

FORMATOS DE VÍDEO DIGITAL

PTC3547 - CODIFICAÇÃO E TRANSMISSÃO MULTIMÍDIA

Guido Stolfi – EPUSP – 04/2023

1. Vídeo Digital × Analógico

Um sinal é denominado *analógico* quando reproduz uma informação através da variação de algum parâmetro, de forma *análoga* à informação original. Um microfone, por exemplo, produz um sinal elétrico cuja tensão V_o é análoga à pressão sonora incidente P_s .

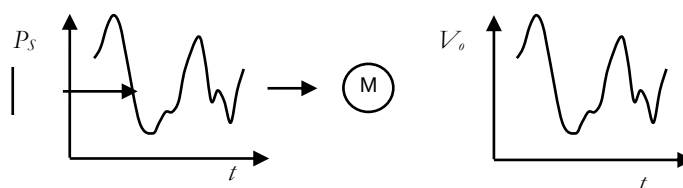


Fig. 1.1 – Sinal Analógico em um Microfone

Como tal, um sinal analógico pode assumir infinitos valores, em infinitos instantes de tempo, dentro de determinados limites. Podemos dizer então que o sinal analógico equivale a uma função contínua $y = f(t)$, definida para todos os valores da variável contínua t .

Por outro lado, um sinal *digital* (do latim *digitus* = dedos) pode assumir apenas valores pertencentes a um conjunto finito de símbolos (p. ex. “0” ou “1” para um sinal binário), e esses valores são atribuídos apenas em determinados instantes dentro de um intervalo de tempo. O sinal digital, portanto, equivale a uma sequência enumerável de símbolos discretos, ou, em outras palavras, a uma sequência de números inteiros.

O processo denominado “digitalização”, pelo qual um sinal analógico é convertido para o domínio digital, compreende assim dois processos independentes, que serão detalhados adiante:

- **Amostragem:** o sinal é representado por valores tomados em determinados instantes de tempo, geralmente de forma periódica;
- **Quantização:** os valores das amostras são arredondados para um número inteiro mais próximo.

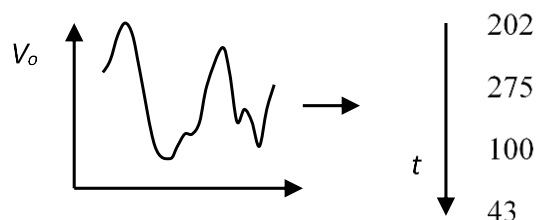


Fig. 1.2 - Digitalização

1.1 Degradações do Sinal Analógico

Uma das principais motivações que levam as técnicas das Telecomunicações em direção ao domínio digital deve-se ao fato de que todo sinal analógico está sujeito a 4 categorias de degradações (Fig. 1.3), que o afetam em maior ou menor grau, sempre que é transportado, processado ou armazenado:

- **Ruído:** aleatório, cumulativo, inerente aos processos físicos envolvidos na circulação dos sinais elétricos;
- **Distorção não-linear:** presente em todos os elementos ativos de circuitos, amplificadores, processadores, etc.;
- **Interferências:** ingresso de sinais indesejados, geralmente de origem conhecida; eventualmente identificáveis, porém de difícil ou impossível eliminação;
- **Perdas:** atenuação, limitação de resposta em frequência, dispersão temporal, defasagem, e outras degradações decorrentes da função de transferência do canal de comunicação.

A codificação digital proporciona muitas vantagens no tratamento e distribuição de sinais de vídeo: além da imunidade às degradações listadas acima, os equipamentos que operam no domínio digital podem realizar, de forma eficiente, funções de processamento de imagem que seriam praticamente impossíveis de executar na forma analógica.

Entre estas funções encontram-se, por exemplo, transformações espaciais e a extração das componentes visualmente relevantes de uma imagem, o que é fundamental para a implementação de processos de compressão de dados.

Uma das primeiras aplicações de processamento digital em TV foi na implementação de *corretores de base de tempo* (“*Time Base Corrector*”, ou TBC), que consistem de linhas de retardo variáveis, usadas para sincronizar o sinal de vídeo na gravação magnética (VT).

À medida que o custo e o desempenho dos circuitos digitais, especialmente dos conversores A/D, foram tornando-se mais vantajosos, foram desenvolvidas outras aplicações de processamento e efeitos visuais em estúdios, como por exemplo:

- **"Frame Store":** armazenamento digital de vídeo em tempo real, para permitir congelamento de imagem, câmera lenta e “*replay*”;
- **"Frame Synchronizers":** sincronizadores de sinais de vídeo com capacidade de compensar defasagens de até 1 quadro completo;
- **Inserção de Caracteres:** para superposição de legendas e informações gráficas;

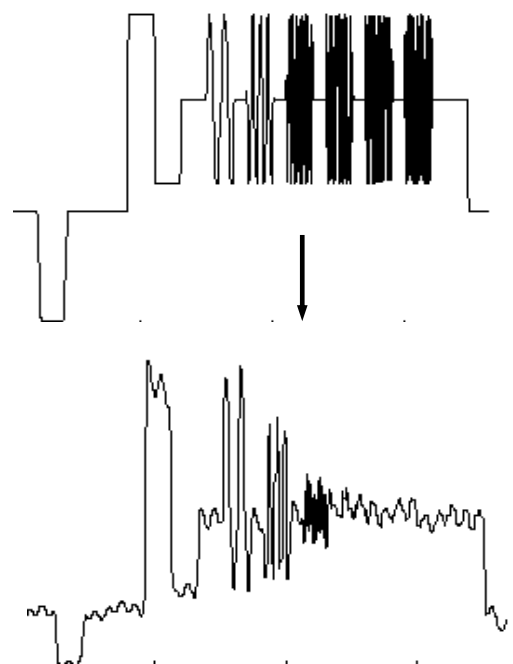


Fig. 1.3 – Degradações de um Sinal Analógico

-
- **Conversores de padrões:** efetuam transcrição entre sistemas de 50 e 60 Hz, implementados através de interpoladores e decimadores espaço-temporais;
 - **Equipamentos de telecinagem:** convertem material cinematográfico, de 24 fotogramas para 60 campos por segundo, através de interpolação;
 - **Efeitos de transformações geométricas:** onde é feito mapeamento espacial de imagens, em tempo real, com mudança de coordenadas, deformações, perspectivas, etc.;
 - **Redutores de ruído:** baseados em filtragem temporal entre linhas e quadros consecutivos;
 - **Gravadores de vídeo digitais:** através da aplicação de códigos corretores de erros, evitam degradação cumulativa da relação sinal/ruído, mesmo que o sinal sofra sucessivas gerações (cópias);
 - **Edição não-linear:** ao invés de utilizar dois ou mais gravadores de VT para finalizar a montagem de programas (pós-produção), utilizam-se meios de armazenamento digital com acesso randômico ao nível de quadros individuais, facilitando o processo de edição.
 - **Servidores de Vídeo:** Sistemas de armazenamento em massa de vídeos compactados, que disponibilizam um acervo de imagens simultaneamente para edição, produção e veiculação.

Finalmente, os avanços das técnicas de compressão de dados, aplicáveis a sequências de imagens em movimento, bem como de códigos de correção de erros e processos eficientes de modulação, tornaram viável a radiodifusão de programas de TV em formato digital.

Hoje existem inúmeros sistemas digitais em operação, seja por satélite, cabo ou radiodifusão terrestre, apresentando vantagens em relação aos canais analógicos, tanto na qualidade de imagem (pela imunidade a ruídos), como na variedade de programas disponíveis (pelo aumento da capacidade do canal de comunicação, conseguido pela compressão), além de permitir a interoperabilidade com serviços adicionais de comunicação de dados.

A aplicação dos processos digitais na compressão, transmissão e radiodifusão de sinais multimídia será detalhada oportunamente em outros capítulos; neste texto vamos nos restringir aos conceitos de amostragem, quantização e processamento digital de sinais de vídeo.

2. Amostragem

O processo de varredura, empregado na TV analógica, consiste de duas amostragens: *temporal* (separação em quadros/campos) e *espacial* na dimensão vertical (divisão da imagem em linhas). A digitalização do sinal de vídeo necessita de mais uma etapa de amostragem, realizada na dimensão horizontal. Disto resulta uma decomposição em **elementos de imagem**, ou **pixels** (*picture elements*), onde cada *pixel* pode ser descrito numericamente pelas amplitudes de suas componentes de luminância e crominância (por exemplo, Y , $R-Y$, $B-Y$) ou primárias (R , G , B).

Em geral (mas não sempre), um sistema amostrado segue o diagrama da fig. 2.1. A função do filtro “anti-aliasing” é evitar a ocorrência de rebatimento espectral, que surge quando a frequência de amostragem é menor que o dobro da maior frequência contida no sinal de entrada (conforme critério de Nyquist / Shannon / Kotelnikov).

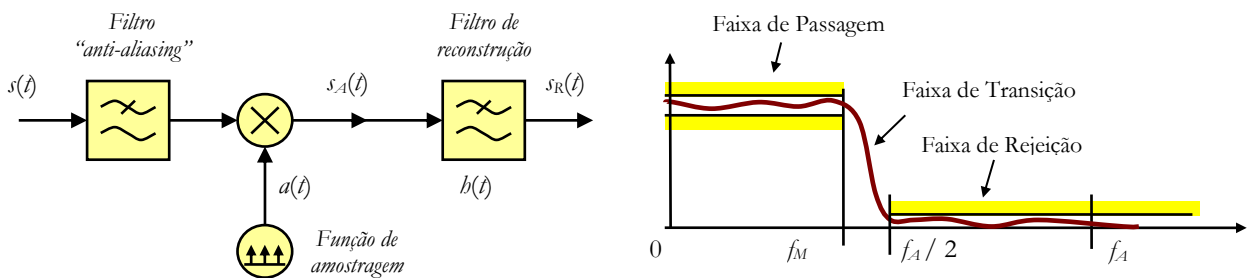


Fig. 2.1 – Sistema Amostrado e Filtro “Anti-aliasing”

2.1 Critério de Kell

O critério de Nyquist representa uma situação limite, que só pode ser atingida com o uso de um filtro ideal. Nos casos em que o filtro de reconstrução não pode ser especificamente implementado, como ocorre frequentemente em sistemas de visualização de imagem, esse critério não proporciona resultados satisfatórios.

A figura 2.2 apresenta um exemplo no qual o critério de Nyquist é atendido ($f_A = 2,2 f_M$) e, no entanto, devido ao uso de uma função de reconstrução retangular, a supressão das componentes espectrais replicadas não é satisfatória, resultando em batimentos, perda de contraste e ambiguidade nas passagens por zero.

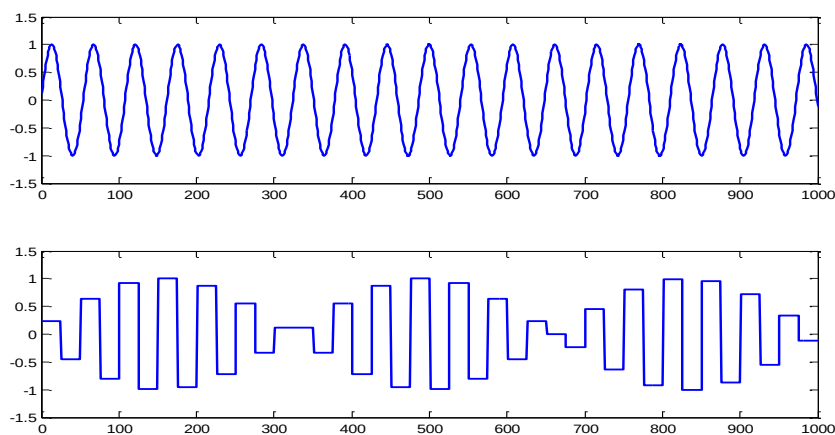


Fig. 2.2 – Amostragem Dentro do Critério de Nyquist ($f_A = 2,2 f_M$)

Em 1934, o engenheiro Raymond D. Kell (RCA) realizou pesquisas tentando determinar a resolução máxima que poderia ser obtida em um sistema de varredura, aplicado à televisão. O “fator de Kell”, obtido experimentalmente, define a relação entre a frequência de amostragem teórica obtida pelo critério de Nyquist e a frequência mínima viável (para reduzir os efeitos de batimento). Corresponde assim ao inverso da ”margem de segurança” aplicada sobre o critério de Nyquist. Valores experimentais estão na faixa de 0,64 a 0,8, e dependem das funções de reconstrução intrínsecas dos processos de reprodução de imagem utilizados.

Podemos combinar o “fator de Kell” com o critério de Nyquist, chegando a um “critério de Kell” (empírico): em valores aproximados, $f_A \geq 3 f_M$.

A figura 2.3 apresenta um sinal amostrado com $f_A = 3,3 f_M$ e reconstruído com pulso retangular. Podemos notar que a ocorrência de máximos e mínimos é preservada para todos os ciclos do sinal original, bem como as passagens por zero. Por causa disso, a amostragem de sinais de vídeo em geral é feita considerando o “critério de Kell”.

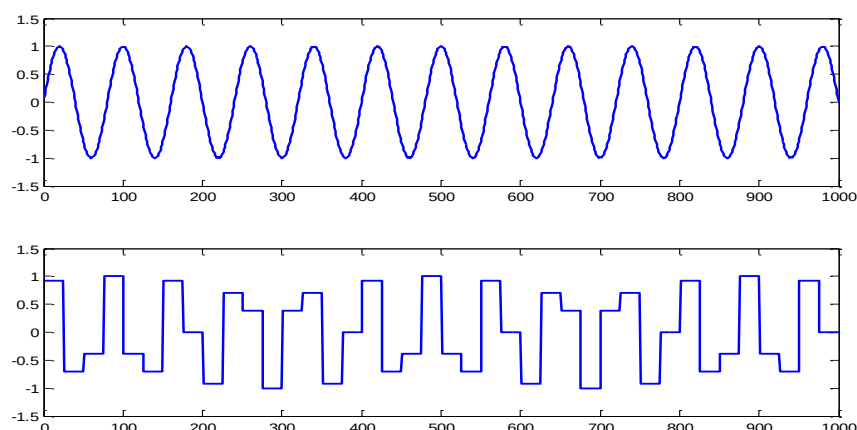


Fig. 2.3 – Amostragem Dentro do Critério de Kell ($f_A = 3,3 f_M$)

2.2 Estruturas de Amostragem Espacial

Considerando que a resolução espacial da visão humana é muito menor para as componentes de crominância do que para a luminância, vários sistemas exploram essa característica utilizando estruturas espaciais diferentes para esses sinais. Essas estruturas utilizam decimação espacial no sentido horizontal e/ou vertical, assumindo que o receptor aplicará uma interpolação inversa para restaurar as amostras primárias (Fig. 2.4).

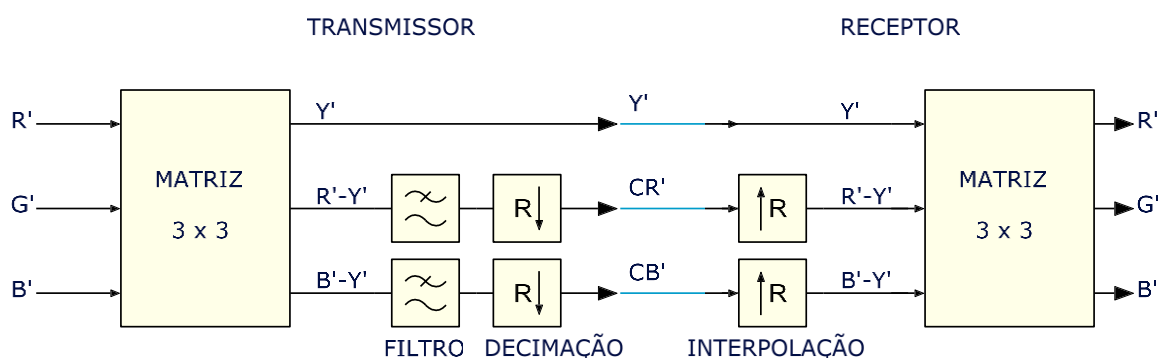


Fig. 2.4 – Amostragem com Resolução Diferenciada

Em geral, a denominação dada a essas estruturas indica a proporção relativa das amostras de luminância (Y) e de crominância (C_R e C_B). A figura 2.5 apresenta algumas estruturas usadas em sistemas de vídeo digital; nessa figura, as amostras espaciais de luminância são representadas por pontos pretos e as amostras de crominância por círculos.

A estrutura “4:4:4”, utilizada em equipamentos profissionais com alta qualidade de imagem, efetua a amostragem das componentes de luminância e crominância nas mesmas posições espaciais, para todos os elementos de imagem. Exemplos dessa estrutura são as interfaces RGB de placas de vídeo, interfaces DVI e HDMI, e processadores de vídeo profissionais.

A estrutura “4:2:2” é utilizada nas interfaces baseadas nas recomendações CCIR-601 e SMPTE-274, nas quais as componentes de crominância (C_R e C_B) são amostradas a cada 2 amostras de luminância (Y).

A estrutura “4:1:1” foi usada em equipamentos digitais domésticos, como o formato de gravação magnética DV-25.

Os padrões de compressão de vídeo MPEG-1 e MPEG-2 usam disposição espacial diferente para as estruturas de amostragem, denominadas “4:2:0”, nas quais é feita interpolação vertical e horizontal das componentes de crominância. Já na estrutura “4:2:0” utilizada em um formato DV, as amostras de crominância (C_R e C_B) ocorrem em linhas alternadas.

Note-se que, para um sistema digital, não é necessário amostrar o sinal de vídeo durante os intervalos de retraço vertical e horizontal, o que proporciona uma economia de cerca de 24% no volume de informação, para o caso da TV no formato “M”.

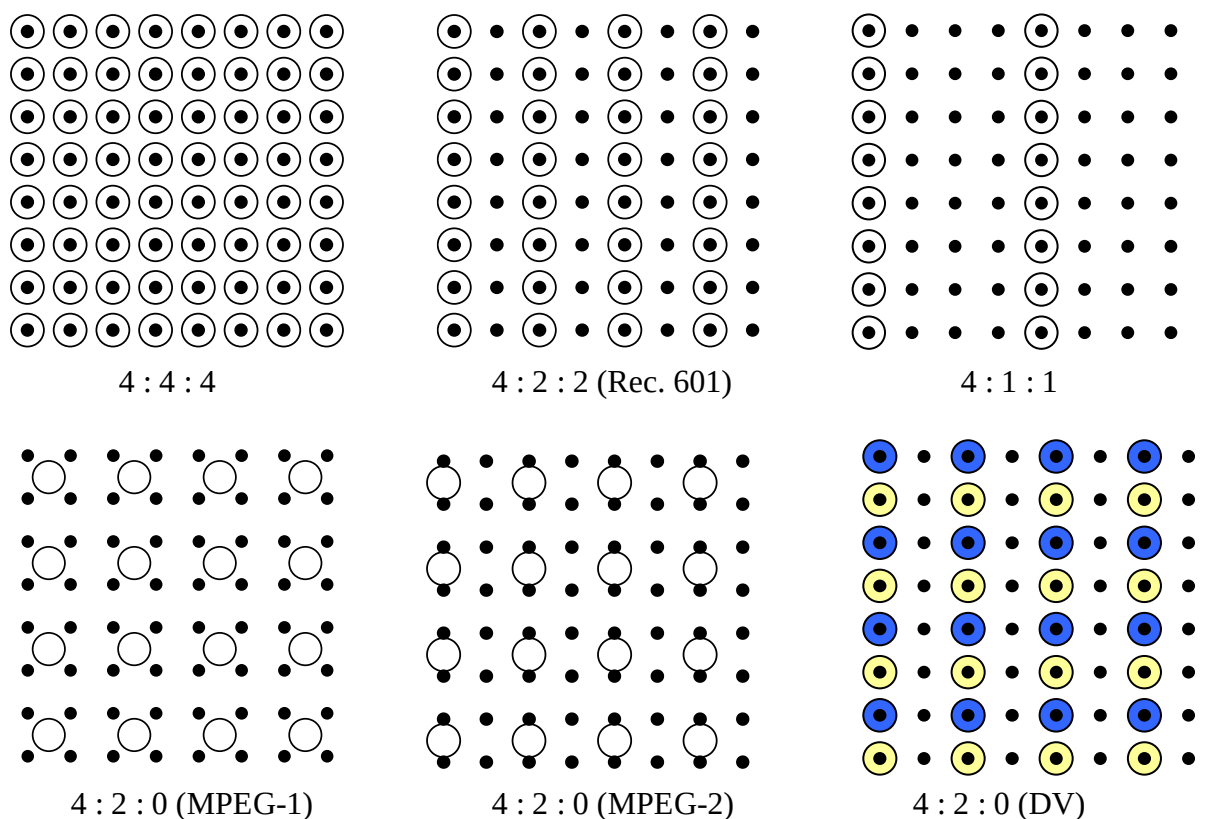


Fig. 2.5 – Estruturas de Amostragem Espacial

3. Quantização

As amostras de um sinal analógico são representadas por números reais, podendo assumir infinitos valores dentro de um determinado intervalo. O processo de quantização corresponde à conversão desses valores reais em números inteiros, atribuindo um mesmo valor numérico S_N a todas as amplitudes dentro de um intervalo $[v_N - Q/2, v_N + Q/2]$. Ou seja, equivale a “arredondar” a amostra para o valor inteiro mais próximo.

A distância Q entre os valores sucessivos de quantização (S_N e S_{N+1}) é denominada *passo de quantização*. Quando o passo é constante, dizemos que a quantização é *uniforme*. Alguns sistemas de digitalização utilizam quantização não linear, onde o passo é variável em função da amplitude do sinal.

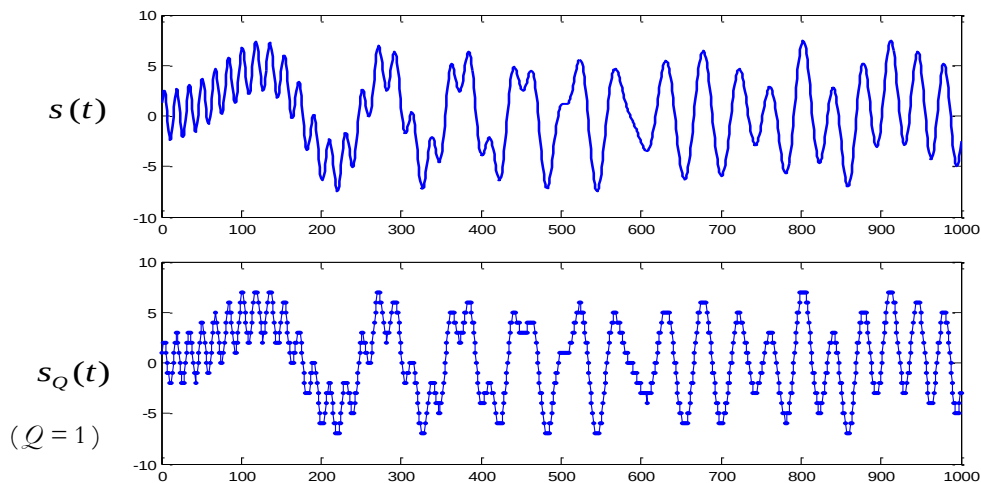


Fig. 3.1 – Quantização de um Sinal Analógico (1000 amostras)

O processo de quantização introduz um erro no sinal resultante, $e_Q(t) = S_Q(t) - S(t)$, tanto maior quanto maior o passo de quantização. Em sistemas com quantização uniforme (desde que não seja ultrapassado o fundo de escala), o erro é limitado em amplitude entre $-Q/2$ e $+Q/2$. Para sinais irregulares, e com Q suficientemente pequeno, o erro de quantização geralmente não é correlacionado com o sinal, e pode ser modelado como um sinal aleatório. Este sinal tem distribuição de amplitude e densidade espectral de potência uniformes, e sua amplitude pico-a-pico é igual a Q .

Em outras palavras, a quantização introduz um ruído aleatório aditivo, que passa a ser denominado *ruído de quantização*. A figura 3.2 apresenta o erro de quantização, seu histograma de amplitudes e o espectro, para o caso exemplificado na fig. 3.1. Nessas condições, a potência do ruído de quantização pode ser expressa por:

$$P_Q = \int_{-\frac{Q}{2}}^{\frac{Q}{2}} s^2 ds = \frac{Q^2}{12}$$

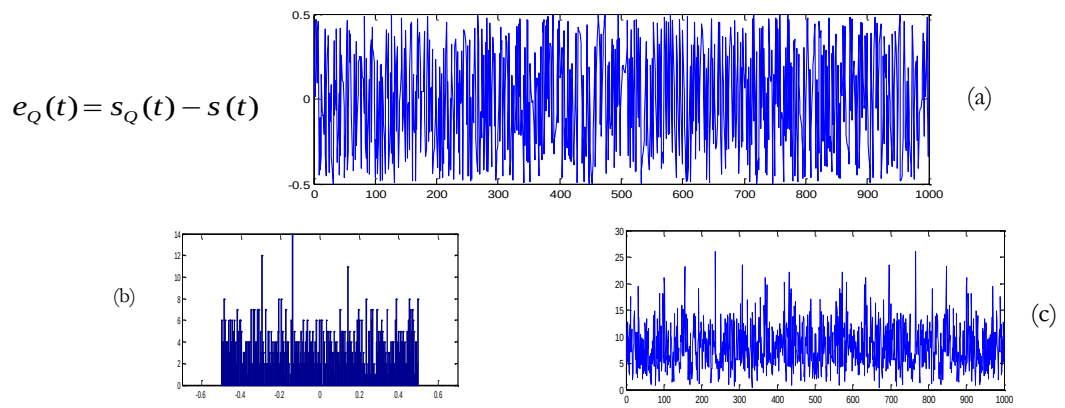


Fig. 3.2 – Exemplo de Erro de Quantização (a), Histograma de Amplitudes (b) e Espectro (c)

3.1 Quantização no Conversor A/D

O dispositivo eletrônico responsável pela quantização de um sinal elétrico é o “Conversor Analógico-Digital”, ou A/D. Na maioria dos casos, este dispositivo converte sinais de entrada (contidos dentro de uma faixa de tensão de pico-a-pico, S_{P-P} , denominada *fundo de escala*) para uma palavra digital expressa na forma binária, com um número n de bits. Desta forma, um conversor A/D usualmente possui 2^n níveis de quantização, em geral uniformemente espaçados, com passo de quantização $Q = S_{P-P} / 2^n$.

Neste caso, podemos estabelecer que a potência de pico do sinal será $P_P = (S_{P-P}/2)^2 = Q^2 \times 2^{2n-2}$. Como a potência do ruído de quantização pode ser modelada por $P_Q = Q^2 / 12$, então a relação Sinal / Ruído máxima que pode ser obtida em um processo de quantização desse tipo será dada por

$$\frac{P_P}{P_Q} = \frac{Q^2 \times 2^{2n-2}}{Q^2 \div 12} = 12 \times 2^{2n-2} = 3 \times 2^{2n}$$

Ou, expressando as grandezas em decibéis, teremos:

$$\begin{aligned} \frac{S}{R} &= 10 \log(2^{2n}) + 10 \log(3) = 20n \log(2) + 10 \log(3) \\ &= 6,02 \times n + 4,77 \text{ dB} \end{aligned}$$

Podemos estabelecer então, como regra prática, que a relação Sinal/Ruído melhora em 6 dB para cada bit adicional fornecido pelo conversor A/D. Por exemplo, conversores A/D com 8 bits (256 níveis de quantização) podem alcançar relação S/R máxima de 52,9 dB. Sistemas de 16 bits (65536 níveis) podem atingir até 101 dB.

3.2 Conversão A/D de Sinais de Vídeo

Uma vez que a **potência** de um sinal de vídeo não tem um significado físico/sensorial direto (como seria o caso do áudio), costuma-se caracterizar o ruído pela relação entre a amplitude **pico-a-pico máxima** do sinal de vídeo (p. ex. 714 mV ou 100 unidades IRE) e a amplitude **eficaz** do ruído de quantização, medida dentro da banda passante (4,2 MHz para NTSC e PAL-M). Desta forma, o ruído de quantização em função do número de bits n seria expresso por:

$$\begin{aligned}\frac{S}{R} &= 10 \log(2^{2n}) + 10 \log(12) = 20n \log(2) + 10 \log(12) \\ &= 6,02 \times n + 10,8 \text{ dB}\end{aligned}$$

Este valor deve ainda ser ajustado em função da relação entre a frequência máxima do sinal f_v e a frequência de Nyquist ($f_A/2$), uma vez que a potência do ruído de quantização é distribuída uniformemente de 0 a $f_A/2$, enquanto que o sinal de interesse atinge f_v , e o ruído entre f_v e $f_A/2$ pode ser eliminado por um filtro passa-baixas:

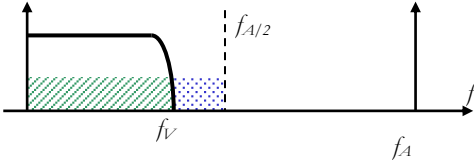
$$\frac{S}{R} = 6,02 \times n + 10,8 \text{ dB} + 10 \log\left(\frac{f_A}{2f_v}\right)$$


Fig. 3.3 – Correção pela Frequência de Amostragem

Outro fator que deve ser considerado é que o fundo de escala V_T do conversor A/D deve prever uma margem de segurança, além da excursão nominal $V_B - V_P$ do sinal de vídeo (margem esta denominada “headroom”). V_B e V_P correspondem aos níveis de branco e preto.

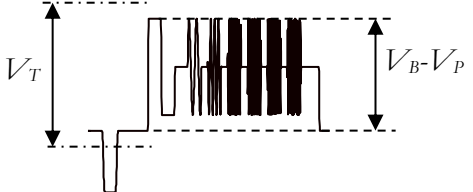
$$\frac{S}{Q_e} = 6,02 \cdot n + 10,8 + 10 \cdot \log\left(\frac{f_A}{2f_v}\right) - 20 \log\left(\frac{V_T}{V_B - V_P}\right) (\text{dB})$$


Fig. 3.4 – Correção pela Margem de “Headroom”

Finalmente, para que a medida tenha significado prático, deve ser levada em conta a resposta em frequência espacial da percepção visual humana, considerando condições de visualização típicas.

A função de transferência utilizada para modelar a resposta da visão, proposta pela CCIR, é:

$$A(f) = 10 \log_{10} \frac{\left[1 + \left(\frac{f}{f_1}\right)^2\right] \cdot \left[1 + \left(\frac{f}{f_2}\right)^2\right]}{\left[1 + \left(\frac{f}{f_3}\right)^2\right]} \text{ dB}$$

onde
 $f_1 = 270 \text{ kHz}$
 $f_2 = 1,37 \text{ MHz}$
 $f_3 = 390 \text{ kHz}$

O efeito desta ponderação espectral, aplicada sobre um ruído com espectro uniforme, equivale a acrescentar 6,81 dB na relação S/R calculada. Como exemplo, a expressão final para a relação S/R de um sistema de conversão A/D de vídeo com 8 bits, usando frequência de amostragem em 14,3 MHz, e com excursão total de 1,22 V, será dada por:

$$\frac{S}{R_Q} = 6,02 \cdot 8 + 10,8 + 10 \cdot \log\left(\frac{13,5}{8,4}\right) - 20 \log\left(\frac{1,22}{0,714}\right) + 6,81 = 63,18 \text{ dB}$$

A conversão com 8 bits é suficiente para aplicações em que o processo de digitalização ocorre apenas uma vez, como é o caso em receptores de TV ou equipamentos de captura de vídeo para

PC's. Aplicações profissionais exigem pelo menos 10 bits, para contemplar a possibilidade de várias conversões A/D e D/A em cascata.

3.2.1 Correção Gama

Deve-se ressaltar que a expressão acima é significativa apenas quando o sinal digitalizado corresponde à percepção visual de luminosidade, ou seja, é obtido a partir da luminância corrigida através do fator gama ($V = A \times B^\gamma$). Assim, espera-se que cada passo de quantização seja menor que um passo de *diferença apenas perceptível* na luminosidade, tanto para as partes claras como escuras da imagem. Ou seja, quando a quantização é uniforme, pressupõe-se que o sinal amostrado corresponde à luminância corrigida pelo fator gama.

3.2.2 “Dithering”

Nem sempre podemos assumir que o erro de quantização é aleatório; em muitas situações, quando o passo de quantização é perceptível, o observador pode perceber franjas semelhantes a “curvas de nível”, resultantes da correlação entre o sinal e a quantização (Fig. 3.4). Nesses casos, é vantajoso acrescentar um ruído aleatório de baixa amplitude ao sinal, antes da quantização. Este processo, conhecido como “dithering”, causa uma redução na relação Sinal/Ruído, mas torna o erro de quantização descorrelacionado com o sinal, reduzindo sua visibilidade.

Na figura abaixo, a visibilidade do erro de quantização com $Q = 1/16$ é reduzida através da adição de um ruído aleatório d com amplitude igual ao passo de quantização.

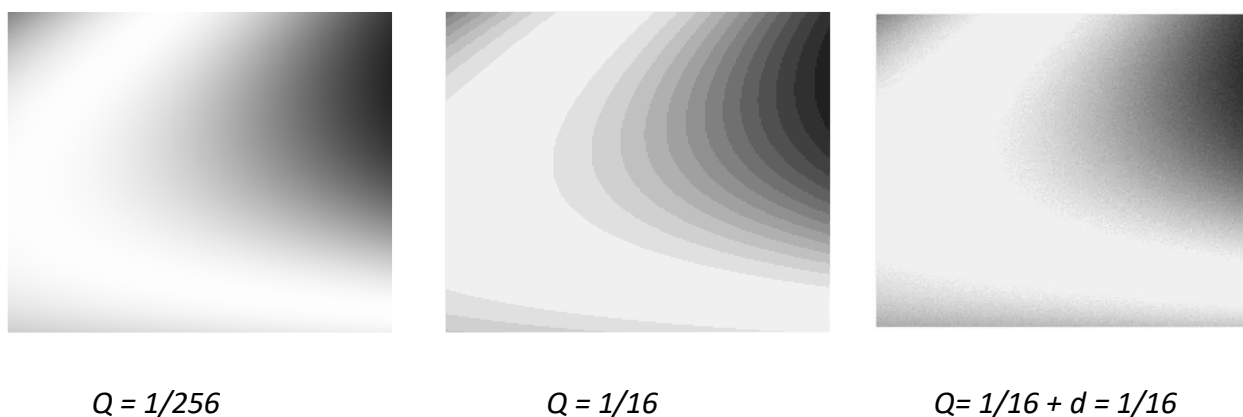


Fig. 3.4 – Amostragem com Dithering

4. Interfaces de Vídeo Digital

Para proporcionar uma interface consistente (evitando assim a degradação cumulativa provocada pela concatenação de conversores A/D e D/A em equipamentos diferentes), foram instituídos alguns formatos padronizados de vídeo digital, compreendendo sistemas de amostragem e interfaces físicas.

Várias entidades de normatização, como o CCIR (*International Radio Consultative Committee*), SMPTE (*Society of Motion Picture and Television Engineers*), ITU (*International Telecommunications Union*), EIA (*Electronics Industry Association*), ARIB (*Association of Radio Industries and Businesses – Japão*), DVB (*Digital Video Broadcasting*) e até o IEEE (*Institute of Electric and Electronics Engineers*), entre outras, contribuíram para a elaboração de recomendações e padrões de digitalização de vídeo.

Hoje, os principais formatos estão reunidos em um conjunto de documentos emitidos pela ITU-R (Setor de Radiocomunicação da ITU), na área denominada BT (*Broadcast – Television*). Estes documentos são:

- ITU_R BT.601: Aplica-se à digitalização de sinais de TV convencional, em 25 ou 30 quadros por segundo;
- ITU_R BT.656: Interfaces físicas para sinais conforme BT.601;
- ITU-R BT.709: Digitalização de TV de alta definição (1080 x 1920);
- ITU_R BT.2020: Aplicável à UHDTV (*Ultra-High Definition Television*), nos padrões “4k” (3840 x 2160) e “8k” (7680 x 4320).

4.1 Padrão D-1 (CCIR-601 / ITU-R-BT.601)

Este padrão é usado na digitalização de sinais de vídeo *Standard Definition*. É também chamado de “*Digital Components*”, ou **4:2:2**, numa referência à proporção entre o número de amostras nas componentes Y' , P_R e P_B . Efetua a amostragem independente das componentes de luminância Y' (com taxa de amostragem de 13.5 MHz) e de crominância P_R e P_B (taxa de amostragem de 6.75 MHz cada uma). A quantização é feita com 8 bits, sendo que o nível de preto (0 IRE) corresponde ao nível de quantização 16, e o branco (100 IRE) corresponde a 219, deixando uma margem de segurança para evitar saturação com sinais de amplitudes maiores. Os códigos 00 e 255 são reservados para sincronismo.

Revisões mais recentes do padrão CCIR-601 permitem digitalização com 10 bits; os valores especificados acima correspondem então aos 8 bits mais significativos.

O padrão D-1 proporciona qualidade de imagem excelente para TV convencional, sendo adequado para aplicações em estúdio. A interface física (definida pelo padrão BT.656) é paralela, balanceada e síncrona (*clock* de 27 MHz), utilizando um conector DB-25 cuja pinagem está indicada na Tabela 4.1. Os níveis de tensão correspondem a níveis lógicos ECL (-0.8 / -1.85 V).

A taxa de amostragem corresponde a 720 pixels de luminância e 360×2 *pixels* de crominância (total de 1440 amostras) durante a parte ativa de cada linha (53.33 μ s), sendo que este valor é adequado tanto para sistemas em 60Hz como para 50Hz. No caso de NTSC e PAL-M, a linha completa corresponde a 858 ciclos de 13,5 MHz, sendo que 138 ciclos correspondem ao retraço horizontal. No caso de sistemas a 50 Hz (PAL europeu), a linha horizontal completa corresponde a 864 ciclos, com 144 ciclos reservados para retraço horizontal.

Pino	Sinal	Pino	Sinal	Pino	Sinal
1	Clock	10	Bit 2	19	Retorno Bit 6
2	Terra A	11	Bit 1	20	Retorno Bit 5
3	Bit 9 (MSB)	12	Bit 0 (LSB)	21	Retorno Bit 4
4	Bit 8	13	Blindagem	22	Retorno Bit 3
5	Bit 7	14	Retorno Clock	23	Retorno Bit 2
6	Bit 6	15	Terra B	24	Retorno Bit 1
7	Bit 5	16	Retorno Bit 9	25	Retorno Bit 0
8	Bit 4	17	Retorno Bit 8		
9	Bit 3	18	Retorno Bit 7		

Tabela 4.1 - Pinagem do Conector SMPTE 125M / BT.656 Paralelo (D-1)

As amostras são intercaladas, na sequência – $C_B C_Y C_R$ – C_Y – $C_B C_Y C_R$ –... onde o grupo $C_B C_Y C_R$ corresponde a amostras espacialmente coincidentes de luminância e cromaticidade (co-localizadas), enquanto que o elemento C_Y seguinte corresponde a uma amostra de luminância avulsa intercalada (ver Figura 4.2). As valores das componentes C_Y , C_B e C_R são quantizados, obtidos a partir das componentes analógicas Y' , P_B e P_R correspondentes. As componentes analógicas Y' , P_R e P_B são calculadas a partir dos estímulos R' , G' e B' (pré-corrigidos pelo fator gama) pelas expressões abaixo:

$$\begin{aligned}
 Y' &= 0,299 R' + 0,587 G' + 0,114 B' \\
 P_B &= -0,169 R' - 0,331 G' + 0,50 B' \\
 P_R &= 0,50 R' - 0,419 G' - 0,081 B'
 \end{aligned}$$

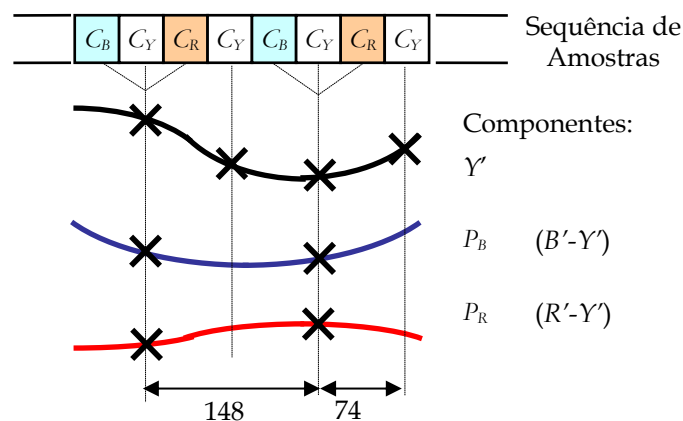


Fig. 4.2 - Estrutura de amostragem 4 : 2 : 2 para componentes no D-1

O intervalo de retraço horizontal é substituído por um bloco de dados ("blanking digital") delimitado por códigos de sincronismo (4 bytes). Estes códigos contêm dados que identificam o campo (par ou ímpar), o período de retraço vertical, e o início e fim do trecho ativo da linha (SAV - Start of Active Video / EAV - End of Active Video). Nas figuras 4.3 e 4.4 temos detalhes temporais da codificação do sinal de vídeo, enquanto que a figura 4.5 mostra os valores de luminância e cromaticidade digitalizados correspondentes a um padrão de barras coloridas com amplitude de 100% (100 IRE pico).

A tabela 4.2 mostra a codificação dos blocos delimitadores do "blanking digital" e da parte ativa de vídeo. Estes blocos consistem de um preâmbulo (3 bytes) e de um byte de dados. A Tabela 4.3 mostra os bits de paridade P3 a P0 que compõem o byte de dados.

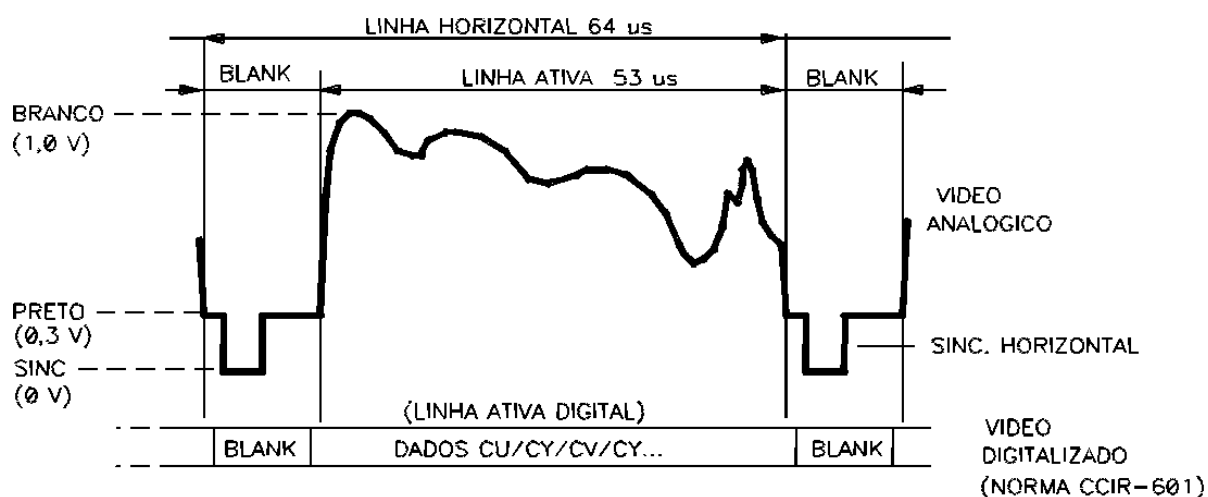


Fig. 4.3 - Formato de Dados e Relações Temporais no Padrão CCIR-601

Bit No.	9	8	7	6	5	4	3	2
Preâmbulo	1	1	1	1	1	1	1	1
	0	0	0	0	0	0	0	0
	0	0	0	0	0	0	0	0
Dados	1	Campo 0/1	Blank Vertical	SAV / EAV	P3	P2	P1	P0

Tabela 4.2 - Bloco Delimitador de amostras ativas

Bits do Byte de Dados									
9	8	7	6	5	4	3	2	1	0
(MSB)	Campo	V. B.	SAV/ EAV	P3	P2	P1	P0	1	(LSB)
1	0	0	0	0	0	0	0	0	0
1	0	0	1	1	1	0	1	0	0
1	0	1	0	1	0	1	1	0	0
1	0	1	1	0	1	1	0	0	0
1	1	0	0	0	1	1	1	0	0
1	1	0	1	1	0	1	0	0	0
1	1	1	0	1	1	0	0	0	0
1	1	1	1	0	0	0	1	0	0

Tabela 4.3 - Cálculo dos bits de paridade (P3 a P0) no Byte de dados

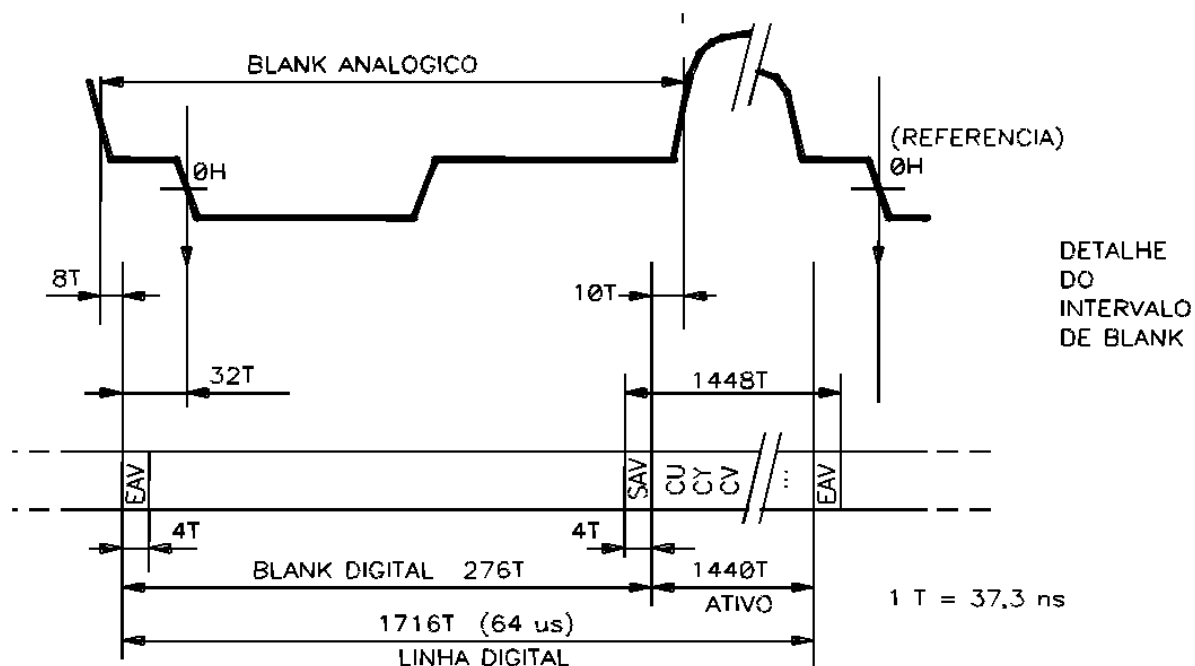


Fig. 4.4 - Formato de Dados e Relações Temporais no Padrão CCIR-601 – Detalhe do Intervalo de Retraço Horizontal

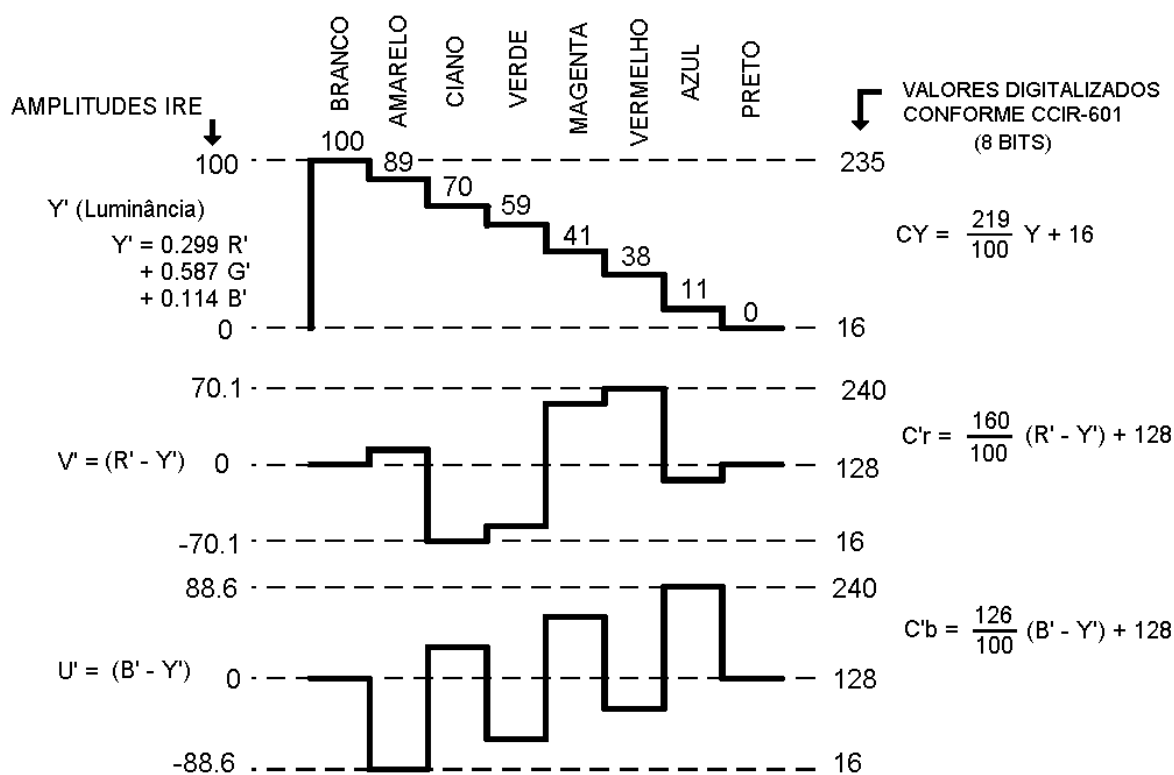


Fig. 4.5 - Padrão de Barras Digitalizado conforme CCIR-601 (D-1)

4.2 Padrão D-2 (Digital Composto)

Este formato, obsoleto, porém mais econômico que o D-1, digitaliza diretamente o sinal de vídeo composto. Para o NTSC, a taxa de amostragem é de 14.31818 MHz (correspondente a 4× a frequência da sub-portadora de cor), e a quantização é de 8 ou 10 bits. A interface física é idêntica à do D-1. A figura 4.6 mostra os níveis de quantização correspondentes para um sinal NTSC.

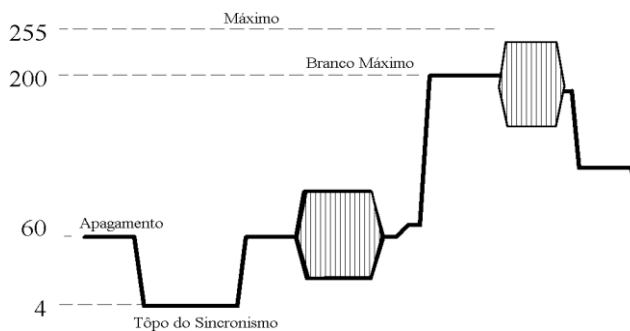


Fig. 4.6 - Níveis de Quantização para D-2

4.3 Padrão D-6 (SMPTE 274M / BT.709 / HDTV Digital)

Para a TV de alta definição, os dois formatos preferenciais de varredura, que são designados respectivamente **1080i** e **720p**, baseiam-se no padrão de vídeo analógico HDTV SMPTE 240M, hoje consolidado na recomendação ITU-R BT.709 (adotada na norma brasileira NBR 15602-1:2007). Suas características estão resumidas na tabela 4.4 abaixo.

Este padrão especifica também a função a ser utilizada na correção “gama” das componentes R, G e B:

$$V' = \begin{cases} 4,5V & 0 \leq V \leq 0,018 \\ 1,099V^{0,45} & 0,018 \leq V \leq 1 \end{cases}$$

Formato:	1080i	720p
Número total de linhas:	1125 (entrelaçadas)	750 (progressivas)
Número de linhas visíveis:	1080	720
Amostras totais por linha:	2200	1650
Amostras visíveis por linha:	1920	1280
Frequência de Amostragem:	74.1758 MHz (Luminância); 37.0879 MHz (Crominância)	
Padrão de Amostragem:	4 : 2 : 2	
Frequência de Amostragem:	74.1758 MHz (Luminância); 37.0879 MHz (Crominância)	
Resolução:	8 ou 10 bits	

Tabela 4.4 – Características de Amostragem para HDTV

As componentes analógicas Y' , P_B e P_R são obtidas pelas expressões:

$$Y' = 0,2126 R' + 0,7152 G' + 0,0722 B'$$

$$P_B = -0,114572 R' - 0,385428 G' + 0,50 B'$$

$$P_R = 0,50 R' - 0,454153 G' - 0,045847 B'$$

A figura 4.7 apresenta os valores de luminância e crominância digitalizados, correspondentes a um padrão de barras coloridas com amplitude de 100% (100 IRE pico), digitalizado com 10 bits.

O código EAV recebe mais duas palavras adicionais (LN0 e LN1), que transportam uma contagem sequencial das linhas, e 4 palavras de CRC (YCR0, YCR1, CCR0, CCR1) para detecção de erros nas sequências de luminância e croma.

A figura 4.8 e a tabela 4.5 mostram as durações dos pulsos de sincronismo para HDTV, após digitalização.

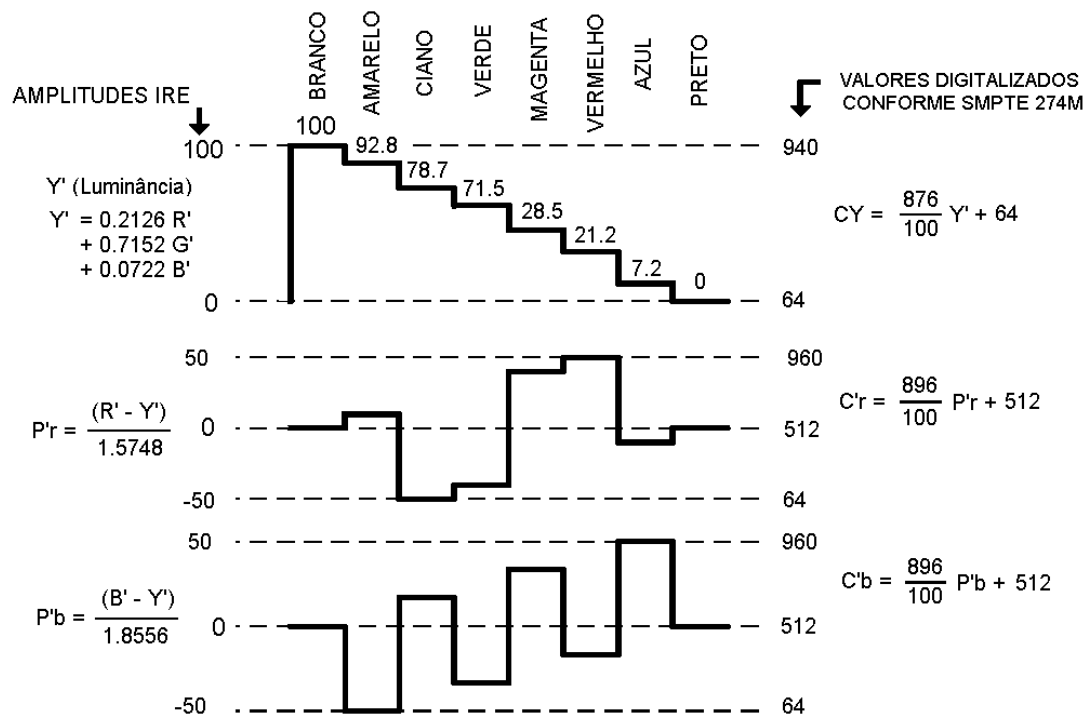


Fig. 4.7 - Padrão de Barras Digitalizado conforme SMPTE 274M (HDTV)

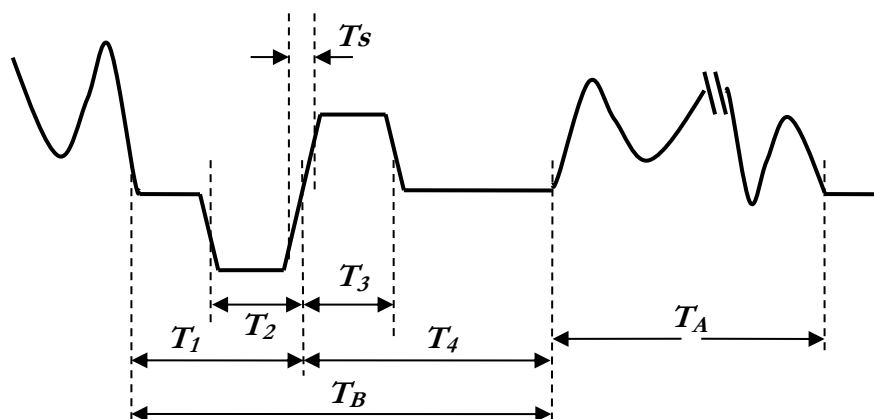


Fig. 4.8 – Digitalização dos sinais de sincronismo para HDTV

Formato:	1080i	720p
T_1 = Pórtico Anterior	88	110
T_2 = Largura do Pulso Negativo de Sincronismo	44	40
T_3 = Largura do Pulso Positivo de Sincronismo	44	40
T_4 = Pórtico Posterior	192	260
T_S = Tempo de Subida / Descida	4	4
T_B = Largura do Apagamento Horizontal	280	370
T_A = largura da Linha Ativa	1920	1280

*Tabela 4.5 – Durações dos sinais de sincronismo para HDTV
(Valores em múltiplos de $1/f_L = 1/74.1758 \text{ MHz} = 13,48 \text{ ns}$)*

4.4 Interface Serial CCIR-656

Para simplificar a interconexão de equipamentos de vídeo, padronizou-se uma interface serial destinada a atender amostragem D-1 a 8 bits, permitindo transmissão por um cabo coaxial de 75 ohms. Esta interface opera a 243 Mb/s, sendo que cada byte de dados é convertido para um símbolo de 9 bits (codificação de grupo 8/9) com a intenção de conservar a componente CC próxima de zero e garantir boa quantidade de transições lógicas para permitir recuperação de *clock*.

4.5 Interface SDI - Serial Digital Interface (SMPTE 259M / ITU-R BT.656-4)

Este padrão foi desenvolvido para permitir transmissão serial de sinais de vídeo, digitalizados com 10 bits, através de cabo coaxial de 75 Ohms, admitindo atenuação de até 30 dB.

Para garantir componente CC zero e presença de transições suficientes para recuperação de relógio (sem, no entanto, aumentar a taxa de bits) é utilizado embaralhamento (“*scrambling*”) através de um randomizador, apresentado na fig. 4.9, além de codificação diferencial (bits “1” são transmitidos como transições). A interface opera a 270 Mb/s para sinais CCIR-601 (D-1), ou 143 Mb/s para Digital Composto (D-2). A amplitude nominal do sinal SDI é de 800 mVpp em 75 Ohms.

Estas características possibilitam, em tese, que o sinal SDI trafegue pelas matrizes de comutação atualmente usadas para distribuir sinais de vídeo composto.

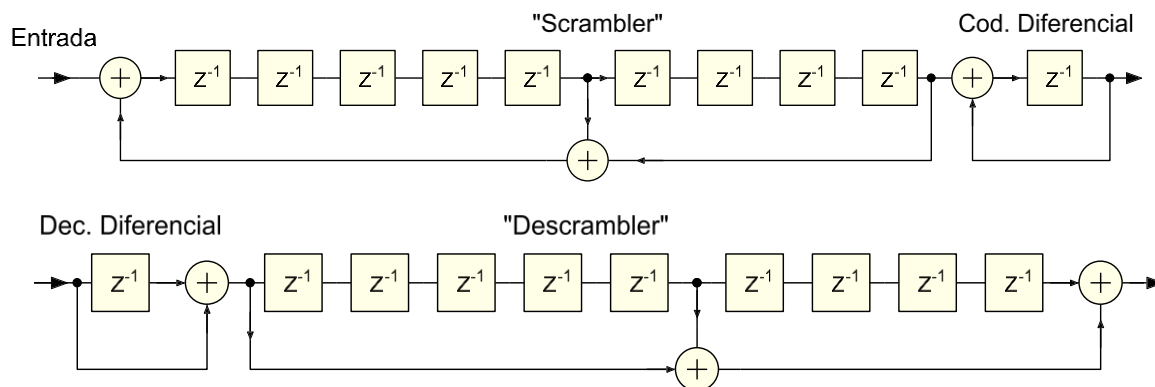


Fig. 4.9 - Randomizador e Codificador / Decodificador SDI

Uma desvantagem do “scrambling” é a multiplicação dos erros de transmissão; a recepção de um bit errado na entrada do circuito acima provocará erros na saída enquanto o bit errado estiver se deslocando dentro dos registradores do “descrambler”. O resultado disso é que a taxa de erros depende fortemente da atenuação no cabo de transmissão; na fig. 4.10 temos valores de BER (“Bit Error Rate”) em função do comprimento do cabo, para um tipo de cabo coaxial comumente usado em estúdios.

O padrão SDI prevê a capacidade de transmissão de áudio digital, na forma de pacotes enviados durante os períodos de apagamento H e V. O sinal de áudio pode ser amostrado a 32 kHz ou 48 kHz, com 16 bits, em até 4 canais.

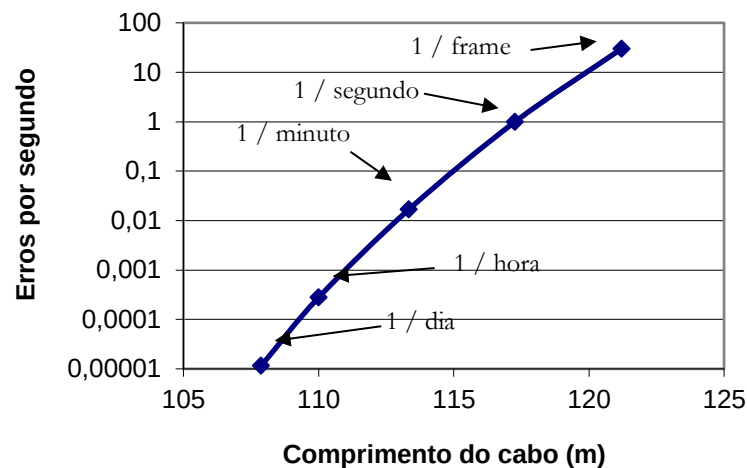


Fig. 4.10 - Taxa de erros para SDI × comprimento de cabo

4.5.1 HD-SDI

A interface serial utilizada para HDTV (HD-SDI, definida pela recomendação SMPTE 292M) oferece uma taxa de transmissão de 1,485 Gb/s para quantização de 10 bits. Neste caso, o embaralhamento (similar ao da figura 4.9) é independente (e intercalado) para as amostras de luminância e croma.

A figura 4.11 mostra a estrutura dos prefixos EAV e SAV, bem como os números de linha e CRC, intercalados para luminância e croma.

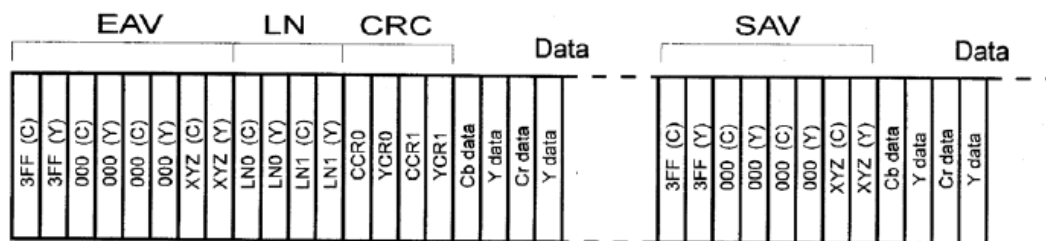


Fig. 4.11 – Intercalamento de Luminância e Crominância na HD-SDI

4.6 Interface DVI (Digital visual Interface)

A interface DVI foi desenvolvida para interligar equipamentos de vídeo com dispositivos de apresentação de imagem que sejam intrinsecamente digitais, como por exemplo, monitores de Plasma ou Cristal Líquido (LCD). Evitam-se assim os problemas decorrentes da conversão D/A no equipamento gerador, seguida de conversão A/D, eventualmente com taxa diferente, no dispositivo de apresentação. Ainda é largamente usada para interligar interfaces gráficas de microcomputadores com monitores digitais.

O padrão DVI utiliza 3 canais em paralelo para transmitir as informações de R, G e B, digitalizadas em 8 bits. A transmissão é serial, balanceada, a taxas que podem alcançar 1,6 Gb/s, sendo que a informação digital é codificada por um processo denominado T.M.D.S. (“*Transistion Minimized Differential Signalling*”). Neste processo, uma palavra de 8 bits é convertida em 10 bits de forma a minimizar a quantidade de transições, o que proporciona menor ocupação espectral e menor irradiação de interferências do sinal transmitido.

Além dos 3 canais de informação de imagem, o padrão DVI utiliza um canal de “clock”, na taxa de amostragem de pixel (1/10 da taxa de bits). As características do PLL a ser utilizado no decodificador, para multiplicar o “clock” e obter a frequência de transmissão de bit (fig. 4.12), são especificadas no padrão, de modo a suportar condições de instabilidade de frequência que podem existir no equipamento codificador.

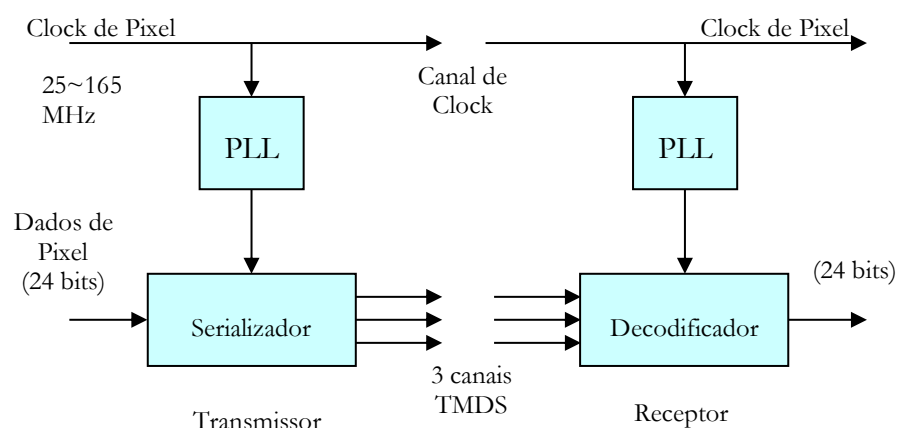


Fig. 4.12 – Interface DVI

4.7 Interface HDMI (High Definition Multimedia Interface)

Esta interface é derivada do padrão DVI, utilizando o mesmo processo de codificação (T.M.D.S.) e de transmissão serial balanceada, porém visando aplicação em equipamentos de entretenimento. Além de utilizar um conector de menores dimensões, acrescenta a possibilidade de transmitir áudio digital, multiplexado com os canais de vídeo RGB.

Além disso, implementa um canal bidirecional de controle (de baixa velocidade), através do qual ambos os equipamentos, gerador e apresentador de imagem, estabelecem as respectivas capacidades de resolução espacial, visando compatibilidade de formatos de geração e de exibição de sinal.

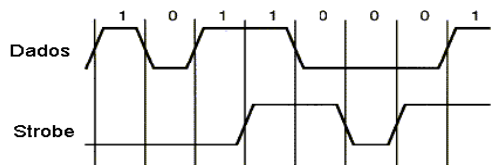
A interface HDMI pode incorporar um sistema de proteção contra acesso indevido, denominado HDCP (*High-bandwidth Digital Content Protection*), que utiliza processos de autenticação e criptografia para transferir os dados entre os equipamentos.

4.8 Interfaces para Vídeo Comprimido

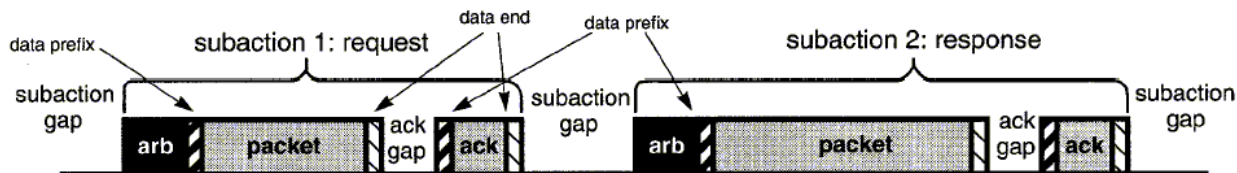
As interfaces descritas a seguir são utilizadas para transmissão de sinais de vídeo que passaram por um processo de compressão (Vídeo e áudio, em geral). A taxa de bits efetiva é muito menor do que nas interfaces sem compressão.

4.8.1 Vídeo Sobre IEEE1394 (“Firewire”, “I-Link”)

O padrão IEEE1394 define uma interface síncrona bidirecional com taxas de transmissão da ordem de 100 a 400 Mb/s (800 Mb/s conforme IEEE1394-B). A interface física consiste de 3 pares de sinais, sendo um par para alimentação CC, e dois pares para dados e “*strobe*”. A codificação usada é apresentada na figura ao lado; os dados são transmitidos em NRZ, balanceado, enquanto que o par de “*strobe*” apresenta uma transição sempre que não houver uma transição no par de dados.



As transações ocorrem através de pacotes de dados, precedidos de etapas de arbitragem e prefixos (endereços de origem e destino, códigos de transação), e finalizados por CRC.



A interface IEEE1394 (“Firewire”) é usada como meio de transporte para vídeo digital comprimido, conforme formatos IEC61883-4 (Transporte de pacotes MPEG-2 DVB), DirecTV e DV (Digital Vídeo – IEC61883-2). No caso do MPEG-2, os pacotes de 188 bytes são acrescidos de um prefixo de descrição de tempo real (data/hora/frame) e depois repartidos em blocos de 24 bytes, que são transportados por pacotes isócronos (Fig. 4.13).

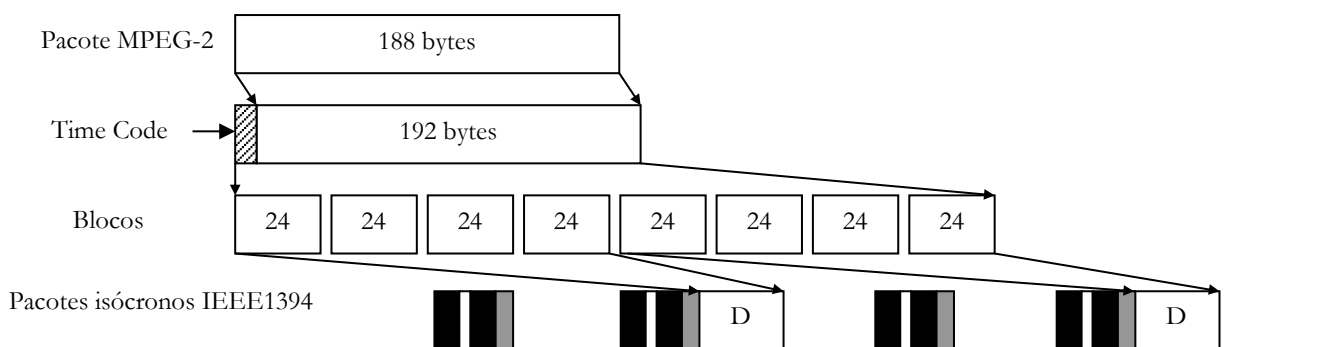


Fig. 4.13 – Decomposição de Pacotes MPEG em Pacotes Isócronos IEEE 1394

4.8.2 Interface DVB-SPI (Synchronous Parallel Interface)

Esta interface é padronizada pelo consórcio europeu DVB (Digital Vídeo Broadcasting) conforme norma ETSI - EN50083-9, com a finalidade de interligar equipamentos de transmissão que utilizam o protocolo de transporte por pacotes (*Transport Stream*) definido no padrão MPEG-2.

Fisicamente, consiste de uma interface paralela, na qual trafegam 8 bits, clock, sincronismo de pacote e um sinal indicando que há dados válidos. Estes 11 sinais utilizam níveis lógicos LVDS (diferenciais).

A figura 4.14 apresenta formas de onda correspondentes a um modo de operação da interface, na qual um pacote de 188 bytes é complementado por 16 bytes de enchimento.

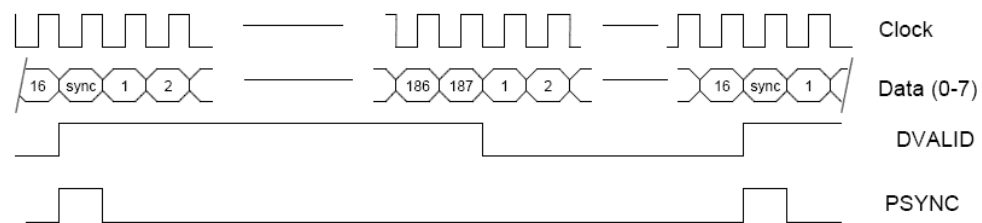


Fig. 4.14 – Sinais na Interface DVB-SPI – Modo 204 Bytes c/ Enchimento

A tabela 4.5 apresenta a pinagem do conector DVB-SPI (conector tipo DB-25).

Pino	Sinal	Pino	Sinal	Pino	Sinal
1	Clock A	10	D0 A (LSb)	19	D4 B
2	Terra	11	DVALID A	20	D3 B
3	D7 A (MSb)	12	PSYNC A	21	D2 B
4	D6 A	13	Blindagem	22	D1 B
5	D5 A	14	Clock B	23	D0 B (LSb)
6	D4 A	15	Terra	24	DVALID B
7	D3 A	16	D7 B (MSb)	25	PSYNC B
8	D2 A	17	D6 B		
9	D1 A	18	D5 B		

Tabela 4.5 – Pinagem do Conector DVB-SPI

4.8.3 Interface ASI – Asynchronous Serial Interface

Esta interface, também especificada na norma ETSI - EN50083-9, é a versão serial da interface DVB-SPI, operando a uma taxa fixa de 270 Mbits/s. As características físicas são semelhantes à interface SDI, com a diferença que os dados relativos ao *Transport Stream* são transmitidos em pacotes. Aqui é aplicada uma conversão de 8 para 10 bits, além da inserção de códigos especiais para enchimento e sincronismo de pacotes.

4.8.4 Vídeo sobre IP (IPTV)

Arquivos de vídeo, comprimidos por MPEG-2 ou outros processos, podem ser transportados por redes de dados utilizando os protocolos UDP, RTP ou TCP/IP. O termo **IPTV** (Televisão sobre

protocolo IP) designa genericamente várias situações pelas quais informações de vídeo podem trafegar pela rede:

- **Transferência de Arquivos:** um vídeo é solicitado a um servidor por transações TCP/IP e transferido ao usuário por FTP, para visualização posterior. Neste caso, o protocolo garante que não ocorrem erros de transmissão, porém o tempo de transferência é indeterminado.
- **Vídeo Sob Demanda (VOD – Video On Demand):** Um usuário solicita um determinado vídeo a um servidor, e a exibição inicia imediatamente. Neste caso, pacotes de dados contendo partes sequenciais do vídeo são direcionados ao usuário através de protocolo UDP (*User Datagram Protocol*) ou RTP (*Real Time Protocol*). Este processo, denominado “*streaming*”, permite a visualização em tempo real, desde que seja implementado, no receptor, um reservatório (“*buffer*”) capaz de amortecer as instabilidades no tempo de propagação dos pacotes na rede. Além disso, os protocolos UDP e RTP não garantem a entrega dos pacotes; assim sendo, podem ocorrer falhas na transmissão.
- **Difusão (Broadcast):** Vários usuários podem se conectar independentemente a um serviço de distribuição de vídeo, que transmite pacotes UDP ou RTP com endereçamento “*multicast*”. Cada novo usuário que se conecta implica em que os roteadores que estejam no caminho liberem o endereço “*multicast*” correspondente ao serviço.

Header IP 12 bytes	Header UDP 8 bytes	Header RTP 12 bytes	Pacote MPEG-2 188 bytes	Pacote MPEG-2 188 bytes	Pacote MPEG-2 188 bytes	Pacote MPEG-2 188 bytes	...	Pacote MPEG-2 188 bytes
--------------------------	--------------------------	---------------------------	-------------------------------	-------------------------------	-------------------------------	-------------------------------	-----	-------------------------------

Fig. 4.15 – Pacote RTP para IPTV

A estrutura de um pacote de vídeo encapsulado em RTP está apresentada na fig. 4.15. O cabeçalho (*Header*) RTP contém, entre outras informações, um campo de 32 bits denominado “*Time Stamp*”, que corresponde a uma amostragem de um contador alimentado por uma frequência de 90 kHz. A amostragem é feita no transmissor a cada novo quadro do vídeo. Todos os pacotes MPEG-2 que correspondem a esse quadro são transmitidos com o mesmo Time Stamp, porém identificados com uma contagem sequencial. Isso permite ao receptor organizar os pacotes recebidos, e aguardar o instante correto para exibir o quadro correspondente.

5. Processamento Digital de Vídeo

5.1 Frame Buffer

Um componente central dos equipamentos processadores de vídeo que operam na forma digital é o “*frame buffer*”, ou memória de armazenamento de imagem.

Este bloco funcional é normalmente constituído de um banco de memória RAM com duplo acesso, ou seja, a escrita e a leitura de informação podem ocorrer simultaneamente e são controladas por unidades de endereçamento independentes. Além disso, o endereçamento da memória é segmentado na forma de coordenadas x e y independentes, onde x = número do *pixel* dentro de uma linha horizontal e y = número da linha no sentido vertical.

Eventualmente, o Frame Buffer pode armazenar mais de um campo ou quadro, acrescentando um ou mais bits de endereçamento para a seleção destes.

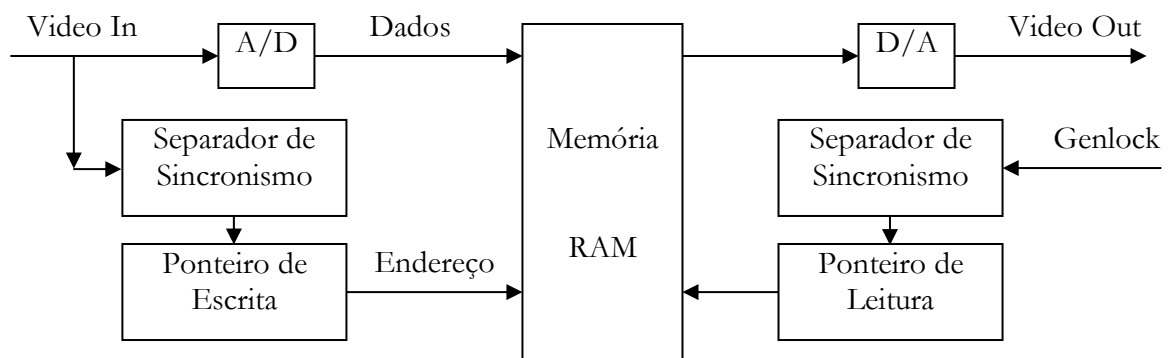


Fig.5.1 – “Frame Buffer”

A figura 5.1 mostra um diagrama de blocos de um “*Frame Buffer*” configurado como sincronizador de vídeo. O sinal digitalizado (por exemplo, em componentes YUV) é armazenado em uma memória com capacidade pelo menos para um quadro completo (~700 kBytes para CCIR-601, ~6 Mbytes para HDTV). Os ponteiros de endereçamento para escrita e leitura são independentes, sincronizados respectivamente pelo sinal de vídeo da entrada e pelo sincronismo geral externo (“*Genlock*”).

5.2 Redutor de Ruído

Um método de redução de ruído, aplicável a imagens de vídeo, explora a alta correlação existente entre *pixels* adjacentes, tanto no sentido espacial como no temporal, para imagens naturais. Equipamentos redutores de ruído digitais efetuam uma média ponderada entre *pixels* próximos, em linhas e quadros consecutivos (fig. 5.2). Utilizam para isso um “*Frame Buffer*” com saídas múltiplas, atrasadas de 1 ou mais *pixels*, linhas e/ou campos, as quais são combinadas de forma linear ou não-linear.

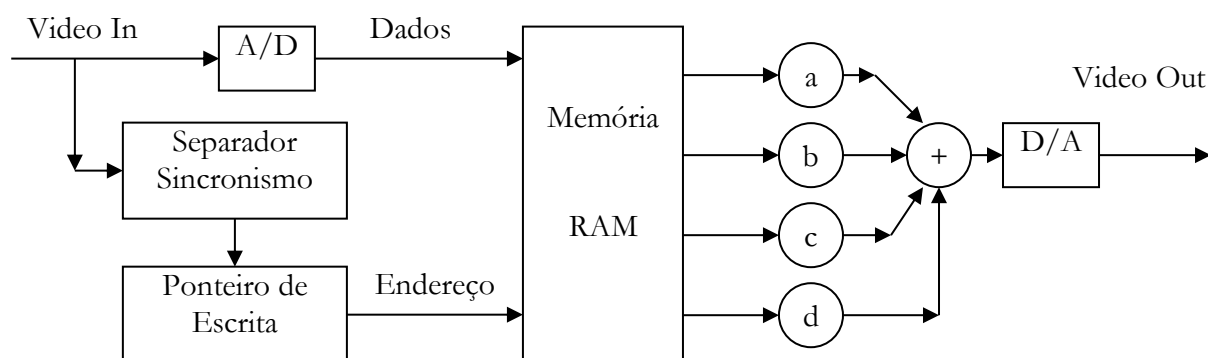


Fig. 5.2 – Redutor Digital de Ruído

5.3 Transformações Geométricas

Para criar efeitos de perspectiva, rotação ou deformação da imagem, é necessário introduzir um processador de geometria no endereçamento de leitura do “Frame Buffer”. Este processador mapeia o sistema de coordenadas da imagem reproduzida (x_R , y_R) sobre as coordenadas da imagem de entrada (x_E , y_E) através de um algoritmo apropriado.

Por exemplo, uma translação da imagem original é realizada pelas seguintes equações:

$$x_E = x_R - T_X$$

$$y_E = y_R - T_Y$$

No exemplo mostrado acima, quando $x_R - T_X < 0$ ou $y_R - T_Y < 0$ (quando as coordenadas de entrada caem fora da imagem), o *frame buffer* fornece um valor zero na saída (imagem preta).

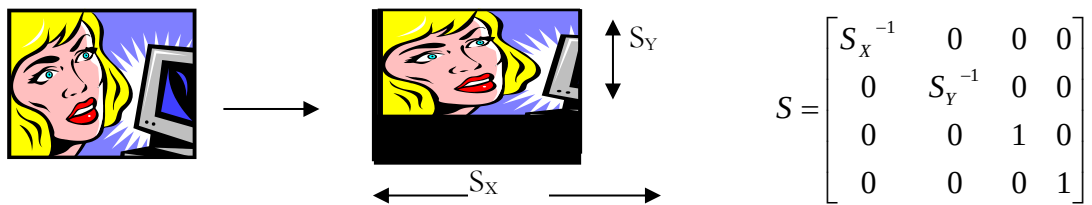
Na forma matricial, podemos expressar estas equações como:

$$\begin{bmatrix} x_E \\ y_E \end{bmatrix} = \begin{bmatrix} 1 & 0 & -T_X \\ 0 & 1 & -T_Y \end{bmatrix} \cdot \begin{bmatrix} x_R \\ y_R \\ 1 \end{bmatrix}$$

Vários efeitos (transformações) podem ser concatenados em uma única matriz; para isso podemos em primeiro lugar generalizar o espaço da imagem como possuindo 3 dimensões, com a coordenada z perpendicular ao plano da imagem (sendo que a imagem reproduzida terá sua coordenada $z_R = 0$), e adotando matrizes quadradas para representar as transformações individuais. Assim sendo, a translação nas dimensões x , y e z será efetuada por uma matriz T dada por:

$$\begin{bmatrix} x_E \\ y_E \\ z_E \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -T_X \\ 0 & 1 & 0 & -T_Y \\ 0 & 0 & 1 & -T_Z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_R \\ y_R \\ z_R \\ 1 \end{bmatrix} \Rightarrow v_E = T \cdot v_R$$

Uma expansão ou compressão da imagem nos eixos x , y e z corresponde ao uso de uma matriz de transformação de escala S dada por:

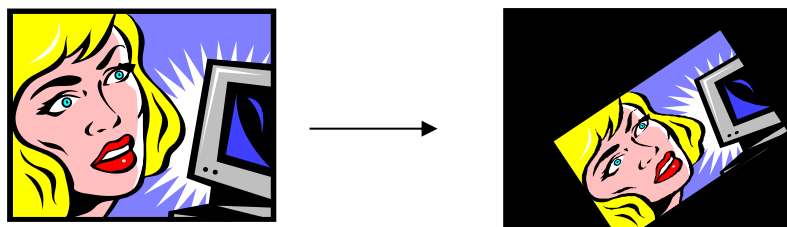


A rotação de uma imagem em torno dos eixos x , y e z é feita pelas seguintes matrizes:

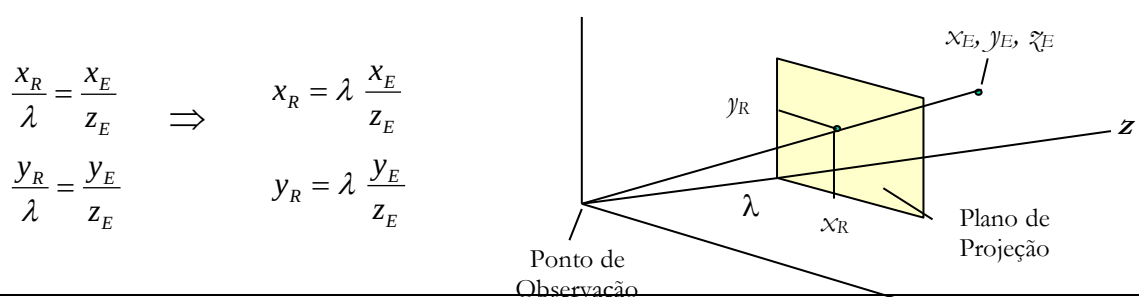
$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha^{-1} & -\sin \alpha^{-1} & 0 \\ 0 & \sin \alpha^{-1} & \cos \alpha^{-1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad R_\beta = \begin{bmatrix} \cos \beta^{-1} & 0 & -\sin \beta^{-1} & 0 \\ 0 & 1 & 0 & 0 \\ \sin \beta^{-1} & 0 & \cos \beta^{-1} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad R_\theta = \begin{bmatrix} \cos \theta & -\sin \theta & 0 & 0 \\ \sin \theta & \cos \theta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Várias transformações podem ser concatenadas calculando-se o produto das matrizes correspondentes a cada transformação, na ordem especificada. Por exemplo, uma sequência de translação, escalamento e rotação no eixo z será definida por

$$v_R = (R_\theta \cdot (S \cdot (T))) \cdot v_E$$



Para criar efeitos de perspectiva, é necessário considerar a projeção de um espaço tridimensional (coordenadas x_E, y_E, z_E) sobre um plano de projeção (coordenadas x_R, y_R, λ), a partir de um ponto de observação. Esta projeção corresponde a uma transformação não-linear, dada por



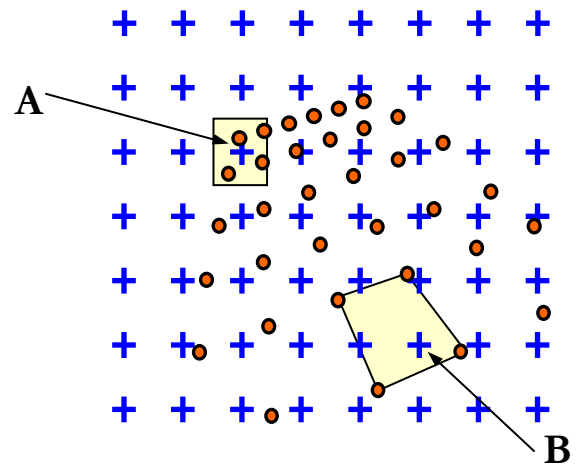
5.4 Interpolação e Decimação

As operações de escalamento podem exigir interpolação de amostras (quando a distância entre os *pixels* nas coordenadas de saída é menor que a distância correspondente na imagem de entrada) ou decimação (no caso inverso).

Na figura ao lado, a grade (+) representa a estrutura de amostragem da imagem de saída, enquanto que (o) representa a estrutura da imagem de entrada, sujeita a uma transformação de perspectiva.

O *pixel* A deve ser calculado a partir da média ou decimação de vários *pixels* da imagem original, enquanto que o *pixel* B deve ser interpolado a partir dos valores dos 4 *pixels* vizinhos na imagem original.

Nestes processos deve ser prevista filtragem para evitar “*aliasing*” espacial.



5.5 Transcrição de Formatos

O processo de transcrição consiste na conversão entre formatos de vídeo diferentes. Por exemplo, a exibição em PAL-M ou NTSC (60 campos por segundo) de um sinal originado em PAL-B (50 Hz) exige que seja feita uma conversão (fig. 5.5):

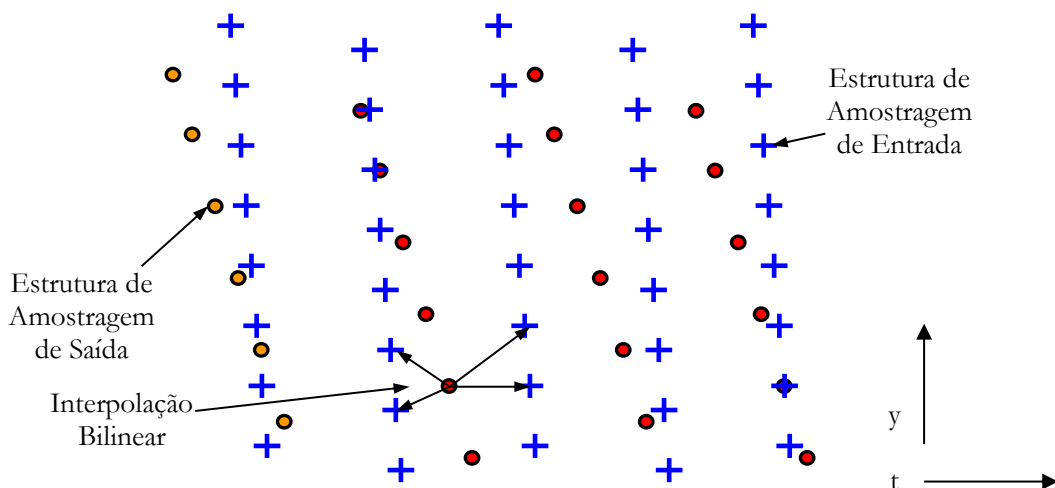


Fig. 5.5 – Conversão de Formatos

Este processo envolve interpolação de quadros e decimação de linhas (o formato PAL-B possui 576 linhas visíveis que devem ser reduzidas para 480 linhas no PAL-M).

Normalmente, é feita uma decimação usando um filtro FIR com 3 a 4 coeficientes no sentido vertical (linhas consecutivas), seguido de uma interpolação linear entre 2 campos consecutivos; ou então, uma interpolação bilinear entre as 4 linhas mais próximas (em 2 campos consecutivos).

5.6 “Chroma Key”

O recurso de “*Chroma Key*” é usado para inserir uma imagem (primeiro plano) sobre outra imagem (plano de fundo), usando a informação de crominância como “chave” para efetuar a comutação entre as imagens.

Uma câmera focaliza os elementos do primeiro plano sobre um fundo uniforme, com uma cor de referência, geralmente azul ou verde. Um circuito compara em tempo real o sinal de vídeo com a cor de referência; se coincidir, a imagem de segundo plano é selecionada. Caso contrário, a imagem do primeiro plano é escolhida.

Como o circuito compara apenas a crominância do sinal, é insensível a sombras e outras variações de luminosidade sobre o fundo uniforme.

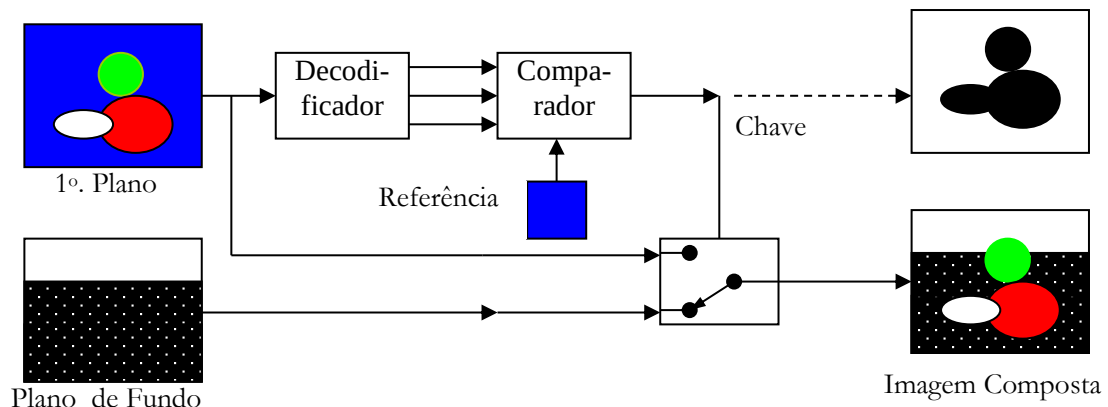


Fig. 5.6 – “Chroma Key”

O mesmo princípio é usado para inserir caracteres, símbolos ou vinhetas, gerados por computador, sobre uma imagem de vídeo. Neste caso, os caracteres ou símbolos são armazenados em um “*Frame Buffer*” através de acesso direto à memória. Um determinado valor de luminância e/ou crominância é interpretado como “transparente” e provoca o chaveamento do sinal de vídeo ao qual os caracteres serão superpostos.

5.7 O receptor de TV Digital

Um receptor moderno deve permitir a exibição de conteúdo multimídia, disponível sob vários formatos, sejam analógicos (para compatibilidade com sistemas antigos) ou digitais. A figura 5.7 mostra o diagrama de blocos de um televisor moderno, que inclui entradas de antena e cabo para radiodifusão (analógica e digital), além de entradas de vídeo composto ou componentes (p. ex. para ligação de gravadores de “vídeo-cassete”, consoles de jogos ou câmeras), entradas digitais (HDMI) para equipamentos modernos, e eventualmente saídas de vídeo composto para gravação de programas em equipamentos analógicos antigos.

O processamento dos sinais de vídeo é feito por um DSP (Processador Digital de Sinais) dedicado, especializado para conversão de formatos, decodificação e recodificação de imagens, compatibilizando todas as entradas com a resolução nativa do display.

5.8 Processamento no receptor de TV: Equalização de Multipercurso

Alguns receptores de TV analógicos incorporam um filtro equalizador que permite reduzir o efeito das distorções por multipercurso, que são resultantes de reflexões em obstáculos, veículos, etc., vulgarmente conhecidas como “*fantasmas*” na radiodifusão analógica. Este filtro baseia-se em um modelo de resposta finita ao impulso para o canal de transmissão (ver figura 5.8). Cada reflexão do sinal transmitido corresponde a um atraso Δt_n e a uma atenuação k_n .

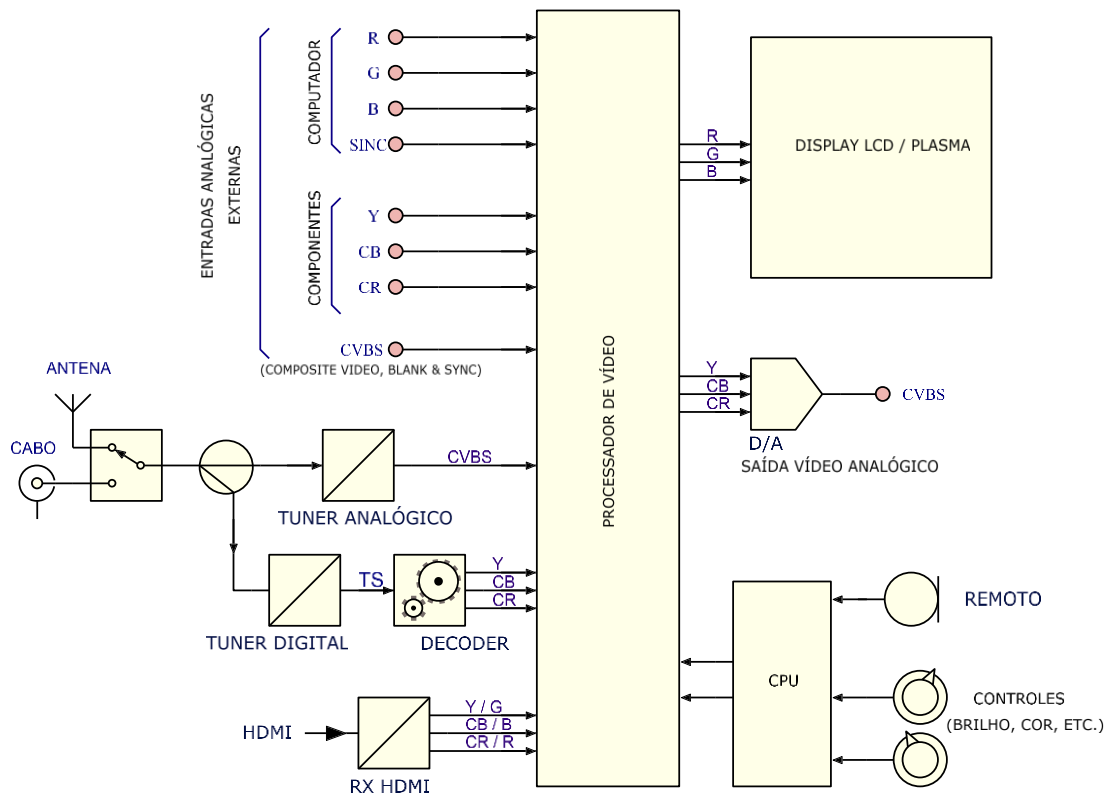


Fig. 5.7 – Diagrama de Blocos de um Receptor de TV Digital

No receptor, é implementado um filtro complementar. O correto funcionamento deste filtro depende da determinação dos valores de Δt_n e k_n , o que pode ser feito por auto-correlação do sinal recebido.

Alguns sistemas dependem da transmissão de um sinal de referência específico (GCR – “*Ghost Cancelling Reference*”) no intervalo de retraço vertical; outros sistemas utilizam a forma de onda da linha 3 (início do sincronismo vertical), na qual ocorre um degrau isolado no centro da linha, que pode ser tomado como referência.

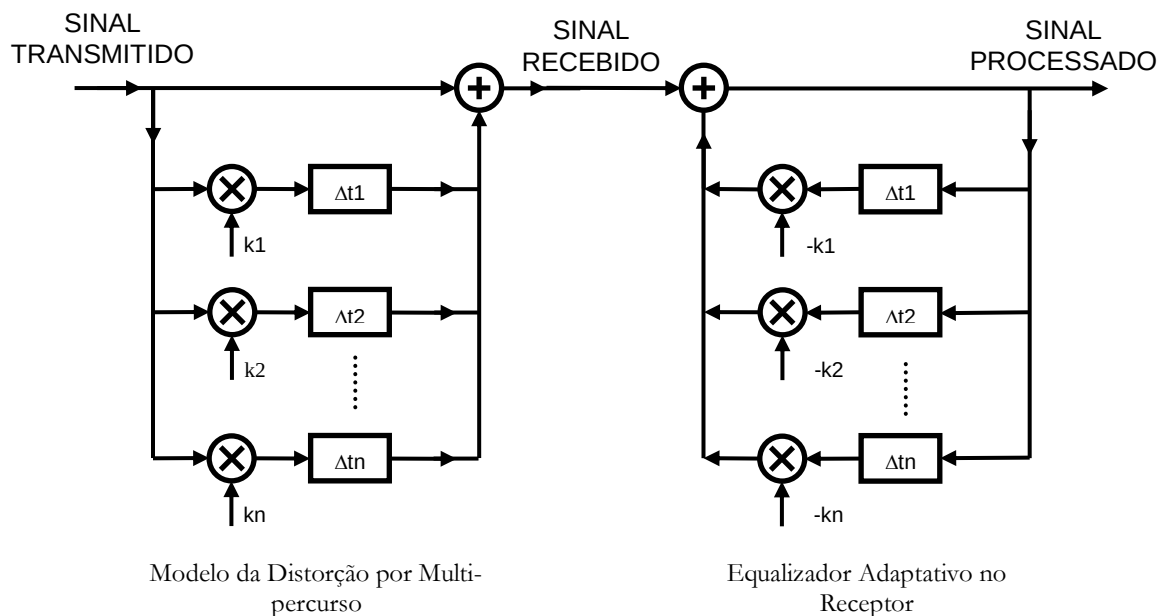


Fig. 5.8 – Equalização de Multipercursos

6. Requisitos de Banda Passante para Transmissão Digital

Na forma digital bruta, a taxa de bits necessária para transmitir um sinal de vídeo é consideravelmente elevada. Por exemplo, os requisitos de transmissão (apenas relativos à parte ativa da imagem) para um sinal de vídeo digital no formato D-1, seriam dados por:

$$1 \text{ linha} = 720 \text{ amostras de Luminância} + 2 \times 360 \text{ amostras de crominância} = 1440 \text{ amostras}$$

$$1440 \text{ amostras/linha} \times 482 \text{ linhas} \times 29,97 \text{ quadros/s} = 20.801.578 \text{ amostras/s}$$

$$20.801.578 \text{ amostras/s} \times 8 \text{ bits} = \mathbf{166,4 \text{ Mb/s}}$$

No caso do padrão D-2, teríamos:

$$1 \text{ linha (parte visível)} = 51 \mu\text{s} ; \text{Taxa de Amostragem} = 14,31818 \text{ MHz}$$

$$14,31818 \text{ MHz} \times 51 \mu\text{s} = 730 \text{ amostras / linha}$$

$$730 \times 482 \text{ linhas} \times 29,97 \text{ quadros/s} = 10.545.244 \text{ amostras/s}$$

$$10.545.244 \text{ amostras/s} \times 8 \text{ bits} = \mathbf{84,36 \text{ Mb/s}}$$

Já para o caso da HDTV em 10 bits no padrão D-6, teríamos:

$$1 \text{ linha} = 1920 \text{ amostras de Luminância} + 2 \times 960 \text{ amostras de crominância} = 3840 \text{ amostras}$$

$$3840 \text{ amostras/linha} \times 1080 \text{ linhas} \times 29,97 \text{ quadros/s} = 124.291.584 \text{ amostras/s}$$

$$124.291.584 \text{ amostras/s} \times 10 \text{ bits} = \mathbf{1242,9 \text{ Mb/s}}$$

Estes valores tornam praticamente inviável a radiodifusão no formato digital, uma vez que exigem canais de comunicação com banda passante muito elevada. Como comparação, as capacidades nominais de alguns canais típicos (atualmente usados para transmissão de TV analógica), caso operassem com modulação adequada para transmissão digital, estão apresentadas na tabela 8.1 abaixo.

Meio	Relação S/R	Modulação	Taxa Bruta	Taxa Líquida
Cabo Coaxial	> 30 dB	256-QAM	42.88 Mb/s	38 Mb/s
Radiodifusão	> 15 dB	8-VSB	38 Mb/s	19 Mb/s
Radiodifusão	> 20 dB	OFDM x 64-QAM	30 Mb/s	22 Mb/s
Satélite (DTH)	> 6 dB	QPSK	45 Mb/s	30 Mb/s

Tabela 8.1 – Capacidades de Canais de Comunicação

(Taxa Bruta = Taxa Líquida + Códigos de Correção de Erros)

Para permitir a radiodifusão de TV digital através dos canais disponíveis, é necessário adotar processos de compressão **com perdas**, baseados na redução de redundância temporal (semelhança entre quadros consecutivos de uma imagem) e espacial (correlação entre *pixels* adjacentes). Estes processos devem obviamente levar em conta as características da percepção visual humana para permitir altas taxas de compressão (maiores que 50:1) apresentando, no entanto, reduzida visibilidade dos erros provocados na reconstrução da imagem.

7. Referências

- Digital Image Processing* - Rafael C. Gonzales, Richard E. Woods - Addison Wesley, 1992
- Digital Video and HDTV* – Charles Poynton – Morgan Kaufmann, 2003
- Digital Visual Interface Rev. 1.0* – Digital Display Working Group Standard, 1999
- International Telecommunications Union: Rec. ITU-R BT.656-4, 1998
- International Telecommunications Union: Rec. ITU-R BT.709-6, 2015
- International Telecommunications Union: Rec. ITU-R BT.2020-2, 2015
- NAB Engineering Handbook* – J. Whitaker, ed. - National Association of Broadcasters, 1999
- Signals & Systems* – Alan Oppenheim, Alan Willsky – Prentice Hall, 1997
- Television Engineering Handbook* - K. Blair Benson, ed. - Mc Graw-Hill, 1985
- The Digital Interface Handbook* – F. Rumsey, J. Watkinson – Focal Press, 1995
- Understanding IPTV* – G. Held – Auerbach Publications, 2007
- Video Editing and Post Production* – J. R. Caruso, A. E. Arthur – Prentice Hall, 1992
- Video Engineering* - Andrew F. Inglis, Arch C. Luther - Mc Graw-Hill, 1996