

GTSpat: um sistema interativo de síntese e espacialização sonora

Gabriel Rimoldi^{1,2}, Adolfo Maia Jr.²

¹Instituto de Artes - Departamento de Música | ²Núcleo Interdisciplinar de Comunicação Sonora (NICS) – Universidade Estadual de Campinas (UNICAMP)
Campinas – SP - Brasil

{gabriel.rimoldi,adolfo}@nics.unicamp.br

Abstract. *This article reports the development of a real-time system of sound synthesis and spatialization of granular flows, named GTSpat. Firstly, we present the general architecture of the algorithm, and then, point out details of the interactive strategies applied to the system. Finally, we present some adaptive schemes by which the system can be modified by interaction with the user.*

Resumo. *Este artigo aborda o desenvolvimento de um sistema de síntese e espacialização sonora de fluxos granulares em tempo real, ao qual denominamos GTSpat. Apresentamos inicialmente a arquitetura geral do algoritmo e, em seguida, detalhamos as estratégias interativas aplicadas ao sistema. Por fim, apresentamos alguns esquemas adaptativos pelos quais o sistema pode ser modificado pela interação com o usuário.*

1. Introdução

O grande desenvolvimento tecnologias voltadas para a difusão multicanal e sua crescente acessibilidade tem, nas últimas décadas, renovado o interesse de compositores e pesquisadores sobre métodos formais para o controle da espacialização sonora [Otondo, 2008]. Neste trabalho, reportamos um método desenvolvido para o controle da distribuição espacial dos sons através das chamadas *Medidas de Similaridade* relacionadas ao timbre, no qual associamos um conjunto de Descritores de Áudio [Puckette e Apel 1998; Foote, 1997] obtidos a partir de uma determinada amostra sonora à sua respectiva localização espacial na difusão sonora com suporte eletrônico digital. Mais especificamente, apresentamos uma solução aplicada à distribuição espacial, em tempo real, de fluxos sonoros gerados por Síntese Granular.

O controle espacial aplicado a sons granulares é um procedimento bastante recorrente na literatura. Desde o surgimento da técnica granular, a atribuição de uma localização aural específica a cada partícula ou grão sonoro em sistema multicanal tem se mostrado como um procedimento bastante idiomático e enriquecedor das paisagens sonoras geradas [Roads, 2001]. Dada a enorme quantidade de grãos usualmente necessários para a formação de um determinado segmento sonoro, é comum à síntese granular o emprego de controle macroestrutural, que determine, dentre outros parâmetros, a localização específica para cada grão no espaço de projeção sonora.

Podemos observar o uso frequente de métodos matemáticos na modelagem da distribuição espacial de síntese granular como, por exemplo, através de distribuições aleatórias, ou estocásticas, dos grãos. Estes métodos, no entanto, apresentam, na maioria das vezes, pouca ou nenhuma relação com os outros aspectos sonoros e/ou musicais em contexto. Trabalhos recentes têm adotado outras estratégias de auralização dos grãos sonoros. Kim-Boyle (2006), Wilson (2008) e Barreiro (2011) utilizaram o algoritmo *boids* para o controle de espacialização do fluxo granular em sistemas multicanal. Souza e Maia (2009) desenvolveram um modelo baseado em Algoritmos Genéticos (G.A.) para o controle de diversos parâmetros da síntese granular, dentre eles, a posição panorâmica de cada grão no sistema estereofônico. Em trabalhos atuais sobre Síntese Concatenativa, observamos o emprego de métodos que associam a posição aural de cada segmento sonoro aos Descritores de Áudio extraídos do mesmo. No sistema *CataRT*, do IRCAM, por exemplo, os *corpus sonoros* são distribuídos em um espaço bi ou tridimensional¹ em função da classificação dos descritores escolhidos pelo usuário [Einbond & Schwarz, 2010].

Em direção próxima aos trabalhos supracitados, nossa pesquisa tratou do desenvolvimento de ferramentas composicionais que tratassem a espacialização sonora como um processo conjugado, ou par a par, com outros aspectos da criação musical. O sistema *GTSpaT* foi então concebido como um modelo de síntese e espacialização do fluxo granular, no qual se estabelece uma função entre qualidades timbrísticas e a posição espacial de cada grão sonoro na difusão em suporte eletroacústico. O sistema foi implementado em *Pure Data* (PD) e utiliza a biblioteca *iem_ambi* desenvolvida por Thomas Musil e colaboradores (2005) para a espacialização sonora com *Ambisonics* [Gerzon, 1973] e algoritmos adaptados da biblioteca *PDescriptors*, desenvolvida por Monteiro (2010), para a extração de características do sinal de áudio.

2. O sistema GTSpaT

A arquitetura do sistema foi concebida em três módulos distintos, sendo o primeiro um módulo de análise e classificação dos segmentos sonoros, o segundo de síntese e espacialização e, por último, o módulo interativo (cf. Figura 1). O usuário seleciona inicialmente uma amostra sonora e define os parâmetros de análise da mesma. No módulo de análise, a amostra de áudio é então seccionada em pequenos quadros (ou grãos), que são analisados segundo um conjunto de descritores acústicos. Os dados obtidos dos descritores para cada grão são alocados na forma de um vetor. A posição espacial de cada grão é então uma função deste vetor denominado *vetor de características*. No módulo de interação o usuário define quais grãos serão enviados ao módulo de síntese e espacialização. Os grãos sonoros selecionados, juntamente com sua posição espacial, são endereçados ao módulo de síntese e espacialização, responsável pela codificação do sinal monofônico em relação à sua posição espacial e, em seguida, pela decodificação dos sinais resultantes para a quantidade de saídas de áudio do sistema.

¹ Neste, a terceira dimensão é representada graficamente através de uma escala de cores.

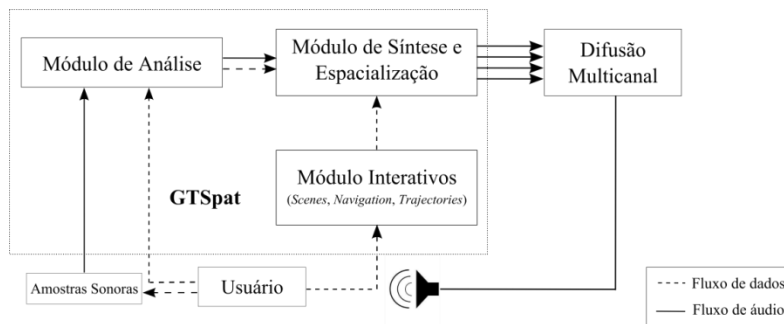


Figura 1: Arquitetura geral do Sistema GTSpac

Da amostra digitalizada selecionada pelo usuário são tomadas sucessivas seqüências de pontos, de tamanho invariável e com uma taxa de sobreposição (interseção das sequências) entre as mesmas. Para cada quadro ou seqüência de pontos, ponderados por uma envoltória de amplitude, atribui-se então um *vetor de características*, resultante de um conjunto de medidas extraídas do sinal, as quais serão denominadas por descritores acústicos. Estes, como o próprio nome indica, representam uma maneira econômica de substituir o objeto por algo, geralmente, muito mais simples, e que é suscetível de uma análise que reflete potenciais propriedades do objeto. Sua utilização torna-se bastante útil, principalmente para sistemas sonoros digitais em tempo real, nos quais processamentos de um sinal podem alcançar tamanhos proibitivos. Optamos então pelo uso de descritores que, embora tenham caráter reducionista em relação ao objeto sonoro analisado, podem ser muito úteis como uma taxonomia para uso imediato em espacialização, por exemplo. No nosso caso, tais descritores permitem o delineamento de comportamentos presentes no sinal através de ponderações estatísticas elementares, tanto no domínio do tempo quanto das frequências, e que possuem correlações diretas com a percepção sonora.

Em nosso modelo, o vetor de características é composto por duas categorias de descritores acústicos: *temporais* e *espectrais*. A primeira consiste de medidas computadas a partir do sinal representado no domínio do tempo, enquanto que a segunda é composta de medidas associadas a informações computadas da representação do sinal no domínio das frequências, através da Transformada Discreta de Fourier. Os descritores utilizados no sistema GTSpac foram: *energia média*, *taxa de cruzamento por zero*, *centroide espectral*, *desvio médio espectral*, *obliquidade espectral*, *curtose espectral*.

A definição do tamanho do quadro de análise e a taxa de sobreposição são previamente definidas pelo usuário. A quantidade de pontos analisados a cada quadro pode variar de 441 a 8820² e a taxa de sobreposição entre quadros pode variar entre 0 a 75 por cento do tamanho da janela temporal de análise. O usuário pode também definir qual tipo de envoltória de amplitude aplicar aos quadros de análise, sendo elas:

² Para os tamanhos de janela não correspondentes a valores de potências de dois, são adicionados zeros até que se complete a quantidade de pontos equivalentes à potência de dois superior mais próxima. Este procedimento é conhecido na literatura como *Zero Padding*.

senoidal, gaussiana, exponencial inversa (*rexpodec*), exponencial (*expodec*), retangular e ADSR³.

Definidos tais parâmetros, cada quadro da amostra sonora é analisado extraindo-se então os dados de interesse, ou seja, os valores numéricos dos descritores do sistema o quais são armazenados em um vetor numérico. A localização espacial de cada grão é dependente das entradas de seu respectivo vetor, as quais são convenientemente normalizadas e ponderadas por pesos definidos pelo usuário.

Matematicamente, em nosso modelo, denotamos por n a quantidade de descritores do sistema e m a quantidade de grãos sonoros obtidos a partir de uma amostra de áudio. Seja d_{kj} o valor do descritor acústico k atribuído ao grão sonoro j , com $k = 1, 2, 3, \dots, n$ e $j = 1, 2, 3, \dots, m$ e ainda p_k, q_k e r_k pesos atribuídos ao descritor k , com $0 \leq p_k, q_k, r_k \leq 1$ e normalizados como $\sum_{k=1}^n p_k = \sum_{k=1}^n q_k = \sum_{k=1}^n r_k = 1$. Definimos a posição espacial (x, y, z) ⁴ para cada grão j como:

$$x_j = \sum_{k=1}^n p_k d_{kj} \quad y_j = \sum_{k=1}^n q_k d_{kj} \quad z_j = \sum_{k=1}^n r_k d_{kj}$$

O módulo de síntese e espacialização é composto por três processos distintos: *granulação, codificação e decodificação espacial*. O primeiro deles é responsável pela seleção do fragmento do áudio amostrado (o grão) e a multiplicação do sinal pela envoltória de amplitude pré-estipulada. Na codificação, o sinal monofônico é processado em relação à distância (r) e direção espacial com ângulos de coordenadas polares (θ, ϑ). Para a codificação de distância, utilizamos filtros do tipo IIR (*Infinite Impulse Response*) para simular a atenuação da energia do espectro de frequências em decorrência da distância de cada fonte em relação ao ouvinte [Daniel, 2003].

Para a codificação direcional, utilizamos um sistema Ambisonics de 2ª ordem. O codificador é responsável pela renderização do sinal monofônico enviado pelo granulador a uma quantidade de canais que simulam o posicionamento espacial virtual do mesmo e que, por sua vez, são enviados ao decodificador. Na codificação, o sinal de cada um desses canais é multiplicado por uma matriz de ganho que determina o quanto cada um destes será enviado para cada canal da difusão sonora. A matriz de ganho varia em relação à configuração de alto-falantes estipulada ao sistema. A Figura 2 ilustra o processo de síntese e espacialização do sistema GTSpac.

3. Estratégias interativas

Através dos módulos de interação, o usuário pode controlar o sequenciamento do fluxo granular que será enviado ao módulo de síntese e espacialização. A fim de explorar

³ Ressaltamos que o uso de janelas com as envoltórias de amplitude acima não são exclusivas. Muitas outras podem ser definidas. Claramente o conteúdo espectral do sinal sonoro é bastante afetado pelos diferentes tipos de amplitude. Por exemplo, ADSR é nada mais que uma aproximação linear por partes da maioria dos sinais sonoros que, de alguma forma, possuem ataque (*attack*), decaimento (*decay*), sustentação (*sustain*) e relaxamento (*release*). Sua simplicidade e plasticidade é a razão de que é muito utilizada em sintetizadores.

⁴ O modelo admite distribuições espaciais bi- ou tridimensionais. Para o primeiro caso, adota-se $z_k = 0$, ou seja, a elevação da fonte sonora na espacialização é considerada constante.

distintas possibilidades de controle da espacialização, construímos diferentes propostas interativas das quais resultam os três módulos de interação: *GTSpat Scenes*, *GTSpat Navigation* e *GTSpat Trajectories*. Cada um destes módulos explora determinadas propriedades de síntese e espacialização sonora. Nos dois primeiros, procuramos explorar a localização espacial dos grãos como elemento de interação. O primeiro baseia-se na metáfora de múltiplas fontes sonoras fixas no espaço, pelo qual o usuário detém o controle paramétrico de fluxos granulares distribuídos em distintas regiões da espacialização. O segundo, por outro lado, baseia-se na metáfora de uma única fonte sonora que se desloca no espaço através do controle gestual do usuário. No terceiro módulo, procuramos salientar as propriedades cinéticas na interação entre o sistema e o usuário, utilizando a *diretividade* e taxa de deslocamento de movimento como parâmetros de controle da síntese e espacialização. Nos tópicos a seguir descreveremos com maior detalhe cada um desses módulos interativos.

3.1. GTSpat Scenes

Neste módulo, o usuário detém o controle de diversos fluxos granulares espalhados através do espaço de difusão sonora. Tendo distribuído os grãos no espaço bi ou tridimensional de controle da espacialização através de seu vetor de características, o usuário pode então determinar sub-regiões deste espaço as quais definirão cada um dos objetos da cena espacial. Estas sub-regiões são representadas na interface por retângulos (cf. Figura 3), que podem ser manipulados pelo usuário (posição e área) a fim de se definir qual grupo de grãos formará o referido *objeto na cena espacial*. O sistema comporta até quatro objetos sonoros simultâneos, cada qual composto por até dez fluxos granulares síncronos (ou quase síncronos). A quantidade de fluxos e intervalo de tempo entre os grãos em cada um dos objetos são definidos pelo usuário, podendo este último variar entre 1 e 10 vezes a duração do grão nesta versão do GTSpat. Cada objeto possui ainda um controle individual de amplitude, que permite que o usuário determine diferentes níveis de intensidade para cada objeto na cena espacial. Selecionada a área de um determinado objeto, o granulador de cada fluxo deste objeto pode sortear de maneira aleatória qualquer grão que esteja contido na área associada a este objeto.

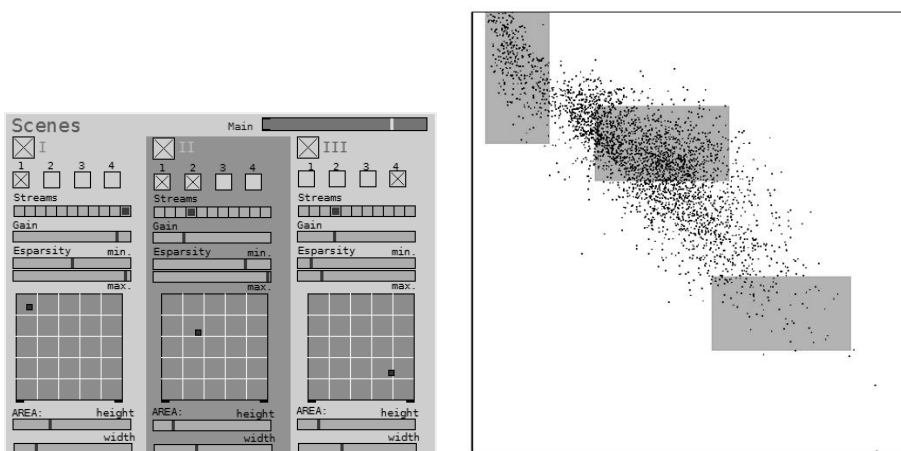


Figura 3: Interface de controle e visualização do GTSpat Scenes

3.2. GTSpat Navigation

Neste segundo módulo interativo, o sequenciamento temporal dos grãos sonoros como também a sua espacialização são controlados pelo usuário através de um “navegador bidimensional”. Este consiste de um quadrado virtual com tamanho definido pelo usuário e posição indicada através de uma interface de controle. A área demarcada pelo navegador no espaço de controle da espacialização, ao qual denominaremos Espaço de Navegação, define a região alvo do sistema. A partir da distribuição dos grãos no Espaço de Navegação, pela ponderação de seus vetores de características, o sistema envia ao módulo de síntese e espacialização os grãos circunscritos na área demarcada pelo navegador.

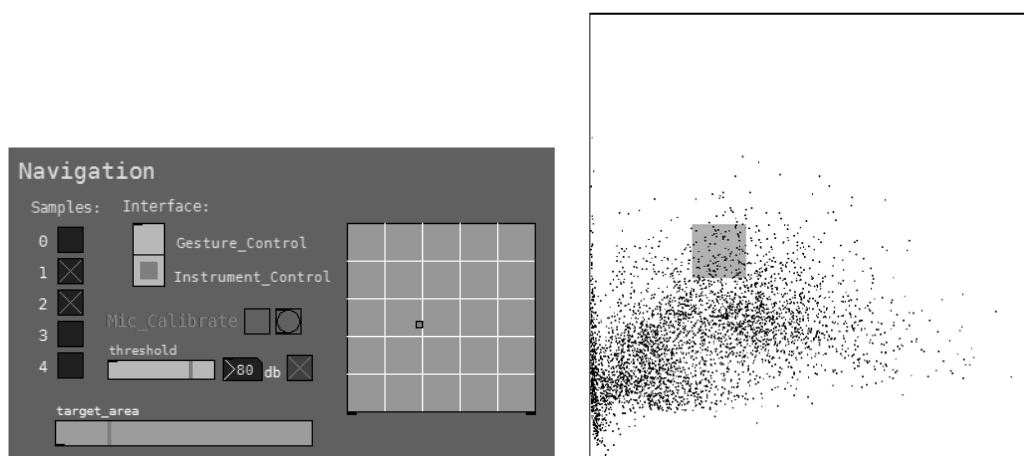


Figura 4: Interface de controle e visualização do GTSpat Navigation

Abordamos dois diferentes tipos de interface de controle. O primeiro consiste em interfaces gestuais capazes de reportar ao sistema informações de posicionamento espacial (*mouse* ou controladores gestuais de jogos, por exemplo). O segundo tipo consiste num microfone, pelo qual o sistema recebe informações de áudio que serão processadas e analisadas pelo próprio sistema e, em seguida, transformadas em informação numérica referente à posição do navegador.

O processamento aplicado à entrada de áudio da interface de controle é similar ao aplicado às amostras sonoras no Módulo de Análise. O áudio captado pelo microfone é analisado em tempo real⁵ pelo mesmo conjunto de descritores do Módulo de análise. O vetor de características extraído em tempo real é ponderado pelos mesmos pesos atribuídos pelo usuário para cada eixo da espacialização e em seguida atribuído à posição bidimensional do navegador no Espaço de Navegação. Os dados absolutos obtidos da análise dos descritores são normalizados através de um procedimento de calibração do sistema.

⁵ A latência do processamento em tempo real é determinada pelo tamanho da janela de análise estipulada pelo usuário para cada amostra analisada previamente, ou seja, pode variar entre 10 a 200 milissegundos.

3.3. GTSpat Trajectories

Com a criação deste módulo, procuramos estabelecer outras possibilidades de interação do sistema, utilizando medidas de segunda ordem das posições registradas pela navegação do usuário. As coordenadas cartesianas do Navegador, ao serem manipuladas pelo usuário no Espaço de Navegação, são registradas no sistema como informações de trajetória, em magnitude (tamanho do deslocamento) e ângulo. A partir desta trajetória, à qual denominaremos por “alvo”, o sistema compara as possíveis trajetórias dentre o conjunto de posições cartesianas atribuídas a cada grão no Espaço de Navegação. O sistema elege então a trajetória, dentre todas possibilidades de deslocamento condicionado pela distribuição dos grãos, que mais se “assemelhe da trajetória alvo” segundo uma conveniente medida, ou função, de comparação. A cada “passo” da trajetória no Espaço de Navegação o sistema compara as possibilidades de deslocamento dos grãos distribuídos no Espaço de Navegação e elege aquele que resultará numa trajetória “próxima” daquela estipulada pelo usuário como “alvo”.

Para a definição das trajetórias resultantes do sequenciamento dos grãos, utilizamos um processo markoviano de decisão, pelo qual o sistema elege o grão subsequente do fluxo em função da posição espacial do grão atual. Dada a coordenada inicial (x_0, y_0) como sendo a posição espacial de um grão emitido pelo sistema. A partir de uma trajetória “alvo” \vec{T} , definida pelo usuário e com magnitude $\|\vec{T}\|$, e ângulo de deslocamento σ_T , o sistema calcula as magnitudes e ângulos de todas as possibilidades de trajetórias em relação ao seu ponto inicial (x_0, y_0) . Dentre as possíveis trajetórias, condicionadas pela quantidade e distribuição dos grãos sonoros no Espaço de Navegação, o sistema escolhe a que mais se assemelhe à trajetória “alvo”. No decorrer de sucessivos “passos” do trajeto, o vetor “alvo” pode manter-se fixo ou também alterar-se a cada novo estado.

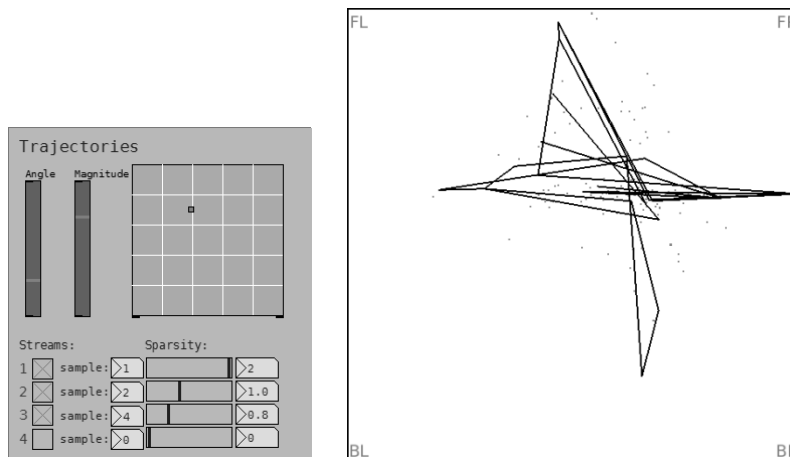


Figura 5: Interface de controle e visualização do GTSpat Trajectories

Em nosso modelo, o sistema sorteia aleatoriamente um grão sonoro inicial de posição definida por (x_0, y_0) para cada fluxo granular (*streams*). Definimos $\vec{U}_j = (x_j - x_0, y_j - y_0)$ como o vetor formado entre a posição (x_0, y_0) do grão inicial emitido e a

posição (x_j, y_j) de cada grão j ⁶, com magnitude $\|\vec{U}_j\|$ e ângulo σ_j . Definimos ainda a trajetória “alvo” do sistema pelo vetor $\vec{T}_n = (x_n - x_{n-1}, y_n - y_{n-1})$ formado pelas diferenças sucessivas⁷ de posição (x_t, y_t) do navegador, com magnitude $\|\vec{T}\|$ e ângulo σ_T . A comparação entre a trajetória alvo e as possíveis trajetórias do sistema é dada pela função:

$$l_{jt} = h \left(1 - \frac{\|\vec{U}_j\| - \|\vec{T}\|}{\sqrt{2}} \right) + \tilde{h} \left(\frac{\sigma_j - \sigma_t}{\pi} \right),$$

onde h e \tilde{h} são pesos que podem ser definidos pelo próprio usuário, a fim de se atribuir maior ênfase ao tamanho ou à direção do trajeto “alvo”.

Nesta fórmula, o sistema calcula dentre as trajetórias possíveis a partir do ponto (x_0, y_0) , com índices $\|\vec{U}_j\|$ e σ_j , a que mais se aproxima da trajetória “alvo” de magnitude $\|\vec{T}\|$ e ângulo σ_T indicada pelo usuário. O sistema seleciona o grão j que obteve maior índice l_{jt} em relação à trajetória alvo e o envia ao granulador. Uma nova medida é então calculada, sucessivamente, a partir da posição espacial do grão selecionado, comparando as possíveis trajetórias em relação à trajetória “alvo”.

4. Estratégias de adaptação

Associados aos módulos de interação do sistema GTSpaT, construímos dois submódulos adaptativos. O propósito da criação destes submódulos foi de estabelecer novas instâncias de interação entre o sistema computacional e o usuário. Para isto, procuramos estabelecer estratégias nas quais o sistema se adapta (ou se auto-regula) em busca de um estado ótimo (distribuição dos grãos) que mais se aproxime das trajetórias do navegador no Espaço de Navegação. Em ambos os submódulos o sistema procura redistribuir os grãos no Espaço de Navegação a partir desses dados cinemáticos extraídos da navegação do usuário.

4.1. Controle adaptativo do vetor de características

No primeiro submódulo, os pesos dos descritores associados a cada um dos eixos da espacialização são redistribuídos de maneira automática, por uma função distância entre a distribuição do conjunto de dados de cada descritor e a distribuição dos pontos das trajetórias do navegador. Adotamos três aspectos como critérios para análise das trajetórias percorridas pelo navegador, sendo eles o ponto médio dentre as posições percorridas em cada eixo do Espaço de Navegação, o grau de dispersão destes pontos em relação ao ponto médio para cada eixo e, por último, o grau de assimetria da distribuição destes pontos em relação ao ponto médio. As mesmas medidas extraídas da navegação do usuário são também realizadas para o conjunto de dados de cada descritor.

⁶ Com $j = \{1, 2, 3 \dots m\}$, referente à quantidade de grãos extraídos da amostra sonora.

⁷ Definimos a taxa de amostragem pela duração temporal equivalente ao *hop-size* aplicado à amostra sonora no Módulo de Análise, conforme descrito na arquitetura geral do sistema.

5. Considerações Finais

Apresentamos neste artigo uma metodologia para controle de espacialização de sons granulares em tempo real. Descrevemos a estrutura algorítmica do sistema GTSpac e, mais detalhadamente, abordamos estratégias de interação e adaptação abordadas. De modo geral, a implementação computacional do modelo comportou-se de maneira bastante estável e apresentou resultados sonoros satisfatórios na distribuição dos grãos pela difusão sonora. Acreditamos que as estratégias interativas por nós adotadas demonstraram-se também bastante eficazes e sintetizam um rico conjunto de potencialidades aplicadas à criação musical. Cada módulo interativo desenvolvido explora diferentes características da espacialização, que são manipuladas através de informações recebidas do usuário. Nos dois primeiros, *Scenes* e *Navigation*, delineamos processos reativos de processamento, enquanto que no terceiro módulo, *Trajectories*, utilizamos mecanismos computacionais de tomada de decisão em função de dados de segunda ordem das informações recebidas. Através dos módulos adaptativos, estabelecemos ainda estratégias de organização interativa do sistema, que pode alterar-se mediante dados de segunda ordem das informações recebidas do usuário.

6. Projeções Futuras

Em sequência ao trabalho aqui apresentado, pretendemos aprimorar o algoritmo de espacialização utilizado, sobretudo no que tange à codificação de informação de distância e na simulação de ambientes acústicos. Pelo custo computacional que isto implicava, optamos nesta primeira versão do sistema restringir a espacialização apenas à simulação de posicionamento das fontes sonoras. A codificação de informações relacionadas às primeiras reflexões e reverberação de fontes sonoras virtuais é elemento bastante rico à espacialização e muito contribui para o grau de realismo da cena acústica criada. Posteriormente, pretendemos utilizar codificadores/decodificadores Ambisonics de ordem superiores, bem como aplicar técnicas híbridas de espacialização, que garantam maior definição da auralização de fontes sonoras virtuais. Com a implementação de sistemas de espacialização mais sofisticados, poderemos ainda explorar interativamente diversos outros atributos espaciais, dentre os quais, destacamos a manipulação em tempo real das propriedades acústicas da sala e mudanças de perspectivas da auralização pela utilização de *sweet spot* dinâmico.

Consideramos ainda que a metodologia de espacialização empregada pode ser ampliada a outros contextos que não propriamente o da síntese granular ou da síntese concatenativa. Procedimentos similares podem, por exemplo, ser utilizados na distribuição espacial de sons de longa duração, que se deslocariam no espaço da difusão em função das medidas extraídas em nível micro-temporal. Em continuidade à nossa pesquisa, pretendemos estender o modelo elaborado pela utilização de diferentes interfaces que, sobretudo, estabeleça um diálogo mais direto entre o espaço físico da difusão e o espaço virtual das fontes sonoras.

7. Agradecimentos

Agradecemos às agências de fomento FAPESP e CAPES pelo financiamento desta pesquisa.

8. Referências

- Barreiro, D. L. (2011) “Manipulação de amostras sonoras em contexto interativo”. *Anais do Encontro Internacional de Música e Arte Sonora (EIMAS)*, Juiz de Fora, Brasil.
- Daniel, J. (2003). “Spatial Sound Encoding Including Near Field Effect: Introducing Distance Coding Filters and a Viable, New Ambisonic Format”. *Proceedings of AES 23rd International Conference*, 23-25.
- Einbond, A; Schwarz, D. (2010) “Spatializing Timbre with Corpus-Based Concatenative Synthesis”. International Computer Music Conference (ICMC) June 1–5, 2010, New York. *Proceedings of the International Computer Music Conference*, NY.
- Foote, J. (1997) “Content-based Retrieval of Music and Audio. Multimedia Storage and Archiving Systems II”. *Proceedings of SPIE*.
- Gerzon, M. A. (1973) “Periphony: With-Height Sound Reproduction”. In: *Journal of Audio Eng. Soc.*, v. 21, p. 2-10, 1973.
- Kim-Boyle, D. (2006) “Spectral and Granular Spatialization with Boids”. *Proceedings of the 2006 International Computer Music Conference*. New Orleans, p.139-142.
- Monteiro, A. (2010) *PDdescriptors*, versão 1.2. Campinas, Brasil.
- Otondo, F. (2008). “Contemporary trends in the use of space in electroacoustic music”. *Organised Sound*, 13 (1), 77-81.
- Musil, T.; Noisternig, M.; Hoeldrich, R. (2005) "A Library for Realtime 3D Binaural Sound Reproduction in Pure Data (PD)", *Proc. Int. Conf. on Digital Audio Effects (DAFX-05)*, Madrid, Spain, September 20-22.
- Puckette, M., Apel, T. (1998) “Real-time audio analysis tools for Pd and MSP”, *Proceedings of International Computer Music Conference*. San Francisco: International Computer Music Association, pp. 109-112.
- Roads, C. (2001) *Microsound*. Cambridge: MIT Press, 2001.
- Schumacher, M.; Bresson, J. (2010). "Spatial Sound Synthesis in Computer -Aided Composition." *Organised Sound*, 15(3), p.271-289.
- Sousa, F. F.; Maia, A. (2009) “Evolutive Processes for Granular Synthesis in Extended Gabor Spaces”, 7º Congresso de Engenharia de Áudio da AES Brasil, 2009, São Paulo. *Anais do 7o Congresso de Engenharia de Áudio da AES Brasil*. São Paulo
- Wilson, S. (2008) “Spatial Swarm Granulation”, *Proceedings of the 2008 International Computer Music Conference*. Belfast: SARC, ICMA