

SeVEM - Separador de Vozes em Execuções Musicais Polifônicas

Rodolfo Daniel Wulfhorst, Eloi Fritsch, Rosa Maria Vicari

Instituto de Informática – Universidade Federal do Rio Grande do Sul (UFRGS)
Caixa Postal 15.064 – 91.501-970 – Porto Alegre – RS – Brazil

{wulfhors, fritsch, rosa}@inf.ufrgs.br

***Abstract.** One of the greatest problems a piano student is confronted with is the interpretation of polyphonic compositions. The SeVEM tool is a software that separates the voices from musicians execution at a digital piano. Listening to what he played in separated voices, the student will be able to realize and to change problems in his execution.*

***Resumo.** Um dos maiores desafios à interpretação pianística é a interpretação de peças polifônicas. Por isso, um dos pilares fundamentais da técnica pianística consiste no aprimoramento da execução de vozes de forma independente. A ferramenta SeVEM propõe isolar as vozes da interpretação do músico, reproduzindo-as separadamente. Ouvindo as vozes em separado, o estudante poderá identificar possíveis problemas na execução das vozes.*

1. Introdução

A interpretação de peças musicais polifônicas têm representado um grande desafio para tecladistas desde o séc. XVIII. A execução de peças com várias melodias simultâneas, considerando que o músico dispõe de apenas duas mãos, exige uma aprimorada técnica do intérprete. A dificuldade reside no fato de todas as melodias terem um significado especial na formação do caráter da peça. Deve haver, portanto, uma preocupação máxima em realizar uma performance que respeite e valorize a linha melódica de cada uma das vozes simultaneamente.

A avaliação da própria execução das melodias por parte do intérprete acaba sendo muitas vezes comprometida pela dificuldade de discernimento das vozes do executante. Na prática da execução, a atenção do intérprete acaba recaindo sobre as vozes que, em dado momento, são julgadas de maior interesse. Uma ferramenta que, a partir da interpretação do pianista, pudesse reproduzir sua própria execução de cada uma das vozes, seria de fundamental interesse para uma auto-avaliação mais detalhada e esmerada de sua interpretação.

Descrevemos neste trabalho a implementação desta ferramenta, denominada SeVEM (Separador de Vozes em Execuções Musicais Polifônicas). O SeVEM possibilita ao intérprete a reprodução de sua interpretação com vozes independentes. Desta forma torna-se possível uma análise exata e detalhada de sua performance. A reprodução de cada voz auxilia na localização de eventuais falhas na linha de canto existente em determinada voz. Outro recurso disponível é a reprodução de apenas algumas vozes, a fim de avaliar a interpretação dada aos diálogos existentes entre as linha melódicas. Uma última alternativa oferecida ao usuário é a reprodução de sua execução através de

instrumentos (timbres) distintos para cada voz. Essa alternativa destina-se principalmente aos iniciantes do estudo de peças polifônicas, oferecendo ao usuário uma visão do todo, facilitando o discernimento das vozes e um acompanhamento horizontal das mesmas.

Tais recursos oferecidos pela ferramenta tem como única função o auxílio à avaliação da interpretação. A percepção de eventuais falhas na execução caberá sempre ao próprio músico.

2. O Problema da Execução Polifônica

O termo “polifonia” já vêm sendo usado desde a Grécia Antiga - *polyphonia*: “que tem muitos sons, muitas vozes” (The New Grove Music Dictionary 1980). Visto como técnica de composição tinha originalmente o significado de “música com muitas vozes”, sendo uma antítese à música monofônica do século XV. A separação de vozes tendo por finalidade a avaliação da interpretação musical ganha sentido num contexto em que não existe uma voz específica que assuma um papel preponderante sobre o significado da música.

Um dos pilares fundamentais da técnica pianística consiste no aprimoramento da execução de vozes de forma independente. Mesmo no repertório romântico, impressionista e principalmente moderno, a independência melódica e rítmica é imprescindível para a performance. Porém, é no repertório barroco, mais especificamente na obra de J. S. Bach, que se encontram peças polifônicas na sua forma mais pura. Para Bach, “os efeitos de sonoridade têm pouca importância, enquanto que as idéias dominam tudo, constituindo uma música absoluta, sobrepassando muitas vezes as possibilidades do instrumento”, afirma Roethlisberger em *O Cravo na obra de Bach*. (Todd 1991).

3. A Separação de Vozes

A separação automática de vozes tem sido estudada com fins diversos. No entanto, para cada aplicação específica usam-se técnicas distintas. De forma geral, separam-se as vozes resultantes de uma performance, podendo ser capturada tanto pela entrada de um dispositivo MIDI quanto pela captura do próprio sinal de áudio. O método de reconhecimento das vozes está fortemente vinculado à forma com que se faz a captura da performance e ao propósito ao qual se destina a separação. Para dar uma visão geral sobre a abrangência desta aplicação e das peculiaridades de cada caso, propomos as seguintes formas de classificação.

Quanto ao propósito da separação

- **Transcrição** - Sistemas de Transcrição musical têm sido abordados com grande ênfase. Constituem ferramentas que extraem uma notação musical a partir de um sinal de áudio. Resultados satisfatórios têm sido obtidos para casos monofônicos. Mais recentemente tivemos resultados para a transcrição polifônica (Shuttleworth 1995).
- **Análise de Performance** - A ferramenta SeVEM propõe-se, inicialmente, a dar um suporte na avaliação da performance das linhas melódicas executadas pelo próprio

músico, separando as vozes executadas. Isto não impede que, no futuro, se desenvolva uma análise melódica de cada voz executada pela própria ferramenta.

Quanto à fonte de captura da performance

- **A partir de Sinal de Áudio** - O reconhecimento de notas tendo como fonte o sinal de Áudio tem sido um problema abordado amiúde por pesquisadores. A maior dificuldade tem sido representar o sinal de forma a facilitar a extração tanto da frequência (altura das notas), quanto das características temporais (tempo de duração das notas). Quando o problema se trata de peça polifônicas, a dificuldade se torna ainda maior, já que o sinal é composto de diversos instrumentos com timbres muitas vezes semelhantes.
- **A partir de eventos musicais** - O SeVEM utiliza-se desta forma de representação, devido à simplicidade de tratamento intrínseca ao protocolo MIDI

Quanto ao método de reconhecimento das vozes

- **Com uso de técnicas heurísticas** - nestes casos a interação com o usuário no processo é praticamente inevitável. Isto se torna mais evidente no caso da ferramenta SeVEM, em que as vozes possuem grande independência, não se podendo contar com uma performance fiel à polifonia da peça, já que o propósito da ferramenta é sua avaliação. Esta técnica, no entanto, é adotada principalmente em sistemas de transcrição, onde naturalmente se supõe não haver uma partitura disponível da peça.
- **Com uso de partituras** - Durante o projeto do SeVEM, optou-se pelo uso de partituras como subsídio na separação de vozes. Considerando a necessidade de uma separação exata das vozes e que o repertório polifônico para teclado do dia-a-dia do pianista é razoavelmente limitado, optou-se pelo uso desta técnica. O grande problema, no entanto, consiste em contornar as diferenças de execução do músico em relação a partituras, como variações de andamento, ornamentos etc.

4. SeVEM

A versão atual da ferramenta SeVEM, desenvolvida em C++, funciona atualmente para o Sistema Operacional Windows 9x. Como entrada do programa temos dois arquivos MIDI contendo a execução da peça. No entanto, uma é gerada pela execução do usuário, através do uso de um software seqüenciador. A outra contém a “partitura” da peça para uso do programa como referência. Estas “partituras” estão disponíveis como uma biblioteca, que contém atualmente peças polifônicas de J. S. Bach. Assim, o primeiro arquivo de entrada contém apenas um canal que compartilha as vozes. Já o segundo possui uma trilha para cada voz, informando ao programa a que voz cada nota pertence.

A saída consiste também num arquivo MIDI com um canal para cada voz. No entanto, as vozes estarão de acordo com a execução do pianista. A saída poderá ser aberta por um seqüenciador, permitindo ao músico a reprodução das vozes da peça executada de forma independente.

O processo de Separação de Vozes consiste de quatro etapas distintas. Na primeira etapa ocorre a leitura dos eventos do arquivo MIDI que contém a partitura da peça executada. As notas presentes em cada uma das trilhas (representando cada uma das vozes) são inseridas em listas separadas conforme a Figura 1.

Durante a próxima etapa, procura-se adaptar estas listas criadas de acordo com a execução do instrumentista (contido no arquivo MIDI gerado pelo usuário). Raramente o músico toca exatamente as notas que constam na partitura. Erros de execução também são tratados aqui. Notas tocadas a mais, a menos ou mesmo erradas, são colocadas numa trilha escolhida pelo algoritmo de acordo com o contexto. É importante ressaltar que no final do processo a estrutura mostrada na Figura 1 estará modificada de forma a conter todas as notas executadas pelo intérprete.

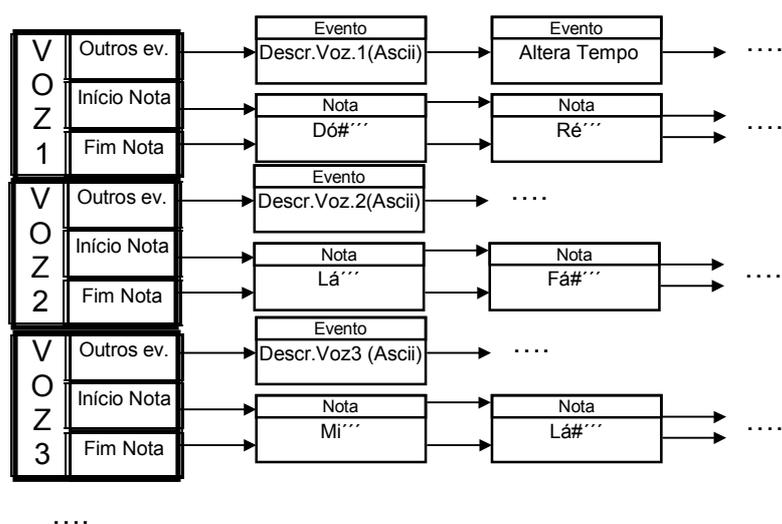


Figura 1: Representação interna da partitura

A rotina principal de *matching* é trivial. Construída a representação da partitura, inicializa-se um ponteiro para cada uma das vozes. Começa então a leitura do arquivo MIDI contendo a execução do usuário. No caso apresentado, espera-se que a primeira nota deste arquivo seja Dó#''', Lá''' ou Mi'''. Se a nota for Lá''', o ponteiro da segunda nota avança. Assim o novo conjunto de notas esperadas se torna {Dó#''', Fá#''', Mi'''}.

O algoritmo de “tolerância” é implementado por uma função recursiva que é chamada quando não se consegue reconhecer determinada nota executada de acordo com a partitura. Supondo que a próxima nota lida não faça parte deste conjunto (Fá'''), busca-se então uma solução por três caminhos distintos:

- Tentar trocar a nota por outra (supondo que tenha tocado por engano Fá''' ao invés de Fá#''')
- Inserir a nota na trilha em que a próxima nota estiver mais próxima (supondo que Fá''' seja um ornamento de Mi''' ou Fá#''' inserida a gosto pelo usuário, ou então que ele tenha tocado por engano uma nota a mais)
- Tentar eliminar a próxima nota de um trilha (a nota foi “esquecida” pelo intérprete). Esta alternativa é chamada para todas as trilhas, já que não se tem com prever de qual trilha ele esqueceu de tocar a nota.

Após um destes procedimentos ser efetuado, a rotina é chamada recursivamente. Se o retorno for verdadeiro, a alteração é confirmada. Caso contrário, tenta-se a próxima alternativa. O objetivo é que através das recursões (suposições) o sistema consiga

ressincronizar o arquivo de execução com a representação interna da partitura, encontrando novamente notas de acordo com a rotina principal de processamento. O espaço de busca é limitado através de um número máximo de níveis de recursão. O nível de tolerância obtido por este algoritmo é razoavelmente satisfatório. Futuramente deveremos testar métodos que utilizem técnicas de IA para efetuar esta etapa.

Na etapa seguinte, são lidos os atributos dos eventos da execução da peça, substituindo os atributos das notas já lidas na primeira etapa. Esta etapa é bastante simples porque a representação interna da partitura já está modificada de acordo com a execução da peça.

Finalmente ocorre a geração do arquivo MIDI de saída, contendo a execução da peça num formato de múltiplas trilhas. Esta etapa dá-se através da leitura trilha-por-trilha da representação da execução com vozes gerada pelo programa.

4.1. Biblioteca de peças do SeVEM

Para realizar a separação de vozes da execução de uma peça, é imprescindível para a ferramenta SeVEM que a mesma peça esteja contida na sua biblioteca.

A biblioteca é composta por arquivos no formato MIDI (SMF - Standard MIDI File) contendo uma representação das partituras das peças polifônicas armazenadas. O requisito básico para poder efetuar a separação de vozes é que cada voz esteja contida numa trilha independente no arquivo MIDI. Para dar flexibilidade ao reconhecimento, é necessário que o arquivo contenha apenas as notas essenciais da peça (sem ornamentos).

Atualmente a biblioteca SeVEM consta de algumas peças polifônicas para teclado de Bach. A extensão dessa biblioteca no futuro é imprescindível para um aproveitamento prático da ferramenta. As peças contidas no momento são:

- Invenção a 3 Vozes No. 5 (3 vozes)
- Gavotte da Suíte Francesa No. 5 (3 vozes)
- Prelúdio em Dó Menor do Cravo Bem Temperado (Vol. II No. 2)
- Fuga a quatro vozes em Mi b Maior do Cravo Bem Temperado (Vol. II No. 7)
- Minueto do Pequeno Livro de Anna Magdalena Bach (Si b Maior)

5. Trabalhos Futuros

Atualmente a ferramenta SeVEM encontra-se num estágio de implementação em que tem aplicação prática, oferecendo valioso auxílio na avaliação da performance pelo próprio músico. Na entanto, existem muitas possibilidades a serem consideradas num futuro desenvolvimento da ferramenta. Podemos expandir a capacidade do SeVEM em duas direções principais:

- Otimização dentro da proposta atual da ferramenta em separar vozes para auto avaliação do músico.
- Oferecer recursos adicionais que avaliem automaticamente a performance.

No que tange à primeira proposta, temos ainda algumas implementações importantes a fazer. A principal delas certamente é o aprimoramento do algoritmo de reconhecimento de diferenças na execução, dando maior flexibilidade à interpretação do músico, sem no entanto comprometer a separação correta das vozes. A expansão da biblioteca é trabalhosa, porém inevitável, para que o usuário não tenha de ser incumbido dela.

A ampliação do escopo atual de atuação do SeVEM, ou seja, inclusão de recursos de avaliação automática da performance, já conta com algumas propostas:

1. Análise de linhas melódicas
2. Análise de contraponto
3. Análise da regularidade da execução (ritmo, andamento, etc..)

6. Conclusão

A ferramenta atualmente implementada já apresentou bons resultados práticos na avaliação de execuções de composições de Bach. Falhas de execução melódicas que muitas vezes eram imperceptíveis ao se escutar o todo, tornaram-se evidentes quando avaliadas com o auxílio da ferramenta SeVEM. Em geral as falhas que não são percebidas durante a execução são aquelas que se encontram nas vozes intermediárias, geralmente relegadas a um segundo plano.

A partir de experiências com usuários, a ferramenta tem ajudado a identificar principalmente problemas em relação à articulação. Em geral tais problemas foram resolvidos através de alterações simples no dedilhado.

É importante ressaltar que não basta o estudante tomar consciência de todas as vozes. Muitas vezes, mesmo conhecendo as vozes, durante a execução acaba preponderando apenas aquela que se julga mais importante num dado momento. Nestes casos esta ferramenta visa justamente fazer com que o próprio executante verifique se ele realmente está conseguindo transmitir suas intenções musicais através de todas as vozes.

7. Bibliografia

Shuttleworth, T. & Wilson, R. (1995) The Recognition of Musical Structures using Neural Networks. *IJCAI-95 Workshop Program Working Notes, Montreal.*

The New Grove Music Dictionary of Music and Musicians. (1980) *London: Macmillan Publishers Limited.*

Todd, P & Loy, G. (1991) Music and Cognition. *The MIT Press: Cambridge, Massachusetts, London, England, 1991.*