



Universidade Federal de Pernambuco
Centro de Informática

Graduação em Ciência da Computação

**Síntese de Áudio Realístico: Acelerando a
computação da Radiação Acústica em
GPU**

Rafael Farias Marinheiro

Trabalho de Graduação

Recife
Julho de 2016

Universidade Federal de Pernambuco
Centro de Informática

Rafael Farias Marinheiro

Síntese de Áudio Realístico: Acelerando a computação da Radiação Acústica em GPU

Trabalho apresentado ao Programa de Graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: *Geber Lisboa Ramalho*

Recife
Julho de 2016

<DIGITE A DEDICATÒRIA AQUI>

Agradecimentos

<DIGITE OS AGRADECIMENTOS AQUI>

Science doesn't purvey absolute truth. Science is a mechanism. It's a way of trying to improve your knowledge of nature. It's a system for testing your thoughts against the universe and seeing whether they match. And this works, not just for the ordinary aspects of science, but for all of life. I should think people would want to know that what they know is truly what the universe is like, or at least as close as they can get to it.

—ISAAC ASIMOV (1988)

Resumo

Neste trabalho nós estudamos e descrevemos o estado-da-arte das técnicas de síntese de som realístico. Também desenvolvemos um método para aproximar eficientemente a Radiação Acústica em GPU. Criamos cenários de teste para comparar a performance e a qualidade do nosso sistema com as abordagens anteriores através de benchmarks, testes numéricos e testes A/B com voluntários. A nossa abordagem obteve um speedup superior a $30\times$, mantendo uma taxa de erro inferior a 5%. A maioria dos voluntários também mostraram preferência pelos áudios sintetizados pela nossa abordagem.

Abstract

In this work we study and describe the state-of-the-art techniques in realistic audio synthesis. We also developed a method to efficiently approximate the Acoustic Radiation on GPU. We created test scenarios in order to compare the performance and the quality of our system against the previous techniques through benchmarks, numerical tests and A/B tests with volunteers. Our method obtained an over $30\times$ speedup while maintaining the error rate smaller than 5%. The majority of the volunteers also showed their preference for the audio synthesized with our method.

Sumário

1	Introdução	1
1.1	Motivação	1
1.2	Objetivos	2
1.3	Estrutura do Trabalho	2
2	Fundamentação Teórica	3
2.1	Som e Acústica	3
2.1.1	Equação da Onda	3
2.1.1.1	Fontes Sonoras e Ambientes Acústicos	3
2.1.2	Radiação acústica e a Equação de Helmholtz	5
2.1.2.1	Aproximação em Far-Field	6
2.2	Vibração de Corpos Rígidos	7
2.2.1	Modelo Elastodinâmico	7
2.2.2	Análise Modal	10
3	Trabalhos Relacionados	13
3.1	Métodos Physically-Based	13
3.2	Métodos Data-Driven	14
4	Algoritmo e Implementação	16
4.1	Discretização e Solução da Equação da Onda	16
4.2	Aproximação da Equação de Helmholtz	17
4.2.1	Expansão Multipolar	18
5	Experimentos e Resultados	21
5.1	Teste de Percepção	22
6	Conclusão	26
6.1	Trabalhos futuros	26

Lista de Figuras

1.1	Artista de Foley em um estúdio de produção	1
2.1	Pressão acústica gerada uma fonte pontual	4
2.2	Radiação acústica gerada uma fonte pontual	6
2.3	Radiação de Near-Field e Far-Field	6
2.4	Sistema elastodinâmico unidimensional	7
2.5	Sistema elastodinâmico bidimensional	8
2.6	Modos de Vibração de objetos	10
2.7	Gráfico da solução da equação $\ddot{q}(t) + (\alpha + \beta \omega_i^2)\dot{q}(t) + \omega^2 q(t) = f(t)$	12
2.8	Composição de modos de vibração	12
3.1	Pipeline de Síntese de Áudio Realístico	14
4.1	Discretização Marker-And-Cell	17
4.2	Harmônicos Esféricos	19
4.3	Comparação entre o Ground Truth da Radiação Acústica e a aproximação Expansão Multipolar	20
5.1	Comparação numérica para o Prato de Cerâmica	23
5.2	Comparação numérica para a Caneca de Cerâmica	24
5.3	Comparação numérica para a Chave de Aço	25

Lista de Tabelas

5.1	Parâmetros de Materiais	21
5.2	Especificações de Hardware	21
5.3	Comparação de Performance entre a nossa solução e abordagens anteriores	22
5.4	Resultado dos Testes de Percepção	22

CAPÍTULO 1

Introdução

1.1 Motivação

A nossa audição é um dos nossos sentidos mais importantes. É ela que nos permite identificar sinais e padrões no nosso mundo. Desde os tempo antigos, utilizamos a nossa audição para identificar situações de ameaças iminentes, como o rugido um predador prestes a atacar ou a buzina de um carro em alta velocidade. É também através dela que desenvolvemos o nossos métodos de comunicação.

As indústrias de entretenimento (filmes, músicas, games e etc) utilizam o som para criar um ambiente imersivo para o usuário, criando experiências mais realistas e divertidas. Uma das técnicas mais utilizadas por essa indústria é chamada de Foley. Nessa técnica, o artista de Foley utiliza elementos e objetos do mundo real para tentar sintetizar as amostras de áudio de uma cena. O artista pode, por exemplo, utilizar pedaços de cocos para simular o som do trote de um cavalo [3] ou folhas metálicas para simular o som de trovões (Ver Figura 1.1).



Figura 1.1: Artista de Foley em um estúdio de produção¹

Essa técnica, porém, tem várias limitações. A prática de Foley requer mão-de-obra especializada. Além disso, o tempo de produção escala linearmente com a quantidade de amostras sonoras desejadas. Essa combinação de fatores faz com que o custo da Técnica de Foley seja alto. Em muitos casos também há limitações físicas envolvidas. Seria inviável, por exemplo,

¹Fonte: http://www.t.sonypicturesstudiotours.com/uploads/img/std_content/soundeffects_05.jpg. Acessado em 3 de Julho de 2016.

usar esse tipo de técnica para sintetizar amostras de áudio geradas por uma estrutura metálica de grandes dimensões ou por uma grande esfera de platina.

Em ambientes dinâmicos, no entanto, as limitações são ainda mais aparentes. O uso de amostras sonoras cria um ambiente acústico repetitivo que não se adapta à realidade do mundo virtual. Esse tipo de técnica não permite que o som dos objetos seja alterado de acordo com o estado da simulação em tempo de execução. O som gerado por um prato ao cair no chão, por exemplo, varia drasticamente de acordo com a posição de contato. Esse tipo de comportamento não pode ser simulado adequadamente com o uso de amostras sonoras.

Para resolver esse tipo de problema, um novo ramo da Computação vem crescendo: o ramo da Computação Acústica. Um dos problemas mais interessantes da área é exatamente o problema da Síntese de Áudio Realístico: Dada a especificação de uma cena (geometria dos objetos, materiais dos objetos) e dados os parâmetros dos materiais presentes, como podemos sintetizar o áudio dessa cena? A solução desse problema resolveria exatamente as dificuldades que temos com as técnicas atuais. Isso nos permitiria criar experiências virtuais mais ricas e fiéis à realidade com apenas uma fração do custo atual de produção.

1.2 Objetivos

O objetivo desse trabalho é descrever os principais conceitos e também o estado-da-arte da área de Síntese de Áudio Realístico. Além disso, também queremos desenvolver um método para acelerar a computação da Radiação Acústica em GPU.

1.3 Estrutura do Trabalho

Esse trabalho está estruturado da seguinte maneira: No Capítulo 2 descrevemos os conceitos matemáticos utilizados no estado-da-arte, o que inclui conceitos de Acústica e de Vibrações. No Capítulo 3 descrevemos os trabalhos relacionados, apresentando a evolução da área e discutindo as diferenças entre métodos Physically-Based e métodos Data-Driven. No Capítulo 4 nós apresentamos o nosso método para calcular a Radiação Acústica em GPU e também os detalhes de implementação. No Capítulo 5 nós descrevemos os experimentos utilizados para avaliar o desempenho do nosso método e também discutimos os resultados obtidos. Finalmente, concluímos esse trabalho no Capítulo 6 apresentando também planos para trabalhos futuros.

CAPÍTULO 2

Fundamentação Teórica

2.1 Som e Acústica

Acústica é a ciência que estuda a som. É ciência que estuda a sua origem e propagação, seja ela em espaços abertos ou espaços fechados [10].

2.1.1 Equação da Onda

No estudo de acústica, a principal equação é a *Equação da Onda*. Considerando um meio dispersante, a pressão do meio é governada pela seguinte equação diferencial:

$$\frac{\partial^2 \rho(u, t)}{\partial t^2} = dc^2 \nabla^2 \rho(u, t) \quad (2.1)$$

onde

$$\begin{aligned} t \in \mathbb{R} &\Rightarrow \text{Tempo em segundos (s)} \\ u \in \mathbb{R}^3 &\Rightarrow \text{Posição em metros (m)} \\ c \in \mathbb{R} &\Rightarrow \text{Velocidade de Propagação (m/s)} \\ \rho \in \mathbb{R} &\Rightarrow \text{Densidade kg/m}^3 \\ \rho(u, t) \in \mathbb{R} &\Rightarrow \text{Pressão acústica (Pa)} \end{aligned}$$

Na equação (2.1), o operador ∇^2 corresponde ao operador Laplaciano Espacial. Isto é:

$$\nabla^2 \rho(u, t) = \Delta \rho(u, t) = \frac{\partial^2 \rho(u, t)}{\partial x^2} + \frac{\partial^2 \rho(u, t)}{\partial y^2} + \frac{\partial^2 \rho(u, t)}{\partial z^2} \quad (2.2)$$

A equação (2.1) é uma equação diferencial parcial linear hiperbólica de segunda ordem. Ela é chamada de equação governante pois, determinadas as condições iniciais e as condições de contorno, os demais resultados são obtidos como consequência dela.

2.1.1.1 Fontes Sonoras e Ambientes Acústicos

Em acústica, existem duas entidades essenciais: a *fonte sonora*, responsável por gerar o som, e o *ambiente*, responsável por transmiti-lo. Matematicamente essas entidades definem exatamente as condições iniciais e as condições de contorno da Equação da Onda.

Uma equação diferencial é definida dentro um domínio, usualmente denotado por Ω . O domínio pode ser um ambiente fechado (sala de estar ou um auditório) ou um espaço aberto

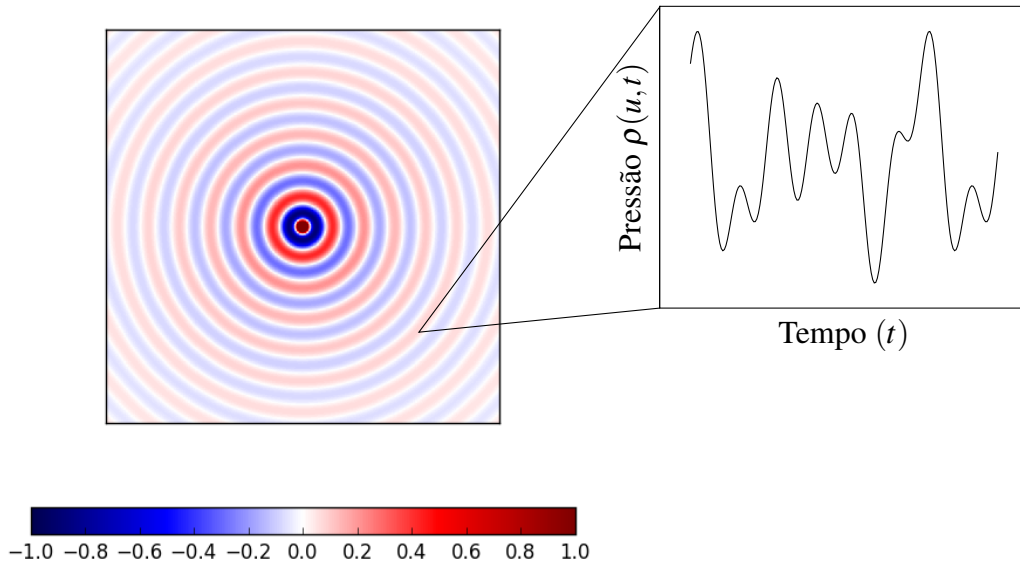


Figura 2.1: Pressão acústica gerada por uma fonte pontual no centro do meio. A fonte e o ambiente determinam as condições de contorno e a pressão acústica nos demais pontos é governada pela Equação da Onda (2.1)

(campo ou um deserto). As condições de contorno são equações que delimitam o comportamento da pressão acústica dentro do domínio Ω . Os dois tipos mais comuns de condições de contorno são as condições de contorno de Dirichlet e as condições de contorno de Neumann.

As condições de contorno de Dirichlet são condições de contorno da forma:

$$\rho(u, t) = f(u, t) \quad (2.3)$$

Essas condições de contorno forçam que o valor de determinados pontos obedeça alguma função já conhecida. Em acústica, as condições de contorno de Dirichlet normalmente são utilizadas para definir fontes sonoras. Se uma determinada fonte sonora localizada no ponto u_f reproduz um som $s_f(t)$, podemos adicionar a condição de contorno $\rho(u_f, t) = s_f(t)$.

As condições de contorno de Neumann são condições de contorno da forma:

$$\frac{\partial \rho(u, t)}{\partial n} = f(u, t) \quad (2.4)$$

onde

$$n \in \mathbb{R}^3 \Rightarrow \text{Vetor normal}$$

Essas condições de contorno forçam que o valor da derivada da pressão em determinados pontos obedeça alguma função já conhecida. As condições de contorno de Dirichlet são utilizadas, por exemplo, para determinar o comportamento na interface entre diferentes meios. Se um ponto u_b encontra-se na borda de uma superfície rígida fixa de normal n_b , podemos adicionar a condição de contorno $\partial \rho(u_b, t) / \partial n_b = 0$.

2.1.2 Radiação acústica e a Equação de Helmholtz

Dadas as condições iniciais e as condições de contorno, a equação da onda nos dá o valor exato da pressão acústica em função da posição e do tempo. A dependência com o tempo, no entanto, nos força a resolver a equação da onda durante a simulação. Esse processo tem um alto custo computacional, portanto é preferível encontrar uma outra solução para esse problema.

O caso mais comum de síntese acústica consiste no problema das fontes sonoras: dada uma fonte sonora na posição u_f emitindo um som $f(t)$, qual a pressão acústica $\rho(u, t)$ num determinado ponto u no tempo t ?

Podemos tentar assumir que a pressão acústica $\rho(u, t)$ depende de duas funções independentes. Isto, é, podemos tentar assumir que:

$$\rho(u, t) = T(u)F(t) \quad (2.5)$$

onde $T(u)$ seria a amplitude no ponto u e $F(t)$ seria uma função que depende apenas som emitido pelas fontes sonoras. Com essa separação, se a fonte passasse a emitir o som $f^*(t)$, precisaríamos apenas computar a função $F^*(t)$ para obter a solução $\rho^*(u, t) = T(u)F^*(t)$.

A função $T(u)$ também é chamada de *Radiação Acústica* ou *Transferência Acústica*, pois ela mede quanta energia acústica foi transferida a partir das fontes sonoras. A equação governante para a Radiação Acústica é a *Equação de Helmholtz*:

$$\nabla^2 T(u) + k^2 T(u) = 0 \quad (2.6)$$

onde

$$\begin{aligned} k \in \mathbb{R} &\Rightarrow \frac{\omega}{c} \rightarrow \text{Número de Onda (1/m)} \\ u \in \mathbb{R}^3 &\Rightarrow \text{Posição em metros(m)} \\ T(u) \in \mathbb{C} &\Rightarrow \text{Amplitude em metros(m)} \end{aligned}$$

Dois fatos são notáveis na Equação de Helmholtz: A constante $k = \frac{\omega}{c}$ depende da frequência ω da fonte sonora. Por tal razão, a Equação de Helmholtz é comumente chamada de Equação da Onda no domínio da frequência.

Note também que a função $T(u)$ tem como contra-domínio o conjunto dos números complexos \mathbb{C} . O uso de números complexos nesse caso é útil para representarmos a fase da onda sonora. Se escrevermos $T(u) = A(u)e^{i\phi(u)}$ e $F(t) = A_0 e^{i\omega t}$, obteremos:

$$\begin{aligned} \rho(u, t) &= \text{Re}(T(u)F(t)) \\ &= \text{Re}\left(A_0 A(u)e^{i(\omega t + \phi(u))}\right) \\ &= \text{Re}\left(A_0 A(u)(\cos(\omega t + \phi(u)) + i \sin(\omega t + \phi(u)))\right) \\ &= A_0 A(u) \cos(\omega t + \phi(u)) \end{aligned} \quad (2.7)$$

Nesse caso, dizemos que a função $\phi(u)$ é a fase da onda sonora.

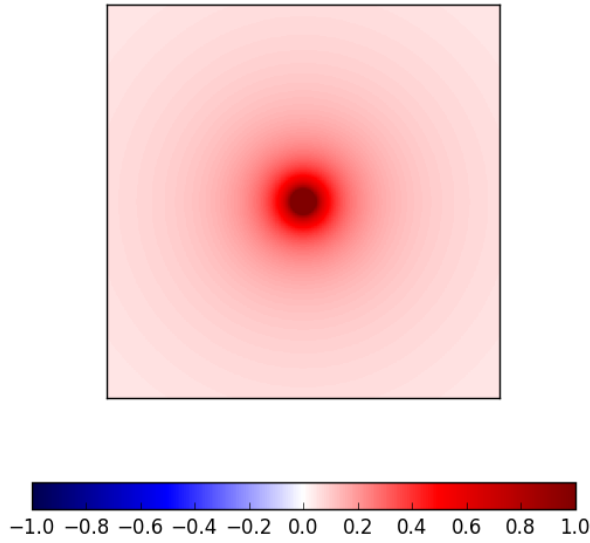


Figura 2.2: Radiação acústica gerada por uma fonte pontual no centro do meio. A fonte e o ambiente determinam as condições do contorno e a radiação acústica nos demais pontos é governada pela Equação de Helmholtz

2.1.2.1 Aproximação em Far-Field

Considere um sistema com algumas fontes sonoras. Seja $B(r)$ a esfera de raio r cujo centro está na origem. Seja r_{near} tal que $B(r_{near})$ contenha todas as fontes sonoras do sistema. Dizemos que a região $B(r_{near})$ é o Near-Field (Campo Próximo) do sistema e que a região externa é o Far-Field (Campo Longínquo) do sistema (Ver Figura 2.3).

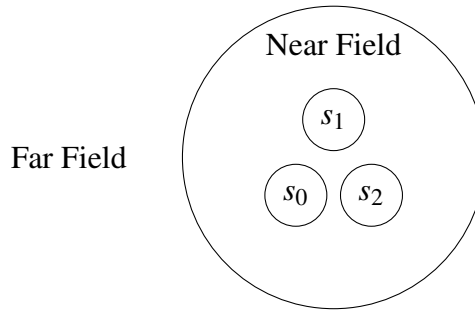


Figura 2.3: Radiação de Near-Field e Far-Field.

Essa separação entre o Near-Field e o Far-Field é importante para efeitos práticos. Toda a radiação acústica é gerada dentro do Near-Field e depois é dissipada para o Far-Field que, por sua vez, não deve gerar radiação acústica. Isso é matematicamente representado pela Condição de Contorno de Sommerfeld, também conhecida por *Condição de Radiação de Sommerfeld*:

$$\lim_{|u| \rightarrow \infty} |u| \left(\frac{\partial}{\partial |u|} - ik \right) T(u) = 0 \quad (2.8)$$

Além disso, também observa-se uma diferença entre a complexidade do Near-Field quando comparado ao Far-Field. Dentro do Near-Field a interação entre as fontes sonoras e os demais objetos faz com que a radiação acústica seja extremamente complexo. Entretanto, a radiação acústica no Far-Field tem uma estrutura muito mais simples.

Seja Γ a superfície do Near-Field e seja u um ponto no Far-Field. A radiação $T(u)$ pode ser calculada pela *Integral de Kirchhoff*:

$$T(u) = \int_{\Gamma} \left[G(u, v) \frac{\partial T}{\partial n}(v) - \frac{\partial G}{\partial n}(u, v) T(v) \right] d\Gamma_v \quad (2.9)$$

onde $G(u, v)$ é a Função de Green da Equação Helmholtz:

$$G(u, v) = \frac{e^{ik\|u-v\|}}{4\pi\|u-v\|} \quad (2.10)$$

Utilizando a fórmula (2.9), vemos que é necessário apenas calcular a Radiação Acústica no Near-Field para definirmos a solução para o Far-Field.

2.2 Vibração de Corpos Rígidos

2.2.1 Modelo Elastodinâmico

O modelo elastodinâmico [23] é um dos modelos mais utilizados no estudo de vibrações. Esse modelo é amplamente utilizado em áreas como Engenharia Civil, Engenharia Mecânica, entre outras.

O objeto de estudo do Modelo Elastodinâmico são os *sistemas elastodinâmicos*. Um sistema elastodinâmico é um sistema composto por corpos i , cada um com massa m_i , e molas (i, j) , cada uma com constante elástica $k_{i,j}$. Um exemplo de sistema elastodinâmico é apresentado na Figura 2.4.

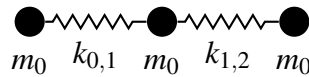


Figura 2.4: Sistema elastodinâmico unidimensional. Cada nó tem massa m_i e cada mola tem uma constante $k_{i,j}$

De acordo com a Segunda Lei de Newton, sabemos que para cada corpo i ,

$$f_{tot,i} = f_{int,i} + f_{ext,i} = m_i \cdot \ddot{p}_i \quad (2.11)$$

onde $f_{ext,i}$ é a força externa, $f_{int,i}$ a força interna, m_i a massa do corpo e \ddot{p}_i é a segunda derivada da posição (aceleração) do corpo i .

Considere um sistema unidimensional com três corpos, como na Figura 2.4, podemos escrever a Equação (2.11) na forma matricial:

$$\begin{bmatrix} f_{tot,0} \\ f_{tot,1} \\ f_{tot,2} \end{bmatrix} = \begin{bmatrix} f_{int,0} \\ f_{int,1} \\ f_{int,2} \end{bmatrix} + \begin{bmatrix} f_{ext,0} \\ f_{ext,1} \\ f_{ext,2} \end{bmatrix} = \begin{bmatrix} m_0 \cdot \ddot{p}_0 \\ m_1 \cdot \ddot{p}_1 \\ m_2 \cdot \ddot{p}_2 \end{bmatrix} = \begin{pmatrix} m_0 & 0 & 0 \\ 0 & m_1 & 0 \\ 0 & 0 & m_2 \end{pmatrix} \begin{bmatrix} \ddot{p}_0 \\ \ddot{p}_1 \\ \ddot{p}_2 \end{bmatrix} = M\ddot{\vec{p}} \quad (2.12)$$

A matriz diagonal M é chamada de *Matriz de Massas* do sistema.

Seja $\vec{p}_{eq} = [p_{eq,0} \ p_{eq,1} \ p_{eq,2}]^T$ o vetor coluna que representa a posição de equilíbrio dos corpos na Figura 2.4. Considere uma outra configuração \vec{p} . Se denotarmos por $\vec{u} = \vec{p} - \vec{p}_{eq}$ o vetor de deslocamento, podemos utilizar a Lei de Hooke para escrever:

$$\vec{f}_{int} = \begin{bmatrix} f_{int,0} \\ f_{int,1} \\ f_{int,2} \end{bmatrix} = \begin{bmatrix} k_{0,1}(u_1 - u_0) & + & 0 \\ k_{0,1}(u_0 - u_1) & + & k_{1,2}(u_2 - u_1) \\ 0 & + & k_{1,2}(u_1 - u_2) \end{bmatrix} \quad (2.13)$$

$$= \begin{pmatrix} -k_{0,1} & k_{0,1} & 0 \\ k_{0,1} & -(k_{0,1} + k_{1,2}) & k_{1,2} \\ 0 & k_{1,2} & -k_{1,2} \end{pmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \quad (2.14)$$

$$= Ku \quad (2.15)$$

A matriz simétrica K é chamada de *Matriz de Rigidez* do sistema. Como $\vec{u} = \vec{p} - \vec{p}_{eq}$, temos que $\ddot{\vec{p}} = \ddot{\vec{u}}$. Desse modo, podemos substituir (2.15) em (2.12):

$$\vec{f}_{ext} + Ku = M\ddot{\vec{u}} \quad (2.16)$$

De modo semelhante, podemos utilizar essa equação para descrever sistemas com maior dimensionalidade. Um exemplo de tal sistema é apresentado na Figura 2.5.

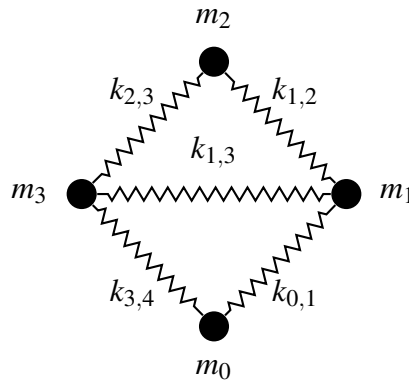


Figura 2.5: Sistema elastodinâmico bidimensional. Cada nó tem massa m_i e cada mola tem uma constante $k_{i,j}$

Para tal sistema, podemos descrever sua Matriz de Massas e sua Matriz de Rigidez. A Matriz de Massas M é a matriz diagonal em blocos:

$$M = \begin{pmatrix} M_0 & 0 & 0 & 0 \\ 0 & M_1 & 0 & 0 \\ 0 & 0 & M_2 & 0 \\ 0 & 0 & 0 & M_3 \end{pmatrix} \text{ onde } M_i = \begin{pmatrix} m_i & 0 \\ 0 & m_i \end{pmatrix} \quad (2.17)$$

A sua Matriz de Rigidez K pode ser escrita como uma soma de matrizes $K_{i,j}$ para cada mola. Cada matriz $K_{i,j}$ tem uma simples estrutura em blocos¹:

$$K_{i,j}[l,c] = \begin{cases} (l,c) = (i,i) \text{ ou } (l,c) = (j,j) \Rightarrow k_{i,j}R_{i,j} \\ (l,c) = (i,j) \text{ ou } (l,c) = (j,i) \Rightarrow -k_{i,j}R_{i,j} \\ \text{Caso contrário} \Rightarrow 0 \end{cases} \quad (2.18)$$

$R_{i,j}$ é a matriz de rotação que alinha o vetor $\overrightarrow{p_j - p_i}$ com o eixo cartesiano x . A matriz $K_{1,3}$, por exemplo, teria a seguinte estrutura:

$$K_{1,3} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & k_{1,3}R_{1,3} & 0 & -k_{1,3}R_{1,3} \\ 0 & 0 & 0 & 0 \\ 0 & -k_{1,3}R_{1,3} & 0 & k_{1,3}R_{1,3} \end{pmatrix} \quad (2.19)$$

A generalização dessa equação é chamada de Equação da Elastodinâmica Linear:

$$M\ddot{u} + C\dot{u} + Ku = f_{ext} \in \mathbb{R}^n \quad (2.20)$$

onde

$$\begin{aligned} u \in \mathbb{R}^n &\Rightarrow \text{Vetor de deslocamento} \\ f_{ext} \in \mathbb{R}^n &\Rightarrow \text{Vetor de forças externas} \\ M \in \mathbb{R}^{n \times n} &\Rightarrow \text{Matriz de Massas} \\ K \in \mathbb{R}^{n \times n} &\Rightarrow \text{Matriz de Rigidez} \\ C \in \mathbb{R}^{n \times n} &\Rightarrow \text{Matriz de Amortecimento} \end{aligned}$$

Para objetos complexos, as matrizes são calculadas utilizando o Método dos Elementos Finitos [7]. O método consiste em subdividir o objeto em pequenas partes, chamados de Elemento. As molas do sistema são utilizadas para unir os elementos que estão em contato. A massa dos Elementos e as constantes da mola são normalmente calculadas a partir das propriedades do material do objeto, como densidade (kg/m^3), Módulo de Young (Pa) e Razão de Poisson [23]. A Matriz de Amortecimento, no entanto, geralmente é calculada utilizando a aproximação de Rayleigh:

$$C = \alpha M + \beta K \quad (2.21)$$

onde α e β são constantes de dependem do material.

¹ $M[l,c]$ representa o elemento da matriz M presente na linha l e coluna c

2.2.2 Análise Modal

Considere um sistema elastodinâmico. A sua Matriz de Massas M é uma matriz diagonal positiva definida². A sua matriz K é a soma de matriz simétricas semi-positivas definidas, portanto ela também é uma matriz simétrica semi-positiva definida³. Por tal razão, o problema dos autovalores generalizados está bem definido para estas matrizes [18]. Isso significa é que possível encontrar uma base formada por vetores v_i , tais que:

$$Kv_i = \lambda_i Mv_i \quad (2.22)$$

Para cada par (λ_i, v_i) , o autovalor $\lambda_i = \omega_i^2$ está relacionado à frequência. De fato, dizemos que $\omega_i = 2\pi f_i$ é a *Frequência Natural de Vibração*. Já o autovetor correspondente é o *Modo Natural de Vibração*. É possível visualizar alguns modos de vibração na Figura 2.6.

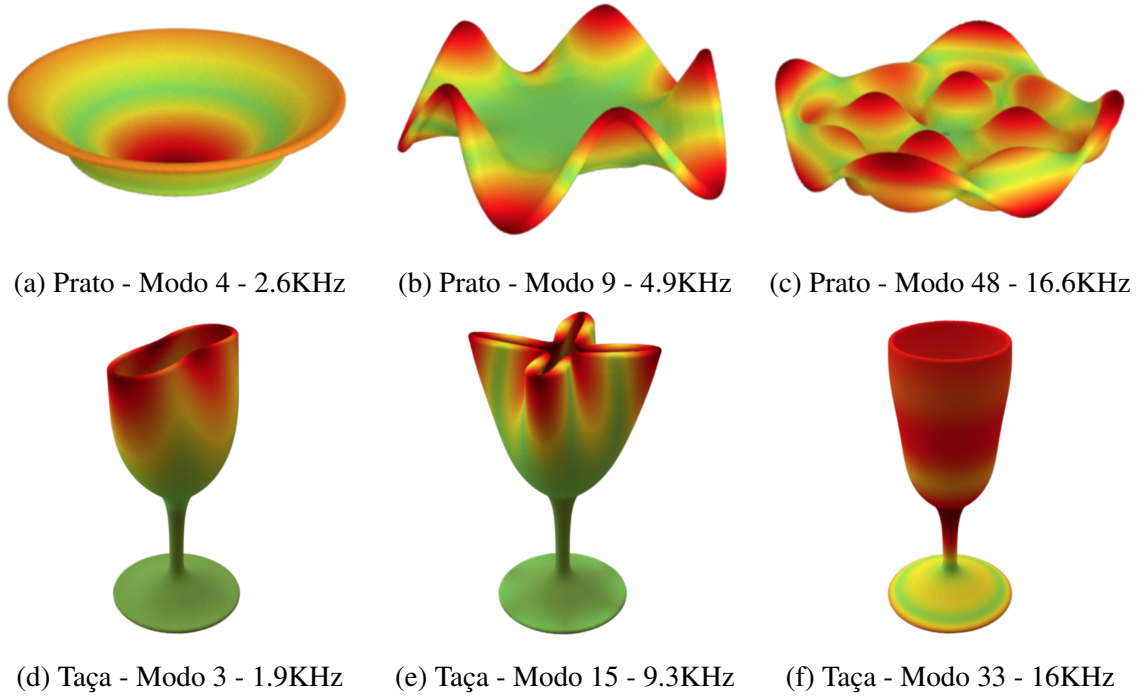


Figura 2.6: Modos de Vibração de objetos. Ao topo: Modos de vibração de um prato de cerâmica. Abaixo: Modos de vibração de uma taça de vidro. Fonte: [11]

Podemos escrever a equação (2.22) também em sua forma matricial:

$$KV = \Lambda MV \quad (2.23)$$

Λ é a matriz diagonal formada pelos autovalores $\lambda_i = \omega^2$ a V é a matriz formada pelos vetores-coluna v_i . Em particular, pode-se escolher V tal que ela seja ortogonal em relação à

²Uma matriz $M \in \mathbb{R}^{n \times n}$ é dita positiva definida se, e somente se, $\forall v \in \mathbb{R}^n \Rightarrow v^T M v > 0$. Uma definição equivalente é a de que todos os autovalores de M são positivos.

³Uma matriz $M \in \mathbb{R}^{n \times n}$ é dita semi-positiva definida se, e somente se, $\forall v \in \mathbb{R}^n \Rightarrow v^T M v \geq 0$. Uma definição equivalente é a de que todos os autovalores de M são não-negativos.

Matriz de Massas, isto é $V^T M V = I$, onde I é a matriz identidade. Nessa forma, dizemos que a matriz Λ é a *Matriz Espectral* do sistema e que V é a *Matriz Modal* do sistema.

Como V representa uma base do espaço, então podemos encontrar um vetor q tal que $u = Vq$. Se fizermos essa substituição na equação da elastodinâmica (2.20) utilizando também a aproximação de Rayleigh (2.21) obteremos:

$$MV\ddot{q} + (\alpha M + \beta K)V\dot{q} + KVq = f_{ext} \quad (2.24)$$

Podemos manipular a equação anterior:

$$f_{ext} = MV\ddot{q} + (\alpha M + \beta K)V\dot{q} + KVq \quad (2.25)$$

$$= MV\ddot{q} + (\alpha MV + \beta KV)\dot{q} + KVq \quad (2.26)$$

$$= MV\ddot{q} + (\alpha MV + \beta \Lambda MV)\dot{q} + \Lambda MV \quad (2.27)$$

$$V^T f_{ext} = V^T MV\ddot{q} + (\alpha V^T MV + \beta V^T \Lambda MV)\dot{q} + V^T \Lambda MV \quad (2.28)$$

$$= \ddot{q} + (\alpha I + \beta \Lambda)\dot{q} + \Lambda q \quad (2.29)$$

Em (2.26) expandimos o termo de \dot{q} . Em (2.27), aplicamos a igualdade (2.23). Em (2.28), multiplicamos ambos os lados da equação por V^T pela esquerda. Em (2.29), utilizamos o fato que V é ortogonal em relação à M e que a multiplicação por Λ é comutativa (pois Λ é diagonal).

Note que na equação final, todas as matrizes associadas aos termos q , \dot{q} e \ddot{q} são matrizes diagonais. Por tal razão, a equação diferencial multidimensional pode ser decomposta em várias equações diferenciais unidimensionais independentes:

$$v_i^T f_{ext} = \ddot{q}_i + (\alpha + \beta \omega_i^2)\dot{q}_i + \omega_i^2 q_i \quad (2.30)$$

Se considerarmos o caso no qual uma força instantânea é aplicada no momento $t = 0$, isto é $v_i^T f_{ext}(t) = \delta(t, 0)$ onde $\delta(x, y)$ é a função Delta de Dirac⁴ e que $q(0) = 0$, a solução da equação é da forma

$$q(t) = k_1 e^{-k_2 t} \sin(k_3 t) \quad (2.31)$$

onde k_1 , k_2 e k_3 são constantes (ver Figura 2.7). Vale a pena salientar que a constante de frequência k_3 tem um valor próximo à frequência natural ω_i . O seu valor é alterado de acordo com as constantes α e β .⁵

A possibilidade de decomposição em sistemas lineares independentes mostra que a vibração de um objeto pode ser expressa como a combinação linear de seus modos de vibração (ver Figura 2.8). Veremos adiante como essa propriedade pode ser explorada para acelerar a simulação.

⁴A função Delta de Dirac é tal que $\delta(x, y) = 1$ se $x = y$ e $\delta(x, y) = 0$ caso contrário.

⁵Em especial, se $\alpha = \beta = 0$, temos que $k_3 = \omega_i$.

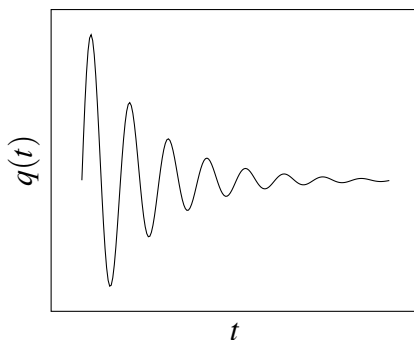


Figura 2.7: Gráfico da solução geral da equação $\ddot{q}(t) + (\alpha + \beta\omega_i^2)\dot{q}(t) + \omega^2 q(t) = f(t)$ considerando $f(t) = \delta(t, 0)$. A frequência de $q(t)$ é próxima à frequência natural ω_i .

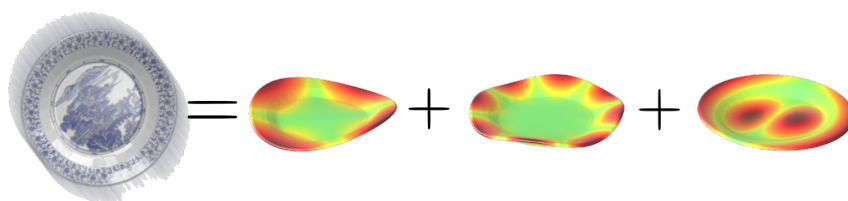


Figura 2.8: Composição de modos de vibração. A vibração de um objeto rígido pode ser expressa como combinação linear de diferentes modos de vibração regidos pela equação (2.30)

Trabalhos Relacionados

O trabalhos da área podem ser divididos em Métodos Physically-Based e Métodos Data-Driven.

3.1 Métodos Physically-Based

Os Métodos Physically-Based, ou métodos baseados em física, são aqueles que tentam, através das propriedades físicas dos materiais, sintetizar o som adequado. Em outras áreas, como Engenharia Mecânica ou Engenharia Aeronáutica, o estudo da radiação acústica gerada por materiais já é uma prática comum [19]. No entanto, os métodos utilizados nessas áreas costumam considerar apenas cenas estáticas, o que inviabiliza a interação com o usuário.

O trabalho de [25] foi o primeiro a considerar o uso de métodos physically-based em aplicações interativas. Os autores desenvolveram um processo simples para aplicações interativas: Em tempo de pré-processamento, os modos de vibração do objetos eram calculados e armazenados. Durante a execução, os objetos eram simulados e o estado $q_i(t)$ de cada modo de vibração era considerado utilizando a equação (2.30). O som final gerado pelo objeto era simplesmente a soma dos modos de vibração:

$$s(t) = \sum_i q_i(t) \quad (3.1)$$

Essa abordagem conseguia gerar som plausível em tempo real, mas muitas nuances do som eram ignoradas. A distância e o ângulo do objeto em relação ao ouvinte, por exemplo, têm uma grande influência no som final. O trabalho de [8] então introduziu o uso de Radiação Acústica no processo de síntese (Ver Figura 3.1). Nesse trabalho, a função de Radiação Acústica $T_i(u)$ era calculada para cada modo de vibração do objeto. Em tempo de execução, a posição do objeto em relação ao ouvinte era considerada. Com essa abordagem, o som final consideraria os modos de vibração ponderados pela radiação acústica:

$$s(t) = \sum_i |T_i(u_{ouvinte} - u_{objeto})| q_i(t) \quad (3.2)$$

O pipeline desenvolvido em [8] é utilizado como base para os trabalhos mais recentes na área.[28], por exemplo, estendeu o modelo para também sintetizar o som de fraturas.

[29] melhorou o modelo de contato utilizado na simulação do objeto. A qualidade do som final depende muito das forças de contato f_{ext} . Para capturar melhor as nuances do contato, eles passaram a utilizar um simulador numericamente estável [9] e também aumentaram a resolução da simulação para a resolução acústica (44KHz). Como resultado, a qualidade do áudio foi drasticamente melhorada mas o tempo necessário para realizar a síntese também aumentou.

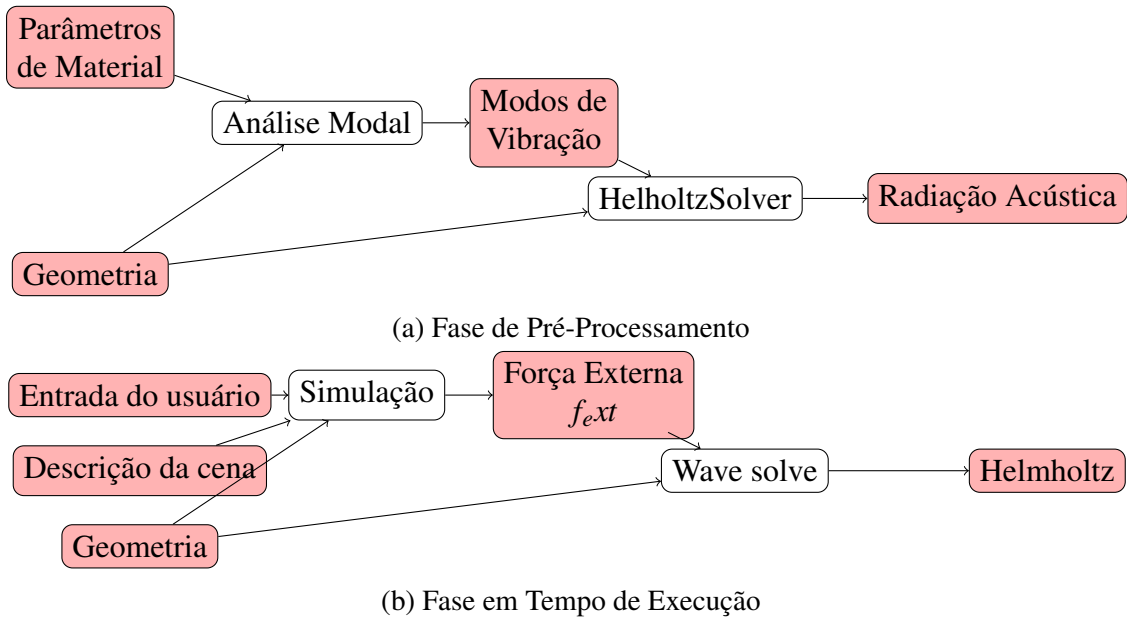


Figura 3.1: Pipeline de Síntese de Áudio Realístico.

Para realizar a síntese, os modos de vibração tem que ser utilizados em tempo de execução. Para objetos maiores, a quantidade de memória necessária para armazená-los pode inviabilizar a simulação. Em [11] um método foi desenvolvido para comprimir os modos de vibração. Isso possibilitou utilizar objetos maiores e mais detalhados sem impacto no tempo de síntese.

3.2 Métodos Data-Driven

Métodos Data-Driven ou métodos orientados à amostras são métodos que utilizam amostras de sons como entrada para o algoritmo de síntese. Um procedimento comum nessa classe de métodos é a extração de características das amostras originais e a extrapolação dessas características na etapa de síntese.

O trabalho de [14], por exemplo, faz uma análise nas amostras sonoras para identificar as frequências dominantes. As frequências dominantes assemelham-se às frequências naturais do objeto. Durante o tempo de execução, as frequências dominantes são somadas poderando cada uma com pesos gerados aleatoriamente. Os sons gerados não são fisicamente plausíveis e também não podem ser estendidos de maneira adequada para objetos distintos. Porém ela é uma abordagem computacionalmente barata que oferece um certo grau de dinamismo na acústica do ambiente virtual. Por tais razões, ela é ideal para ser incorporada à jogos em tempo real.

Os trabalhos de [22] e de [24] utilizam uma abordagem mista. Os dois utilizam as amostras sonoras para encontrar os parâmetros do material. Esses parâmetros são então utilizados como entrada em métodos baseados em física. Para aprender esses parâmetros, o som e o modelo virtual de um objeto real são utilizados como entrada do algoritmo. Os resultados gerados são plausíveis embora fisicamente incorretos.

Redes neurais também são comuns na área de síntese de áudio. Embora o seu uso seja mais comum na área de síntese de fala [21, 12], trabalhos recentes tem explorado técnicas para gerar sons de objetos. O trabalho de [17] utiliza Deep Learning para extrair características do áudio de amostras de vídeos. Com essas características, os autores conseguem sintetizar áudio plausível para outras cenas sem áudio.

Algoritmo e Implementação

4.1 Discretização e Solução da Equação da Onda

A equação da onda (2.1) é a equação governante da pressão acústica $\rho(u, t)$. Na nossa solução, no entanto, gostaríamos de especificar as condições de contorno em termo da velocidade das partículas. Para fazer isso, podemos separar a equação em um sistema de equações diferenciais:

$$\frac{\partial^2 \rho(u, t)}{\partial t^2} = dc^2 \nabla^2 \rho(u, t) \Leftrightarrow \begin{cases} \frac{\partial v(u, t)}{\partial t} = -\frac{\nabla \rho(u, t)}{d} \\ \frac{\partial \rho(u, t)}{\partial t} = c^2 \nabla \cdot v(u, t) \end{cases} \quad (4.1)$$

onde

$t \in \mathbb{R}$	\Rightarrow	Tempo em segundos (s)
$u \in \mathbb{R}^3$	\Rightarrow	Posição em metros (m)
$c \in \mathbb{R}$	\Rightarrow	Velocidade de Propagação (m/s)
$c \in \mathbb{R}$	\Rightarrow	Densidade kg/m^3
$\rho(u, t) \in \mathbb{R}$	\Rightarrow	Pressão acústica (Pa)
$v(u, t) \in \mathbb{R}^3$	\Rightarrow	Velocidade das Partículas (m/s)

Nesse sistema de equações, $v(u, t)$ é a velocidade das partículas do meio na posição u e no tempo t . O primeiro passo para resolver esse sistema de equações, é discretizar o seu domínio. Na nossa implementação, nós utilizamos o esquema de discretização Marker-And-Cell (MAC) [26]. Nesse esquema, o espaço é discretizado em elementos cada elemento contém a pressão acústica $\rho(u, t)$ no centro e contém as componentes da velocidade $v(u, t)$ nas bordas (Ver Figura 4.1).

Podemos facilmente mapear as derivadas de (4.1) para esse domínio:

$$\begin{aligned} \frac{\partial v_d(u, t)}{\partial t} &\Rightarrow \frac{1}{\Delta t} (v_d[x, y, z, t + \Delta t] - v_d[x, y, z, t]) \\ \frac{\partial \rho(u, t)}{\partial t} &\Rightarrow \frac{1}{\Delta t} (\rho[x, y, z, t + \Delta t] - \rho[x, y, z, t]) \\ \frac{\partial \rho(u, t)}{\partial x} &\Rightarrow \frac{1}{\Delta x} (\rho[x + \Delta x/2, y, z, t + \Delta t] - \rho[x - \Delta x/2, y, z, t]) \\ \frac{\partial v(u, t)}{\partial x} &\Rightarrow \frac{1}{\Delta x} (v_d[x + \Delta x/2, y, z, t] - v_d[x - \Delta x/2, y, z, t]) \end{aligned}$$

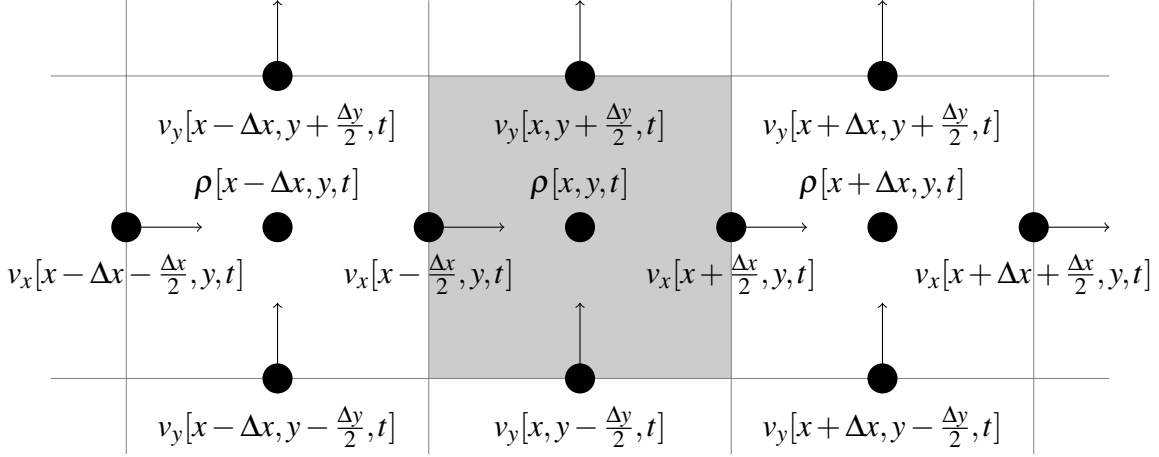


Figura 4.1: Discretização Marker-And-Cell

Com isso, podemos substituir essas fórmulas na equação (4.1) para obter uma regra de atualização dos campos de velocidade e pressão acústica:

$$\begin{cases} \forall n \in (x, y, z) \Rightarrow & v_n[x, y, z, t + \Delta t] = v_n[x, y, z, t] - \frac{\Delta t}{d} \frac{\partial \rho(u, t)}{\partial n} \\ & \rho[x, y, z, t + \Delta t] = \rho[x, y, z, t] + c^2 \left(\sum_{n \in (x, y, z)} \frac{\partial v(u, t)}{\partial n} \right) \end{cases} \quad (4.2)$$

Nós utilizamos as fórmulas em (4.2) alternadamente para avançar a nossa solução. Primeiro atualizamos o campo da velocidade e utilizamos esses valores para atualizar o campo da pressão acústica. Esse tipo de solução para sistemas de equações diferenciais também é conhecido como *Leapfrog*.

As condições de contorno são calculadas para cada modo de vibração. Considere objeto vibrando no modo de vibração v_i na sua frequência natural ω_i . Dado um ponto p na superfície do objeto, digamos que o vetor de vibração correspondente seja $\vec{n}_p \in \mathbb{R}^3$. Nesse caso, forçamos que a velocidade das partículas nesse ponto seja igual à velocidade da superfície do objeto adicionando a condição de contorno:

$$v(p, t) = \vec{n}_p \omega_i \sin(\omega_i t) \quad (4.3)$$

Para lidar com as bordas do domínio, utilizamos as condições de contorno conhecidas como Perfectly Matched Layers [13]. Desse modo, as ondas são absorvidas ao invés de serem refletidas para o domínio.

Na nossa implementação, a atualização dos campos de pressão acústica e de velocidade é executada em GPU. Nós utilizamos a abordagem descrita em [16] para explorar o paralelismo.

4.2 Aproximação da Equação de Helmholtz

O método descrito na seção anterior mostra como podemos calcular a pressão acústica $\rho(u, t)$. Porém, desejamos na verdade calcular a Radiação Acústica gerada pela vibração do objeto.

Nessa seção mostraremos como aproximamos a Radiação Acústica na nossa implementação. Para cada ponto u do domínio, desejamos calcular a Radiação Acústica $T(u) = A(u)e^{i\phi(u)}$. Consideremos, então, o sinal de um ponto fixo. Ao entrar em equilíbrio, o sinal $s(t)$ desse ponto poderá ser representado por $s(t) = A \cos(\omega t + \phi)$. Se calcularmos o valor médio¹ \bar{s} do valor absoluto sinal ao longo de um certo período, obteremos $|\bar{s}| = \frac{2A}{\pi}$. Logo, se soubermos $|\bar{s}|$ e $s(t)$ para um certo t , podemos calcular A e ϕ :

$$\begin{cases} A = \frac{\pi|\bar{s}|}{2} \\ \phi = \arccos\left(\frac{s(t)}{A}\right) - \omega t \end{cases} \quad (4.4)$$

Na nossa implementação, o valor médio do sinal de cada elemento do domínio é calculado utilizando uma *Weighted Moving Average* (WMA), isto é:

$$\bar{s}_t = \alpha \bar{s}_{t-1} + (1 - \alpha)s(t) \quad (4.5)$$

A amplitude A e a fase ϕ são calculados a cada timestep utilizando a fórmula (4.4). O valor médio de A e ϕ também são calculados utilizando uma WMA. Em todos os casos, utilizamos $\alpha = 0.95$.

4.2.1 Expansão Multipolar

A nossa solução da equação da onda é calculada dentro de um domínio limitado. Não seria adequado limitar a síntese acústica para esse domínio. Já vimos, no entanto, que é possível calcular a transferência acústica para Far-Field utilizando a Integral de Kirchoff (2.9).

Calcular a Integral de Kirchoff (uma integral de superfície) em tempo de execução é um procedimento custoso em termos de tempo e memória. Para evitar isso, exploramos a estrutura da Equação de Helmholtz.

A função de Green $G(u, v)$ da equação de Helmholtz pode ser decomposta numa base de funções [5]:

$$G(u, v) = ik \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(u - u_0) R_n^{-m}(v - u_0) \quad (4.6)$$

O ponto u_0 é o ponto no centro do Near-Field. A função S_n^m é a Função Esférica Singular da Equação de Helmholtz e R_n^m é a Função Esférica Singular da Equação de Helmholtz. Elas são definidas de acordo com as coordenadas polares $u = (r, \theta, \phi)$:

$$\begin{cases} S_n^m(u) \Rightarrow S_n^m(r, \theta, \phi) = h_n^{(2)}(kr) Y_n^m(\theta, \phi) \\ R_n^m(u) \Rightarrow R_n^m(r, \theta, \phi) = j_n(kr) Y_n^m(\theta, \phi) \end{cases} \quad (4.7)$$

¹O valor médio \bar{f} de uma função $f(t)$ é definido por

$$\bar{f} = \frac{\int_a^b f(t) dt}{b - a}$$

As funções $j_n : \mathbb{R} \rightarrow \mathbb{R}$ são as Funções Esféricas de Bessel. As funções $h_n^{(2)} : \mathbb{R} \rightarrow \mathbb{C}$ são as Funções Esféricas de Hankell de segunda espécie. As funções $Y_n^m : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{C}$ são os Harmônicos Esféricos (Ver Figura 4.2). Essas funções aparecem constantemente em diversos ramos da ciência² por elas formam uma base para o espaço de funções em $\mathbb{R}^3 \rightarrow \mathbb{R}$. A sua estrutura permite explorar a simetria radial e também permite aproximar funções com apenas alguns termos.

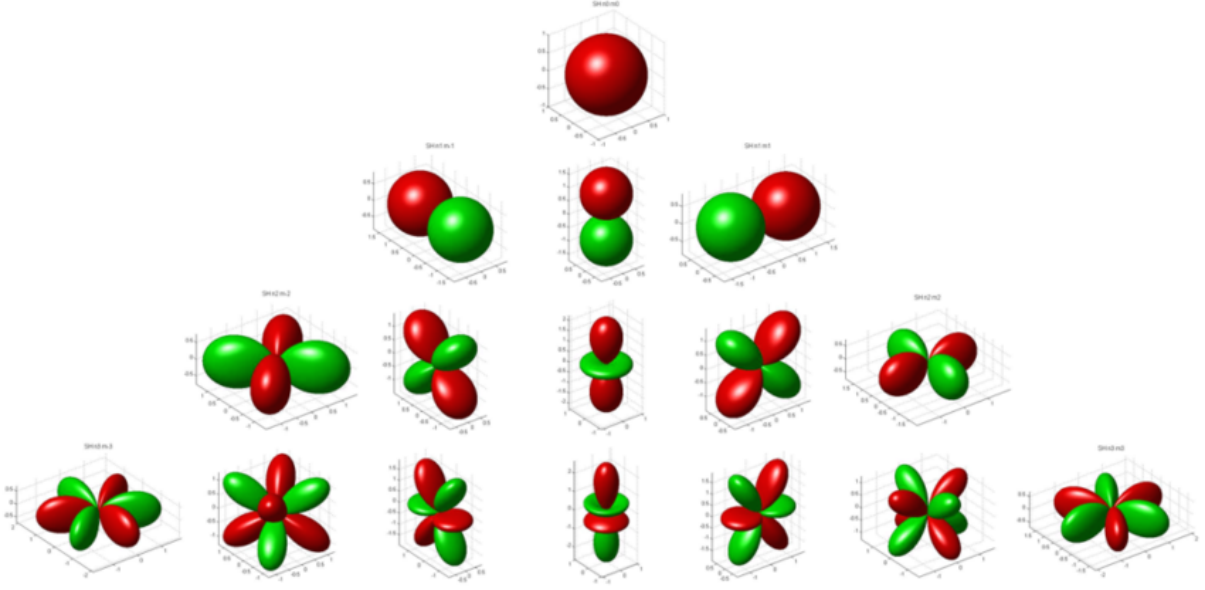


Figura 4.2: Harmônicos Esféricos

Se aplicarmos a expansão da Função de Green (4.6) na Integral de Kirchoff (2.9), obtemos:

$$\begin{aligned}
 T(u) &= \int_{\Gamma} \left[G(u, v) \frac{\partial T}{\partial n}(v) - \frac{\partial G}{\partial n}(u, v) T(v) \right] d\Gamma_v \\
 &= \int_{\Gamma} \left[ik \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(u - u_0) R_n^{-m}(v - u_0) \frac{\partial T}{\partial n}(v) \right. \\
 &\quad \left. - ik \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(u - u_0) \frac{\partial R_n^{-m}}{\partial n}(v - u_0) T(v) \right] d\Gamma_v \\
 &= \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(u - u_0) ik \int_{\Gamma} \left[R_n^{-m}(v - u_0) \frac{\partial T}{\partial n}(v) - \frac{\partial R_n^{-m}}{\partial n}(v - u_0) T(v) \right] d\Gamma_v \\
 &= \sum_{n=0}^{\infty} \sum_{m=-n}^n S_n^m(u - u_0) M_n^m
 \end{aligned} \tag{4.8}$$

²O uso de Funções Esféricas é comum em ramos como Eletromagnetismo, Mecânica Quântica, Acústica e Computação Gráfica.

Os termos M_n^m são chamados de *Coeficientes Multipolares*. Eles podem ser precomputados utilizando a fórmula:

$$M_n^m = ik \int_{\Gamma} \left[R_n^{-m}(v - u_0) \frac{\partial T}{\partial n}(v) - \frac{\partial R_n^{-m}}{\partial n}(v - u_0) T(v) \right] d\Gamma_v \quad (4.9)$$

Esses termos podem ser utilizados para calcular a Radiação Acústica $T(u)$ em tempo de execução utilizando a fórmula (4.8). Na nossa implementação, os coeficientes são calculados em paralelo na GPU. Nós utilizamos as recursões apresentada em [20] para calcular a Funções Esféricas. A Figura 4.3 apresenta uma comparação entre o Ground Truth da Radiação Acústica e o valor aproximado com a Expansão Multipolar.

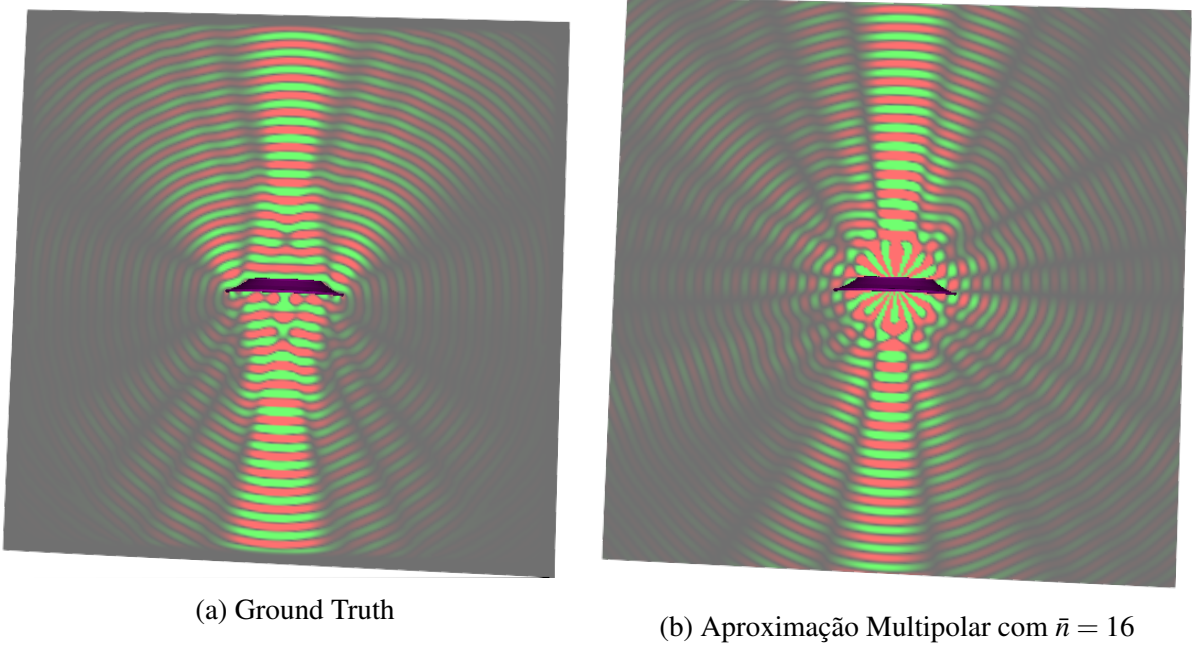


Figura 4.3: Comparação entre o Ground Truth da Radiação Acústica e a aproximação Expansão Multipolar. O objeto é um Prato de Cerâmica oscilando no Modo 23 a 10215.97Hz

CAPÍTULO 5

Experimentos e Resultados

Para comparar a qualidade e a performance da nossa solução com as abordagens anteriores, criamos três cenários de teste: (1) Prato de Cerâmica, (2) Caneca de Cerâmica, (3) Chave de Aço. Os parâmetros de materiais utilizados estão descritos na Tabela 5.1.

Material	Densidade (kg/m^3)	Módulo de Young (GPa)	Razão de Poisson	α	β
Cerâmica	2700	7.4×10^{10}	0.19	6	1×10^{-7}
Aço	1050	3.5×10^9	0.34	30	8×10^{-7}

Tabela 5.1: Parâmetros de Materiais

Para cada cenário de teste, computamos os Coeficientes Multipolares usando a nossa abordagem e também computamos os Coeficientes Multipolares utilizando a abordagem descrita em [28]. Essa última que é a mesma utilizada em [28, 29, 11], utiliza o software comercial FastBEM [1]. Vale notar que tivemos acesso ao código original dos autores.

Os testes foram todos executados na mesma máquina. A Tabela 5.2 apresenta a especificação do Hardware utilizado.

Componente	Especificação
Placa-Mãe	SuperMicro X8DAH
CPU	Intel Xeon X5570 - 8-Core @2.93GHz
Memória RAM	12x4GB DDR3 @1333MHz
GPU	GeForce TITAN BLACK 6GB DDR5

Tabela 5.2: Especificações de Hardware

A Tabela 5.3 apresenta a comparação de performance entre a nossa abordagem e as abordagens anteriores. A nossa abordagem obtém um speedup de mais de 30x em todos os cenários de teste.

Também comparamos os valores numérico da nossa solução com os valores calculados pela abordagem anterior (Ver Figura 5.1, Figura 5.2 e Figura 5.3). A diferença numérica entre os métodos ficou abaixo de 5%.

Cenário	# de modos	Nosso Método		FastBEM	Speedup
		Tempo (min)	Memória de GPU (MB)	Tempo (min)	
1	29	8.3 min	578.1MB	327.2 min	39.42x
2	39	18.6 min	998.7MB	686.5 min	36.90x
3	4	3.54 min	1586.3MB	102.9 min	29.06x

Tabela 5.3: Comparação de Performance entre a nossa solução e abordagens anteriores

5.1 Teste de Percepção

Para testar a qualidade do áudio final, um teste A/B foi feito com voluntários. Nesse teste, o voluntário assistia ambas as versões do vídeo (identificadas apenas por A ou B) e deveria responder à pergunta: “Qual dos dois pareceu mais realista? A ou B?”. O voluntário tinha como opções: “A”, “B” ou “Não sei”. A Tabela 5.4 contém o resultado da pesquisa.

Cena	Nosso Método	FastBEM	Não sei	Total
Prato de Cerâmica	70.5%	24%	5.5%	100%
Caneca de Cerâmica	64.6%	24%	11.4%	100%
Prato de Cerâmica	46.9%	20.5%	32.7%	100%
Total				254 respostas

Tabela 5.4: Resultado dos Testes de Percepção

Em todos os testes, a maioria dos usuário considerou que o som gerado pelo nosso método era mais realista que ou tão realista quanto o som gerado pelos métodos anteriores. Considerando que o nosso objetivo era apenas acelerar o método, esses resultados excedem as nossas expectativas.

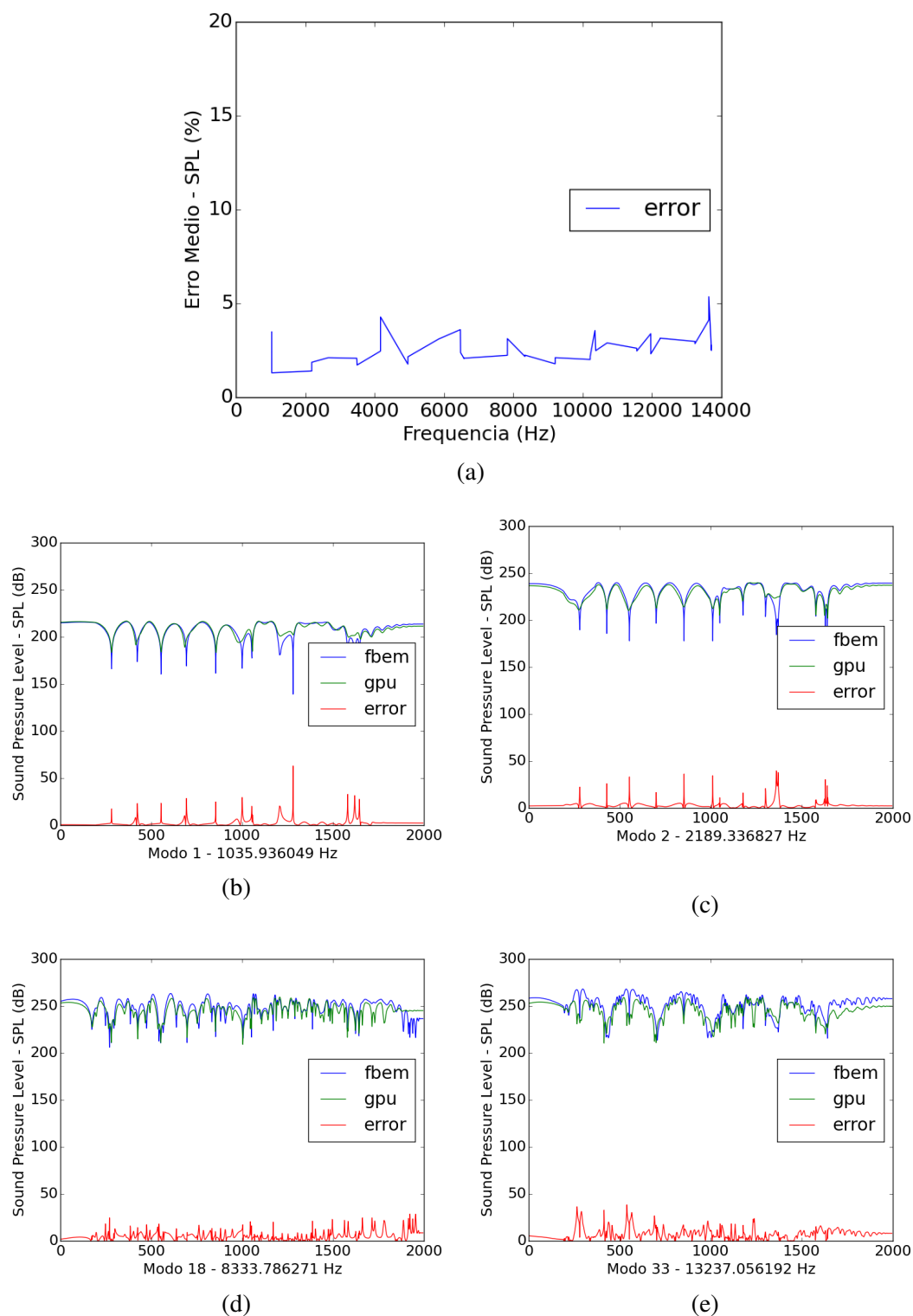


Figura 5.1: Comparação numérica para o Prato de Cerâmica. A figura ao topo apresenta o Erro Médio por frequência. As demais figuras apresentam a comparação para modos de vibração específicos.

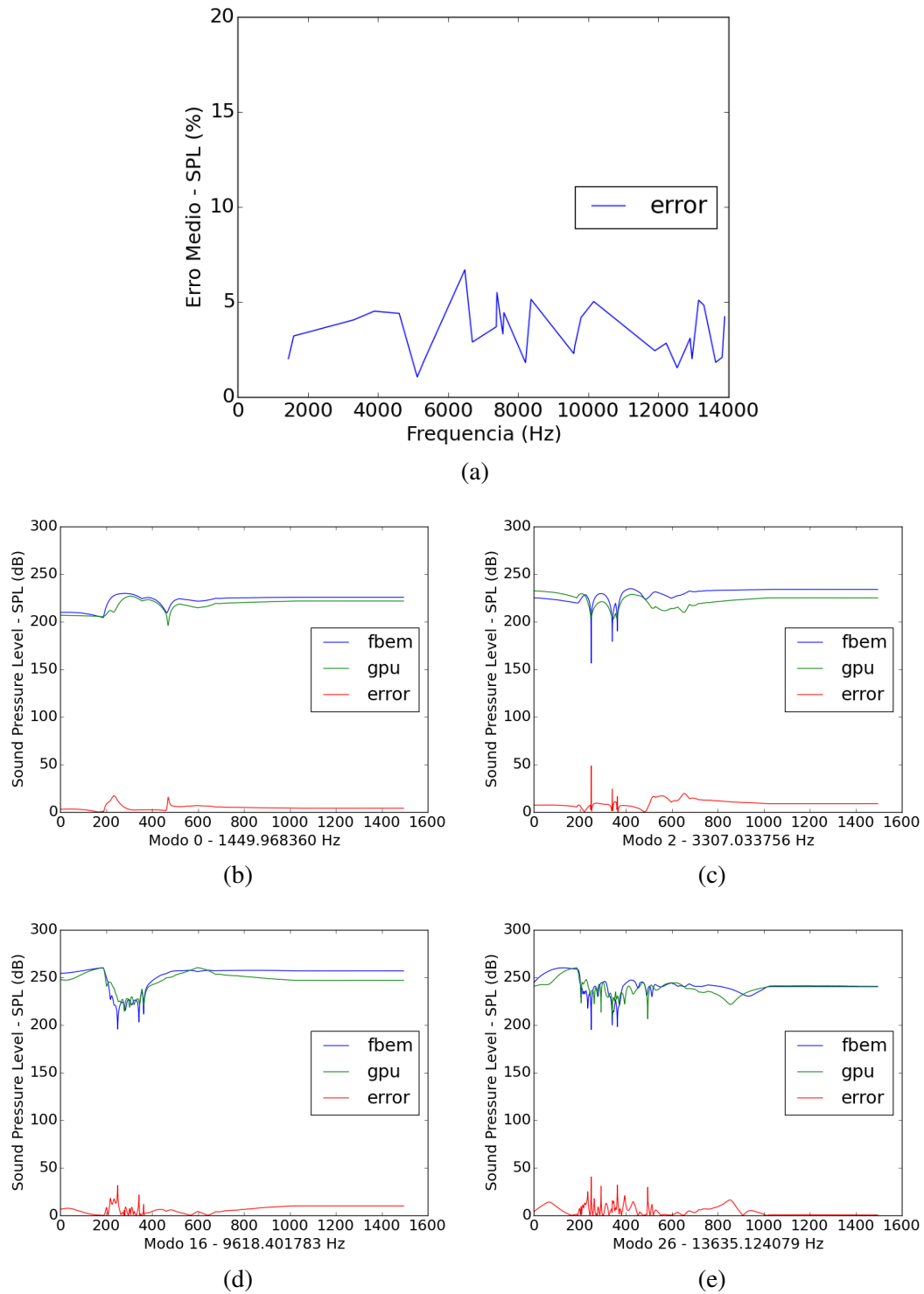


Figura 5.2: Comparação numérica para a Caneca de Cerâmica. A figura ao topo apresenta o Erro Médio por frequência. As demais figuras apresentam a comparação para modos de vibração específicos.

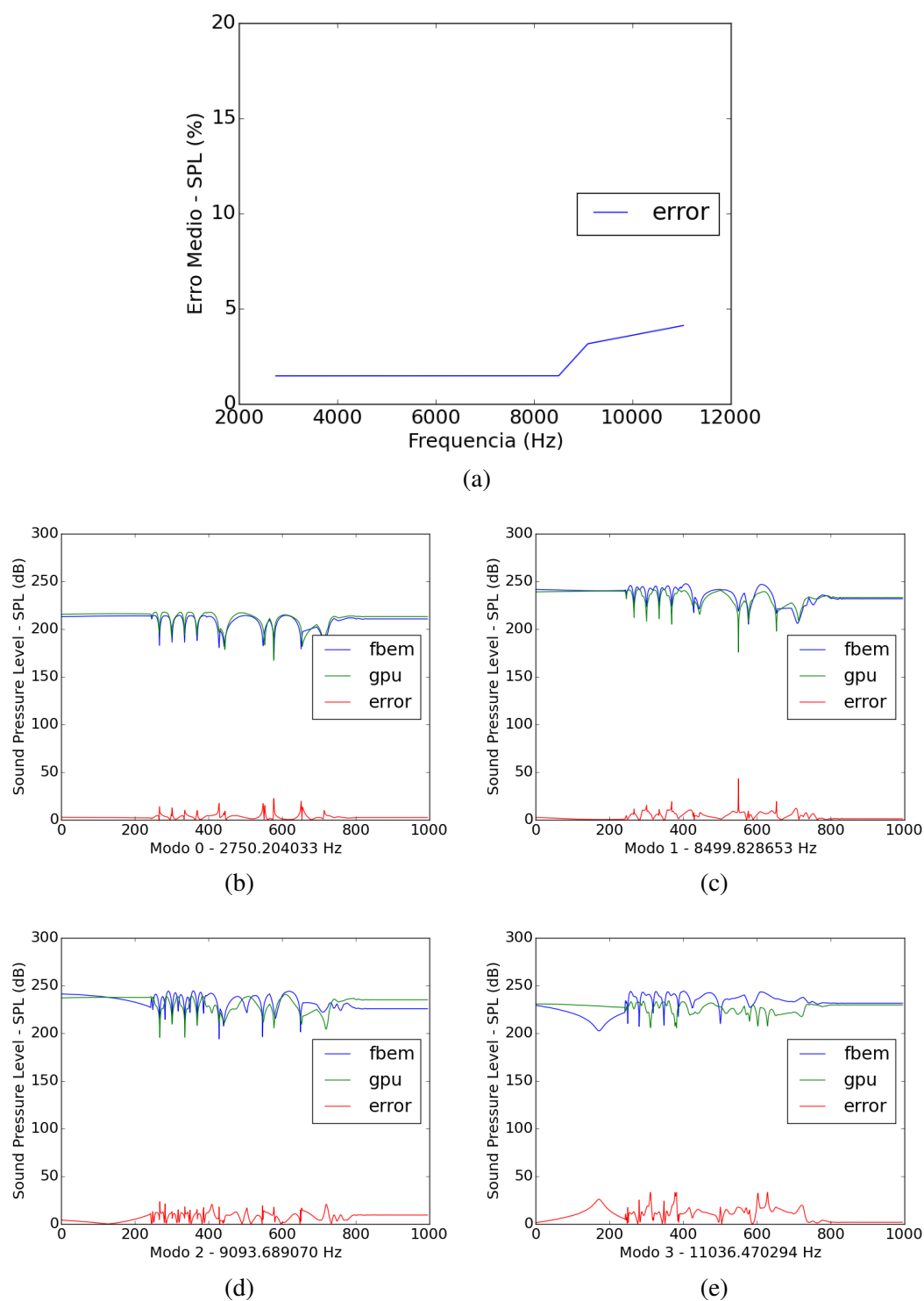


Figura 5.3: Comparação numérica para a Chave de Aço. A figura ao topo apresenta o Erro Médio por frequência. As demais figuras apresentam a comparação para modos de vibração específicos.

Conclusão

Nesse trabalho nós estudamos o estado-da-arte de Síntese de Áudio Realístico. Nós desenvolvemos um método para aproximar a solução da Equação de Helmholtz em GPU e comparamos com as soluções anteriores obtendo resultados bastante satisfatórios. Mais importante do que esses resultados é a fundamentação dos conceitos envolvidos. O conhecimentos de equações diferenciais e de métodos eficientes de solução pode ser utilizado nas mais diversas áreas. O mesmo pode ser dito para a experiência com o uso de GPUs.

6.1 Trabalhos futuros

Ainda existem muitas limitações com os métodos do estado-da-arte. Os métodos ainda tem um alto custo computacional em tempo de pré-processamento e também em tempo de execução. Isso inviabiliza a simulação de objetos maiores e também inviabiliza a simulação em tempo real. Trabalhos futuros podem tentar explorar técnicas de discretização adaptativas para lidar com domínios maiores e também pode tentar explorar técnicas de GPU para paralelizar as demais etapas do pipeline.

Referências Bibliográficas

- [1] FastBEM. <http://www.fastbem.com>, 2007-2016.
- [2] T. M. Apostol. *Calculus, vol. II*. Ed Reverté, 1969.
- [3] T. Bonebright. Were those coconuts or horse hoofs? Visual context effects on identification and perceived veracity of everyday sounds. In *The 18th International Conference on Auditory Display (ICAD2012) Atlanta,(GA)*, volume 2, 2012.
- [4] J. N. Chadwick, S. S. An, and D. L. James. Harmonic shells: a practical nonlinear sound model for near-rigid thin shells. In *ACM Transactions on Graphics (TOG)*, volume 28, page 119. ACM, 2009.
- [5] N. A. Gumerov and R. Duraiswami. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier, 2005.
- [6] B. Hamilton and C. J. Webb. Room acoustics modelling using GPU-accelerated finite difference and finite volume methods on a face-centered cubic grid. In *Digital Audio Effects (DAFx), Maynooth, Ireland*, 2013.
- [7] T. J. Hughes. *The finite element method: linear static and dynamic finite element analysis*. Courier Corporation, 2012.
- [8] D. L. James, J. Barbič, and D. K. Pai. Precomputed acoustic transfer: output-sensitive, accurate sound generation for geometrically complex vibration sources. In *ACM Transactions on Graphics (TOG)*, volume 25, pages 987–995. ACM, 2006.
- [9] D. M. Kaufman, S. Sueda, D. L. James, and D. K. Pai. Staggered projections for frictional contact in multibody systems. In *ACM Transactions on Graphics (TOG)*, volume 27, page 164. ACM, 2008.
- [10] H. Kuttruff. *Acoustics: an introduction*. CRC Press, 2007.
- [11] T. R. Langlois, S. S. An, K. K. Jin, and D. L. James. Eigenmode Compression for Modal Sound Models. *ACM Transactions on Graphics (TOG)*, 33(4):40, 2014.
- [12] Z.-H. Ling, S.-Y. Kang, H. Zen, A. Senior, M. Schuster, X.-J. Qian, H. M. Meng, and L. Deng. Deep learning for acoustic modeling in parametric speech generation: A systematic review of existing techniques and future trends. *IEEE Signal Processing Magazine*, 32(3):35–52, 2015.

- [13] Q.-H. Liu and J. Tao. The perfectly matched layer for acoustic waves in absorptive media. *The Journal of the Acoustical Society of America*, 102(4):2072–2082, 1997.
- [14] D. B. Lloyd, N. Raghuvanshi, and N. K. Govindaraju. Sound synthesis for impact sounds in video games. In *Symposium on Interactive 3D Graphics and Games*, pages PAGE–7. ACM, 2011.
- [15] R. Mehra, N. Raghuvanshi, L. Savioja, M. Lin, and D. Manocha. An efficient time-domain solver for the acoustic wave equation on graphics processors. *Applied Acoustics*, 73(2), 2012.
- [16] P. Micikevicius. 3D finite difference computation on GPUs using CUDA. In *Proceedings of 2nd workshop on general purpose processing on graphics processing units*, pages 79–84. ACM, 2009.
- [17] A. Owens, P. Isola, J. McDermott, A. Torralba, E. H. Adelson, and W. T. Freeman. Visually indicated sounds. <https://arxiv.org/abs/1512.08512v2>, 2015.
- [18] B. N. Parlett. *The symmetric eigenvalue problem*, volume 7. SIAM, 1980.
- [19] A. D. Pierce et al. *Acoustics: an introduction to its physical principles and applications*, volume 20. McGraw-Hill New York, 1981.
- [20] W. H. Press. *Numerical recipes 3rd edition: The art of scientific computing*. Cambridge university press, 2007.
- [21] M. G. Rahim, C. C. Goodyear, W. B. Kleijn, J. Schroeter, and M. M. Sondhi. On the use of neural networks in articulatory speech synthesis. *The Journal of the Acoustical Society of America*, 93(2):1109–1121, 1993.
- [22] Z. Ren, H. Yeh, and M. C. Lin. Example-guided physically based modal sound synthesis. *ACM Transactions on Graphics (TOG)*, 32(1):1, 2013.
- [23] A. A. Shabana. *Theory of vibration: Volume II: discrete and continuous systems*. Springer Science & Business Media, 2012.
- [24] A. Sterling and M. C. Lin. Interactive modal sound synthesis using generalized proportional damping. In *Proceedings of the 20th ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, pages 79–86. ACM, 2016.
- [25] K. Van Den Doel, P. G. Kry, and D. K. Pai. FoleyAutomatic: physically-based sound effects for interactive simulation and animation. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 537–544. ACM, 2001.
- [26] J. E. Welch, F. H. Harlow, J. P. Shannon, and B. J. Daly. *The MAC method*. Los Alamos Scientific Laboratory of the University of California, 1966.

- [27] C. Zheng and D. L. James. Harmonic fluids. *ACM Transactions on Graphics (TOG)*, 28(3):37, 2009.
- [28] C. Zheng and D. L. James. Rigid-body fracture sound with precomputed soundbanks. In *ACM Transactions on Graphics (TOG)*, volume 29, page 69. ACM, 2010.
- [29] C. Zheng and D. L. James. Toward high-quality modal contact sound. *ACM Transactions on Graphics (TOG)*, 30(4):38, 2011.