

Conversão de imagens para sons através de três classes do OpenMusic

Image-to-Sound Conversion Through Three OpenMusic Classes

por Luiz E. Castelões

RESUMO

A intenção deste artigo é analisar três classes do aplicativo OpenMusic (versão 6.4), o “BPF”, “BPC” e “3DC”, no que se refere ao seu uso para a conversão de pontos de imagens 2D e 3D para parâmetros musicais, além de avaliar possíveis aperfeiçoamentos dos resultados sonoros/musicais obtidos por conversões de imagens para sons em patches específicos do OpenMusic.

Palavras-chave *OpenMusic; Conversão de Imagens para Sons; Composição Musical*

ABSTRACT

The main purpose of this article is twofold: (a) to analyze three classes from the application OpenMusic (version 6.4), “BPF”, “BPC” and “3DC”, as regards their use for the conversion of 2D and 3D image points to musical parameters, and (b) to assess possible improvements of musical results obtained by the conversion from image to sound in specific OpenMusic patches.

Keywords *OpenMusic; Image-to-sound Conversion; Music Composition*

Introdução

O objetivo deste artigo é descrever as possibilidades em conversão de imagens para sons, com fins composicionais, oferecidas por três classes do aplicativo OpenMusic/Ircam, ou OM¹ (o “BPF”, “BPC” e “3DC”), analisando-as em seguida segundo a natureza dos dados gráficos e/ou numéricos de entrada, suas limitações técnicas e a relação estabelecida entre imagem e som (ou seja, especificidades da transcrição de imagem para som em cada caso). Em seguida, logo da inserção destas classes em exemplos específicos de programações (“patches”) de OM, serão abordados os agenciamentos necessários à conversão de seus dados (numéricos, gráficos) para parâmetros sonoros/musicais, além de sugestões gerais de refinamento de algoritmos para esta conversão.

No contexto do tema mais amplo da conversão de imagens para sons com fins composicionais, o qual pode incluir, por ex., a conversão de cores para ritmos e harmonias através do sistema HSV, como é o caso em obras recentes de Hector Parra como “L’aube assaillie” (PARRA I ESTEVE, 2005), a presente investigação se debruça especificamente sobre a conversão de pontos de uma imagem imóvel 2D ou 3D (o que inclui, por ex., fotografias, desenho livre, ou listas de coordenadas de objetos 3D escaneados) para parâmetros musicais, ou seja, alturas, *onsets* (momentos de ataque), durações, dinâmica e timbre/orquestração (este último via atribuição de canais de MIDI).

O presente artigo é estruturado em duas partes centrais, das quais a primeira apresenta uma tabela comparativa das classes elencadas para análise e avaliação e a segunda demonstra o funcionamento destas classes em exemplos de patches de OM. As conclusões deste artigo serviram de ponto de partida para as atividades do COMUS – Grupo de pesquisa em composição musical da UFJF no primeiro semestre de 2011.

A carência de literatura sobre OpenMusic em português (contribuindo para seu pouco uso no Brasil, o que de maneira nenhuma reflete suas potencialidades para a composição musical), e mesmo em outros idiomas quando se trata de temas mais específicos (DIAKTINE, 2011; AGON et al., 2006), motiva e justifica o presente artigo, a partir do qual se esperam discussões mais frequentes acerca do OM em periódicos brasileiros.

BPF, BPC e 3DC: uma visão geral

Resumo das características básicas de cada classe analisada:

¹ No âmbito desta pesquisa, utilizei o OM 6.4 (disponível em <http://forumnet.ircam.fr/363.html?&L=1>). A recém-lançada versão 6.5 não traz nenhuma novidade em relação aos conteúdos deste artigo.

Tabela 1 – Quadro comparativo das classes BPF, BPC e 3DC

Classe OM Características	BPF	BPC	3DC
Natureza dos dados gráficos e/ou numéricos de entrada	Coordenadas XY. Um só traço sem interrupções e com apenas um valor de Y para cada X (ou seja, o traço evolui em apenas um sentido do gráfico)	Coordenadas XY. Um só traço sem interrupções e com possibilidade de múltiplos valores de Y para cada valor de X (ou seja, o traço pode ir e voltar livremente no plano XY)	Coordenadas XYZ. Aceita tanto listas de números (divididos em três entradas, XYZ), quanto coordenadas isoladas ou pareadas de BPFs.
Limitações técnicas	O fato de oferecer apenas um traço sem interrupções faz com que a conversão de imagens com muitos traços e/ou interrupções exija um grande número de BPFs, reunidos posteriormente através de um “list” e um “BPF-Lib” (para visualização integral)	A possibilidade de o traço ir e voltar livremente no plano XY o faz menos limitado graficamente do que o BPF. Ainda assim, possíveis interrupções nas imagens convertidas exigirão múltiplos BPCs	Nos casos em que os modelos 3D sejam de alta resolução (como nos arquivos PLY mais pesados da Stanford), o limite de processamento é ultrapassado
Especificidades da transcrição de imagens para sons	O BPF tem um correspondente facilmente identificável em música: o traço sem possibilidade de duplo valor para Y é análogo a uma linha melódica na escrita musical ocidental (se tomamos Y para as alturas e X para o fluxo do tempo)	A conversão de um BPC será tão diversa e complexa quanto o próprio BPC, tendo como resultado desde uma linha melódica (em cujo caso a função do BPC não diferiria da de um BPF) até texturas hetero- ou polifônicas e complexas (de acordo com a multiplicidade de idas e vindas do traço do BPC) e quaisquer situações/texturas intermediárias	A conversão de modelo 3D para parâmetros musicais se revela menos previsível do que em 2D, já que será necessário agenciar criativamente a distribuição dos pontos XYZ entre parâmetros musicais (algo sem analogia na escrita musical ocidental, essencialmente 2D). De forma geral, convém pré-analisar qual lista de pontos (X, Y ou Z) se prestará a cada parâmetro musical

Exemplos de patches no OM para conversão de imagens para sons

Com BPF

No patch a seguir, os traços de uma imagem (desenho livre, foto) são reproduzidos através de BPFs, reunidos (para visualização integral) em uma BPF-lib e inseridos em um esquema de algoritmos para conversão de pontos do traçado desta imagem para valores de parâmetros musicais (alturas, momentos de ataque, durações e dinâmicas).

O primeiro aspecto a ser ressaltado é a reunião dos traços dos BPFs em uma BPF-lib:

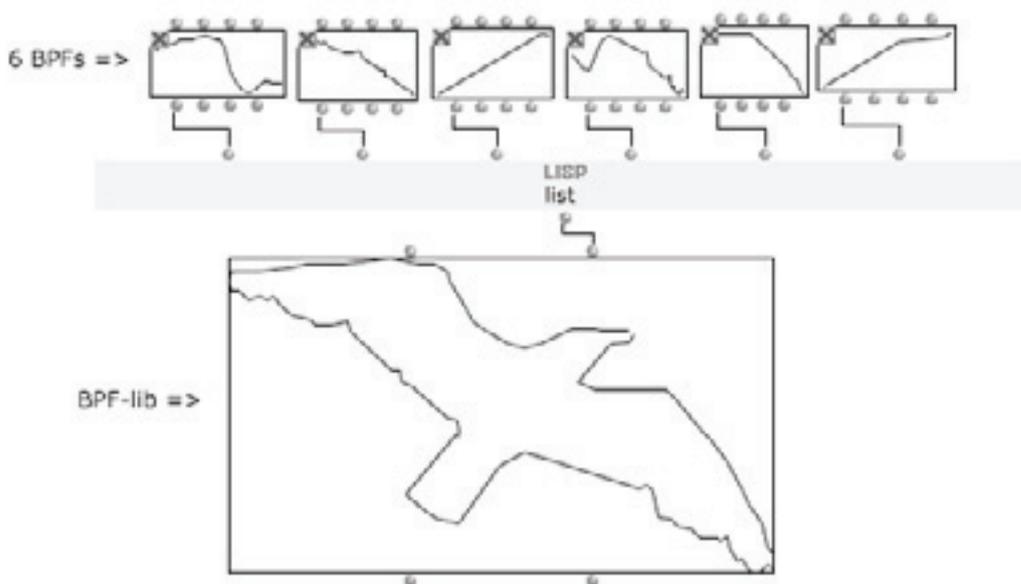


Figura 1 – 6 BPFs reunidos em uma BPF-lib

O objeto “list” reúne cada um dos BPFs que se encontram na parte superior da Fig. 1 no BPF-lib que se encontra na parte inferior da mesma. Esta é a única forma de se realizar com BPFs uma imagem que contenha vários traços (ou “curvas” no dizer do OM), já que o BPF só permite um ponto Y para cada ponto X. O circuito acima serve apenas para visualização integral da imagem a ser convertida, não influenciando o algoritmo para conversão de imagem para som a ser usado futuramente. Do ponto de vista estritamente gráfico, o BPF não constitui uma ferramenta eficiente para reproduzir imagens com muitas curvas, o que é demonstrado na Fig. 1 pelo fato de que o traçado de uma simples silhueta requereu nada menos do que 6 BPFs.

Para conversão de imagens para sons usando vários BPFs, deve-se conectar um

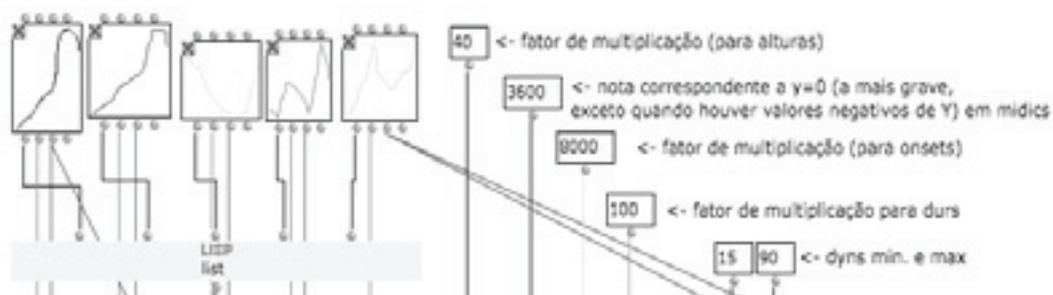


Figura 3 – Coeficientes inseridos na conversão de imagem para som via BPFs (onsets = momentos de ataque; durs = durações; dyns = dinâmicas)

Com BPC

A diferença fundamental que os BPCs oferecem em relação aos BPFs é a possibilidade de múltiplos pontos de Y para cada ponto de X, o que graficamente corresponde ao potencial de se ir e voltar na imagem a ser convertida (permitindo, por ex., a conversão de formas geométricas fechadas, imagens de letras e, portanto, um detalhismo mais acentuado).

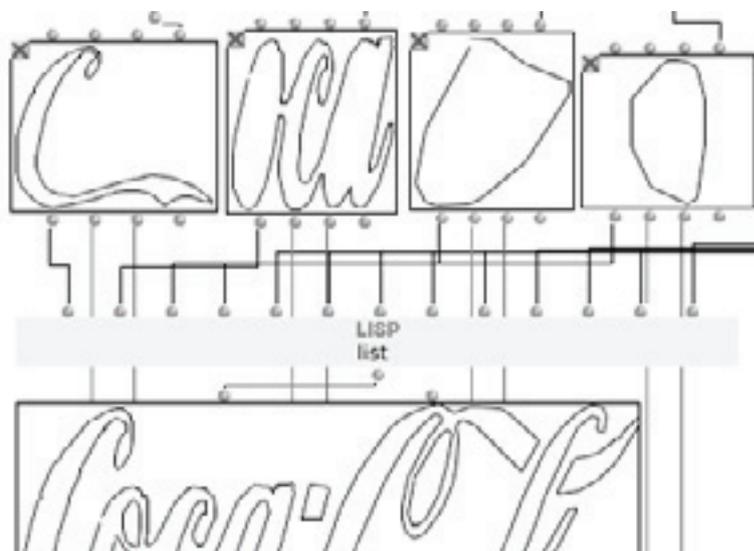


Figura 4 – BPCs reunidos em uma BPC-lib (maior eficiência do BPC, em relação ao BPF, aqui demonstrada por sua capacidade de realizar formas geométricas fechadas)

Do ponto de vista gráfico, a conversão com BPCs vai requerer um número significativamente menor de caixas BPC do que seria o caso com BPFs, ou seja, sob o ponto de vista estritamente gráfico trata-se de uma ferramenta mais eficiente de conversão de imagem para som. Sob a perspectiva sonora/musical, da mesma forma, a conversão de apenas um BPC pode gerar um repertório mais amplo de

resultados: desde uma só linha melódica, de modo análogo ao BPF, até texturas complexas, variando no tempo, dependendo da complexidade do traçado original.

Com 3DC

Na importação de arquivos de imagens 3D ao OpenMusic, para conversão de imagem para som, alguns aspectos que não estavam presentes nas conversões 2D com BPFs e BPCs precisam ser levados em conta.

Em arquivos nos formatos PLY (usado na base de dados da Stanford²) e OFF (usado na base de dados da Princeton³), será necessário editar o texto de modo a suprimir qualquer informação que não corresponda aos pontos de X, Y e Z. No âmbito desta pesquisa, tal tarefa foi feita rápida e diretamente sobre os próprios arquivos (convertidos em arquivos de texto), mas nada impede que se criem algoritmos que realizem tais edições dentro do próprio OM. A análise dos formatos e conteúdos de arquivos 3D e de sua utilização no OM merece um artigo à parte, aliás. Em auxílio ao leitor e usuário (e evitando maiores digressões ao tema central deste artigo), pode-se dizer brevemente, em relação aos formatos manipulados no âmbito desta pesquisa, que: o formato .XYZ só contém coordenadas e, portanto, já vem pronto para uso no OM (embora seu repertório de imagens disponíveis online não pareça especialmente promissor); os formatos .PLY e .OFF requerem alguma edição, a qual pode ser feita em um programa simples de edição de texto; e o formato .3DS abre como binário em um editor de texto e precisa, portanto, passar por uma conversão (para .PLY, por ex.) via programas como o Blender⁴ ou o MeshLab⁵.

Para a separação das coordenadas X, Y e Z (de modo a converter cada uma delas para diferentes parâmetros musicais), foi usado o expediente mostrado na figura da página seguinte.

O objeto “om-scale” posicionado antes do 3DC é aqui utilizado para normalizar as coordenadas de modo a que elas sejam visíveis no visor do 3DC (no caso, muito frequente aliás, em que as coordenadas inseridas contenham números significativamente menores que 1). Embora do ponto de vista gráfico seja mais rico utilizar uma definição melhor da imagem, durante a presente pesquisa, as melhores conversões (sob a perspectiva do resultado sonoro) foram obtidas com resoluções menores, as quais produzem menos redundância nos valores para conversão (e, subsequentemente, menor repetitividade nos valores dos parâmetros musicais). Isso se consegue diminuindo-se o valor do quinto input do 3DC (que, no exemplo acima, está com 6 casas decimais).

2 <http://graphics.stanford.edu/data/3Dscanrep/>

3 <http://shape.cs.princeton.edu/benchmark/>

4 Disponível para download no link: <http://www.blender.org/download/get-blender>

5 Disponível para download no link: <http://meshlab.sourceforge.net/>

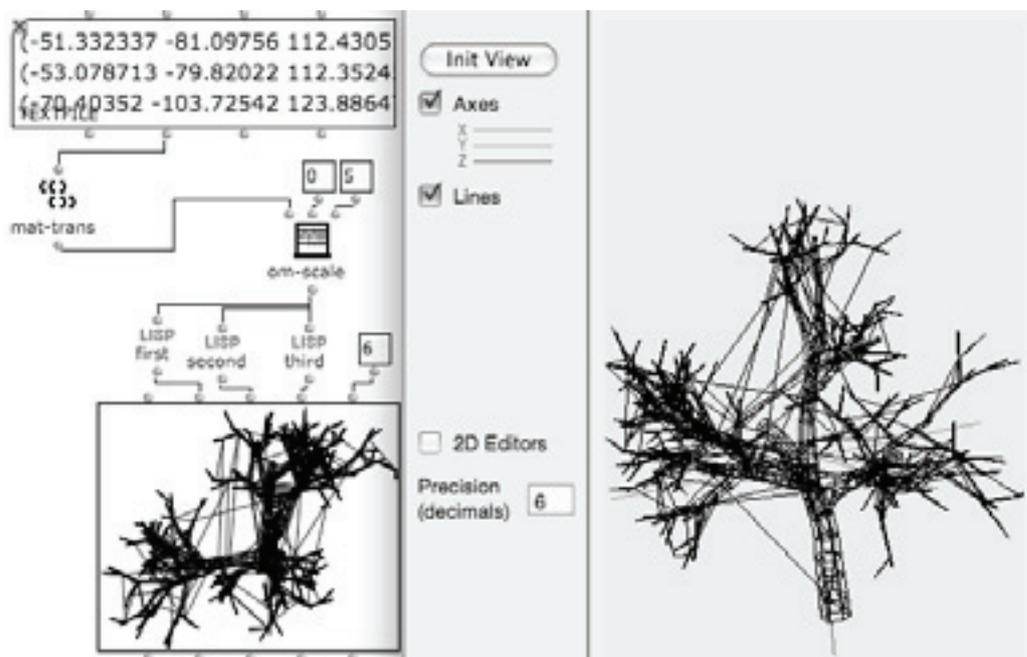


Figura 5 – Expediente para importação de coordenadas XYZ de um arquivo de texto (gerado a partir de um arquivo .OFF)

Na conversão de imagem 3D a som, o primeiro desafio criativo diz respeito ao agenciamento das três listas de pontos X, Y e Z em termos de parâmetros musicais, embora neste caso, ao contrário do 2D, a analogia com a escrita musical tradicional seja mais remota. Elezovic (2008, p. 7) sugere uma solução preliminar a este problema ao afirmar que a “terceira dimensão” em música corresponderia, em Scelsi por ex., a todos os parâmetros musicais que não a altura e a duração. Ele os lista como: articulação (clusters microtonais, trilos, tremolos, glissandos, vibratos largos, frulato), instrumentação, dinâmica e textura. Embora possamos partir desta premissa, a criação de algoritmos para a conversão de imagens 3D para parâmetros sonoros/musicais aqui demonstrada requer um nível significativamente maior de determinação do papel de cada lista de coordenadas em relação ao papel que irá exercer no parâmetro sonoro/musical correspondente. No que diz respeito às supracitadas bases de dados de arquivos de coordenadas 3D aqui utilizadas (de Stanford e Princeton), este papel só pôde ser determinado caso a caso, já que estes arquivos não possuem uma uniformização dos dados de X, Y e Z no sentido de que cada eixo tenha uma posição clara em relação ao objeto escaneado (por ex., largura, altura e profundidade). Ao menos no que se refere a estas bases de dados, portanto, o compositor/usuário deverá analisar para cada arquivo de coordenadas 3D qual eixo se adequa melhor a cada parâmetro musical no contexto de seu projeto composicional. O exemplo de patch a seguir ilustra um encaminhamento possível para a conversão de imagem 3D a som usando arquivos de imagem destas bases de dados.

Observe, primeiramente, que o patch da Fig. 6 apresenta uma segunda solução

para a necessidade de se normalizar os valores da lista de pontos da imagem para valores compatíveis com as unidades usadas pelo OM para os diversos parâmetros musicais; trata-se de uma alternativa à solução de coeficientes para os diversos parâmetros musicais usados na conversão com BPFs (Fig. 3). Desta vez, o “om-scale” permite substituir a livre multiplicação dos valores dos diversos parâmetros musicais pelo estabelecimento de valores mínimos e máximos para cada um destes parâmetros (um recurso útil ao se trabalhar, por ex., com instrumentos musicais, com registros e limitações técnicas específicas). Na solução anterior, com coeficientes, constatou-se alguma dificuldade em se obter ao mesmo tempo um conteúdo harmônico satisfatório aliado ao controle da tessitura a ser utilizada – dificuldade dirimida a partir do uso do “om-scale”.

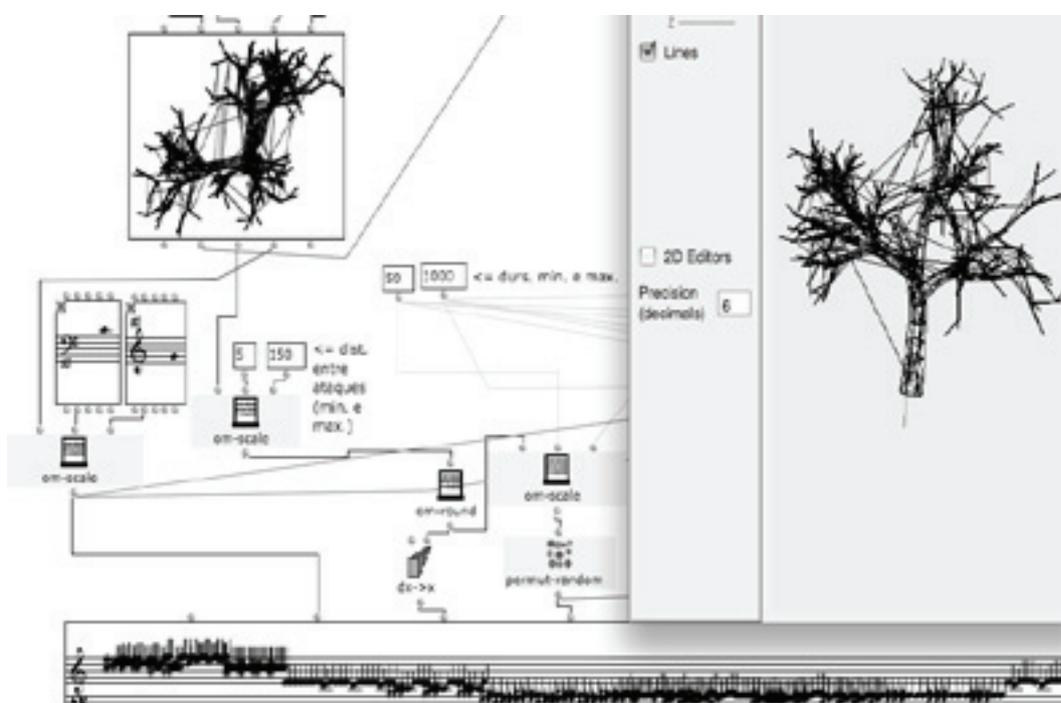


Figura 6 – “Om-scale” utilizado para alturas, momentos de ataque e durações

Para o estabelecimento de uma lista de valores de dinâmica (lembre-se que o OM faz corresponder dinâmicas como *p*, *mf* e *fff* a valores de 0 a 127), concorreram não só uma lista de pontos (do eixo de Z, no caso), mas também dois outros fatores: o inverso proporcional das durações (para que as notas tenham suas dinâmicas proporcionalmente reduzidas quanto mais longas fôrem) e o inverso proporcional das alturas (para que as notas também tenham suas dinâmicas proporcionalmente reduzidas quanto mais agudas fôrem). Este desenvolvimento já constitui um passo além da mera conversão de imagens para sons, na medida em que se trata de uma alteração do algoritmo preliminar de conversão com vistas à produção de alterações no resultado sonoro final – alterações estas determinadas fundamentalmente pela escuta do usuário/compositor (que segue, por assim dizer, uma lógica

musical/auditiva) e não mais pelos agenciamentos preliminares da conversão de imagens para sons (os quais seguem uma lógica formal/matemática).

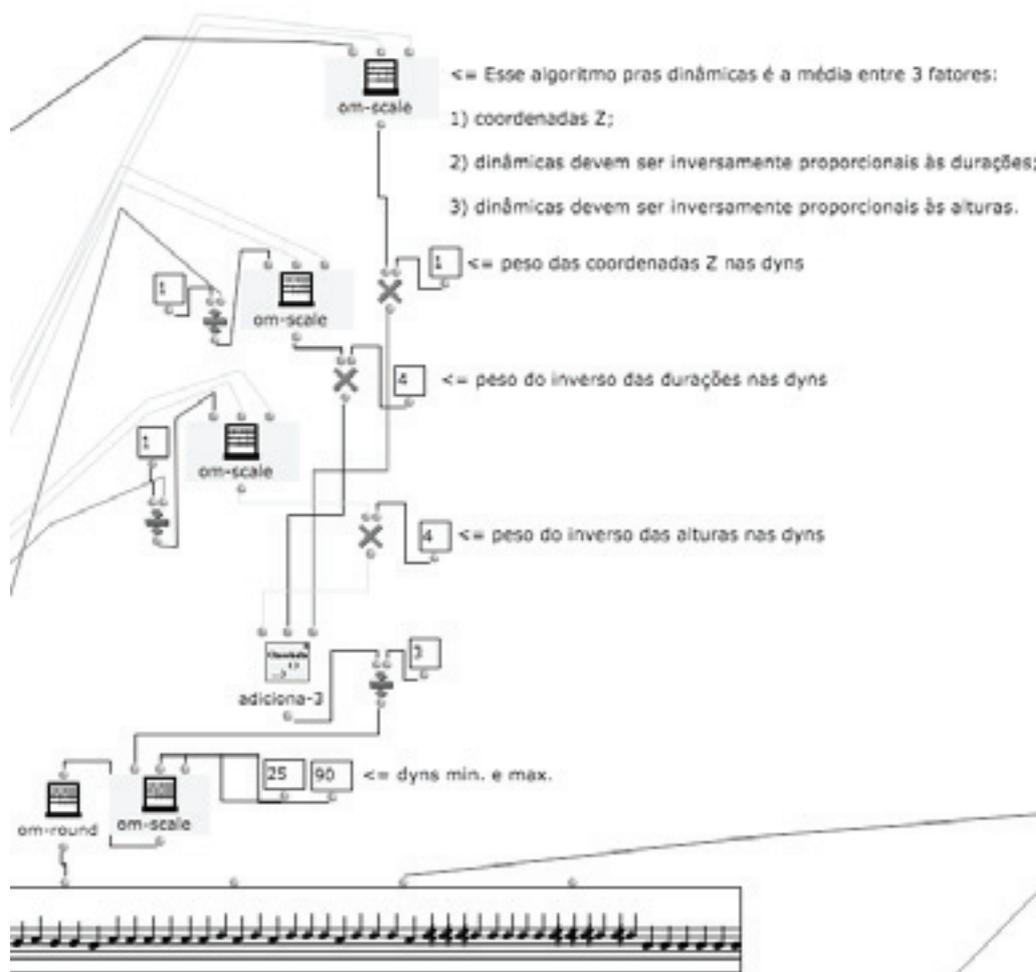


Figura 7 – Fatores concorrendo para a conversão de pontos de uma imagem 3D em uma lista de dinâmicas (“dyns”)

Em etapas posteriores do desenvolvimento de algoritmos para os momentos de ataque (*onsets*) foram usadas curvas sobrepostas às listas de valores geradas pelas coordenadas através do uso da função *sinus-ser* (a qual produz curvas análogas a senóides). Foi também incluída no algoritmo a possibilidade de se dosar pesos distintos tanto desta curva senoidal “macro” quanto da lista de coordenadas sobre o resultado sonoro final. É preciso observar novamente, entretanto, que a partir deste ponto do processo composicional, não se está mais operando na esfera da conversão de imagens para parâmetros sonoros/musicais, mas no sentido da distorção do resultado sonoro preliminar desta conversão a fim de se obter um resultado mais satisfatório à escuta. Ou seja, a partir daqui discutir-se-iam subjetividades relacionadas à obtenção de resultados sonoros específicos e individuais, e não mais aspectos objetivos de como se realizar a conversão de imagens para sons – estes são generalizáveis para outras obras musicais, aqueles apenas discutíveis caso a caso.

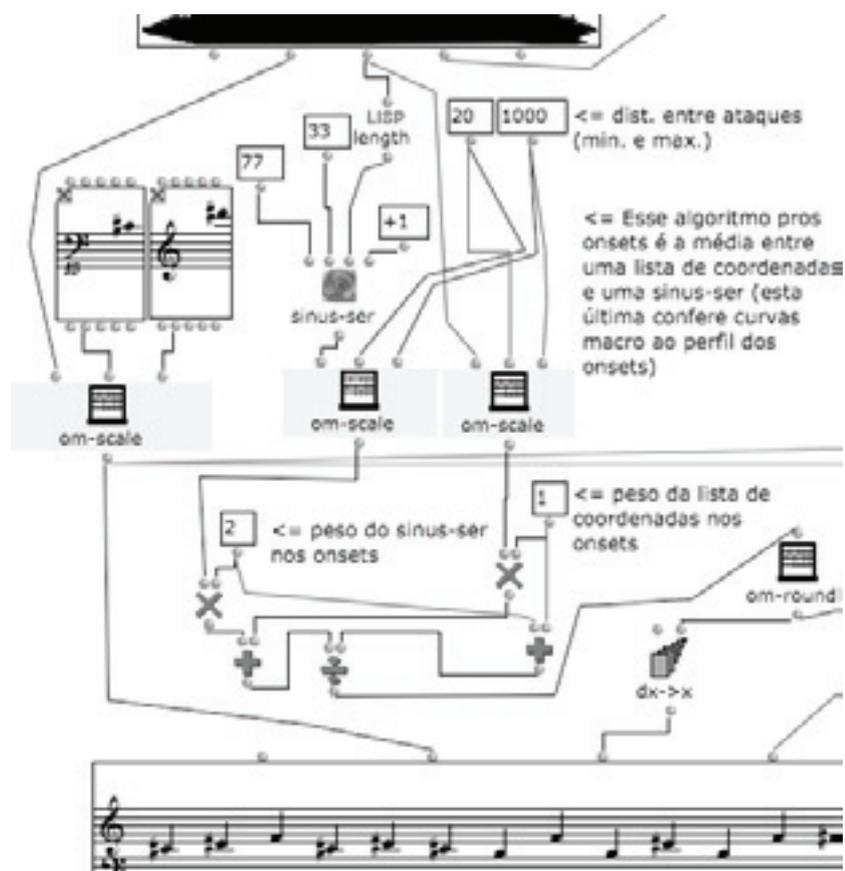


Figura 8 - Uso de uma sinus-ser no algoritmo dos momentos de ataque (*onsets*) de modo a sobrepôr uma curva “macro” aos valores da lista de coordenadas (e com o intuito de se refinar o resultado sonoro obtido)

Por fim, vale ressaltar um último aspecto da conversão de imagens para sons com as referidas classes que exerce um papel fundamental sobre os resultados sonoros obtidos a partir desta conversão. Trata-se da existência de padrões detectáveis como “triviais” pela escuta entre os valores de uma lista de coordenadas (presume-se aqui que o usuário está geralmente à procura de um resultado sonoro mais “humano” e menos “mecânico”, isto é, com alguma dose de irregularidade e imprevisibilidade).

O problema pode ser resolvido (ou pelo menos evitado) de duas formas distintas, de acordo com o procedimento de conversão que se esteja usando.

No caso do uso de modelos 3D previamente escaneados (das bases de dados de Stanford e Princeton, por ex.), qualquer regularidade no padrão de escaneamento em 3D resultará em regularidade em uma ou mais dimensões da lista de coordenadas (X, Y e/ou Z) e subsequentemente em regularidade no parâmetro musical que porventura se utilizar desta lista de coordenadas. O problema é facilmente detectável pela escuta do resultado sonoro obtido a partir da conversão – o usuário perceberá padrões “triviais” no parâmetro musical em questão (por ex., listas de

onsets [momentos de ataque] com um só intervalo de tempo, ou com muito poucos intervalos de tempo). Outra solução que pode se revelar bem mais trabalhosa (levando-se em consideração que os arquivos de coordenadas 3D aqui utilizados chegam a ter milhões de pontos) seria a análise prévia das coordenadas, ou do próprio processo de escaneamento, ou ainda do desenvolvimento de um algoritmo no próprio OM que “recortasse” a lista de coordenadas de forma suficientemente irregular, embora fazê-lo de modo a não desfigurar a imagem original provavelmente demande algum artesanato, ou talvez arruine o propósito inicial de conversão, com alguma fidelidade, de imagens para sons.

Já no caso em que o usuário esteja usando o mouse para fazer desenho livre ou contornando os traços de uma foto fixada sobre o fundo de um BPC ou BPF, a regularidade/previsibilidade excessiva em uma lista de pontos pode ser evitada com o uso da ferramenta “seta”, distribuindo-se os pontos sobre o plano de maneira irregular, em vez do uso da ferramenta “lápis” com o clique do mouse continuamente apertado (o que resultaria invariavelmente em uma lista de pontos regular em demasiado).

Conclusão

No presente artigo, foram resumidos aspectos essenciais da comparação entre as classes BPF, BPC e 3DC na conversão de traços de imagens imóveis em 2D e 3D para parâmetros musicais.

Embora, do ponto de vista da eficiência da conversão e do potencial de complexidade das imagens de origem, o 3DC seja uma ferramenta mais eficaz que o BPC e o BPF (e o BPC por sua vez seja mais eficaz que o BPF), a maior complexidade dos traços de uma imagem não representa necessariamente um maior interesse no resultado musical, o que equivale a dizer que o sucesso da conversão de imagens para sons (sob a perspectiva musical) não é proporcional às potencialidades da ferramenta OM utilizada – o 3DC não será necessariamente uma ferramenta mais bem-sucedida para esta conversão do que o BPF. O trabalho prático de refinamento de cada patch demonstrou muitas vezes o contrário, ou seja, que no processo de se obterem resultados musicalmente “satisfatórios”, foi necessário diminuir a quantidade de dados convertida e trabalhar com imagens menos detalhadas e “realistas”, seja através da redução da resolução da imagem (i.e., menos casas decimais no input do 3DC), do uso de arquivos 3D em formato .PLY de baixa resolução, ou ainda da preferência pelo trabalho de tracejo de pontos com o mouse sobre uma foto colada ao fundo de um BPF ou BPC (em 2D).

Foi também demonstrado que etapas posteriores do processo composicional envolvendo conversão de imagens para sons implicam a passagem de uma lógica formal/matemática na construção dos algoritmos para uma lógica musical/auditiva, a qual impõe mudanças aos algoritmos preliminares que não estão mais associadas

à tarefa de conversão de imagens para sons, mas respondem apenas a transformações que a escuta de um ouvinte em particular (o compositor/programador/usuário) deseja impor aos resultados sonoros preliminarmente obtidos.

Duas das implicações mais interessantes das três classes analisadas são (1) a possibilidade de fusão entre a atividade da composição musical e da criação gráfica (com possíveis repercussões na relação entre som e imagem na criação de música para produtos audiovisuais) e (2) a possibilidade de que a atividade da composição musical esteja acessível a artistas visuais sem formação musical.

Futuras etapas da pesquisa incluirão colaborações com pesquisadores/criadores das artes visuais (animadores, designers, etc.) que usam outros aplicativos em suas produções (AutoCad, por ex.) no sentido de articular produtos gerados nestes programas com os patches de conversão de imagens para sons criados no âmbito desta investigação para o OM, além da pré-análise dos traços de uma imagem a ser convertida a som em termos de suas possibilidades musicais, antevendo (ou anteuovindo) o potencial musical de uma imagem antes mesmo de sua conversão, inclusive para melhor selecioná-la e manipulá-la. Uma espécie de sinestesia entre imagem e som encarada de forma técnica, composicional.

Referências Bibliográficas

- > AGON, Carlos et al (eds.). 2006. The OM Composer's book, volume 1. Paris: Editions Delatour.
- > DIATKINE, Coralie. 2011. OM 6.4 – User Manual. Disponível em: <<http://support.ircam.fr/forum-ol-doc/om/om6-manual/co/OM-Documentation.html>>. Acesso em: 6 jun 2011.
- > ELEZOVIC, Ivan. 2008. "Scelsi's Approach to the 'Third Dimension' of Sound." Search: Journal for New Music and Culture (Fall 2008, Issue 3). Disponível em: <<http://www.searchnewmusic.org/elezovic.pdf>>. Acesso em: 2 jul 2011.
- > PARRA I ESTEVE, Hector. 2005. Pour une approche créatrice des interrelations structurelles entre les espaces acoustiques et visuels. Dissertação de D.E.A. em Artes da Cena e do Espetáculo – Opção Música. Université de Paris VIII, Vincennes – Saint-Denis.
- > SHILANE, Philip et al. 2004. The Princeton Shape Benchmark. Disponível em: <<http://shape.cs.princeton.edu/benchmark>>. Acesso em: 6 abr 2011.
- > THE STANFORD 3D SCANNING REPOSITORY. 2011. Disponível em <<http://graphics.stanford.edu/data/3Dscanrep/>>. Acesso em: 30 mar 2011.

*Este artigo é um sub-produto das atividades do "COMUS – Grupo de Pesquisa em Composição Musical da UFJF", sediado no Instituto de Artes e Design da Universidade Federal de Juiz de Fora (<http://www.ufjf.br/comus>) e financiado pelo edital CNPq Universal 2010.

Luiz Eduardo Castelões, professor adjunto – nível 2 no Instituto de Artes e Design – Universidade Federal de Juiz de Fora (UFJF)

lecasteloes@gmail.com