

Sistema de Reconhecimento Automatizado de Notas Musicais Isoladas

M. M. Cordeiro Jr.

Universidade Federal de Pernambuco, Departamento de Eletrônica e Sistemas.
Recife, Pernambuco, 50.711-970, Brasil E-mail: mauriciojunior19@hotmail.com

H.M. de Oliveira

Universidade Federal de Pernambuco, Departamento de Eletrônica e Sistemas.
Recife, Pernambuco, 50.711-970, Brasil E-mail: hmo@ufpe.br

RESUMO

Este trabalho apresenta uma técnica alternativa para detecção e reconhecimento de notas isoladas produzidas por instrumentos musicais [1-2]. O método explora a estrutura “ataque”, “platô” e “decrecendo” para construir um diagrama de estados. Os arquivos de análise são de formato “.wav”, monofônicos, *signed*, com taxa de amostragem de 44.100 kHz. Usa-se uma FFT para blocos de comprimento 2.205 (correspondendo a uma janela de 50 mseg), determinando-se o espectro discreto, retendo-se apenas uma única componente espectral de maior energia. O reconhecimento das notas [3] se faz em escala temperada (incluindo # e b), através de um mapeamento de identificação tonal, introduzido aqui. O sistema para reconhecimento de notas musicais isoladas consiste em um programa desenvolvido em Matlab® capaz de identificar uma nota musical pura (DÓ-RÉ-MI-FÁ-SOL-LÁ-SI), assim como, seus acidentes musicais, Sustenido ou Bemol, sendo um total de 12 notas em qualquer escala que se repete à medida que se dobra a frequência fundamental da mesma [1]. Uma função interna disponível no Matlab® avalia a densidade espectral via uma transformada rápida de Fourier, retornando as amplitudes e frequências, respectivamente, do sinal. Para a análise utilizou-se uma precisão de 16 bits por amostra, ou seja, 65.536 níveis de quantização. Como o espectro é expresso pelas componentes da transformada discreta de Fourier armazenadas em um vetor, determina-se o ponto de maior amplitude e sua frequência referenciada. A identificação da nota

é feita via a fórmula
$$\chi(f) := \text{floor} \left[12 \cdot \left[\left(\frac{\ln(f)}{\ln(2)} \right) - \text{floor} \left(\frac{\ln(f)}{\ln(2)} \right) \right] \right]$$
 em que *floor* é a função piso, i.e., o maior inteiro menor que o argumento.

A variável codmusical $\chi(f)$ recebe um valor de 0 a 11. Por exemplo, a frequência $f=575$ Hz corresponde a qual nota? Tem-se $\chi(f)=2$, donde nota = Ré.

Tabela 1 Função identificadora de notas musicais em função da frequência detectada.

Cifra	C	C#	D	D#	E	F	F#	G	G#	A	A#	B
Nota	Dó	Dó#	Ré	Ré#	Mi	Fá	Fá#	Sol	Sol#	Lá	Lá#	Si
CodMusical	0	1	2	3	4	5	6	7	8	9	10	11

Cada janela foi definida com o comprimento de 50 ms (2205 amostras) que foi subdividida em três partes iguais de 16,67 ms (735 amostras) denotadas por Borda de Subida, Parte Central e Borda de Descida, como segue (Figura 1):



Figura 1 Estrutura da janela de 50 ms para o método de detecção nota a nota (borda inicial, parte central, borda de descida). A idéia é explorar o “ataque”, “platô” e “decrecendo”.

Cada janela é analisada individualmente a partir da função (*Condx.m*), calculando-se alguns parâmetros como média da amplitude das amostras de toda janela, média da borda de subida e média da borda de descida. O programa principal já com a variável que contém as amostras

absolutas e normalizadas do sinal carregado na memória, chama a função COND1 do Matlab[®] que percorre o arquivo janela a janela, testando se a média amostral de cada janela é menor ou igual 0,1; se menor que este valor, considera-se que a janela está percorrendo uma região de silêncio ou baixo ruído, caso contrário, COND1 não sendo satisfeita, o sistema avança para o *Estado B* e neste momento é marcado o início da nota. No *Estado B* é carregado a COND2 que agora testa se a média amostral é maior ou igual a 0,1 e a razão da média da borda de descida sobre a da borda de subida é menor ou igual a 2, permanecendo neste estado enquanto esta condição for satisfeita. Do contrário, avança-se para o *Estado C* que chama COND3 a fim de testar se a média amostral é maior ou igual a 0,05 e a razão da média da borda de descida sobre a da borda de subida é menor ou igual a 1,25. O sistema permanece no *Estado C* enquanto a condição for satisfeita; caso contrário, o sistema retorna ao *Estado A*, marcando agora o final da nota. A variação da duração dos tons produzidos constitui um dos obstáculos no sucesso do reconhecimento dos tons. Para a identificação do *Estado A* (silêncio ou ausência de notas), o qual corresponde à concatenação de notas sucessivas isoladas, propõe-se uma nova abordagem.

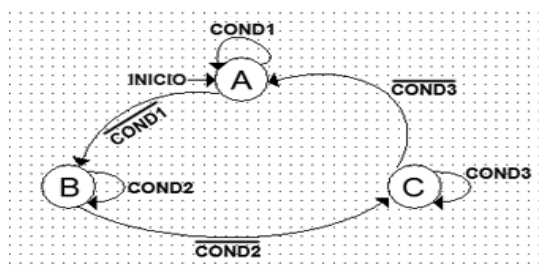


Figura 2 Diagrama de estados para as condições de identificação do trecho. O estado A corresponde ao "ataque" (início de uma nota). O estado B corresponde ao "plató" (parte de estado permanente em que a nota soa) e o estado C corresponde ao "decrecendo" (final de uma nota). A duração de uma nota particular, oriunda de um instrumento específico, tem tempos de estada diferentes em cada estado. A maneira de controlar o tamanho corresponde ao tempo de visitação em um estado da cadeia.

Uma grande quantidade de testes foi conduzida com arquivos gerados através do *software VirtualPiano*[®], simulando diversos instrumentos (com taxa de acerto acima de 99%). Implementou-se toda a abordagem usando plataforma Matlab[®], com interface gráfica, permitindo gerar, gravar e analisar arquivos de tons produzidos por diferentes instrumentos musicais. Aqui, explora-se a estratégia de diagrama de estado para lidar com a duração variável das notas emitidas. A definição (e uso) do mapeamento de identificação tonal $\chi(f)$ é também inédita. A abordagem é bem mais simples do que aquela usando cepstrum [4] ou cadeias de Markov [5], comuns também em identificação de instrumentos. Atualmente, investiga-se uma extensão desta técnica para o reconhecimento de acordes e para transcrição de partituras.

Palavras-chave: *análise musical, reconhecimento de notas, automação musical.*

Referências

- [1] F.R. Moore, "Elements of Computer Music", Prentice Hall, Englewood Cliffs, 1990.
- [2] A. Mitre, M. Queiroz, Um sistema automático de transcrição melódica, Simpósio Brasileiro de Computação Musical, 2005.
- [3] L. Ortiz, F. Javier, Polyphonic transcription using piano modeling for spectral pattern recognition, Int. Conf. on Digital Audio Effects, 2002.
- [4] A. Eronen, A. Klapuri, Musical instrument recognition using cepstral coefficients and temporal features. IEEE Proc. Acoustics, Speech, and Signal Processing. ICASSP 2000.
- [5] J-J. Aucouturier, M. Sandler, Segmentation of musical signals using hidden Markov models, AES 110th Convention, May 12–15 Amsterdam, The Netherlands, 2001.