

## 5

### Conclusões e Sugestões para Trabalhos Futuros

Nesta dissertação foi analisado um método de ajuste de parâmetros LSF's, para o qual foram propostas melhorias. O objetivo do ajuste é buscar aproximar as LSF's interpoladas das originais de cada sub-quadros de voz. Foram utilizadas medidas de distorção objetivas (distorção espectral, outliers, e distância LSF euclidiana ponderada) para a avaliação das melhorias nos parâmetros LSF, e o teste PESQ, para a avaliação da qualidade de voz sintetizada. Para o teste PESQ foi empregado um codificador de voz operando a taxas médias de 2 kb/s. A seguir tem-se uma descrição resumida das principais contribuições desta dissertação e algumas sugestões para trabalhos futuros.

#### 5.1

##### Resumo e principais conclusões

No Capítulo 2 foi feita uma breve descrição do processo de codificação de voz, indicando o ponto de inserção do processo de ajuste de parâmetros LSF estudado nesta dissertação. Foi feita uma apresentação resumida do processo de análise linear preditiva (LPC), dando ênfase à função do filtro preditor e às características dos coeficientes preditores. Por último, foram apresentados os parâmetros LSF, explicando como são obtidos, e quais as características que os tornam interessantes para o processo de codificação de voz.

No capítulo 3 foi analisado o método de ajuste de parâmetros LSF, proposto em [5]. Inicialmente foi feita uma breve explicação sobre dois temas fundamentais para a compreensão do processo de ajuste: a interpolação de LSF e a distância LSF euclidiana ponderada. A seguir, foram apresentadas deduções detalhadas das expressões de ajuste contidas no trabalho publicado em [5]. É importante observar que tais deduções foram feitas para dois modos, com e sem introdução de atraso no processo de codificação da voz. Logo após, é feita a apresentação da distorção espectral (DE) e do processo de obtenção dos coeficientes de otimização que minimizam a DE e a  $d_{LSF}$ . De modo a estabelecer

uma referência para comparação de medidas foram feitas simulações com o codificador sem nenhuma alteração, isto é, sem utilizar o processo de ajuste de LSF's. A partir de simulações realizadas com os coeficientes apresentados em [11], verificou-se a redução das medidas de distorção, DE e  $d_{LSF}$ , quando minimizadas através dos respectivos coeficientes de otimização. Tais medidas foram realizadas antes e depois da quantização, constatando-se a melhoria de desempenho em ambas as situações. Também ocorreu uma pequena melhoria no teste PESQ quando minimizada a DE. É importante notar que em [5] e [11] não é apresentada nenhuma avaliação direta da qualidade de voz.

Na seqüência do capítulo 3 foram treinados novos coeficientes de otimização, tanto para a minimização da DE, quanto para minimização da  $d_{LSF}$ . Tais coeficientes foram utilizados em novas simulações através das quais foi possível verificar a redução das medidas de distorção. Uma conclusão importante é que no caso de se desejar obter uma redução dos outliers, a minimização da  $d_{LSF}$  se mostrou mais eficiente do que a da DE. Em geral, o teste PESQ também apresentou melhoria, exceto nos casos “no lookahead” e “lookahead” com 1 subquadro à frente, minimizando a  $d_{LSF}$ .

O capítulo 4 apresenta a maior contribuição dessa dissertação na medida em que propõe melhorias no processo de ajuste de LSF's apresentado no capítulo 3. Inicialmente foi abordado um método de aproximação da DE pela  $d_{LSF}$ , definido em [5], para a qual são determinados os parâmetros, a partir de um conjunto de treinamento. Através de tal aproximação foi possível deduzir uma expressão para o cálculo da DE aproximada em função diretamente de parâmetros LSF. Os novos parâmetros para a aproximação da DE e a expressão para a DE aproximada em função dos LSF's podem ser considerados contribuições iniciais deste capítulo. Sem dúvida, o ponto de maior interesse desta dissertação é o que se segue no capítulo 4, onde são apresentadas deduções detalhadas de um novo método de ajuste de LSF's realizado através da expressão da DE aproximada. Tal ajuste é proposto para ambas situações com e sem a introdução de atraso. A grande contribuição desta nova proposta é eliminar o treinamento de coeficientes de otimização, processo que exige grande esforço computacional, podendo ser extremamente demorado. Além disso, por realizar a minimização diretamente através de uma expressão da DE, ainda que aproximada, foi possível obter uma

redução ainda maior da medida. Isso foi confirmado através de simulações que mostraram a redução da DE em todos os casos, exceto no modo “no lookahead”, quando medida antes da quantização. O teste PESQ apresentou uma pequena melhoria, mas foi possível estabelecer uma relação direta entre a evolução do número de sub-quadros à frente utilizados e o aumento da medida. Testes de escuta informais revelaram que a diferença obtida na percepção subjetiva de qualidade de voz é pequena. Isso possivelmente se deve ao fato de que o codificador utilizado apresenta baixa qualidade de voz, devido ao fato de operar a baixas taxas, fazendo que pequenas mudanças praticamente não sejam percebidas.

## 5.2

### Sugestões para Trabalhos Futuros

A seguir são apresentadas algumas sugestões para trabalhos futuros que podem ser realizados de modo a dar seguimento no que foi exposto nesta dissertação. Muitas dessas sugestões são linhas de investigação previstas no início deste trabalho, outras são possibilidades que surgiram ao longo do estudo, mas que foram descartadas devido a limitação de tempo. São elas:

- É possível que diferentes esquemas de janelamento, aplicado logo após a segmentação do sinal do voz para a análise LPC, tenham efeito sobre o processo de ajuste, na medida em que podem alterar a contribuição de cada sub-quadro individualmente.
- Testar novos codificadores de voz, principalmente a taxas maiores, como por exemplo o G723.1, pois é possível que neste caso a qualidade de voz sintetizada possa ser melhor avaliada pelo teste PESQ.
- Treinar novos coeficientes de otimização usando conjuntos de treinamento ainda maiores e mais diversificados, de modo a tornar o ajuste ainda mais preciso.
- Testar novos processos de interpolação além da linear, como por exemplo a polinomial.

- O treinamento do quantizador dos codificadores em geral não utiliza LSF's ajustados. Portanto, seria interessante realizar um novo treinamento usando LSF's modificados pelo processo de ajuste.
- Aplicar o ajuste de LSF's no reconhecimento de voz.
- Analisar a interação entre o processo de ajuste de LSF's e os pós-filtros de saída do decodificador, pois é possível que novas combinações de filtros proporcionem melhores resultados.
- Expressar a relação entre DE e  $d_{LSF}$  através de outra função que não a logaritmica (por exemplo a quadrática) e aplicar essa nova expressão no processo de ajuste.