

## **A escritura de uma aquarela: orquestração para técnicas estendidas com suporte de descritores de áudio**

### COMUNICAÇÃO ORAL: COMPOSIÇÃO

*Ivan Eiji Simurra*

*Instituto de Artes – IA/UNICAMP, NICS/UNICAMP e Mus3/UFPB- ieysimurra@gmail.com*

*Jônatas Manzolli*

*Instituto de Artes – IA/UNICAMP e NICS/UNICAMP- jotamanzo@gmail.com*

**Resumo:** Apresentamos a aplicação de descritores de áudio para planejamento orquestral e de escritura. Utilizamos os descritores denominados como Centróide Espectral, Desvio Padrão Espectral e Chroma, para produzir pontos de referência estruturantes à obra "*Lana Tai – no dia em que nasceu uma aquarela*", para orquestra de cordas. Implementamos um ambiente computacional em Pure Data (PD), o qual possibilitou estudar configurações orquestrais, denominadas *Marcos Sonoros* (MS). Os MS foram construídos com amostras de um banco sonoro digital e se relacionaram com os pontos de referência.

**Palavras-chave:** Composição Musical. Composição Assistida por Computador. Orquestração Assistida por Computador. Descritores de Áudio.

#### **The watercolor composition: Music Orchestration for Extended Techniques with the support of audio features.**

**Abstract:** It is presented an audio features usage towards music composition and orchestral planning. We worked on this path to create the string orchestra piece "*Lana Tai – no dia em que nasceu uma aquarela*". We applied the following features: Spectral Centroid, Spectral Deviation and Chroma. A virtual environment in Pure Data (PD) software enabled to analyze some potential orchestral settings called "Sound Marks". They are created through a sound database that was related with the structural compositional landmarks in "*Lana - Tai*".

**Keywords:** Music Composition. Computer-Aided-Composition. Computer-Aided-Orchestration. Audio Features.

### **1. Introdução**

Durante o seu planejamento composicional, o compositor dispõe de um sem número de recursos sonoros para construir a sua obra musical, seja para instrumento solo ou para uma orquestra sinfônica. No contexto deste artigo, esses recursos são provenientes dos diversos e particulares modos de execução de cada instrumento musical aos quais alteram a sonoridade característica de cada nota musical. Sinestesticamente, tal como numa pintura em aquarela, o "colorido sonoro" é controlado pela forma como o compositor concatena essas técnicas instrumentais, podendo criar "manchas sonoras" que recriam o timbre associado a um instrumento ou a uma combinação deles. Essa ideia de aquarela, deu origem ao nome da obra cujo processo composicional é aqui reportado; "*Lana Tai – no dia em que nasceu uma aquarela*", para orquestra de cordas.

O desenvolvimento da composição musical está vinculado à utilização de suportes tecnológicos, especialmente para a música produzida a partir do século XX. No âmbito desse trabalho, os “suportes tecnológicos” foram utilizados como ferramentas de assistência ao processo criativo, da mesma forma que um compositor poderia utilizar a sua voz, um instrumento musical ou um recurso de outra área do conhecimento. Em particular, a utilização de modelagem matemática em processos de orquestração reporta-nos à pesquisa realizada em (MANZOLLI, 1988). Os recentes métodos de análise e composição com suporte computacional, possibilitam a extração de vários parâmetros ou curvas que descrevem características sonoras a partir do espectro de frequências. Muitos desses modelos computacionais focam a “recuperação da informação musical”, área de estudo que é denominada na literatura como “Music Information Retrieval” ou simplesmente MIR (BYRD; FINGERHUT, 2009). Essa área utiliza recursos de processamento de sinais para realizar a extração de características sonoras a partir de medidas estatísticas aplicadas sobre o espectro do sinal digitalizado. Essas medidas são denominadas de “descritores de áudio”. Apresentamos aqui um estudo sobre a sua utilização no planejamento e nos procedimentos de escritura e orquestração quando o compositor faz uso de técnicas instrumentais estendidas (PADOVANI & FERRAZ, 2011).

Uma importante referência sobre a utilização de descritores de áudio voltada à caracterização da sonoridade de técnicas estendidas instrumentais é a biblioteca Zsa.Descriptors desenvolvida para o ambiente MAX/MSP (MALT & JOURDAN, 2009). Uma das principais fontes de material bibliográfico acerca de descritores de áudio são os Anais dos Congressos da Sociedade Internacional para Recuperação da Informação Musical (ISMIR<sup>1</sup>) e o repositório na Internet denominado de MIREX. Para desenvolver os procedimentos composicionais de “*Lana Tai*”, utilizamos a biblioteca PDescriptors, desenvolvida por (MONTEIRO, 2012) no NICS/UNICAMP, a partir de abstrações do ambiente computacional Pure Data<sup>2</sup>.

## **2. Objetivo do Projeto Composicional**

“*Lana Tai*” foi elaborada a partir de um diálogo com suportes tecnológicos advindos, principalmente, dos descritores de áudio. O ponto de partida vinculou-se a dois universos: a) as técnicas de execução estendidas e b) as ferramentas computacionais para avaliar o conteúdo espectral dos sons gerados por essas técnicas. Desenvolveu-se então um método para auxiliar o compositor a relacionar: a) descrições de alto nível ou simbólicas como as noções de “sonoridade” com b) os modos específicos de execução instrumental

estendida. Utilizou-se diversas possibilidades de análise via descritores de áudio, tais como: “Centróide Espectral”, “Desvio Padrão Espectral” e “Chroma” ou Perfis de Classe de Alturas da escala cromática. Apresentamos, a seguir, uma breve definição dos descritores usados e, em seguida, descrevemos o processo criativo utilizado na confecção de “*Lana Tai*”.

### 3. Descritores de Áudio

A referência que utilizamos nesta seção foi (PEREIRA, 2009) que apresenta um conjunto de definições de descritores de áudio. Os descritores de áudio partem das análises do conteúdo espectral de sinais de áudio digitalizados as quais são realizadas via *Short-Time Fourier Transform* (STFT), que é definida a seguir (SHEH; ELLIS, 2003):

$$X_{STFT} = [k, n] = \sum_{m=0}^{N-1} x[n-m] * w[m] * e^{-j2\pi km/N} \quad (1)$$

onde,  $k$  indexa o eixo de frequência (*bin*) dentro do intervalo  $0 < k < N - 1$  do  $n^{\text{ésimo}}$  quadro;  $w[m]$  é uma janela de  $N$  amostras.

#### 3.1 Centróide e Desvio Padrão Espectral

Uma das medidas estatísticas mais utilizadas na extração de características do áudio é denominada *Centróide Espectral* (THEODORIDIS & KOUTRUMBAS, 2006). Esse descritor pode ser descrito como o centro de massa ou baricentro do espectro analisado. A literatura associa o Centróide com a percepção de “brilho” do sinal sonoro analisado: quanto maior o valor do centróide, dado em Hz, maior será a concentração de energia nas componentes espectrais de frequências altas, i.e. esse fator estaria associado à percepção de um som mais “brilhante”. O Centróide é definido pela seguinte equação:

$$SC_{(i)} = \frac{\sum_{k=1}^{K-1} f(k) * |X_i(k)|^2}{\sum_{k=1}^{K-1} |X_i(k)|^2} \quad (2)$$

as definições dos parâmetros da equação acima encontram-se em (PEREIRA, 2009, p. 16)

$$f[k] = \frac{\text{taxa\_de\_amostragem}}{2K} k.$$

O Desvio Padrão caracteriza a variação em torno do Centróide da distribuição de energia dos pontos analisados na Eq. 1.

$$SP_i = \sqrt{\frac{\sum_{k=1}^{K-1} \{f(k) - SC_i\}^2 |X_i[k]|^2}{\sum_{k=1}^{K-1} |X_i[k]|^2}} \quad (3)$$

onde  $SC$  é o Centróide Espectral definido anteriormente (PEREIRA, 2009, p. 16).

### 3.2 Chroma ou Perfis de Classe de Alturas Musicais

O modelo perceptivo do Chroma baseia-se em (SHEPARD, 1982) que aponta duas características importantes para reconhecer *hierarquias de alturas*: a) a configuração rotacional denominada de Chroma ou “*Pitch Class*”, ou seja, a equivalência das classes de alturas musicais entre oitavas (Figura 1a); b) a posição vertical do índice do Chroma, o qual Shepard chama de “*height*” ou registro Figura 1b. Por exemplo, na notação “C<sub>3</sub>” (dó três), o letra “C” caracteriza a classe de altura, e o índice “3” o registro.

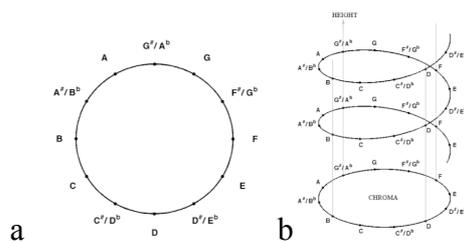


Figura 1: Representação do comportamento rotacional (a) e espiral das alturas em oitavas diferentes (b).

A partir de FUJISHIMA (1999), (Gomez, 2006) desenvolveu o *Harmonic Pitch Class Profile* (HPCP) com as seguintes características: a) ponderação  $\log_2$  para caracterizar a propriedade logarítmica da percepção de alturas musicais; b) ênfase na contribuição dos componentes harmônicos de cada classe de altura; c) melhoria da resolução da distribuição da energia entre os diferentes registros. É possível configurar o HPCP para várias sub-divisões de bandas de frequência tais como semitons, quartos de tom e sextos de tom. Para subdivisão caracterizada por 12 semitons, cria-se uma relação entre os *bins*  $k$  da STFT com os *bins*  $p$  (vide equação 1) de acordo com:

$$P[k] = \left[ 12 \log_2 \left( \frac{k f_{sr}}{N f_{ref}} \right) \right] \bmod 12 \quad (3)$$

com  $f_{sr}$  = frequência de amostragem,  $f_{ref}$  = frequência de referência que, pela convenção deste algoritmo, estabelece o valor de 440 Hz para a classe da nota “*Lá*” (GOMEZ, 2006, pg. 46).

Para cada amostra analisada, são calculados os valores dos componente do vetor Chroma que são dados pela soma das magnitudes de todos os *bins* de frequência que correspondem a cada uma das 12 classes de alturas:

$$Chroma[j] = \sum_{k:P(k)=j}^{12} |X[k]|^2 \quad (4)$$

onde  $j$  é o índice relacionado à resolução das classes de alturas. Para semitom  $j = 0, 1, \dots, 11$ , quartos de tom  $j = 0, \dots, 23$  e para sexto de tom  $j = 0, \dots, 35$ .

#### 4. Processo Criativo

O processo criativo foi dividido em duas etapas: a) planejamento composicional baseado na análise de protótipos sonoros, denominados “Pontos de Referência”, com 03 descritores de áudio (Centróide Espectral, Desvio Padrão Espectral, Chroma) e b) elaboração de “Marcos Sonoros”, que são demarcações ou pilares estruturais da obra que foram construídos a partir dos Pontos de Referência. Ou seja, os Marcos Sonoros foram gerados sobrepondo-se várias amostras sonoras de técnicas estendidas de instrumentos de cordas. O espectro sonoro resultante da mistura das amostras foi analisado com os 3 descritores como objetivo de caracterizar um campo variacional para ação criativa do compositor.

##### 4.1 Pontos de Referência

No planejamento inicial, determinou-se que a estrutura formal de “Lana Tai” teria 14 Pontos de Referência. Esses partiriam de um uníssono, a nota  $D\#_4$ ,  $\sim 311.13$  Hz, em direção gradual a sonoridades complexas, nas quais se agregariam diferentes técnicas estendidas relacionadas aos modos de ataque das cordas, expandindo gradativamente o número de notas e a tessitura de cada naipe. Os valores médios do Centróide e Desvio Padrão estabeleceram um campo exploratório dentro de um espaço bidimensional. Ou seja, a partir dos valores para o uníssono, estabelecemos outros pontos de referência, elaborando, assim, a estrutura formal de “Lana Tai”, como mostram as Figura 2 e Figura 3:

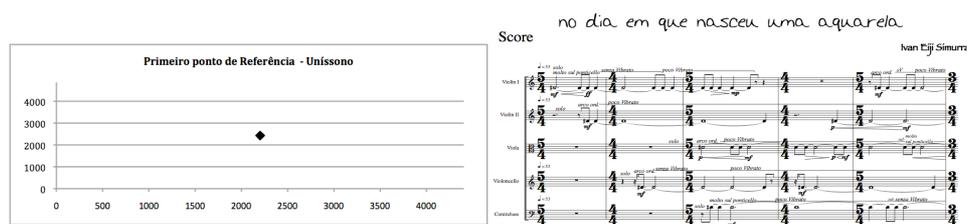


Figura 2: À esquerda, primeiro ponto de referência da obra. O eixo “x” representa o Centróide Espectral. O eixo “y” representa o Desvio Padrão. À direita, a abertura de “Lana Tai”, realizada pelos instrumentos solistas da orquestra de cordas.

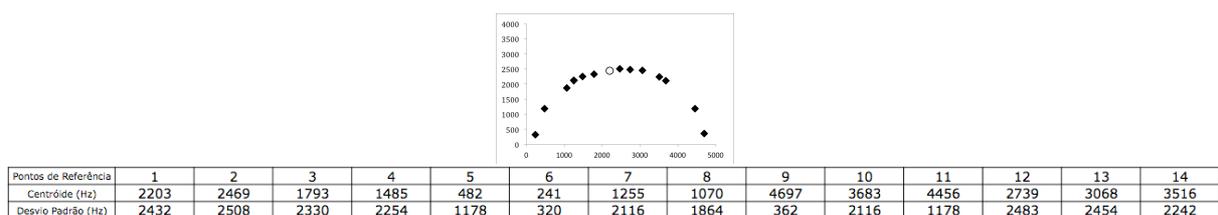


Figura 3: Referências estruturais de “Lana Tai”. No gráfico, o eixo “x” representa o Centróide Espectral e o eixo “y” representa o Desvio Padrão. O círculo branco indica a referência do uníssono.

## 4.2 Marcos Sonoros

Na segunda fase, fizemos novas medidas que foram comparadas aos Pontos de Referência. Ao se relacionar com os dados de Centróide e Desvio Padrão de cada Ponto de Referência, os MS apresentaram comportamentos distintos de Chroma. As alturas musicais com maior destaque em magnitude foram definidas como o “espaço de sonoridades” de cada MS. E as relações entre MS e os pontos de referência, via centróide espectral e desvio padrão, foram realizadas neste espaço de sonoridades. O desenvolvimento do trabalho composicional de “Lana Tai” concentrou-se em estabelecer relações de proximidade e distanciamento entre Marcos Sonoros e Pontos de Referência. No total, foram confeccionados 15 Marcos Sonoros, apresentados na matriz a seguir:

Marcos Sonoros	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Centróide (Hz)	2469	1793	1485	1877	1872	482	1473	1255	1311	1272	2739	3068	3516	4298	4597
Desvio Padrão (Hz)	2508	2330	2254	2203	2490	1178	2298	2116	1913	2239	2483	2454	2242	3133	3594

Figura 4: Matriz com os valores de Centróide e Desvio Padrão Espectral dos Marcos Sonoros.

Criamos transformações timbrísticas, uma trajetória que descreve a sonoridade global de “Lana Tai” em 3 etapas (vide figura 5): a) inicia-se numa frequência comum a todos os instrumentos, ou seja, o  $D\#_4 \sim 311$  Hz, b) desloca-se para uma região grave, até o MS 6 executado por contrabaixos e violoncelos,  $F\acute{a}\#_1 \sim 46$  Hz e, por fim, c) move-se para a região mais aguda, até o MS 15, com os violinos executando  $R\acute{e}\#_7 \sim 2489$  Hz.

The image shows three parts of a musical score labeled a), b), and c). Part a) shows the beginning of the piece with various instruments (Violin I, Violin II, Viola, Violoncello, Contrabaixo) playing a melody. Part b) shows a section for Violoncello and Contrabaixo. Part c) shows a section for Violoncello and Contrabaixo. The score includes dynamic markings like *pp*, *f*, and *mf*, and performance instructions like *tutti mst sV con sordina*.

Figura 5: Partitura das três etapas estruturais da sonoridade global de “Lana Tai”.

O concerto do qual “Lana Tai” participou foi realizado e gravado pela Orquestra Sinfônica da Unicamp, em outubro de 2012, sob a regência do maestro Akira Miyashiro. Os dados obtidos pela execução orquestral, denominados Excertos Sonoros, foram comparados com os Marcos Sonoros e são apresentados a seguir:

	Centróide (Hz)	Desvio Padrão (Hz)		Centróide (Hz)	Desvio Padrão (Hz)		Centróide (Hz)	Desvio Padrão (Hz)		Centróide (Hz)	Desvio Padrão (Hz)		Centróide (Hz)	Desvio Padrão (Hz)
MS I	2469.56	2508.36	MS IV	1877.79	2203.95	MS VII	1473.59	2298.86	MS X	1272.02	2239.72	MS XIII	3516.03	2242.57
EXCERTO I	1788.35	2980.95	EXCERTO IV	2564.02	3058.38	EXCERTO VII	1542.88	2633.87	EXCERTO X	1714.51	2698.73	EXCERTO XIII	2993.33	3203.88
MS II	1793.01	2330.93	MS V	1872.15	2490.45	MS VIII	1255.13	2116.79	MS XI	2739.19	2483.38	MS XIV	4298.51	3133.88
EXCERTO II	1804.2	2544.86	EXCERTO V	1836.44	2639.74	EXCERTO VIII	1757.65	2577.95	EXCERTO XI	2336.49	2869.14	EXCERTO XIV	3480.45	4218.4
MS III	1485.9	2254.63	MS VI	482.05	1178.93	MS IX	1311.39	1913.5	MS XII	3068.07	2454.58	MS XV	4597.51	3594.54
EXCERTO III	1588.85	2362.95	EXCERTO VI	1584.51	3221.42	EXCERTO IX	1611.16	2599.91	EXCERTO XII	2696.45	3005.77	EXCERTO XV	4493.97	4488.5

Figura 6: Tabela com os dados estatísticos dos excertos gravados pela orquestra e os MS.

Para uma análise consistente, as amostras foram configuradas com a mesma duração. Do ponto de vista local de cada MS e o seu respectivo Excerto Sonoro, a relação entre Centróide e Desvio Padrão Espectral evidenciou diferentes desvios entre dados coletados. O Excerto III apresentou os menores desvios com relação ao seu respectivo MS. De outro modo, o Excerto VI apresentou o maior desvio. Diversos fatores podem justificá-los: a) os desvios foram causados por diferentes tipos de afinação dos instrumentos, alterando o seu comportamento espectral e; b) ambiente acústico no momento da gravação. A próxima etapa da pesquisa se concentrará na análise e discussão destes dados.

## **6. Conclusões**

Segundo Grégoire Carpentier (CARPENTIER, 2008, pg. 37), dentro da História da Música, as questões às quais circundam a problemática sobre as estratégias de orquestração sempre encontraram um subterfúgio na subjetividade, empirismo e nas “práticas comuns” desenvolvidas pelos próprios compositores. Mesmo os já consagrados tratados de instrumentação e orquestração, como os de Berlioz (BERLIOZ, 1844) e Rimsky-Korsakov (RIMSKY-KORSAKOV, 1912), seguem essa metodologia. Naturalmente, os procedimentos composicionais e metodológicos, aqui reportados, não objetivam estabelecer um tratado contemporâneo de orquestração e instrumentação auxiliado por suporte computacional. Mas, este trabalho propõe o desenvolvimento de ambientes para criar objetos sonoros/musicais que podem ser observados a partir dos métodos de recuperação das informações musicais. Os próximos passos deste trabalho se concentrarão na expansão das análises de interação entre Pontos de Referência, Marcos Sonoros e Excertos Gravados pela orquestra. Esse aspecto será abordado em publicações futuras, principalmente porque a quantidade de dados gerados e a análise requerem tempo, domínio de várias técnicas de processamento de sinais e, por extensão, novos desenvolvimentos de programação para construir um conhecimento composicional que possa conviver com o planejado, o gravado, sua análise e a volta para o processo criativo.

## **7. Agradecimentos**

Os autores gostariam de agradecer o apoio da FAPESP à bolsa de doutorado, processo: 2011/23972-2, e ao CNPq aos projetos no. 477343/2010-4 e no. 304064/2010-6.

## Referências

- BERLIOZ, H. *Treatise on Instrumentation*. Paris: Schoenberger, n.d. (1843)
- BYRD, D; FINGERHUT, M. *The History of ISMIR – A Short Happy Tale*. D-lib Magazine, Vol. 8, No. 11, 2002.
- CARPENTIER, G. – Approche computationelle de l’orchestration musicale. Ph.D. Thesis, University Pierre et Marie Curie, Paris. 247 pgs.
- FUJISHIMA, T. Realtime Chord Recognition of Musical Sound: a System Using Common Lisp Music. CCRMA, Stanford University, 1999. 4 pgs.
- GOMEZ, E. Tonal description of music audio signals, Ph.D. dissertation, MTG, Universitat Pompeu Fabra, Barcelona, Spain, 2006, <http://mtg.upf.edu/egomez/thesis>.
- MALT, M.; JOURDAN, E. *Zsa.Descriptors: a library for real-time descriptors analysis*. In 5<sup>th</sup> Sound And Music Computing Conference, Berlin, Allemagne, 31th july to August, 3rd, 2008.
- MANZOLLI, J. Um Modelo Matemático para Timbre Orquestral, dissertação de mestrado, IMECC, Unicamp, 1988.
- MONTEIRO, A. MANZOLLI, J. *Análise Computacional de Texturas Sonoras via Mapas de Poincaré*. XXI Congresso da Associação Nacional de Pesquisa e Pós-Graduação em Música – Uberlândia. 2011.
- MONTEIRO, A. Criação e Performance Musical no Contexto dos Instrumentos Digitais. Dissertação de Mestrado. Núcleo Interdisciplinar de Comunicação Sonora. Universidade Estadual Campinas. 2012. 159 pgs.
- PADOVANI, J.; FERRAZ, S. Proto-História, Evolução e Situação Atual das Técnicas Estendidas na Criação Musical e na Performance. *Música Hodie*, Vol.11. Nº2. 2011.
- PEREIRA, E. Estudos Sobre uma Ferramenta de Classificação Musical. Dissertação de Mestrado – Universidade Estadual Campinas, Faculdade de Engenharia e de Computação. 2009. 65 pgs.
- RIMSKY-KORSAKOV. *Principles of Orchestration*. 1912. Dover Publications, 1964.
- SHEH A, ELLIS DPW. Chord segmentation and recognition using EM-trained hidden Markov models. In: International Conference On Music Information Retrieval, 2003, Baltimore, USA. 2003. p. 185–191.
- SHEPARD, R. N. Structural representations of musical pitch. In Deutsch, D., editor, *The Psychology of Music, First Edition*. Swets & Zeitlinger. 1982.
- THEODORIDIS, S.; KOUTROUMBAS. *Pattern Recognition*. Academic press, Third Edition. 2009. 967 pgs.

<sup>1</sup> Disponível em <http://www.ismir.net/>. Data de acesso: 24/02/2013.

<sup>2</sup> Disponível em <http://www.puredata.info>. Data de acesso: 26/02/2013.