PROCESSAMENTO DIGITAL DE SINAIS

Márcio Portes de Albuquerque e Marcelo Portes de Albuquerque
{mpa,marcelo}@cbpf.br

Centro Brasileiro de Pesquisas Físicas (CBPF)
Rua Dr. Xavier Sigaud n^o 150, Urca
Rio de Janeiro, julho de 2002

Resumo

Esta apostila fornece uma visão geral de sinais e métodos de processamento de sinais. A caracterização matemática dos sinais é primeiramente abordada em conjunto com a classificação de sinais. Alguns sinais típicos serão apresentados em detalhes e as informações carregadas por eles serão descritas. Será feita uma revisão em algumas operações utilizadas em processamento de sinais. Discutiremos as vantagens e desvantagens do processamento digital de sinais.

1 Introdução

Sinais realizam um papel importante em nosso cotidiano. Exemplos de sinais que encontramos freqüentemente são voz, música, fotografias e vídeo. Um sinal é uma função de variáveis independentes como tempo, distância, posição, temperatura, pressão. Uma música ou um sinal de voz, por exemplo, representam a pressão do ar como função do tempo em um ponto no espaço. Uma fotografia em preto e branco representa a intensidade de luz como função de duas coordenadas espaciais. Um sinal de vídeo de uma televisão consiste em seqüências de imagens, chamadas quadros, como função de três variáveis: duas coordenadas espaciais e o tempo.

A maioria dos sinais que encontramos são gerados de modo natural. No entanto, um sinal pode ser gerado sinteticamente ou por simulação. Um sinal carrega informação, e o objetivo do processamento de sinais é extrair a informação carregada por este. O método de extração de informação depende do tipo de sinal e da natureza da informação carregada por ele. Portanto, superficialmente falando, processamento de sinais está interessado com a representação matemática de sinais e com as operações algoritmicas realizadas para a extração da informação presente. A representação de sinais pode ser em termos de funções base no domínio das variáveis independentes originais ou em termos de funções base no domínio da transformada. Do mesmo modo, o processo de extração de informação pode ser representado no domínio original do sinal ou no domínio de sua transformada. Neste curso estaremos interessados nos sinais representados no tempo discreto e seu processamento discreto no tempo.

Esta apostila fornece uma visão geral de sinais e métodos de processamento de sinais. A caracterização matemática dos sinais é primeiramente abordada em conjunto com a classificação de sinais. Alguns sinais típicos serão apresentados em detalhes e as informações carregadas por eles serão descritas. Será feita uma revisão em algumas operações utilizadas em processamento de sinais. Discutiremos as vantagens e desvantagens do processamento digital de sinais.

2 Caracterização e classificação de sinais

Dependendo da natureza das variáveis independentes e dos valores das funções que definem o sinal, vários tipos de sinais podem ser definidos. As variáveis independentes, por exemplo, podem ser contínuas ou discretas. Do mesmo modo, os sinais podem ser tanto funções contínuas ou discretas das variáveis independentes. Além disso, os sinais podem ser tanto funções de variáveis reais como de variáveis complexas.

Um sinal pode ser gerado por uma fonte única ou por múltiplas fontes. No primeiro caso, é um sinal escalar e no segundo teremos um vetor de sinais também chamado de sinal em multicanais. Um sinal de uma dimensão(1-D) é uma função de uma única variável independente. Um sinal bidimensional (2-D) é uma função de duas variáveis independentes. Um sinal multidimensional (M-D) é uma função de mais de uma variável. O sinal de voz é um exemplo de sinal unidimensional (1-D), onde a variável independente é o tempo. Um sinal de imagem, como uma fotografia, é um exemplo de sinal bidimensional (2-D), onde as variáveis independentes são as duas coordenadas espaciais. Um sinal de vídeo em preto e branco é um sinal 3-D, onde cada quadro é uma imagem 2-D em função de duas variáveis espaciais e uma temporal. O sinal de vídeo colorido, por sua vez, é um sinal de três canais cada um com três dimensões, onde um canal é o sinal 3-D das intensidades de vermelho (R), o segundo é um canal com as intensidades do verde (G) e o terceiro é um canal com as intensidades do azul (B). Para a finalidade de transmissão, o sinal RGB da televisão é transformado em outro tipo de sinais de três canais, composto pela luminância e por duas componentes de crominância.

O valor do sinal em valores específicos das componentes independentes é chamado de *amplitude*. A variação da amplitude em função das variáveis independentes é chamado *forma de onda*. Para um sinal 1-D, a variável independente normalmente é chamada de tempo. Se a variável independente for contínua, o sinal é chamado *sinal contínuo no tempo*. Se a variável independente for discreta, o sinal é chamado *sinal discreto no tempo*. Um sinal contínuo no tempo é definido a cada instante de tempo. De outra forma, um sinal discreto é definido em instantes discretos do tempo e, portanto, é uma seqüência de números.

Um sinal contínuo no tempo com amplitudes contínuas normalmente é chamado de *sinal analógico*, como, por exemplo, um sinal de voz. Sinais analógicos são comumente encontrados em nosso cotidiano e são normalmente gerados por meios naturais. Um sinal discreto no tempo com valores discretos de amplitudes são representados por um número finito de dígitos e é referenciado como *sinal discreto*. Um exemplo de sinal digital são músicas digitalizadas e armazenadas em discos de CD-ROM. Um sinal discreto no tempo e contínuo em amplitude é chamado de *sinal de dados amostrados*. Este último tipo de sinal ocorre em circuitos de chaveamento de capacitores (SC). Um sinal digital é portanto uma quantização de um sinal de dados amostrados. Finalmente, um sinal contínuo no tempo com valores discretos de amplitudes é chamado de *quantizer boxcar signal*[1]. A Figura 1 ilustra os quatro tipos de sinais.

A representação matemática da dependência funcional é freqüentemente mostrada explicitamente. Para um sinal 1-D contínuo no tempo, as variáveis independentes contínuas são normalmente denotadas por t, visto que para um sinal 1-D discreto no tempo, as variáveis independentes discretas

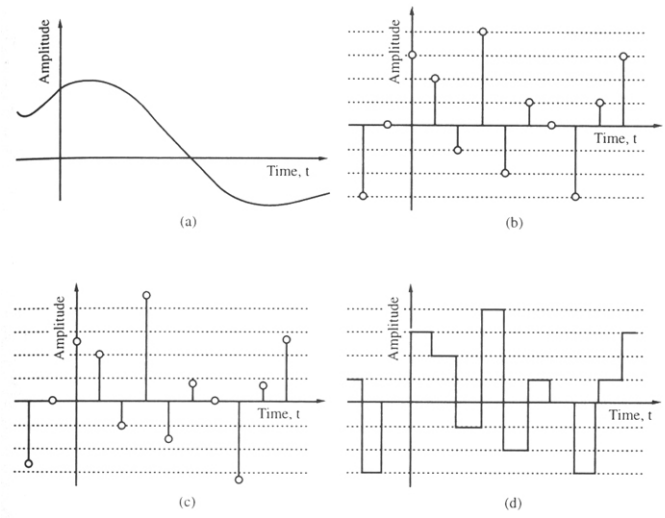


Figura 1: (a) sinal contínuo no tempo, (b) sinal digital, (c) sinal de dados amostrados e (d) *quantizer boxcar signal*

são normalmente denotadas por n. Por exemplo, $u(t)$ representa um sinal 1-D contínuo no tempo e $\{v[n]\}$ representa um sinal 1-D discreto no tempo. Cada membro, $v[n]$, de um sinal discreto no tempo é chamado *amostra*. Em muitas aplicações, sinais discretos no tempo são gerados a partir da amostragem de sinais contínuos no tempo. Se os instantes discretos no tempo estiverem eqüidistantes, ou seja, uniformemente separados, as variáveis independentes discretas n podem ser normalizadas para assumir valores inteiros.

No caso de sinais 2-D contínuos no tempo, as duas variáveis independentes são as coordenadas espaciais, que são normalmente denominadas por x e y. As intensidades de luminosidade em imagens em níveis de cinza, por exemplo, podem ser expressas por $u(x, y)$. De outro modo, uma imagem digital é um sinal discreto 2-D, e suas duas variáveis independentes são discretizadas no espaço normalmente denotado por m e n. Desta forma, uma imagem digital pode ser representada por $v[m, n]$. Do mesmo modo, uma seqüência de vídeo em preto e branco é um sinal 3-D que pode ser representado como $u(x, y, t)$, onde x e y representam as duas variáveis espaciais e t a variável temporal. Um sinal de vídeo colorido é um vetor de sinais composto por três sinais que representam as três cores primárias, que são o vermelho (R), o verde (G) e o azul (B):

$$\mathbf{u}(x, y, t) = \begin{bmatrix} r(x, y, t) \\ g(x, y, t) \\ b(x, y, t) \end{bmatrix} \tag{1}$$

Existe outra classificação de sinais que depende da certeza na qual os sinais são unicamente descritos. Um sinal *determinístico* é aquele definido por um processo como uma expressão ou regra matemática. A geração aleatória de um sinal, no entanto, torna o mesmo imprevisível no tempo e

desta forma é chamado de *sinal aleatório*. Inicialmente iremos ver o processamento de sinais determinísticos discretos no tempo. No entanto, como na prática sistemas discretos no tempo utiliza "palavras´´ finitas para o armazenamento dos sinais e a implementação dos algoritmos de processamento de sinais, é necessário o desenvolvimento de ferramentas para análise dos efeitos de palavras de comprimento finito na performance de sistemas discretos no tempo. Para esta finalidade, foi conveniente representar alguns sinais pertinentes como sinais aleatórios e aplicar técnicas estatísticas em suas análises. Algumas operações típicas de processamento de sinais serão revisadas na próxima seção.

3 Operações típicas de processamento de sinais

Diversos tipos de operações de processamento de sinais são utilizadas na prática. No caso de sinais analógicos, a maioria das operações de processamento de sinais são normalmente aplicadas no domínio do tempo, enquanto que no caso de sinais discretos, são utilizadas operações tanto no domínio do tempo como no da freqüência. Em ambos os casos, as operações desejadas são implementadas como uma combinação de algumas operações elementares. Estas operações são usualmente implementadas em tempo real, em algumas aplicações, elas podem ser implementadas "*off-line´´*.

4 Operações elementares no domínio do tempo

As três operações mais elementares no domínio do tempo são: adição, atraso e escalamento. O *escalamento* consiste na multiplicação do sinal por uma constante positiva ou negativa. No caso de sinais analógicos esta operação normalmente é chamada amplificação se a magnitude da constante de multiplicação, chamada ganho, for maior que 1. Caso contrário a operação é chamada de atenuação. Portanto se $x(t)$ for um sinal analógico, a operação de escalamento gera um sinal $y(t) = \alpha x(t)$ onde α é a constante de multiplicação. Duas outras operações elementares podem ser mencionadas, a integração e a diferenciação. A integração de um sinal aleatório $x(t)$ gera o sinal $y(t) = \int_{-\infty}^{t} x(\xi)d\xi$, enquanto sua diferenciação resulta em um sinal $w(t) = dx(t)/dt$.

A operação de *atraso* gera um sinal idêntico ao original e deslocado no tempo. Para sinais analógicos $x(t)$, $y(t) = x(t-t_0)$ é o sinal obtido pelo atraso de t_0 (positivo) em $x(t)$. Se t_0 for negativo ocorrerá uma operação de *avanço* no sinal.

Muitas aplicações necessitam operações com mais de dois sinais para a geração de um sinal. Por exemplo, $y(t) = x_1(t) + x_2(t) - x_3(t)$ é o sinal gerado pela adição de três sinais analógicos $x_1(t)$, $x_2(t)$ e $x_3(t)$. Outra operação elementar é o produto de dois sinais. Portanto, o produto de dois sinais $x_1(t)$ e $x_2(t)$ geram o sinal $y(t) = x_1(t)x_2(t)$.

As operações elementares mencionadas acima também podem ser utilizadas em sinais discretos no tempo, como discutiremos mais tarde no texto. Em seguida, vamos rever algumas operações complexas bastante utilizadas em processamento de sinais, que são uma combinação de duas ou mais operações elementares.

5 Filtragem

Uma das operações mais utilizadas em processamento de sinais é a filtragem. Sua utilização é passar algumas componentes de freqüência de um sinal através de um sistema, sem que haja nenhuma distorção, e bloquear outras componentes de freqüência. Os sistemas que implementam estas operações são chamados de filtros. Vários tipos de filtros podem ser definidos, dependendo da natureza da operação de filtragem. Na maior parte dos casos, a operação de filtragem para sinais analógicos é linear e é descrita pela integral de convolução:

$$y(t) = \int_{-\infty}^{\infty} h(t-\xi)x(\xi)d\xi \tag{2}$$

onde $x(t)$ é o sinal de entrada e $y(t)$ a saída do filtro caracterizada pela resposta ao impulso $h(t)$.

Um filtro passa-baixa permite a passagem de todas as componentes de freqüência abaixo de f_c, denominada de freqüência de corte, e bloqueia todas as altas freqüências acima de f_c. Um filtro passa-alta realiza a tarefa oposta, deixando passar as freqüências superiores a f_c e rejeitando as freqüências abaixo de f_c. Um filtro passa-banda permite a passagem de freqüências dentro do intervalo definido por f_{c1} e f_{c2} e rejeita qualquer freqüência fora deste intervalo. O filtro rejeita-banda realiza a operação contrária ao filtro passa banda, rejeitando somente as freqüências definidas dentro da faixa f_{c1} e f_{c2} e permitindo a passagem de quaisquer outras componentes de freqüência. A Figura 2(a) apresenta um sinal composto por três componentes de freqüência com 50Hz, 110Hz e 210Hz, respectivamente. As Figuras 2(b) a (e) apresentam os resultados das quatro operações de filtragem mencionadas acima.

O filtro rejeita-banda permite a passagem de somente uma componente de freqüência, sendo conhecido como filtro $notch$. Um filtro multi-banda é aquele que possui mais de um filtro passa-banda e rejeita-banda. Um filtro *comb* é projetado para rejeitar freqüências que são múltiplos inteiros de freqüências baixas.

Um sinal pode ser corrompido por interferências indesejáveis chamadas ruído. Em muitas aplicações os sinais desejados ocupam bandas de freqüências baixas desde DC até f_LHz, são corrompidos por ruídos de altas freqüências acima de f_HHz com $f_H > f_L$. Nestes casos, podemos remover o ruído aplicando um filtro passa-baixa com freqüência de corte f_c onde $f_L < f_c < f_H$. Uma fonte comum de ruído são radiações de campos elétricos e magnéticos. Os sinais gerados pelas linhas de força aparecem como uma senóide com freqüência de 60Hz corrompendo o sinal desejado. Este ruído pode ser removido por um filtro $notch$ de 60Hz.

6 Geração de sinais complexos

Os sinais podem possuir valores reais ou valores complexos. Todos os sinais gerados naturalmente são sinais de valores reais, no entanto desejamos desenvolver um sinal complexo a partir de um sinal real. Este sinal complexo pode ser obtido a partir da *transformada de Hilbert* do sinal real, que é caracterizada pela resposta ao impulso $h_{HT}(t)$ dada por [2, 3]:

$$h_{HT} = \frac{1}{\pi t} \tag{3}$$

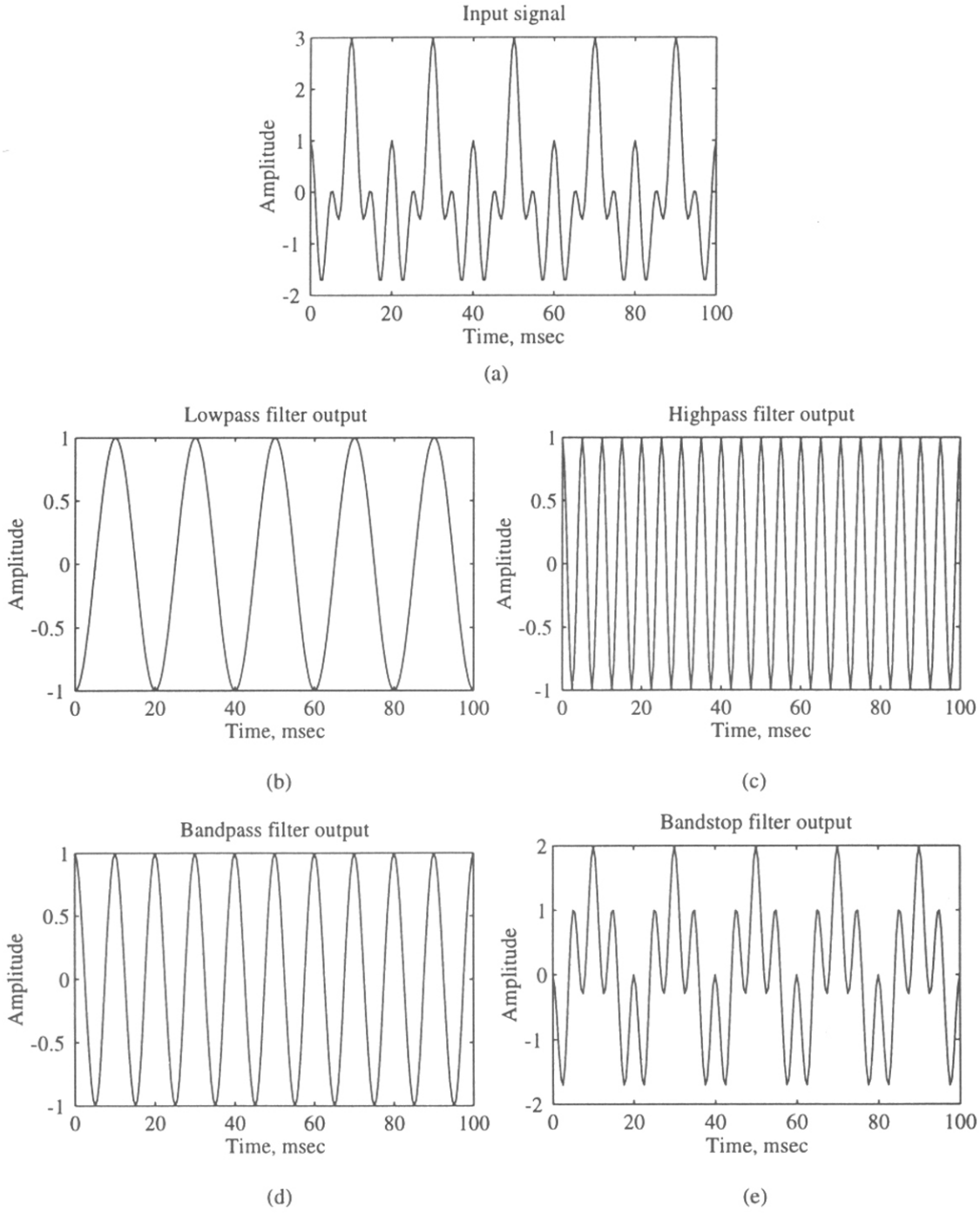


Figura 2: (a) sinal de entrada, (b) saída do filtro passa-baixa com freqüência de corte de 80Hz, (c) saída do filtro passa-alta com $f_c = 150$Hz , (d) saída de um filtro passa-banda com $f_{c1} = 80$Hz e $f_{c2} = 150$Hz, (e) saída de um filtro rejeita-banda com $f_{c1} = 80$Hz e $f_{c2} = 150$Hz.

Para ilustrar o método, considere um sinal real $x(t)$ com a transformada de Fourier contínua no tempo $X(j\Omega)$ dada por:

$$X(j\Omega) = \int_{-\infty}^{\infty} x(t)e^{-j\Omega t}dt \tag{4}$$

$X(j\Omega)$ é chamado de espectro de $x(t)$. A magnitude do espectro de um sinal real apresenta a simetria par enquanto o espectro de fase apresenta a simetria ímpar. Portanto, o espectro $X(j\Omega)$ de um sinal real $x(t)$ contém tanto freqüências positivas como negativas e por isso pode ser expresso como:

$$X(j\Omega) = X_p(j\Omega) + X_n(j\Omega) \tag{5}$$

onde $X_p(j\Omega)$ é a porção de freqüências positivas de $X(j\Omega)$ e $X_n(j\Omega)$ é a porção de freqüências negativas de $X(j\Omega)$. Se realizarmos a transformada de Hilbert em $x(t)$ a saída $\hat{x}(t)$ será dada pela convolução linear de $x(t)$ com $h_{HT}(t)$:

$$\hat{x}(t) = \int_{-\infty}^{\infty} h_{HT}(t-\xi)x(\xi)d\xi \tag{6}$$

O espectro $\hat{X}(j\Omega)$ de $\hat{x}(t)$ é dado pelo produto da transformada de Fourier de $x(t)$ e $h_{HT}(t)$ da equação (3), que é dado por:

$$H_{HT}(j\Omega) = \begin{cases} -j, & \Omega > 0 \\ j, & \Omega < 0 \end{cases} \tag{7}$$

Desta forma,

$$\hat{X}(j\Omega) = H_{HT}(j\Omega)X(j\Omega) = -jX_p(j\Omega) + jX_n(j\Omega) \tag{8}$$

as funções pares e ímpares correspondem a magnitude e fase, respectivamente. Considere o sinal $y(t)$ formado pela soma de $x(t)$ e $\hat{x}(t)$:

$$y(t) = x(t) + \hat{x}(t) \tag{9}$$

os sinais $x(t)$ e $\hat{x}(t)$ são chamados de componentes em fase e quadratura de $y(t)$. Utilizando as Equações (5) e (8) na transformada discreta de Fourier de $y(t)$, nós obtemos

$$Y(j\Omega) = X(j\Omega) + \hat{X}(j\Omega) = 2X_p(j\Omega) \tag{10}$$

Em outras palavras, o sinal complexo $y(t)$, chamado de sinal analítico, possui somente componentes com freqüências positivas.

A Figura 3 apresenta um diagrama de blocos que ilustra o processo de obtenção de um sinal analítico a partir de um sinal real.

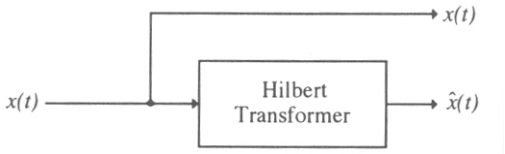


Figura 3: Geração de um sinal analítico utilizando transformada de Hilbert.

7 Modulação e demodulação

Para transmissão de sinais em longas distâncias são utilizados meios de transmissões como cabo, fibra óptica ou a atmosfera. Cada um desses meios possui uma banda que é mais eficiente na transmissão de sinais de alta freqüência. Como resultado, para transmitirmos um sinal de baixa freqüência em um canal, é necessário transformar este sinal para alta freqüência através da operação de modulação. Ao chegar no receptor o sinal tem que ser demodulado para obtenção do mesmo em baixa freqüência. Existem quatro tipos de modulação de sinais analógicos: modulação em amplitude, modulação em freqüência, modulação em fase e modulação em quadratura. Discutiremos, nesta apostila, os conceitos da modulação em amplitude [2, 3].

No esquema de modulação em amplitude, a amplitude de um sinal de alta freqüência senoidal $A\cos(\Omega_0 t)$, chamado de sinal da portadora, é variado pelo sinal de baixa freqüência $x(t)$, chamado de sinal de modulação, gerando um sinal de alta freqüência $y(t)$, chamado de sinal modulado de acordo com:

$$y(t) = Ax(t)\cos(\Omega_0 t) \tag{11}$$

Desta forma, a modulação em amplitude pode ser implementada através do produto do sinal de modulação com o sinal da portadora. Para demonstrar a propriedade de translação em freqüência do processo de modulação por amplitude, tomaremos $x(t) = \cos(\Omega_1 t)$ onde Ω_1 é muito menor que a freqüência Ω_0 da portadora, i.e., $\Omega_1 << \Omega_0$. Pela Equação (11) obtemos:

$$\begin{aligned} y(t) &= A\cos(\Omega_1 t).\cos(\Omega_0 t) && (12)\\ &= \frac{A}{2}\cos((\Omega_0 + \Omega_1)t) + \frac{A}{2}\cos((\Omega_0 - \Omega_1)t) && (13) \end{aligned}$$

Portanto, o sinal modulado $y(t)$ é composto por dois sinais senoidais de freqüências $\Omega_0 + \Omega_1$ e $\Omega_0 - \Omega_1$, que estão próximas a Ω_0 uma vez que assumimos que Ω_1 é muito menor que a freqüência da portadora Ω_0.

Através das propriedades da transformada de Fourier podemos mostrar que o espectro $Y(j\Omega)$ de $y(t)$ é dado por:

$$Y(j\Omega) = \frac{A}{2}X(j(\Omega - \Omega_0)) + \frac{A}{2}X(j(\Omega + \Omega_0)) \tag{14}$$

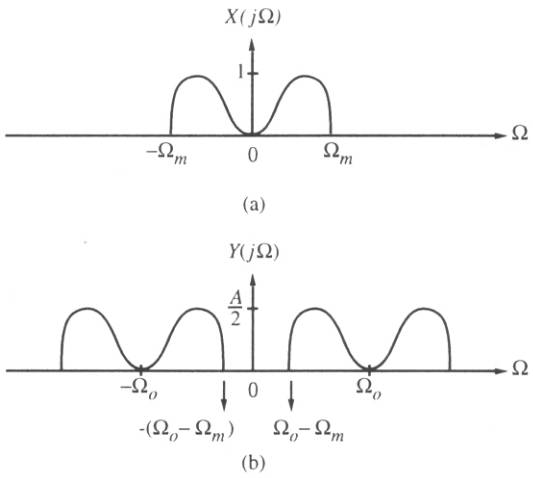


Figura 4: (a) Espectro do sinal modulado $x(t)$, e (b) espectro do sinal modulado $y(t)$. Por conveniência, ambos espectros são de funções reais.

onde $X(j\Omega)$ é o espectro do sinal modulado $x(t)$. A Figura 4 apresenta o espectro do sinal de modulação e do sinal modulado assumindo que a freqüência da portadora Ω_0 é muito maior que Ω_m, que é a freqüência mais alta contida em $x(t)$. Como ilustrado na figura $y(t)$ é agora limitado por banda de alta freqüência e centrado em Ω_0. Este processo de modulação é chamado de *Double Sideband Supressed Carrier* (DSB-SC), devido as porções de freqüência $\Omega_0 + \Omega_m$ e $\Omega_0 - \Omega_m$.

A demodulação de $y(t)$, é feita em dois estágios. Primeiro, o produto de $y(t)$ com um sinal senoidal de mesma freqüência da portadora é formado:

$$r(t) = y(t)\cos(\Omega_0 t) = Ax(t)\cos^2(\Omega_0 t) \tag{15}$$

que pode ser reescrito como

$$r(t) = \frac{A}{2}x(t) + \frac{A}{2}x(t)\cos(2\Omega_0 t) \tag{16}$$

Este resultado indica que o produto do sinal é composto pela modulação do sinal senoidal escalado por um fator 1/2 e uma modulação do sinal em amplitude com uma freqüência de portadora de $2\Omega_0$. O espectro $R(j\Omega)$ de $r(t)$ está indicado na Figura 5. O sinal modulado original pode ser recuperado a partir de $r(t)$ através da utilização de um filtro passa baixas com freqüência de corte Ω_c satisfazendo a relação $\Omega_m < \Omega_c < 2\Omega_0 - \Omega_m$. A saída do filtro é então uma réplica escalada do sinal de modulação.

A Figura 6 apresenta a representação em diagrama de blocos dos esquemas de modulação e demodulação em amplitude. O princípio fundamental na assunção do processo de demodulação descrito acima é que o sinal senoidal deve ser idêntico ao sinal da portadora e pode ser gerado no receptor. Em geral, é difícil garantir que o sinal senoidal de demodulação possui freqüência igual ao da portadora

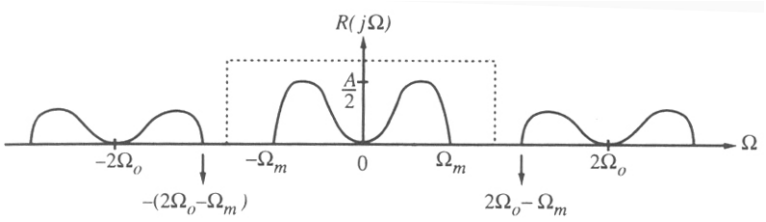


Figura 5: Espectro do produto do sinal modulado e da portadora.

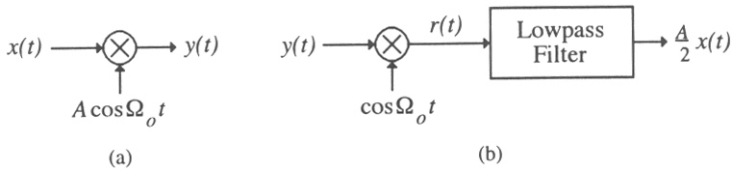


Figura 6: Representação esquemática da modulação e demodulação em amplitude: (a) modulador e (b) demodulador.

todo o tempo. Para solucionar o problema, na transmissão do sinal modulado em amplitude, é incluído o sinal da portadora. Isto é feito através da redefinição da operação de modulação como segue:

$$y(t) = A[1 + mx(t)]\cos(\Omega_0 t) \tag{17}$$

onde m é o número escolhido para assegurar que $[1 + mx(t)]$ seja positivo para todo t. A Figura 7 apresenta a forma de onda do sinal de modulação senoidal de freqüência 2Hz e modulação de amplitude por portadora obtida de acordo com a Equação (17) para uma freqüência de portadora de 40Hz e $m = 0.5$. Note que a envoltória do sinal modulado pela portadora é a forma de onda do sinal de modulação. A portadora também está presente no sinal modulado, o processo é chamado de *Double Sideband Modulation* (DSB).

8 Multiplexação e demultiplexação

Para uma utilização eficiente do canal de transmissão em banda larga, vários sinais de banda estreita e baixas freqüências são combinados formando um sinal de banda larga que é transmitido como um único sinal. O processo de combinação destes sinais é chamado de *multiplexação* que é implementado de forma a assegurar que uma réplica dos sinais originais de baixa freqüência e banda estreita podem ser recuperados no receptor. O processo de recuperação destes sinais no receptor é chamado de *demodulação*.

Um método amplamente utilizado que combina diferentes sinais de voz em um sistema de comunicação é chamado *Frequency Dvision Multiplexing* (FDM) [4, 3]. Cada sinal de voz tipicamente é limitado por uma banda de baixa freqüência com largura $2\Omega_m$, a freqüência é transladada em alta

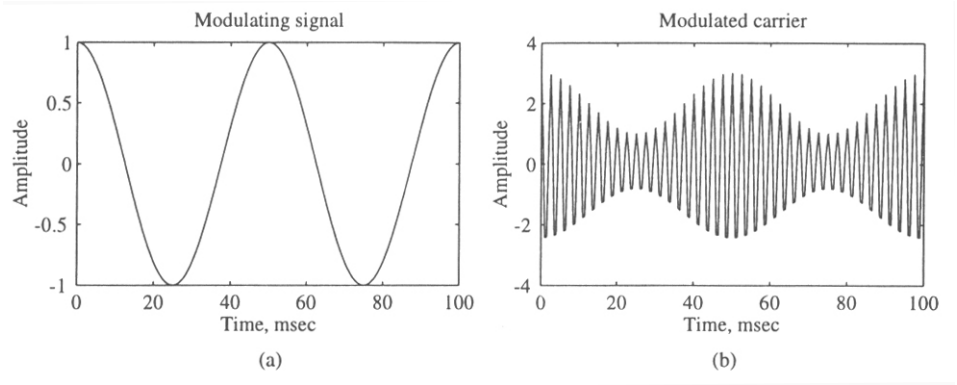


Figura 7: (a) Um sinal de modulação senoidal de 2Hz, e (b) sinal da portadora com a freqüência de 40Hz baseado na modulação DSB.

freqüência utilizando o método de modulação por amplitude da Equação (11). A freqüência da portadora do sinal adjacente modulado por amplitude é separada por Ω_0, com $\Omega_0 > 2\Omega_m$ para assegurar que não haverá nenhuma sobreposição nos sinais individuais modulados. Depois o sinal é modulado em uma portadora principal e transmitido. A Figura 8 ilustra o esquema do processo de multiplexação por divisão de freqüência.

No receptor, a decomposição dos sinais em banda é feita por demodulação FDM. Depois cada sinal individual transladado em freqüência é demultiplexado através de filtros passa banda com banda um pouco maior que $2\Omega_m$. A saída do filtro passa banda é demodulado utilizando o método da Figura 6 recuperando uma réplica escalada do sinal original de voz.

No caso da modulação em amplitude convencional, como apresentado na Figura 4, o sinal modulado possui largura de banda $2\Omega_m$. Para aumentar a capacidade de transmissão média, uma forma modificada da modulação em amplitude é freqüentemente empregada, onde são transmitidas as bandas altas ou as baixas do sinal. Este método é chamado de *Single Sideband Modulation* (SSB) para distinguir do método DSB representado no esquema da Figura 6.

Uma maneira de implementar a modulação em amplitude SSB está indicado na Figura 9, onde a transformada de Hilbert é definida pela equação (7). Os espectros referentes aos sinais da Figura 9 são apresentados na Figura 10.

9 Por quê processamento digital de sinais?

A origem das técnicas de processamento digital de sinais foi no século XVII quando métodos de diferenciação finita, métodos de integração numérica e métodos de interpolação numérica foram desenvolvidos para resolver problemas físicos envolvendo variáveis e funções contínuas. O interesse mais recente de processamento de sinais digitais apareceu em 1950 com a disponibilização de um número crescente de computadores digitais. As aplicações iniciais foram a simulação de métodos de

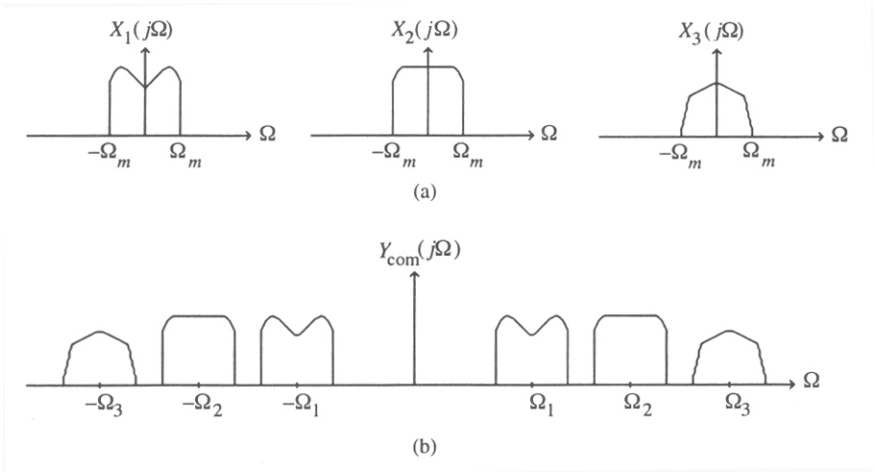


Figura 8: Ilustração da operação de multiplexação por divisão de freqüências.

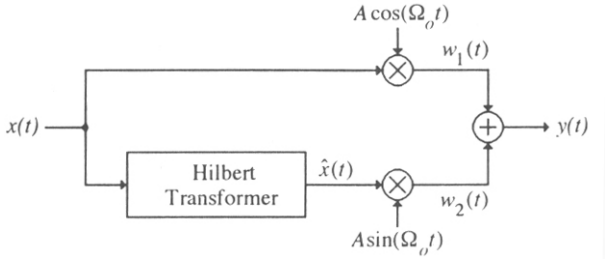


Figura 9: Esquema de modulação SSB empregando a transformada de Hilbert.

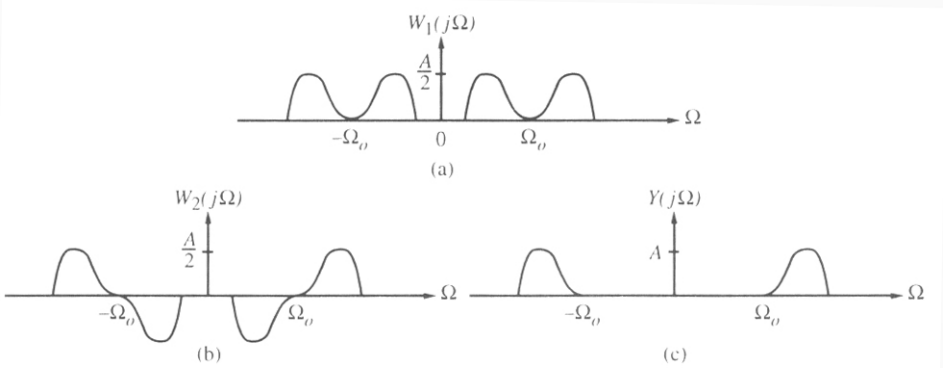


Figura 10: Sinais pertinentes ao espectro da Figura 9.



Figura 11: Diagrama de blocos dos passos de um processamento digital de um sinal analógico.

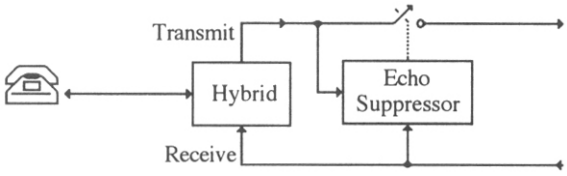


Figura 12: Esquema de supressão de eco.

processamento analógico de sinais. No início da década de 1960, pesquisadores passaram a considerar o processamento digital de sinais como uma nova área disciplinar. Desde então ocorreram descobertas e desenvolvimentos significantes tanto na teoria quanto nas aplicações de processamento digital de sinais.

O processamento digital de sinais analógicos consiste basicamente de três passos: conversão do sinal analógico para digital, processamento da versão digital e a conversão do sinal digital de volta em analógico. A Figura 11 ilustra o diagrama de blocos desse processo.

Como a amplitude do sinal analógico varia com o tempo um circuito de *sample-and-hold* (S/H) é utilizado para amostrar o sinal analógico em intervalos periódicos congelando o valor da amostra com valor constante na entrada do circuito de conversão analógico-digital (A/D), permitindo, desta forma, uma maior fidelidade na conversão dos dados. A entrada de um conversor A/D é um sinal analógico do tipo escada se o circuito S/H segurar a amostra até o instante seguinte de conversão. A saída de um conversor A/D é uma cadeia de dados binários que será processada em seguida pelo algoritmo desejado de processamento digital de sinais. A saída do processador digital é outra cadeia binária de dados que será convertida em um sinal analógico tipo escada através do conversor digital-analógico (D/A). O filtro passa-baixa na saída do conversor D/A remove todas as componentes em altas freqüências indesejáveis e é obtido o sinal analógico na saída através do processamento digital. A Figura 13 apresenta as formas de onda dos sinais em cada um dos estágios do processamento digital, os sinais binários foram representados por pulsos positivos e negativos.

O processamento analógico de sinais analógicos é conceitualmente muito mais simples e envolve somente um único processador, como ilustra a Figura 14. É natural perguntarmos quais as vantagens do processamento digital de sinais analógicos.

Diferentemente dos circuitos analógicos, as operações nos circuitos digitais não dependem de valores precisos dos sinais digitais. Como resultado, um circuito digital é menos sensível as tolerâncias

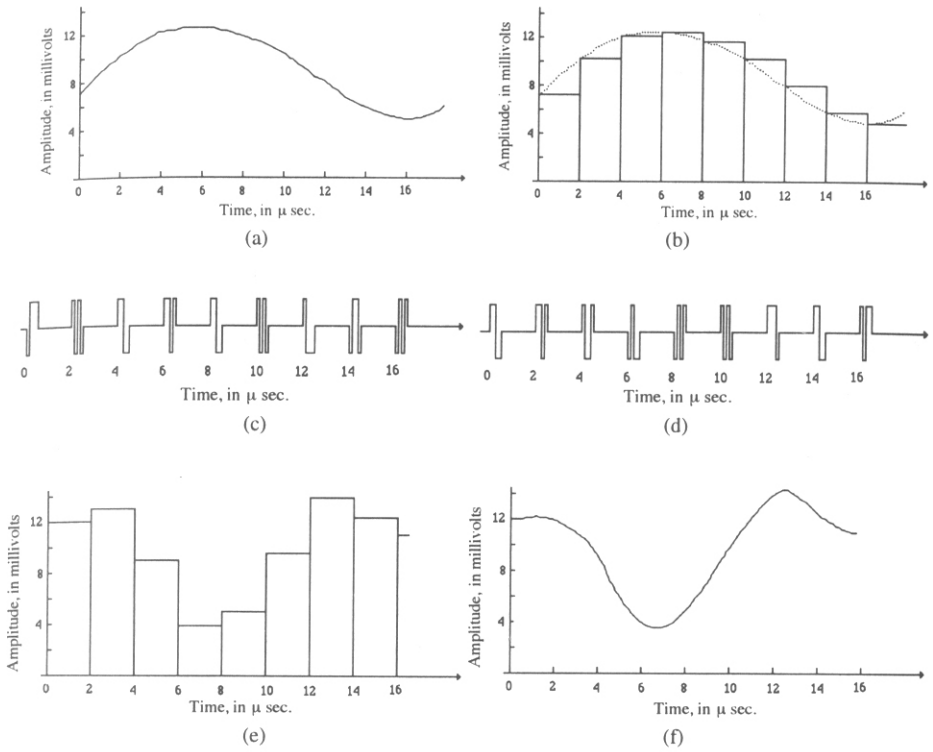


Figura 13: Formas de onda típicas dos sinais nos diversos estágios da Figura 12. (a) Sinal de entrada analógico, (b) circuito de saída S/H, (c) saída do conversor A/D, (d) saída do processador digital, (e) saída do conversor D/A, e (f) saída do sinal analógico. Em (c) e (d), os níveis digitais ALTO e BAIXO são os pulsos positivos e negativos.

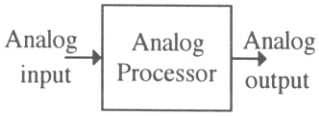


Figura 14: Processamento analógico de sinais analógicos.

dos valores dos componentes, é razoavelmente independente da temperatura, envelhecimento e da maioria dos parâmetros externos. Um circuito digital pode ser reproduzido facilmente em larga escala e não requer nenhum ajuste durante a fabricação ou em sua utilização. Além disso, é simples sua integração completa como, por exemplo, em circuitos VLSI, o que foi possível uma integração sofisticada e complexa em sistemas de processamento digital de sinais em um único chip.

Em um processador digital, os sinais e os coeficientes que descrevem as operações de processamento são representados por palavras binárias. Desta forma, a precisão desejada pode ser alcançada variando o tamanho da palavra, estando sujeito a variações nos custos. Além disso, a faixa dinâmica para sinais e coeficientes pode ser aumentada, se necessário, utilizando aritmética de ponto flutuante.

A implementação digital permite a realização de certas características que não são possíveis na implementação analógica, como, por exemplo, processamento multitaxa. Circuitos digitais podem ser cascateados sem que haja problemas de carga, diferentemente dos circuitos analógicos. Os sinais digitais podem ser armazenados indefinidamente sem que haja nenhuma perda de informação em vários tipos diferentes de mídia como, por exemplo, fitas e discos magnéticos e discos ópticos. Estes sinais armazenados podem ser processados posteriormente "*off-line´´*. Os sinais analógicos, por sua vez, são deteriorados rapidamente com o tempo o não podem ser recuperados em sua forma original.

Os processadores, utilizados nas implementações digitais, podem ser ajustados facilmente o que é fundamental em diversas implementações como, por exemplo, na filtragem adaptativa. Estes ajustes podem ser realizados pela mudança dos coeficientes do filtro nos algoritmos executados pelos processadores. Outra vantagem da aplicabilidade do processamento digital está nos sinais em baixa freqüência, como nos sinais sísmicos, onde indutores e capacitores necessários no processamento analógico seriam fisicamente muito grandes para serem utilizados.

O processamento digital de sinais possui desvantagens também. Uma desvantagem óbvia é o aumento na complexidade dos sistemas com processamentos digitais de sinais analógicos, onde são necessários dispositivos de pré e pós-processamento como, por exemplo, os conversores A/D, D/A e seus filtros associados.

Uma segunda desvantagem dos processamentos digitais está na limitação das freqüências disponíveis para processamento, o que limita, em particular, sua aplicação no processamento de sinais analógicos. Em geral, sinais analógicos contínuos no tempo devem ser amostrados com freqüências superiores a duas vezes a maior componente de freqüência existente no sinal[3]. Se esta condição não for satisfeita, as componentes do sinal acima da metade da freqüência de amostragem apareceram como freqüências abaixo deste valor, o que irá distorcer a forma de onda do sinal analógico. A faixa máxima de freqüência que pode ser processada digitalmente é estipulada pelo circuito de amostragem, *i.e.*, pelo conversor A/D.

Uma terceira desvantagem é o fato dos sistemas digitais serem constituídos por dispositivos ativos que consomem energia elétrica. O processamento analógico, por outro lado, utiliza indutores, capacitores e resistores, que são elementos passivos, *i.e.*, não necessitam de energia elétrica. Além disso, dispositivos ativos são de menos confiança que os componentes passivos.

No entanto as vantagens compensam as desvantagens em inúmeras aplicações, principalmente com a contínua queda no preço dos processadores digitais podemos observar um número crescente de aplicações em processamento digital de sinais.

Referências

[1] K. Steiglitz. *Mathematical foundations of signal processing*. In S.K. Mitra and J.F. Kaiser, editors, *Handbook for Digital Signal Processing*, chapter 2, pages 57-59. Wiley-Interscience, New York, 1993.

[2] M.E. Frerking, *Digital signal procesing im communication systems*. Van Nostrand Reinhold, New York, 1994

[3] A.V. Oppenheim and A.S. Willsky. *Signals and sistems*. Prentice-Hall, Englewood Cliffs, 1989.

[4] L.W. Cooley and J.W Turkey. *An algorithm for the machine calculation of complex Fourier series*. Macmillan, New York, 1983