

# Composição de Paisagens Sonoras Digitais Através da Síntese Evolutiva

José Fornari<sup>1</sup>, Adolfo Maia Jr.<sup>1,2</sup>, Jônatas Manzolli<sup>1,3</sup>

<sup>1</sup>Núcleo Interdisciplinar de Comunicação Sonora – Universidade Estadual de Campinas (UNICAMP)

Caixa Postal 6166 – 13.091-970 – Campinas – SP – Brazil

<sup>2</sup>Departamento de Música, Instituto de Artes (DM/IA)

<sup>3</sup>Departamento de Matemática Aplicada (IMECC)

{fornari,adolfo,jonatas}@nics.unicamp.br

**Abstract.** *We present here the late development of the digital sonic structure system implemented in Pd (Pure Data) that uses Evolutionary Computation as a systemic and organizational strategy. The method here developed is called Evolutionary Sound Synthesis and is here described in its latest version where psychoacoustic descriptors are used to determine and manipulate its sonic typology. The synthesis results are here linked with the concept of soundscape design, as previously described by others and mentioned below. Finally, the late implementation is described together with its graphic interface.*

**Resumo.** *Apresentamos neste artigo o desenvolvimento final de um sistema de síntese de estruturas sonoras digitais desenvolvido em Pd (Pure Data) que utiliza-se de Computação Evolutiva como estratégia sistêmica e organizacional. O método que desenvolvemos, denominado de Síntese Evolutiva (SE), é aqui descrito, bem como as últimas etapas da pesquisa onde são utilizados descritores psicoacústicos para determinar e manipular a tipologia sonora. O resultado sonoro deste processo é por nós relacionado com o conceito de paisagem sonora digital, no sentido já discutido por outros autores e apresentados a seguir. Finalmente, descrevemos a implementação final do sistema e a sua interface gráfica.*

## 1. Introdução

Computação Evolutiva, CE, é uma metodologia de otimização computacional que busca ótimos locais para um problema definido por uma população de possíveis soluções que se transformam dinamicamente [Koza,97]. Os métodos de computação evolutiva não possuem qualquer condição inicial ou campo fixo de definição, teoricamente tais métodos poderiam ser usados na resolução de qualquer tipo de problema [Bäck,96]. A utilização de CE em Computação Musical e Computação Gráfica tem sido ampliada nos últimos anos e na literatura há vários estudos sobre tais aplicações. [Miranda,03] investigou o potencial de algoritmos denominados de "vida artificial" (Alife) no contexto da criatividade musical. [Soddu,02] usou algoritmos generativos aplicados à multiplicidade de possíveis soluções descritas como estruturas arquitetônicas e representadas por computação gráfica. [Garcia,00] usou algoritmos genéticos na



automação de projetos de técnicas de síntese sonora. [Gibbs,96] desenvolveu animações gráficas com figuras articuladas usando programação genética. [Biles,94] usou algoritmos genéticos para criar o software GemJam, um programa que articula improvisações de jazz em tempo-real. [Horner,93] desenvolveu um método evolutivo para a geração automática de algoritmos aplicados à síntese FM.

No trabalho aqui apresentado, a nossa motivação para utilizar CE como método de construção de paisagens sonoras baseia-se em três premissas:

1. CE permite que a geração de formas de onda, seu sequenciamento temporal e espacialização sejam dinamicamente transformados num contexto adaptativo.
2. A utilização de um conjunto alvo e formas de onda pré-escolhidas pelo usuário/compositor pode tornar-se um controle eficiente pelo qual guia-se a evolução do processo e seu consequente resultado sonoro.
3. O resultado sonoro combina variedade e similaridade adaptativa.

O estudo aqui reportado vem se desenvolvendo desde 1999 no NICS. O primeiro projeto foi denominado “*VoxPopuli*”, um ambiente evolutivo para gerar seqüências musicais através do protocolo MIDI (*Musical Instrument Digital Interface*) [Moroni et al,00], posteriormente, baseado nos resultados do primeiro sistema foi implementado o JaVOX, um sistema que ampliou as funcionalidades do primeiro e foi desenvolvido em JAVA ([http://www.geocities.com/Artemis\\_Moroni/JaVox/](http://www.geocities.com/Artemis_Moroni/JaVox/)). O uso de CE aplicada à síntese sonora, denominado de Síntese Evolutiva de Segmentos Sonoros, iniciou-se com um método baseado em curvas psicoacústicas [Fornari,01]. Dentro do contexto, criou-se uma metodologia baseada em um sistema imunológico artificial apresentada em [Caetano,05].

Recentemente, estudamos a utilização de funções de localização sonora ITD (*inter-aural time difference*) como um parâmetro de controle do genótipo sonoro [Fornari,07]. A utilização de localização espacial junto com a capacidade da síntese evolutiva de gerar sons com similaridade variante mostrou-se uma abordagem para a criação e controle dinâmico de paisagens sonoras digitais.

Nas próximas seções apresentamos a idéia geral do método, destacando as conexões entre a Síntese Evolutiva e descritores sonoros. Apresentamos também o modelo matemático desenvolvido, seguido da implementação em Pd.

## 2. Método: Síntese Evolutiva

Iniciamos com um breve resumo do processo de síntese que desenvolvemos. A seguir apresentamos um modelo que se utiliza da localização espacial através do método ITD (*inter-aural time difference*) como fator de adequação do processo ou critério de “*fitness*”.

### 2.1 Síntese Evolutiva de Segmentos Sonoros (SE)

No princípio da Síntese Evolutiva (SE), descrito em [Fornari,01], segmentos sonoros digitais (waveforms) são utilizados como indivíduos de uma população que sofre a ação evolutiva. O grau de adequação de cada segmento sonoro é medido pela distância em relação a um *conjunto alvo* de outros segmentos sonoros o qual representa o fator condicionante do meio.

O incremento temporal desta evolução artificial é dado pela *geração*, equivalente à um ciclo computacional do processo onde a população evolui, gerando novos segmentos sonoros que são condicionados pelo conjunto alvo. Os processos de *seleção* e *reprodução* agem na população a cada geração. A *seleção* compara os indivíduos da população em relação à sua semelhança com os indivíduos do conjunto alvo. Esta semelhança é dada pela função de adequação (i.e. medida de distância) que, a cada geração da população, determina o indivíduo mais próximo ao alvo, o qual é chamado de *melhor indivíduo*. A reprodução gera novos indivíduos na população através de dois operadores genéticos: *crossover* que, a cada geração, substitue todos os indivíduos da população por descendentes que são obtidos por cruzamento com o *melhor indivíduo* da geração atual e mutação que produz variações randômicas nestes indivíduos.

## 2.2 SE e Localização Espacial

Com a inclusão de localização espacial no processo, constatamos que o método de síntese poderia ser expandido conceitualmente. Neste artigo, nós restringimos, por simplicidade, a localização espacial às duas dimensões do plano horizontal em relação aos ouvidos do usuário. Não consideramos a altura da fonte sonora em relação ao ouvinte. Claramente o modelo pode ser estendido com a inclusão de localização 3D. Desta forma, estudamos a utilização de funções de localização sonora como um novo parâmetro do genótipo sonoro. Isso nos levou a considerar a síntese evolutiva não apenas para a transformação dos fatores psicoacústicos, mas também para a manipulação da interpretação do som, descrita por alguns aspectos relacionados à cognição de estímulos sonoros (descritores apresentados em 4.1).

A localização espacial utilizada como função de adequação permite a criação de uma seqüência de amostras sonoras posicionados no espaço que apresentam uma similaridade convergente como característica do próprio processo de síntese. Estes sons com posicionamento espacial acabam, ao longo do tempo, compondo um ambiente sonoro que apesar de variar continuamente é reconhecido por sua similaridade. Deste modo, supomos que este comportamento sonoro da síntese evolutiva poderia ser utilizado para a criação do que chamamos de *paisagens sonoras digitais*.

## 2.3 SE e Paisagem Sonora Digital

Conceitualmente o termo paisagem sonora (i.e. *soundscape*) denota um ambiente complexo no qual a informação sonora nunca se repete e no entanto é cognitivamente auto-similar, uma vez que é facilmente reconhecido pela percepção auditiva humana. Paisagens sonoras podem ser naturais (ex: som de uma floresta) ou não (ex: som de um canteiro de obras em atividade). Segundo [Schafer,77], tais sistemas são caracterizadas por três tipos de referências sonoras: *keynotes* - correspondente ao centro de referência sonora, um ruído de fundo cujo equivalente, em obras tonais, seria a tônica, *signals* - são os sinais sonoros conscientemente percebidos e presentes no ambiente sonoro que alertam e chamam a atenção do ouvinte, e *soundmarks* - são os sinais sonoros únicos ao contexto de um soundscape que caracterizam por sua similaridade e aspectos culturais como enfatiza [Truax,78].

Dentre estes três aspectos fundamentais, percebemos que o resultado sonoro da síntese evolutiva se inseria no paradigma descrito acima especialmente no que tange à “*similaridade variante*” da população de formas de ondas que produzem *soundmarks* que são manipuladas pela informação espacial. A síntese evolutiva que desenvolvemos

passou assim a ter uma aplicação criativa: composição de paisagens sonoras digitais. A implementação em Pd permite que o usuário controle a geração das paisagens, não apenas pela manipulação dos parâmetros de síntese e do conjunto alvo, mas também através de um arquivo de texto, chamado de *score*, que organiza o processo de composição.

### 3. Modelagem

Inicialmente foi necessário o desenvolvimento de uma representação de genótipo sonoro, ou seja, do conjunto de parâmetros que caracterizam amostras sonoras num dado contexto musical, que não apenas permitisse o reconhecimento e discriminação de sua natureza psicoacústica, mas que também fosse computacionalmente viável. Iniciamos com o estudo de três modelos do genótipo sonoro descritos em [Fornari,03]: 1) *genótipo como forma de onda (waveform)*, 2) *genótipo como um conjunto de curvas psicoacústicas*, e 3) *genótipo como espectrograma* sonoro e culminamos por descrever o genótipo através de descritores cognitivos não-contextuais, os quais são descritos a seguir na seção 4.

#### 3.1 Função ITD

Conforme descrito em [Murray,04] a função ITD simula o mecanismo pelo qual a audição percebe a localização de uma fonte sonora pela diferença de tempo de recepção do som entre os ouvidos. Se não houver atraso (se ambos os ouvidos receberem a mesma informação sonora simultaneamente) a localização da fonte é percebida como estando em frente ao ouvinte, caso contrário, percebe-se a fonte sonora deslocada horizontalmente, num dado ângulo de azimute. A função ITD emula esse efeito através da inserção de atraso entre dois canais que contém a mesma informação sonora.

##### 3.1.1 SoundMarks no Campo de Localização Sonora

Sob o ponto de vista matemático, este modelo é constituído de um espaço de triplas  $\mathbf{G}=\{(W, I, L)\}$ , chamado de Espaço Genotípico, onde  $0 \leq I \leq 1$  é o fator de intensidade sonora, e  $-1 \leq L \leq 1$  é o fator de localização espacial (2D) ITD, dado pelo ângulo de azimute  $\theta$ , onde  $L = (90^\circ - \theta) / 90^\circ$  e  $0^\circ \leq \theta \leq 180^\circ$ . Para maiores detalhes, veja [Fornari,07]. O conjunto de todos os possíveis valores do par (I,L) é chamado de “*Campo de Localização Sonora*” (SLF). Em nosso modelo, este é descrito como um semi-círculo e o par (1,0) é associado com o som de maior intensidade e localizado em frente ao ouvinte. A dispersão espacial no SFL é caracterizada pela distribuição dos conjuntos finitos de pares  $S = (I, L)$  conforme mostrado na Figura 1.

Definimos como População um sub-conjunto finito de elementos de  $\mathbf{G}$ . Em nosso modelo iniciamos com uma população  $P^{(0)}$  e conjunto alvo  $\mathbf{T}$  e, interativamente, construímos uma seqüência de um número  $r$  de populações  $G^{(1)}, G^{(2)}, \dots, G^{(r)}$ , onde a  $k$ -ésima população é um sub-conjunto de  $\mathbf{G}$  com  $N$  indivíduos (elementos)  $G^{(k)} = \{G_1^{(k)}, G_2^{(k)}, \dots, G_N^{(k)}\}$  e os indivíduos são dados pelas triplas  $G_i^{(k)} = (W_i^{(k)}, I_i^{(k)}, L_i^{(k)})$ . O conjunto alvo tem  $M$  indivíduos  $\mathbf{T} = \{t_1, t_2, \dots, t_M\}$  cujo  $j$ -ésimo indivíduo é dado por  $t_j = (W_j^{(T)}, I_j^{(T)}, L_j^{(T)})$ . A dispersão espacial na SLF é caracterizada pela distribuição do conjunto de pares  $S_i = (I_i, L_i)$  conforme mostra a Figura 1. Estes pares, onde os Operadores Genéticos são aplicados são chamados de Genótipos de Espacialização Sonora (SSG). O conjunto alvo  $\mathbf{T} = \{T_k = (I_k, L_k), \text{ onde } k=1, \dots, M\}$  pode, em princípio, ser gerado por uma série de controladores gestuais associados a posição e movimento

do usuário. Isto permitiria que impressões perceptuais guiem interativamente o processo evolutivo de distribuição espacial do som.

Uma vez que  $\mathbf{G}^{(k)}$  e  $\mathbf{T}$  são subconjuntos de  $\mathbf{G}$  nós definimos a distância entre estes dois como se segue:

$$d_{ij}(\mathbf{G}^{(k)}, \mathbf{T}) = \frac{1}{2} \frac{|I_i^{(k)} - I_j^{(k)}|}{A} + \frac{1}{4} \frac{|L_i^{(k)} - L_j^{(k)}|}{B} \quad (1)$$

onde as constantes A e B são respectivamente dadas como os fatores máximos de intensidade e localização, e cuja distância é normalizada dentro do intervalo [0,1].

A distância  $d_k$  entre  $\mathbf{G}^{(k)}$  e  $\mathbf{T}$  é definida por:

$$\mathbf{d}_k \equiv d(\mathbf{G}^{(k)}, \mathbf{T}) = \min_{i,j} d_{ij}(\mathbf{G}^{(k)}, \mathbf{T}) \quad (2)$$

onde  $i=1, \dots, N$  e  $j=1, 2, \dots, M$ . Observe que a função de distância usa apenas dois fatores de SSG.

O melhor indivíduo da k-ésima população  $\mathbf{G}^{(k)}$ ,  $\mathbf{G}_{i^*}^{(k)} = (W_{i^*}^{(k)}, I_{i^*}^{(k)}, L_{i^*}^{(k)})$  é aquele que apresenta a distância  $\mathbf{d}_k \equiv d(\mathbf{G}^{(k)}, \mathbf{T})$ . Este novo indivíduo (o melhor da população) pode ser usado pelo processo evolutivo interativo, conforme descrito em 4. Para controlar a saída sonora nós usamos a distância descrita acima para definir a *Similaridade Espacial* como segue:

Dado dois indivíduos da população  $\mathbf{G}_i^{(k)}$ ,  $\mathbf{G}_j^{(k)}$ , eles são similares se  $\mathbf{d}(\mathbf{G}_i^{(k)}, \mathbf{G}_j^{(k)}) \leq \varepsilon$ , onde  $\varepsilon$  é um número arbitrário e a distância  $\mathbf{d}$  é definida pela Eq. (1).

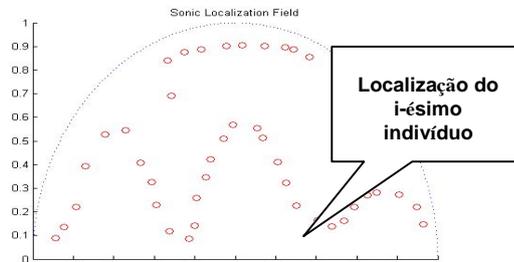


Figure 1. Campo de localização sonoro (SLF).

## 3.2. Operadores Genéticos na Espacialização Sonora

### 3.2.1 Crossover

Dado o melhor indivíduo da k-ésima geração  $\mathbf{G}_{i^*}^{(k)} = (W_{i^*}^{(k)}, I_{i^*}^{(k)}, L_{i^*}^{(k)})$  e a taxa de crossover  $\alpha$ , onde  $0 \leq \alpha \leq 1$ , os parâmetros de localização dos indivíduos da População serão transformados como se segue:

$$\begin{aligned} I_i^{(k+1)} &= \alpha I_{i^*}^{(k)} + (1-\alpha) \cdot I_i^{(k)}, \text{ e} \\ L_i^{(k+1)} &= \alpha \cdot L_{i^*}^{(k)} + (1-\alpha) \cdot L_i^{(k)} \\ \text{para } 1 \leq i \leq N, \text{ e } k &= 0, 1, \dots, R. \end{aligned} \quad (3)$$

onde R é o número de interações.

### 3.2.2 Mutação

Similarmente, dado o melhor indivíduo da  $k$ -ésima geração  $G_{i^*}^{(k)} = (W_{i^*}^{(k)}, I_{i^*}^{(k)}, L_{i^*}^{(k)})$  e a taxa de mutação  $\beta$ , onde  $0 \leq \beta \leq 1$ , a operação de mutação é definida como se segue:

$$\begin{aligned} I_i^{(k+1)} &= \beta_1 \cdot (\text{rand}) + (1-\beta_1) \cdot I_i^{(k)} \quad \text{e} \\ L_i^{(k+1)} &= \beta_2 \cdot (\text{rand}) + (1-\beta_2) \cdot L_i^{(k)} \end{aligned} \quad (4)$$

para  $1 \leq i \leq N$ , e  $k=0,1,\dots, r$ .

onde “rand” é o parâmetro randômico no intervalo  $C$  e as taxas  $\beta_1$  e  $\beta_2$  controlam o grau de aleatoriedade da operação de mutação. Nesta implementação nós consideramos  $\beta_1 = \beta_2$  por simplicidade.

## 4. Implementação

A primeira implementação do SE utilizando a linguagem de programação Pd serviu para salientar os pontos que deviam ser abordados pelo método inicial [Fornari, 01]. Dos diversos aspectos percebidos, viu-se a necessidade de se incorporar uma série de extensões ao método de modo a torná-lo mais semelhante à evolução biológica. Entre eles, destacam-se as seguintes extensões.

- *Conceito de variação demográfica, dado pela implementação de uma população de tamanho variável onde os indivíduos apresentam um determinado tempo de vida.*
- *Conceito de reprodução sexuada, dado pela criação de indivíduos com gênero sexual.*
- *Conceito de localização e mobilidade geográfica de indivíduos através da implementação de genótipos com funções de localização espacial sonora.*

Estas três extensões serviram para re-orientar não só o desenvolvimento do SE, mas também a sua aplicação. Dada a sua crescente proximidade com modelos de evolução biológica, a SE passou a se aproximar da geração de estruturas mais complexas, ou seja, ao invés da síntese de sons, a produção de paisagens sonoras. Percebemos que a grande diversidade de formas de onda e a similaridade entre elas, era uma característica importante do método.

### 4.1 Descritores Percepção Sonora

No desenvolvimento da SE, a primeira e a terceira extensão descritas acima estão diretamente relacionadas à geração de paisagens sonoras. Neste tipo de aplicação o número de fontes sonoras pode variar (população de tamanho variável) bem como as fontes sonoras podem se deslocar no espaço (localização espacial). Apesar de não relacionada diretamente a paisagens sonoras, achamos interessante implementar também a segunda extensão descrita acima: *a reprodução sexuada*. Para isso, criamos o conceito de gênero no indivíduo da síntese evolutiva. Partindo do conceito de reprodução sexuada uma nova extensão do método da síntese evolutiva foi acrescida do conceito de genes diplóides. Assim, em cada genótipo tem-se que cada cromossomo é formado por genes dominantes e/ou recessivos, tal e qual na biologia.

Os genes são dados por descritores cognitivos não-conceituais, tal como apresentados em [Leman,03]. Existem diversas representações de descritores, tal como ataque,

harmonicidade, inharmonicidade, rugosidade, entropia, flutuação, etc. Para essa implementação da SE, utilizamos quatro descritores: *onsetness*, *loudness*, *pitchness* e *location*. Estes descritores foram definidos neste trabalho com base nos estudos de [Leman,04]. Cada um dos descritores são normalizados em escalas entre zero (nulidade ou ausência do descritor em questão) e um (presença máxima do descritor).

Assim definimos o descritor *onsetness* como aquele representando a quantidade de ataque (*onset*) presente no som, sendo que *onsetness* é próximo de zero para sons que não possuem ataque perceptual, tal como é o caso de sons da família das cordas friccionadas (não *pizzicato*) quando este efeito é desejado e pode ser efetivamente conseguido com músicos com bom treinamento. Do mesmo modo, *onsetness* seria próximo de um para sons que sejam praticamente puro ataque, tal como o som de um estalo (sem reverberação), ou um pulso unitário.

O descritor *pitchness* representa o grau de tonalidade de um som. Para sons ruidosos, que não apresentam qualquer definição perceptual de *pitch*, tem-se *pitchness*=0. Para sons com espectro harmônico, como o som de uma nota musical de um instrumento melódico, sem qualquer presença de ruído (como o som de um diapasão) tem-se *pitchness*=1. O descritor *loudness* determina o grau de percepção de intensidade sonora, lembrando-se que este não está unicamente relacionado com a intensidade do sinal sonoro, mas também com o ataque, a frequência do parcial fundamental e a composição espectral do som em questão.

No processo de seleção existem quatro parâmetros correspondentes aos três descritores (definidos acima) e à localização sonora. Estes definem a máxima distância permitida para cada descritor do genótipo do indivíduo em relação ao alvo. Por exemplo, se num dado processo de seleção o parâmetro de distância de um dado descritor for  $dd=0,5$  então será eliminado da população qualquer indivíduo cujo mesmo tipo de descritor  $d$ , comparado aos seus correspondentes no conjunto alvo satisfizer  $|d - da| \geq dd$  a todos os genótipos do conjunto alvo (i.e.  $da$  é um descritor do conjunto alvo).

O processo de reprodução é aqui determinado por quatro fatores. Os dois primeiros são descritos no genótipo do indivíduo: *localização espacial e gênero*. Apesar de não ser uma característica genotípica, por motivos de conveniência computacional a localização do indivíduo é descrita no arquivo texto que representa o genótipo do indivíduo. Este é dado por duas variáveis normalizadas, I (intensidade) e L (ângulo de azimute), dada por  $I=[0,1]$  e  $L=[1,1]$  (respectivamente correspondente aos ângulos de  $0^\circ$  e  $180^\circ$ ). Nesta implementação definimos quatro tipos de gênero dos indivíduos:  $m$  (*macho*),  $f$  (*fêmea*),  $s$  (*estéril*) e  $b$  (*ambos*). As únicas reproduções possíveis são as que obedecem às seguintes regras:

$$m \& f \Rightarrow m | f | b$$

$$[m | f | b] \& b = s \text{ (reprodução com } b \text{ sempre gera indivíduos estéreis).}$$

**onde:**  $\&$  = AND,  $|$  = OR e  $x \& y$  é a reprodução entre  $x$  e  $y$ .

## 4.2 Implementação em Pure Data

A saída de som da SE também foi expandida. Ao invés de ser apenas formada pelo melhor indivíduo de cada geração, esta passou a ser definida pelo usuário através de dois parâmetros:  $m$  (melhores, ou mais próximos do alvo) e  $p$  (piores, ou mais distantes do alvo) seguidos por uma variável  $[0,1]$  referente a porcentagem de indivíduos na

população. Exemplificando,  $m=0.3$  corresponde à saída dada pela mistura de 30% dos indivíduos da população ordenados do melhor ao pior. A implementação inicial foi feita na forma do arquivo Pd, se00.pd (versão 0.0). A melhoria desta implementação inicial resultou no se01.pd (versão 0.1), cuja figura abaixo mostra sua interface gráfica (*front-end*):

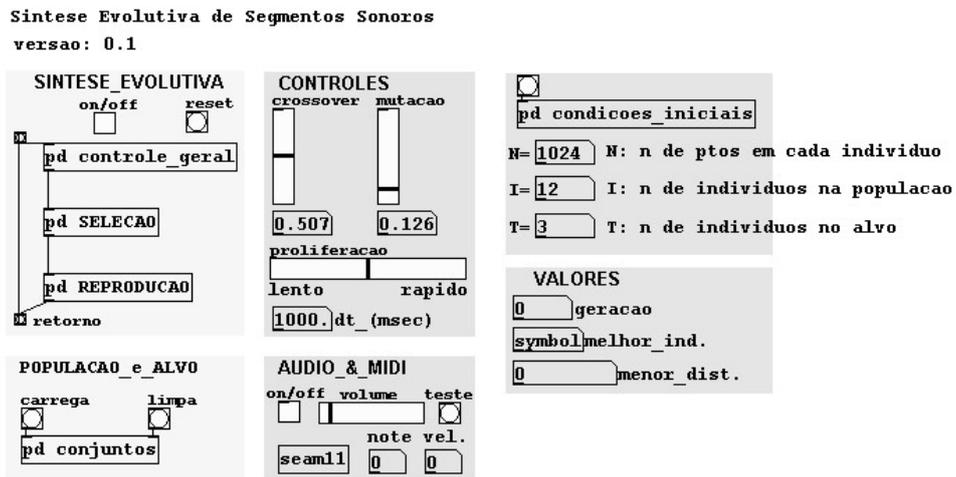


Figura 2. Primeira implementação do SE, o se01.pd

Ao iniciar a síntese, o *se01.pd* vai selecionar o indivíduo da população que é mais próximo dos indivíduos do conjunto alvo. Este será o primeiro melhor indivíduo e em seguida o processo de reprodução modifica todos os 12 elementos do conjunto população através das operações genéticas entre cada indivíduo da população e o melhor indivíduo escolhido anteriormente.

#### 4.2.1 Arquivo de Controle Paramétrico

Na fase final da implementação em Pd, o método da síntese evolutiva foi remodelado com base nos novos critérios apresentados na seção anterior e o sistema foi assim expandido para a geração de paisagens sonoras digitais. Obteve-se um sistema mais complexo, que inclusive utiliza-se de um arquivo texto, que chamamos de *score*. Este arquivo contém instruções ordenadas no tempo que organizam o processo de geração de paisagens sonoras. Além do arquivo *score*, a SE é também controlada pelos parâmetros contínuos das taxas dos operadores genéticos (crossover e mutação), a taxa de proliferação populacional, bem como pela modificação dinâmica dos indivíduos do conjunto alvo pelo usuário em tempo real. No modelo de *score*, tem-se toda a informação inicial e as linhas de comando são organizadas de forma sequencial. Na figura abaixo tem-se um exemplo típico do arquivo *score*:

Table 1 – *Score* utilizado para controle da Síntese Evolutiva

Title	<i>Title</i> : o nome do <i>score</i> .
ParteTeste;	<i>instructions</i> : instruções da organização deste <i>score</i> que irão aparecer ao usuário, sob a forma e comentário, na janela do programa.
instructions	<i>num-pop</i> : número inicial de indivíduos na população, que deve estar em conformidade com o número de WAV files no arquivo <i>pop</i> .
Esta eh a sintese evolutiva de \\	<i>num-alvo</i> : número inicial de indivíduos no conjunto alvo que, igualmente, deve estar em conformidade com o número de WAV files no arquivo <i>alvo</i> .
soundscapes com score ParteTeste\\	

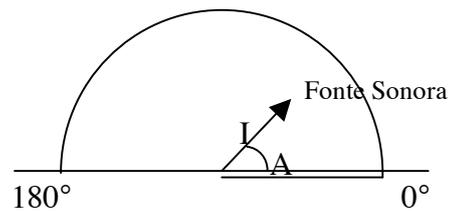
<pre> \ Clique SPACE para prosseguir\ \ modos do score: \ time selec onset loud pitch location \ time repro gender proximity\ output proxim x distant y all\ \ global: selecao\ step: reproducao\ state: output\ ; num-pop 7; num-alvo 3; proli .5; cros .5; mut .1; 00 sel .5 .5 .5 .5; 00 rep .5; 00 out p 1; 01 out p 5; 05 out l 1; </pre>	<p><i>proli</i>: valor inicial de proliferação da reprodução de indivíduos.</p> <p><i>cros</i>: valor inicial da taxa de crossover.</p> <p><i>mut</i>: valor inicial da taxa de mutação.</p> <p>Logo em seguida, tem-se as linhas de comando, ou <i>command-lines</i>, dada em uma única linha encerrada por ";". Estas são ordenadas com as seguintes instruções:</p> <p><i>Time</i>: tempo de atraso para a execução da linha, dado em segundos.</p> <p><i>Mode</i>: determina o modo de operação: <i>sel</i>/processo de seleção, <i>rep</i> reprodução e <i>out</i> saída sonora.</p>
--	---

O genótipo de cada segmento sonoro da população é também dado por um arquivo de texto. Seguindo o padrão do objeto “*textfile*” do Pd, onde cada linha é encerrada pelo símbolo “;”. A primeira linha do arquivo de texto do genótipo especifica a localização espacial do indivíduo (*location*). Esta é dada através de dois parâmetros, Intensidade, que varia entre  $I = [0,1]$ , e ângulo de azimute que varia entre  $A = [-1,1]$  respectivamente correspondendo à uma variação de ângulo entre  $[-180^\circ, 0^\circ]$ . A segunda linha arquivo de texto do genótipo determina o “tempo de vida” (*lifespam*) deste indivíduo, ao longo do processo evolutivo. Este tempo é dado em segundos. Após terminado este tempo pré-estabelecido, o indivíduo é automaticamente eliminado da população. Para representar indivíduos “imortais” usa-se “zero” como parâmetro (conforme é mostrado no exemplo da figura 3).

A terceira linha determina o gênero do indivíduo (*gender*), podendo este ser: *m* (*masculino*), *f* (*feminino*), *b* (*ambos*), *s* (*estéril*). Conforme foi mostrado, existem duas regras que determinam a reprodução anteriormente, sendo que se o indivíduo for *s* não pode reproduzir e permanecerá na população até terminar seu tempo de vida ou ser eliminado por distanciamento do conjunto alvo.

As próximas linhas do genótipo determinam os três descritores cognitivos escolhidos para representar o genótipo em si. Estes formam o cromossomo diplóide do indivíduo onde existe o conceito de dominância e recessividade. Então, em cada linha tem-se o valor para o gene dominante dado pelo parâmetro *d*, seguido de seu valor numérico (entre  $[0,1]$ ) para o gene recessivo, dado por *r*, seguido de seu valor normalizado. O conceito de hereditariedade e caracterização fenotípica seguem as simples regras dispostas pela genética de *Gregor Mendel*. Um exemplo típico é dado abaixo:

```
location l -1;
lifespan 0;
gender m;
onsetness d .5 r .5;
loudness d .1 r .9;
pitchness d .9 r .1;
```



**Figura 3. Esquerda: Exemplo de um genótipo da SE, representado por um arquivo de texto. Direita: Localização do indivíduo da SE.**

#### 4.2.2 Interface Gráfica

O *front-end* da implementação da SE é visto na figura 4. Nesta janela inicial o usuário é convidado a digitar o nome de um *score* que guiará o processo evolutivo. Ao digitar o nome do *score* uma nova janela aparece com este novo nome no título e com outros parâmetros evolutivos. Na figura 4, centro, é mostrada um exemplo desta janela. Caso não haja um *score* definido, a próxima janela é similar àquela localizada à esquerda da figura 4.

Como se percebe, não é necessária a utilização de um *score* na síntese evolutiva. Este funciona como uma partitura musical que descreve as passagens de uma composição musical. Analogamente, a não utilização de um *score* assemelha-se à improvisação musical, pois deixa a cargo do usuário o controle em tempo real de todos os parâmetros do processo de síntese.

### 5. Discussão

Do ponto de vista qualitativo sonoro, estamos agora desenvolvendo projetos composicionais onde poderemos verificar o potencial musical desta nova aplicação. Todavia, a nova implementação da SE está disponível em: <http://www.nics.unicamp.br/~fornari/se>. Esperamos com isso que uma grande quantidade de pesquisadores da área avaliem esta implementação qualitativamente, gerem amostras sonoras e emitam seus pareceres. É nosso intento que tenhamos conseguido chegar a uma ferramenta musical interessante para o compositor de música eletroacústica. Coerente com o método de implementação, buscamos agora uma estratégia populacional para testar a variedade e aplicabilidade do método de síntese que desenvolvemos.



Figura 4. (esquerda) Janela front end da interface do programa. (centro) Janela com o nome do *score* e os respectivos controles para a síntese evolutiva. (direita) Janela da síntese evolutiva feita sem a presença de um *score*.

## 6. Conclusão

Apresentamos a implementação do sintetizador evolutivo em Pd, com processamento e controle em tempo-real. Esta implementação abre as portas para um novo conceito de processamento não-determinístico de ambientes sonoros, ou paisagens sonoras digitais, uma vez que o método gera sons que se modificam adaptativamente, convergindo com similaridade variante.

A interface gráfica (vide figura 4) da SE também apresentou grande melhoria em relação à primeira versão implementada [Fornari,01]. No entanto, o Pd não oferece grandes recursos gráficos em sua plataforma atual. Teoricamente, é possível utilizar apenas o *engine* do Pd processando em *background* controlado por uma plataforma gráfica mais desenvolvida, por exemplo, em JAVA. Isto tornaria a visualização do programa mais aprazível e protegeria o código do algoritmo.

Do ponto de vista de aplicações musicais, estamos interessados em utilizar o método com dispositivos de localização espacial para obtermos um processo interativo em tempo real onde o conjunto alvo poderia ser substituído por uma *buffer* dinâmico. A segunda aplicação, é justamente voltada à composição de paisagens sonoras utilizando uma grande variedade de segmentos sonoros, onde poderíamos testar, então, a capacidade do sistema como ferramenta composicional.

## Agradecimentos

José Fornari teve o apoio da FAPESP através de bolsa de pós-doutoramento desenvolvida no NICS, processo 04/00499-6R. Jônatas Manzolli tem o apoio do CNPq através de bolsa de produtividade em pesquisa.

## Referências

Koza, John R., "Genetic Programming", Encyclopedia of Computer Science and Technology, 1997.

- Bäck, T. "Evolutionary Algorithms in Theory and Practice: Evolution Strategies, Evolutionary Programming, Genetic Algorithms", Oxford Univ. Press. 1996.
- Miranda, E. R.. "On the evolution of music in a society of self-taught digital creatures." Digital Creativity, Vol. 14, No. 1, 29-42. 2003.
- Soddu C. O. "A Generative Approach to Art and Design" Leonardo, 1 June 2002, vol. 35, no. 3, pp. 291-294. MIT Press. 2002.
- Garcia, Ricardo A. "Automatic Generation of Sound Synthesis Techniques". Proposal for degree of Master of Science. MIT - Fall. 2000.
- Gibbs, J. "Easy Inverse Kinematics using Genetic Programming". Genetic Programming 1996: Proceedings of the First Annual Conference. MIT Press. <http://www.red3d.com/cwr/evolve.html>. 1996.
- Biles, J. A., "Gen Jam: A Genetic Algorithm for Generating Jazz Solos", Proceedings of the 1994 International Computer Music Conference, (ICMC'94), 131-137. <http://www.it.rit.edu/~jab/GenJam.html>. 1994.
- Horner, A., Beauchamp, J., and Haken, L., "Machine Tongues.16. Genetic Algorithms and Their Application to FM Matching Synthesis," Computer Music Journal, vol. 17, pp. 17-29. 1993.
- Moroni, A., Manzolli, J., Von Zuben, F., Gudwin, R., "Vox Populi: An Interactive Evolutionary System for Algorithmic Music Composition", Leonardo Music Journal, San Francisco, USA, MIT Press, Vol. 10. 2000.
- Fornari, J., Manzolli, J., Maia Jr., A., Damiani F., "The Evolutionary Sound Synthesis Method". SCI conference. Orlando, USA. 2001.
- Caetano, M. F., Manzolli, J., Von Zuben, F. J. "Application of an Artificial Immune System in a Compositional Timbre Design Technique". In: Jacob, C., Pilat, M.L., Bentley, P.J., Timmis, J.. (Org.). Artificial Immune Systems, Lecture Notes in Computer Science. Berlin: Springer-Verlag, v. 3627, p. 389-403. 2005.
- Fornari, J., Manzolli, J., Maia Jr., A., Damiani F., "The Evolutionary Sound Synthesis Method". SCI conference. Orlando, USA. 2001.
- Schafer, R. M. (1977). "The Soundscape". ISBN 0-89281-455-1.
- Truax, B.. (1978) "Handbook for Acoustic Ecology". ISBN 0-88985-011-9.
- Fornari, José Eduardo. "Síntese Evolutiva de Segmentos Sonoros". Dissertação de Doutorado. DSIF/FEEC/UNICAMP. 2003.
- Leman, M., Vermeulen, V., De Voogdt, L., Moelants, D., & Lesaffre, M. (2004). Correlation of Gestural Musical Audio Cues. *Gesture-Based Communication in Human-Computer Interaction:5th International Gesture Workshop, GW 2003*, 40-54.
- Murray, J. C., Erwin, H. R., and S. Wermter, "Robotic sound source localization using interaural time difference and cross-correlation", presented at KI-2004, September 2004.
- Fornari, J.; Maia Jr. A.; Manzolli, J.. "Generating Soundscapes Using Evolutionary Spatial Control". In Proceedings of EvoWorkshops, EvoMusarts, Springer-Verlag: Berlin. pg 517-526, Valencia. 2007.