

5

Conclusões e Sugestões para Trabalhos Futuros

Nesta dissertação foi analisada a qualidade da voz em codecs a baixas taxas operando em redes IP. Na simulação de perda de quadros em rajada, típica de redes IP, adotou-se um modelo Markoviano de dois estados, também conhecido como Modelo de Gilbert. Para quantizar os parâmetros LFS usou-se três esquemas de quantização vetorial: QVPC4, QVPCP2 e QVSM. Esses esquemas foram comparados em diversas taxas de perda de quadros e como medida de desempenho utilizou-se a distorção espectral, percentual de *outliers* e a recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ). Posteriormente, os experimentos foram refeitos adotando os mesmos parâmetros iniciais e as mesmas medidas de desempenho, mas com vozes em presença de 3 tipos de ruído (falatório, fábrica e branco). E, por fim, utilizou-se um algoritmo de supressão de ruído baseado em *wavelets* e avaliou-se a qualidade da voz decodificada utilizando o PESQ e testes comparativos de escuta A/B.

A seguir é apresentado um resumo das principais contribuições contidas na dissertação e algumas sugestões para trabalhos futuros.

5.1

Resumo da Dissertação

No Capítulo 2 foram apresentados os esquemas de quantização vetorial utilizados nessa dissertação e o codec a baixa taxa de bits utilizado nos experimentos. Ainda neste capítulo, discutiu-se os problemas envolvidos na transmissão e recuperação dos parâmetros LSF em redes IP, e apresentou-se uma solução adotada na recuperação de perda de quadros de LSF, que é a interpolação dos parâmetros LSF. No Capítulo 3 foi avaliado o desempenho do codec operando em rede IP, utilizando 3 esquemas de quantização vetorial das LSFs. Em uma primeira etapa, as simulações foram feitas com frases sem ruído ambiente. Nessas simulações, verificou-se que o esquema de quantização vetorial QVPC4 proposto por de Lamare e Alcaim,

obteve melhor desempenho em todas as taxas de perda de quadros quando comparados aos esquemas QVPCP2 e QVSM. O QVSM, foi o que forneceu o pior desempenho nas condições simuladas. Verificou-se, ainda, que até a uma taxa de perda de quadros de 9% o desempenho dos quantizadores não foi muito afetado, ficando próximo do desempenho conseguido em redes sem perdas. Posteriormente, adicionou-se 3 tipos de ruído nas vozes a serem codificadas e repetiu-se os experimentos. Verificou-se que em RSR baixas, menores que 10dB, a qualidade da voz processada pelo codec a baixas taxas, fica bastante comprometida, tendo elevados valores de DE e percentual de *outliers* e pontuações muito baixas no PESQ. Essa situação é agravada quando a rede opera com taxas de perdas altas como 28,6% e 38,5%. De qualquer modo, o melhor desempenho, mesmo em ambientes muito hostis, é atingido pelo QVPC4. No Capítulo 4, aplicou-se uma técnica de supressão de ruído baseada em transformadas *wavelet - wavelet denoising* - com o objetivo de reduzir o ruído das vozes antes destas serem processadas pelo codec. Isso permitiu avaliar a qualidade da voz codificada com e sem a utilização de *wavelet denoising*. Neste capítulo optou-se por utilizar apenas um esquema de quantização vetorial, pois o que queríamos avaliar era a qualidade da voz codificada utilizando a técnica de supressão de ruído em questão. O QVPC4, foi o esquema escolhido por ter obtido o melhor desempenho nas simulações anteriores. Neste capítulo, usou-se dois critérios de avaliação da qualidade da voz, a medida PESQ e testes subjetivos de comparação A/B. Através dos resultados obtidos com o PESQ, verificou-se que o desempenho do codec utilizando *wavelet denoising* não melhorou segundo essa medida. Entretanto, o teste de comparação A/B mostrou que a maioria dos ouvintes preferiram a voz codificada com a utilização do método de supressão de ruído. Mesmo o algoritmo de supressão de ruído causando uma leve deterioração da voz, a maioria dos ouvintes preferem escutar as vozes codificadas com menos ruído de fundo. Isso mostra que apesar da recomendação ITU-T P.862 de avaliação perceptiva de qualidade da voz (PESQ) ser uma boa medida objetiva de avaliação, em alguns casos é importante também avaliar os resultados através de testes de escuta subjetivos.

5.2

Sugestões para Trabalhos Futuros

Seria interessante utilizar outros codificadores, como padrões de telefonia móvel, a fim de avaliar o impacto de perda de quadros de LSF e

ambientes ruidosos nesses codificadores. E comparar esses resultados com os obtidos nessa dissertação.

Para combater os efeitos de perda de quadros de LSF, uma solução adotada nessa dissertação, foi a interpolação linear dos parâmetros LSF. Entretanto, é interessante pesquisar outros métodos para combater os efeitos dessas perdas.

Na dissertação, considerou-se apenas a perda de quadros de LSFs. É interessante também, verificar os efeitos da perda de outros parâmetros da voz, na qualidade da voz codificada por um codec operando em redes de pacotes. E, além disso, pesquisar meios de recuperação desses parâmetros.

Nesse trabalho, utilizou-se apenas um cenário para redes IP, seria interessante simular outros cenários e simular o codec operando em redes ATM e Frame Relay.

Nos últimos anos, os sistemas de reconhecimento de voz têm despertado um interesse crescente, principalmente devido a oportunidade de utilização desses sistemas em comunicações móveis e na internet. Então, seria interessante verificar os efeitos de perda de quadros de LSF, ruído ambiente e utilização de *wavelet denoising* em reconhecedores de voz distribuídos.