DESCRIÇÃO

O objetivo inicial do projeto era realizar a distorção de sinais de áudio de forma inteiramente digital em tempo real. O processo de distorção de sinais de áudio consiste em passar o sinal por uma função não linear, que diferentemente de uma função linear, gera como harmônicos das frequências fundamentais do sinal, além de combinações das frequências fundamentais presentes nele. Pela natureza digital do sinal, as frequências só podem ser corretamente representadas até a taxa de Nyquist, portanto os harmônicos gerados pela distorção que ultrapassarem essa frequência sofrem rebatimento, sendo representado por alguma frequência digital relacionada à taxa de amostragem.

Para evitar esse fenômeno indesejado, se faz necessário aumentar artificialmente a taxa de amostragem, utilizando o processo de interpolação. Entretanto, isso aumenta consideravelmente o custo computacional, fazendo com que o processamento em uma CPU fique inviável para os requisitos de tempo real do sistema.

Dessa forma o processamento em CPU utilizado em computadores convencionais é substituído pelo desenvolvimento em GPU, que permite descentralizar as operações dos núcleos do processador e distribuí-los em pequenos processamentos paralelos que torna o processo mais rápido e dinâmico. O projeto é baseado no kit de desenvolvimento Jetson TX2 da NVIDIA, que permite a utilização de um módulo dedicado para essa finalidade que apresenta um grande poder de processamento em GPU.

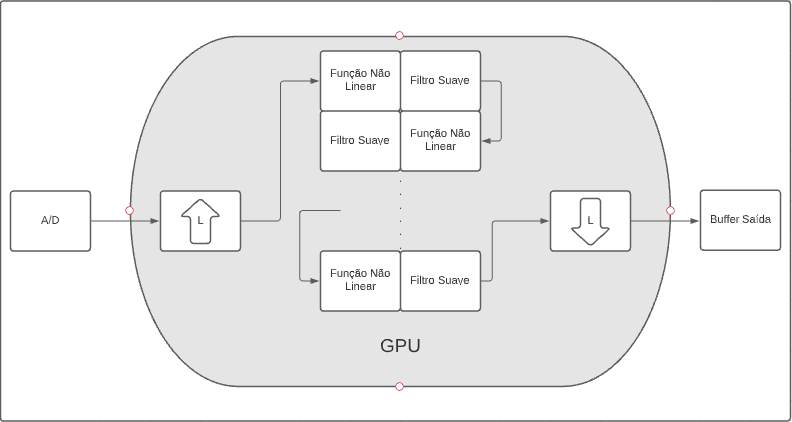
Inicialmente, a entrada de áudio do nosso sistema não estava implementada e era esperada a sua implementação durante o projeto. Entretanto, isso não foi possível devido a alguns problemas com o sistema operacional do Jetson. Portanto, decidimos demonstrar que a implementação em GPU é viável com processamentos offline, usando arquivos de áudio previamente gravados como entrada do sistema e escrevendo o resultado do sistema em um novo arquivo.

Devido ao fato do processamento ser feito de forma offline, utilizar o tempo de processamento do áudio completo como métrica comparativa entre a GPU e CPU não se justifica, pois a leitura das amostras não é feita no tempo de amostragem do sinal de áudio gravado. As amostras são lidas do arquivo a medida que o processador da CPU passa pelo comando de leitura dentro de um loop interno do programa, ou seja, mesmo que uma amostra devesse ser lida com uma periodicidade fixa, ela só é lida no próximo momento que o comando for executado.

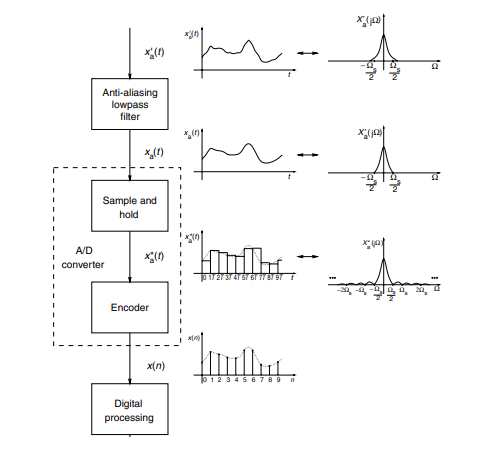
Em um sistema em tempo real, a leitura das amostras na correta taxa de amostragem é controlada pelo driver de áudio do sistema. Sendo assim, para justificar a viabilidade na GPU para processamentos em tempo real, utilizaremos como métrica o delay que o processamento gera desde a leitura da amostra do arquivo de entrada até a escrita da mesma processada no arquivo de saída.

METODOLOGIA - DIAGRAMA DO PROJETO E DIGITALIZAÇÃO

Do ponto de vista do processamento digital, o trabalho segue o diagrama de blocos da Figura 7. Como todo sistema digital que depende de um sinal analógico, o primeiro passo a ser seguido é a digitalização de um sinal proveniente de um instrumento musical. Em seguida, tendo em vista possíveis rebatimentos indesejados durante a distorção, aumenta-se a taxa de amostragem do sistema pelo processo de interpolação. Depois do pré-processamento do sinal mencionado, são aplicadas as distorções no sinal através de funções não lineares. A distorção não é concentrada em apenas um estágio, ela é implementada em uma série de passos, que incluem não só funções mas também filtros suaves, com a finalidade de ter um maior controle sobre o espectro do sinal e tirar proveito do paralelismo. Por fim, o sinal deve ser convertido de volta à sua taxa de amostragem final, com o número de amostras original.

*Figura 7: Diagrama de blocos do processo*

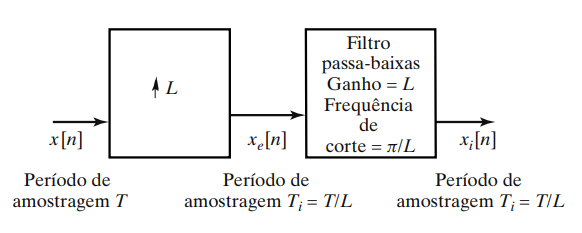
O processo de digitalização consiste na passagem por um filtro anti-aliasing, da amostragem e da quantização. De modo geral, esses três passos são feitos de forma única em um hardware especializado. Tendo em vista as características do ouvido humano, frequências acima de 20 KHz de forma geral não são percebidas, portanto seria suficiente de forma a respeitar o critério de Nyquist amostrar o sinal a 40 Khz. Todavia, comumente sinais de áudio são amostrados a uma taxa maior, de 48 KHz, isso traduz-se no seguinte benefício. Implementações de filtros ideais são não causais e, portanto, impossíveis de serem implementados. No entanto se existir uma margem de segurança acima do critério de Nyquist na taxa de amostragem, a especificação do filtro pode ser relaxada, podendo apresentar uma faixa de transição menos abrupta do que de um filtro ideal, tornando não só o filtro realizável como diminuindo a complexidade computacional do filtro cada vez mais para transições mais brandas. Como um todo, o processo de digitalização segue o diagrama de blocos da Figura 8, onde o bloco de Sample and hold e encoder são a amostragem e a quantização respectivamente.



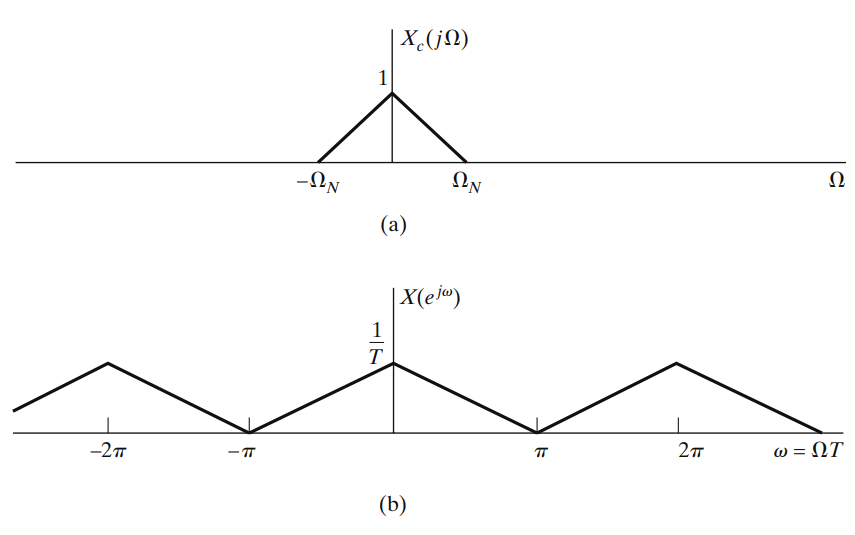
*Figura 8: Processo de Digitalização* (OPPENHEIM e SCHAFER, 2012)*.*

Como segundo passo é realizada a interpolação. O sinal vindo da conversão analógica-digital é limitado em banda pelo filtro do processo de digitalização. Deseja-se realizar a distorção desse espectro, por conta disso novas componentes espectrais serão acrescentadas, como novas harmônicas, possivelmente tornando esse espectro de banda ilimitada. Desse fato, percebe-se que o rebatimento consequente da distorção é inevitável se ele tornar o espectro limitado ou se as novas componentes espectrais estiverem acima da metade da frequência de amostragem, ou seja, 24 KHz, o que limita muito a qualidade da distorção. Uma mitigação do rebatimento pode ser feita de forma bastante razoável através da interpolação. Basicamente, o processo de interpolação por um fator L consiste em superamostragem, inserção de L-1 amostras nulas entre cada amostra do sinal e prosseguido por um filtro de reconstrução que irá redefinir as L-1 amostras nulas com base nas amostras originais; no domínio da frequência este filtro consiste em um passa-baixas. Como resultado, temos que a interpolação “libera” banda pelo menos desde até , sendo que corresponde à metade da frequência de amostragem após a interpolação, que é L vezes maior que a frequência de amostragem antes da interpolação. Assim o sinal a ser distorcido possui mais banda para adicionar novas componentes espectrais e espera-se que perto de , no limiar do rebatimento, elas já são praticamente desprezíveis, tornando o rebatimento também insignificante.

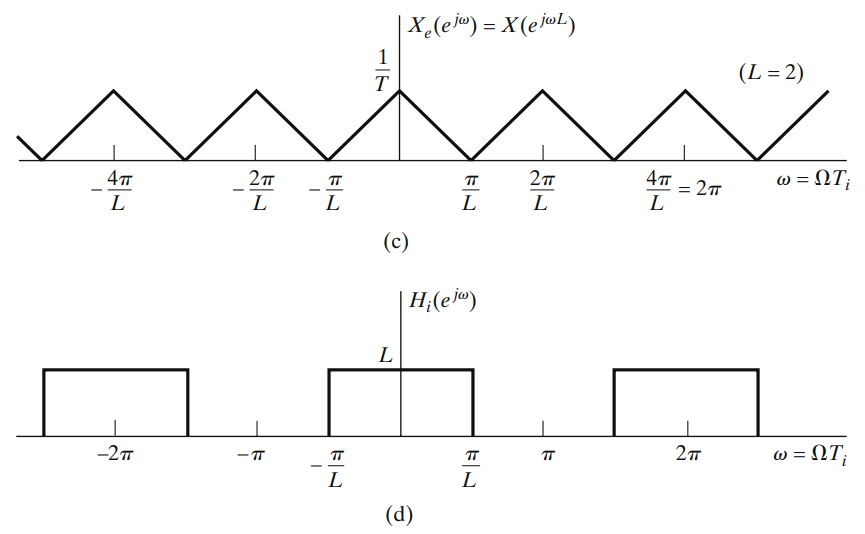
Definimos que componentes abaixo de 90 dB da componente máxima do sinal já são considerados desprezíveis, pois os aparelhos de reprodução de áudio não conseguem reproduzir componentes com intensidades tão distintas. Um diagrama de blocos do processo de interpolação é apresentado na Figura 9. Nesse projeto definimos que a frequência do sinal vindo do conversor A/D deve ser aumentada para valores superiores a 500 KHz, frequência essa difícil obter simplesmente por um processo de amostragem, portanto o fator de interpolação L deve ser maior que 500/48, ou seja, maior que 11. As imagens da Figura 10 mostram como o espectro altera durante o processo de superamostragem, porque um filtro de reconstrução deve ser implementado para recuperar o espectro original do sinal e como o método como um todo libera banda até .



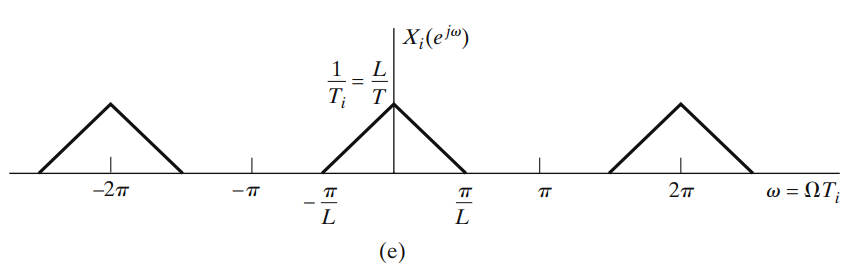
*Figura 9: Diagrama de blocos interpolação* (OPPENHEIM e SCHAFER, 2012)*.*



*Figura 10.1: Espectro do sinal em tempo contínuo e tempo discreto* (OPPENHEIM e SCHAFER, 2012)*.*



*Figura 10.2: Alteração do espectro no processo de sobreamostragem em fator L e o filtro passa-baixas com corte em π/L* (OPPENHEIM e SCHAFER, 2012)*.*



*Figura 10.3: Espectro final após a filtragem* (OPPENHEIM e SCHAFER, 2012)*.*