

# Um Ambiente Computacional Interativo e Imersivo para a Síntese Evolutiva de Cantos de Pássaros

José Fornari<sup>1</sup>, Francisco Barretto<sup>2</sup>

<sup>1</sup> Núcleo Interdisciplinar de Comunicação Sonora (NICS) – Rua da Reitoria, 165 - Cidade Universitária "Zeferino Vaz" – CEP: 13.083-872 – fone/fax: +55 (19) 3521 7923 – Universidade Estadual de Campinas (UNICAMP) – Campinas – SP – Brazil

<sup>2</sup> Laboratório de Pesquisa em Arte Computacional (Midialab). Departamento de Artes Visuais. Instituto de Artes. Universidade de Brasília (UnB). Brasília – DF – Brasil

{tutifornari,kikobaretto}@gmail.com

**Abstract.** *Birdsongs are a ubiquitous sound component in several natural soundscapes. They can be seen as agents, in an ecological network, where their interaction self-organizes a sonic texture, with similar cognitive features and distinct acoustic aspects. This article presents a computer model for the generation of artificial soundscapes of birdsongs. Such synthetic environment is created by an Evolutionary Algorithm, the ESSynth model, that continuously generates control parameters for models of syrinx sound synthesis, which is capable of emulating a broad variety of birdsongs. The result is an interactive soundscape coupled with an interactive computer animation, both influenced by the instant text messages transmitted from the micro-blog Twitter.*

**Resumo.** *Os cantos de pássaros são parte integrante de diversas paisagens sonoras naturais. Estes podem ser vistos como agentes de uma rede ecológica cuja interação auto-organiza um sistema sonoro de características cognitivas similares e aspectos acústicos dinamicamente originais. Este artigo apresenta um sistema computacional desenvolvido para a geração de paisagens sonoras artificiais de cantos de pássaros. Tal ambiente artificial é criado por um algoritmo evolutivo, o ESSynth, que gera continuamente parâmetros de controle para modelos computacionais de siringe, capazes de emular uma ampla variedade de cantos de pássaros. O resultado é uma paisagem sonora e animação digital interativa; ambas dinamicamente influenciadas por mensagens instantâneas (tweets) do microblog: Twitter.*

## 1. Introdução

É espantosa a quantidade e variedade de ambientes onde é possível escutar o canto de pássaros. Estes trocam informações sonoras com outros indivíduos, através de seus característicos cantos, formando sua rede de comunicação social que constitui para nós uma paisagem sonora [Schafer 1994]. Tal textura de sons orgânicos, dinâmicos e similares, muitas vezes, se mescla com os ruídos urbanos de máquinas e pessoas, criando um ambiente imersivo cibernético [Wiener 1968], composto da interação entre agentes orgânicos e maquinais, criando um complexo sistema sonoro, emergente, acusticamente auto-organizado e cognitivamente auto-similar.

A estrutura fisiológica, que permite aos pássaros gerar cantos tão diversos e complexos, é de sofisticação ímpar. Existem modelos computacionais já desenvolvidos de tal aparato orgânico, em essência, a siringe [Clarke 2004]. Porém, os modelos computacionais da siringe possuem um grande número de parâmetros de controle, o que, muitas vezes, torna a sua exploração manual dificultosa e até mesmo inarticulável.

Controlar um grande número de parâmetros simultaneamente é uma das tarefas de extrema complexidade que o cérebro realiza com larga facilidade. Um exemplo disto é o controle da posição e deslocamento corporal, e consequentemente, da expressão gestual, que é realizado pela maioria das pessoas praticamente de modo automático e até mesmo inconsciente. Uma alternativa de simular tal controle é através de modelos computacionais inspirados nas estratégias naturais e biológicas da resolução de tais problemas. Este é o ramo tratado pela Inteligência Artificial, inspirados em dois processos naturais de resolução de problemas complexos: o Cérebro, de onde advêm metodologias tais como as redes neurais; e a Evolução Natural, que trata do desenvolvimento não-supervisionado de sistemas complexos (tais como o próprio cérebro) e que motivou a criação de Algoritmos Evolutivos (AE) [Eiben 2007]; onde está baseada a abordagem deste trabalho. Para tanto, desenvolvemos um modelo computacional de AE que permite não apenas o controle de um modelo de canto de pássaros, mas da simulação de um pequeno ecossistema destes, formado por um conjunto População, com tamanho variável de indivíduos, que são modelos computacionais de siringes. Cada indivíduo é o instanciamento de uma sub-rotina em PD ([www.puredata.info](http://www.puredata.info)) de síntese sonora procedural, que realiza o modelamento físico do processo acústico onde o pássaro gera e modula o som do seu canto, ao invés de utilizar amostras de áudio de sons gravados dos pássaros. Tais indivíduos têm seu próprio genótipo, constituído por um arquivo de texto contendo os parâmetros de cada síntese sonora. Dentro da População, os indivíduos nascem, reproduzem e morrem, compondo assim uma paisagem sonora evolutiva, dinâmica, cognitivamente similar e acusticamente original.

Para tornar este sistema aberto, foi estabelecido um segundo grau de interatividade. Além de permitir a livre evolução da população de indivíduos, este sistema também permite a interveniência externa de múltiplos usuários, que podem inserir dinamicamente novo material genético na População evolutiva, conforme explicado adiante. Para viabilizar a interatividade simultânea de diversos usuários com o sistema evolutivo, utilizamos o famoso serviço de *microblog*, conhecido por *Twitter* ([www.twitter.com](http://www.twitter.com)) [Barretto 2010]. *Twitter* é uma palavra da língua inglesa que significa “silvar”, ou seja, a sucessão de pequenos sons emitidos pelos pássaros canoros. A metáfora aferida por tal nome, que compara o canto de pássaros com a comunicação de pequenos textos numa rede social, encaixa-se perfeitamente na contextualização deste sistema, posto que descreve com precisão a filosofia por detrás deste microblog. Os silvos dos pássaros inicialmente podem soar como algo aparentemente desprovido de significado ou ordem. Porém, o seu contexto é dinamicamente inferido pela cooperação entre as aves, que os emitem e os escutam. O mesmo se aplica ao Twitter, onde as inúmeras mensagens, vistas isoladamente, podem parecer completamente aleatórias e/ou desprovidas de sentido, porém dependem ativamente do seu receptor para estabelecer um significado global que as unifique em um único contexto [Dorsey 2009].

Utilizou-se aqui a metáfora dos pássaros como base fundamental para o desenvolvimento deste trabalho. Neste trabalho, o Twitter fornece o canal de inserção de novo material genético para a síntese evolutiva do AE. Cada modelo de canto de pássaros é visto como um indivíduo na população do AE. Cada modelo de siringe utiliza 16 parâmetros de controle. Tais parâmetros correspondem ao gene do DNA. Por simplicidade, optamos por utilizar o código ASCII dos caracteres das mensagens do Twitter, como o DNA dos indivíduos da População. O sistema AE é implementado em PD. Já o cliente Twitter foi implementado em Java. Ao receber uma nova mensagem, este a transforma num vetor de números inteiros, correspondendo aos caracteres ASCII, normalizado entre 0 e 1, e representando o DNA utilizado pela síntese evolutiva.

A cada instante que uma nova mensagem é recebida, um novo indivíduo, contendo este respectivo DNA é inserido na população evolutiva. Este indivíduo corresponde a um novo objeto sonoro, ou seja, um novo canto de pássaro. Cada objeto sonoro tem um

período de vida limitado, onde pode reproduzir, ser selecionado e, por fim, morrer. A resultante sonora de todos os indivíduos ativos (vivos) a cada momento da evolução da População, conforme explicado adiante, se auto-organiza na forma de uma Paisagem Sonora (PS). Em paralelo, um outro modelo computacional, desenvolvido em *Processing* ([www.processing.org](http://www.processing.org)) [Reas 2007] foi implementado, para a geração de uma Paisagem Visual (PV) em correspondência à PS. Esta é controlada pelos mesmos genótipos, ou seja, parâmetros dos modelos de síntese da siringe, que são compartilhados através do protocolo OSC ([www.opensoundcontrol.org](http://www.opensoundcontrol.org)) e complementam a experiência imersiva de criação dinâmica e interativa da PS, coletivamente composta e interativamente experimentada. Diversos algoritmos computacionais do modelamento físico da siringe, foram desenvolvidos. Aqui, nos baseamos no modelo computacional de Hans Mikelson, inicialmente desenvolvido para o ambiente de programação *Csound* [Mikelson 2000]. Este algoritmo foi posteriormente aperfeiçoado e implementado em PD por Andy Farnell, que criou um modelo computacional que simula a siringe e também o processamento sonoro dado pela traquéia e pelo bico [Farnell 2010].

## 2. Discussão e Conclusões

Este trabalho apresentou os fundamentos de criação de um sistema computacional para a criação de paisagens sonoras através do algoritmo evolutivo ESSynth [Fornari 2008], que se estende do campo sonoro para o visual, através da implementação de um ambiente de simulação gráfica dos agentes formadores da paisagem sonora, e sua representação sincrônica num ambiente computacional de paisagem visual. A paisagem sonora criada por tal sistema é inspirada na rede social da natureza, gerada pelos pássaros, através de seus cantos. Uma das características marcantes do ESSynth é criar uma textura sonora similar e variante. Isto é dado por um conjunto de objetos sonoros que evoluem ao longo do tempo, através dos processos computacionais que simulam a reprodução e a seleção, conforme observadas na natureza, em populações biológicas. O som resultante de todos os objetos sonoros, é cognitivamente similar, ao mesmo tempo que é acusticamente distinto. Pelo que consideramos, esta é uma das características fundamentais das paisagens sonoras naturais. Assim, o ESSynth possibilita gerar paisagens sonoras artificiais. Os objetos sonoros são os indivíduos pertencentes ao conjunto de tamanho variável; a População evolutiva. Os indivíduos têm um código genético que parametriza o processo de síntese sonora do modelamento físico da siringe. Este genótipo é modificado pelo processo de reprodução e medido pelo processo de seleção. Assim, estes indivíduos reproduzem em pares, gerando novos indivíduos, com características genóticas similares a de seus predecessores. Estes são selecionados com base na similaridade de seus genótipos, com o da média da população. Os indivíduos têm um tempo de vida limitado. Quando um indivíduo morre, seu genótipo é eliminado da população, e nunca mais se repete. Assim, não existem clones na população evolutiva do ESSynth.

Neste trabalho, os indivíduos são instanciamentos do modelo de síntese sonora procedural de cantos de pássaros. Utilizamos a adaptação em PD do modelo inicialmente criado por Hans Mikelson, que simula a siringe. O modelo foi estendido por Andy Farnell, para incorporar a geração de frases melódicas aleatórias, nos moldes que observamos os pássaros cantarem na natureza. Cada indivíduo da População evolutiva de ESSynth é o instanciamento destes modelos físicos do canto de pássaros. Os genótipos são dados por vetores de parâmetros destes modelos. Estes advêm de mensagens (*tweets*) do Twitter, o que permite que a População de objetos sonoros de ESSynth passe de um sistema fechado para aberto, caracterizando assim um CAS (*Complex Adaptive Systems*). Este é um sistema complexo e aberto, que apresenta características auto-similares, formado por agentes (indivíduos) independentes, mas que interagem entre si. O CAS apresenta propriedades emergentes, ou auto-organizadas, como são os sistemas naturais, criados pela evolução natural [Holland 2006].

Deste modo, o sistema aqui apresentado permite a interatividade de múltiplos usuários externos, que enviam mensagens para o sistema evolutivo, através do Twitter. Para criar a realimentação entre o usuário e o sistema, foi desenvolvido um modelo de animação gráfica dos indivíduos. Estes podem ser vistos pelos usuários que enviaram mensagens, na forma da metáfora visual de uma animação lembrando o vôo de um pássaro, cujas asas são formadas pela mensagem enviada. Assim, o usuário pode identificar a sua inserção no *pool* genético da população evolutiva. Com isso, temos duas camadas de interatividade sistêmica; a interna e a externa. A interna, se dá pelos indivíduos do ESSynth, que compõem a paisagem sonora, formada pela síntese dos diversos cantos de pássaros, similares e variantes, e engendradas pelos processos evolutivos de reprodução e seleção. A externa, é formada pelos *tweets* dos múltiplos usuários que alteram a composição genética da população e (eventualmente) podem identificar visualmente os indivíduos que criaram e os correlacionar com seus objetos sonoros correspondentes. Ambas esferas de interação são conectadas pela metáfora do microblog Twitter; cujo desenvolvimento também se baseou na rede social formada através dos cantos dos pássaros, que se auto-organizam em paisagens sonoras. Alguns vídeos e exemplos sonoros deste sistema evolutivo em funcionamento estão disponibilizados no seguinte link: <http://sites.google.com/site/nicsunicamp/sbcm2011>

## 9. Referências

- Barretto, F. P. (2010) "Arte Inteligente: Abordagens para o Desenvolvimento de Sistemas Criativos." In Anais do 9º Encontro Internacional de Arte e Tecnologia, edited by Suzete Venturelli, 171-180. Brasília.
- Clarke, Julia A. (2004) "Morphology, Phylogenetic Taxonomy, and Systematics of Ichthyornis and Apatornis (Avialae: Ornithurae)". Bulletin of the American Museum of Natural History 286: 1–179.
- Dorsey, Jack. (2009) "Twitter creator Jack Dorsey illuminates the site's founding document." L.A. Times. David Sarno. February 18, 2009. <http://latimesblogs.latimes.com/technology/2009/02/twitter-creator.html> (accessed May 17, 2010).
- Eiben, A.E., Smith, J. E., (2007) "Introduction to Evolutionary Computing". 2nd Ed., Springer, Natural Computing Series.
- Farnell, A. (2010) "Designing Sound". MIT Press, Cambridge, Massachusetts, London, England.
- Fornari, J., Maia, A., Manzolli J., (2008) "Soundscape Design Through Evolutionary Engines". Journal of the Brazilian Computer Society, 2008, Volume 14, Number 3, Pages 51-64
- Holland, J. (2006) "Studying Complex Adaptive Systems." Journal of Systems Science and Complexity 19 (1): 1-8.
- Mikelson, Hans. (2000) Bird calls. Csound Magazine, Winter, 2000.
- Reas, C., Fry, B. (2007) "Processing: a programming handbook for visual designers and artists". 1st Edition. London: The MIT Press.
- Schafer, M., R. (1994) "The soundscape: our sonic environment and the soundscape". Destiny Books
- Schmeder, A., A. Freed, and D. Wessel. (2010) "Best Practices for Open Sound Control." Linux Audio Conference '10.
- Wiener, N. (1968) "Cibernética e sociedade: o uso humano de seres humanos". São Paulo: Cultrix.