A Digital IF-Modulator for Video and Associated Audio Signals

Dietmar Wenzel, Joachim Speidel

Institute of Telecommunications
University of Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart, Germany

ABSTRACT - **In this paper, a digital intermediate frequency modulator is presented. A composite colour video signal and two associated audio signals are converted into a digital composite TV signal with spectrum at intermediate frequency. The modulator is designed for the high quality requirements that exist in professional equipment for digital CATV head ends. The architecture of the digital modulator shown is optimized for an implementation with ASICs. The signal processing with reduced complexity is described in detail and the simulated results are presented.**

1. INTRODUCTION

The following fundamental requirements have to be met by the digital intermediate frequency modulator (IFM):

- Predistortion of the group-delay of the colour video blanking and synchronizing signal (CVBS signal)
- Vestigial sideband modulation
- An overall passband ripple of less than ±0.1 dB for the CVBS signal
- Signal to noise ratio (SNR) for the video signal of more than 60 dB
- Preemphasis filter for the associated audio signals
- Frequency modulation of the audio signals
- SNR for the audio signal of more than 68 dB
- Output sampling frequency 28.27 MHz
- Spectrum of IF output signal centered at a quarter of the sampling frequency
- IF output signal compatible to the digital CATV distribution system described in [1, 2, 3, 4]
- Architecture with symmetric FIR filters suited for implementation with the ASICs described in [5]

The presented modulator is designed for TV signals according to the colour television standard PAL B/G [6] with two associated audio signals according to the ITU Rec. [7, 8]. However, the signal processing architecture for the CVBS signal is also suited for other colour television systems with a video bandwith equal or less than 5 MHz, e.g. the NTSC system.

The IFM will be used for next generation digital TV distribution. In the CATV head end the output signals of several IFM will be multiplexed together and fed into an SDH optical transmission system (SDH synchronous digital hierarchy). For this professional application the requirements given above are more stringent than [6].

2. OVERVIEW

There are several solutions to generate a digital composite TV signal with spectrum at intermediate frequency. A simple method is to use analog modulation and sampling of the analog TV IF signal. A more sophisticated solution is the application of digital signal processing. The principle operation and the main building blocks of a digital IFM using complex signal processing are shown in Fig. 1.

After clamping, amplifying and analog filtering the CVBS signal is sampled with a rather high sampling frequency of f_S = 28.27 MHz in order to reduce the requirements for the analog anti-aliasing lowpass filter. As the digital IFM will also be used for video transmission in SDH networks, f_S is derived from the clock frequency $f_{SDH} = 155.52$ MHz of the SDH transmission system according to $f_S / f_{SDH} = 2/11$. Investigations of picture quality have shown that 12 bit/sample are sufficient to suppress picture artifacts caused by sampling at a frequency which is not an integer multiple of the TV line frequency. The sampling rate for the audio signals is 47.12 kHz which is 600 times less. It is close to the 48 kHz standard. Therefore high quality low-cost A/D converters can be used.

2.1 Principle of the video signal processing

For the video path in Fig. 1, first a pre-correction of the group-delay is carried out with a filter that has a nonlinear phase (GDPF group-delay pre-correction filter). A vestigial sideband video signal is generated by an asymmetric low pass filter with complex impulse response (CVSF). Another solution using filters with a real impulse response is described later. A real-valued offset is added to the inverted complex-valued video signal. In combination with the following complex mixer oscillating with $\Omega_{IF1} = 2\pi \, f_{IF} / f_S$, a negative amplitude modulation with a complex-valued output signal is obtained. Due to the modulation the offset is converted into the carrier signal. All operations are based on a clock frequency equal to the video sampling rate. In order to reduce the computational overhead a sampling rate conversion in the video processing path is implemented as described in section 3.

2.2 Principle of the audio signal processing

The audio signal processing paths in Fig. 1 are nearly identical. The main tasks are preemphasis, sampling rate conversion and frequency modulation (FM). The audio signals are sampled with at least 14 bit/sample. In order to improve the SNR of the audio signal after demodulation a preemphasis filter (PRF) with a time constant of $\tau_P = 50\,\mu s$ increases the intensity of the higher spectral parts. There are three types of audio signals: mono, stereo and two-channel signals. A sound identification signal (sound-ID) is added to audio signal 2. It depends on the ID whether the stereo matrix is active or not. The system B/G requires two sound carriers [6, 8]. Therefore both audio signals are frequency modulated separately. Before the FM

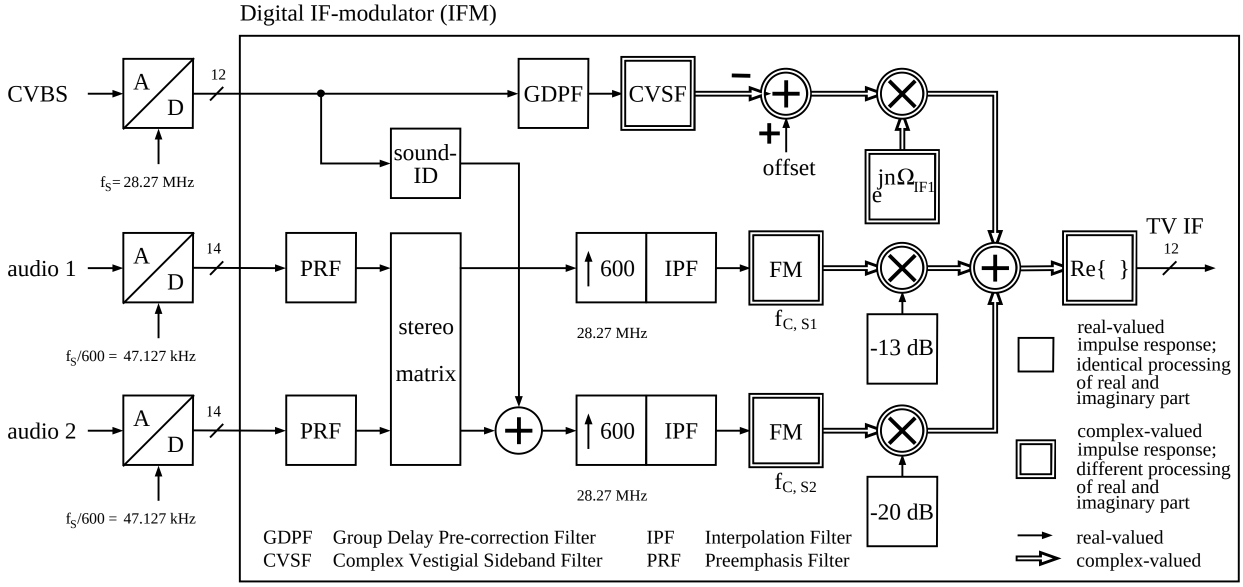

Fig. 1. Main building blocks of the digital IF-modulator

the sampling rate must be increased, because the bandwidth is increased by the modulation. Also an upconversion by factor 600 has to be performed to be able to add the FM audio signals to the video signal. It is not advantageous to carry out this conversion in one go. This is the reason why there are several stages performing the sampling rate conversion as described below. Now both signals are fed into FM modulators. The center frequencies f_C are adjusted so that the corresponding sound carriers are placed in defined distances to the vision carrier. In Fig. 1 the output of the modulators is a complex-valued signal and after an attenuation it is added to the modulated video signal. The real part of the sum is taken and the output is referred to as digital TV IF signal.

3. Video signal processing

This section deals with the architecture which is used to reduce the complexity and the amount of operations in the video path of the IFM. The structure is designed for an implementation with an application-specific integrated circuit (ASIC). This configurable ASIC contains several programmable FIR filters with symmetric impulse responses and a complex-valued numerically controlled oscillator (NCO) [5]. As these elements can be combined in a lot of different ways various signal processing tasks can be implemented. Fig. 2 shows the block diagram and in Fig. 3 the corresponding magnitude responses of the filters for the video signal are shown.

An offset is added to the analog CVBS signal to adapt the signal to the dynamic range of the A/D converter. The decimation filter DECF and a downsampling by factor two yields a decimation of the video signal. The decimation is done to reduce the number of coefficients in the subsequent filters. The i-th coefficient of the impulse response of a filter *x* is denoted as $h_x(i)$. The output signal is

$$b(n) = \sum_{i=0}^{N_{DECF}-1} h_{DECF}(i) \cdot x(2n-i) \qquad (1)$$

m represents discrete time with the higher sampling rate f_S while n corresponds to the lower sampling rate f_S'. The *z*-transform of $x(m)$ and of the impulse response of the FIR filter $h_X(m)$ are given by

$$X(z) = \sum_{m=0}^{\infty} x(m) z^{-m} \text{ and } H_X(z) = \sum_{m=0}^{N_X-1} h_X(m) z^{-m} \,. \qquad (2)$$

The corresponding transfer functions in Fig. 3 are given by

$$H_x(z) = H_x'(f)\,,\ z = e^{j\Omega}\,,\ \Omega = 2\pi\, f / f_S \,. \qquad (3)$$

The *z*-transform of $b(n)$ is

$$B(z') = \frac{1}{2} \sum_{\nu=0}^{1} X\left(z'^{\frac{1}{2}} W_2^{\nu}\right) \cdot H_{DECF}\left(z'^{\frac{1}{2}} W_2^{\nu}\right) \qquad (4)$$

with $$W_2^{\nu} = e^{-j2\pi\nu/2} = e^{j\pi\nu} \qquad (5)$$

where $z' = z^2$. The fat line in Fig. 4 shows the magnitude response of the implemented DECF. The coefficients are quantized with 16 bit each. The thin line indicates the ideal case without quantization. The tolerance scheme is given by the shaded area. The stopband attenuation is greater at the positions where the main spectral parts of the luminance and chrominance signal occur after decimation. The passband ripple, which is only about ±0.02 dB, is drawn enlarged.

The next processing task in Fig. 2 is the pre-correction of the group-delay with an FIR filter (GDPF) which has an unsymmetric impulse response. As a consequence of their inherent stability, their small contribution to the noise and the ability of linear phase only FIR filters with fixed point arithmetic are used in the IFM. In order to implement the GDPF with the same ASICs used for all other symmetric filters the impulse response is split into its even and odd part. If the symmetric property of the two partial impulse

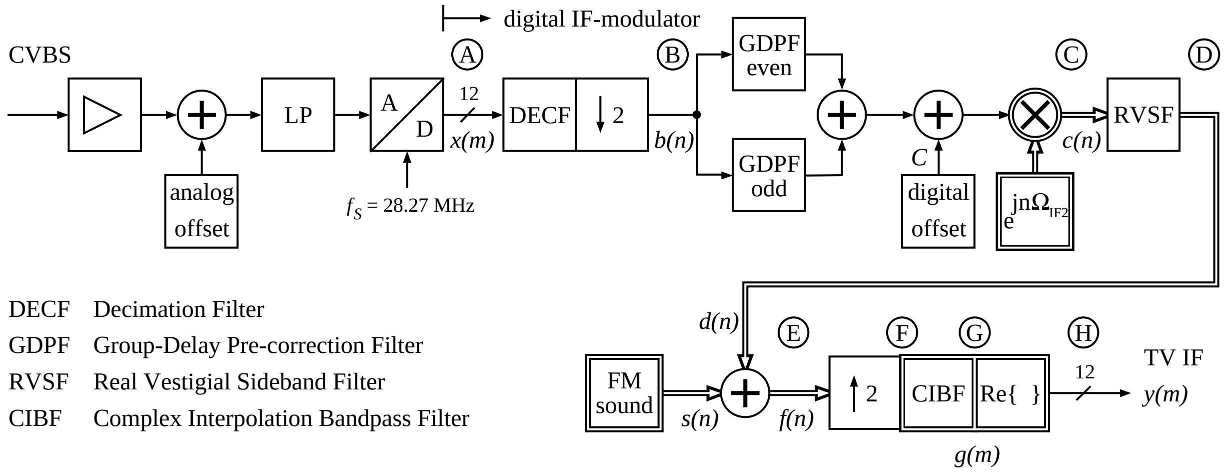

Fig. 2. Block diagram for video signal processing within the digital IF-modulator

responses is used to reduce the number of multiplications in each partial filter the total number of multiplications in the entire filter is equal to the number of coefficients in the GDPF. For this reason the split of the GDPF in even and odd part causes almost no additional computational overhead. The magnitude response in dB and the group-delay τ_{GDPF} in ns is shown in Fig. 5. The group-delay is compliant to the tolerance scheme given in [6]. Combining the DECF and the GDPF in a single unsymmetric filter would require a considerably higher number of coefficients. Fig. 5 shows an increase of the magnitude response outside the passband. This is an effect caused by the design process of the filter and the maximal number of available coefficients but nevertheless the overall magnitude response of the entire IFM is in the given tolerance scheme as will be shown in section 6, Fig. 13.

In order to perform a negative vestigial sideband modulation the signal has to be inverted and scaled. A detailed investigation which is not shown here results in a scaling factor x and an offset C to be

$$x = -\frac{y_2}{y_1} \cdot 0.616 \text{ and } C = A \cdot y_2 \cdot 0.508 . \qquad (6)$$

A denotes full scale and the maximum signal level before scaling is $y_1 A$ and the maximum signal level after scaling is $y_2 A$. The scaling can be done easily by modifying the coefficients of the GDPF. Simulation of the IFM has shown that a maximum SNR is achieved if

$$20 \lg y_1 = -1\,\mathrm{dB} \text{ and } 20 \lg y_2 = -3\,\mathrm{dB} \qquad (7)$$

are used. The frequency-shifting is now done with the help of a complex-valued carrier that is provided by a digital os-

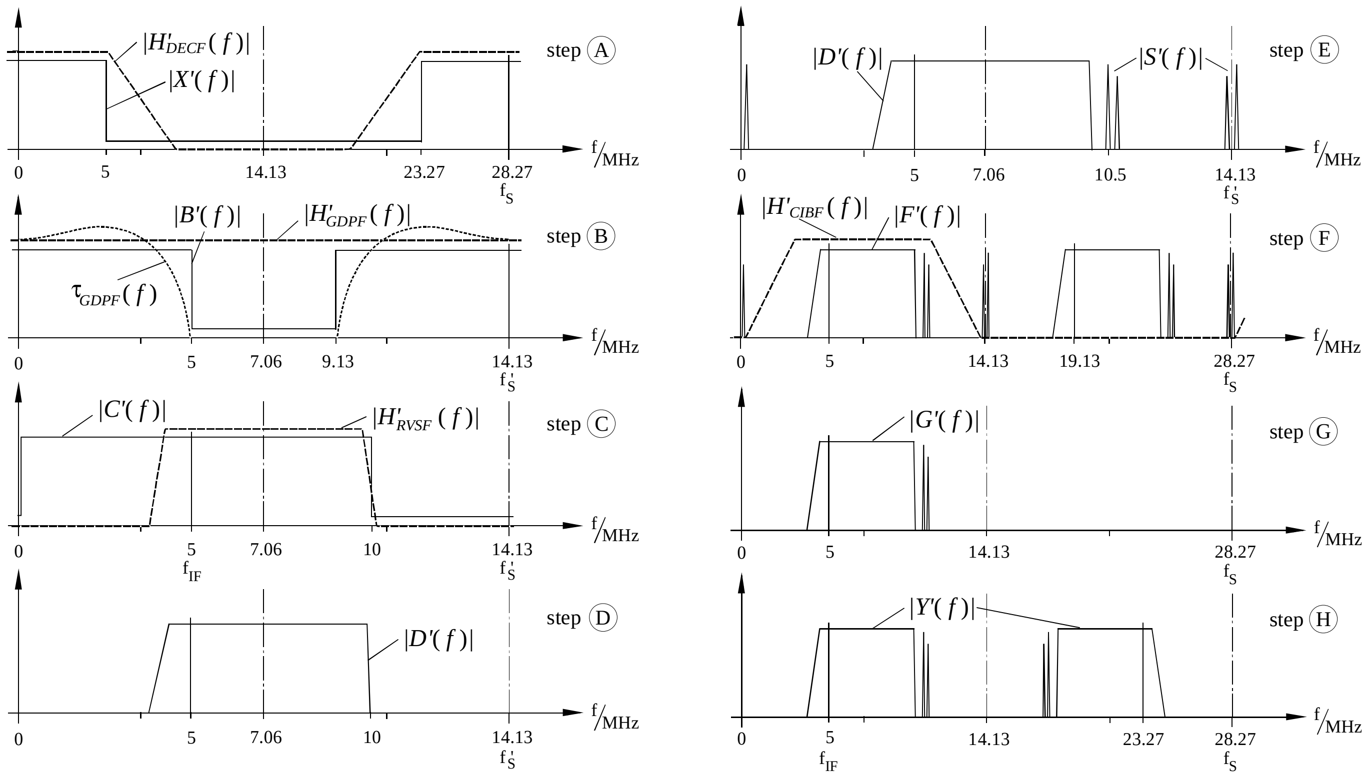

Fig. 3. Video signal processing within the digital IF-modulator in frequency domain (principles); symbols correspond to (3)

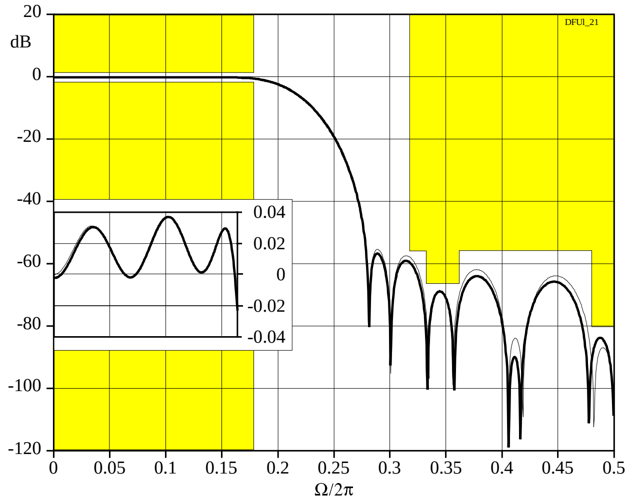

Fig. 4. Magnitude response of the DECF with 28 coefficients

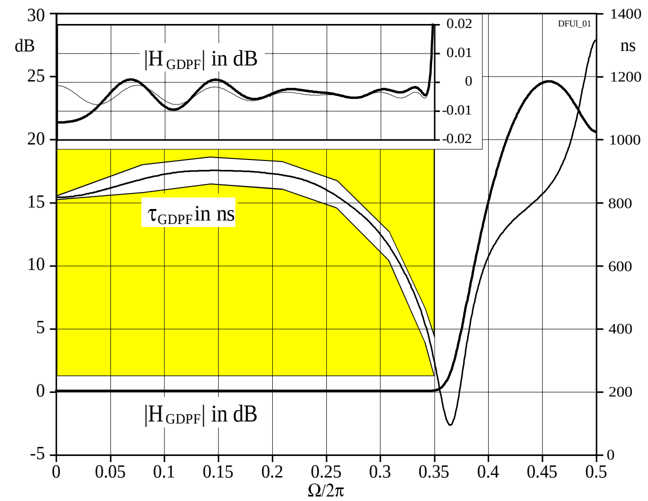

Fig. 5. Magnitude response and group-delay of the GDPF with 24 coefficients

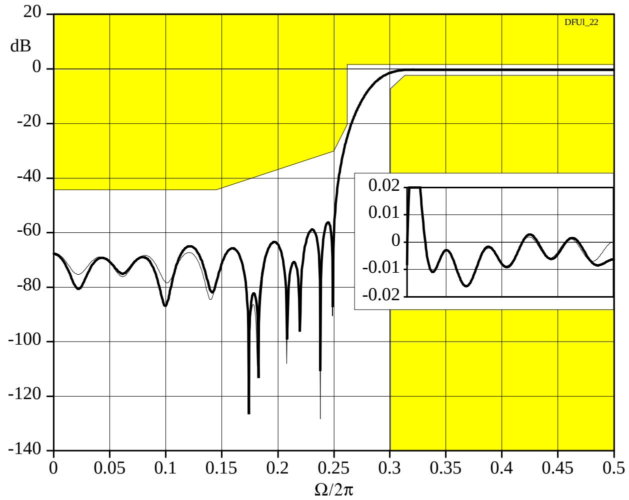

Fig. 6. Magnitude response of the RVSF with 51 coefficients

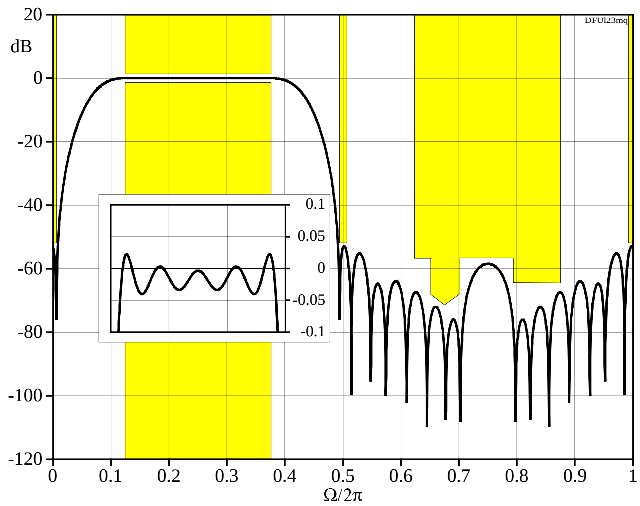

Fig. 7. Magnitude response of the CIBF with 27 coefficients

cillator with

$$\Omega_{IF2} = 2\pi \cdot 4.994\,\text{MHz}/f_S' . \tag{8}$$

The specifications for the vestigial sideband filter in [6] allow the combination of the requirements in the tolerance scheme for the lower and upper edge frequency of the unsymmetric bandpass filter in a tolerance scheme for a symmetric bandpass filter. In the discrete time domain a highpass filter can also fulfill the function of this specific bandpass filter. The carrier frequency in (8) was chosen to place the signal spectrum at a position where the real vestigial sideband filter (RVSF) can be a simple highpass filter with real coefficients as shown in Figs. 2, 3C and 6, so the same hardware that is also suited for processing complex-valued signals can be used for the DECF, GDPF and RVSF.

In Fig. 2 the FM sound signal $s(n)$ which is discussed in section 5 is added. The magnitude response of $S'(f)$ is shown in Fig. 3E. It also contains spectral parts near to f_S' and $f = 0$. The resulting signal is

$$\begin{aligned} f(n) = &\sum_{i=0}^{N_{DECF}-1} \sum_{j=0}^{N_{GDPF}-1} \sum_{k=0}^{N_{RVSF}-1} \\ &\left\{ h_{DECF}(i)\, h_{GDPF}(j)\, h_{RVSF}(k)\, x\big(2(n-j-k)-i\big) \right. \\ &\left. + C \cdot h_{RVSF}(k) \right\} e^{j(n-k)\Omega_{IF2}} + s(n) \end{aligned} \tag{9}$$

Applying the modulation theorem of the *z*-transform and the correspondence

$$Z\{C\} = C \frac{z'}{z'-1} \tag{10}$$

to (9), the *z*-transform of the TV signal $f(n)$ can be written as

$$\begin{aligned} F(z') = &\left\{ B\left(z' \cdot e^{-j\Omega_{IF2}}\right) \cdot H_{GDPF}\left(z' \cdot e^{-j\Omega_{IF2}}\right) \right. \\ &\left. + C \frac{z' \cdot e^{-j\Omega_{IF2}}}{z' \cdot e^{-j\Omega_{IF2}} - 1} \right\} \\ &\cdot H_{RVSF}(z') + S(z') \end{aligned} \tag{11}$$

Finally an interpolation by factor two is applied to the signal $f(n)$ to return to the initial sampling frequency f_S and to compute the real-valued IF TV signal (Fig. 2). As the video signal was frequency-shifted with (8), its spectrum is centered around ¼ f_S after interpolation. The interpolation filter (CIBF) must be a bandpass filter to select the spectral parts centered at ¼ f_S. At the same time the undesired parts of the sound signals at ½ f_S and $f = 0$ must be attenuated (see Figs. 3F and 7). The TV IF output signal is

$$y(m) = \mathrm{Re}\left\{ \sum_{l=0}^{N_{CIBF}-1} h_{CIBF}(l) \cdot f'(m-l) \right\}$$
$$\text{with } f'(m) = \begin{cases} f\left(\frac{m}{2}\right) & ,m = 2n \\ 0 & ,m = 2n+1 \end{cases} . \tag{12}$$

In z domain we obtain for the complex-valued signal after interpolation

$$G(z) = F\left(z^2\right) \cdot H_{CIBF}(z) \tag{13}$$

and finally the z-transform of the TV IF signal is

$$Y(z) = Z\left\{\mathrm{Re}\left\{g(m)\right\}\right\} = \frac{1}{2}\left(G(z) + G^*\left(z^{-1}\right)\right). \tag{14}$$

The magnitude response is shown in Fig. 3H. For further processing of $y(m)$ either the spectral part in original or in mirrored position can be used.

4. Filter design

Whereas the DECF, the RVSF and the CIBF with constant group-delay can be designed applying the well-known method of Parks/McClellan, the GDPF necessitates another design method.

4.1 Design of the GDPF

As is well known, the relation between the group-delay $\tau(\Omega)$ and the phase $\varphi(\Omega) = \mathrm{arc}\, H(\Omega)$ of the transfer function $H(\Omega)$ is given by

$$\varphi(\Omega) = -\int \tau(\Omega)\, d\Omega + \Phi \tag{15}$$

where Φ is a constant that is used to obtain a causal impulse response. Let $H(\Omega)$ be specified by a vector of M complex values in the frequency range of interest

$$\mathbf{H} = \begin{bmatrix} H_0 & H_1 & \dots & H_{M-1} \end{bmatrix}^T \text{ with} \tag{16}$$

$$H_i = \left|H(\Omega_i)\right| \cdot e^{j\varphi(\Omega_i)} \ , \ \Omega_i = 2\pi \frac{f_i}{f_S} \quad i = 0 \dots M-1 . \tag{17}$$

The task is to approximate $H(\Omega)$ by an FIR filter with N coefficients. The transfer function $I(\Omega)$ of this filter is determined at the M ($M > N$) frequency points of interest by

$$\mathbf{I} = \begin{bmatrix} I_0 & I_1 & \dots & I_{M-1} \end{bmatrix}^T \tag{18}$$

$$\text{with} \qquad I_i = \sum_{n=0}^{N-1} h_n \cdot e^{-jn\Omega_i} \quad i = 0 \dots M-1 . \tag{19}$$

The approximation error is defined as

$$\mathbf{e} = \mathbf{I} - \mathbf{H} . \tag{20}$$

If $\mathbf{I}$ is written in matrix form

$$\mathbf{I} = \mathbf{W} \cdot \mathbf{h} \text{ with} \tag{21}$$

$$\mathbf{W} = \begin{bmatrix} 1 & e^{-j\Omega_0} & \dots & e^{-j(N-1)\Omega_0} \\ 1 & e^{-j\Omega_1} & \dots & e^{-j(N-1)\Omega_1} \\ \dots & \dots & \dots & \dots \\ 1 & e^{-j\Omega_{M-1}} & \dots & e^{-j(N-1)\Omega_{M-1}} \end{bmatrix} \text{ and } \mathbf{h} = \begin{bmatrix} h_0 \\ h_1 \\ \dots \\ h_{N-1} \end{bmatrix}$$

the following system of linear equations is obtained

$$\mathbf{e} = \mathbf{W} \cdot \mathbf{h} - \mathbf{H} . \tag{22}$$

A good approximation is achieved by minimizing

$$e_{\min} = \min_{\mathbf{h}} \left\{ \mathbf{e}^* \mathbf{e} \right\} = \min_{\mathbf{h}} \left\{ (\mathbf{W} \cdot \mathbf{h} - \mathbf{H})^* \cdot (\mathbf{W} \cdot \mathbf{h} - \mathbf{H}) \right\} \tag{23}$$

where ‘*’ denotes the complex conjugated and transposed matrix or vector, respectively. With the matrix

$$\mathbf{R} = \mathbf{W}^* \cdot \mathbf{W} = \mathbf{R}^* \tag{24}$$

we obtain from (23)

$$e_{\min} = \min_{\mathbf{h}} \left\{ \mathbf{H}^*\mathbf{H} - \mathbf{H}^*\mathbf{W}\mathbf{R}^{-1}\mathbf{W}^*\mathbf{H} + \left(\mathbf{W}^*\mathbf{H} - \mathbf{R}\mathbf{h}\right)^* \mathbf{R}^{-1} \left(\mathbf{W}^*\mathbf{H} - \mathbf{R}\mathbf{h}\right) \right\} \tag{25}$$

which can be proven using $(\mathbf{X} \cdot \mathbf{Y})^* = \mathbf{Y}^* \cdot \mathbf{X}^*$. Since $e_{\min}$ is always positive the minimum is given by setting the quadratic term in (25) containing $\mathbf{h}$ to zero. The result is

$$\mathbf{h} = \left(\mathbf{W}^* \cdot \mathbf{W}\right)^{-1} \cdot \mathbf{W}^* \cdot \mathbf{H} . \tag{26}$$

The resulting transfer function $I(\Omega)$ depends on the proper selection of the arrangement of the M frequency points. The result which meets the requirements is shown in Fig. 5.

4.2 Design of the CIBF

In case of the CIBF a convolution of two complex-valued signals must be carried out. Normally this requires four real-valued convolutions simultaneously. But if an FIR filter is used that has the following properties:

- magnitude response with symmetry about the axes at ¼ f_S and ¾ f_S (Fig. 7),
- odd number of coefficients,
- linear phase,
- interpolation by factor two,
- only real-valued output,

the CIBF can be split into two polyphases with real-valued symmetric impulse responses $h_0(n)$ and $h_1(n)$ as shown in Fig. 8. Only one filter operation for each output sample must be computed [10].

5. Audio signal processing

The audio signal processing will now be discussed in detail for one audio channel. In Fig. 9 the various stages for the

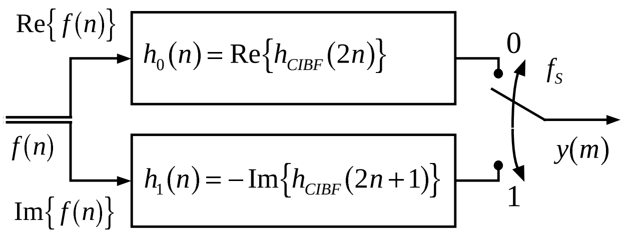

Fig. 8. Filter structure for the CIBF

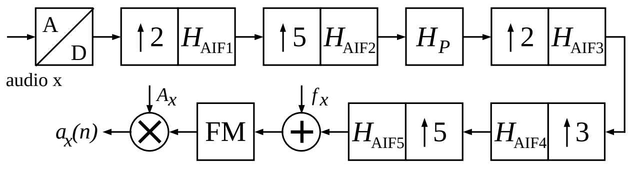

Fig. 9. Block diagram for audio signal processing

sampling rate conversion to f_S' are presented. The total interpolation factor of 300 is split into its prime factors. This reduces the number of coefficients and operations per clock cycle. The first anti-imaging filter (AIF1) is a half-band filter with 23 coefficients. The AIF2 is a fifth-band filter where every fifth coefficient is zero except for the middle coefficient. The magnitude response is shown in Fig. 10. This kind of filter can be implemented in a structure shown in Fig. 11 with high efficiency. The values for the multipliers in the filter structure are derived from the filter coefficients according [9]. AIF3, AIF4 and AIF5 have 15 coefficients each.

The pre-emphasis filter H_P is placed after the second interpolation stage considering that the implementation of the preemphasis filter is less complex if the sampling frequency is about 10 times higher than the initial sampling frequency of the audio signals. The stereo matrix and the sound-ID signal for audio channel two are not shown in Fig. 9.

In contrast to Fig. 1 the FM for the two audio signals is only real-valued because this requires only two real-valued digital oscillators. As is well known the FM is achieved by manipulating the phase accumulator of the digital oscillator every clock cycle. The frequency swing of 50 kHz is adjusted by scaling the coefficients in one AIF. The FM signals $a_x(n)$ (x = 1, 2) are attenuated by multiplication with the constant A_x to achieve the desired level of the sound carriers [7].

Since the signals $a_x(n)$ are real-valued it is impossible to carry out the FM with the desired sound carrier frequencies of about 10.5 MHz that are shown in Fig. 3E directly, because the mirrored spectral parts would overlap with the spectrum of the video signal. For this reason the two real-valued frequency-modulated sound signals are placed at intermediate frequencies f_x which are 1/8 of the actual sampling frequency f_S' away from this sound carriers. These specific intermediate frequencies are chosen because the mirrored spectral parts of the real-valued sound signals are placed at positions where the distance from the spectrum of the video signal is maximal after the frequency-shifting according to (27) (Fig. 3E).

$$s(n) = [a_1(n) + a_2(n)] \cdot e^{-jn\pi/4} \qquad (27)$$

The frequency-shift is carried out with the structure in Fig. 12 which needs only one multiplier, an inverter and few multiplexers. It can easily be implemented with an FPGA. Since the signal $s(n)$ is clipped to a word length of 12 bit/sample before adding to the modulated video signal, the SNR is almost only limited by this clipping and not by the noise produced during audio signal processing.

6. Performance

If the symmetry of the filters is considered the total number of multiplications and additions per clock cycle $1/f_S$ for the video signal processing is 62 and 114.5, respectively. This results in a total of 4.9 billon operations per second (GOPS). The audio signal processing requires about 7 multiplications and 12 additions per the same clock cycle. The total is 0.52 GOPS which is about 10% of the total processing. The overall magnitude response for the video signal is shown in Fig. 13 (bottom). The passband ripple is ±0.03 dB if the complex-valued IF TV signal $g(m)$ is looked at after the CIBF. The passband ripple of the IFM is only limited by the number of coefficients. The simulated signal spectrum is shown in Fig. 13 (top). An SNR of 65 dB is achieved for the video signal, 73 dB for audio signal one and 70 dB for audio signal two [10].

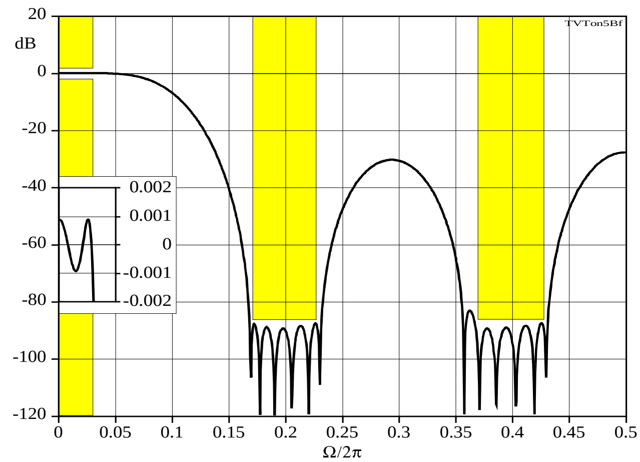

Fig. 10. Magnitude response of fifth-band filter AIF2

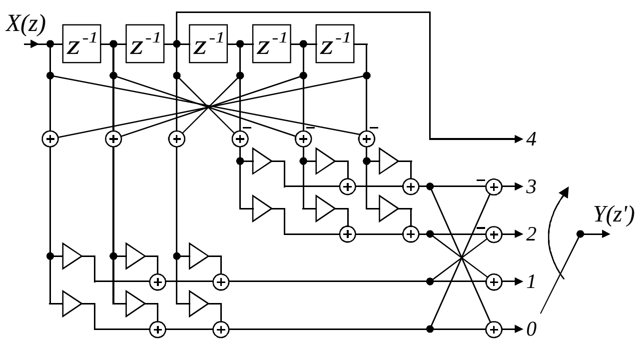

Fig. 11. Polyphase fifth-band filter with 29 coefficients

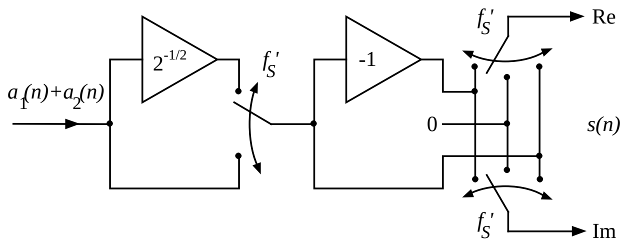

Fig. 12. Mixer for FM sound signals

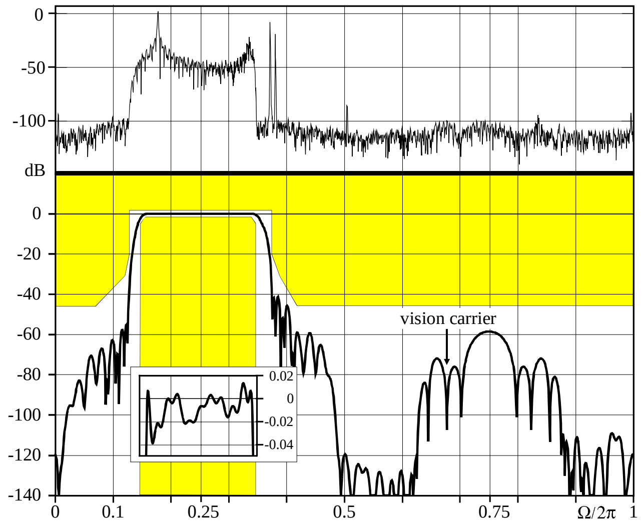

Fig. 13. TV IF signal spectrum including sound (top), Overall magnitude response for the video signal (bottom)

7. Conclusion

This paper presents an efficient method for the conversion of PAL video signals with two associated audio signals into a digital IF TV signal using digital signal processing techniques. It is shown that all required filter operations can be carried out with identical FIR filter structures that have a symmetrical impulse response to simplify the hardware complexity. In addition, an overall passband ripple of less than ±0.03 dB is better than required in the ITU specification. The proposed system architecture is a further step towards an approach to an all digital TV transmission systems for feeding and distributing.

References

[1] W. Bambach, G. Berthes, H. Lohscheller, J. Mayer, "Digital optical trunk system for CATC networks", International Television Symposium Montreux, 1997.

[2] W. Bambach, K. Grotz, "First installation of a novel all-digital access network system for distributive and interactive services", Optoelectronics and Communications Conference, Tokyo, 1996.

[3] H. G. Göckler, K. Grotz, "Diamant - All digital frequency division multiplexing for 10 Gbit/s fibre-optic CATV distribution systems", European signal processing Conference (EUSIPCO), Edinburgh, pp. 999-1002, Sep. 1994.

[4] H. Feilhauer, K. Grotz, J. Speidel, "All digital optical CATV distribution system", Proc. ITG-Workshop Optische Teilnehmerzugangsnetze, Dresden, 1995.

[5] "ASICs DFUlong and DFUint, Documentation", Bosch Telecom, Backnang, 1997.

[6] ITU-R Rec. BT.470-5 "Conventional television systems", Genf, 1998.

[7] ITU-R Rec. BS.707-3 "Transmission of multisound in terrestrial television systems PAL B, G, H and I", Genf, 1998.

[8] CCIR Rep. 793-3 "Transmission of two or more sound programs or information channels in television", Genf, 1990.

[9] Z.-J. Mou, "Symmetry exploitation in digital interpolators/decimators", IEEE Transactions on Signal Processing, Vol. 44, no. 10, pp. 2611-2619, Oct. 1996.

[10] H. Hofmann, "Design of a digital IF-Modulator", Diploma Thesis 1890 D, Institute of Telecommunications, University of Stuttgart, 1997.