Active filters for video

Originally, video filters were passive LC circuits surrounded by amplifiers. Smaller, more efficient designs can currently be achieved by combining the amplifier with an RC filter. Sensitivity analysis and predistortion methods developed in the 1960s have, moreover, overcome the poor performance that gave early video filters a bad reputation.

High-performance op amps and specialized software for the PC enable the design of wide-bandwidth active filters, but those advantages do not address the requirements of any specific application. For video filters, the particular application and signal format add nuance to each circuit design. The two major video applications follow.

Antialiasing filters: These devices are placed before an analog-to-digital converter (ADC) to attenuate signals above the Nyquist frequency, which is one half the sample rate of the ADC. These filters are usually designed with the steepest possible response to reject everything above the cutoff frequency[1]. For ITU-601 applications and others, such performance is achieved using analog filters combined with digital filters and an oversampling ADC. For applications such as PC graphics, very little filtering is required.

Reconstruction filters: Also called (sinx)/x or zero-order-hold correctors, these filters are placed after a digital-to-analog converter (DAC) to remove multiple images created by sampling, though not to remove the DAC clock. Reconstruction filters are seldom as selective as antialiasing filters, because the DAC's hold function also acts as a filter—an action that lowers the required selectivity, but introduces loss in the response. The available video formats are RGB, component video, composite video, and RGB PC graphics.

All applications and formats require a video filter to be "phase linear," a condition specified by the parameter called group delay (delay versus frequency). The degree of phase linearity required depends on the application and the video format. For example, antialiasing filters and component formats are more tightly specified than are reconstruction applications and composite video. Requirements for the various applications and formats are specified by NTSC, PAL/DVB, ITU, SMPTE, and VESA.

This article compares different filters to determine the optimum design for a given application or format. Rauch and Sallen-Key realizations are compared for their GBW-to-cutoff ratios, using predistortion and element sensitivity techniques to achieve accuracy in the design. Those filters to be considered are:

- An ITU-601 antialiasing filter
- A 20MHz antialiasing and reconstruction filter
- An HDTV reconstruction filter

Filters and their characteristics

Whether used for antialiasing or reconstruction, the filter must have a lowpass characteristic to pass the video frame rate. One should, therefore, be wary of AC-coupling. Lowpass filters are categorized by their amplitude characteristic or by the name of the polynomial that describes it (Bessel, Butterworth, Chebyshev, or Cauer). **Figure 1** shows these characteristics normalized to a 1-rad bandwidth. Typically, a filter with the best selectivity and the minimum number of poles (to minimize cost) would be chosen, but the additional need for phase linearity limits the available choices.

Phase linearity and group delay

A filter's phase linearity is specified as envelope delay or group delay (GD) versus frequency. A flat group delay indicates all frequencies are delayed by the same amount, which preserves the shape of the waveform in the time domain. Thus, absolute group delay is not as important as the variation in group delay. A separate specification called channel-to-channel variation, which is specified as "time coincidence," should not be confused with group delay.

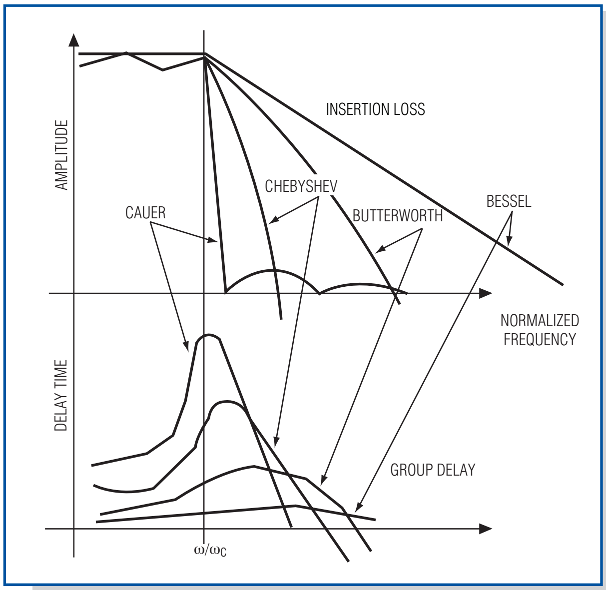


Figure 1. Amplitude and group delay vs. frequency for various filter types are normalized to a 1-rad bandwidth.

Though not desirable for video, how much group-delay variation is acceptable, and why? The answer depends on the application and the video format. For example, ITU-470 specifies group delay very loosely for composite video. However, ITU-601 specifies it tightly to ensure generational stability, both for MPEG-2 compression and to control phase jitter before serialization. So, what filter characteristics are considered necessary to ensure phase linearity?

The group-delay curves in Figure 1 show a peak near the cutoff frequency ($\omega/\omega_c = 1$). That is a problem caused by the steep phase change near the cutoff frequency. To get an idea of scale, a 3-pole, 6MHz Butterworth filter has a group-delay variation of 20ns–25ns over its bandwidth. Increasing the number of poles or the filter's selectivity increases that variation. Other, more exotic filters[2] used to minimize group-delay variation include Bessel, phase approximation, Thompson-Butterworth, and LeGendre. Nevertheless, the Butterworth characteristic is most often used for video.

Group-delay problems with component video

All formats and applications are sensitive to group-delay variation. The degree of sensitivity depends on the number of signals and their bandwidths. Composite NTSC/PAL has only one signal, with group delay specified in ITU-470. Those requirements are easily met. RGB and component video each have multiple signals. The RGB signals have equal bandwidths while component-video signals do not, making group-delay matching easy with RGB, but difficult with component video.

Because Pb and Pr signals have half the bandwidth of the luma (Y) signal, their group delay is double that of the Y signal[3]. One solution is to slow down the Y signal by adding delay stages. Another solution is to equalize bandwidths by doubling the sample rates of Pb and Pr, which raises the 4:2:2 sampling rate to 4:4:4[4], allowing the signal to be treated as RGB. The additional Pb and Pr samples are discarded during antialiasing or averaged in reconstruction applications.

The other component-video format, S-VHS, can be somewhat confusing. The Y channel is the same as in YPbPr, but the chroma signal (C) looks like it should be bandpass filtered rather than lowpass filtered. As for YPbPr signals, bandpass filtering causes group-delay and timing problems and, therefore, should not be implemented. Unless analog encoding is done, Y and C can be lowpass filtered with the same filter. S-VHS is more forgiving of bandwidth than of problems caused by trying to equalize the delay. S-VHS is typically seen in reconstruction applications, for which the main concern is correct timing between Y and C.

Choosing an op amp

After choosing a filter characteristic, the next step is to implement it with an actual circuit. The most commonly used, single-op-amp circuits are the Sallen-Key configuration in noninverting form and the Rauch configuration in inverting form. An important consideration for op amps operating in the wide bandwidths of video applications is the minimum gain-bandwidth (GBW). Video signals are large, typically $2V_{P-P}$, so the large-signal GBW is referenced. This parameter is not to be confused with the $2V_{P-P}$ 0.1dB GBW, which is much lower.

For filter circuits, how much larger than the filter's cutoff frequency does the op-amp GBW have to be? For a Rauch (inverting) filter, the phase argument of the characteristic is:

$$\mathrm{Arg}[K(j\omega)]_{inv} = -(\omega_c/GBW_{rad})(1+Rf/Ri) \quad \text{(Eq 1)}$$

For a Sallen Key (noninverting) filter:

$$\mathrm{Arg}[K(j\omega)]_{noninv} = \pi -(\omega_c/GBW_{rad})(1+Rf/Ri) \quad \text{(Eq 2)}$$

where R_f and R_i are the gain-set resistors in ohms, GBW_{rad} is the op amp's gain-bandwidth product, and ω_c is the filter's cutoff frequency in radians per second. Set the gain by introducing values for R_f and R_i[5], and solve for (ω_c/GBW_{rad}). A unity-gain Rauch circuit has $R_f/R_i = 1$, and a Sallen-Key circuit has $R_f/R_i = 0$. Thus, for the same phase error, a Sallen-Key requires half the GBW of a Rauch circuit. As the required gain increases, they converge and leave little advantage for the Sallen-Key in terms of GBW, but other issues must be considered as well.

Predistortion, bandwidth, Q, and element sensitivity

Anything less than an infinite GBW_{rad}/ω_c ratio causes the closed-loop poles of a filter to move. That is why an actual filter often exhibits a lower bandwidth (ω_c) than does the paper design[6]. This can be compensated for by increasing the design bandwidth, which is known as predistortion. Formulas for the Sallen-Key and Rauch circuits (listed in **Tables 1** and **2**) allow us to calculate a design bandwidth that provides the actual bandwidth needed. Component tolerance must then be taken into account.

To determine component tolerance, a sensitivity function[7] is needed: $S_X{}^Y$ gives the ratio between a change in the value of part X and the consequent change in parameter Y. For example, Table 1 shows that the Q in a Sallen-Key circuit (vs. a Rauch circuit) has a large sensitivity to variations in C1 and C2. That means a Sallen-Key is less tolerant of parasitics than is a Rauch. The point is that $S_X{}^Y$ lets the effect be predicted, and then the design can be created accordingly. Next, some typical designs are considered.

Table 1. Component sensitivities including BW and Q predistortion formulas (Sallen-Key realization, ω_o = 1rad/sec)

Sensitivity Function S_X^Y	Gain K = 3 - 1/Q (R1 = R2 = C1 = C2 = 1)	Gain K = 1 (R1 = R2 = 1)	Gain K = 2 (R1 = C1 = 1)
S_x^ω (x = R1, R2, C1, C2)	-1/2	-1/2	-1/2
S_K^Q	14	50	10
S_{R1}^Q	4.5	0	4.5
S_{R2}^Q	-4.5	0	-4.5
S_{C1}^Q	9.5	1/2	5.5
S_{C2}^Q	-9.5	-1/2	-5.5
S_{Ra}^K	-9/14	N/A	-1/2
S_{Rb}^K	9/14	N/A	1/2
ω_C (actual)	ω_C (design)[1 - 1/2(3 - 1/Q)$^2\omega_C$ / GBW]	ω_C (design)[1 - ω_CQ / GBW]	—
Q (actual)	Q (design)[1 + 1/2(3 - 1/Q)$^2\omega_C$ / GBW]	Q (design)[1 + ω_CQ / GBW]	—

Table 2. Component sensitivities including BW and Q predistortion formulas (Rauch realization, ω_o = 1rad/sec)

Sensitivity Function S_X^Y	Gain K = 1 (R1 = R2 = R3 = 1)	Gain K = 2 (R1 = 1, R3 = H_O, R2 = (H_O / 1 + H_O))	Gain K = 2 (C1 = 1, C2 = C1 / 100)
S_x^ω(x = R2, R3, C1, C2, S_{R1}^ω = 0)	-1/2	-1/2	-1/2
S_{R1}^Q	1/3	1/3	1/3
S_{R2}^Q	-1/6	0	0
S_{C1}^Q	1/2	1/2	1/2
S_{C2}^Q	-1/2	-1/2	-1/2
S_{R3}^K	1	1	1
S_{R1}^K	-1	-1	-1
S_{R3}^Q	1/6	0	0
ω_C (actual)	ω_C (design)[1 - 3ω_CQ / 2GBW]	—	—
Q (actual)	Q (design)[1 + 3ω_CQ / 2GBW]	—	—

Design of antialiasing filters

For antialiasing filters, selectivity is determined by a template for ITU-601 like the one in **Figure 2**. The specified bandwidth is 5.75MHz ±0.1dB, with an insertion loss of 12dB at 6.75MHz and 40dB at 8MHz, and with a group-delay variation of ±3ns over the 0.1dB bandwidth. Such performance is too difficult for an analog filter alone, but 4x oversampling modifies the requirements to 12dB at 27MHz and 40dB at 32MHz.

Using software or normalized curves[8], one can find that a 5-pole Butterworth filter with -3dB bandwidth of 8.45MHz satisfies the requirement for selectivity, though not for group delay. For the latter, a delay stage is needed, for which the important op-amp parameter is the 0.1dB, $2V_{P-P}$ bandwidth[9]. That number should be used in equations 1 and 2 to get an accurate design. A schematic for this application, with curves showing its gain and group-delay characteristics, is based on 4x oversampling (**Figures 3a** and **3b**).

PC video is considered next. VESA does not specify templates for antialiasing or reconstruction filters. The XGA resolution (1024 x 768 at 85Hz) has a sampling rate of 94.5MHz and a Nyquist frequency of 47.25MHz. For >35dB attenuation at the Nyquist frequency, a Rauch realization of a 20MHz, 4-pole Butterworth filter (**Figures 4a** and **4b**) is used. Again, the MAX4450/4451 are chosen for their excellent transient response and large-signal bandwidth (175MHz at $2V_{P-P}$).

Reconstruction filters

Reconstruction filtering after a DAC is among the more poorly understood applications. Some designers think reconstruction filters are introduced to remove the sample

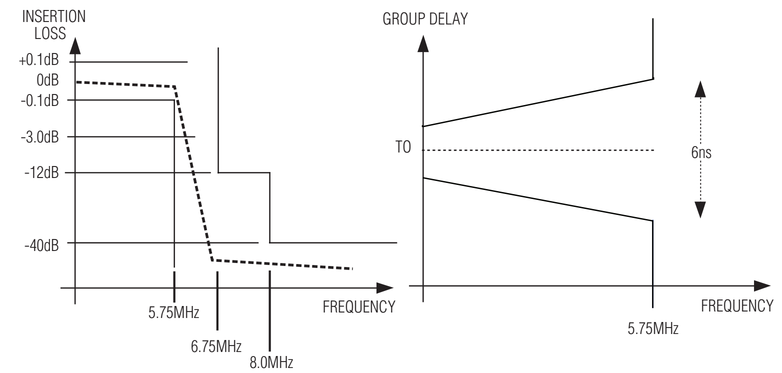


Figure 2. This filter template illustrates antialiasing requirements in accordance with the ITU-R BT.601-5 standard.

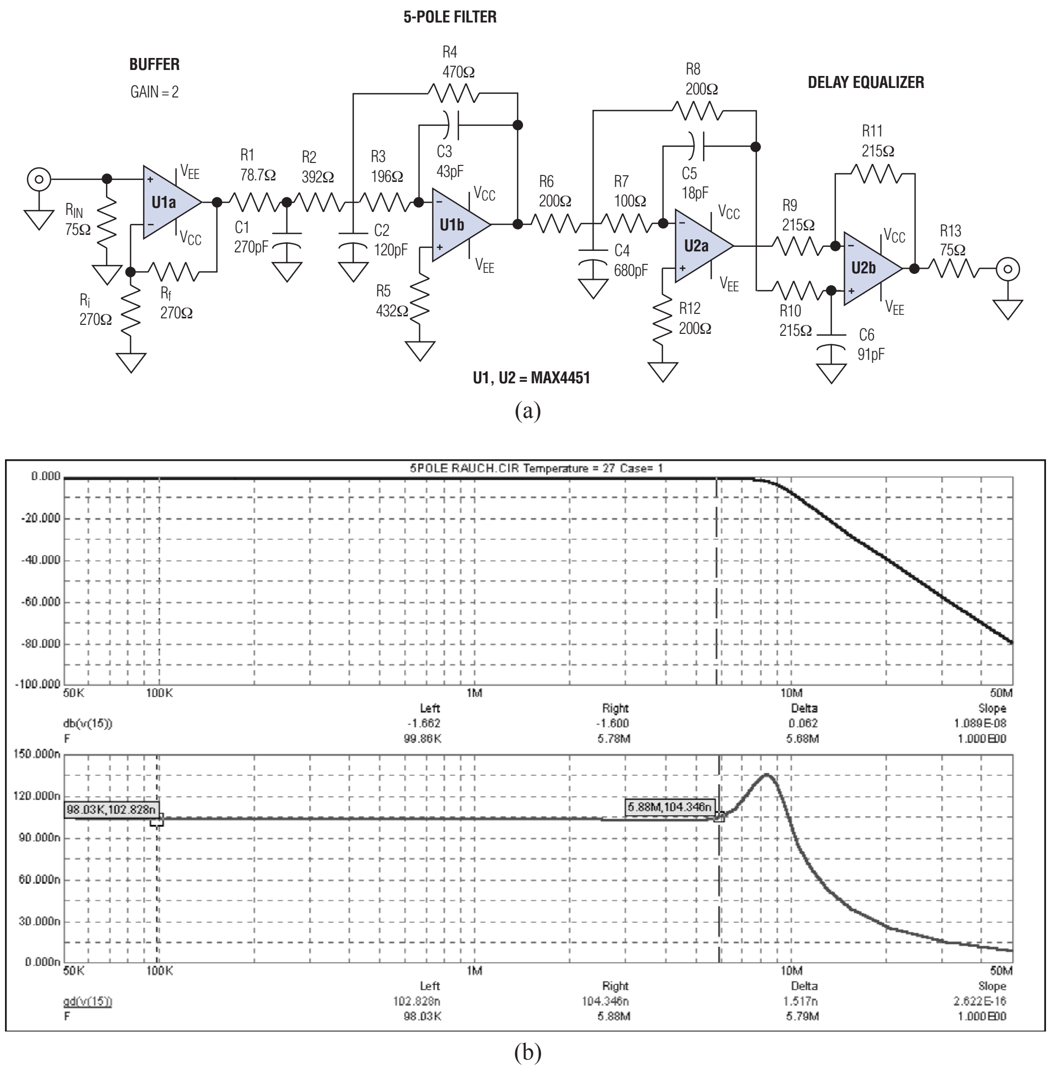


Figure 3. This schematic (a) and output response (b) represent a 5-pole, 5.75MHz Butterworth filter for ITU-601 antialiasing, using a Rauch circuit with a delay equalizer.

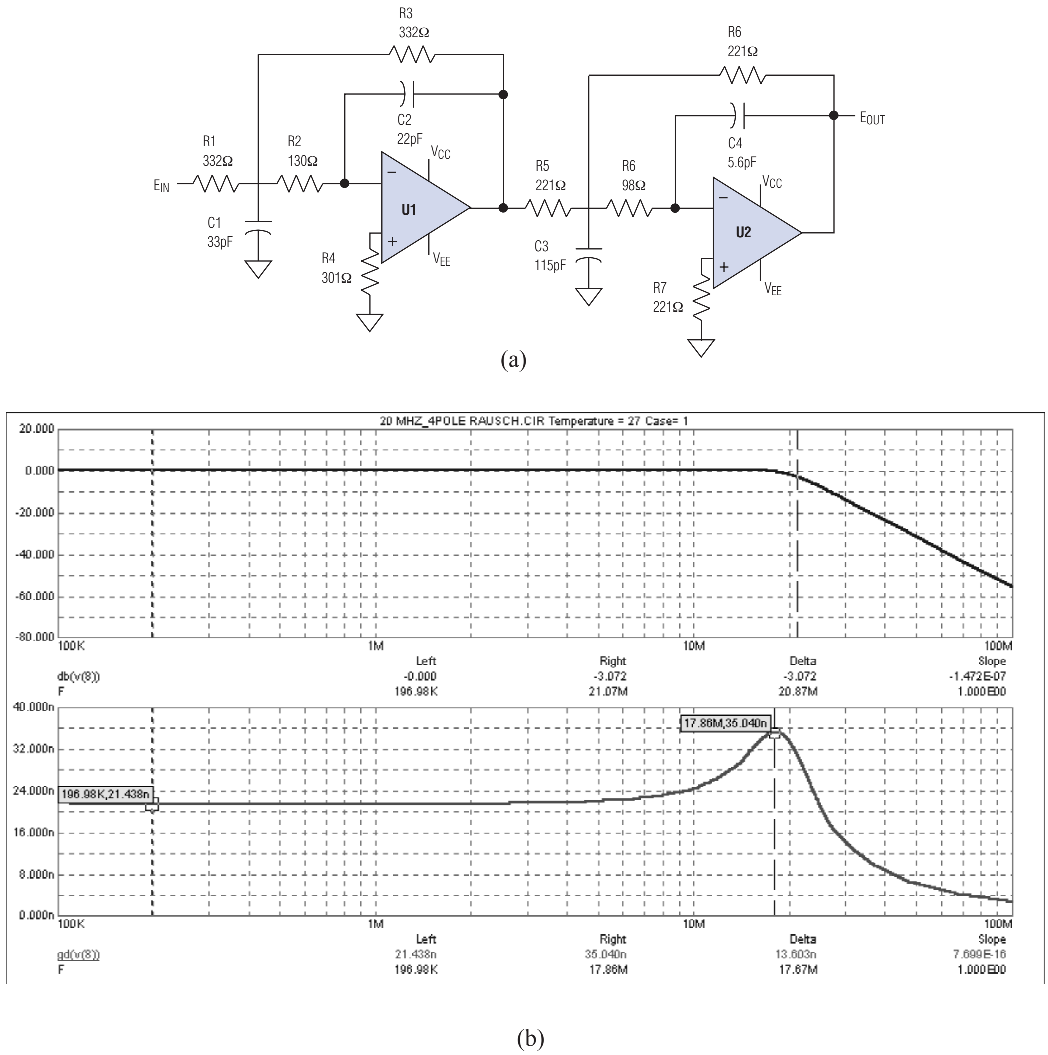


Figure 4. This schematic (a) and output response (b) depict a 4-pole, 20MHz Butterworth filter for XGA graphics antialiasing, using a Rauch circuit.

clock, but nothing is further from the truth. When a signal is sampled, the samples are composed of multiple recurring signal images centered on harmonics of the sample clock. A reconstruction filter removes all but the baseband sample. If the antialiasing filter has served its purpose, the DAC output looks like image A in **Figure 5**, and then all samples to the right of it should be removed. Thus, reconstruction is similar to antialiasing except that, because each sample exists only for an instant, the DAC holds each for one clock period, thereby creating the familiar staircase approximation to a sloping line.

The hold function corresponds to a digital filter whose characteristic[10] is similar to that of a Butterworth or Bessel filter (**Figure 6**). Notice that the response is decreased by 4dB at half the sample frequency. The second objective of a reconstruction filter is to restore that loss, which requires an amplitude equalizer like the circuit shown in **Figure 7a**. The equalizer is based on a delay stage and has a response like a Bessel filter. It can be designed from the DAC sample rate (F_S). **Figure 7b** shows the DAC's frequency response with and without an amplitude equalizer. Like the delay stage, it can be included in any reconstruction filter.

The hold response also has a pole centered on the sample clock, which completely removes the clock. Nevertheless, most reconstruction applications refer to clock attenuation as a figure of merit. Now that the function of a reconstruction filter is understood, one can be designed.

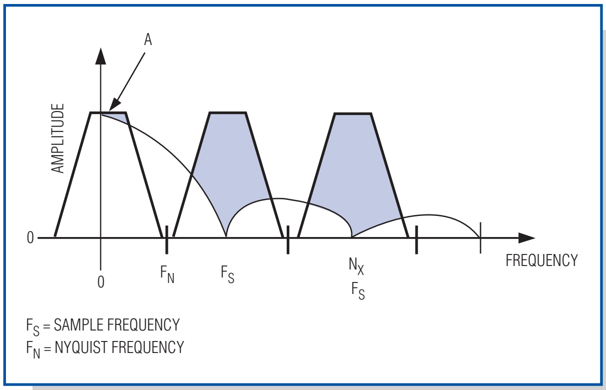


Figure 5. A typical DAC output spectrum is shown in terms of the sampling (F_S) and Nyquist (F_N) frequencies.

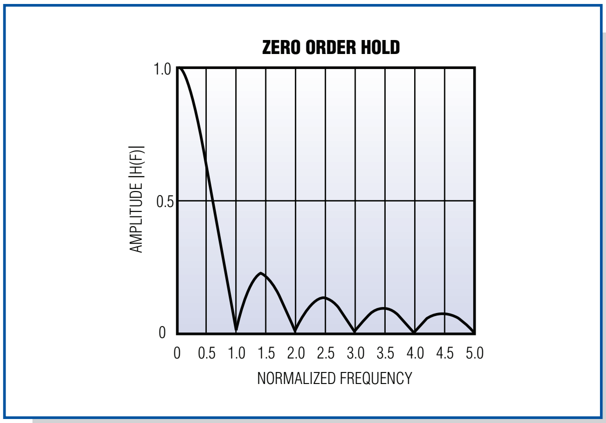


Figure 6. The "hold" function of a DAC produces a (sinx)/x response with nulls at multiples of the sampling frequency.

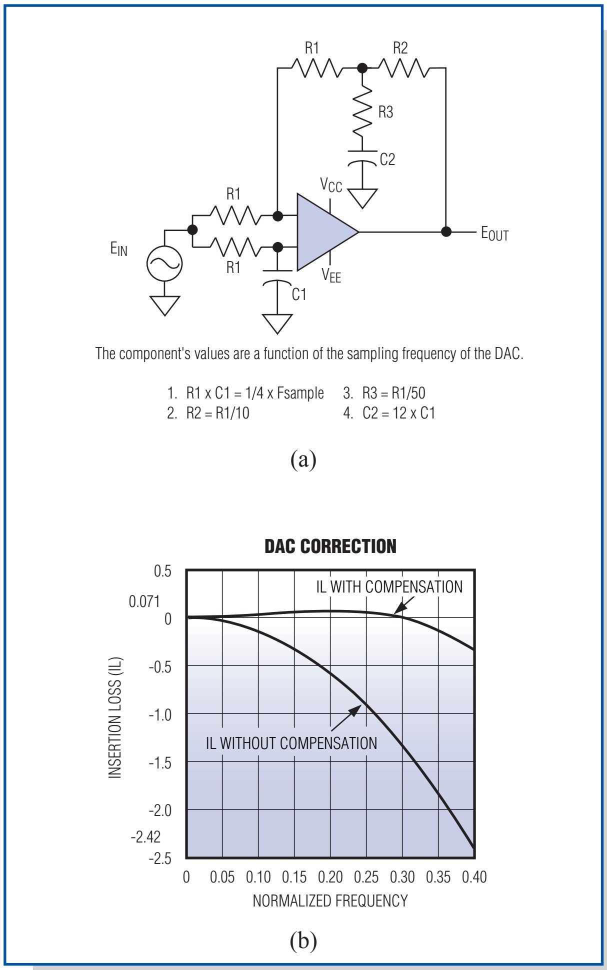


Figure 7. A DAC output (b) is shown with and without the (sinx)/x correction provided by an amplitude equalizer circuit (a).

The most common requirement for NTSC/PAL reconstruction is an attenuation of >20dB at 13.5MHz and >40dB at 27MHz, where ω_c depends on the applicable video standard. A 3-pole Butterworth with Sallen-Key configuration is chosen for two reasons. First, its gain (+2) drives a back-terminated cable. Second, the group-delay variation can be adjusted to optimize performance without a delay equalizer. (**Figures 8a–8d** show NTSC and PAL designs, including gain and group-delay characteristics.) These applications usually include digital amplitude correction for the DAC, which can easily be added, if necessary.

Illustrating a circuit for XGA, a 20MHz, 3-pole Butterworth filter in the Sallen-Key configuration includes the Figure 8 circuit for amplitude correction (**Figures 9a** and **9b**). Complementing the antialiasing filter of Figure 4, this filter has a gain of +2 to drive a back-terminated 75Ω coaxial cable.

The last application is a reconstruction filter for HDTV. Based on the templates in the SMPTE 274 and 296M, it has a center frequency of $\omega_c = 0.4 \times F_S = 29.7$MHz. Amplitude correction for the DAC is usually included, but group-delay compensation must be added. The resulting 30MHz, 5-pole Sallen-Key filter (**Figure 10**) has >40dB attenuation at 74.25MHz, as well as a group-delay stage with a +2 gain to drive a back-terminated 75Ω coaxial cable.

Practical aspects of active video-filter design

Whether filters are designed by hand, with the aid of software, or with a combination of these approaches, the actual response may not be exactly what is wanted. One cause is the discrepancy between a calculated response and the actual response obtained using standard component values.

That error can be minimized by choosing standard (5%) capacitor values and deriving resistor values from them. The reason is practical—capacitors with 1% or 2% tolerance can be acquired, but only at 5% values, although

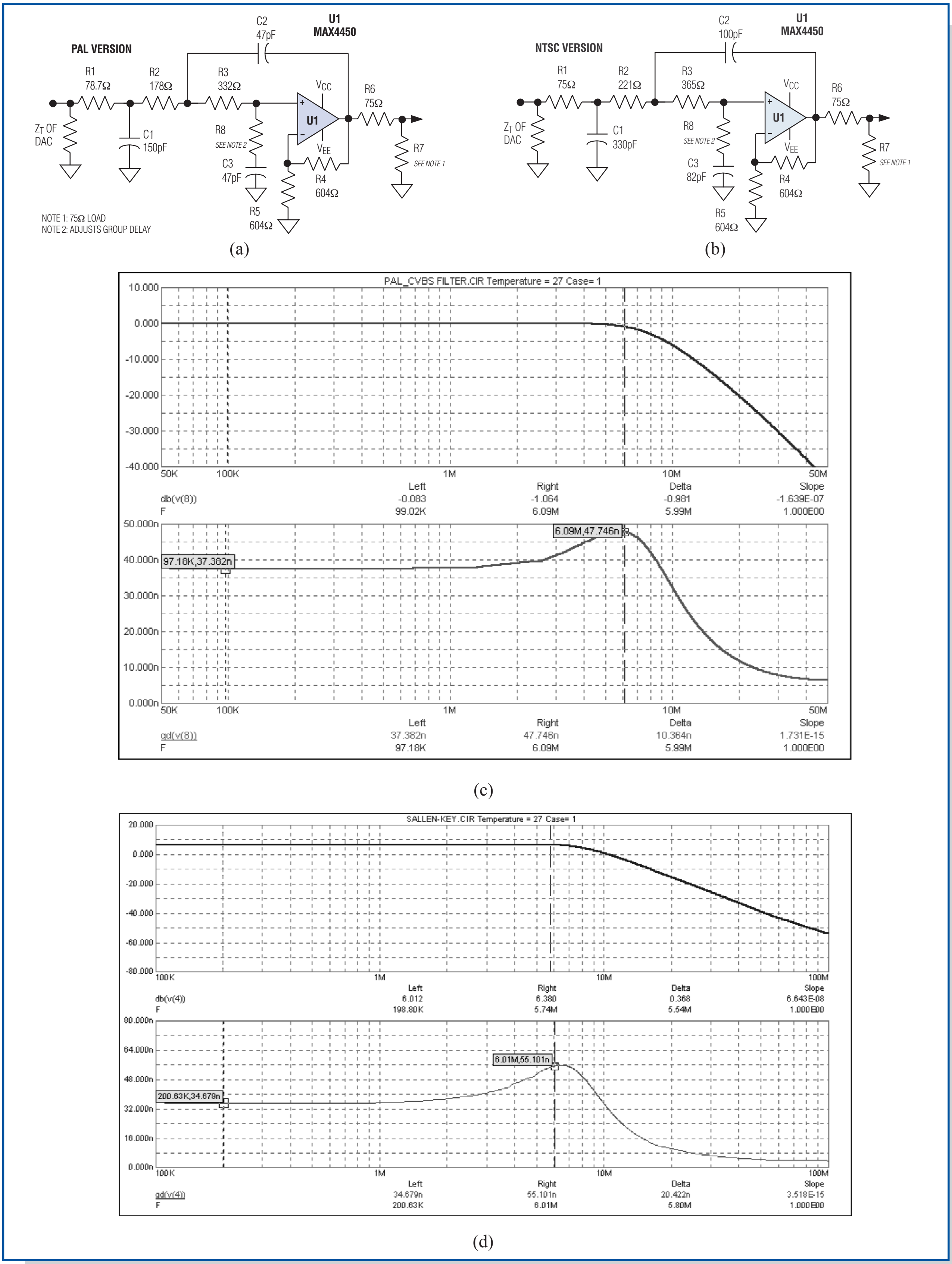
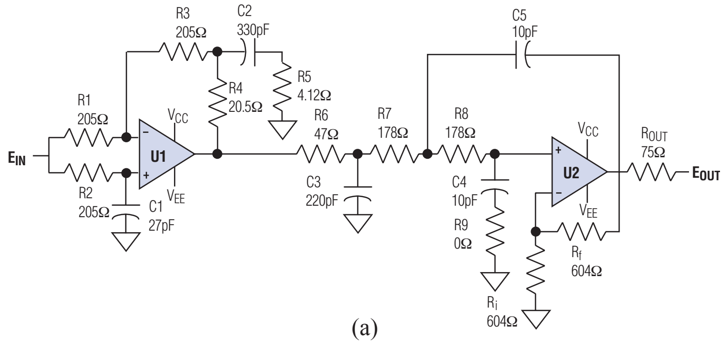
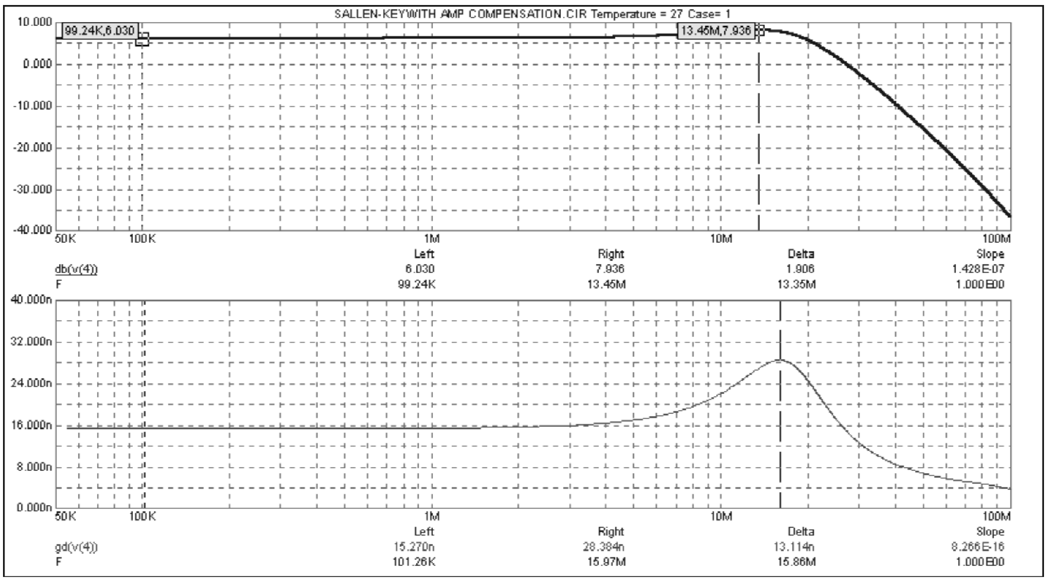


Figure 8. For reconstruction filters with group-delay adjustment, the PAL version (a) has the amplitude and group-delay responses shown in (c), and the NTSC version (b) has the amplitude and group-delay responses shown in (d).



(a)



(b)

Figure 9. This 3-pole, 20MHz Butterworth filter for XGA reconstruction (a) includes (sinx)/x compensation. Its output response is shown in (b).

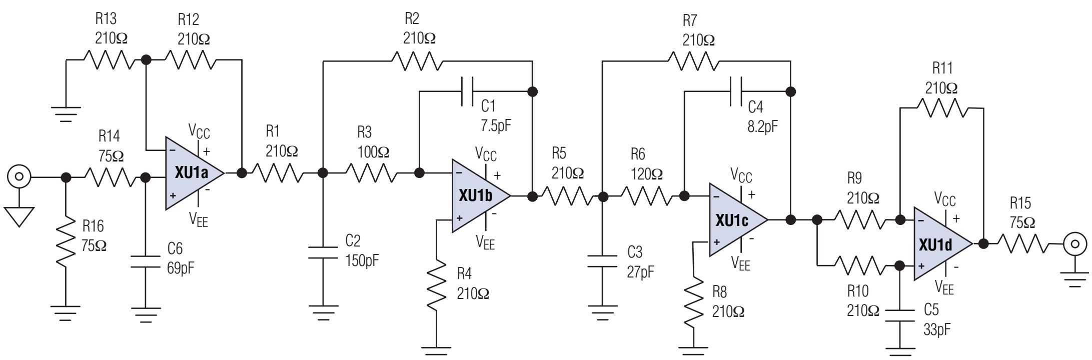


Figure 10. A 5-pole, 30MHz reconstruction filter for HDTV includes amplitude correction for the DAC.

resistors that combine 1% values with 1% tolerance are available. Such components give the best approximation and the most precise amplitude response.

Once built, a filter can be unstable and oscillate. In that case, short the input to ground and see if it continues to oscillate. If it stops, the impedance is too high. Lowering the design impedance should eliminate the oscillation. If it continues, note whether the oscillation is near the filter cutoff frequency or just below. In that case, the oscillation is probably due to components or parasitics. If the oscillation is above the cutoff frequency, it is probably due to the op amp or the circuit layout.

Good layout seems an art, but it is based on a few simple principles. It is important to have a clean supply voltage and a solid ground, meaning filtration with low-ESR capacitors and sometimes with a regulator. The loop formed by the bypass-capacitor connections must be small, or the resulting parasitic inductance will resonate with the capacitance. A good ground plane is essential to good analog design, but as bandwidth increases, it may add parasitic capacitance that can detune the filter. To avoid that problem, remove the ground plane beneath the offending part(s) and traces.

References

[1] A filter response drops by -3dB at the cutoff frequency.

[2] Taylor and Williams, Electronic Filter Design Handbook, McGraw Hill, ISBN 0-07-070441-4.

[3] Ibid.

[4] The 4:2:2 sampling originally indicated the number of times the color subcarrier was oversampled. ITU-601 replaced the subcarrier frequency with 3.375MHz. The 4:2:2 is sampled at 13.5MHz and 6.75MHz, respectively.

[5] For the noninverting case, $R_f/R_i = 0$. For the inverting case, $R_f/R_i = 1$.

[6] E.J. Kennedy, Operational Amplifier Circuits: Theory and Applications.

[7] Defined by H.W. Bode in Network Analysis and Feedback Amplifier Design (D. Van Nostrand, Princeton NJ, 1945).

[8] Taylor and Williams, op. cit.

[9] MAX4450/51 data sheet is available at www.maxim-ic.com.

[10] The Sinc function in mathematics is (sinx)/x.

A similar article appeared in the June 2003 issue of the EDN.