

Universidade do Minho

Processamento Digital de Sinal

Caracterização espectral de sons vozeados



Ricardo Jorge Martins de Amorim
61998 MIEEIC

Índice

Introdução.....	3
Desenvolvimento.....	4
“Ricardo”	5
“Piano”	6
“Maria”	7
“farinha”	8
“lápis”	9
“agoniado”	10
Análise e estatística de resultados	11
Conclusão/Crítica	13

Introdução

No âmbito da disciplina de Processamento Digital de Sinal foi realizado um trabalho visando estudar e tirar as devidas conclusões acerca de frequências predominantes que caracterizam os sons vozeados. Este estudo foi feito à letra “i” visto ser a primeira vogal do meu nome.

Mais concretamente este trabalho visa tirar ilações sobre a consistência estatística (valor médio e variância) das frequências predominantes do espectro da letra “i” em diferentes palavras.

O principal objectivo deste trabalho é aplicar o algoritmo FFT (“fast Fourier transform”) leccionado durante as aulas da Unidade Curricular, recorrendo ao software Matlab.

Para que seja possível perceber e dar sentido ao trabalho realizado foi necessário fazer um pequeno estudo sobre alguns conceitos que até então eram desconhecidos. Segundo a teoria, cada vogal tem um conjunto de frequências predominantes (grande amplitude) denominadas formantes, que a caracterizam. Pode dizer-se que estas frequências são como que uma “assinatura” que permite identificar/caracterizar uma vogal.

Por definição, a informação necessária para que o ouvido humano consiga distinguir vogais pode ser representada quantitativamente pela frequência do som da vogal.

Desenvolvimento

Inicialmente foi necessário fazer uma selecção de palavras que tivesse na sua constituição a letra “i”. Para o efeito decidiu-se pelo seguinte grupo de palavras:

- “Ricardo”, “Maria”, “piano”, “farinha”, “lápis” e “agoniado”.

A frequência de amostragem usada foi de 8kHz e para cada palavra recolheu-se um total de 16000 amostras o que se reflecte num tempo de 2 segundos de sinal. Assim foi possível recolher e representar por completo o sinal respectivo de cada palavra. Em linguagem do software Matlab traduz-se em.

```
sinal=wavrecord(16000,8000)
```

Para cada sinal foi necessário encontrar o intervalo de tempo em que a letra em estudo se encontrava. Para o efeito, reproduziu-se o sinal em diferentes intervalos de tempo até se encontrar o pretendido.

```
sound(sinal(t1:t2)); em que t1 e t2 representam 2 instantes de tempo.
```

Uma vez encontrada a parte do sinal pretendida passou-se para o domínio das frequências. Para isso utilizou-se o algoritmo da fft (fast Fourier transform) do Matlab

```
fsinal=abs(fft(sinal(t1:t2)))
```

Com isto, foram reunidas as condições para retirar informação sobre o espectro do sinal sobre as formantes. Para visualizar o espectro:

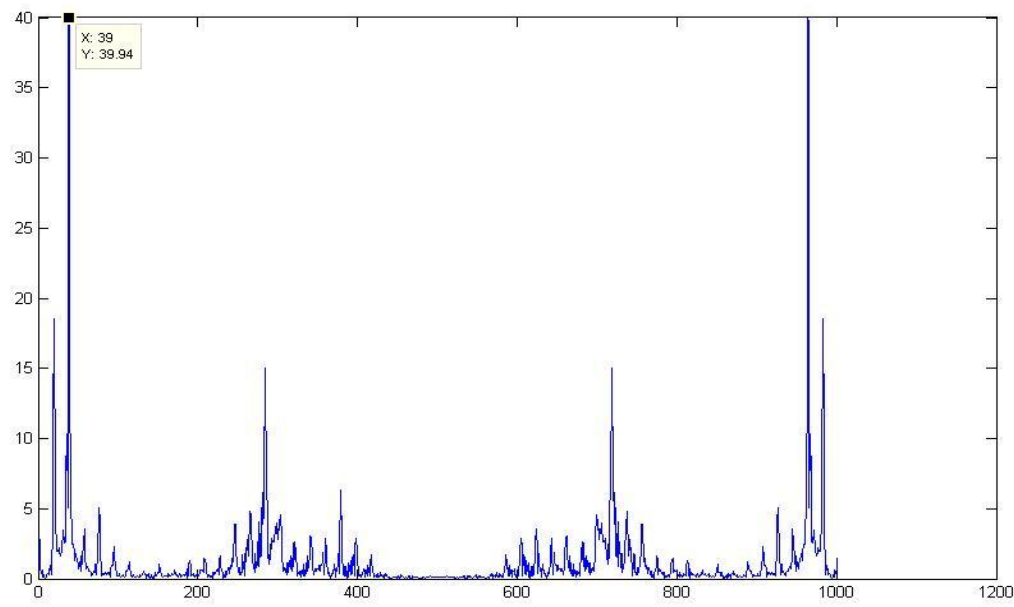
```
plot(fsinal)
```

Os máximos poderiam ter sido retirados correndo à função *max* do Matlab. No entanto, visto que os máximos eram facilmente identificáveis, recorreu-se ao *data cursor* do software para que o processo fosse mais rápido.

Todos estes passos foram repetidos para as diferentes amostras. Desta maneira foi possível identificar as formantes do sinal bem como determinar o seu valor no que diz respeito à frequência.

“Ricardo”

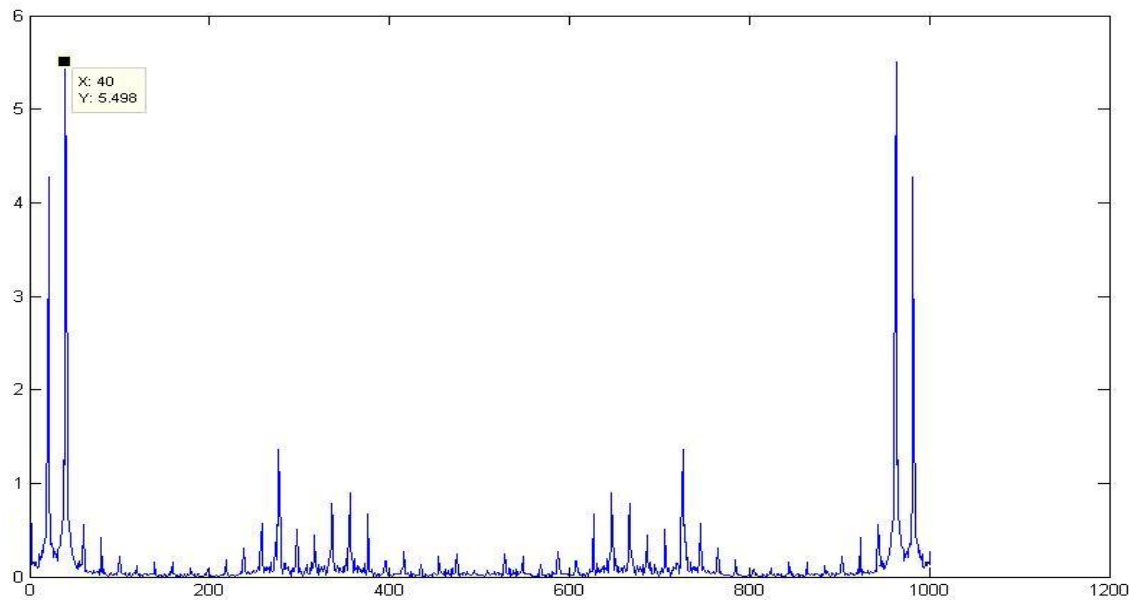
Ilustração 1 - Espectro da letra "i" em Ricardo



	Frequência no gráfico	Frequência convertida (Hz)
F1	39	312
F2	284	2272
F3	379	3032

“Piano”

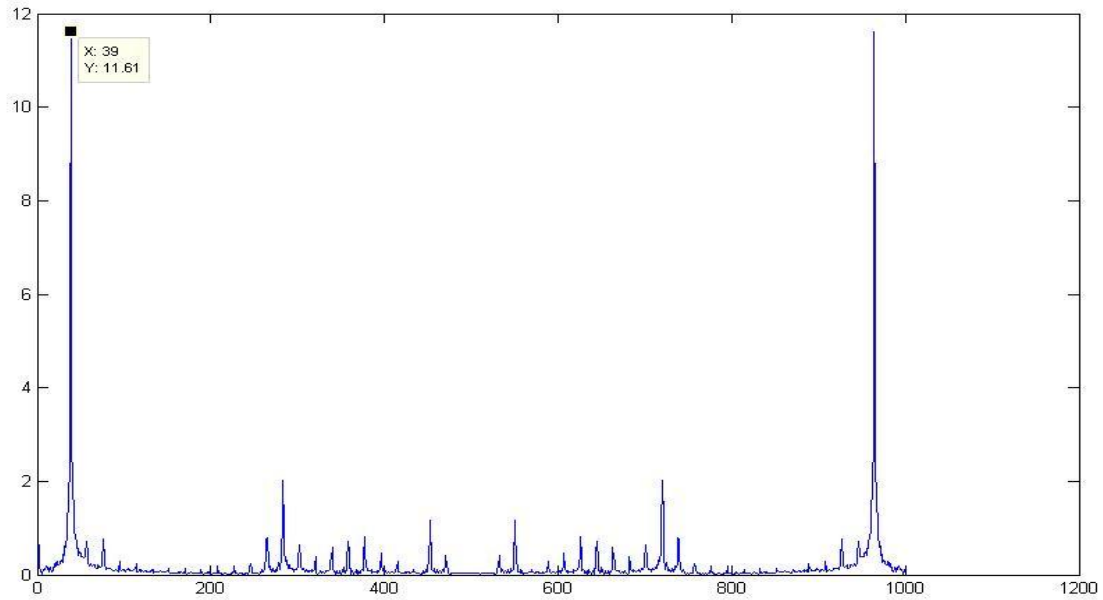
Ilustração 2 - Espectro da letra "i" em Piano



	Frequência no gráfico	Frequência convertida (Hz)
F1	40	312
F2	277	2216
F3	356	2848

“Maria”

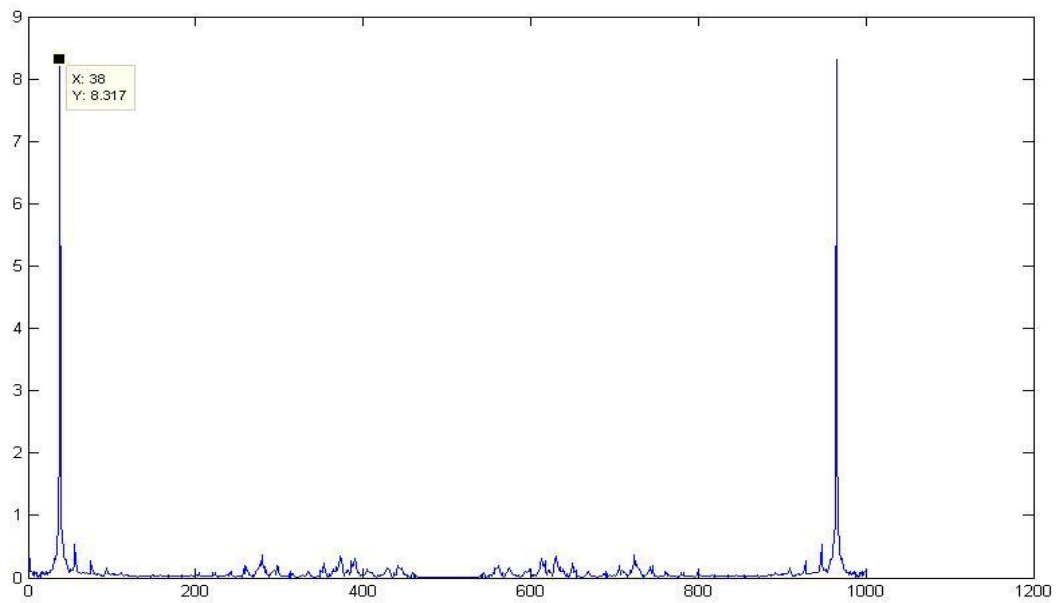
Ilustração 3 - Espectro da letra "i" em Maria



	Frequência no gráfico	Frequência convertida (Hz)
F1	39	312
F2	283	2264
F3	377	3016

“Farinha”

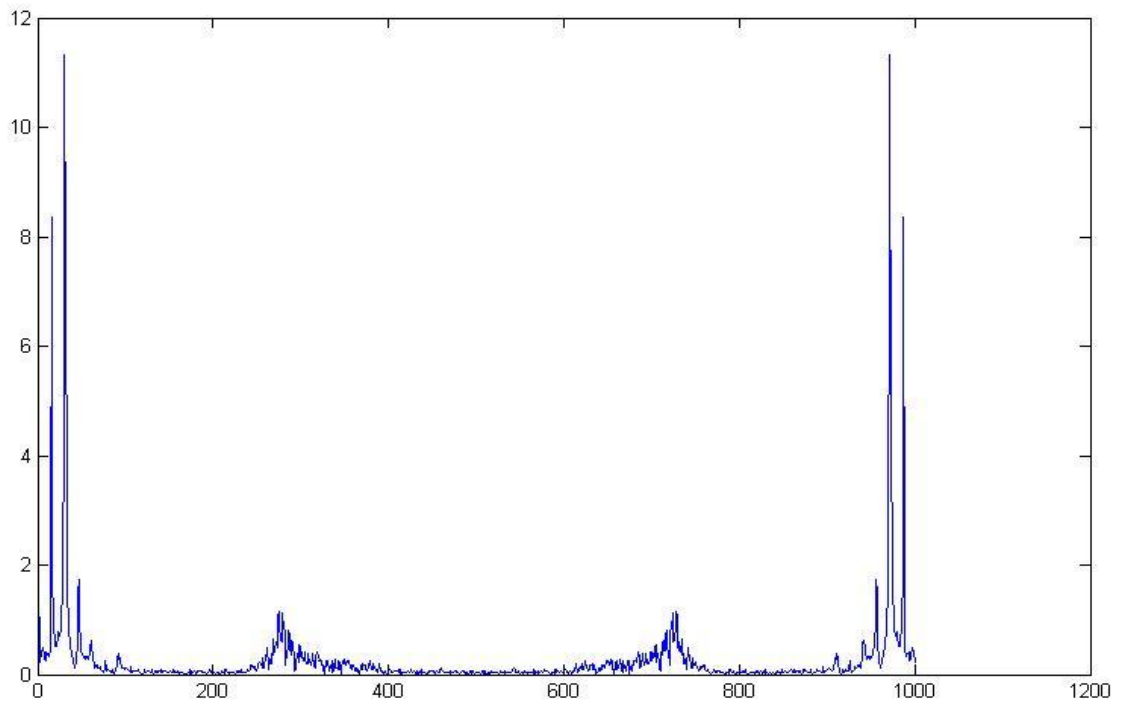
Ilustração 4 - Espectro da letra "i" em farinha



	Frequência no gráfico	Frequência convertida (Hz)
F1	38	304
F2	279	2232
F3	373	2984

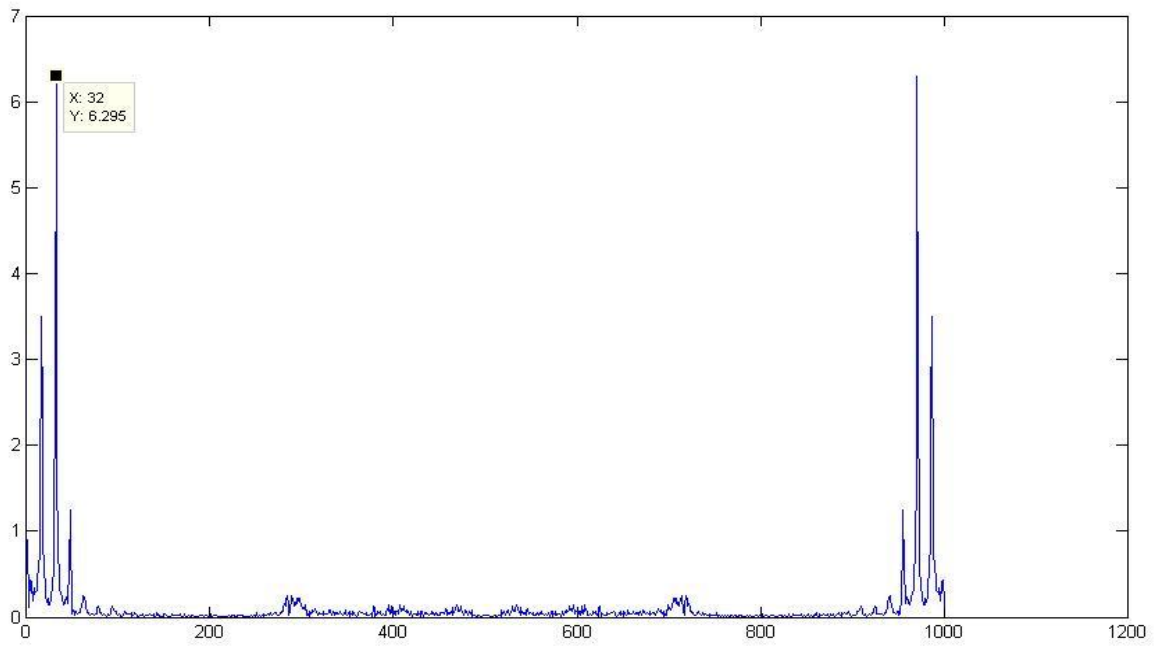
“Lápis”

Ilustração 5 - Espectro da Letra "i" em lápis



	Frequência no gráfico	Frequência convertida (Hz)
F1	31	248
F2	275	2200
F3	379	3032

“Agoniado”



	Frequência no gráfico	Frequência convertida (Hz)
F1	32	256
F2	284	2272
F3	380	3040

Análise e estatística de resultados

	1ª Formante	2ª Formante	3ª Formante
Ricardo	312 Hz	2272 Hz	3032 Hz
piano	320 Hz	2216 Hz	2848 Hz
Maria	312 Hz	2264 Hz	3016 Hz
farinha	304 Hz	2232 Hz	2984 Hz
lápiz	248 Hz	2200 Hz	3032 Hz
agoniado	256 Hz	2272 Hz	3040 Hz
VALOR MÉDIO	292 Hz	2242,67 Hz	2992 Hz
VARIÂNCIA	992	964,27	6240

Olhando para os dados da tabela, verifica-se que, no que diz respeito à 1ª formante, as palavras estudadas podem dividir-se em 2 grupos:

- “Ricardo”, “piano”, “Maria”, “farinha”.

- “lápiz”, “agoniado”

No primeiro grupo a frequência da primeira formante é maior. Analisando com alguma atenção verifica-se que em “Maria” e “farinha” a letra “i” faz parte da sílaba acentuada da palavra. Nos caso de “Ricardo” e “Piano” a letra “i” não faz parte da sílaba acentuado, no entanto em ambos os casos encontra-se no início da palavra.

Já no segundo grupo é evidente que a frequência da primeira formante das 2 palavras se desvia da gama dos 300-320Hz do grupo anterior. Uma das razões que pode forçar a que isto aconteça é o facto da letra “i” se encontrar a meio e no fim respectivamente das palavras “agoniado” e “lápiz”. Em ambos os casos a letra “i” não faz parte da sílaba acentuada.

No que diz respeito à segunda formante e estando a falar em frequência dos 2200-2300Hz, verifica-se que as variações são pequenas, não sendo possível fazer uma apreciação rigorosa de factores que possam condicionar os respectivos valores.

Olhando para as frequências da 3ª formante mais uma vez não se verificam grandes variações à excepção do caso “piano”. Uma possível razão para explicar o fenómeno pode ser o facto de “i” anteceder outra vogal (“a”).

Tabela 3. Médias dos valores das frequências dos harmónicos correspondentes aos três primeiros formantes (F1, F2, F3), em Hz, para cada vogal, para ambos os sexos.

VOGAIS	MULHERES			HOMENS		
	F1	F2	F3	F1	F2	F3
/a/	1002,90	1549,95	2959,70	753,87	1278,70	2483,44
/ε/	672,45	2242,93	3018,60	588,44	1745,11	2566,00
/e/	437,03	2429,76	3087,09	406,63	1955,60	2540,33
/i/	361,90	2583,89	3378,14	297,80	2150,85	2925,14
/ɔ/	715,34	1073,27	2981,69	580,15	947,25	2525,52
/o/	444,89	914,26	2899,80	411,62	832,84	2376,13
/u/	461,82	763,41	2902,55	345,27	799,51	2351,50

A tabela acima representa um conjunto de valores de frequências dos formantes de diversas vogais. Esta tabela foi retirada de um estudo conjunto realizado por diversos professores de várias Universidades Brasileiras. Como é possível verificar os valores das frequências dos formantes realizados neste trabalho são muito aproximados com os da tabela.

Conclusão/Crítica

Com este trabalho foi possível concluir que olhando para o espectro do sinal da vogal “i” se verifica que as frequências denominadas de formante tomam valores próprios que caracterizam a letra. Sendo assim, pode dizer-se que é possível identificar/distinguir a vogal olhando para o seu espectro assumindo que as outras vogais irão também ter frequências formantes próprias.

Foi também perceptível neste trabalho a importância e a utilidade do algoritmo fft (fast Fourier transform) que foi referido nas aulas teóricas. Na necessidade de analisar um sinal na gama das frequências é uma ferramenta indispensável.

Apesar do balanço do trabalho ser bastante positivo, é necessário salientar alguns aspectos menos bons que podem condicionar a qualidade do trabalho e/ou fiabilidade dos resultados. Para realizar a recolha dos sinais foi utilizado um microfone do computador portátil. Assim, não se pode esperar que o sinal recolhido não tenha algum ruído proveniente da falta de qualidade do microfone bem como do ambiente onde foi gravado.