



Universidade Federal de Pernambuco
Centro de Informática

Graduação em Ciência da Computação

Síntese de Áudio Realístico: Acelerando a Equação de Helmholtz 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 Equação de Helmholtz 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

<DIGITE O RESUMO AQUI>

Palavras-chave: <DIGITE AS PALAVRAS-CHAVE AQUI>

Abstract

Keywords: <DIGITE AS PALAVRAS-CHAVE AQUI>

Sumário

1	Introdução	1
1.1	Foley	1
2	Fundamentação Teórica	2
2.1	Som e Acústica	2
2.1.1	Equação da Onda	2
2.1.1.1	Fontes Sonoras e Ambientes Acústicos	2
2.1.2	Radiação acústica e a Equação de Helmholtz	4
2.1.2.1	Aproximação em Far-Field	5
2.2	Vibração de Corpos Rígidos	6
2.2.1	Modelo Elastodinâmico	6
2.2.2	Análise Modal	9
3	Trabalhos Relacionados	12
3.1	Métodos Baseados em Física	12
3.2	Métodos Orientados à Amostras	13
4	Algoritmo e Implementação	15
4.1	Análise Modal	15
4.2	Aproximando da Equação de Helmholtz	15
4.3	Síntese	15
5	Experimentos e Resultados	16
5.1	Benchmark	16
5.2	Teste Numérico	16
5.3	Teste de Percepção	16
6	Conclusão	21

Lista de Figuras

2.1	Pressão acústica gerada uma fonte pontual	3
2.2	Radiação acústica gerada uma fonte pontual	5
2.3	Radiação de Near-Field e Far-Field	5
2.4	Sistema elastodinâmico unidimensional	6
2.5	Sistema elastodinâmico bidimensional	7
2.6	Modos de Vibração de objetos	9
2.7	Gráfico da solução da equação $\ddot{q}(t) + (\alpha + \beta \omega_i^2)\dot{q}(t) + \omega^2 q(t) = f(t)$	11
2.8	Composição de modos de vibração	11
3.1	Overview do pipeline	13
5.1	Modos de Vibração de objetos	18
5.2	Modos de Vibração de objetos	19
5.3	Modos de Vibração de objetos	20

Lista de Tabelas

5.1	Parâmetros de Materiais	16
5.2	Resultado dos Testes de Percepção	17

CAPÍTULO 1

Introdução

Introduzindo as paradas

1.1 Foley

CAPÍTULO 2

Fundamentação Teórica

2.1 Som e Acústica

Som é uma coisinha bonitinha. Som é uma coisinha bonitinha. Som é uma coisinha bonitinha. Som é uma coisinha bonitinha. Som é uma coisinha bonitinha. Som é uma coisinha bonitinha.

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 [9].

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:

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

onde

$$t \in \mathbb{R} \Rightarrow \text{Tempo em segundos(s)}$$

$$u \in \mathbb{R}^3 \Rightarrow \text{Posição em metros(m)}$$

$$c \in \mathbb{R} \Rightarrow \text{Velocidade de Propagação (m/s)}$$

$$\rho(u, t) \in \mathbb{R} \Rightarrow \text{Pressão acústica (Pa)}$$

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

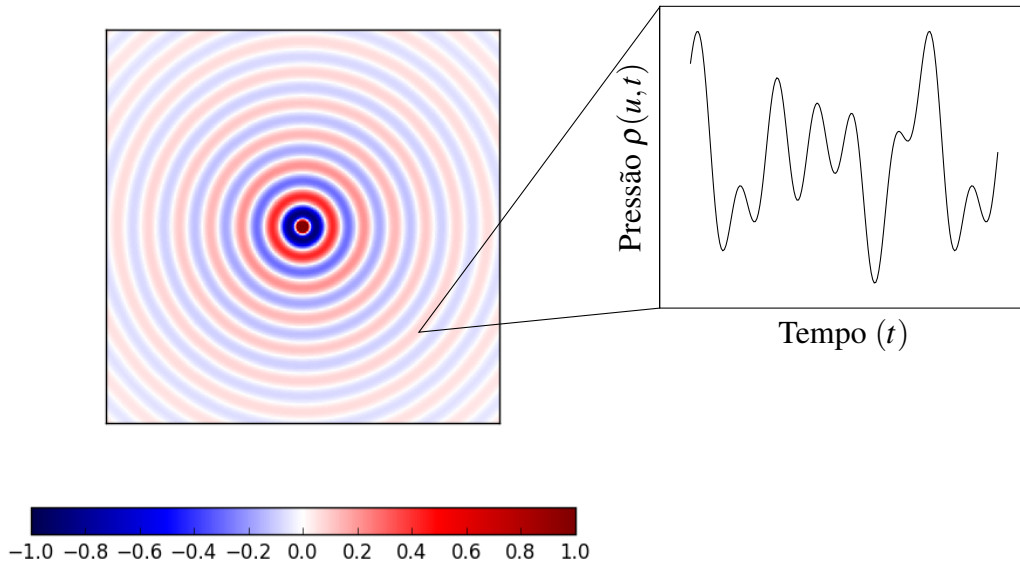


Figura 2.1: Pressão acústica gerada por uma fonte pontual no centro do meio. A fonte e o ambiente determinam as condições de contorno e a pressão acústica nos demais pontos é governada pela Equação da Onda (2.1)

(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:

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

onde $R(u)$ 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 $\rho^*(u, t) = R(u)F^*(t)$.

A função $R(u)$ 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*:

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

onde

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

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 $R(u)$ 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 $R(u) = A(u)e^{i\phi(u)}$ e $F(t) = A_0 e^{i\omega t}$, obteremos:

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

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

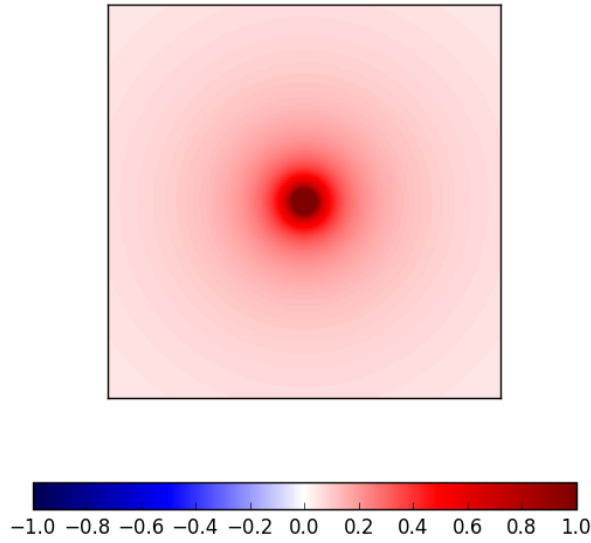


Figura 2.2: Radiação acústica gerada por uma fonte pontual no centro do meio. A fonte e o ambiente determinam as condições do contorno e a radiação acústica nos demais pontos é governada pela Equação de Helmholtz

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).

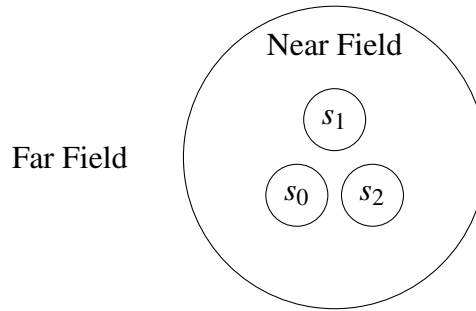


Figura 2.3: Radiação de Near-Field e Far-Field.

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*:

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

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 $R(u)$ pode ser calculada pela *Integral de Kirchhoff*:

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

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

O estudo das vibrações de estruturas rígidas é uma prática comum entre vários ramos da ciência. (explicar porque)

2.2.1 Modelo Elastodinâmico

O modelo elastodinâmico [20] 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.

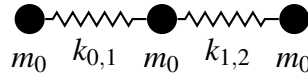


Figura 2.4: Sistema elastodinâmico unidimensional. Cada nó tem massa m_i e cada mola tem uma constante $k_{i,j}$

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.

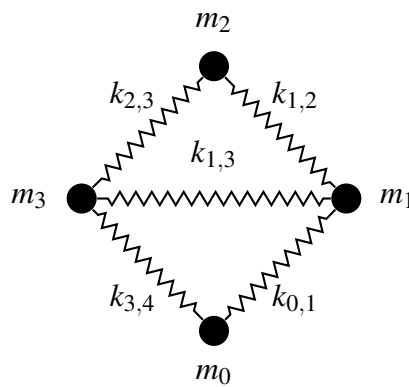


Figura 2.5: Sistema elastodinâmico bidimensional. Cada nó tem massa m_i e cada mola tem uma constante $k_{i,j}$

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 [6]. 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 [20]. 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 [16]. 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.

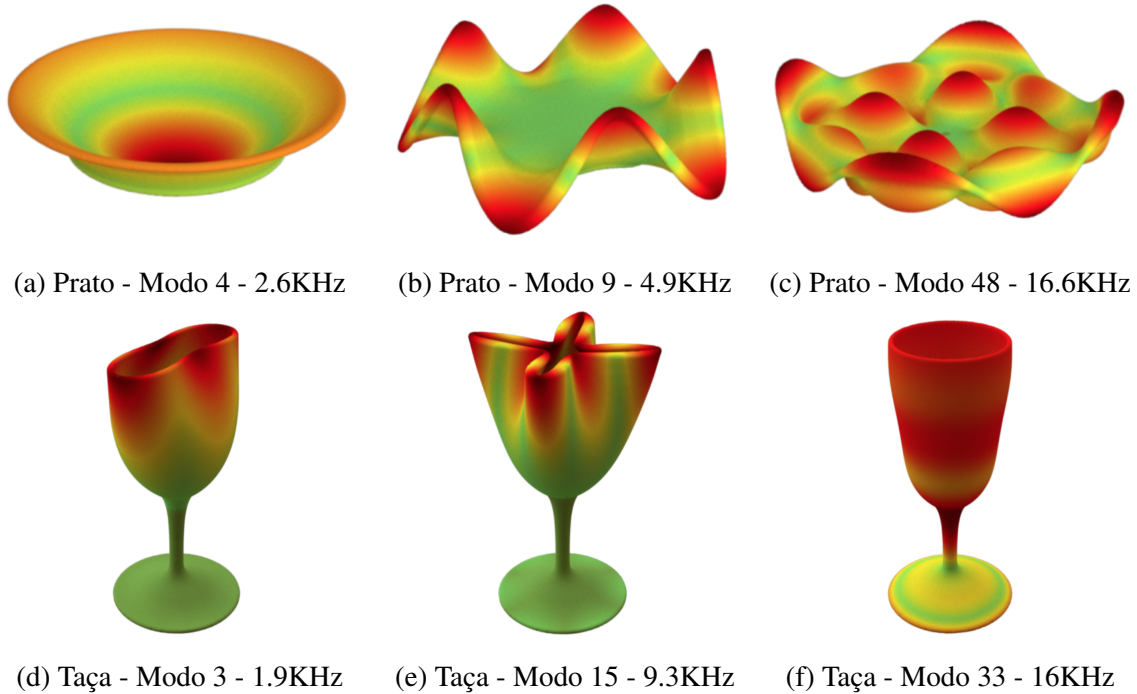


Figura 2.6: Modos de Vibração de objetos. Ao topo: Modos de vibração de um prato de cerâmica. Abaixo: Modos de vibração de uma taça de vidro. Fonte: [10]

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 $\lambda_i = \omega^2$ e 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 $k_3 = \omega_i$.

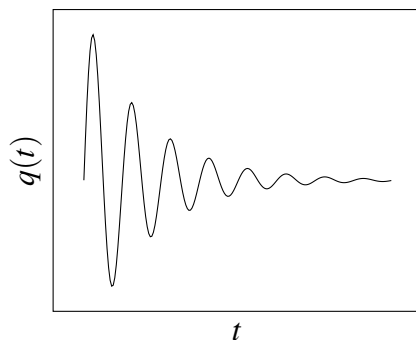


Figura 2.7: Gráfico da solução geral da equação $\ddot{q}(t) + (\alpha + \beta\omega_i^2)\dot{q}(t) + \omega^2 q(t) = f(t)$ considerando $f(t) = \delta(t, 0)$. A frequência de $q(t)$ é próxima à frequência natural ω_i .

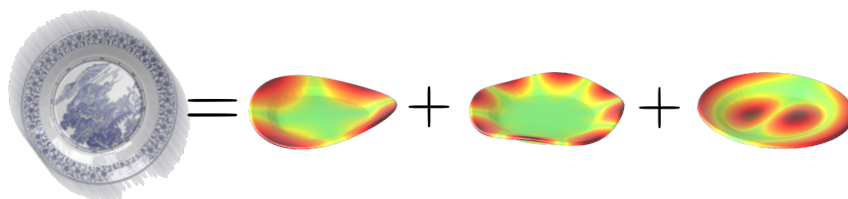


Figura 2.8: Composição de modos de vibração. A vibração de um objeto rígido pode ser expressa como combinação linear de diferentes modos de vibração regidos pela equação (2.30)

Trabalhos Relacionados

A gente basicamente se divide em métodos analíticos e métodos data-driven

3.1 Métodos Baseados em Física

Os 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 [17]. No entanto, os métodos utilizados nessas áreas costumam considerar apenas cenas estáticas, o que inviabiliza a interação com o usuário.

O trabalho de [22] 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 objeto 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 [7] então introduziu o uso de Radiação Acústica no processo de síntese (Ver Figura 3.1). Nesse trabalho, a função de Radiação Acústica $R_i(u)$ 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:

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

O pipeline desenvolvido por [7] é utilizado como base para os trabalhos mais recentes na área.[24], por exemplo, estenderam o modelo para também sintetizar o som de fraturas.

[25] melhoraram 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 [8] 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.

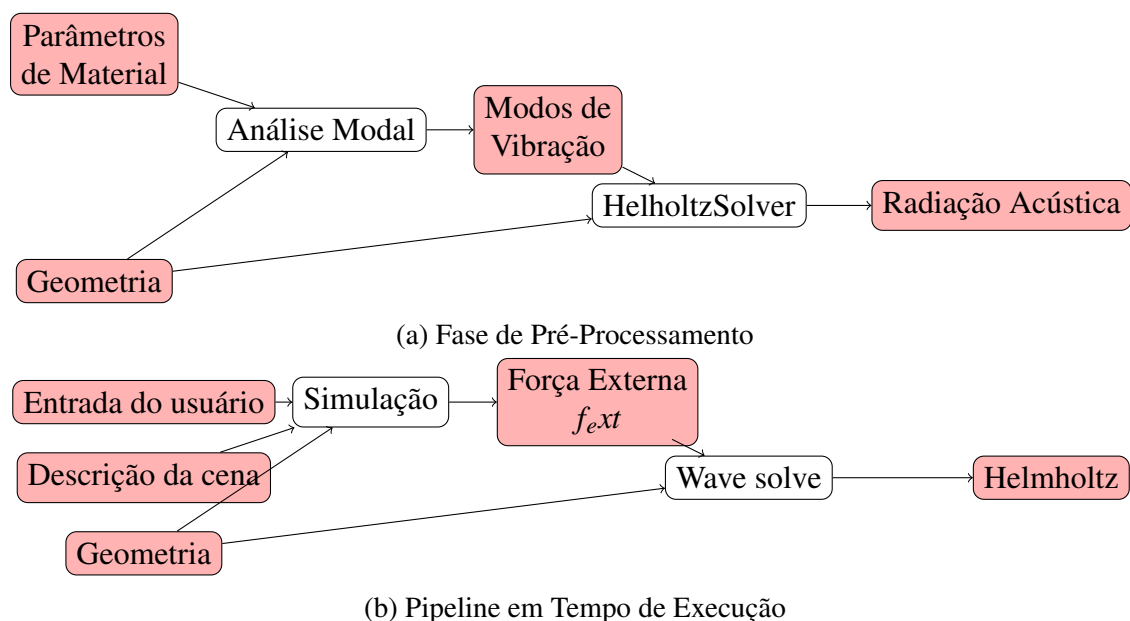


Figura 3.1: Figura 3.1a plots of....

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. [10] desenvolveram um método para comprimir os modos de vibração. Isso possibilitou utilizar objetos maiores e mais detalhados sem impacto no tempo de síntese.

3.2 Métodos Orientados à Amostras

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 [12], 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 [19] e de [21] 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 [18, 11], trabalhos recentes tem explorado técnicas para gerar sons de objetos. O trabalho de [15] 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 Análise Modal

4.2 Aproximando da Equação de Helmholtz

Para cada ponto u do domínio, desejamos calcular $R(u) = A(u)e^{i\phi(u)}$. 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 :

$$\begin{cases} A = \frac{\pi|\bar{s}|}{2} \\ \phi = \arccos\left(\frac{s(t)}{A}\right) - \omega t \end{cases} \quad (4.1)$$

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

$$\bar{s}_t = \alpha \bar{s}_{t-1} + (1 - \alpha)s(t) \quad (4.2)$$

A amplitude A e a fase ϕ são calculados a cada timestep utilizando a fórmula (4.1). O valor médio de A e ϕ também são calculados utilizando uma WMA. Em todos os casos, utilizamos $\alpha = 0.95$.

4.3 Síntese

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

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

CAPÍTULO 5

Experimentos e Resultados

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

5.1 Benchmark

5.2 Teste Numérico

Para testarmos numericamente o nosso método

Para testarmos numericamente o nosso método

Para testarmos numericamente o nosso método

Para testarmos numericamente o nosso método

5.3 Teste de Percepção

Para testarmos a qualidade final do áudio sintetizado, geramos três cenários distintos. Para cada cenário, geramos um áudio com os coeficientes calculados pelo FastBEM e outro áudio com os coeficientes calculados com o nosso método. O restante do pipeline não foi modificado.

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 5.2 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 5.2: Resultado dos Testes de Percepção

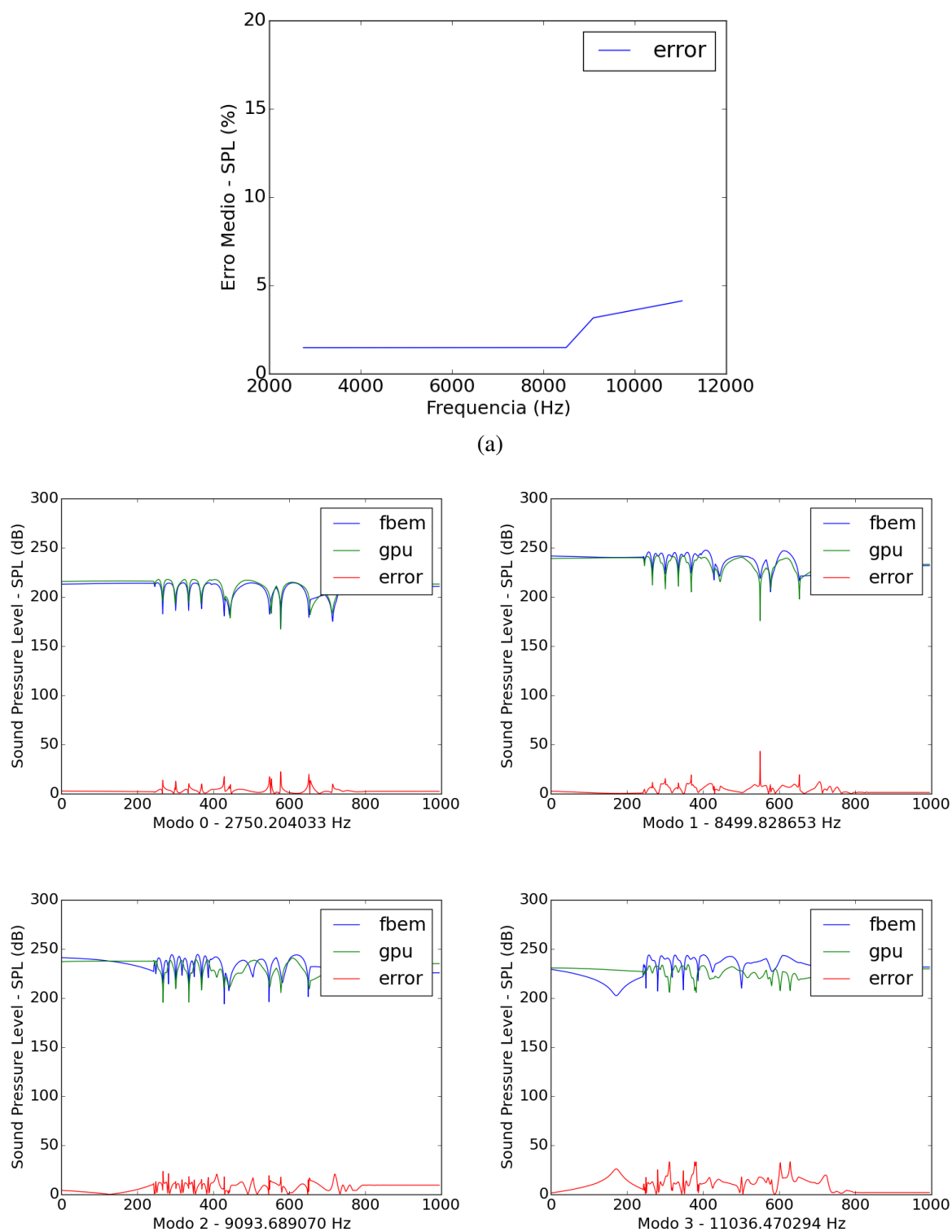


Figura 5.1: Modos de Vibração de objetos. Ao topo: Modos de vibração de um prato de cerâmica. Abaixo: Modos de vibração de uma taça de vidro. Fonte: [10]

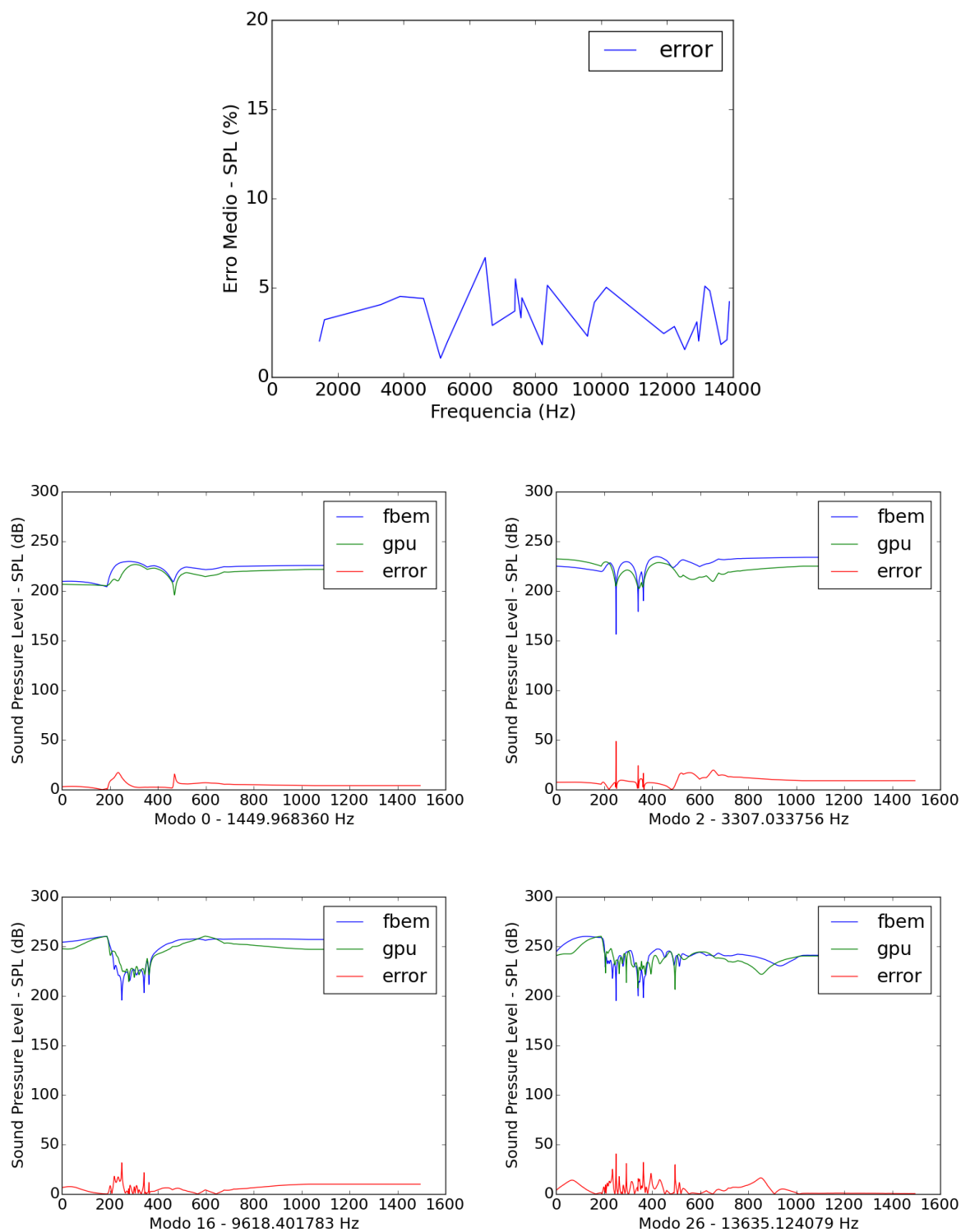


Figura 5.2: Modos de Vibração de objetos. Ao topo: Modos de vibração de um prato de cerâmica. Abaixo: Modos de vibração de uma taça de vidro. Fonte: [10]

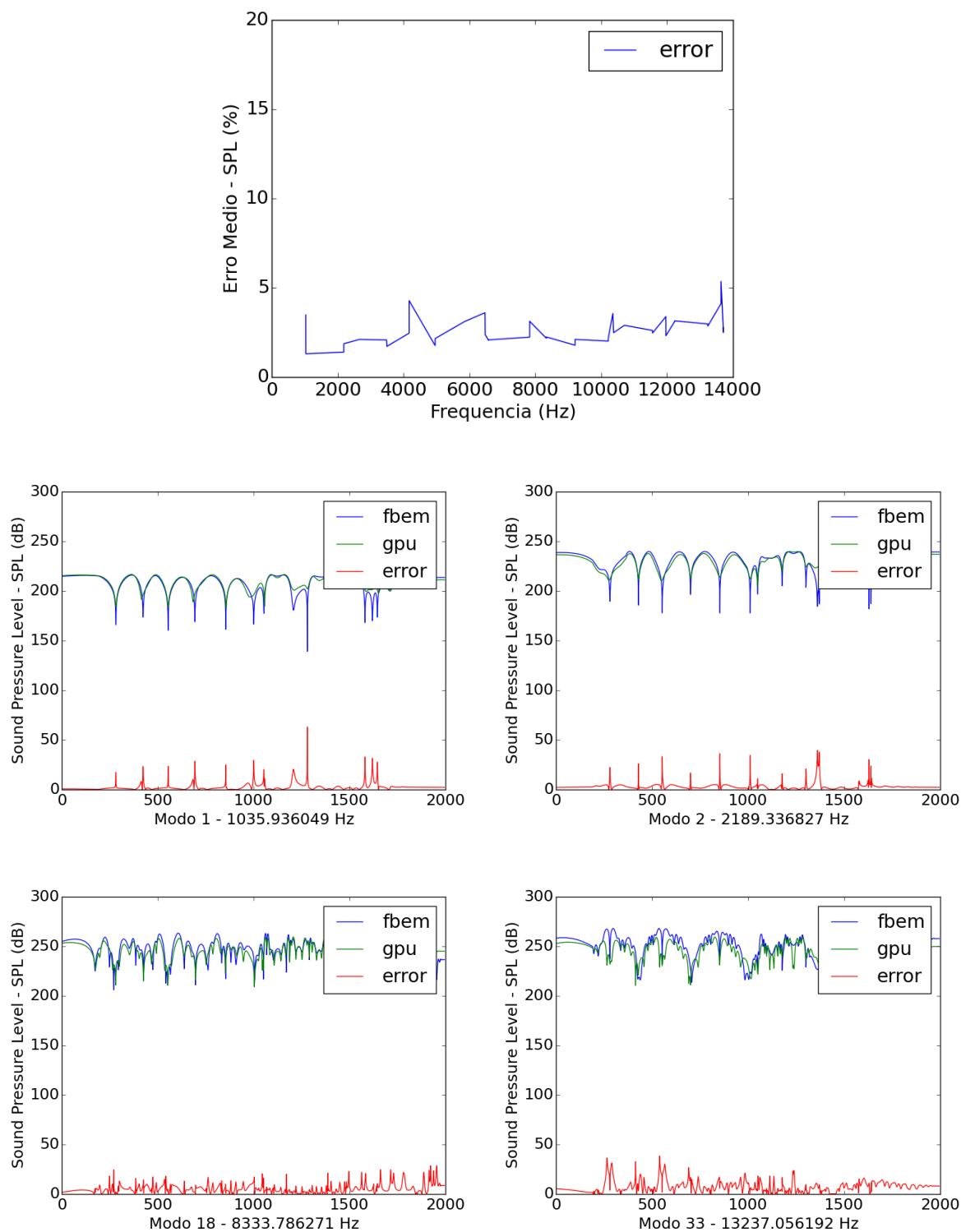


Figura 5.3: Modos de Vibração de objetos. Ao topo: Modos de vibração de um prato de cerâmica. Abaixo: Modos de vibração de uma taça de vidro. Fonte: [10]

CAPÍTULO 6

Conclusão

Concluindo as paradas [15]

Referências Bibliográficas

- [1] Tom M Apostol. *Calculus, vol. II*. Ed Reverté, 1969.
- [2] Jeffrey N Chadwick, Steven S An, and Doug L James. Harmonic shells: a practical non-linear sound model for near-rigid thin shells. In *ACM Transactions on Graphics (TOG)*, volume 28, page 119. ACM, 2009.
- [3] Y. Cui, E. Poyraz, K. B. Olsen, J. Zhou, K. Withers, S. Callaghan, J. Larkin, C. Guest, D. Choi, A. Chourasia, Z. Shi, S. M. Day, P. J. Maechling, and T. H. Jordan. Physics-based Seismic Hazard Analysis on Petascale Heterogeneous Supercomputers. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '13*, pages 70:1–70:12, New York, NY, USA, 2013. ACM.
- [4] Nail A Gumerov and Ramani Duraiswami. *Fast multipole methods for the Helmholtz equation in three dimensions*. Elsevier, 2005.
- [5] Brian Hamilton and Craig 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.
- [6] Thomas JR Hughes. *The finite element method: linear static and dynamic finite element analysis*. Courier Corporation, 2012.
- [7] Doug L James, Jernej Barbič, and Dinesh 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.
- [8] Danny M Kaufman, Shinjiro Sueda, Doug L James, and Dinesh K Pai. Staggered projections for frictional contact in multibody systems. In *ACM Transactions on Graphics (TOG)*, volume 27, page 164. ACM, 2008.
- [9] Heinrich Kuttruff. *Acoustics: an introduction*. CRC Press, 2007.
- [10] Timothy R Langlois, Steven S An, Kelvin K Jin, and Doug L James. Eigenmode Compression for Modal Sound Models. *ACM Transactions on Graphics (TOG)*, 33(4):40, 2014.
- [11] Zhen-Hua Ling, Shi-Yin Kang, Heiga Zen, Andrew Senior, Mike Schuster, Xiao-Jun Qian, Helen M Meng, and Li 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.
- [12] D Brandon Lloyd, Nikunj Raghuvanshi, and Naga K Govindaraju. Sound synthesis for impact sounds in video games. In *Symposium on Interactive 3D Graphics and Games*, pages PAGE–7. ACM, 2011.
- [13] Ravish Mehra, Nikunj Raghuvanshi, Lauri Savioja, Ming Lin, and Dinesh Manocha. An efficient time-domain solver for the acoustic wave equation on graphics processors. *Applied Acoustics*, 73(2), 2012.
- [14] Paulius 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.
- [15] Andrew Owens, Phillip Isola, Josh McDermott, Antonio Torralba, Edward H. Adelson, and William T. Freeman. Visually Indicated Sounds, 2015.
- [16] Beresford N Parlett. *The symmetric eigenvalue problem*, volume 7. SIAM, 1980.
- [17] Allan D Pierce et al. *Acoustics: an introduction to its physical principles and applications*, volume 20. McGraw-Hill New York, 1981.
- [18] Mazin G Rahim, Colin C Goodyear, W Bastiaan Kleijn, Juergen Schroeter, and M Mohan Sondhi. On the use of neural networks in articulatory speech synthesis. *The Journal of the Acoustical Society of America*, 93(2):1109–1121, 1993.
- [19] Zhimin Ren, Hengchin Yeh, and Ming C Lin. Example-guided physically based modal sound synthesis. *ACM Transactions on Graphics (TOG)*, 32(1):1, 2013.
- [20] Ahmed A Shabana. *Theory of vibration: Volume II: discrete and continuous systems*. Springer Science & Business Media, 2012.
- [21] Auston Sterling and Ming 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.
- [22] Kees Van Den Doel, Paul G Kry, and Dinesh 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.
- [23] Changxi Zheng and Doug L James. Harmonic fluids. *ACM Transactions on Graphics (TOG)*, 28(3):37, 2009.
- [24] Changxi Zheng and Doug L James. Rigid-body fracture sound with precomputed soundbanks. In *ACM Transactions on Graphics (TOG)*, volume 29, page 69. ACM, 2010.

- [25] Changxi Zheng and Doug L James. Toward high-quality modal contact sound. *ACM Transactions on Graphics (TOG)*, 30(4):38, 2011.