



Universidade Federal de Uberlândia

– Trabalhos de Processamento Digital de Sinais –

Prof. Alan Petrónio Pinheiro

Faculdade de Engenharia Elétrica

Curso de Engenharia Eletrônica e de Telecomunicações (*campus* Patos de Minas)

Versão 1.0 – janeiro de 2015

1) Sobre os projetos de PDS

A fim de se demonstrar algumas das aplicações de Processamento Digital de Sinais (PDS) e proporcionar ao discente uma possibilidade extra e prática de refinar seus conhecimentos, é proposto ao aluno ao final da disciplina um projeto envolvendo o processamento (digital) de algum tipo de sinal(is) em uma aplicação qualquer. Além dos objetivos apontados anteriormente, espera-se com que com este projeto o aluno possa buscar - por meios próprios - novas informações (quer dizer: pesquisar!) que possam complementar o conteúdo aprendido e ao mesmo tempo consolidar estes conhecimentos e incentivá-lo ao treino das “ferramentas” aprendidas. Vale destacar mais uma vez que a maior parte dos projetos devem exigir conhecimentos que vão além da ementa da disciplina e por isto exigem uma pesquisa. Com isto espera-se averiguar não só a capacidade de pesquisa do aluno para solução de problemas, mas principalmente a sua capacidade de **auto-aprendizagem**.

O aluno tem liberdade de propor ao professor qualquer tipo de projeto que empregue processamento digital de sinais (em aplicações multidisciplinares ou não). Os projetos podem envolver *hardware/software* ou somente simulações (em Matlab). Considerando a dificuldade de implementar as sofisticadas técnicas de DSP em processadores comuns (e este não é o foco desta disciplina), a maior parte das sugestões podem ser feitas em Matlab permitindo averiguar (por simulação) com relativa facilidade a técnica e a solução do problema proposta. De qualquer forma, o aluno entusiasta que desejar aplicar alguma destas técnicas usando microeletrônica, pode fazê-la através de consulta ao professor.

2) Sugestões de projetos

Na sequência desta são feitas algumas sugestões de projetos. O aluno deve escolher uma delas e indicar ao professor que deverá dar maiores detalhes e requerimentos. Durante o projeto, o docente se disponibiliza a ajudar dando sugestões sobre leituras, avaliações de resultados e indicação de possíveis erros. Contudo, cabe ao aluno a execução do projeto e a pesquisa.

- a) **Decodificador DTMF**: este projeto deve converter um registro de “tons de discagem” de um telefone real nos correspondentes dígitos. Para testar sua técnica, você pode usar os sons disponíveis na página da disciplina ou em www.ee.columbia.edu/~dpwe/sounds/. Se desejar, incremente o projeto fazendo a análise em tempo real: ou seja, cada teclada digitada em um telefone deve ser decodificada logo após a tecla ser digitada.
- b) **Equalização de canal**: neste projeto um sinal é submetido a alguma distorção desconhecida. Esta distorção pode ser representada por um filtro que representa as características de um canal que distorce algumas componentes de frequência do sinal que passa pelo canal de comunicação. O

objetivo é avaliar o que este canal (ou filtro, uma vez que ele pode ser representado matematicamente por um comportamento de filtro) faz no sinal e tentar compensar seus efeitos. Deste modo, você deve fazer um filtro que compensa as distorções do canal e este filtro de compensação é usado no canal de comunicação para tentar desfazer algumas das distorções. Para isto, você pode usar alguns dos exemplos de sons disponíveis na página da disciplina (ver arquivos “sinal 0? - versao original.wav” e “sinal 0? - versao distorcida?.wav”) contendo o som original e algumas versões diferentes do som distorcido. É importante você conhecer o som original (i.e., sem distorções) e seu espectro e também o espectro do sinal corrompido e projetar um filtro para fazer o espectro do sinal corrompido tornar-se o mais próximo possível do original desfazendo assim a maior parte dos efeitos do canal.

- c) **Supressão de ruído dos sinais:** sinais podem ser corrompidos por diferentes meios e fontes. Na página da disciplina tem-se alguns sinais de voz que foram artificialmente corrompidos por diferentes tipos de ruídos. A ideia básica deste projeto é minimizar este ruído ao máximo a partir da avaliação do espectro do ruído. No caso de ruídos de amplo espectro, é necessário o uso de filtros específicos (como, por exemplo, o filtro de Wiener) que buscam otimizar a relação sinal-ruído (SNR). Além das técnicas tradicionais, existem outras mais sofisticadas (e muitas vezes mais adequadas) para a supressão de ruídos (principalmente as de largo espectro) de sinais. Uma destas abordagens é a descrita no livro da referência [1], especialmente na seção 8.3.2.
 - d) **Extração de pitch:** um problema muito conhecido (e ainda estudado) na área de processamento de sinais (especialmente os de fala) e diz respeito ao reconhecimento do “pitch” da voz de uma pessoa. O pitch de um sinal refere-se a periodicidade de um determinado som sustentado (constante, como o de uma vogal pronunciada repetidamente) que caracteriza muitas vezes a voz de uma pessoa. Ele indica não só a frequência fundamental, mas as próprias variações de período e amplitude (conhecidas como jitter e shimmer) deste sinal que são indicadores importantes para a caracterização da voz. Há vários modos de se fazer isto. O mais clássico deles é a auto-correlação. Outras abordagens podem ser consultadas na internet ou em um livro clássico de processamento de sinais de voz referenciado em [2].
 - e) **Estimação do atraso de tempo e ângulo de chegada de um som:** nós usamos os 2 ouvidos para detectar a direção de um som que chega. A pequena diferença de tempo acontece devido a velocidade finita do som que viaja em um determinado sentido a alcança as orelhas em diferentes instantes de tempo. Através desta pequena diferença de tempo, é possível identificar o sentido do som através de sua abertura angular. Porém, para isto, é preciso sinais gravados por pelo menos 2 microfones com alguma distância entre si gerando assim sinais stereo (ou [binaural](#)). Procure por registros de sons binaural simulados e algoritmos para estima a diferença de tempo entre os canais e então a direção da fonte de som. Veja o vídeo disponível na referência [3].
 - f) **Reconhecimento de fonemas:** neste trabalho é usado alguns sinais de sons e fonemas cuja análise da frequência fundamental (e algumas harmônicas) possibilita determinar se o falante é adulto do sexo feminino ou masculino ou criança. Acredita-se que falantes da língua portuguesa, a frequência fundamental para homens podem variar de 80 a 150 Hz. Já as femininas estão na faixa de 150 a 250 Hz e as infantis encontram-se acima de 250 Hz. Obviamente que está não é uma regra que se aplique a todos os casos. Uma das formas mais empregadas para obter automaticamente a frequência fundamental e usando a transformada “cepstrum” que, de forma resumida, é a transformada inversa do logaritmo da amplitude do espectro do sinal. Se o espectro
-

do sinal tiver componentes harmônicos espaçadas regularmente, sua transformada cepstrum apresentará um pico no valor que indica o espaçamento entre estas harmônicas produzindo assim o período fundamental da voz e conseqüentemente sua frequência. Outra forma de caracterizar o som dos vogais (para homens e mulheres) é através dos formantes segundo a Figura 1. Neste trabalho deve-se classificar (de forma semi-automática) as vogais e o tipo de voz (masculino e feminino) usando algum critério como, por exemplo, a análise de frequência fundamental e formantes. Faça uma pesquisa para estudar as características dos sons da língua portuguesa e o empregue para esta classificação. Sugestão: se for usar formantes, procure uma técnica para suavização do espectro (“spectral smoothing”) para extrair sua envoltória ou consulte [4].

VOGAIS	MULHERES			HOMENS		
	F1	F2	F3	F1	F2	F3
/a/	1002,90	1549,95	2959,70	753,87	1278,70	2483,44
/u/	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
/o/	715,34	1073,27	2981,69	580,15	947,25	2525,52
/ai/	444,89	914,26	2899,80	411,62	832,84	2376,13
/au/	461,82	763,41	2902,55	345,27	799,51	2351,50

Figura 1 – Tabela descrevendo a média dos valores das frequências dos harmônicos correspondentes aos três primeiros formantes F1, F2 e F3. Um formante é um dos picos harmônicos do espectro de uma vogal e é dado pela relação $F_n=(2n+1).f_s$ onde $n=1,2,3...$ e f_s é a frequência fundamental da vogal.

Outras sugestões que também podem ser boas alternativas de projeto. São elas:

- o “Detecção de início e fim de palavras para reconhecimento de voz”: para maiores informações, procure ler o artigo da referência [5].

Observações relevantes:

- 1) Para facilitar, em alguns projetos já é disponibilizado ao aluno arquivos para testes (de som, principalmente) que podem ser obtidos na página desta disciplina: www.alan.eng.br/disc_dsp/
- 2) A maior parte dos exemplos anteriores utiliza sinais sonoros. Isto porque, pelo fato de serem audíveis e por estarmos familiarizados com eles, a análise e interpretação é facilitada. Contudo, as técnicas exploradas nas sugestões anteriores podem ser aplicadas a qualquer tipo de sinal temporal (*e.g.*, com uma dimensão).

3) Apresentação dos projetos

Os projetos devem ser entregues no formato digital e devem conter: (i) os arquivos Matlab utilizados e (ii) um reporte técnico simplificado. Embora o aluno tenha a liberdade de fazer o reporte como desejar, sugere-se algumas das seções abaixo:

- *Introdução*: breve texto descrevendo o que faz seu projeto e seus objetivos.
- *Especificação do problema*: uma breve descrição do problema e as formulações gerais que dêem informações e subsídios gerais para o leitor (que pode ser leigo no assunto) entender a questão abordada. Pode-se usar de referências bibliográficas para enriquecer com informações o documento para o leitor mais interessado.

- *Metodologia*: descrição de como você deverá resolver o problema passo a passo. Aconselha-se o uso de figuras, tabelas ou códigos para facilitar esta tarefa.
- *Resultados e análises*: descrição do que aconteceu quando você aplicou seu método nos seus dados. Utilize gráficos, análises estatísticas, numéricas e quaisquer outros métodos para mostrar seus resultados. Uma vez demonstrados, faça seus comentários indicando, quando for o caso, sua interpretação e sua avaliação dos resultados indicando seu significado e tecendo os comentários que achar oportunos.
- *Referências*: nesta seção liste as referências (use algum dos modelos de referências que desejar desde que obedeçam algum padrão) empregadas durante o texto. Você pode citar artigos de revistas ou congressos, livros, sites e outros documentos científicos quaisquer.

Destaca-se que seu reporte técnico não será avaliado pela quantidade de texto, mas sim pela sua qualidade. Por isto, recomenda-se fortemente textos sucintos, coesos e objetivos. Espera-se que através de seu texto qualquer possa reproduzir seu trabalho sem muitas dificuldades. Se isto não puder ser feito, indica que seu texto não cumpriu com os objetivos! Ainda, o projeto deve ter muitas escolhas de ordem pessoal e técnica. Isto ajuda a diferenciar um projeto de uma cópia. Pense nisto! (projetos com indícios de cópia terão sua nota anulada)

4) Avaliação

Os projetos serão avaliados sob dois critérios: (i) qualidade dos resultados que devem ser avaliados a partir dos seus arquivos de Matlab e (ii) o reporte técnico. A intenção do reporte vai muito além da mera documentação dos seus resultados. Ela demonstra sua capacidade de comunicação escrita, de organização das ideias, metodologia e sua capacidade de interpretação que também devem fazer parte de sua formação universitária.

Seus resultados, se corretos, devem compor 60% da sua nota. Os demais pontos (i.e., 40%) referem-se à avaliação do seu reporte onde será avaliada não só o texto, mas principalmente os resultados (quantitativamente e qualitativamente) e as interpretações dadas.

Os projetos devem ser entregues PESSOALMENTE ao professor (em pen-drive ou equivalente) até o último dia de aula (ou outra data especificada pelo professor em sala de aula) que fará a avaliação de seu projeto durante o ato da entrega. O aluno deverá explicar suas escolhas técnicas que diferenciem os projetos e ajudarão a demonstrar seus conhecimentos técnicos e demonstrar que seus resultados não constituem mera réplica uma vez que a algumas aplicações propostas são comumente encontradas na internet e em outras boas universidades. Resumindo: convença-me!

5) Referências bibliográficas

- [1] John R. Deller Jr., John H. L. Hansen e John G. Proakis. *Discrete-time processing of speech signals*. Wiley-IEEE, 1999.
 - [2] Rabiner e Schafer. *Digital processing of speech signals* (capítulo 4). Prentice-Hall, 1978.
 - [3] Vídeo “Virtual Barber Shop”. Disponível em: www.youtube.com/watch?v=IUDTlvagjJA. Site acessado em 3 fev. 2015.
 - [4] Fernando Villavicencio, Axel Röbel e Xavier Rodet. *Improving LPC spectral envelope extraction of voice speech by true-envelope estimation*. International Conference on Acoustics, Speech and Signal Processing, pp. 869-872, 2006.
-

- [5] Jia-lin Shen, Jieh-weih Hung, Lin-shan Lee. "Robust Entropy-based Endpoint Detection for Speech Recognition in Noisy Environments". Proc. Int. Conf. on Spoken Lang. Processing, Sydney, 1998.
-