

Usual sentence derivation from a grammar will automatically generate the desired structure for the generated music. The user may impose restrictions, e.g. on the maximum number of fragments, of fragments of a given type, of repetitions of certain sequences of types. The expected result is the automatic generation of music with sound similar in quality to the previous ones, and showing the structure created by the system under user's specification.

**6.4-Phase 3** - to be started in short term - the next stage refers to obtaining structuring rules automatically. The underlying grammar of the structure imposed to automatically generated compositions will be itself deduced by the system. A musician must help in this activity by supplying structural analyses for a significant set of works to be used as models, identifying and classifying the fragments in each work. Then, an algorithm is executed to make a grammatical inference that generalize the information extracted from the given samples, in order to infer a formation rule to be used as a grammar describing the desired structure. This phase will result in a complete synthesis program, for music will then be obtained automatically both in contents and in formal structure. User's participation will be restricted to inserting information on the structure of the samples used as models.

**6.5-Phase 4** - this phase is foreseen to begin in some years. It is much more complex, because it should include heavier artificial intelligence resources; in order to operate with no human-supplied information concerning the structure of the pieces used as models. Structural analysis should be made by programs with enough built-in knowledge on musical analysis to allow automatic identification of the musical elements present in the score. This phase will produce software that will be able to perform musical analysis and synthesis, without direct participation of the user.

#### 7-References

- COPE, DAVID. **Experiments Musical in Intelligence (EMI): Non-linear Linguistic-Based Composition** Interface, vol, 18, p. 117-139, 1989.
- COPE, DAVID. **Computer Modeling of Musical Intelligence in EMI**. Computer Music Journal, vol 16, no. 2, Summer, p. 69-83, 1992.
- COPE, DAVID. **Music & Lisp**. OH Expert, March, 1988, p.26-33.
- DODGE, CHARLES; JESSE, THOMAS A.. **Composition With Computers - Computer Music - Synthesis, Compositions and Performance**. NY, Schirmer, London, Collier MacMilan, 381 pg., p.265-321, 1985.
- GOGINS, MICHAEL. **Iterated Functions Systems Music**. Computer Music Journal, vol. 15, no. 1, Spring, 1991, p.40-48
- HUDAK, PAUL; MAKUCEVICH, TONE; DDE, SYANGA; WONG,BO. **Haskore Music Notation—An Algebra of Music** -. Journal of Functional Language, vol 6, May, 1996 p.465-483
- LIMA, LUCIANO VIEIRA. **Um Sistema de Composição Musical Dirigido por Estilo**. Tese de Doutorado, Universidade de São Paulo, São Paulo - Brasil, 1998.
- SMOLIAR, STEPHEN W. Book Review and Response - **Models of Musical Communication and Cognition**. Interface, vol. 18, 1990, p. 361-371
- TODD, PETERSOM. **Connectionist Approach to Algorithmic Composition**. Computer Music Journal, vol. 13, no.4, Winter, 1989, p.27-43.

## SOUND FUNCTORS APPLICATIONS

JÔNATAS MANZOLLI<sup>2</sup>, ADOLFO MAIA Jr.<sup>1,2</sup>

*Mathematics Department<sup>1</sup>*

*Interdisciplinary Nucleus for Sound Communications (NICS)<sup>2</sup>*

*University of Campinas (UNICAMP) - Brazil*

*[Jonatas@nics.unicam.br](mailto:Jonatas@nics.unicam.br), [adolfo@nics.unicamp.br](mailto:adolfo@nics.unicamp.br)*

### ABSTRACT

*The concept of mathematical structures called Functors can be useful to develop a large number of compositional algorithms. We introduce here concepts such as Categories and Functors as generalised sound construction tools. We define two applications: a functor between plane curves and sound events based on the MIDI protocol and a functor between the class of continuous curves  $C(x)$  defined in a finite interval  $I \subset \mathcal{R}$  and the class  $\Psi$  of Fourier Spectra  $A(v)$ .*

### INTRODUCTION

Music composition is plenty of procedures that can be related to mathematical symmetries, it is presented in (Hofstadter 1989), a study relating music, design and mathematical structures. On the other hand, music symmetries can be better understood and handled if we search for underlined or hidden mathematical structure that generates them. In this work we show that the intuitive use of symmetry and associations in music can be studied and explored through the mathematical concepts of Category and Functor.

This approach is interesting, not only because it can be used to classify sound material, but mainly for it furnishes new relations pointing out to an enormous variety of sound generation methods. In short, we claim that Functors, in the same way it was firstly devised in mathematics (MacLane & Birkhoff 1953, 1979; MacLane 1971), can be generalised as universal tools to construct mathematical models for music composition and sound synthesis.

There has been a series of approaches applying mathematics to build sound structures such the use of 1/f noise fractal distribution (Voss & Clark 1978; Bolognesi 1983), non-linear dynamical systems and iterated function systems (Pressing 1988; Scipio 1990; Gogins 1991). There is a study about these systems in (Manzolini, 1993a).

Our group has been worked with applied mathematics to produce sound design machines focusing new methods for sound synthesis using non-linear dynamics (Manzolini 1993b), mathematical models for algorithm composition such as Markov Chains and

Boundary Functions (Manzoli & Maia 1995) and development of interactive desktop and gesture interface for composition in real time (Manzoli & Ohtsuki 1996).

Below, we discuss Functors as generalised Algorithmic Compositional tools for sound construction defining Class, Morphism, Category and Functor. It follows applications in the sound domain.

## CATEGORIES AND FUNCTORS

Functor is a kind of function, as suggested by its own name, but it is not an ordinary one because it carries the underlined structures of the sets in which it is applied. These sets are called Categories. More precisely, a **Category**  $\Phi$  is defined by three proprieties:

- A class of objects **A, B, C...**
- For each pair of objects **A, B**  $\in \Phi$  we have a set of applications (morphisms) **M(A,B)** from **A** to **B**.
- For each triple of objects **A, B, C**  $\in \Phi$  we have a composition law for the morphisms

$$\begin{aligned} \mathbf{M(A,B)} \times \mathbf{M(B,C)} &\rightarrow \mathbf{M(A,C)} \\ (f,g) &\longrightarrow g \circ f \end{aligned}$$

Now, for the proprieties above it follows three axioms, which are primary properties of these categories:

**A<sub>1</sub>**) The sets of morphisms **M(A,B)** and **M(C,D)** are mutually disjoint unless **A = C** and **B = D**.

**A<sub>2</sub>**) Associative Law: **h(gf) = (hg)f**.

**A<sub>3</sub>**) Existence of Identity: for each object **A** there exists a morphism identity **1<sub>A</sub>: A → A** such that for any **f: A → B** and **g: C → A** we have **f ∘ 1<sub>A</sub> = f** and **1<sub>A</sub> ∘ g = g**.

Usually, the theory of categories and functors is mathematically involved. Here we use only basic properties of the definition above to show that underlined structures can be mapped from categories of mathematical objects to categories of sound objects and the last undertake formal properties of their mathematical counterparts via a functor.

Given two categories  $\Phi$  and  $\Psi$ , a functor **F** between  $\Phi$  and  $\Psi$  is a map which associates each object **A**  $\in \Phi$  to an object **F(A)**  $\in \Psi$

$$\begin{aligned} \mathbf{F: \Phi} &\longrightarrow \mathbf{\Psi} \\ \mathbf{A} &\longrightarrow \mathbf{F(A)} \end{aligned}$$

and for each morphism **f**  $\in \mathbf{M(A,B)}$  associates a morphism **F(f)**  $\in \mathbf{M(F(A), F(B))}$  with the properties

$$\begin{aligned} \mathbf{F(gf)} &= \mathbf{F(g) F(f)} \\ \mathbf{F(1_A)} &= \mathbf{1_{F(A)}} \end{aligned}$$

Functor **F** operates on morphisms as well on elements of the category  $\Psi$ . In this way, the structure of a product (or composition) between two morphisms in the category  $\Phi$  is transported to morphisms in the category  $\Psi$  via functor **F**.

There are too many examples of mathematical functors. Since we are focusing music application only, two simple examples of functors applied to sound categories are presented in the next section.

## APPLICATIONS

### EXAMPLE 1 - Sequence of MIDI Events

We begin with the mathematical category  $\Phi = \{\text{continuous finite curves in a bounded region } U \subset \mathbb{R}^2\}$ . Given two curves **C<sub>1</sub>, C<sub>2</sub>**  $\in \Phi$ , a morphism in  $\Phi$  as an application

$$\begin{aligned} \mathbf{f: C_1} &\longrightarrow \mathbf{C_2} \\ \mathbf{x} &\longrightarrow \mathbf{f(x)} \end{aligned}$$

which means to deform **C<sub>1</sub>** into **C<sub>2</sub>**, i.e. **f**  $\in \mathbf{M(C_1, C_2)}$ . The product **gf** is defined as the composition **gf = gof**

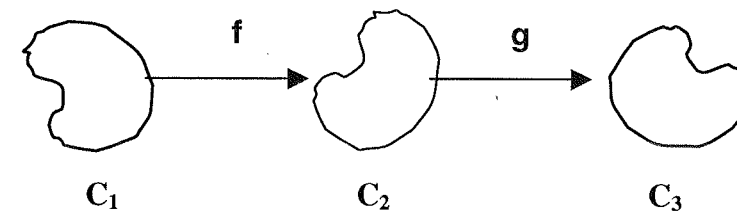


Figure 1. Composition of Plane Curves

We define  $\Psi$  as a category of functions to control sound parameters described by the MIDI (Music Instrument Digital Interface) protocol. Thus, **F(f)**  $\in \mathbf{M(F(C_1), F(C_2))}$  and **F(f)** deforms **F(C<sub>1</sub>)** to **F(C<sub>2</sub>)**. Given a curve **C** in  $\Psi$ , the function **F(C)**  $\in \Psi$  can be defined in several different manners. Particularly, **F(C)** can be expressed by the Distance Function between a fixed point in the plane to the points of the curve **C**. It is easy to see that the functor property is satisfied, namely

$$\mathbf{F(g(f(C)))} = \mathbf{F((gf)(C))} = \mathbf{F(g)F(f)(F(C))}$$

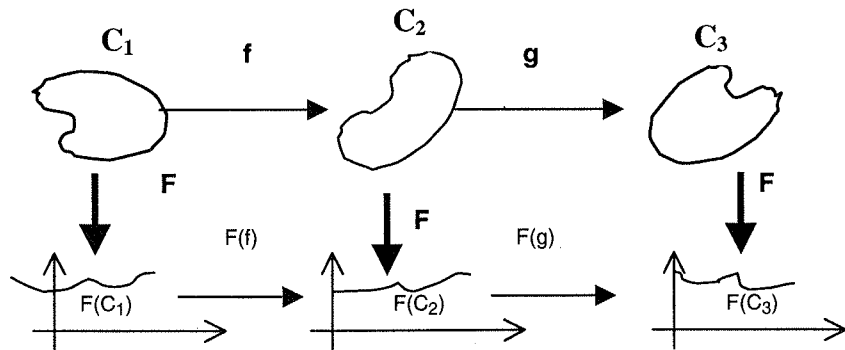


Figure 2. Diagram of the Functor operation

We have implemented this functor of plane curves to MIDI parameters in a computer program called CurvaSom. This software uses a graphic interface running in a MS Windows NT environment. Here we are just introducing main theoretical aspects of the Sound Functor approach.

#### EXAMPLE 2 - Sound Synthesis

In this example, an application of a functor is used to build Spectral Envelopes. We take as mathematical category, a class  $\Phi$  of smooth functions  $C(x)$  defined in a finite interval  $I \subset \mathbb{R}$  and as sound category, a class  $\Psi$  of Fourier Spectra  $A(v)$  of a finite set of sounds. Here the function  $C(x)$  acts as a shaper of a spectral set. Starting from a fixed spectrum, which we call Input Spectrum, we use deformation of curves as morphisms in both classes of functions. We construct the following Functor:

$$\begin{aligned} \mathbf{F}: \Phi &\longrightarrow \Psi \\ C &\longrightarrow \mathbf{F}(C) = C(A(v)) \end{aligned}$$

Where the parameter  $x = A(v)$ , with  $0 \leq x \leq A_{max}$  and  $A_{max}$  is the maximal amplitude considered.

The morphisms in  $\Phi$  are deformations of smooth functions and the morphism in  $\Psi$  can be chosen as  $\mathbf{F}(f)=f$ . This means that the morphisms in  $\Phi$  are used as spectral deformations. Using the above definition it is easy to see the property  $\mathbf{F}(g) \mathbf{F}(f) = \mathbf{g}f = \mathbf{F}(gf)$  is satisfied.

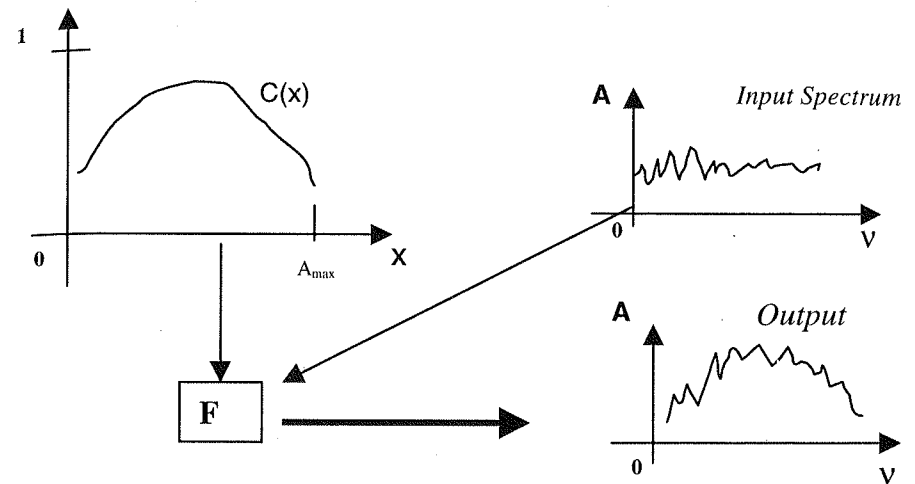


Figure 3. Spectrum Functor Diagram

#### CONCLUSION

We can construct several examples similar to the above ones. Another example using a different mathematical construction is presented in (Maia, Valle & Manzolli 1998). Other mathematical structures like groups, lattices, algebras, several different geometrical and topological structures can be used as mathematical categories (see MacLane 1971) in order to get sound outputs through a suitable choice of functors. Functors allow us to construct a huge quantity of sound outputs reflecting the structure (or symmetry in several cases) of external mathematical categories used to generate them. Musicians who work on Computer Music, aware of the properties of functors, can expand their sound tools in a way, certainly, not yet explored by the authors of this paper.

Mathematical tools in composition brings new possibilities for composers envision development of Compositional Systems. Research in Computer Music provides powered tools for constructing these systems. This union of artistic and mathematical knowledge creates a framework for investigation and music production, an environment for applying mathematics in order to manipulate sound structures. Mathematical models presented here could be expanded using graphic interfaces to create new musical performance situations, making mathematical design to produce computer music instruments.

## REFERENCES

- Bolognesi, T. (1983). Automatic Composition: Experiments with Self-Similar Music. *Computer Music Journal* 7(1):25-36.
- Gogins, M. (1991). Iterated Functions Systems Music. *Computer Music Journal* 15(1):40-48.
- Hofstadter, D. R. (1989). *Gödel, Escher, Bach: an eternal golden braid*. Vintage Books, New York, ISBN 0-394-75682-7.
- MacLane, S. (1971). *Categorias for the working mathematician*. New York, Springer-Verlag.
- MacLane, S. & G. Birkhoff (1953). *A Brief Survey of Modern Algebra*. New York, the MacMillan Company.
- MacLane, S. & G. Birkhoff (1979). *Algebra*. New York, Macmillan Company.
- Manzoli, J. (1993a). Non-linear Dynamics and Fractals as a Model for Sound Synthesis and Real Time Composition. PhD Thesis at the University of Nottingham, England.
- Manzoli, J. (1993b). *Musical Applications Derived From FracWave Sound Synthesis Method. Proceedings of the 94th Audio Engineering Society Convention*, Berlin.
- Manzoli, J. (1994). FracWave: Non-linear Dynamics as Timbral Constructs. Proceedings of the of the XIV Congress of the Computer Brazilian Society, I Symposium on Computer Music, Caxambú - Brasil.
- Manzoli, J. & A. Maia Jr (1995). Interactive Composition Using Markov Chain and Boundary Functions. *Proceedings of the XV Congress of the Computer Brazilian Society, II Symposium on Computer Music*, Canela, Brazil.
- Manzoli, J. & W. Ohtsuki (1996). INTERASOM: a desktop for Algorithmic Composition. *Proceedings of the XVI Congress of the Computer Brazilian Society, III Symposium on Computer Music*, Recife, Brazil.
- Maia Jr., A., R. do Valle & J. Manzoli (1998). Estruturas Matemáticas como Ferramenta Algorítmica para Composição. Submitted to the XI Congress of the Brazilian Association for Research in Music, ANPPOM, University of Campinas.
- Pressing, J. (1988). Non-linear Maps as Generators of Musical Design. *Computer Music Journal* 12(2):35-45.
- Scipio, A. (1990). Composition by Exploring of Non-linear Dynamic System. *Proceedings of the 1991 International Computer Music Conference*.
- Voss, R. F. & J. Clarke, (1978). 1/f noise music: Music from 1/f noise. *Journal of the Acoustical Society of America* 63(1):258-263.

## ACKNOWLEDGMENT

We thank Prof. Raul do Valle for fruitful discussions and the Foundation for Research of State of São Paulo - FAPESP for supporting this work.

## Busca e Recuperação de Informação Musical

Ana Miccolis

Universidade Federal do Rio de Janeiro  
Coordenação dos Programas de Pós-Graduação em Engenharia - COPPE/UFRJ  
Programa de Engenharia de Sistemas e Computação  
Caixa Postal 68 511- CEP 21945-270 - Rio de Janeiro - RJ - Brasil  
E-mail miccolis@cos.ufrj.br

### Abstract

The process of musical composition can be benefit by the use of resources extracted from a database of musical passages. For this purpose, the composer may have access to the sound resources and be able to manipulate them in the compositional process. The traditional musical representation employed at the conventional music compositions allow the composer to retrieve the necessary information, to process it and to restore it in the same representation. The musical data in non conventional composition can't be represented only by the conventional written. In this case, the musician needs an resource that let the composer select music passages based on musica characteristics. One of the goals of information retrieval of musical data is to aid the research in compositional resources, like selection of timbre in a database. The database with this capability might provide the access to all the conventional attributes like alphanumeric data and the musical data. This paper investigates this kind of search and retrieval and which criteria one may consider at the musical information retrieval process.

### 1 - Introdução

Sistemas de recuperação de informação processam requisições de informações sobre arquivos identificando e recuperando destes certos registros que atendem aos requisitos exigidos. A recuperação de alguns registros em particular depende da similaridade entre estes e as consultas. Esta similaridade pode ser medida através da comparação de valores de certos atributos dos registros com os valores requisitados pela consulta. Em muitos sistemas de gerenciamento de banco de dados os arquivos contém registros homogêneos, com um conjunto de atributos especificado para caracterizar cada item do arquivo e com valores dos atributos que conseguem descrever univocamente e de forma completa os registros armazenados. Nestas circunstâncias a recuperação de informação depende de uma comparação exata entre os valores do arquivo e os da consulta. O nome de um autor ou intérprete, a data de nascimento, a sua nacionalidade e os títulos das obras musicais são tratados como dados não ambíguos e o processo de recuperação baseado nestes critérios é

simples. Contudo outros tipos de informação exigem a construção de modelos abstratos para representá-los e linguagens avançadas para definição e manipulação de dados. Num sistema de informações musicais, a recuperação de dados convencionais é importante, mas carece de outros elementos que auxiliem o músico no processo de composição ou análise. Seria de interesse para este tipo de usuário que além da seleção de obras a partir de atributos convencionais (autor, título da obra, instrumento de execução, etc.), a busca por trechos ou amostras sonoras pudessem ocorrer a partir da comparação entre elementos característicos do dado sonoro. A busca e recuperação poderia ocorrer em três níveis de abstração:

- Busca e recuperação de um trecho específico - Equivalente à comparação exata entre os dados de uma obra armazenada e os de uma consulta formulada. Poderia retornar por exemplo, uma Sonata de Mozart que possuísse no corpo da obra o trecho musical oferecido como argumento de pesquisa.
- Busca e recuperação por atributos acústicos mensuráveis - Equivalente à comparação em textos utilizando técnicas Fuzzi para selecionar trechos musicais que possuam atributos acústicos mensuráveis como altura, duração e volume, dentro de uma faixa de aceitação.
- Busca e recuperação por propriedades subjetivas do som - Correspondente à comparação entre elementos determinantes do timbre de um som como o envelope de amplitude, harmonicidade e envelope espectral.

Para atender às necessidades deste tipo de usuário, com conhecimentos de música e até de acústica musical, um sistema de informações musicais deveria oferecer busca e recuperação de informação de áudio em vários níveis de abstração. Os níveis 1 e 2 que são comuns à busca de texto e o nível 3 que é específico de áudio.

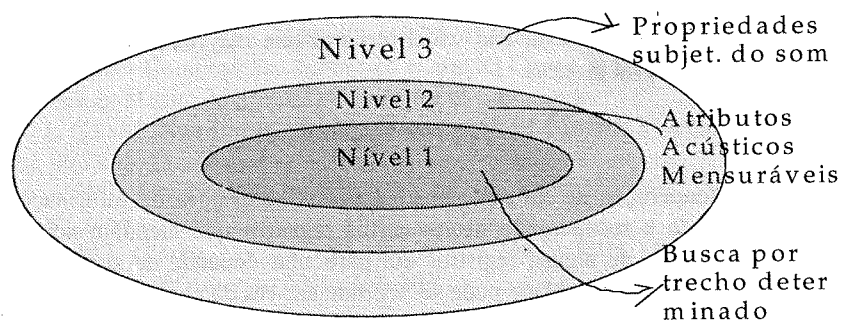


Figura 1 – Níveis de abstração de uma consulta em um sistema de informações musicais

Um sistema de informações musicais deve atender a diversos tipos de demanda de informação permitindo o acesso de usuários com diferenciados níveis de conhecimento no assunto. A grande quantidade de dados obriga o desenvolvimento de métodos de filtragem, seleção e interpretação dos dados que possibilitem um real aproveitamento de toda a informação disponibilizada. O objetivo da busca e recuperação de informações musicais é

prover diferentes níveis de acesso aos interessados em música. Entre os fins profissionais pode-se citar a pesquisa de recursos para composição (banco de timbres, por exemplo), o auxílio à pesquisa musicológica, transformação de exemplos sonoros e a possibilidade de busca de música não convencional através de uma representação própria. Para os diletantes um objetivo seria o de propiciar o conhecimento gradativo do universo musical disponível a partir de características próprias fornecidas por este usuário.

## 2 - Descrição dos Atributos da Informação Musical

### 2.1. A Altura

Na música convencional a identificação da altura de um som pode ser bem definida a partir da frequência da fundamental, mas na música não convencional este conceito sofre interferência de outros parâmetros, que passam a ser manipulados neste tipo de composição. Experiências de Schaeffer (descritas no Solfejo do Objeto Sonoro) apontam para uma nova noção de altura na qual a compreensão de outros fatores torna-se tão ou mais importante que a medida de frequência dada pela fundamental. Em seu trabalho, a exclusão da fundamental e a permanência dos demais parciais de um som possibilita a correta reconstrução da altura original. A possibilidade de supressão de harmônicos somente foi possível na música contemporânea e por esse motivo, as definições clássicas deste atributo começam a ganhar outros conhecimentos anteriormente desconsiderados.

### 2.2. A Intensidade

A intensidade traduz as sensações de forte ou fraco de um mesmo som e depende da amplitude sendo medida em decibéis. Outros termos que exprimem a noção de intensidade são a amplitude, volume, loudness e nos equipamentos midi, a velocidade. Segundo alguns autores a intensidade é influenciada não somente pela amplitude mas também pelas características temporais e espectrais do som (Moore 1989; Pierce 1992). O compositor pode empregar este recurso como forma de destacar um som, dando-lhe mais intensidade que aos demais. Este recurso é muito utilizado para formar a dinâmica da música permitindo criar planos sonoros distintos, uma vez que ele está relacionado com a percepção de distância.

### 2.3. A Duração

Em geral na escrita convencional, a música é disposta em compassos de duração determinada pelo andamento da obra, o qual define a duração da unidade de tempo de um trecho musical. Para isso o autor da obra seleciona a unidade com que irá trabalhar (semínima, colcheia, etc.) e quantas unidades de tempo deseja que estejam contidas num único compasso. Na música não convencional, onde parâmetros como ataque de um som, podem ser manipulados, verifica-se que a noção de tempo pode ser alterada em função deste recurso. Invertendo-se o ataque de um instrumento, por exemplo, pode-se conseguir alterar a percepção de duração do mesmo. Isto acontece quando a parte final da duração do

som, que teria uma consonância conhecida do ouvinte passa a ter uma inversão. Desta forma o observador não se desliga da conclusão do som, passando a perceber a duração do mesmo como mais longa que a do mesmo som sem a inversão.

#### 2.4. O Timbre

O timbre é a qualidade do som que permite a diferenciação entre os diferentes instrumentos musicais. Ele pode ser um valioso recurso para as composições contemporâneas, mas sua utilização ainda é dificultada pelo fato deste atributo ser fisicamente difícil de descrever e usar. Os primeiros estudos significantes nesta área foram realizados no século XIX por Hermann Helmholtz(1863) apontando para uma definição de timbre que o considerava apenas em função da composição de seus harmônicos. Depois da primeira síntese digital do som de Max Mathews em 1958 ficou aparente que o timbre não poderia depender apenas de distribuição harmônica do som, mas também de características temporais. Outras pesquisas de Risset e Mathews revelam que a proporção de harmônicos em certas classes de instrumentos musicais aumenta à medida que a intensidade cresce, dando início a uma nova abordagem do timbre na qual não se pode descrevê-lo por valores instantâneos de parâmetros numéricos, mas sim pela evolução deles, sugerindo que características espectrais e temporais são inseparáveis para definição do timbre. Nos estudos de John Grey(1977) ele aponta para uma solução tridimensional para a percepção de timbre. Uma interpretação física desses três eixos foi feita a partir da distribuição de energia do espectro harmônico, assim como das propriedades temporais dos sons. Em (Hourdin 1997) uma nova abordagem para o estudo do timbre é utilizada, aplicando técnicas multidimensionais à descrição física do som, com recursos do software Csound (Vercoe 1992). Neste trabalho a representação do timbre inclui tanto características espectrais como aspectos temporais do som permitindo uma classificação por ordem de importância das propriedades físicas. A vantagem desta abordagem reside principalmente na similaridade com que os tons são percebidos pelos humanos. A percepção de timbre aparece diretamente associada com a trajetória do som no espaço físico. Pode-se considerar para efeito de se extrair as componentes significativas do timbre, as características como envelope de amplitude e envelope espectral. A utilização da fórmula denominada "Fast Fourier Transform"(FFT) possibilita a observação do espectro harmônico de um som. O timbre sofre alterações durante o tempo de duração de uma nota. Por exemplo, se uma nota tocada no oboé iniciar com intensidade fraca e for progressivamente crescendo em intensidade o espectro harmônico irá se modificar com o volume. Novos harmônicos são adicionados ao som e ele ganha mais volume. Este é o motivo pelo qual em um som eletrônico não se consegue um efeito de "crescendo" convincente apenas aumentando a sua intensidade. O ataque de um tom é freqüentemente essencial para a identificação do timbre. Timbres com uma mesma distribuição de energia no espectro tendem a ser considerados similares. Em (Blum 1996A) o timbre pode ser analisado a partir de seus componentes: envelope de amplitude, harmonicidade e envelope espectral. O envelope de amplitude traduz a característica acústica de brilho. O brilho de um som é calculado como a centróide do espectro de magnitude da STF(Short-time Fourier). Exemplificando, quando coloca-se a mão sobre a boca, o brilho do som é reduzido

e também a sua intensidade. O envelope espectral traduz a característica acústica conhecida como largura de banda que é calculada como sendo a média ponderada da magnitude das diferenças entre os componentes espectrais e a centróide. Exemplificando, uma simples onda senoidal tem uma característica de largura de banda equivalente a zero, enquanto um ruído branco tem um valor infinito para esta característica. A harmonicidade permite a distinção entre um espectro harmônico(como vozes e sons musicais), um espectro inarmônico(como sons metálicos), e ruídos(espectros que variam aleatoriamente em freqüência e tempo). Ela é calculada medindo-se o desvio do espectro de um som de um espectro harmônico perfeito.

### 3 - Estrutura para Armazenamento dos Atributos do Som

A informação sonora deve poder ser recuperada a partir de similaridade acústica, sendo necessário assim medir esta similaridade. Alguns métodos tratam a busca de dados sonoros independente de suas características particulares, dando-lhes o mesmo tratamento utilizado na recuperação de imagens (Foote 1997). Contudo a pesquisa de timbres não consegue ser corretamente atendida por estes métodos, em virtude da priorização das alturas sonoras como fator de indexação. O timbre ao contrário das demais características acústicas deve ser definido a partir das suas componentes como foi visto no tópico 2. A pesquisa sobre a composição do timbre considera a análise de vários componentes, entre eles, o envelope de amplitude, o envelope espectral e a harmonicidade. Para permitir a recuperação de dados sonoros a partir de diferentes critérios, deve-se primeiramente medir uma variedade de atributos acústicos para cada som. Este conjunto de "N" atributos pode ser representado em uma estrutura "N-vetorial". Em um sistema de busca e recuperação de informações de áudio, o usuário deve poder comparar as informações similares para um dado som ou que tenham mais ou menos uma determinada propriedade. Para permitir esta busca o espaço vetorial criado deve satisfazer a seguinte restrição para cada propriedade a ser utilizada na recuperação do dados. Primeiramente os sons divergentes em relação a uma dada propriedade devem ser mapeados em diferentes regiões do N-espaço para permitir que o banco de dados possa fazer distinção entre diferentes valores desta propriedade. Além disso deve garantir que todas as propriedades relevantes tenham sido consideradas na construção do vetor para atender a um número máximo de propriedades que o usuário possa especificar na consulta. Os diversos aspectos do som como altura, sonoridade, brilho e harmonicidade podem variar com o tempo e sua trajetória deve ser calculada e armazenada em função de parâmetros como: média, desvio padrão, autocorrelação, máximos e mínimos, pontos críticos, tempo de ataque, etc. Um vetor deve ser utilizado para medir a duração e os parâmetros citados para cada um dos aspectos do som (altura, harmonicidade, intensidade).

#### 4 - Classificação de Dados Sonoros

Bancos de dados com recursos multimídia armazenam gravações de áudio frequentemente indexadas de forma muito pobre, no sentido de utilização de classificação. Geralmente a indexação é realizada a partir de atributos textuais aos quais o dado sonoro se relaciona, acarretando uma série de dificuldades no momento da busca. Uma delas é a ambigüidade na classificação do dado. Mesmo que o mantenedor do recurso tenha associado palavras chaves para classificá-lo, esta associação é bastante subjetiva e pode não ser compreensível para outra pessoa. A busca por um som particular ou uma classe de sons requer um processo de classificação que permita identificar dados sonoros a partir de seus atributos próprios, permitindo uma comparação entre eles tal que sons similares segundo uma determinada propriedade estejam mapeados numa mesma região do espaço de classificação. Uma vez que se tenha conseguido armazenar os sons numa estrutura adequada ao tipo de classificação que se pretende (Blum 1996B), como por exemplo, um Vetor\_N, pode-se utilizar as propriedades definidas (altura, intensidade, duração, etc) para realizar a classificação de dados sonoros. Para cada novo som inserido no banco de dados, o vetor é calculado e quando ele precisa ser classificado, este processo é realizado a partir do cálculo da distância deste em relação ao modelo existente, levando em consideração o peso de cada característica para a classe tratada. A distância é comparada para determinar se o som está dentro ou fora da classe. Se ele se enquadrar em mais de uma classe, ele será colocado naquela da qual ele for menos distante. Para bancos de dados de pequeno volume torna-se fácil calcular a distância para todos os sons e depois procurar aqueles que atendem a um certo resultado. Para aumentar a velocidade de busca em banco de dados muito grandes, pode-se indexar os sons por todas as características acústicas.

#### 5 - Consultas por Similaridade Melódica

A consulta por similaridade melódica pode atender vários tipos de necessidade. Para responder a um usuário que precise extrair as obras que contenham uma melodia exatamente igual a da consulta, o sistema de busca deverá selecionar trechos musicais construídos a partir de uma sucessão de notas da mesma altura que a da melodia informada. Este tipo de consulta, pode ser feito inclusive a partir da escrita convencional de partitura, como na representação e recuperação baseada em conteúdo de partituras musicais em bases de dados orientadas a objetos proposta por M. Figueiredo no IV SBCM/1997. Contudo a extração da informação a partir da partitura convencional limita o universo de consulta, pois muitas vezes o usuário não contém conhecimento sobre escrita musical e mesmo para aqueles que o possuem, o interesse na recuperação de obras musicais a partir de similaridade melódica requer uma flexibilidade quanto à igualdade de frequências. Um segundo grupo de usuários seria aquele interessado em formular consultas num banco de dados de música para obter obras cuja melodia fosse similar à informada, porém escritas num certo intervalo acima ou abaixo do original.

Consulta	Frequência(Hz)	Nota recuperada	frequência(Hz)	Intervalo
Do 3	130,81	Sol 3	196,00	(do3 a sol3) - 5a. justa
Sol 2	98,00	Ré 3	146,83	(sol2 a re3) - 5a. justa
Lá 2	110,00	Mi 3	164,81	(lá2 a mi3) - 5a. justa
Do 4	261,63	Sol 4	392,00	(dó4 a sol4) - 5a. justa

Tabela 1 – Relação de frequências e intervalos de notas de um trecho de uma consulta e da obra recuperada, a partir de similaridade melódica

Desta forma, se o usuário informasse um trecho melódico contendo as notas do, sol, la e do, músicas com uma melodia que contivessem estas notas exatamente ou por exemplo, sol, re, mi e sol deveriam ser recuperadas. No exemplo dado, um intervalo de quinta justa é mantido entre cada das quatro notas da melodia. Ainda que a frequência de cada nota comparada seja diferente, a diferença de altura permanece a mesma.

Uma terceira classe de consulta por similaridade melódica seria aquela de aproximação das alturas. Neste caso, tomaria-se como base a primeira nota da seqüência e a partir dela as demais seriam classificadas segundo a sua altura em relação à nota imediatamente anterior. Para esse fim a escala de sons pode ser reduzida a um alfabeto restrito que indique a altura relativa de cada nota em relação à anterior. Considere como exemplo, a conversão da introdução do tema da Quinta Sinfonia de Beethoven na seguinte seqüência: - B B A C B B A. A primeira nota é ignorada por não possuir uma anterior na seqüência que permita a comparação. Excetuando a primeira nota, à cada uma das demais da seqüência é atribuída uma das letras A, B, ou C conforme a altura da nota seja inferior, idêntica ou superior à anterior. Se utilizarmos um alfabeto de três letras para indicar as alturas das notas da seqüência dada no exemplo de consulta anterior (do3, sol2, la2, do4) chegaremos à seguinte seqüência: (-, A, C, C). Utilizando este método para representar um trecho da melodia de uma música, o sistema para recuperação de informação musical em banco de dados de áudio proposto por Ghias (1995) mostrou-se capaz de recuperar muitas melodias de doze notas dentre as 183 armazenadas. Para um volume pequeno de músicas a ser pesquisado, este método pode trazer a vantagem da rapidez, em detrimento da precisão. Contudo em banco de dados de grandes volumes de melodias, uma resolução maior seria necessária. Para isso o alfabeto inicialmente considerado deveria ser expandindo. Outras pesquisas em técnicas para reconhecimento de padrão melódico como em (Crawford 1997) sugerem o uso de técnicas empregadas no reconhecimento de padrão de cadeias de caracteres, devido a similaridade entre a análise da música tradicional e os problemas computacionais de reconhecimento de seqüências de caracteres. O músico interessado em análise musical está frequentemente preocupado em encontrar ocorrências de padrões ou repetições de um mesmo padrão, possivelmente com variações, da mesma forma que os algoritmos de computador têm que executar tarefas similares sobre cadeias de caracteres, as quais são seqüências de símbolos de algum alfabeto. Muitos objetos podem ser considerados como seqüências de caracteres: um arquivo texto ou também uma partitura musical. Ela pode ser vista como uma seqüência de símbolos de um alfabeto de notas. Desta

forma podemos trabalhar com estruturas polifônicas, formulando os seguintes tipos de consultas:

- i. Dado um conjunto de seqüência de notas (uma para cada voz), descubra se um subseqüência exata ocorre em uma das vozes.
- ii. Dado um conjunto de seqüência de notas (uma para cada voz), descubra se uma subseqüência exata ocorre em uma das vozes sem preservar o tempo de duração de cada padrão.
- iii. Dado um conjunto de seqüência de notas (uma para cada voz), identifique padrões de repetição em vozes diferente ou repetidos na mesma voz.
- iv. Dado um conjunto de seqüência de notas (uma para cada voz), e um padrão, descubra se ele ocorre em uma das seqüências dadas tanto na forma original, invertida, retrógrada ou com inversão retrógrada.

Utilizando o conhecimento rítmico da música pode-se apurar o reconhecimento de melodias similares, acrescentando à informação de altura, o acento característico do tempo em que a nota é ouvida. Segundo (Bakhmutova 1997) quase sempre fragmentos similares de uma música podem ser descobertos em outras a partir das chamadas repetições de segundo tipo ou adaptações. Através da representação de intervalos entre notas, um trecho musical pode ser representado como uma seqüência do tipo  $I_k S_k$ ,  $K = 1, 2, \dots, N - 1$ , onde  $I_k$  é o intervalo de alturas entre a  $K$ -ésima nota e a  $K+1$  e  $S_k$  representa a acentuação métrica dos sons. Esta acentuação é caracterizada pela divisão em compassos, que distribui a melodia em tempos fortes e fracos. Nesta representação tem-se a possibilidade de comparar duas melodias e extrair padrões de repetições não somente na variação de altura, mas também no elemento rítmico, que ajuda a definir melhor a melodia. Utilizando este método para comparar as duas seguintes seqüência de notas N2 e N3 à seqüência original N1, num compasso 6/8, teremos:

N1 : (nota)	do3,	sol2,	la2,	sol3,	la2,	sol2,	do3
(tempo)	1,2,	3	1,2,	3,	4,5,	6	1,2,3,4,5,6
(acentuação)	(F)		(F)		(F)		(F)

N2 : (nota)	sol3,	re3,	mi3,	re4,	mi3,	re3,	sol3
(tempo)	1,2,	3	1,2,	3,	4,5,	6	1,2,3,4,5,6
(acentuação)	(F)	-	(F)	-	(F)	-	(F)

N3 : (nota)	sol3,	re3,	mi3,	re4,	mi3,	re3,	sol3
(tempo)	1,2,	3	1,2,	3,	4,5,	6	1,2,3,4,5,6
(acentuação)	(F)	-	(F)	-	(F)	-	(F)

N1 :	(3-+)	(1+-)	(6++)	(6--)	(1-+)	(3+-)
N2 :	(3-+)	(1+-)	(6++)	(6--)	(1-+)	(3+-)
N3 :	(3-+)	(1+-)	(6++)	(6--)	(1-+)	(3+-)

A representação (3-+) presente no início dos três trechos musicais N1, N2 e N3, significa que a distância entre a primeira nota desta passagem (do3) e a próxima (sol2) equivale a três notas (si, la, sol). O sinal negativo indica a direção do intervalo, que no caso passa do mais agudo (do3) para o mais grave (sol2). O sinal positivo indica que a acentuação forte do compasso não recai sobre a última nota (sol2) do intervalo representado. No compasso de seis tempos utilizado no exemplo acima, o próximo tempo forte do trecho N1 cairá sobre a terceira nota (la2). Por esse motivo a representação do intervalo (sol2-la2) terá o sinal de acentuação negativo (1+-).

Neste tipo de representação mesmo utilizando-se a mesma seqüência de notas com variação apenas na distribuição rítmica, as seqüências N1 e N3 são consideradas apenas parcialmente similares, pois os dois últimos intervalos de N3 não apresentam semelhança rítmica com os de N1.

## 6 - Consultas por similaridade de atributos sonoros

Para usuários que estejam interessados em outros atributos da música diferentes da melodia, como compositores preocupados com aspectos timbrísticos do som, por exemplo, o sistema deve permitir a formulação de consultas por valores, levando em consideração as propriedades desejadas. Utilizando um vetor para armazenar as propriedades desejadas da consulta, pode-se excluir aquelas que o usuário considere desprezíveis em termos da sua finalidade e montar tipos de consultas para recuperação de sons a partir das suas propriedades.

- Consulta por Valor: Este tipo de consulta deve ser capaz de recuperar todos os sons cujos valores de uma certa propriedade  $P_i$  sejam maior, igual ou menor que um determinado valor.

Ex.: Recupere todos os sons cujos valores da propriedade  $P_0 \geq 0.5$  e propriedade  $P_1 < 0.3$

- Consulta por exemplos: Este tipo de consulta deve poder recuperar todos os sons similares a uma certa amostra, em relação à propriedade  $p_0$ . A similaridade neste caso pode ser obtida recuperando todos os sons dentro de uma faixa de valores para a propriedade  $p_0$ .

Ex.: Recupere todos os sons cuja propriedade  $P_0 \geq 0.3$  e  $P_0 < 0.5$

## 7 - Conclusão

Empregando formas de classificação que utilizam pesos diferenciados aos atributos da seleção, se soubermos a priori que algumas características acústicas não são importantes para a classe em questão, podemos ignorá-las ou atribuir-lhes um peso pequeno no cálculo da distância. Utilizando métodos que não dependem das características particulares de áudio e que podem ser aplicados à recuperação de imagem, como o proposto por (Foote 1997), consegue-se um bom resultado quanto à classificação de dados sonoros em relação à sua altura. Contudo a classificação segundo uma determinada característica, por exemplo, o



timbre, fica prejudicada, uma vez que este método não utiliza a atribuição de pesos a cada uma das propriedades tratadas. Na classificação de um som de oboé, por exemplo, este método tende a aproximar sons de alturas similares, podendo colocar na mesma classe outros instrumentos como trombone ou violoncelo, baseando-se na altura. Outra linha de pesquisa (Feiten 1994) tem abordado o problema da classificação em banco de dados de áudio utilizando redes neurais. Contudo, esta prática ainda apresenta alguns problemas devido ao fato de ser difícil olhar dentro da rede depois que ela é treinada ou durante a operação para determinar como o treinamento ocorre ou quais aspectos dos dados sonoros são similares a outros. Desta forma, torna-se difícil ao usuário especificar quais as características do som são importantes e quais aquelas que podem ser ignoradas para fins de classificação.

## 8 - Agradecimentos

Esta pesquisa vem sendo apoiada pelo CNPQ e pela Escola de Música da UFRJ. Meu especial agradecimento ao professor Rodolfo Caesar por seus valiosos comentários.

## 9 - Referências

- Bakhmutova, I.; Gusev, V. & Titkova, T. (1997). The Search for Adaptations in Song Melodies. *Computer Music Journal*, 21:1.
- Blum, Thom; Keislear, Doug; Wheaton, James & Wold, Erling. (1996). Audio Analysis for Content-Based Retrieval. *Tech. rep., Muscle Fish LLC, 2550 Ninth St., Berkeley, CA, USA.*
- Blum, Thom; Keislear, Doug; Wheaton, James & Wold, Erling (1996). Content-based classification, search and retrieval of audio. *IEEE Multimedia.*
- Crawford, Tim; Iliopoulos, C. S & Raman, Rajeev. (1998). String Matching techniques for Musical Similarity and Melodic Recognition. [http://www.kcl.ac.uk/kis/schools/hums/music/ttc/String\\_matching.html](http://www.kcl.ac.uk/kis/schools/hums/music/ttc/String_matching.html)
- Feiten, Bernard & Günzel, Stefan (1994). Automatic Indexing of a Sound Database Using Self-organizing Neural Nets. *Computer Music Journal*, 18:3.
- Foote, Jonathan. (1997). Content-Based Retrieval of Music and Audio, *Multimedia Storage and Archiving System II*, Proc. of SPIE, Vol 3229, pp 138-147.
- Ghías, A.; Logan, J.; Chamberlin, D. & Smith, B. (1995), Query by Humming - Musical information retrieval in an audio database. *ACM Multimedia*, Electronic Proceedings, San Francisco, Califórnia, 1995.
- Hourdin, Christophe; Charbonneau, Gérard & Moussa Tarek (1997). A Multidimensional Scaling Analysis of Musical Instruments's Time-Varying Spectra. *Computer Music Journal*, 21:2.
- Opren, Keith & Huron, David (1991). The Measurement of similarity in music: A quantitative approach for non-parametric representations. *Computers in Music Research*, n. 4.

- Lindsay, Adam (1998). Using Contour as a Mid-Level Representation of Melody. *S. M. Thesis*, MIT Media Laboratory, Cambridge, MA.
- McNab, Rodger; Smith, Lloyd; Bainbridge, B. & Witten, Ian (1997). The new Zealand Digital Library Melody Index. *In D-Lib Magazine.*
- Pierce, J. R. (1984). *Le Son Musical*. Editions Pour la Science - Paris.
- Schaeffer, Pierre & Guy Reibel. *Solfège de l'objet sonore*. Production du Groupe de Recherches Musicales de l'O.R.T.F., Editions du Seuil.
- Vercoe, Barry L. (1992). *Csound - Manual for the Audio Processing System and Supporting Programs*. M.I.T., Cambridge, Massachusetts.