component count.

4.4 Integrated Circuit Embodiments

There have been several attempts to integrate all or part of the NTSC decoding function onto a single or a set of integrated circuits. In general, they have resulted in large chips and performance that sits between low-cost discrete circuits found in television receivers and the high-performance studio grade circuits. For purposes of providing some indication of the type of work reported, the following cases are presented as points of reference:

(1) *Digital Signal Processors for Decoding/Encoding Color TV Signals* [31] - A 1986 work which is a single chip NTSC decoder which performs all the functions in the digital domain. No A/D or anti-aliasing is performed on chip, thus, a high speed (15 Ms/s minimum) 8 bit converter is required in addition to this chip. Moreover, no sync signal processing is performed. All filtering is performed using FIR filters, thus insuring excellent phase response. A simple 1H comb filter is used for Y/C separation. The line memory is implemented using a specialized array of latches designed to minimize power. However, the luminance channel frequency response is down 6 dB at f_{sc} , which precludes the use of this chip in very high-performance video systems. Total chip area is 119 mm² in a 2 μ m P-well CMOS technology, dissipating 1.5W. From the die photograph provided, the Y/C separator appears to occupy an area of about 34.12 mm², and the luminance low pass filter occupies about 15.1 mm².

(2) *A Single-Chip CMOS Analog/Digital Mixed NTSC Decoder* [32] - A 1990 work which also performs the NTSC decoding function on one chip. Sync signal processing is performed in the analog domain, with the remainder of the signal processing being performed digitally; an A/D converter is provided on-chip. However, as the sampling rate is only $4f_{sc}$, a

Chapter 4 Existing Circuit Embodiments of NTSC Decoders

high order external anti-aliasing filter is required. Moreover, the quantization is performed only to 6 bits, although it is reported that novel techniques overcome this somewhat low resolution and result in high-quality images. The Y/C separation is performed using a simple notch in the luminance path. A dematrixing circuit to convert YPrPb to RGB is provided. The overall chip size is 121 mm^2 in a $1.2 \text{ }\mu\text{m}$ double metal, double poly CMOS twin-tub technology. Power dissipation is 980 mW. The DSP core, which contains the notch filter and chrominance demodulators appears to occupy about 72 mm^2 based on the provided die photograph.

(3) *The Digital Television Signal Processing System from Philips/Signetics, SAA9051, etc.* [33, 34] - This is a complete digital television signal processing system split up over eleven VLSI chips introduced in 1989. It is intended to operate on all three color television standards (NTSC, PAL, and SECAM), while providing high quality images. Of particular interest are the TDA8708 A/D Converter, the SAA9057 Clock Generator and the SAA 9051 Chroma Demodulator and Decoder, as these three chips most closely duplicate the functionality of the proposed chip. One fundamental difference of this system is that the sampling rate is 13.5 Mhz, to conform with the CCIR 601 digital video standards. The A/D Converter is an 8 bit flash converter implemented in bipolar technology. Unlike the circuits above, this system utilizes double oversampling at the quantizer to relax the anti-alias filter requirements. The clock generator chip is also implemented in bipolar technology. The chroma processing chip is fabricated in $1.5 \text{ }\mu\text{m}$ NMOS technology with a dissipation of 1.5 W. The Y/C separation is performed using a 1H line delay made of a 868×8 bit FIFO RAM. As the sampling frequency is not $4f_{sc}$, the sine and cosine demodulating signals are generated through the use of a ROM look up table. The chip area is 125 mm^2 and is reported to include 200,000 transistors. No information on the area of the remaining two chips are available.

(4) *A CMOS-CCD Comb Filter with Dropout Compensation for a VCR* [35] - A 1988 work which performs the signal processing in the analog domain utilizing CCDs as the delay element. This chip only integrates the delay element and comb filter circuit. Operating at $4f_{sc}$, it achieves a S/N of 56 dB, with a die size of 13.9 mm^2 . Its frequency response is down 1.3 dB at f_{sc} though, and its DG and DP are three percent and three degrees respectively. No mention of technology is made, however, it is estimated at $1.5 \text{ }\mu\text{m}$.

(5) *A CMOS-CCD Video Delay Line* [36] - An early (1984) work which implemented just the delay line to form a comb filter on a single chip. Again, the actual technology used was not

Optimal Architectures for an Integrated NTSC Decoder

reported, but it is estimated to be 3 μm . It operates at a clock frequency of 10.7 Mhz, and utilized 680 registers to provide a 1H delay line. The chip consumes 100 mW using a +5 and a +9 volt supply. Overall size is 8.75mm².

4.5 Comparisons with the Proposed Chip

As mentioned earlier, it is difficult, if not impossible to make a direct comparison of prior work with the proposed chip. However, using extrapolative techniques, a comparison on two fronts will be made. The first, is the size taken for a common sub-block, the Y/C separator which includes the delay lines if applicable for comb filtering. The second comparison will be made on the basis of total chip area to perform the NTSC decoding process. For this comparison, an additional $3.89 \times 10^6 \mu\text{m}^2$ will be added to the proposed chip to account for the quantization step after the demodulation of the chrominance signal, making the total area for the proposed chip $54.99 \times 10^6 \mu\text{m}^2$ (85,260 mil²).

To facilitate an equal basis of comparison, the areas quoted for the above prior work are scaled to an equivalent area in 1 μm technology. Figure 4.1 shows the area for a single 1H (single - ended) delay line for the various architectures. Figure 4.2 shows the total chip area to perform the NTSC decoding and quantization process. Chip areas for the work cited will be extrapolated to match the functionality provided in the proposed chip (A/D, anti-alias, 2H comb filter, etc.)

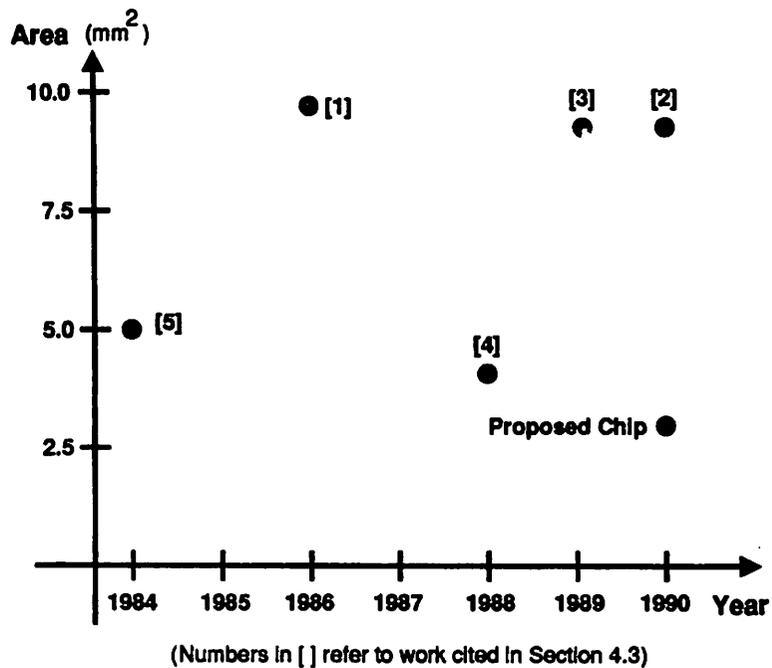


Figure 4.1: Silicon Area for a 1H Video Delay Line for Above Cited Work

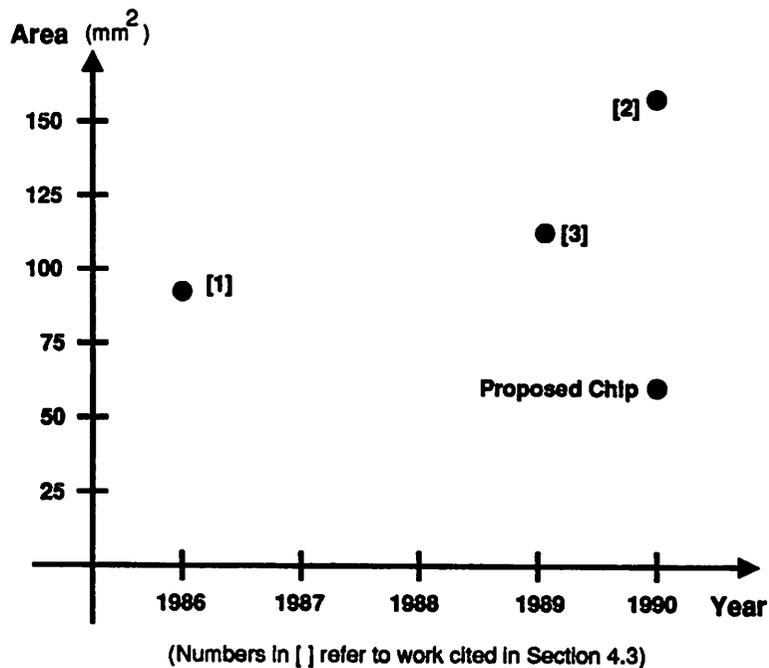


Figure 4.2: Silicon Area for NTSC Decoder Function of Proposed Chip

Chapter 5

Architecture and Circuit Implementation of the Proposed Chip

This chapter will describe the design of each of the sub-blocks described in chapter three that are required to perform the NTSC decoding task. As stated in the introduction, the motivation for this project is to determine whether judicious use of analog sampled data signal processing will result in circuits that perform better and/or consume less silicon area than an all digital implementation.

5.1 Overall Chip Architecture

The overall architecture of the chip closely follows the description of the NTSC decoding process given in chapter three. A block diagram of the architecture is given in Figure 5.1, and shows a logical division of the entire task into 9 separate blocks. Prior to a detailed discussion of each block, a brief outline is given to help place each block in perspective with the other blocks.

- (1) DC Restoration - The first task of any video processing circuit is to restore the DC level. This is especially critical in this chip, as the dynamic range of the signal is limited by the power supply voltages. Thus, proper placement of the signal common mode voltage is critical to prevent saturation of the subsequent circuitry.
- (2) Buffer and S/E to Differential Conversion - Signal processing via analog sampled data structures benefits greatly from the use of differential signals. Among the many advantages are increased noise immunity, power supply rejection, and increased dynamic range. The buffer insures that the remainder of the circuitry will be driven from a low-impedance source.
- (3) Coarse Sync Strip - This block detects the sync tip and generates an active flag during the sync period. This signal is then fed back to the DC Restore block to allow it to properly locate the blanking interval.
- (4) Fine Sync Strip - This block takes the output of the coarse sync stripper and generates a voltage equal to 50 percent of the sync tip amplitude. A comparator then generates an accurate signal that represents the start of every line. As this chip will lock onto subcarrier as opposed to H , the accuracy of this block is not overly critical.
- (5) Phase Locked Loop and Clock Generators - This section contains the crystal VCO used to lock onto the incoming signal colorburst and generate the sampling and related clocks.

As will be seen, the initial sampling is done at a rate of $16 f_{sc}$, with the chroma separation and demodulation done at $4 f_{sc}$. The specific architecture of the chroma demodulator is very sensitive to the phase of the sampling clock. Thus, this block will take a feedback signal from the chroma demodulator to allow adjustment of the sampling phase as necessary.

(6) **4x Oversampling Sample and Hold with 1st Stage Decimator** - For reasons explained in detail later, the chip will initially sample the signal at 16 times the color subcarrier frequency. Built into the sample and hold circuitry is a 3 tap FIR filter that performs a first stage decimation to lower the sampling rate to 8 times f_{sc} . A feedback signal from the next stage filter is used to reduce the number of integrators necessary for the entire filter and to simplify the switching waveforms.

(7) **Second Stage Decimation Filter** - This block takes the signal at a rate of $8 f_{sc}$ and applies the final 5 Mhz low pass filter while at the same time lowering the sampling rate to $4 f_{sc}$. A switched capacitor architecture is used with additional circuitry to preserve the phase response of the filter.

(8) **Luminance/Chrominance Separator** - This block performs the Y/C separation described in section 3.5. In order to obtain the highest performance possible, a 3 line, adaptive comb filter has been designed for this block. The delay lines are constructed using an innovative modification of switched capacitor techniques, which is designed to offer the flexibility of CCDs while eliminating the major source of noise associated with CCDs.

(9) **Chroma Demodulator** - This final block actually performs the quadrature amplitude demodulation and filtering to produce the YPrPb signals from the modulated chrominance signal. An optional dematrixing circuit may be included to transform the YPrPb signals to RGB as may be necessary

The design process for each of the above sub-sections for the proposed chip is outlined below.

5.2 DC Restoration and Input Buffer

The design and proper operation of the DC Restoration circuit is vital to insure that the signal remains within the linear input range of the circuit. Video signals by convention have an amplitude of approximately one volt peak to peak. The chip, operating with ± 2.5 volt power supplies can be expected to have at most a three volt common mode range, beyond which the circuitry will cease to operate properly.

The proposed circuit is outlined in Figure 5.2. The principle of operation is a slight modification of the common practice of varying the DC bias on the input coupling capacitor. In this circuit, a controlled current is either sourced or sunk out of the input buffer summing node. As the amplifier is operating as a shunt feedback device, currents forced into the summing node translate directly into changes in the output common mode voltage.

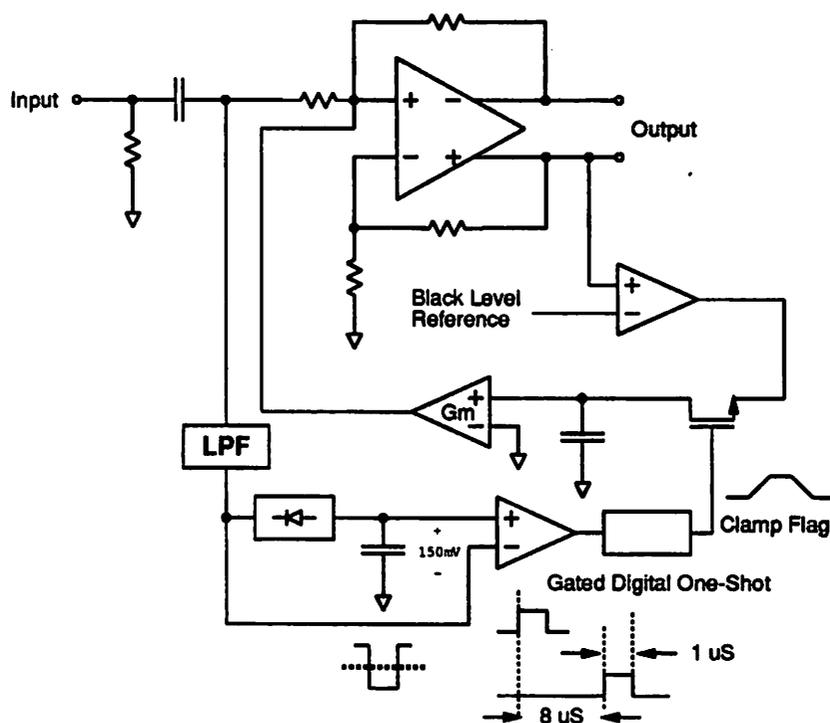


Figure 5.2: Block diagram of input circuit

In addition to performing DC restoration, the amplifier also performs the single-ended to differential signal conversion. This is accomplished by the commonly used resistive feedback network to force the outputs of the amplifier in equal and opposite directions about a common mode voltage. This type of circuit usually does not lend itself well to MOS circuits, as the resistors place a actual current load on the output of the amplifier rather than a capacitive charge load. As a result, the amplifier, which is a folded cascode amplifier, is modified to drive resistive loads with the addition of class A output stage.

5.2.1 CMOS Operational Amplifiers for Resistive Loads

Techniques to allow resistive loads on the output of CMOS operational amplifiers often suffer serious drawbacks in the areas of swing and frequency response. Source followers, are commonly used as output devices; while they perform well in the area of frequency response, they limit the output swing dramatically, especially in processes with a large body effect coefficient (γ). CMOS inverter stages, often seen on standalone CMOS operational amplifiers, swing from rail to rail, but due to the Miller effect, require substantial compensation to insure stability. This results in a rather poor frequency response [37]. In video applications, the phase shift due to the operational amplifier will result in color distortion, color hue is encoded as the phase of the chrominance signal. Thus, the addition of a pole below 100 Mhz is unacceptable.

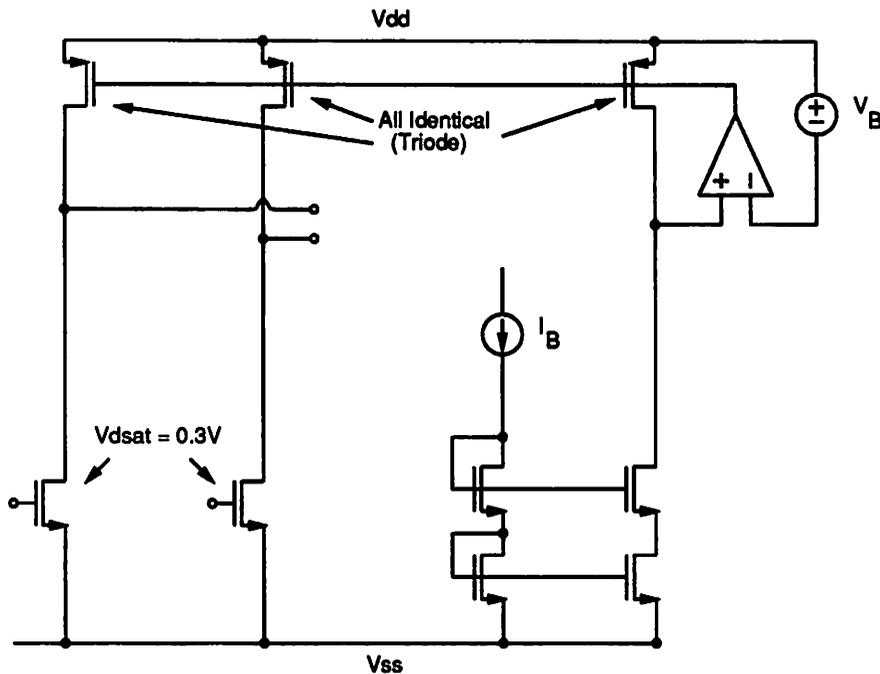


Figure 5.3: Class A output stage with triode region PMOS loads.

The class A output stage shown in Figure 5.3, combines the ability of the inverter stage to swing near the supply rails, while preserving the high bandwidth of source followers. The circuit makes use of triode region PMOS devices acting as the load to a NMOS common source stage. Biasing of the PMOS loads is through the use of a replica biasing stage [38]. The incremental resistance of the PMOS devices is low, thus keeping the gain of the stage low, preventing degradation of the frequency response due to the Miller effect. Simulations of the stage show a

frequency response of greater than 450 Mhz. A schematic of the entire operational amplifier with the output stage is shown in Figure 5.3a.

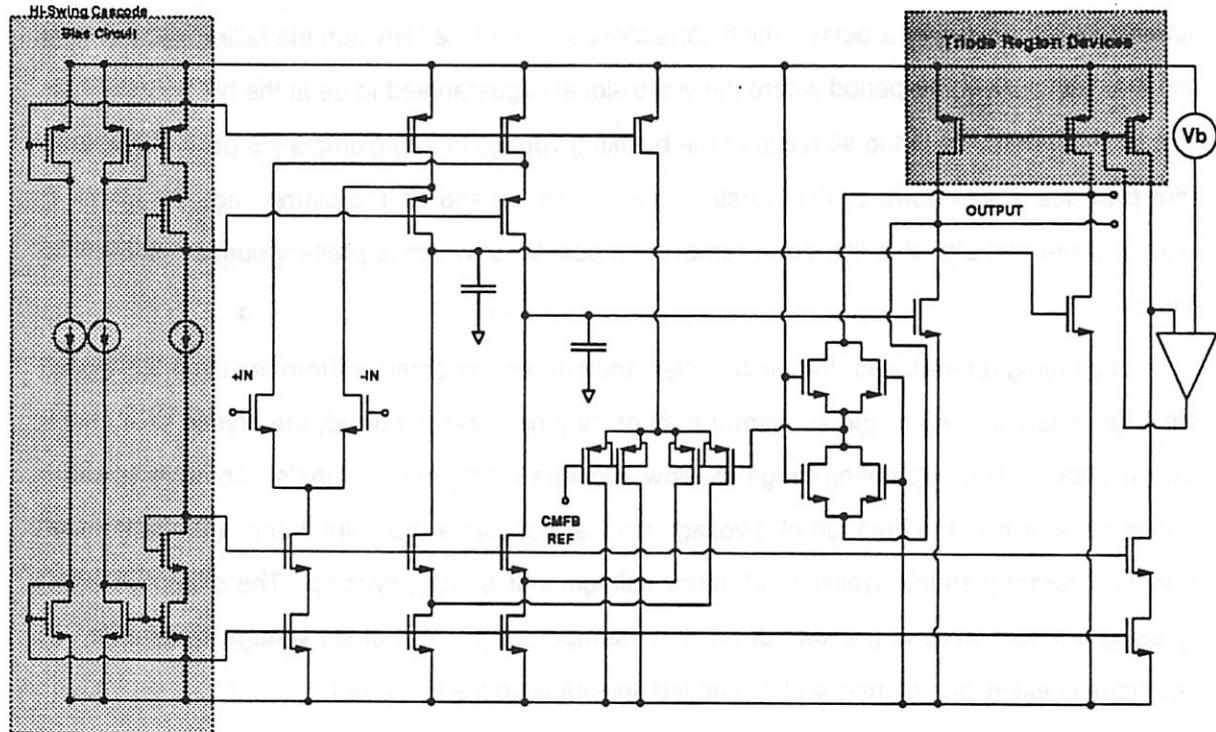


Figure 5.3a: Schematic of Buffer Amplifier

5.2.2 Operation of the Clamping Circuit

One output of the amplifier is sampled by a simple operational amplifier which compares the output with a *black level reference* voltage. This voltage corresponds to the signal voltage which represents the blanking level within the chip, currently set at ground. During the blanking interval, a clamp flag is activated which closes a feedback loop. Any error in the output of the main amplifier causes a voltage to appear at the input of a transconductance stage, which then feeds an proportional error current into the summing node thereby correcting the error. A holding capacitor is provided to preserve the correction current during periods other than blanking. The common mode voltage of the signal is adjusted internally by a common mode feedback circuit internal to the op-amp.

The clamp flag used above is generated by the coarse sync strip circuit. An active diode (diode in an op-amp feedback loop), and a capacitor are used to form a peak hold circuit. An op-

amp with a "built-in" offset (see below) is used to detect the falling edge of sync. To prevent active video information from falsely triggering the clamp flag, the circuit is gated to allow triggering only once per line period. Following the detection of the falling edge of sync, a digital delay is activated which provides a delay, which corresponds to the time between the falling edge of sync and the beginning of the period where the video signal is guaranteed to be at the blanking voltage. This period where the video signal is at the blanking voltage is approximately 5 μ s; a digital one-shot provides a waveform of this duration which then is used as the clamp flag. As stated in chapter three, the signal is filtered to remove the colorburst which is present during the blanking period.

The timing reference for the digital delay and one-shot is obtained from the clock generator. Although actual locking to the subcarrier may or may not have occurred, the crystal VCO has a narrow enough of an operating range to allow accurate timing of the signals. An op-amp with a built in offset allows the creation of a voltage comparator with a trip point that is a predetermined amount different than the available reference voltage, in this case, sync tip. The offset is created by using unequal sizes in the load devices in the input stage. The offset voltage created by this technique is easily determined, well controlled and stable to the first order.

The differential signal at the output of this sub-block is intended to drive the subsequent circuitry. The DC value of the signal is controlled to insure that saturation of circuits does not occur.

5.3 Sync Processing Circuits

The coarse sync stripping circuitry is integrated with the DC restoration subsection, leaving only the fine sync strip function to perform. As this chip needs only lock to subcarrier rather than to H, the only function of this block is to determine the start of every line by detecting the 50 percent point of sync. This is done by storing the blanking level and the sync tip level on two capacitors. The voltages are then averaged and used as a reference for a gated comparator.

As with the coarse sync strip, gating is necessary to prevent false triggering by active video. The gate used for this comparator can be the same as that used for the coarse sync strip, thus reducing circuit complexity even further.

An additional output required from this block is a burst flag — the chroma demodulator and sampling phase adjustment circuitry need to know what period of the signal corresponds to the colorburst. This flag is generated by utilizing another digital one-shot to provide a delay between

the falling edge of sync and the beginning of the colorburst period.

5.4 Phase Locked Loop and Clock Generators

Successful sampling of the video signal for subsequent processing requires that the sampling operation occur at accurate time intervals. Achieving this goal requires a stable oscillator and circuitry to generate the multitude of clocks that are needed to perform the sampling and filtering operations within the chip.

The most important output signals from this block are the clocks used to drive the sampling gates in the sample and hold circuit. These signals ultimately determine the overall sampling rate and processing speed of the chip. A judicious choice of the sampling rate will prove to greatly simplify the signal processing while minimizing speed requirements of the actual circuits.

5.4.1 Choice of Sampling Rate

The choice of the sampling rate to be used in this chip is somewhat arbitrary as long as it is above the Nyquist rate of the signal of interest. NTSC video signals are bandlimited to 4.2 Mhz, although it is common to extend the response out to 5.0 Mhz. Therefore, any sampling rate over 10 Ms/s will satisfy the Nyquist sampling theorem. However, it is found that sampling at an integral multiple of the subcarrier frequency minimizes the introduction of artifacts into the signal, especially those affecting the chrominance signal. This is especially true when the sampled signal is eventually quantized, as the quantization tends to have an adverse effect on the color subcarrier signal if the sampling rate is not an integral multiple of f_{sc} [39].

The lowest multiple of the subcarrier which is above 10 Ms/s is three times the subcarrier frequency, or $3f_{sc}$. This corresponds to a sampling rate of approximately 10.74 Ms/s. While this sampling rate is an integral multiple of the subcarrier and satisfies the Nyquist criterion, it suffers from what is known as non-vertically aligned pixels. Given that there are exactly $455/2$ cycles of subcarrier per line, sampling at three times the subcarrier frequency results in a non-integral number of samples per line. As a result, the samples from one line to the next will not line up vertically in space.

While the original image can be reconstructed from such samples, certain operations, such as comb filtering, are made substantially more complex by the lack of vertical alignment. This drawback can be eliminated by a process known as Phase Alternation Line Encoding (PALE) [40],

where the sampling phase is inverted each line. However, the circuitry necessary to handle this, and undo the inversion after signal processing makes this solution undesirable.

The use of a sampling rate corresponding to the $4f_{sc}$, or approximately 14.32 Ms/s, provides vertically aligned samples as a result of producing exactly 910 samples per line. Moreover, $4f_{sc}$ sampling provides an added advantage in the chroma demodulation process. Demodulation of the chrominance signal involves multiplying the chrominance signal by sine and cosine signals at the subcarrier frequency. Traditionally, performed in the continuous time domain via a four quadrant multiplier and an oscillator locked to the subcarrier frequency, sampled data processing utilizes a sampled data representation of the sinusoid and cosinusoid at the subcarrier frequency. In a digital signal processing environment, generation of sampled sinusoids is usually accomplished via a cyclic counter and a look up ROM. However, in the analog sampled domain, such methods are not feasible. Applying a sample and hold to the output of an oscillator would generate the required signal, however, the question of how to implement an accurate analog multiplier in the sampled data domain still remains.

However, given that the sampling phase is adjusted properly, a sinusoid sampled at four times its frequency results in the periodic set of samples $\{0, 1, 0, -1, 0, \dots\}$, while the cosine signal would be represented by the set $\{1, 0, -1, 0, 1, \dots\}$. Thus, under the proper conditions, a sampled sinusoid and the accompanying analog multiplier are no longer necessary. All that is required is a circuit to perform sign inversion, which in a differential topology is merely a cross coupling of the signals at the appropriate intervals. Thus, a chroma demodulator can be constructed using the a circuit similar to that shown in Figure 5.4 operating at a sampling rate of $4 f_{sc}$.

Successful demodulation of a QAM signal requires that the phase of the demodulator oscillator be locked to that of the oscillator of the modulator. Usually, the phase of the local demodulating oscillator is varied to match the phase of the modulating oscillator. However, the method of chroma demodulation described above is equivalent to using a demodulating oscillator whose phase is fixed by virtue of the fixed samples used to represent the demodulating sinusoids. Thus, adjustment of the subcarrier phase must be done when the video signal is sampled by varying the instant the samples are taken. Therefore, adjustment of the sampling phase is critical for proper operation of the chroma demodulation circuit. A novel feedback loop has been designed to insure that the sampling phase is accurate and will be discussed in the clock generator section of this chapter.

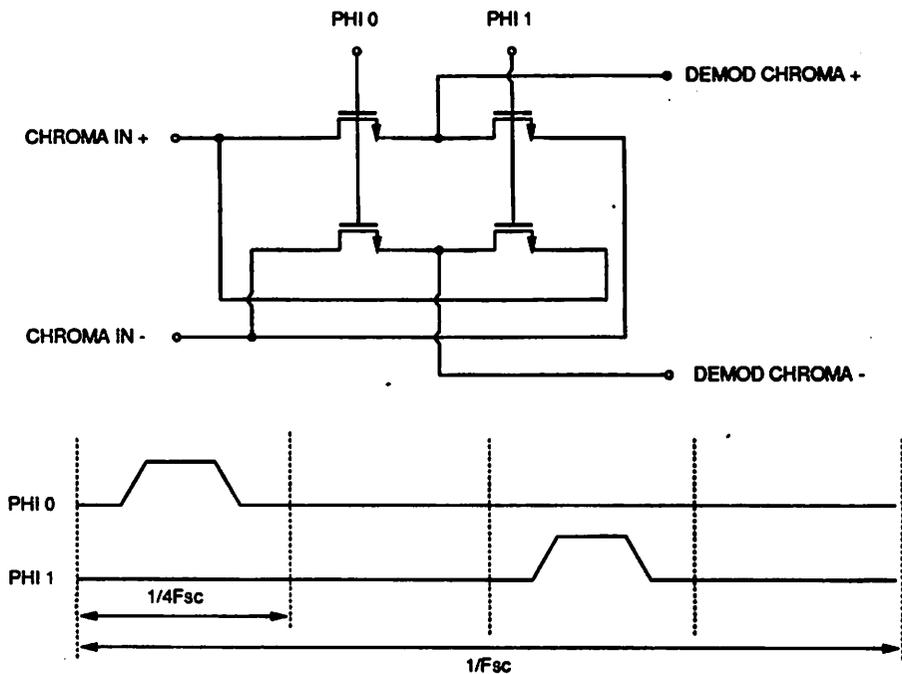


Figure 5.4: $4 f_{sc}$ Chroma Demodulation (Conceptual Diagram)

5.4.2 Anti-Aliasing and Oversampling

While the $4f_{sc}$ sampling rate provides vertically aligned samples and provides a particularly easy method of chroma demodulation, a sampling rate of 14.32 Ms/s requires that an anti-alias (AA) filter be placed in the signal path prior to sampling which removes all extraneous energy above 7.16 Mhz to avoid aliasing distortion. Video signals are in general well behaved, and the amount of energy in the signal above 5 Mhz is typically only a small fraction of the total signal energy. However, an attenuation of 25 dB or more is needed starting at 7.16 Mhz [41]. Thus, a lowpass filter structure which is flat from DC to 5 Mhz, while providing at least 25 dB of attenuation at 7.16 Mhz is required. Moreover, this structure must have good group delay characteristics, as delay distortion produces visible ringing in images. The combination of a relatively narrow transition band and the requirement of linear phase result in a very complex filter structure for the AA filter. Typically, this filter will be a fifth order elliptic filter followed by two second order allpass delay equalization stages. This filter is relatively expensive to manufacture, as each filter requires hand alignment to adjust the response due to component variations. Moreover, the individual filter components, in particular the inductors, are relatively expensive. Therefore, a method of relaxing the requirements of the input AA filter would result in significant reduction in the cost and

complexity of the filter.

By using oversampling techniques at the front end, and subsequently reducing the sampling rate to $4f_{sc}$ using decimation filters, the advantages of $4f_{sc}$ sampling can be maintained while shifting the burden of anti-alias filtering from the continuous time filter in front of the chip to an on-chip filter. The optimum oversampling ratio is a function of the inherent circuit speeds attainable and the complexity of the decimation filters. On the other hand, the higher the oversampling ratio, the less stringent the requirements on the input AA filter. For current $1\ \mu\text{m}$ analog CMOS circuits, sampling rates are limited to about 100 Ms/s. Furthermore, the size and complexity of the decimation filters increase in proportion to the oversampling ratio. Oversampling by a factor of 4 provides a substantial relaxation in the AA filter requirements, while at the same time results in a manageable sampling rate of 57.27 Ms/s.

5.4.3 Monolithic MOS Crystal Voltage Controlled Oscillator

Generation of the sampling clocks to perform the oversampling requires an oscillator that is locked to the subcarrier, and a frequency multiplier, along with a multi-phase clock generator. The sampling scheme chosen for this chip (see below), utilizes a six-phase clock with a 50 percent duty cycle. In addition, a three-phase clock at eight times the subcarrier frequency is used by the first integrator for the filtering function.

As the subcarrier frequency is very well controlled in most all video systems, a crystal oscillator with a nominal frequency of four times the subcarrier frequency is used as a system time-base. However, as QAM demodulation requires exact phase locking to the subcarrier, a method of altering the frequency of the oscillator is still necessary.

Crystal oscillators operate on the same principle as most other oscillators, by utilizing a feedback path with a loop gain greater than unity, which has zero phase shift at the frequency of oscillation. A very common oscillator configuration is based on the Pierce oscillator, where the crystal is placed across the terminals of an inverting gain stage [42]. Quartz crystals can be modeled by an equivalent circuit shown in Figure 5.5(a), and have an impedance curve similar to that shown in Figure 5.5(b). The two frequencies of resonance are commonly denoted ω_s and ω_p for series and parallel resonance respectively. The separation of ω_s and ω_p is on the order of 0.5 percent. Over this range of frequencies, the crystal appears as a frequency dependent inductance.

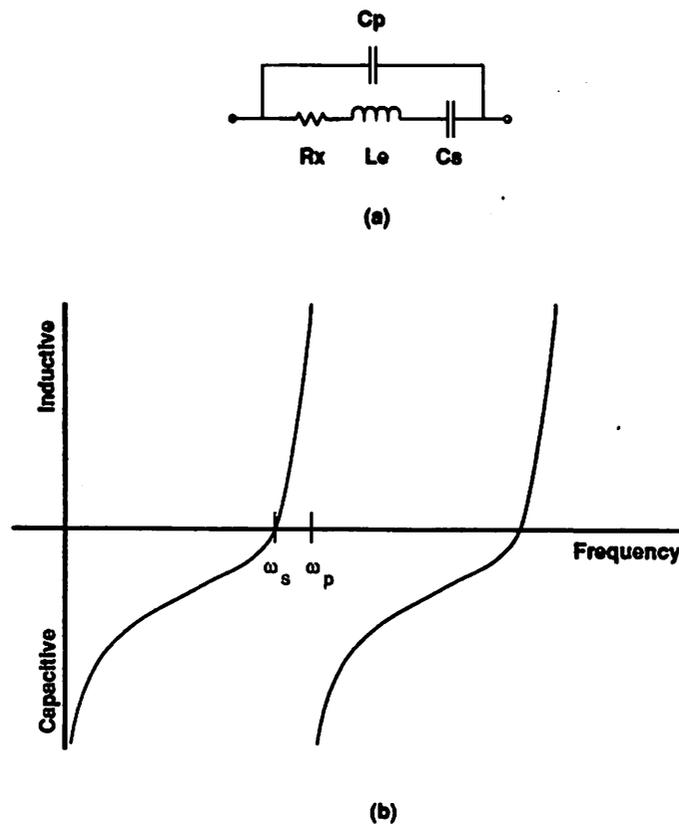


Figure 5.5: (a) Equivalent circuit of a quartz crystal.
 (b) Impedance of quartz crystal vs. frequency

Pierce crystal oscillators operate in the parallel resonance mode, where the crystal acts as an inductive element. The actual frequency of operation is determined by the load capacitance seen by the crystal, as oscillation occurs when the load capacitance and the synthesized inductance resonate to produce a 180 degree phase shift (parallel resonance). The magnitude of inductance required is a function of the load capacitance — correspondingly, as the value of the synthesized inductance is a strong function of the frequency, the frequency of oscillation is determined by the load capacitance.

Thus, a first approach to design a crystal VCO could make use of a voltage dependent capacitance, similar to a junction capacitance. Specialized devices known as varactor diodes are manufactured specifically for purposes of tuning oscillators. Incorporation of a varactor diode as a load element as shown in Figure 5.6, produces a well behaved crystal oscillator. However, this circuit contains several components (highlighted in gray on the figure), that are not easily included on a CMOS integrated circuit, such as the crystal, inductors, and the actual varactor itself.

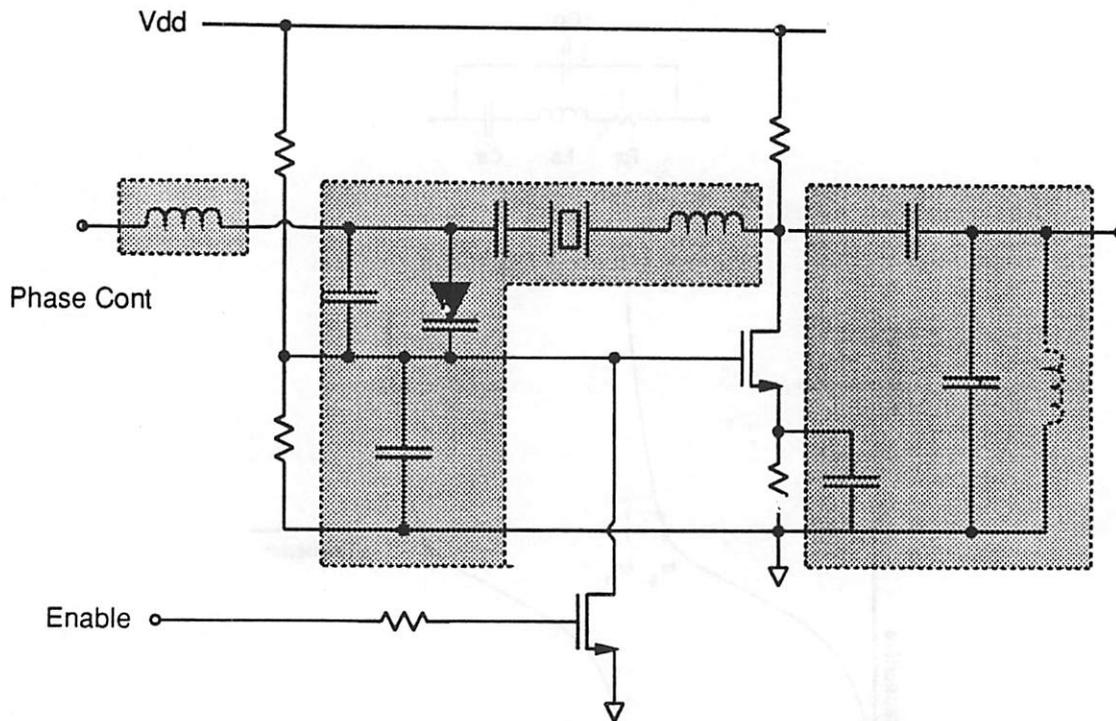


Figure 5.6: Crystal VCO using a varactor as a tuning element

Thus, a different type of oscillator has been designed, attempting to minimize the number of external components, ideally requiring only the crystal as an external component. The circuit shown in Figure 5.7 accomplishes this task. This circuit utilizes the crystal in the series resonant mode as will be seen. Operation of this circuit is best analyzed by breaking the loop at the gate of M2, which functions as an inverting gain stage. The drain current of M2 represents a signal which is 180 degrees out of phase with the signal at the gate. The high impedance node at the drain of M2 coupled with the capacitor produces a voltage that is nearly 90 degrees out of phase with the current. A transconductance (G_m) stage is then used to reconvert the voltage into a current. The drain current of M1, which forms the other half of a source coupled pair, is out of phase with I_{d2} , or in phase with the gate of M2. I_{d1} is fed into a variable gain stage made of another source coupled pair, M3 and M4. By varying the voltage across the gates of M3 and M4, I_{d1} can be shunted in varying proportions into I_{d3} and I_{d4} . The portion flowing in M4 is summed with the current flowing from the G_m stage. As the signals are sinusoidal and of the same frequency, the resulting voltage across the load resistor, R1, is simply a vector phase addition of I_{d4} and the phase shifted I_{d1} . Therefore, the voltage seen at this summing node will have a phase which varies with the voltage applied across the gates of M3 and M4.

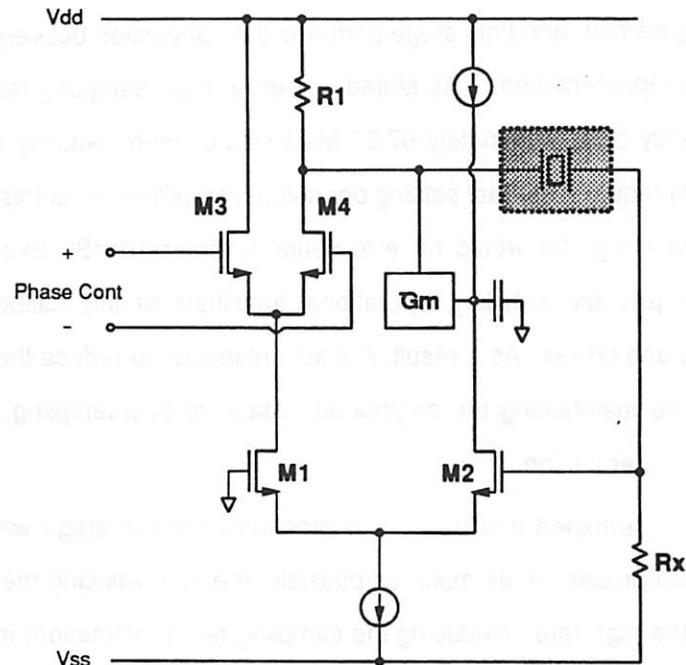


Figure 5.7: Crystal Voltage Controlled Oscillator

The summing node voltage is therefore nominally in phase with the gate of M2. Thus, to complete the feedback loop with zero phase, the crystal must operate under this condition at the series resonant frequency. Altering the voltage on the gates of M3 and M4 will change the loop phase from the gate of M2 to the drain of M4. Thus, the frequency of oscillation must move to allow the crystal to compensate for the phase change.

Note that this circuit achieves the goal of having no external components except the quartz crystal itself. Moreover, since the crystal is being used in the series resonant mode, it is relatively insensitive to stray capacitances that may be seen at the leads of the crystal. The tuning range afforded by this circuit is fairly narrow, as the phase characteristics of a crystal change rapidly over a very narrow frequency range. The range of operation can be widened at the expense of frequency stability by increasing the value of the crystal load resistor, R_x . This has the effect of decreasing the resonant Q of the crystal and lowering the sensitivity of the phase function with respect to frequency.

5.5 Oversampling S/H Stage with Integral FIR Decimation Filter

The oversampling sample and hold stage performs the conversion between continuous time and discrete time data representations. As stated earlier, a initial sampling rate of 16 times the color subcarrier frequency or approximately 57.27 Ms/s was chosen. Analog sampled data processing at these speeds require very fast settling operational amplifiers — at this data rate, a standard switched capacitor integrator would have to settle in under 10 nS. Even with the use of scaled technologies (1 μm and smaller), operational amplifiers of this caliber consume large amounts of silicon area and power. As a result, it is advantageous to reduce the sampling rate as quickly as possible while maintaining the original advantage of oversampling, that of a relaxed requirement in anti-alias filter design.

Therefore, it was determined that the proper procedure for this stage was to combine the sampling and decimation processes as much as possible, thereby reducing the number of circuit elements operating at the high rate. Reducing the sampling rate (decimation) in general requires that the actual downsampling operation be preceded by a low pass filtering operation to avoid aliasing distortion. The main advantage here is that this filtering can be done in the sampled data domain, rather than in continuous time as would have been the case without oversampling.

The design of decimation filters necessitates a choice of filter topology. Finite Impulse Response (FIR) filters are attractive as they provide linear phase, an important criteria for video. However, FIR filters usually require a higher order structure to meet a certain rolloff characteristic compared with the Infinite Impulse Response (IIR) and related filters. Moreover, the length or order of the FIR filter necessary for a certain frequency response grows linearly with the decimation ratio. Simulations showed that a 4:1 decimation FIR filter with at least 25 dB out of band rejection requires a filter of at least 50 taps. This would require far too many memory elements and multipliers than would be practical in an analog signal processor of this type. Thus a two-stage filter approach was taken.

5.5.1 First Stage Filter Design

The first stage of the decimation filter reduces the sampling rate by a factor of two from $16 f_{sc}$ to $8 f_{sc}$, while providing adequate anti-aliasing in the critical frequency bands. By reducing the sampling rate to $8 f_{sc}$, signals in the frequency band of $8 f_{sc} \pm 5 \text{ Mhz}$, will be aliased back into the baseband. Thus, prior to decimation, a 3 tap FIR structure is used to nominally place a double

zero at $4f_{sc}$. This theoretically results in 23 dB of rejection at the 5 Mhz bandedge and 25.6 dB of rejection at the 4.2 Mhz NTSC limit. By virtue of the FIR structure, this filter contributes no delay distortion. Figure 5.8 shows the computed frequency response of this filter.

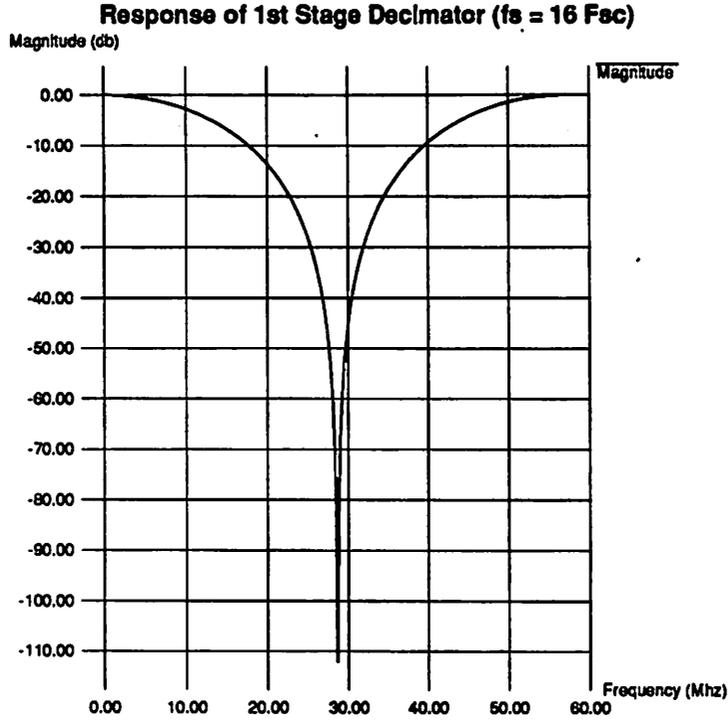


Figure 5.8: Computed frequency response of a double zero filter.

5.5.2 Circuit Implementation of the First Stage Filter

Implementation of this filter in MOS technology is accomplished via a modified double sampling technique (Fig. 5.9). This structure is a transversal FIR filter, which accomplishes its function by utilizing capacitors as both its memory elements and multiplier stages. To help understand the equivalence of this transversal MOS circuit and the traditional delay line approach of representing FIR filters, consider the circuits shown in Figure 5.10. The upper circuit is the conventional delay line (shift register) FIR structure with multipliers. The transfer function is given by

$$H(z) = a + bz^{-D} + cz^{-2D} + dz^{-3D} + ez^{-4D}$$

In the lower structure, the multi-phase clock takes samples on capacitors *a* through *e* at time intervals of *D*. The gain from each sampling capacitor to the output is given by the capacitance ratio of the sampling capacitor to the integrating capacitor, *X*. At the end of the fifth sampling operation,

the charge stored in the five capacitors are dumped into the integrating capacitor, forming the weighted sum. This results in the same transfer function as the circuit above. Note however, that in order to realize an output sample for each clock interval, five parallel stages are needed. As a result, long length FIR filters quickly result in very complex structures.

In the circuit of Figure 5.9, a six phase clock is used to operate the sampling gates, *a - e*, which take samples at intervals corresponding to the 57.27 Ms/s rate (17.46 nS). As the filter length is three, and the decimation ratio is two, every other incoming sample will be a constituent of two output samples (Fig. 5.11). As a result, during every other sampling interval, the incoming signal is stored in two separate memory locations. A three phase clock at half the rate is used to integrate the accumulated charges. The arrangement of the clock phases allows the amplifier to integrate during all phases of the half rate clock; decimating by a factor of two reduces the number of output samples by one-half. This effectively increases the allowable settling time to 34.9 nS. A high speed op-amp using 1 μ m technology is expected to meet this requirement.

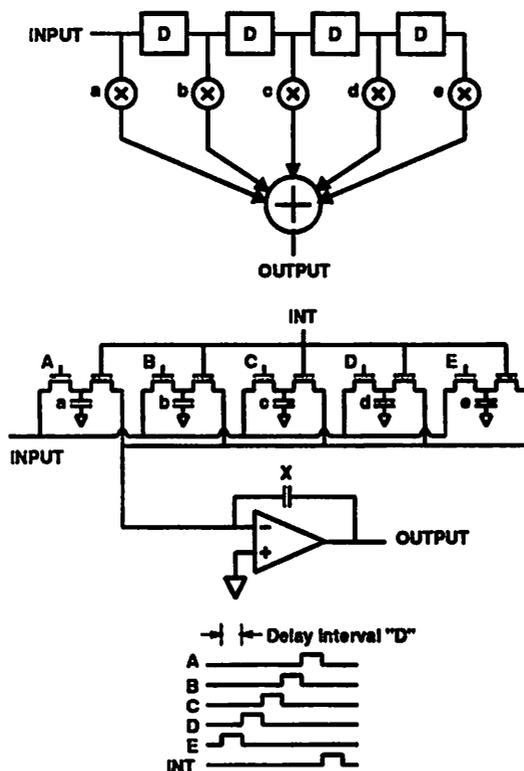


Figure 5.10: Equivalence between delay line and transversal FIR filters

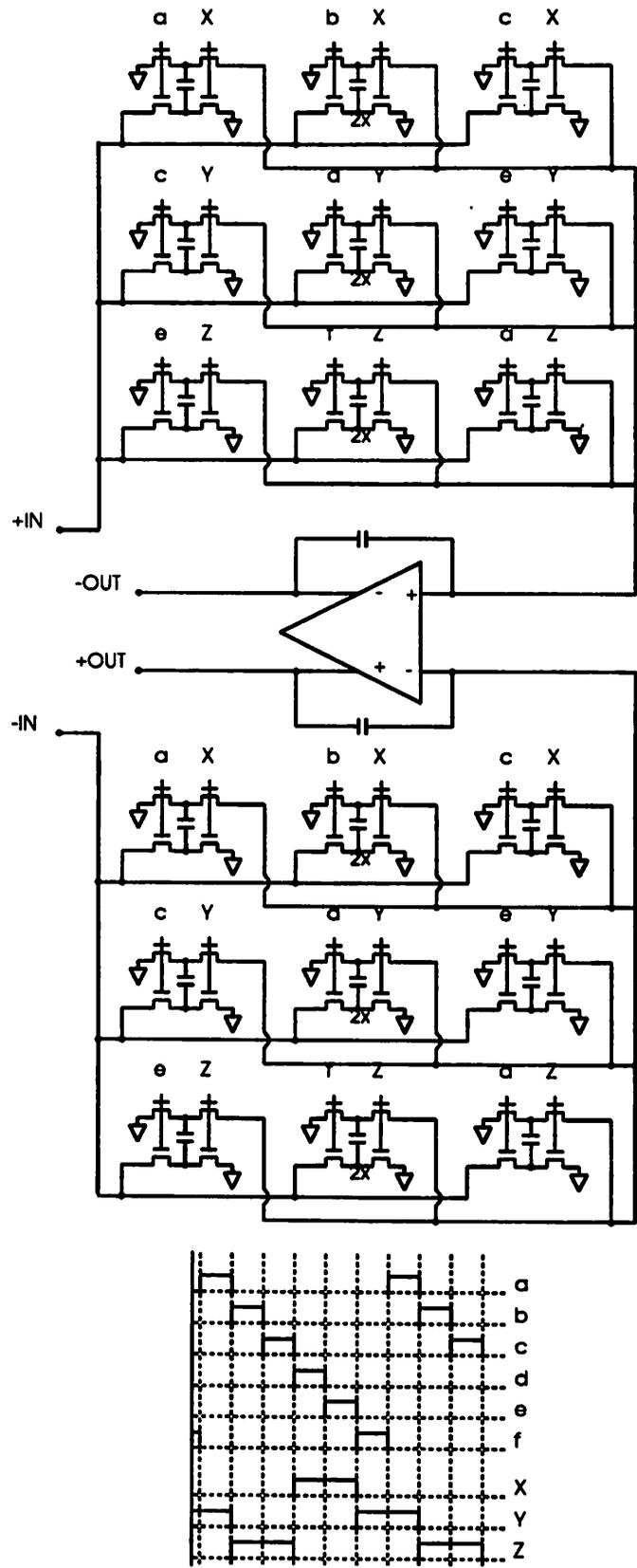


Figure 5.9: Input sampling stage and first decimation filter

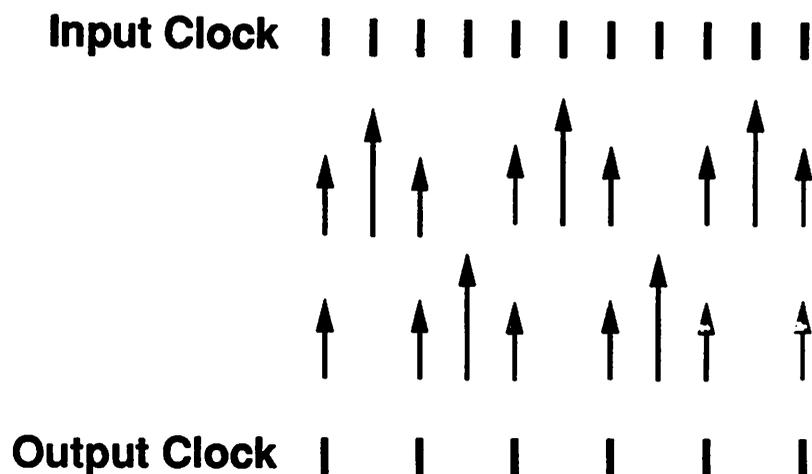


Figure 5.11: Time relationship between FIR impulse response and input/output clocks.

5.5.3 Effects of Circuit Non-Idealities on Filter Performance

The frequency response of the first stage filter is determined by the effective ratios of the sampling capacitors to the integrating capacitors. In any integrated circuit process, some variations in capacitance between identical structures is to be expected. The magnitude of matching errors in capacitors has been studied to some extent, as it is a major criteria in switched capacitor filter design [43]. The general trend is that larger capacitors tend to match better, with the variance of the mismatch being roughly proportional to the inverse of the square root of the capacitance. Shyu reports mismatches of about 0.75% with 300 fF capacitors. As the dynamic range of video signals is relatively narrow, the noise from kT/C effects are not a concern in determining capacitor sizes. Moreover, the sampling mode bandwidth and area of the filter are determined in large part by the unit capacitance value due to the large number of capacitors used in this structure. In order to determine the maximum allowable tolerance in capacitor matching, the frequency response of a three tap FIR filter will be examined.

As stated earlier, the goal of this first stage filter is to lowpass filter the incoming signal prior to decimation. Of particular importance is adequate rejection (≈ 25 dB or more) in the zone of $8 f_{sc} \pm 5$ Mhz, or from approximately 23.6 to 33.6 Mhz, as signals in this band will alias directly into the signal baseband after decimation. Minimal disturbance to the signal passband by this filter is desirable; however, slight degradation in the passband response can be compensated by the second stage filter. A first order system has been found to have inadequate attenuation at the

edges of the alias band; a second order system placing a double zero at half the sampling rate fulfills the frequency response requirements for this application (Fig. 5.12 (a)).

Consider a generalized three tap FIR filter with tap coefficients a , b , and c . The resulting z -transform transfer function is given as:

$$H(z) = A + Bz^{-1} + Cz^{-2}$$

The magnitude squared response of such a filter is given by:

$$|H(e^{j\omega T})|^2 = [A + B \cos \omega T + C \cos 2\omega T]^2 + [B \sin \omega T + C \sin 2\omega T]^2$$

The zeros of such a filter will be located at:

$$\frac{-B \pm \sqrt{B^2 - 4AC}}{2A}$$

(1) Consider first the case when the argument inside the square root is greater than zero. Then the zeros lie on the real axis, on either side of the point $(-1, 0)$ in the z -plane at a distance determined by the argument inside the square root (Fig. 5.12 (b)). If the goal is to place a double zero at the frequency corresponding to half the sampling rate prior to decimation, then minimizing $B^2 - 4AC$ is critical. The tap weights for this type of filter nominally have a ratio of $1 : 2 : 1$ ($A : B : C$); thus, if the product AC is less than one, real axis zeros will occur. This is especially undesirable as the frequency response will contain no zeros. A one percent reduction in the values of A and C with respect to B causes a reduction in the notch depth from nearly infinite to -46 dB, with a corresponding worsening of the response at other frequencies.

(2) If the argument inside the square root is zero, then a perfect double zero "cosine" filter results (Fig. 5.12 (a)).

(3) If the argument inside the square root is greater than zero, then complex zeros will result. The zeros are then given by:

$$\text{Re}\{z\} = \frac{-B}{2A}$$

$$\text{Im}\{z\} = \pm \frac{\sqrt{4AC - B^2}}{2A}$$

Two sub-conditions exist. If a and c are essentially equal, then the zeros will lie on the unit circle, as the sum of the squares of the real and imaginary parts of the zeros equals one

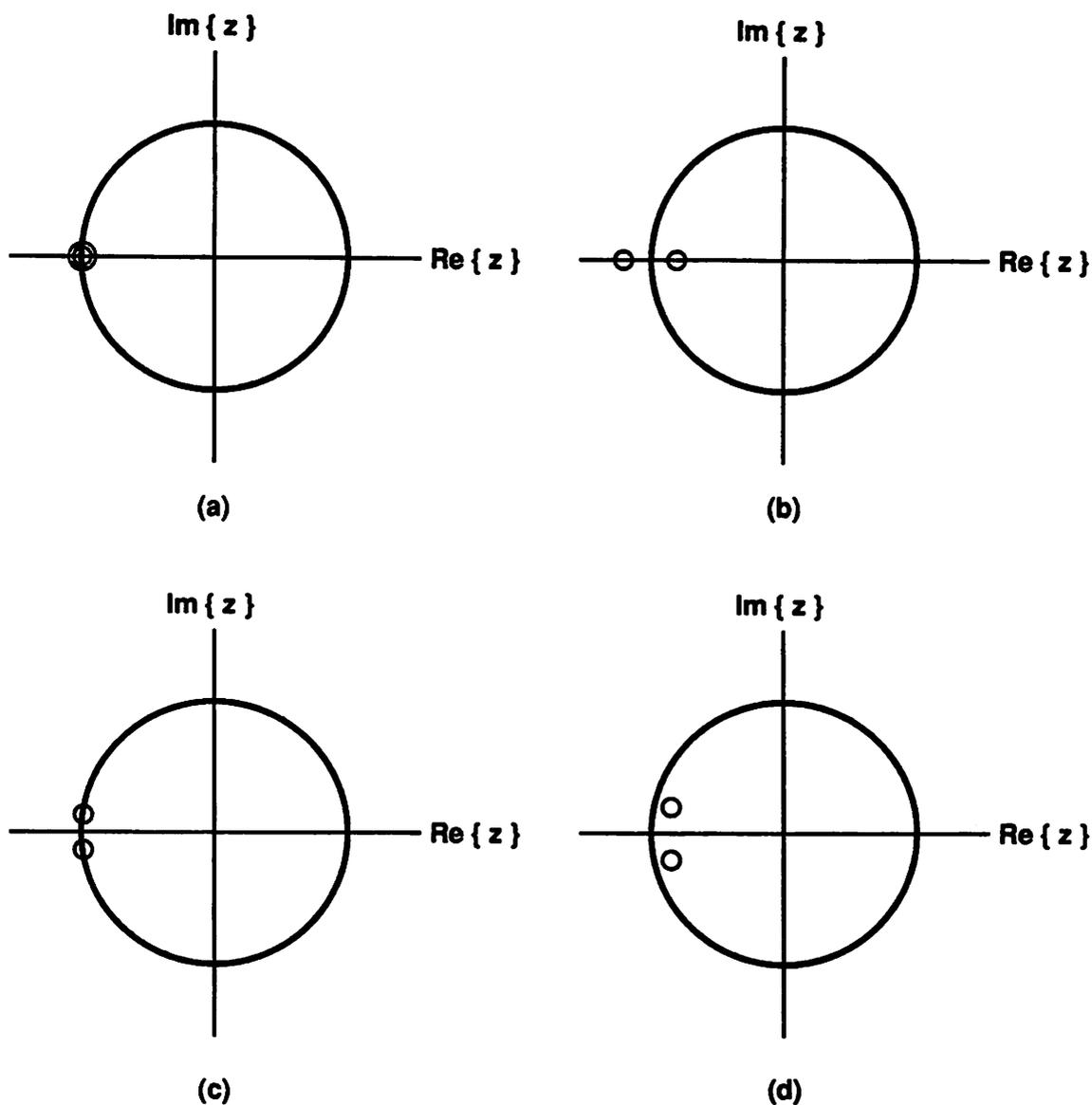


Figure 5.12: (a) Ideal Double Zero at $(-1, 0)$, (b) Real Axis Zeros, (c) Complex Zeros on Unit Circle, (d) Complex Zeros off Unit Circle

(Fig. 5.12 (c)). Otherwise, the complex zeros will lie off the unit circle at a distance proportional to the ratio of C to A (Fig. 5.12 (d)). Complex zeros lying on the unit circle are actually advantageous, as it gives two nulls in the response of the filter, spread slightly apart in frequency, with a zone of high attenuation in between. This results in a higher degree of attenuation at the alias band edge than achieved with condition (2) above. The same holds true to some extent with complex zeros off the unit circle, with the caveat that the nulls will no longer be true zeros. However, the effect of mismatches is lessened, as the error is

distributed between the real and imaginary components of the zeros. A one percent error results in nulls of about -63 dB, compared with -46 dB as was the case with real zeros (Fig. 5.13).

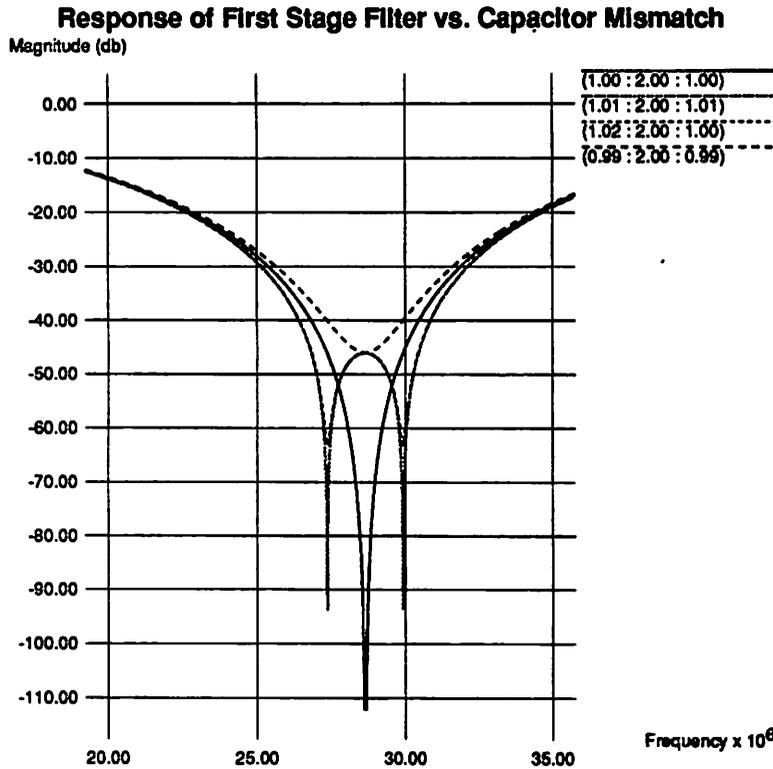


Figure 5.13: Frequency response of first stage filter near alias band as a function of capacitor mismatches.

The conclusion of this analysis is that complex zeros are very much preferable to real zeros. As this is the case, the design of the filter will use tap ratios of (1.01 : 2 : 1.01) to increase the likelihood of complex zeros and minimize the chance of real zeros occurring.

5.6 Second Stage Decimator and Filter

The first stage filter is designed to interface directly into the second decimation filter. (Note the lack of an integrator dump function on the first amplifier). The role of the second filter is to provide a lowpass function rolling off all signals above 5 Mhz, and to provide at least 25 dB of attenuation at 7.16 Mhz. Concurrently, this filter will reduce the sampling rate from $8 f_{sc}$ to the final $4 f_{sc}$ signal processing rate. Furthermore, the frequency response of this filter should be pre-

distorted to compensate for the slight attenuation in the passband caused by the first decimation filter.

Investigation into possible filter structures yields two potential architectures. The first is a FIR structure similar to the first stage, but with many more taps. Simulations reveal that a minimum of 21 taps are necessary to achieve the proper frequency response characteristics. A filter structure of this type would require a large number of sampling capacitors and a very complicated switching structure. Thus, it is not as desirable as the second solution, a switched capacitor analog of a continuous time elliptic ladder filter. Elliptic filters afford very sharp transition bands with a minimum number of filter sections. A fifth order filter has been shown to provide a more than adequate magnitude response. However, elliptic filters exhibit very poor group delay characteristics, especially near the passband edge. Thus, a series of delay equalizers are necessary to correct this fault. In order to adequately compensate for the delay distortion introduced by the magnitude section of this filter, two second order delay equalizers are used to restore an adequate step response. As an example, a fifth order elliptic filter is shown in Figure 5.14(a). Note that a doubly terminated design is used, which has the property of minimizing the sensitivity of the transfer function with respect to component variations [44]. The magnitude response is shown in Figure 5.14(b), with the phase response shown in Figure 5.14(c). The presence of a transmission zero at 7.355 Mhz insures a sharp rolloff through the transition band. The phase plot shows that the non-equalized phase response is decidedly non-linear, which would result in a the poor step response of the filter. However, the addition of two second order allpass sections with poles at $-2.416 \times 10^6 \pm 1.243 \times 10^6 j$ and $-2.145 \times 10^6 \pm 3.712 \times 10^6 j$ results in the equalized phase response, which is nearly linear over the passband. Finally, the slope of the phase vs. frequency curve is much greater, indicating that the overall group delay of the filter has increased — however, this is not a concern in this system.

The switched capacitor realization of a fifth order elliptic filter has been demonstrated on numerous occasions. The two allpass sections, however, require a specialized form of a biquad circuit topology to realize zeros outside the unit circle. Topologies that accomplish this task have been reported by several sources [45, 46]. Thus, by cascading these two sections, a suitably delay corrected fifth order decimation filter can be constructed.

The design of a switched capacitor filter from a continuous time equivalent ladder network has been documented by several sources. It is commonly agreed that the type of digital integrator used to perform the 1/s function strongly determines the characteristic of the filter. The most

The filter in this chip is designed to have a cutoff frequency that is a fairly large percentage of the clock rate. Thus, the shortcomings of the LDI mapping are expected to degrade the performance of the filter considerably. As a result, an alternative method of performing discrete time integration corresponding to the trapezoidal approximation, known as the bilinear transform, is proposed for use in this chip. Historically, bilinear switched capacitor architectures have been plagued by sensitivity to parasitics within the circuit. However, recent work has demonstrated methods which implement the bilinear transform while maintaining the parasitic insensitivity that is associated with LDI integrators [47, 48]. A generic circuit topology for a third order elliptic low-pass filter is shown in Figure 5.15. In the actual chip, the first integrator in the magnitude section of the filter would be common to the integrator in the first stage filter, thus reducing the op-amp count by one.

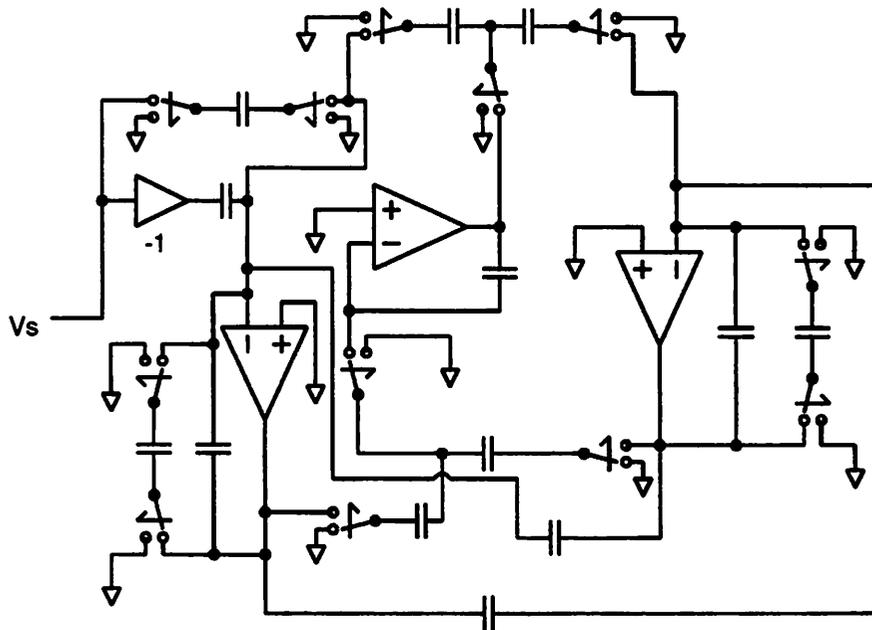


Figure 5.15: Generic Topology of Second Stage Decimator
(Single Ended Circuit Shown) (From [48])

5.7 Luminance/Chrominance Separator

The design goal of the luminance/chrominance separator for use in this chip was to achieve the highest quality separation possible while remaining within the bounds of a monolithic circuit in the analog sampled data domain. As was discussed in chapter three, there are a number of

different topologies for the Y/C separator, each with varying degrees of performance. Prior monolithic implementations of Y/C separators usually were of the bandpass type or recently, of the single line comb type. However, for this chip, an attempt is being made to implement an adaptive 2H comb filter for the highest performance with the aid of a new type of delay line structure.

The actual architecture of the comb filter follows very closely with that shown in Figure 3.5, as it was determined that the additional delay line introduces less complexity than would the extra bandpass filters. A rudimentary adaptive switching system is planned for this filter — although more elaborate schemes would increase the performance of this filter, the circuitry required to implement the complex switching algorithms would unnecessarily complicate the system. Moreover, the goal of this chip is to demonstrate novel circuit techniques, not to develop adaptive switching algorithms. Readers interested in adaptively switched comb filters are directed to work done specifically in that field.

5.7.1 Development of an Analog RAM Delay Line

As was alluded to in chapter three, the most important component in a comb filter is the delay line. As such, it is necessary to design a delay line structure that offers high performance while processing signals in the analog sampled data domain. CCDs have been used extensively for this application, but as stated earlier, they suffer from noise and require a specialized process to integrate them on an analog CMOS chip.

Capacitors formed on a CMOS chip are known to be about the closest realization of an ideal capacitor in terms of leakage current. Thus, a voltage or charge can be stored on a very small capacitance over a period of time with little loss in accuracy. Subsequently, a sampled data delay element can be formed by storing the signal on a capacitor, waiting the delay period and "reading" the signal out of the capacitor. This concept can be expanded using a set of these elements to form a delay line, with each incoming sample being stored on a separate capacitor. A predetermined number of clock cycles later, each capacitor would be "read" in succession, thus recovering the waveform stored earlier in the capacitor array. The individual capacitors, after being read, can be re-used to store current incoming samples. Thus, N capacitors are required to delay a signal for a period equal to N sampling intervals.

This method of performing a delay has an inherent advantage over CCD techniques. Rather than transferring a charge N times through N individual elements at the clock rate, this method

transfers a signal twice, once to place it on the storage capacitor, and once to read it back out, regardless of the length of the delay. This advantage is significant when the delay length is near 1000 clock cycles as is the case with a 1H delay line operating at $4f_{sc}$. Moreover, the noise introduced by this circuit is effectively determined by the kT/C noise introduced by the sampling operation and any noise added by the write/read circuit described below. For signals in the region of 1 V_{p-p} , 60 dB of dynamic range can be achieved with capacitors smaller than 1 pF.

5.7.2 Write/Read Circuit for Analog RAM Cells

The primary disadvantage of this method of producing a delay line is the complexity of the write/read circuit. Rather than the simple voltage to charge converter placed at the head of a CCD line, and the inverse converter at the tail of the CCD, this system requires an elaborate switching scheme to transfer the data into individual capacitors and read them back at appropriate intervals. As there will be a large number of capacitors spread out over a sizable silicon area, the issue of parasitics and their effect on circuit operation is important.

A choice exists whether to store the data samples as a charge or as a voltage. Although the incoming data is represented as set of voltage samples, converting them into a charge prior to storage provides several advantages over storing data as a voltage. First and foremost, storing data as a charge negates the effect of voltage non-linearity of the capacitors which would be a serious concern in MOS capacitors. This also gives the flexibility of using capacitors structures with a higher specific capacitance that have poorer linearity compared with polysilicon-polysilicon capacitors, for example. The second advantage is that the absolute size of the storage cell capacitors is not critical. Although for a fixed input signal, the stored voltage on the cells will differ with varying capacitances, as charge is the quantity of interest, the voltage differences will not contribute an error term to the signal.

As a result, storage of the signal as a charge is strongly preferable to storing voltages. Thus, the write circuit, in addition to switching the proper capacitor into place for each pixel, must convert the incoming voltage samples into a charge. The most straightforward method of doing this would be to sample the voltage on a linear capacitor, and then use an integrator to transfer the resultant charge on the storage capacitor. The inverse operation can be performed by using another integrator to transfer the charge on the storage capacitor onto a linear capacitor, thus reconverting the signal into a voltage for subsequent processing. This is the approach taken by this chip.

5.7.3 Implementation of the 2H Comb Filter

Implementation of the 2H comb filter using an analog RAM delay line described above takes form as shown in Figure 5.16. This circuit implements the 0.25 : 0.5 : 0.25 summing of the signals from the current line and the two previous lines. Clocking of the filter is provided by a two phase non-overlapping clock operating at $4f_{sc}$. The actual delay lines consist of hundreds of storage elements of the type shown as Elements A and B that are highlighted in gray.

Operation of the circuit can be traced by assuming that the array containing Element A is initialized with information from the previous line, and Element B is initialized with information from two lines before. During ϕ_0 , the input is sampled onto capacitors C1 and C2. At the same time, the charge stored in Element A is dumped into integrator U3 resulting in a voltage being sampled onto C3 and C4. Similarly, the charge stored in Element B is dumped into integrator U1, with the resultant voltage being sampled on C5. During ϕ_1 , the signals stored on C1, C4, and C5 are integrated by U2. As C1 and C5 are half unit size and the integrating capacitor is double size, the output of U2 represents the comb filtered output. At the same time, the signal stored in C2, which reflects the input from a half clock cycle before is stored on Element A using U5 as the integrator. The charge in C3, representing the signal that was in Element A is now transferred to Element B by U4. This cycle now repeats for the next pixel during ϕ_0 with the next storage elements in the storage arrays.

A review of the above operation shows that during the first clock phase, the stored signals and the current input are moved onto temporary memory locations, C1 through C5. The second clock phase computes the weighted sum, and also shifts the current input into Element A, and moves what was in Element A into Element B. Parasitic insensitivity is maintained by using bottom plate sampling techniques at all nodes, especially those involving the storage elements.

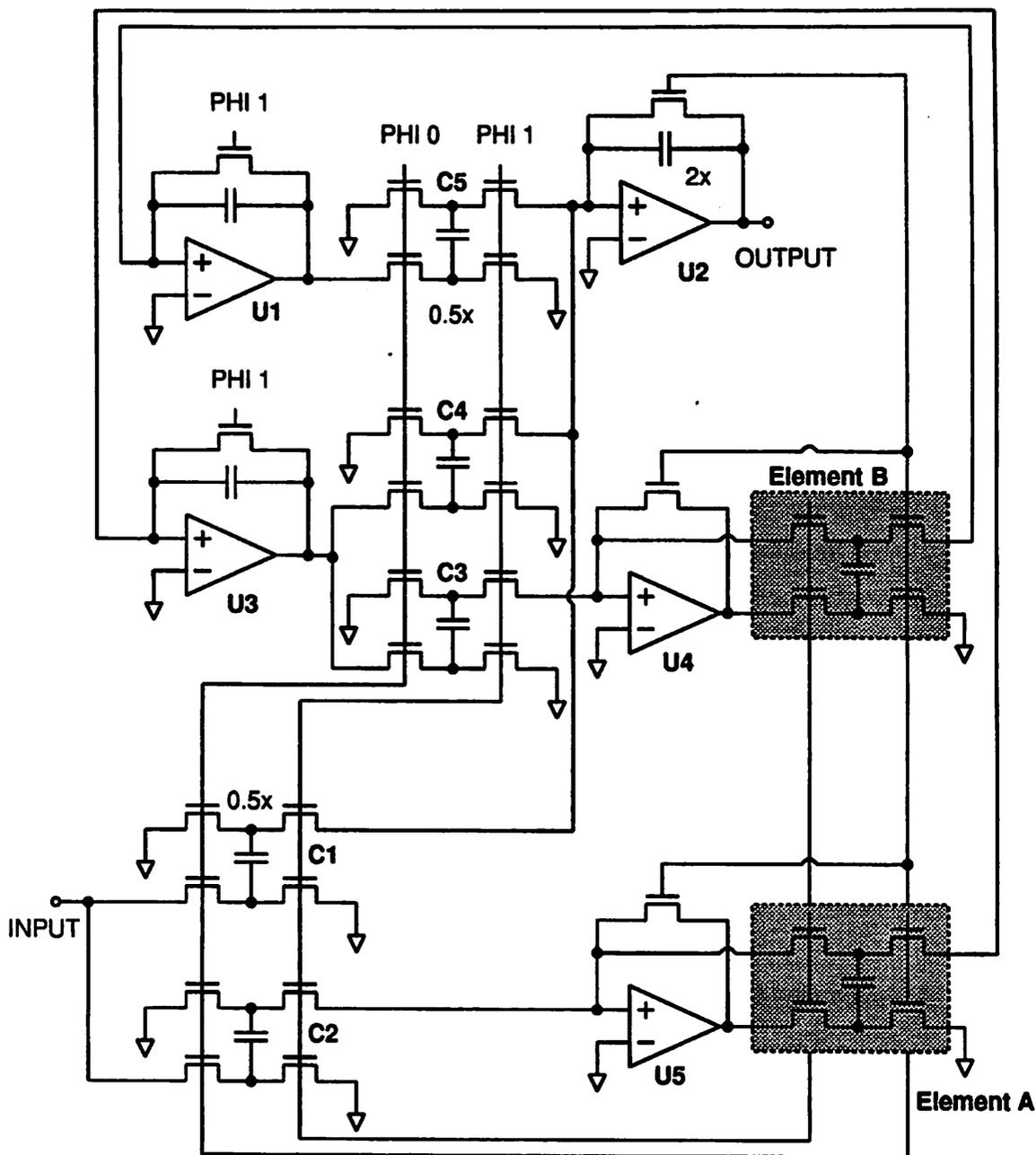
Actual operation of this comb filter would require that each storage element be clocked by a different phase of a multi-phase clock signal to insure that the proper storage element is switched in at the right point in time. In the actual chip, it is planned to manufacture this delay line in two to four sections per 1H delay segment as the parasitics associated with the individual delay elements and metal lines, while they do not affect the accuracy of the delay structure, result in a prolonged settling time for the op-amps.

5.7.4 Effects of Circuit Non-Idealities

Circuit non-idealities will affect the delay line discussed above in certain ways that are of significance in video applications. Of primary concern, is the gain of the operational amplifiers used in the integrator stages. A review of the signal paths shows that a given data sample is passed through five integrators. Finite operational amplifier gain will result in the loss of signal charge that manifests itself as a gain factor of less than unity through the filter. While this is similar to a non-unity charge transfer efficiency of CCDs, for a given integrator the gain is a fixed value, and thus will tend to apply the same transfer efficiency to all data samples that pass through the stage. Thus, rather than multiplying the signal by a random variable that changes once per clock cycle, the structure above will reduce the amplitude of the signal by a small unknown, but fixed amount. Therefore, the amount of noise added by this effect is minimal. However, if the magnitude of the loss is great enough, it will affect the effective filter coefficients used in the comb filter. This will result in a distortion in the frequency response very similar to that seen in the first stage decimator as a result of capacitor mismatches. As was with the decimator filter, complex zeros are more preferable to real axis zeros. Hence, a small compensating change may be made to capacitor C5 by increasing its value to help promote formation of complex zeros over real axis zeros.

The finite settling time of the operational amplifiers will also reduce the accuracy of the filter as the signals that are sampled onto capacitors C3 through C5 do not represent the actual signal value, but a quantity slightly different due to the settling of the operational amplifier. As this error is not directly proportional to the signal and will be a function of the previous data sample, it will constitute a distortion component. As a result, it is required that the operational amplifiers settle to within 0.5 percent in 40 nS.

The other major circuit non-ideality that will adversely affect the video signal is mismatch of the capacitors in the storage arrays. Ideally, the capacitance of each of the elements is identical. Mismatch of the capacitors will tend to create a fixed pattern noise in the image as the mismatch pattern is repeated once per line. As the video signals are stored as charges, this is not a primary concern. As a further precaution against fixed pattern noise, the addressing order of the individual cells will be reversed periodically. That is, cell 1 will correspond to the first pixel of a line during some periods of time, and the last pixel of the line during other periods. This has the effect of breaking up the pattern so as to make it much less visible.



NOTE: ALL CAPACITORS UNIT VALUE UNLESS OTHERWISE DESIGNATED

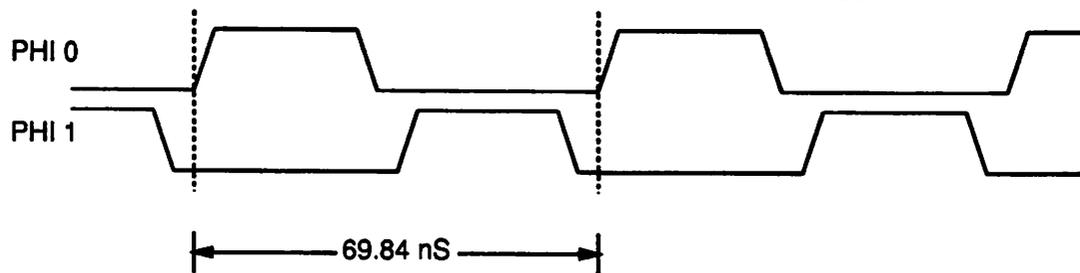


Figure 5.16: Comb Filter using Analog RAM cells

5.8 Chrominance Signal Demodulator

This section of the chip takes the separated chrominance signal from the Y/C separator, and performs a QAM demodulation to yield the two color components. Maximizing the performance of the chip requires that I/Q or "wideband" demodulation be performed, which uses all the information in the NTSC signal to reconstruct the original RGB values. However, due to the crosstalk component caused by the vestigial sideband of the modulated I channel, filters with different cutoff frequencies are required in the I and Q channels to remove extraneous out of band signals. Associated with the unequal cutoff frequencies are unequal delays through the filters, which would require that two delay lines of unequal length, one in the luminance channel and one in the I channel be added to insure that the total delay of the signals through the chip remain the same.

Replacing the I/Q demodulator with a Pr/Pb demodulator would result in significant hardware savings. The two filters for the Pr and Pb baseband channels could be of identical design; thus only one delay line for the luminance channel would be needed. Moreover, the extra dematrixing step of converting the I/Q signals to Pr/Pb would be eliminated simplifying both the signal path and the phase feedback circuitry (see below).

For the purposes of this report, both methods of demodulation will be discussed — at the time of publication, a final decision as to the method of demodulation has not been made. However, the methods are similar in that they both involve a quadrature multiplication followed by a filtering operation.

5.8.1 Multiplication Stage

Due to the $4 f_{sc}$ sampling rate of the system, the two signals necessary to demodulate the chrominance signal, $\sin 2\pi f_{sc}t$ and $\cos 2\pi f_{sc}t$, when sampled at the $4 f_{sc}$ rate, produce a repetitive pattern of samples. Moreover, if the sampling phase is set properly, the magnitude of the samples will be either zero or unity. This great simplification is used in the chip to demodulate the sampled chrominance signal. The chrominance signal is split into two parallel paths. The samples in the first path will be multiplied by a chain of coefficients $\{ \dots, 0, 1, 0, -1, 0, 1, 0, \dots \}$, which represents multiplication by $\sin 2\pi f_{sc}t$, while the samples in the second path will be subjected to the same chain of coefficients, but shifted in phase 90 degrees to perform the quadrature operation, $\{ \dots, 1, 0, -1, 0, 1, 0, -1, \dots \}$ ($\cos 2\pi f_{sc}t$).

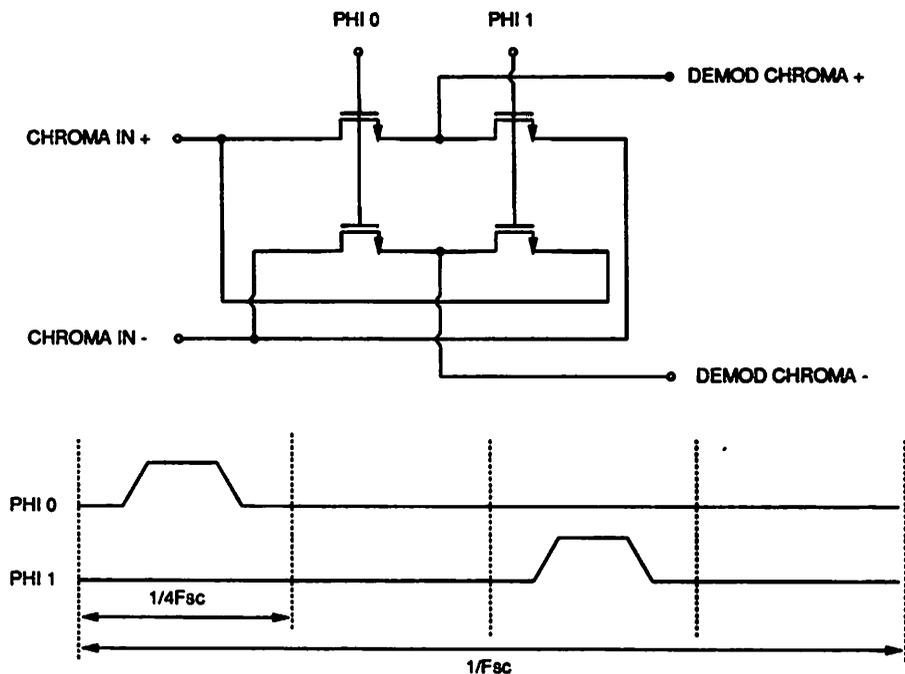


Figure 5.17: Chrominance Signal Demodulator for $4f_{sc}$ sampling rate

Circuitry to accomplish this task is relatively straightforward with a conceptual diagram given in Figure 5.4. Figure 5.16 shows a parasitic insensitive switching scheme with complementary switches to insure that common mode voltages near the supply rail do not cause the circuit to cease operating.

5.8.2 Sampling Phase Adjustment

One limitation of this method of demodulation is that the demodulating oscillators are fixed in phase. Thus, adjustment of the phase of the demodulating oscillators on chip relative to the modulating oscillator must be made during the initial sampling operation. This requirement places a demand on the sample and hold stage to be able to take a feedback signal from the chroma demodulator to insure that a proper relationship between the sampling phase and the demodulator exists.

Fortunately, the NTSC signal provides a very convenient method of locking the sampling phase to the proper value. Recall that at the start of each line, a colorburst signal corresponding to eight or nine cycles of the subcarrier waveform are inserted into the signal. More importantly, the phase of the colorburst signal is fixed, corresponding to a chrominance signal with 100 percent

-Pb component and no Pr component. Thus, if the sampling phase is offset from the proper value, a non-zero value of Pr will be demodulated when the colorburst signal is used as an input.

The option of demodulating along the I/Q axes instead of the Pr/Pb axes adds a slight complication. Adjusting the sampling phase until the Pr output nulls would result in demodulating the chrominance signal along the Pr/Pb axes. Processing the signal along the I/Q axes requires that the sampling phase be shifted by 33 degrees (with respect to f_{sc}). Rather than attempting to include a discrete time phase shift network, a dematrixing operation is performed before checking for null output. That is, assume that the sampling phase is properly adjusted for I/Q demodulation. Then, the two outputs are the I and Q signals. By feeding these two signals into an I/Q to Pr/Pb matrix as defined in chapter three, the Pr output during the colorburst interval should be zero; if not, it indicates that the sampling phase is incorrect and should be adjusted. The error voltage generated by this phase comparison is fed back to the crystal VCO described in section 5.4.3. In the event of Pr/Pb demodulation, the Pr output would be used directly as an error signal.

An option that is being considered is to add a separate demodulation circuit (multiplier only) strictly for sampling phase adjustment. A concern is that the group delay through the Y/C separators and associated filters may make the feedback loop from the chroma demodulator back to the sample and hold have excess delay contributing to loop instability. Thus, by tapping the signal during the burst period immediately after the sample and hold and performing the demodulation, an error signal can be generated without accumulating the delay through the signal processing circuits.

5.8.3 Chrominance Signal Filtering

The double frequency by-products of QAM demodulation along with any crosstalk in the Q channel must be removed using a set of low pass filters. As the crosstalk component in the Q channel starts at the frequency where the upper sideband of the modulated I channel is cut off, the stopband of the baseband Q channel should begin at that frequency. The NTSC channel is limited to 4.2 Mhz, hence the upper sideband of the I channel is cut off at (4.2 - 3.58) Mhz or 620 Khz. Thus, a filter with a passband of about 500 Khz with a fairly sharp cutoff providing approximately 20 dB of attenuation at 800 Khz is required. Although FIR filters are always preferable in video work due to their linear phase characteristics, the stopband requirements of this filter would mandate excessively long filters lengths. As such, a switched capacitor filter realization of a continuous time elliptic filter is the design of choice. As with the second stage decimation circuit, the

filter must be followed with some type of allpass network to equalize to some extent the delay of the filter. As the performance demanded through the chrominance channel is not as stringent as through the luminance channel, a third order elliptic (one transmission zero) with a biquad allpass filter proves to be sufficient. The structure used is very similar to that in Figure 5.14.

5.9 Projected Performance of the Chip

Here, an attempt will be made to quantify the performance of the chip and to compare this embodiment with those presented earlier. As stated in the introduction, the design goal of this chip is to achieve near broadcast performance in the areas of frequency response and Y/C separation. This chip is the first known to implement a multiple line adaptive comb filter for chrominance separation. The chip will be designed to meet or exceed 56 dB of SNR. In order to justify the claim that this level of performance can be met with less silicon area, an estimate of the total silicon area will be made for key sub-blocks of the chip.

5.9.1 Oversampling Front-End and Decimators

The key components of the oversampling front end and first stage filter are the sampling capacitors, the switches and the integrating operational amplifier. There are 20 capacitors used, of which 8 are double size. Assuming a worst case unit capacitor size of 600 fF based on kT/C considerations and a specific capacitance of $0.4 \text{ fF}/\mu\text{m}^2$, the area of the capacitors is $42,000 \mu\text{m}^2$. The operational amplifier used must be high-speed, settling in under 35 nS. Such amplifiers have been designed and occupy approximately $60,000 \mu\text{m}^2$ in a $1 \mu\text{m}$ technology. Therefore, allowing for a factor of two for wiring and metalization, the total area for this stage is approximately $210,000 \mu\text{m}^2$.

The clock generation for the filter requires the generation of a six phase clock along with a half rate three phase clock. Area estimates for this circuitry range from $50,000$ to $80,000 \mu\text{m}^2$.

The second stage decimator consists of a fifth order elliptic filter and two second order delay equalizers. This can be realized using eight operational amplifiers plus capacitors and switches as one op-amp is shared between the first and second stages. A fifth order elliptic filter for PCM voiceband has been fabricated in $1.25 \mu\text{m}$ technology with the goal of minimizing the area in 510 mil^2 [49]. Therefore allowing for wasted area, a reasonable estimate for this filtering stage is $1.28 \times 10^6 \mu\text{m}^2$. Thus, the total front end sample and hold with decimation filters is approximately

$1.57 \times 10^6 \mu\text{m}^2$, or 1.57 mm^2 .

5.9.2 Delay Lines for use in the Y/C Separator

The Y/C separator for this chip makes use of multiple arrays of delay line elements to realize the 1H delays that are necessary. As three 1H delay lines are necessary to realize the comb filter, combined with the fact that this is a differential chip, a total of 6 1H delay lines are required. Each delay line consists of 910 sampling capacitors and the associated switches. Assuming that the sampling capacitors are 300 fF each, and the switches are of minimum size, a reasonable unit cell area would be $3,500 \mu\text{m}^2$ including an allowance for metalization and wiring. This results in an array size of $3.185 \times 10^6 \mu\text{m}^2$ per 1H delay. Moreover, each pair of delay lines require at least three driver (read/write) stages, which contain five op-amps each. Thus, the total area per differential 1H delay line is approximately $7.1 \times 10^6 \mu\text{m}^2$, or 7.1 mm^2 (11,000 mil²). Three delay line pairs are required for the comb filter, thus making the total area approximately 21.3 mm^2 (33,000 mil²).

Additional area is required for the rather complex clock generation system used to address the individual storage elements. A rough estimate places the number of gates required at 100. Assuming that each gate takes $500 \mu\text{m}^2$ and allowing 50 percent for wiring overhead, the total area for clock generation is $75,000 \mu\text{m}^2$.

5.9.3 Adaptive Switching Circuit for Comb Filter

The circuitry necessary to implement the adaptive switching involves the use of short delay lines and non-linear function circuits such as absolute value and RMS calculation. Although the circuitry has not been finalized as of the printing of this report, it is estimated that the circuit will consist of ten op-amps or their area equivalents resulting in an area of approximately $650,000 \mu\text{m}^2$ (1010 mil²).

5.9.4 Chroma Low-Pass Filters

This stage requires a delay line for the luminance path, and two low pass filters for the chrominance channels. Two third order elliptic filters are required with the luminance path delay equalization being performed via a switched capacitor delay line. A total of nine op-amps are projected for this stage, with additional area being used for switching and clock generation for the delay elements. A very preliminary estimate of this stage places the area at $1.5 \times 10^6 \mu\text{m}^2$ (2,326

mil²).

5.9.5 Overall Chip Area

The area for the complete chip consists of the sum of the above areas plus probably another factor of fifty percent for miscellaneous circuits not included above such as the sync processor, buffer and oscillator. In addition, another factor is required to account for inter-block wiring and related non-active area. As such, a preliminary estimate of the total chip area is $51.1 \times 10^6 \mu\text{m}^2$ (79,225 mil²).

Chapter 6 Conclusions

6.1 Results

A proposed architecture for an integrated NTSC decoder with an on-chip 2H adaptive comb filter using analog sampled data signal processing techniques has been described. By use of novel circuit techniques and making full use of the characteristics of the NTSC signal, an efficient use of silicon area is achieved while providing a very high quality signal at the output. These include the use of the blanking interval as a reference, sampling phase adjustment via the color-burst signal, and the use of $4f_{sc}$ sampling rate processing. Obstacles to full integration of the decoding function such as anti-alias filtering are lifted by the inclusion of an oversampling front-end, which greatly relaxes the specifications for the continuous time filter required.

Special consideration was given to architectures required to provide the functionality of a DSP core while remaining in the analog sampled data domain. A new block, the analog RAM cell, was introduced as a means to store large amounts of analog sampled data in a reasonably small area. Issues such as component mismatch and circuit non-idealities were addressed and circuit techniques to minimize their effect were developed. The oversampling front-end was optimized for minimum circuit complexity by combining a novel transversal FIR filter with a standard switched capacitor filter structure.

Several conclusions can be drawn from the development of the architecture for the chip. First and foremost is the fact that straight DSP processing leads to a circuit which consumes more silicon area than a mixed analog/digital design. Second, circuit functions that have previously been relegated to the digital signal processing domain have found new implementations in the analog domain thus extending the utility of analog signal processing techniques. Finally, circuit non-idealities associated with analog processing can be overcome with careful layout, use of parasitic insensitive techniques, and pre-compensation of certain circuit elements.

6.2 Future Work

One obvious area of expansion of this work would be to design and fabricate the inverse function, that of a NTSC coder. The concepts of two-dimensional filtering can be extended to the coder as well. In fact, systems can be designed to incorporate filtering in both the coder and decoder to achieve higher levels of performance than would be attainable with filtering in either one alone. Many of the circuit concepts developed for this work can be applied to a coder chip.

Within the realm of the decoder, the optimization of filters remains as a potential area of improvement. Work needs to be done to ascertain whether the two stage FIR/IIR filter used for decimation is optimum in terms of silicon area and performance. An all FIR technique was considered early in this work and was abandoned due to the large area. However, a novel circuit topology combined with proper selection of FIR filter coefficients has a large chance of resulting in a superior filter to that proposed for this chip. Other areas of potential improvement include the Y/C separator and chroma demodulator. To limit the complexity of this chip, only a very simple adaptive algorithm was implemented. However, as was mentioned, this algorithm has the tendency to suffer from false triggering resulting in reduced performance of the overall Y/C separator. Inclusion of a more complex filtering system with a more intelligent method of controlling the various separation processes is a definite area of interest for this application.

Related to the Y/C separation problem is the lack of good simulation tools. Currently, algorithms are developed through actual implementation in hardware. As the mathematical model of an NTSC signal is available, simulation tools for the design of optimum Y/C separators can in theory be made available.

Finally, better simulation tools need to exist for large system oriented chips such as this work. Circuit level simulators are inappropriate to verify the proper operation of an entire chip such as this. While such simulators work well for each individual sub-block, they are incapable of handling the large number of circuit elements that comprise the entire chip. Simulators which are capable of modeling a circuit sub-block as a macro-cell are needed to verify operation of an entire system level circuit, along with methods to verify that each sub-block interfaces properly with all the other sub-blocks.

References

- [1] R. Luke, 1989, private communication.
- [2] H. E. Ennes, *Television Broadcasting: Equipment, Systems, Operating Fundamentals*, Indianapolis IN, H. W. Sams, 1979.
- [3] Y. Faroudja and J. Roizen, "Improving NTSC to Achieve Near-RGB Performance," *SMPTE Journal*, vol. 96, no. 8, pp. 750-761, August 1987.
- [4] S. J. Auty, D. C. Read, and G. D. Roe, "PAL Color Picture Improvement Using Comb Filters," *BBC Engineering*, 108:28, Dec. 1977.
- [5] H. E. Ennes, *Television Broadcasting: Equipment, Systems, Operating Fundamentals*, Indianapolis IN, H. W. Sams, 1979.
- [6] D. G. Fink, *Color Television Standards: Selected Papers and Records of the National Television System Committee*, New York NY, McGraw-Hill, 1955.
- [7] F. G. Stremler, *Introduction to Communication Systems*, Reading PA, Addison-Wesley, 1982.
- [8] A. B. Carlson, *Communication Systems*, New York NY, McGraw-Hill, 1986.
- [9] A. N. Netravali, B. G. Haskell, *Digital Pictures*, New York NY, Plenum Press, 1988.
- [10] D. G. Fink, *Color Television Standards: Selected Papers and Records of the National Television System Committee*, New York NY, McGraw-Hill, 1955.
- [11] G. B. Townsend, *PAL Colour Television*, London, Cambridge University Press, 1970.

Optimal Architectures for an Integrated NTSC Decoder

- [12] F. G. Stremier, *Introduction to Communication Systems*, Reading PA, Addison-Wesley, 1982.
- [13] M. E. Van Valkenburg, *Analog Filter Design*, New York NY, Holt, Rinehart and Winston, 1982.
- [14] V. Gopinathan, Y. Tsvividis, "A 5V 7th-Order Elliptic Analog Filter for Digital Video Applications," in *ISSCC Digest of Technical Papers*, pp. 208-209, 1990.
- [15] W. F. Schreiber, "Improved Television Systems: NTSC and Beyond," in *SMPTE Journal*, vol. 96, no. 8, pp. 734 - 744, August, 1987.
- [16] Y. Faroudja and J. Roizen, "Improving NTSC to Achieve Near-RGB Performance," *SMPTE Journal*, vol. 96, no. 8, pp. 750-761, August 1987.
- [17] J. O. Drewery, "The Filtering of Luminance and Chrominance to Avoid Cross-Colour in a PAL Colour System," *BBC Engineering*, vol. 8, pp. 8 - 39, Sept. 1976.
- [18] J. Rossi, "Comb Filter for TV Signals," U. S. Patent No. 4,050,084, July 14, 1976.
- [19] L. B. Jackson, *Digital Filters and Signal Processing*, Boston MA, Kluwer Academic Publishers, 1989.
- [20] Asahi Glass Co, "Technical Data Sheet for Delay Lines for Video Applications."
- [21] R. Luke, 1989, private communication.
- [22] S. D. Wagner, 1989, private communication.
- [23] S. M. Sze, *Physics of Semiconductor Devices*, 2 ed., New York NY, John Wiley & Sons, 1981.

References

- [24] Y. Maki, T. Kondo, A. Izumi, et. al., "A CMOS-CCD Comb Filter with Dropout Compensation for a VCR," in *ISSCC Digest of Technical Papers*, pp. 46 - 47, 1988.

- [25] Y. Faroudja, "Adaptive Comb Filtering," U. S. Patent No. 4,179,705, Mar. 13, 1978.

- [26] Y. Faroudja and J. Roizen, "Improving NTSC to Achieve Near-RGB Performance," *SMPTE Journal*, vol. 96, no. 8, pp. 750-761, August 1987.

- [27] Y. Faroudja and J. Campbell, "Processing Methods Using Adaptive Threshold For Removal of Chroma/Luminance Cross-Talk in Quadrature-Modulated Subcarrier Color Television Systems," U. S. Patent No, 4,731,660, Mar. 15, 1988.

- [28] S. D. Wagner, 1990, private communication.

- [29] International Radio Consultative Committee (CCIR), Recommendation No. 601, 1986.

- [30] W. N. Sproson, *Colour Science in Television and Display Systems*, Belfast, Universities Press, 1983.

- [31] M. Nagatani, H. Yoshimura, T. Tsuchiya, and Y. Suzuki, "Digital Signal Processors for Decoding/Encoding Color TV Signals," in *IEEE Journal of Solid-State Circuits*, vol. SC-21, no. 6, pp. 964 - 970, Dec. 1986.

- [32] M. Ohta, K. Kohiyama, N. Tahara, et. al., "A Single-Chip CMOS Analog/Digital Mixed NTSC Decoder," in *ISSCC Digest of Technical Papers*, pp. 118 - 119, 1990.

- [33] Philips/Signetics, *Technical Descriptions of Digital Television Signal Processing System and Demonstration Board Schematics*, July 1989.

- [34] H. Kniess, 1990, private communication.

- [35] Y. Maki, T. Kondo, A. Izumi, et. al., "A CMOS-CCD Comb Filter with Dropout Compensation for a VCR," in *ISSCC Digest of Technical Papers*, pp. 46 - 47, 1988.
- [36] M. Sato, T. Hashimoto, S. Ogasawara, K. Suzuki, "A CMOS CCD Video Delay Line," in *ISSCC Digest of Technical Papers*, pp. 120 - 121, 1984.
- [37] D. M. Monticelli, "A Quad CMOS Single-Supply Op Amp with Rail-to-Rail Output Swing", in *IEEE Journal of Solid-State Circuits*, vol. SC-21, no. 6, pp. 1026 - 1034, Dec. 1986.
- [38] B. Kim, P. R. Gray, "A 30 Mhz High-Speed Analog/Digital PLL in 2 μ m CMOS" in *ISSCC Digest of Technical Papers*, pp. 104 - 105, 1990.
- [39] R. M. Dorwood, "Aspects of the Quantization Noise Associated with the Digital Coding of Colour-Television Signals," *Electronic Letters*, pp. 5 - 7, Jan. 8, 1970.
- [40] U. S. Patent No. 3,946,432
- [41] International Radio Consultative Committee (CCIR), Recommendation No. 601, 1986.
- [42] R. G. Meyer, "EECS 242 Class Notes", University of California, Berkeley, 1989.
- [43] J. B. Shyu, G. C. Temes, F. Krummenacher, "Random Error Effects in Matched MOS Capacitors and Current Sources" in *IEEE Journal of Solid-State Circuits*, vol. SC-19, no. 6, pp. 948 - 955, Dec. 1984.
- [44] E. A. Guillemin, *Synthesis of Passive Networks*, New York NY, Wiley, 1957.
- [45] R. Gregorian, "Switched-Capacitor Filter Design Using Cascaded Sections" in *IEEE Trans. Circuits Syst.*, vol CAS-27, pp. 515 - 521, June, 1980.
- [46] P. E. Fleischer and K. R. Laker, "A Family of Active Switched Capacitor Biquad Building Blocks" in *Bell Syst. Tech. J.*, vol. 58, pp. 2235 - 2269, Dec. 1979.

References

- [47] G. C. Temes, H. J. Orchard, "Switched-Capacitor Filter Design Using the Bilinear z-Transform" in *IEEE Trans. Circuits Syst.*, vol. CAS-25, pp. 1039 - 1044, Dec. 1978.
- [48] M. S. Lee, G. C. Temes, et. al., "Bilinear Switched-Capacitor Ladder Filters" in *IEEE Trans. Circuits Syst.*, vol. CAS-28, pp. 811 - 821, Aug. 1981.
- [49] S. P. Shieh, C. K. Wang, R. Castello, P. R. Gray, "A Scalable Switched-Capacitor Filter Implemented in 1.25 μm Technology" in *IEEE Journal of Solid-State Circuits*, vol. 24, no. 1, Feb. 1989.

Additional References

- (1) E. Dubois and W. F. Schreiber, "Improvements to NTSC by Multidimensional Filtering", *SMPTE Journal*, vol. 97, no. 6, pp. 446 - 463, June, 1988.
- (2) D. Teichner, "Adaptive Filter Techniques for Separation of Luminance and Chrominance in PAL TV Signals," in *IEEE Trans. Consumer Elect.*, vol. CE-32, pp. 241 - 250, Aug. 1986.
- (3) J. Rossi, "Digital TV Comb Filter with Adaptive Features," in *Proc. IERE Conf. on Video and Data Recording*, pp. 267 - 281, 1976.
- (4) Y. Faroudja and J. Roizen, "A Progress Report on Improved NTSC", in *SMPTE Journal*, vol. 98, no. 11, pp. 817 - 822, November, 1989.

Optimal Architectures for an Integrated NTSC Decoder

The NTSC color system is a complex of interrelated processes. The primary challenge in its implementation is the efficient and accurate conversion of the color signals into a form suitable for transmission and subsequent decoding. This involves a series of operations, including chrominance-to-luminance conversion, which is a critical step in the overall decoding process. The architecture of the decoder must be designed to handle these operations in a way that minimizes distortion and maximizes efficiency.

One of the key considerations in the design of an integrated NTSC decoder is the choice of the color subcarrier frequency. This frequency is determined by the relationship between the line rate and the frame rate, and it plays a crucial role in the synchronization and timing of the color signals. The decoder must be able to accurately track this frequency, which is often done using a phase-locked loop (PLL) circuit. The design of the PLL is therefore a critical part of the overall decoder architecture.

In addition to the color subcarrier frequency, the decoder must also be able to handle the timing of the color signals. This is typically done using a color burst, which is a short burst of color information that is transmitted at the beginning of each line. The decoder must be able to detect and track this burst, which is often done using a color burst detector circuit.