

PROPOSTA E AVALIAÇÃO DE UM MÉTODO DE MEDIDA OBJETIVA DE QUALIDADE DE CODECS DE VOZ

Jayme G. A. Barbedo e Amauri Lopes

Departamento de Comunicações - FEEC - UNICAMP
Caixa Postal 6101, CEP: 13.083-970, Campinas - SP - Brasil
Tel: (19) 3788-3703; {jgab, amauri}@decom.fee.unicamp.br

SUMÁRIO

Apresenta-se neste trabalho um método de avaliação objetiva de qualidade de codecs na faixa de telefonia (300-3400 Hz) que reúne as principais características dos métodos presentes na literatura, denominado Medida Objetiva de Qualidade de Voz (MOQV). Apresenta-se uma análise de seu desempenho, tendo como base os experimentos presentes em um banco de dados da *International Telecommunication Union* (ITU).

A análise de desempenho oferece uma ampla caracterização do método, permitindo determinar-se em quais circunstâncias seu uso é ou não adequado, além de fornecer subsídios para sua utilização frente às diferentes situações encontradas na prática.

1. INTRODUÇÃO

O avanço das técnicas de processamento digital de sinais e da tecnologia proporcionou um crescente interesse em métodos e dispositivos de codificação e decodificação de voz (codecs) mais eficientes. A avaliação da qualidade de codecs é necessária para o desenvolvimento destes dispositivos e também para o projeto de redes digitais de telecomunicações.

Os métodos objetivos clássicos de avaliação de desempenho, como a taxa de erro e a relação sinal-ruído, não oferecem medidas suficientemente correlacionadas com as impressões dos usuários dos sistemas de telecomunicações. Assim, as medidas subjetivas de qualidade são ainda amplamente empregadas. Porém, seu custo, complexidade e tempo demandado motivam fortemente a busca por métodos eficazes para a realização de medidas objetivas que estimem a qualidade subjetiva de maneira satisfatória.

Nesse contexto, a ITU-T incentivou o desenvolvimento de um método objetivo que atendesse às muitas exigências de desempenho. Diversas propostas foram submetidas a uma série de testes e, após sua comparação, chegou-se à conclusão de que o método que melhor estimava os valores subjetivos era o *Perceptual Speech-Quality Measure* (PSQM) [1,2], o qual foi adotado pela ITU-T, dando origem à Rec. P.861 [3].

Apresenta-se neste artigo as principais características inerentes ao método em questão, bem como uma implementação do mesmo através de um programa de computador. Esta implementação incorpora também algumas propostas que melhoram o desempenho do PSQM, além de recursos adicionais que o tornam mais completo e versátil [4]. Este conjunto é aqui

denominado de método MOQV.

Por fim, são apresentados os testes realizados com o MOQV, cujos resultados são utilizados na descrição do comportamento e aplicabilidade deste método em situações práticas.

2. MÉTODOS PSICOACÚSTICOS

Os melhores métodos objetivos desenvolvidos baseiam-se no modelamento matemático do ouvido humano [2,5,6,7]. A figura 1 ilustra as características básicas comuns a estes métodos.

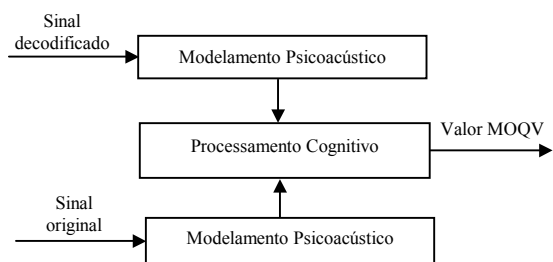


Fig. 1 - Esquema básico dos métodos psicoacústicos

2.1. Modelamento Psicoacústico

O pré-processamento realizado pelo ouvido sobre o sinal acústico é uma atividade objetiva, pois envolve a transformação do sinal acústico que chega ao ouvido externo em impulsos elétricos nos feixes de neurônios distribuídos ao longo da cóclea no ouvido interno. O processamento subjetivo será realizado pelas funções superiores do córtex cerebral, baseado neste sinal condensado gerado pelo ouvido [8].

Um dos conceitos fisiológicos preponderantes para o modelamento do ouvido humano é o fenômeno das *bandas críticas*. Elas definem um conjunto de faixas de frequência, cada uma delas em torno de uma frequência central [9,10,11]. Em cada banda, um sinal de maior energia dominará a percepção e mascarará outro de menor intensidade, fenômeno este conhecido como *mascaramento*. Este efeito pode ser sentido tanto no domínio da frequência, em que só se distingue dois tons quando ambos estão em bandas críticas diferentes, quanto do tempo, em que um sinal de grande amplitude mascara outro de pequena amplitude por um certo tempo após ter sido ouvido.

No modelamento do ouvido, é ainda necessário levar em conta as transformações necessárias entre as escalas de amplitude de sinal e a escala de sonoridade no ouvido interno. O

modelamento matemático nos métodos objetivos consiste dos seguintes passos [12]:

- primeiro, relaciona-se as frequências acústicas à resolução perceptual de frequências, utilizando o conceito de bandas críticas, obtendo-se o *padrão de excitação*;
- como nem todas as bandas críticas têm o mesmo ganho, faz-se a ponderação do padrão de excitação pela sensibilidade do ouvido a cada faixa de frequência, o que equivale à conversão de um sinal que está numa escala de intensidade para uma escala de *excitação* ou *sonoridade*;
- foi verificado ainda que a escala de sonoridade não é linear em relação à sonoridade percebida pelo ouvido humano, o que acarretou a necessidade de se transformar esta escala não-linear em uma escala linear, através de uma função de compressão não-linear, obtendo-se assim a sonoridade subjetiva, representando a densidade espectral de energia perceptual, aqui denominada *representação interna*.

2.2. Processamento Cognitivo

Consiste de todos os processamentos que são realizados após o cálculo do sinal de diferença entre as representações internas do sinal fonte e do sinal na saída do codec. A qualidade do sinal decodificado é calculada através deste sinal de diferença processado. Para os métodos PSQM e MOQV, em particular, o nível médio deste sinal, calculado ao longo do tempo através das suas componentes de frequência, é relacionado diretamente com a qualidade da voz fornecida pelo codec.

As diferenças inaudíveis, dentro das precisões de representação e julgamento destes métodos, entre o sinal original e aquele resultante do codec, não influem na pontuação resultante da medida objetiva. Em particular, se os sinais são idênticos, será estimada uma qualidade perfeita, independente das características dos sinais de entrada.

3. MÉTODO PROPOSTO

O programa de computador original do método PSQM foi implementado na linguagem ANSI C. Este serviu de base para o desenvolvimento do programa MOQV, o qual foi implementado em Matlab® (versão 5.3), a fim de se ter maior flexibilidade e versatilidade para a realização de modificações e testes. A figura 2 mostra a estrutura adotada para o programa.

Este programa oferece diversas facilidades não encontradas naquele implementado para o método PSQM: alinhamento temporal dos arquivos de voz fonte e decodificado via correlação cruzada; opção de várias taxas de amostragem; detalhamento das características dos arquivos de voz, como sua visualização, audição e nível do sinal; ajuste do nível dos arquivos; diversas opções de mapeamento para os valores subjetivos; processamentos adicionais visando a melhora de seu desempenho [4,13]. Tais melhorias resultam numa maior interatividade, tornando seu uso adequado tanto para usuários com larga experiência na área quanto para iniciantes.

No início do programa, o usuário deverá fornecer os nomes dos arquivos de voz original e decodificado, a taxa de amostragem (em Hertz), a opção de ter ou não mais detalhes a

respeito dos arquivos analisados e, por fim, as opções de mapeamento.

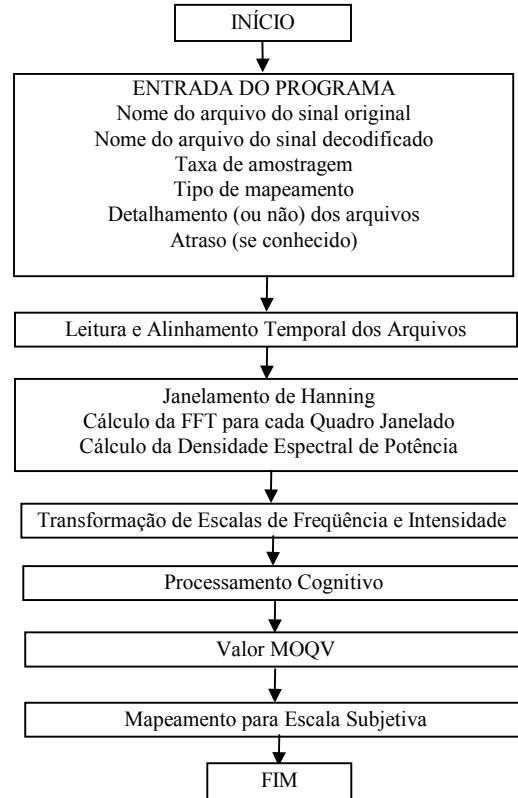


Fig. 2 - Fluxograma do Programa Implementado

Se o usuário optar por ter mais detalhes, entrará em funcionamento a rotina “geral”, onde o arquivo de voz original é executado, ou seja, faz-se uso dos recursos de áudio do computador; além disso, são traçadas as curvas (no domínio do tempo) dos dois sinais para efeito de comparação entre ambos. Por fim, são fornecidos os níveis, em dB, dos sinais. Este nível pode ser mudado através de uma terceira rotina, a qual foi implementada separadamente.

O alinhamento temporal entre os arquivos é realizado através da correlação cruzada entre eles, calculada através de um algoritmo de transformada rápida de Fourier (FFT).

O índice do valor máximo da correlação representa o deslocamento entre os sinais, e o alinhamento dos mesmos é feito automaticamente.

O janelamento foi feito através da rotina “enframe”, onde os arquivos são divididos em quadros de 256 ou 512 amostras, para uma taxa de amostragem de 8 kHz ou 16 kHz, respectivamente. Há uma superposição de 50% entre os quadros. Após, é feito o cálculo da FFT individualmente para cada quadro, através da rotina “mfft”, a qual chega a ser 10 vezes mais rápida que a rotina “fft” normal, pois considera apenas as componentes espectrais relevantes no cálculo das medidas MOQV. Após, há o cálculo da densidade espectral de potência.

As transformações de escalas de frequência e intensidade seguem aquelas utilizadas no método PSQM [2].

No processamento cognitivo, calcula-se a diferença entre os sinais resultantes da transformação de intensidade. Após, são levados em conta efeitos de assimetria na percepção sonora (e.g. a inserção ou supressão de componentes estranhas pelo codec) e, por fim, realiza-se uma ponderação dos intervalos de silêncio para condicionar sua contribuição para o valor objetivo final.

O programa fornece algumas opções para o mapeamento da medida MOQV:

1) para as medidas MOS ou CMOS, que são medidas subjetivas padronizadas [14], a primeira representando um sistema de pontuação que define a qualidade de pequenos grupos de sentenças descorrelacionadas e a segunda representando um sistema de pontuação que fornece um desempenho relativo de um sinal em relação a outro de referência;

2) para as línguas francesa, inglesa, japonesa, ou mapeamento genérico;

3) opção de mapeamento por polinômios de 1ª a 6ª ordem, os quais foram determinados através de um banco de dados contendo diversos arquivos de voz, como será descrito na próxima seção.

O programa foi testado em um computador PC com processador Intel Pentium® III 650 MHz, com 128 Mbytes de memória RAM e sistema operacional WindowsME®. Para arquivos contendo 120.000 amostras, o tempo de execução é de 4,78 segundos, bem próximo àquele encontrado para o programa PSQM em linguagem C testado em estações Sun UltraSparc® e ambiente Unix. Para arquivos de voz contendo 30.000 amostras, este tempo ficou abaixo de 1 segundo.

4. ANÁLISE DE DESEMPENHO

4.1. Banco de Dados Utilizado

Todos os testes realizados utilizaram os arquivos de voz contidos no banco de dados “Supplement 23” da ITU, o qual é constituído por arquivos de voz em inglês, francês e japonês, associados aos testes de vários codecs submetidos a algumas condições de teste [15]. Cada arquivo foi gravado em três diferentes versões: original, pré-processado (arquivos filtrados e equalizados) e decodificado. A cada teste é fornecido o valor MOS ou CMOS correspondente, o qual deve ser estimado pelo método MOQV. Tal material está dividido em três grupos de experimentos:

- 1º Experimento: os arquivos de voz presentes neste experimento foram submetidos a diversos tipos de codecs padrão da ITU e de telefonia celular. Para cada uma das três línguas (francês, japonês e inglês), dois locutores de cada sexo foram utilizados, cada qual enunciando uma sentença. Há 44 condições de teste para cada locutor, perfazendo um total de 528 arquivos para o experimento (176 por língua).

- 2º Experimento: aqui os arquivos de voz foram submetidos a diversos tipos de ruídos de ambiente - ruídos de escritório, veículo, rua, branco e música, com uma SNR de 10 ou 20 dB. Dois locutores de cada sexo foram utilizados para as primeiras 28 condições, e um de cada sexo para as últimas 12. Há então 40 condições, com um total de 136 arquivos de teste para cada uma das 3 línguas (perfazendo um total global de 408 arquivos para o experimento).

- 3º Experimento: os arquivos deste experimento simulam os efeitos de transmissão do sinal codificado através de um canal de comunicações que introduz erros, em particular erros aleatórios e erros FER (burst frame erasure – erros de surto). Dois locutores de cada sexo foram utilizados, havendo 50 condições para cada um deles, perfazendo um total de 200 arquivos para teste por língua. Aqui, além das línguas já citadas, há também o italiano, o que resulta num total de 800 arquivos de teste.

4.2. Mapeamento entre as Medidas

No processo de mapeamento das medidas objetivas para as subjetivas, foi utilizada a expressão [16]:

$$y_i = \beta_1 + \beta_2 \cdot x_i + \dots + \beta_p \cdot x_i^p \quad (1)$$

onde x_i representa as medidas objetivas e y_i representa as medidas subjetivas. Os parâmetros β_i são obtidos pela minimização de

$$Q = \sum_{i=1}^n [y_i - (\hat{\beta}_1 + \hat{\beta}_2 x_i + \dots + \hat{\beta}_p x_i^p)]^2 \quad (2)$$

Vários valores para a ordem polinomial p serão testados.

O cálculo da correlação entre as medidas objetivas obtidas e as subjetivas correspondentes é realizado segundo [17]:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (3)$$

onde \bar{x} e \bar{y} representam, respectivamente, as médias para as medidas objetivas e as medidas subjetivas.

Assim, quanto mais próximo de 1 for o módulo de r , melhor é o desempenho do método de avaliação objetiva.

4.3. Tipos de Testes Realizados

Nos testes, procurou-se abordar o maior número de situações possíveis, a fim de se traçar um perfil do desempenho do método MOQV. Para isso, considerou-se:

- as características de cada um dos três experimentos submetidos aos testes, como descrito na seção 4.1;

- mapeamento para cada língua e mapeamento geral médio em relação a todas as línguas presentes no banco de dados. O mapeamento individual para cada língua resulta em uma melhor correlação entre as medidas objetivas e os valores subjetivos, já que esta abordagem faz com que os resultados estejam mais bem

adaptados às peculiaridades daquela língua e cultura. Para efeito de comparação, utilizaram-se os dados obtidos, através deste critério, pela *Ericsson* [18]; o segundo critério consiste em se fazer um mapeamento geral, levando-se em consideração todas as amostras de todas as línguas disponíveis, o que faz com que a correlação diminua. Este é a abordagem adotada pela *KPN Research* da Holanda [19];

- a influência do valor da ordem polinomial p , permitindo assim determinar as ordens mais adequadas para a realização do mapeamento.

4.4. Ilustração dos Resultados Alcançados

Uma grande quantidade de testes foi realizada, dando origem a um total de 156 curvas, as quais descrevem graficamente o desempenho do método para cada experimento considerado [13]. Apresenta-se aqui apenas alguns resultados típicos para exemplificar estes resultados.

A figura 3 ilustra os resultados obtidos no estudo da correlação entre as medidas MOQV e aquelas subjetivas para as diversas línguas.

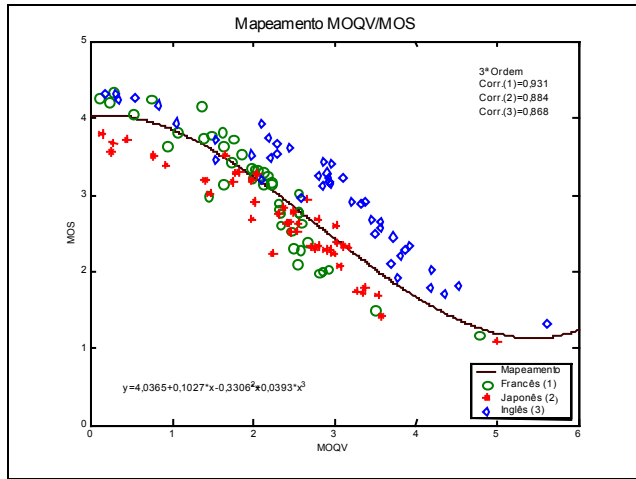


Fig. 3 - Mapeamento MOQV/MOS

Esta figura foi gerada para o primeiro experimento, usando a segunda abordagem, isto é, o mapeamento geral, e com polinômio de 3ª ordem. No eixo x tem-se os valores objetivos MOQV, enquanto que no eixo y tem-se os valores subjetivos medidos. Os coeficientes da curva de mapeamento, obtidos através da equação 3, são mostrados no canto inferior esquerdo. Os valores das correlações entre os valores objetivos e subjetivos para cada língua, obtidos através da equação 4, são apresentados no canto superior direito do quadro. Os demais gráficos obtidos seguem as mesmas prerrogativas.

A figura 4 ilustra os resultados obtidos no estudo sobre o valor adequado da ordem polinomial p a ser utilizada, obtida para o primeiro experimento, com mapeamento geral. Observa-se, tipicamente, que a partir da 4ª ordem polinomial, começa a haver uma quebra significativa de monotonicidade, o que não é desejável. Além disso, os ganhos nos valores das correlações vão se tornando desprezíveis conforme se aumenta o valor de p .

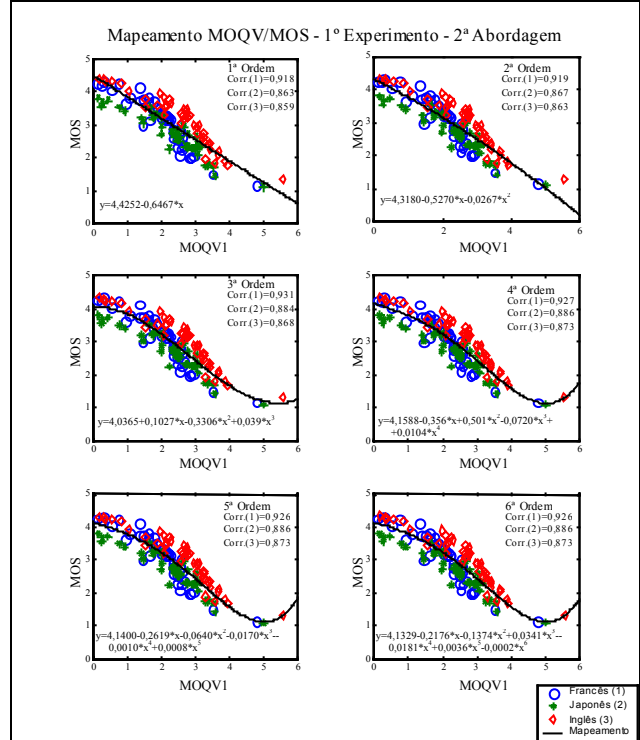


Fig. 4 - Estudo do Parâmetro p

A figura 5 ilustra o desempenho típico observado para cada um dos experimentos com o mapeamento geral e polinômio de terceira ordem. Como se pode observar, os melhores desempenhos foram obtidos para o primeiro e segundo experimento. As correlações obtidas para o terceiro experimento foram ruins, o que é ilustrado pelo grande espalhamento dos valores mapeados no gráfico. Isso indica que o método não é indicado para situações em que há a presença de erros de transmissão. Note-se também que o mapeamento realizado para o segundo experimento foi o CMOS, pois esta é a medida subjetiva oferecida pelo banco de dados, enquanto para os outros dois utilizou-se o MOS.

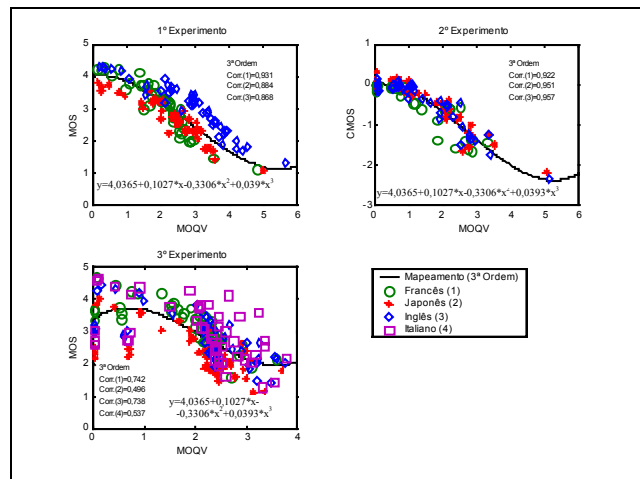


Fig. 5 - Desempenho do Método MOQV para cada Experimento

Além do cálculo da correlação entre as medidas objetivas MOQV e as medidas subjetivas, fez-se também uma análise estatística que incluiu a construção de histogramas descrevendo o comportamento dos desvios dos valores objetivos em relação aos valores subjetivos correspondentes. Foi ainda elaborada uma tabela com os valores de desvio máximo para determinados intervalos de confiança [13].

4.5. Aplicabilidade

Apresenta-se aqui um resumo das conclusões resultantes da análise de todos os testes realizados, permitindo obter as informações necessárias para o uso adequado do método MOQV frente às diversas situações encontradas na prática:

1 - A primeira abordagem (mapeamento específico) é a mais adequada, pois se adapta melhor às peculiaridades de cada língua e cultura. Idealmente, cada língua deveria ter seus valores de mapeamento. Porém, devido à limitação do banco de dados utilizado, o qual continha apenas arquivos nas línguas inglesa, japonesa e francesa, determinou-se um mapeamento geral, utilizando em conjunto as três línguas citadas, o qual pode ser utilizado para as demais línguas, incluindo o português. Entretanto, é preciso ter presente que tal mapeamento não oferece os melhores resultados. Por isso, sugere-se que ou sejam feitos experimentos para obter o melhor mapeamento para a língua específica, ou faça-se uso dos valores MOQV brutos, ou seja, analise-se a qualidade diretamente a partir dos valores MOQV obtidos, sem a realização do mapeamento. Esta opção está disponível no programa desenvolvido e é condicionada ao grau de experiência apresentado pelo usuário.

2 - A monotonicidade é uma característica desejável neste tipo de mapeamento. Os resultados mostram que a utilização de ordens de mapeamento acima da terceira não é aconselhável por introduzir, em certos casos, grandes quebras na monotonicidade. A ordem de mapeamento mais apropriada para a maioria dos tipos de arquivos é a terceira, fato este observado para todas as línguas presentes nos testes. Isto ocorre devido à sua capacidade de modelar satisfatoriamente o comportamento aproximado dos avaliadores nos testes subjetivos, ou seja, representa bem a tendência de perda de linearidade nos extremos de qualidade (sinal muito bom ou muito ruim), pois os ouvintes tendem a saturar a avaliação nestes pontos. Pequenas quebras na monotonicidade foram observadas para este caso, porém com efeito desprezível no resultado final, pois os poucos casos testados que eventualmente se encontrem na pequena faixa do mapeamento em que houve quebra na monotonicidade apresentarão valores mapeados bem próximos do desejado. O ganho na correlação representado pela passagem da primeira para a terceira ordem chegou a 0,030 no caso do primeiro e segundo experimentos e 0,065 no caso do terceiro. Os ganhos observados na análise dos intervalos de confiança também se mostraram bastante expressivos, diminuindo os limites de desvio do valor correto em mais de 0,1 em muitos casos.

3 - Pela análise dos gráficos, pode-se afirmar que o MOQV funcionou adequadamente para os dois primeiros experimentos, mostrando que este método se adapta bem a casos em que o sinal é submetido a um ou mais tipos de codecs, bem como a situações em que há a presença de ruído de fundo. Já para o terceiro experimento, os resultados foram ruins, denotando uma clara inadequação a situações em que os sinais estão sujeitos a condições de erro severas. As correlações foram sempre ruins,

especialmente para as línguas japonesa e italiana. Porém, levando-se em consideração todas as ressalvas e restrições inerentes ao processo, este método pode ser usado, nestas condições, para as línguas francesa e inglesa. No caso de as condições de erro não serem muito severas (abaixo de 3%), o MOQV pode ser utilizado para as demais línguas, desde que observadas as limitações deste procedimento, já que os resultados revelarão apenas uma tendência do desempenho do codec, e não resultados precisos. Esta alternativa, contudo, fica condicionada a estudos preliminares e também à aplicação desejada para o codec a ser avaliado.

4 - Como já foi comentado, seria desejável um mapeamento individual para cada língua. Uma consideração importante é o fato de que a variação dos valores de mapeamento de língua para língua seguem um certo padrão, o qual pode ser explicado pelas diferenças culturais de cada país. Dentre as línguas analisadas, pode-se perceber claramente que os japoneses são os mais exigentes, seguidos de perto pelos franceses e, com um grau de exigência bem menor, os norte-americanos.

5 - Para a determinação de um mapeamento adequado às peculiaridades da língua portuguesa, em particular aquela falada no Brasil, seria necessário um banco de dados com arquivos de voz e valores de medidas subjetivas, o que ainda não foi feito devido à falta de estrutura laboratorial para este fim. A utilização do mapeamento geral, no entanto, se mostra suficiente para grande parte das aplicações desejadas. Alternativamente, pode-se utilizar diretamente o valor MOQV.

6 - De uma maneira geral, as restrições na utilização do MOQV dependem fortemente da aplicação desejada para o codec a ser avaliado, bem como da qualidade da avaliação que se deseja fazer. No caso de codecs a serem utilizados em aplicações menos exigentes, como comunicação militar ou o VoIP (telefonia sob protocolo da Internet), por exemplo, este programa pode ser utilizado sem restrições importantes. Já no caso de aplicações mais exigentes, em que a qualidade é um fator muito importante e decisivo e há a necessidade de uma avaliação mais precisa, pode ser necessário se fazer uso das avaliações subjetivas. Nestes casos, contudo, o método MOQV pode ser utilizado com o objetivo de fornecer uma tendência para o comportamento dos codecs testados, o que pode ser valioso em situações em que não se dispõe de tempo e/ou recursos suficientes para a realização de testes subjetivos.

5. CONCLUSÃO

Apresentou-se neste trabalho uma medida objetiva de qualidade de codecs de voz para telefonia baseada no método PSQM [2], incorporando, ao mesmo tempo, características que melhoram seu desempenho. O programa de computador que a implementa incorpora recursos de manipulação dos sinais, formando um produto completo que facilita e agiliza o uso do método. Foram descritas suas características básicas e apresentadas ilustrações dos resultados alcançados em testes exaustivos, utilizando o banco de dados ITU Supplement 23 [15], o que permitiu traçar um perfil da aplicabilidade do método.

Os resultados gerais destes testes indicam que:

- o método MOQV funciona bem para casos em que o sinal é submetido a um ou mais tipos de codecs, bem como a situações

em que há a presença de ruído de fundo, sendo porém inadequado na presença de erros de transmissão;

- deve-se utilizar mapeamentos de 1ª a 3ª ordem, sendo esta última a mais adequada para a maioria das aplicações;

- quando possível, deve-se realizar um mapeamento individual para a língua em questão a fim de se respeitar suas peculiaridades; não sendo possível, pode-se utilizar o mapeamento geral ou o valor MOQV não mapeado;

- devido à falta de dados, não foi possível determinar um mapeamento para a língua portuguesa falada no Brasil; para a maioria das aplicações, porém, o mapeamento geral é suficiente.

Todos os resultados obtidos são também válidos, de forma aproximada, para o método PSQM, e indicam a necessidade da realização de mais pesquisas, a fim de que se possa ampliar a aplicabilidade e melhorar o desempenho de ambos os métodos.

6. REFERÊNCIAS BIBLIOGRÁFICAS

[1] Royal PTT, *Measuring the quality of audio devices*, CCITT Contribution COM XII-114, Geneva, December 1991.

[2] Beerends, J.G., Stererdink, J.A.: *A Perceptual Speech-Quality Measure Based on a Psychoacoustic Sound Representation*, J. Audio Eng. Soc., Vol. 42, No. 3, pp. 115-123, March 1994.

[3] ITU-T Recommendation P.861 (1996), *Objective quality measurement of telephone-band (300 - 3400 Hz) speech codecs*.

[4] KPN, *Improvement of the P.861 Perceptual Speech Quality Measure*, The Netherlands, December 1997.

[5] Atkinson, D.J., *Proposed Annex to Recommendation P.861*, NTIA, ITU Study Group 12 - Contribution COM 12-24-E, December 1997

[6] Voran, S., *Objective Estimation of Perceived Speech Quality - Development of the Measuring Normalizing Block Technique*, IEEE Trans. on Speech and Audio Processing, 7(4):371-390, July 1999

[7] Berger, J., *TOSQA - Telecommunication Objective Speech Quality Assessment*, ITU Study Group 12, Contribution COM12-34-E, December 1997

[8] Wang, S.; Sekey, A. and Gersho, A., *An Objective Measure for Predicting Subjective Quality of Speech Coders*, IEEE Journal on Selected Areas in Communications, 10(5):819-829, June 1992.

[9] Fourcin, A.J. et al, *Speech processing by man and machine: Group Report*, pages 307-351, Bullock, T., 1977.

[10] Zwicker, E., Zwicker, U.T., *Audio engineering and psychoacoustics: Matching signals to the final receiver, the human auditory system*, J. Audio Engineering Society, Vol. 39 No. 3, pp. 115-126, March 1991.

[11] Hermansky, H., *Perceptual linear predictive (PLP) analysis of speech*, J. Acoust. Soc. Am., 87(4):88-94, April 1990.

[12] Campos Neto, S. F. e Violaro, F., *Metodologias de Avaliação de Algoritmos de Codificação de Voz*, Dissertação de Mestrado, Unicamp, Julho de 1993.

[13] Barbedo, J.G.A., Lopes, A., *Medida Objetiva de Codecs de Voz na Faixa de Telefonia (300 - 3400 Hz)*, Dissertação de Mestrado em Fase de Revisão de Redação, Unicamp, Campinas, 2001

[14] ITU-T Rec. P.830 (1996), *Subjective performance assessment of telephone-band and wideband digital codecs*.

[15] *Subjective test plan for characterization of an 8 kbit/s speech codec*, ITU-T Study Group 12 – Speech Quality Experts Group – Issue 2.0, 25 September 1995.

[16] Guthrie, W., *Engineering Statistics Handbook, Data Analysis for Process Modeling*, Nist/Sematech, 1999;

[17] Rix, A., Hollier, M., *Performance Metrics for Objective Quality Assessment Systems in Telephony*, ITU Study Group 12 - Delayed Contribution D.79, November 1998

[18] Folkesson, M., Karlsson, A., *Results of Processing the Supplement 23 Speech Database for Development of the Extended P.861*, Ericsson, ITU Study Group 12 – Delayed Contribution D.89, November 1998

[19] Beerends, J.G., Hekstra, A.P., *Comparison of the ITU-T P.861 PSQM, PSQM+ and MNB Objective Speech Quality Measurement Methods*, ITU Study Group 12, Contribution COM 12-58-E, September 1998