

Transformações sonoras através de operações timbrais

JOSÉ EDUARDO FORNARI¹, FURIO DAMIANI^{1,2}

1. Depart. Semicondutores Instrumentos e Fotônica - DSIF/FEE/UNICAMP
 2. Núcleo Interdisciplinar de Comunicações Sonoras - NICS/UNICAMP
 CEP. 13.081-970 Campinas-São Paulo - Brasil
 e-mail: Fornari@dsif.fec.unicamp

Resumo:

Descreve-se o projeto de um sistema de processamento digital para áudio (ADSP), em tempo-real. O sinal de áudio de entrada é digitalizado, tem seu espectro de frequência determinado, é submetido a operadores que modificam seu espectro e é transformado em sinal analógico de saída. O espectro de frequência é obtido pela transformada rápida de Fourier (FFT), sendo sua inversa (IFFT) utilizada para gerar o sinal de saída. Os operadores timbrais manipulam as magnitudes dos harmônicos que compõem o espectro, e podem ser parametrizados por controladores externos, bem como pelo "passado" do sinal sonoro de entrada.

1. Do som ao timbre:

O som é constituído por vibrações mecânicas, expansões e compressões longitudinais que se propagam em um meio material com uma velocidade característica e pode ser caracterizado por três grandezas físicas (frequência, intensidade e duração) (Parker). A variação temporal da intensidade pode ser mostrada *macroscopicamente*, pela envoltória da intensidade no tempo (ataque, sustentação e decaimento) e *microscopicamente* pelas variações temporais da amplitude dentro da envoltória, com frequência aproximadamente periódica, percebendo-se assim o efeito de seus componentes harmônicos [Alles].

O sentido da audição humana possui limites. Limites de intensidade: ouvem-se sons com intensidades entre: 10^{12} W/m² e 10^{-1} W/m², correspondentes aos limiares de percepção e de dor. A percepção da variação da intensidade é de cerca de 10% da intensidade sonora total. Limites de frequência: aproximadamente entre 20 e 20.000 Hz (um piano produz sons com frequências fundamentais entre 27 Hz e 3480 Hz). A maior sensibilidade à variação de frequência fica na faixa de 500 a 4.000 Hz. para a qual é de cerca de 0,3% da frequência sendo ouvida, ou 1/20 de semitom [Culver]. Limite de duração: Conhecido como persistência auditiva, é de aproximadamente 50ms.

O timbre é uma percepção subjetiva, e sua conceituação procura definir o que é percebido pela audição que permite reconhecer e distinguir um som. O timbre depende de vários fatores, como a envoltória do espectro de frequência e de suas variações temporais, as variações periódicas da amplitude, a frequência fundamental, se o som for melódico ou percussivo" [Plomp]. A *American Standards Association* o define: "considerando estritamente o estado de regime sonoro, timbre é: o atributo da sensação auditiva pelo qual um indivíduo pode julgar se dois sons complexos e sustentados, com a mesma altura e loudness são diferentes entre si".

Uma maneira de representar o som de modo a tornar as informações que constituem o timbre observáveis, é através de uma superfície, chamada aqui de superfície sonora ou SS, que representa a variação temporal das magnitudes dos componentes do espectro do som. O conjunto das componentes de frequência de um som num dado instante é o seu *espectro de frequência*. De maneira geral, esse espectro é limitado e contínuo (formado por infinitos componentes dentro dos limites de frequência da audição), sendo a SS uma superfície contínua com tres dimensões: frequência, amplitude e tempo. Uma seção da SS ao longo de uma mesma frequência gera um plano onde a magnitude dessa frequência é representada em função do tempo, ou *plano morfológico*; uma seção para um dado instante de tempo representa o espectro de frequência deste instante, ou *plano harmônico*.

A SS pode ser discretizada, dando origem a uma matriz que a pode representar adequadamente desde que os erros de discretização estejam dentro das limitações da percepção auditiva. Isto torna possível o processamento digital de manipulação do timbre.

2. As transformações sonoras

As transformações são operações matemáticas para o mapeamento, ou representação, de um conjunto de pontos de um espaço vetorial R^n em outro espaço vetorial R^m . Podem ser lineares e não-lineares. Uma transformação linear deve satisfazer a: $F(u+v) = F(u) + F(v)$, e $F(k.v) = F(v)$, caso contrário, ela é dita não-linear [Boldrini]. Os equipamentos para transformações sonoras podem ser agrupados em classes, pela grandeza física que transformam. Transformações da duração: É a classe dos equipamentos que atuam na envoltória temporal do som, como reverberadores, retardadores, câmaras de eco. Transformações do espectro: É a classe do que atuam na envoltória de frequência do som, como os filtros, deslocadores espectrais, *Chorus*. Transformações da intensidade: É a classe dos equipamentos que atuam na amplitude do som, como os prolongadores, equalizadores, compressores de nível. As transformações sonoras nunca transformam exclusivamente uma única grandeza sonora, elas atuam predominantemente nela. As transformações que agem, simultânea e correlacionadamente, em mais de uma grandeza, como as transformações não-lineares, são classificadas como *transformações mistas*.

O som tanto pode ser transformado tanto analógica quanto digitalmente. Respeitando-se os limites de percepção da audição humana, a representação digital do som o pode representar adequadamente (sem perdas perceptíveis) [Chamberlin]. Os sistemas de transformação sonora digital estão sendo cada vez mais utilizados devido a sua grande flexibilidade na manipulação de dados, ao crescente e significativo aumento de suas velocidades de processamento e à diminuição de seu custo. O processador digital de sinais, DSP (Digital Signal Processor), quando especializado para processamento sonoro, é chamado de processador digital de áudio, ADSP (Audio DSP).

Um som contínuo possui um espectro de frequência com um número infinito de componentes da análise de Fourier, obtíveis com o uso da transformada de Fourier. No caso de sua representação digital, para uma amostra de N pontos, obtém-se um espectro discreto com N/2 componentes, calculados pela transformada discreta de Fourier. O algoritmo otimizado que os calcula é a transformada rápida de Fourier, FFT. A representação digital da SS será uma matriz cujas linhas correspondem a um dado instante discreto no tempo e cujas colunas correspondem às frequências discretas componentes do espectro, sendo os valores numéricos dos elementos da matriz iguais às magnitudes das componentes desse espectro. Esta matriz será chamada de matriz harmônica, ou MH.

3. A operação timbral

Definimos como transformação do timbre a transformação de cada instante discretizado de um som, re-parametrizada a cada novo instante de tempo. Uma operação timbral realiza uma transformação do espectro, através do controle e da parametrização desta operação em tempo-real. Esta transformação resulta na modificação do timbre sonoro. O operador proposto é um sistema ADSP que realiza operações timbrais pela manipulação de um espectro discreto contido em um vetor. A operação timbral é aplicada ao espectro do som de entrada e parametrizada por um bloco de controle externo. Este recebe e processa informações de interfaces externas e da MH. O operador timbral, ou OT, realiza todas as operações vetoriais necessárias num intervalo de tempo suficientemente curto para ser considerado pela audição como tempo-real.

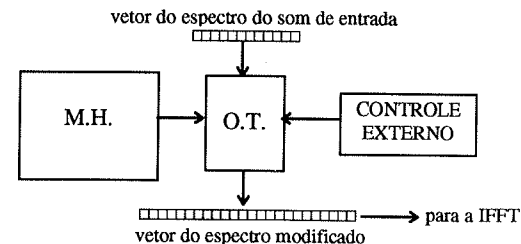


Fig. 1 Estrutura básica de um operador timbral

4. A estrutura do sistema

A entrada e saída são analógicas, sendo o processamento interno digital. Na entrada faz-se o pré-processamento e conversão A/D do sinal, $s(t)$. O espectro de $s(n)$ é obtido pelo algoritmo FFT para seqüências de amostras com 50ms. A magnitude da saída complexa da FFT é alocada em um vetor de elementos reais, que representa o controle do operador timbral. O operador timbral manipula o vetor que representa um novo espectro, de acordo com seus parâmetros. Este novo espectro é formado pelas operações timbrais. Todas as operações timbrais ocorrem num ciclo de operação do OT, cujo intervalo de tempo é sincrono com a IFFT. Após cada ciclo do OT, o vetor de saída é escrito na MH e encaminhado para o processamento de sua IFFT e conversão D/A, gerando a saída analógica $s'(n)$. O sistema possui um atraso entre a entrada $s(t)$ e a saída $s'(t)$ aproximadamente igual a $(\tau_1 + \tau_2 + \tau_3) \leq 50ms$.

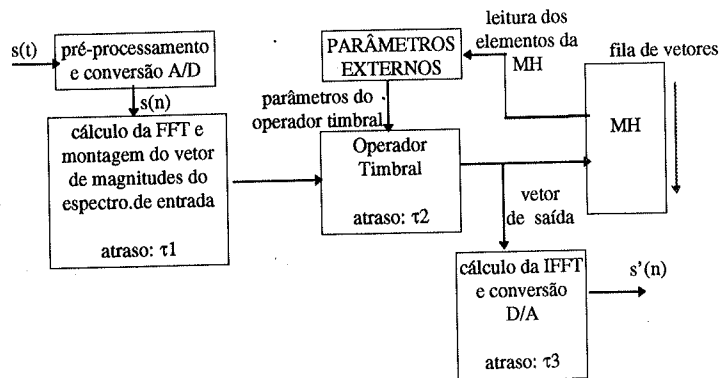


Fig. 2 Estrutura do sistema de transformação sonora

5. A simulação do OT

O ambiente de simulação do MATLAB 4.0 foi utilizado para o desenvolvimento de programas de simulação das transformações sonoras através de operações timbrais. Uma vez que se trata de um software, as operações timbrais não ocorrem em tempo-real.

As operações timbrais foram simuladas pelos programas descritos abaixo. Utilizaram-se os seguintes valores: taxa de amostragem $F_s=8.192$ Hz., pontos de entrada da FFT: $N = 256$, resolução do eixo da frequência: $F_s/N = 32$ Hz. e intervalo de tempo correspondente a um espectro: $N/F_s = 31ms$. Cada vetor do espectro representa a magnitude dos harmônicos em $N/2 = 128$ pontos.

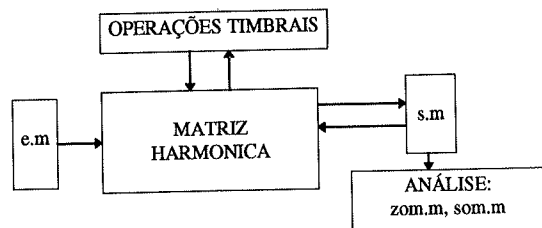


Fig. 3 Estrutura para a simulação do operador timbral.

e.m: lê um arquivo de som no padrão u-law, calcula a MH através da FFT e monta gráficos: do som e da SS. **s.m:** monta um arquivo de som através da IFFT aplicada na MH e monta gráficos da SS modificada e da envoltória temporal do som. **zom.m:** mostra um detalhe de um intervalo de tempo do som modificado e seu espectro. **som.m:**

permite a audição do arquivo de som original e modificado. As operações timbrais foram feitas na forma de programas que utilizam operações vetoriais básicas sobre as linhas de MH, tais como: adição, subtração, produto, divisão, inversão, deslocamento dos seus elementos. Através destes foram elaboradas operações timbrais, como: compressão dos níveis máximo e mínimo de superfície sonora, expansão (ou compressão) da magnitude do espectro do som, modificação da envoltória da SS, redução do número de harmônicos do espectro (preserva apenas os n maiores), modificação da altura (frequência fundamental). Observa-se que a transformação sonora realizada através dos operadores timbrais permite uma grande flexibilidade de manipulação e visualização dos resultados no timbre do som. O resultado de algumas simulações feitas é mostrado a seguir.

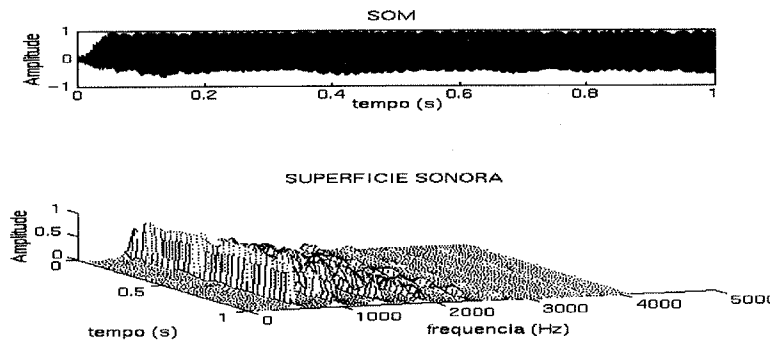


Fig. 4 Amostra de som de violino. Envoltória temporal da amplitude e SS.

Utilizou-se o som amostrado de um violino, emitindo continuamente a nota A4 (440 Hz) durante um intervalo maior que 1 segundo. Observa-se a sua envoltória no tempo e superfície sonora para o primeiro segundo de som. Nota-se a grande possibilidade de observação dos detalhes macroscópicos e microscópicos som, como: pouco ataque, presença de muitos harmônicos, centralização do harmônico de maior magnitude (fundamental) sobre o ponto do eixo da frequência correspondente a 440 Hz, variação da envoltória do espectro ao longo do tempo e oscilações da intensidade do som ao longo do tempo, entre outros.

Aplicou-se sobre a matriz harmônica desse som uma operação timbral que expande o espectro do som, na razão de 1 para 2, ou seja, os harmônicos componentes do espectro são distanciados linearmente entre si e a sua envoltória passa a ocupar o dobro do intervalo em frequência que originalmente ocupava.

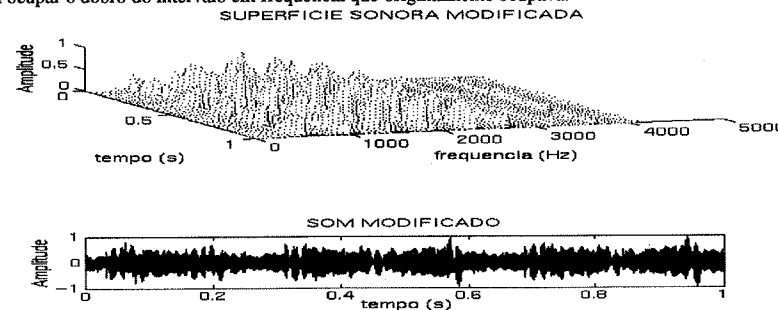


Fig.5 Violino. Transformação por expansão dos espectros (2:1). Envoltória temporal da amplitude e SS.

O resultado é mostrado tanto na forma de superfície sonora, quanto de envoltória do som no tempo. Observa-se o espalhamento do espectro, que antes ocupava uma região menor de intervalo de frequência. Passam a existir componentes em frequências mais graves e também mais agudas.

A seguir utilizou-se como arquivo de entrada o som de uma nota contínua emitida por uma flauta, por mais de 1 segundo.

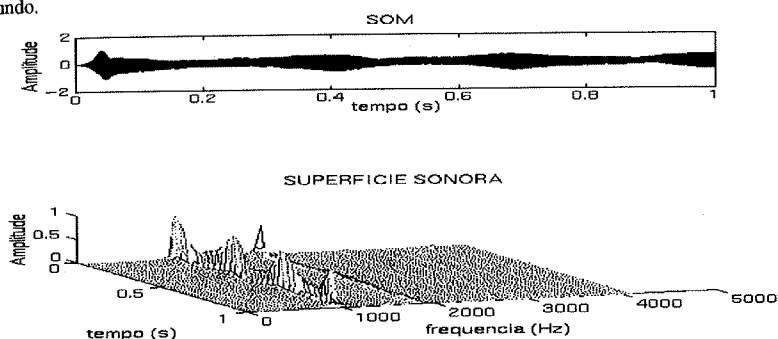


Fig. 6 Amostra de som de flauta. Envoltória temporal da amplitude e SS.

Pode-se observar uma maior quantidade de ataque e menor quantidade de harmônicos presentes do que na amostra de violino. Durante o ataque, ocorrem harmônicos que depois desaparecem que são relacionados ao som "assoprado" do instante inicial da emissão de uma nota deste instrumento.

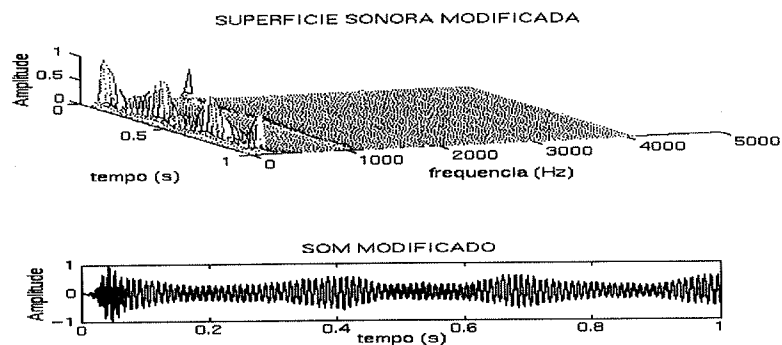


Fig.7 Flauta. Transformação por deslocamento da fundamental de 100 Hz para 200 Hz. Envoltória temporal da amplitude e SS

Aplicou-se um deslocamento linear dos harmônicos do espectro para frequências mais graves. O resultado que se observa na envoltória é que o tempo de ataque e das ondulações da intensidade do som permanecem inalterados. No entanto o som todo tornou-se bem mais grave, podendo até se observar na envoltória do tempo a periodicidade correspondente aos harmônicos de baixa frequência.

6. Resultados e possibilidades

A simulação do sistema foi feita através de programas que fazem o papel de operações timbrais, e dessa maneira, modificam os elementos contidos na matriz harmônica. Em tempo-real, as operações manipulam o espectro e armazenam o resultado na MH.

Também é importante salientar que as simulações foram feitas sobre uma superfície sonora, cujos eixos de amplitude, frequência e tempo são todos lineares. No entanto, a percepção auditiva, para a intensidade e a frequência do som, são aproximadamente logarítmicos. Portanto, as operações timbrais, programadas no sistema em tempo-real, devem levar em conta a natureza da audição humana para estas grandezas. A resolução e os limites máximo e

mínimo de cada eixo da superfície sonora deve corresponder a percepção auditiva para as variações da sua grandeza correspondente. Para o tempo, a escala é linear, com resolução menor que 50 ms. Para a amplitude, a escala é logarítmica (dB), com resolução melhor que 3dB. No caso do som amostrado, deve-se ter uma conversão A/D com resolução de 20 bits, afim de garantir $SQNR = 20 \cdot \log_{10} 2^{20} = 120$ dB [Chamberlin]. Para a frequência, a escala deve ser modelada de acordo com a sensibilidade auditiva $\Delta f/f$, que equivale a 0,3% entre 500 a 4.000 Hz e aumenta (torna-se menos sensível) abaixo e acima dessa região.

Através da simulação do sistema proposto foi possível observar algumas possibilidades de transformações sonoras que as operações timbrais permitem. Deve-se levar em conta que o sistema é uma ferramenta sonora, e que as operações timbrais poderão ser programadas por software, e executadas em tempo-real. Isto mostra que as suas possibilidades de utilização são muito grandes. Pode-se citar algumas:

Composição timbral Através de um material sonoro inicial, o artista pode criar novos timbres, tendo a SS como interface visual e pelos resultados sonoros.

Modelamento virtual do timbre O modelamento timbral pode também ser feito através de interfaces controladoras para realidade virtual, uma vez que existe a possibilidade de observação em tempo-real das transformações timbrais.

Performances artísticas Podem-se variar as transformações no timbre de um instrumento durante uma performance. Isto pode ser feito tanto por controladores que o próprio instrumentista opere, quanto por um outro artista, que as promova ou ainda por programas pre-estabelecidos em software.

Tratamentos e análises sonoras Este sistema pode ser útil para estudos de gravação e em laboratórios de análise sonora. A operação em tempo-real, dispensa o uso de grandes capacidades de memória, para o armazenamento do trecho sonoro que vai ser processado ou analisado. Todos os tratamentos sonoros, como: compressões, equalizações inteligentes, correções de pitch, detecções de padrões sonoros, podem ser feitos através da programação de operadores timbrais adequados. O som passa apenas uma vez através do sistema e na saída tem-se todas as transformações, ou detecções, executadas.

Bibliografia

- Alles G. Harold. *Music Synthesis Using Real Digital Techniques*. Proceedings of IEEE, vol. 68, n.4, april. 1980.
 Boldrini, José Luis. *Algebra Linear*. 3.ed. cap. 5. São Paulo : Harper & Row do Brasil. 1980.
 Chamberlin, Hal. *Musical applications of microprocessors*, cap. 2. 2ed. Indianapolis, Hayden Books. 1985.
 Culver, C. A. *Musical Acoustics*. 4.ed. cap. 5. McGraw-Hill. 1968.
 Parker, Sybil P. *Acoustics source book*. McGraw-Hill. San Francisco. 1987.
 Roseberg, Martin E. *Sound and Hearing. Studies in Biology n.145*. cap. 2. Afd Sonologie. 1982.
 Plomp, Reiner. *Aspects of the tone sensation, A psychophysiological study*. Academic Press. London, New York, San Francisco. A subsidiary of Hancourt Brace Favonovich Publisher. 1976.