



Keiko Verônica Ono Fonseca
Alexandre de Almeida Prado Pohl
(Editores)

Novas Técnicas e Ferramentas Aplicadas ao Sistema Brasileiro de TV Digital



omnipax
editora

www.omnipax.com.br

Keiko Verônica Ono Fonseca
Alexandre de Almeida Prado Pohl
(Editores)

**Novas Técnicas e Ferramentas
Aplicadas ao
Sistema Brasileiro de TV Digital**



2013

Copyright ©2013 Omnipax Editora Ltda
Caixa Postal: 16532 - 81520-980 Curitiba, PR



A editora disponibiliza por acesso livre a versão eletrônica deste livro no *site*: <http://www.omnipax.com.br>, sob uma licença *Creative Commons* 3.0 Atribuição-Não Comercial-Sem Derivados (CC BY-NC-ND 3.0 BR).

Digital Object Identifier (DOI): 10.7436/2013.sbtvd.0

Capa:

Sérgio Alexandre Prokofiev

Supervisão editorial:

Heitor Silvério Lopes

Projeto gráfico e editoração:

Omnipax Editora Ltda

Ficha catalográfica:

Juliana Farias Motta (CRB7/5880)

Dados Internacionais de Catalogação na Publicação

N936 Novas técnicas e ferramentas aplicadas ao sistema brasileiro de TV digital./ Editores e colaboradores: Keiko Verônica Ono Fonseca, Alexandre de Almeida Prado Pohl; Célio Lúcio Vasco [et al.]. — Curitiba, PR: Omnipax, 2013
146 p.: il.

Inclui bibliografias
eISBN: 978-85-64619-14-2

1. Sistema brasileiro de televisão (Paraná). 2. Telecomunicações. 3. Processamento de sinal – Técnica digital. 4. Inovações tecnológicas – Telecomunicações. 5. Redes – Telecomunicações. I. Fonseca, Keiko Verônica Ono. **Ed.** II. Pohl, Alexandre de Almeida. **Ed.** II. Vasco, Célio Lúcio [et al.]. **col.** III. Título.

CDD (22. ed.) 621.3888

Apoio



Biografia dos Colaboradores

Alexandre de Almeida Prado Pohl: é graduado e mestre em Física (Universidade Estadual de Campinas, 1982 e 1987, respectivamente), e doutor em Engenharia Elétrica (Universidade Técnica de Braunschweig, Alemanha - TUBS, 1994). Atualmente é professor associado da Universidade Tecnológica Federal do Paraná na área de telecomunicações. É membro da Sociedade Brasileira de Microondas e Optoeletrônica (SBMO), da Sociedade Brasileira de Telecomunicações (SBrT) e da *Optical Society of America* (OSA). É bolsista de Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora do CNPq, membro do Conselho Universitário da UTFPR e vice-presidente de finanças da SBrT. Desde 2005 tem liderado e participado de projetos na área de TV digital com ênfase em desenvolvimento de métricas para avaliação de vídeo digital.

Célio Lúcio Vasco é graduado em Engenharia Industrial Elétrica e mestre em Ciências (Universidade Tecnológica Federal do Paraná, 1985 e 2009, respectivamente) e Especialista em Regulação da Agência Nacional de Telecomunicações (ANATEL).

Christofer Schwartz é graduado em Engenharia Elétrica (Centro Universitário de Jaraguá do Sul, 2008) e mestre em Engenharia Elétrica (Universidade Federal do Paraná, 2011). Foi bolsista da CAPES no projeto *Formação de Pessoal Qualificado em TV Digital no Paraná*. Atualmente é doutorando em Engenharia Eletrônica no Instituto Tecnológico de Aeronáutica.

Evelio Martín García Fernández é graduado em Engenharia Eletrônica (Universidad Central de Las Villas, Cuba, 1985), mestre e doutor em Engenharia Elétrica (Universidade Estadual de Campinas, 1997 e 2001, respectivamente). Atualmente é professor adjunto da Universidade Federal do Paraná. Tem experiência na área de Engenharia Elétrica, ênfase em Sistemas de Telecomunicações, atuando nos seguintes temas: comunicações digitais, teoria de informação e codificação, comunicações sem fio e rádio cognitivo.

Fabrcio Enembreck é bacharel em Informática (UEPG, 1997), mestre em Informática (PUC-PR, 1999) e doutor em Tecnologias de Informação e Sistemas (Université de Technologie de Compiègne, França, 2003). Atualmente é professor adjunto da PUC-PR.

Keiko Verônica Ono Fonseca: é graduada em Engenharia Elétrica (Universidade Federal do Paraná, 1985), mestre em Engenharia Elétrica (Universidade Estadual de Campinas, 1988) e doutor em Engenharia Elétrica (Universidade Federal de Santa Catarina, 1997). É professor associado 4 na Universidade Tecnológica Federal do Paraná, onde coordenou a execução do projeto multi-institucional *Formação de Pessoal Qualificado em TV Digital no Paraná* no período de set/2008–set/2012 com recursos CAPES do Edital RH-TVD 01/2007. É membro da Sociedade Brasileira de Computação (SBC), do *Institute of Electrical and Electronics Engineers* (IEEE) - *Communications Society* e do *Institute of Electronics, Information and Communication Engineers* (IEICE Japan). Atualmente é bolsista do CNPq em pós doutoramento na Alemanha na Faculdade de Informática da Universidade Técnica de Dresden (TUD).

Manoel Camillo Penna: é graduado em Engenharia Elétrica (Universidade do Estado do Rio de Janeiro, 1981), mestre em Ciências – Sistemas e Computação (Instituto Militar de Engenharia, 1986) e doutor em Ciência da Computação (Université de Paris VI – Pierre et Marie Curie, 1992). Atualmente é Professor titular da Pontifícia Universidade Católica do Paraná e membro do conselho diretor da empresa Visionnaire Informática S.A. Atua principalmente nos seguintes temas: engenharia de tráfego, gerência de redes e redes sem fio.

Marcelo Eduardo Pellenz: é graduado em Engenharia Elétrica (Universidade Federal de Santa Maria, 1993), mestre e doutor em Engenharia Elétrica (Universidade Estadual de Campinas, 1996 e 2000, respectivamente). Atualmente é professor do Programa de Pós-Graduação em Informática da Pontifícia Universidade Católica do Paraná. Atua principalmente nos seguintes temas: transmissão digital, codificação de fonte, codificação de canal, criptografia, modelagem de tráfego, redes de sensores, redes ad-hoc e redes sem fio.

Marco Antônio Costa de Borba: é graduado em Engenharia Elétrica (Universidade Federal do Rio Grande do Sul, 1980) e mestre em Ciências (Universidade Tecnológica Federal do Paraná, 2011), e coronel da reserva da Aeronáutica. Atualmente é professor da PUC-PR.

Richard Demo Souza: é graduado e doutor em Engenharia Elétrica (Universidade Federal de Santa Catarina, 1999 e 2003, respectivamente), com estágio sanduíche na University of Delaware (2003). Atualmente é Professor Associado da Universidade Tecnológica Federal do Paraná. Atua principalmente nos seguintes temas: comunicações sem fio, comunicação cooperativa, sistemas MIMO, códigos corretores de erros, ARQ e codificação da fonte.

Tânia Lúcia Monteiro: é graduada em engenharia civil (Sociedade de Ensino Superior Augusto Motta, 1987), mestre em Ciência da Computação (Universidade Federal de Santa Catarina, 2003) e doutor em Informática na área de Redes de Computadores e de Telecomunicações (Pontifícia Universidade Católica do Paraná, 2012), com estágio sanduíche na Université de Paris VI – Pierre et Marie Curie. Foi bolsista da CAPES no projeto *Formação de Pessoal Qualificado em TV Digital no Paraná*. Áreas de interesse: ciência da computação, com ênfase em metodologia e técnicas da computação, redes sem fio, conectividade, redes de computadores e protocolos de comunicação.

Wyllian Bezerra da Silva: é graduado em Física e mestre em Engenharia Elétrica (Universidade Federal de Uberlândia, 2004 e 2008, respectivamente), e doutor em Engenharia Elétrica (Universidade Tecnológica Federal do Paraná – UTFPR, 2013). É também graduando em Engenharia da Computação (UTFPR). Foi bolsista da CAPES no projeto *Formação de Pessoal Qualificado em TV Digital no Paraná*. Tem experiência na área de redes de computadores, atuando como gerente/administrador de redes e professor.

Prefácio

A transmissão de sinais de vídeo e áudio por meio de radiodifusão apresenta um rico e longo histórico de desenvolvimento. Concebida como uma técnica para transmissão de imagens à distância, a televisão se tornou um meio de entretenimento e de comunicação de massa imprescindível para a interação entre os seres humanos. Desde a demonstração do primeiro sistema de transmissão analógico, totalmente eletrônico, no final dos anos vinte do século passado, até o desenvolvimento presente, pautado pela televisão de alta definição, pela recepção em dispositivos móveis e pela interatividade, verifica-se uma enorme evolução tecnológica.

O desenvolvimento presente é ditado pelo emprego da tecnologia digital. No início dos anos 80 o Japão já se destacava por utilizar os primeiros sistemas de transmissão digital, desenvolvimento este logo seguido por europeus e americanos. À época desejava-se a adoção de um sistema que permitisse trazer o telespectador mais próximo à situação experimentada por ele em uma sala de cinema. Assim, os sistemas de transmissão e de visualização de sinais em alta definição se tornaram uma obsessão perseguida por pesquisadores e empresas. Os sistemas atuais adotam resoluções de 1920x1080 pixels, em varredura entrelaçada, ou 1280x720 pixels, em varredura progressiva. Contudo, sistemas com o dobro e o quádruplo dessas resoluções têm sido demonstrados. A evolução das técnicas de processamento de sinal e de comunicação sem fio também tem contribuído para agregar valor tecnológico aos sistemas de televisão. Particularmente, as técnicas de codificação de fonte permitiram vencer o desafio de transmitir sinais de alta resolução em canais com banda de transmissão limitada. O desenvolvimento de sistemas e padrões de codificação, do qual os sistemas MPEG-2 e H.264 são os exemplos mais conhecidos, tornou-se objeto de entidades reguladoras e de inúmeros grupos de pesquisa ao redor do mundo. A introdução em sistemas televisivos de técnicas robustas de transmissão, como é o caso da técnica OFDM (*Orthogonal Frequency Division Multiplexing*), permitiram não somente levar ao telespectador um sinal com elevada relação sinal-ruído como também viabilizou a recepção em dispositivos móveis, permitindo que o conteúdo televisivo pudesse ser assistido em meios de transporte públicos e privados. A terceira vertente do desenvolvimento em sistemas de televisão digital diz respeito à interatividade, concebida como forma de os usuários interagirem entre si e acessarem a Internet e as redes sociais. De um sistema projetado inicialmente para transmitir sinais de forma unidirecional, as plataformas de televisão atuais procuram empregar as tecnologias de acesso sem fio e por cabo para viabilizar o emprego de aplicativos específicos para as áreas de educação, saúde, entretenimento e

de comércio.

No Brasil estudos iniciais para implantação do sistema digital datam do início da década de 1990. Como premissa, foram considerados os três padrões de televisão digital existentes no mundo: o norte-americano ATSC (*Advanced Television System Committee*), o europeu DVB (*Digital Video Broadcasting*) e o japonês ISDB (*Integrated Services Digital Broadcasting*). Em 1998 um convênio realizado entre a Universidade Presbiteriana Mackenzie, a SET (Sociedade Brasileira de Televisão) e a ABERT (Associação Brasileira de Emissoras de Rádio e Televisão), com apoio da Fundação CPqD, permitiu a realização de ensaios com os três sistemas já citados, que foram executados entre outubro de 1999 e julho de 2000. Em 1998 a Rede Globo de Televisão transmitiu em alta definição a partida de futebol entre Brasil e Escócia, na Copa do Mundo da França, que marcou a primeira transmissão de um sinal digital por uma concessionária de televisão brasileira. Em novembro de 2003 o ato presidencial 4.901 instituiu o Sistema Brasileiro de TV Digital (SBTVD), com a finalidade de criar um modelo de referência para a transmissão de TV digital terrestre no Brasil. No ano de 2006, após a realização de diversos estudos e análise de resultados obtidos com os três padrões mencionados, o governo brasileiro adota o padrão japonês, ISDB-T, como base de sua televisão digital terrestre. Finalmente, no dia 2 de dezembro de 2007, é inaugurada oficialmente pela Rede Globo de Televisão a transmissão digital terrestre.

Desde o início dos estudos de viabilidade para implantação da plataforma de transmissão de televisão digital as universidades brasileiras, através de seus grupos de pesquisa e desenvolvimento, estiveram envolvidas em diversas etapas do projeto. Entre 2001 e 2006, com apoio da agências de fomento à pesquisa (CNPq e FINEP) foram desenvolvidos vários projetos que abrangiam as áreas de transmissão e recepção, codificação de fonte e de canal, camada de transporte, modulação, *middleware*, interatividade e serviços e aplicações. O resultado desses projetos auxiliaram o comitê gestor do SBTVD, do qual participavam membros dos ministérios das Comunicações, Ciência e Tecnologia, Fundação CPqD, academia e entidades empresariais, a realizar a escolha do padrão brasileiro. A variante brasileira do sistema japonês adotou a codificação H.264 e o áudio HE AAC v2, além de conceber o GINGA como uma arquitetura para o desenvolvimento de aplicações e prover as funcionalidades para a implementação da interatividade.

A adoção do sistema brasileiro de televisão digital, com prazo de implantação de dez anos, até 2016, em todo o território nacional, impôs diversos desafios, entre os quais a necessidade de capacitar pessoal qualificado para auxiliar no desenvolvimento, implantação e operação do sistema. O Brasil é carente, especialmente de engenheiros, com conhecimentos relativos ao sistema de transmissão digital, seja nos seus aspectos construtivos como nos de desempenho, cobertura de sinal e aplicações. Cientes desse

desafio, agências do governo, como a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior) têm fomentado projetos com o objetivo de qualificar pessoal de alto nível que domine as técnicas necessárias para aplicação no SBTVD e promova o desenvolvimento tecnológico brasileiro nesta área.

Foi dentro desse escopo que a Universidade Tecnológica Federal do Paraná (UTFPR), a Universidade Federal do Paraná (UFPR) e a Pontifícia Universidade Católica do Paraná (PUCPR) formaram um consórcio com o objetivo de contribuir para a formação de recursos humanos qualificados e gerar novos conhecimentos e aplicações que possam ser empregados no SBTVD. Este livro relata os resultados dessa iniciativa, descritos em quatro capítulos que compõem este volume.

O primeiro capítulo, *Ferramenta de Previsão de Desempenho do Sistema ISDB-T de TV Digital*, apresenta uma ferramenta de previsão de desempenho para o sistema de transmissão no padrão brasileiro de TV digital. A ferramenta permite, através de um software de interface amigável, avaliar o impacto causado pelas diferentes combinações de parâmetros que o sistema de transmissão pode assumir. A ferramenta é de grande utilidade para engenheiros de radiodifusão na correta configuração do sistema de transmissão e tem como o objetivo o uso mais eficiente dos recursos disponíveis, de forma que estes possibilitem um aumento na qualidade do serviço oferecido.

O segundo capítulo, *Previsão e Medição da Propagação de Sinais de TV Digital na Região Metropolitana de Curitiba*, aborda os aspectos relacionados à área de cobertura no sistema de televisão digital terrestre. Discute-se como conhecer a região geográfica servida por emissões neste sistema. Métodos de medição são propostos e discutidos com a finalidade de validar as predições através de modelos computacionais, bem como identificar precisamente áreas descobertas e orientar na sua cobertura. A região metropolitana de Curitiba é tomada como exemplo para o estudo e aplicação dos modelos de propagação. Os resultados dos modelos de cobertura foram validados através de medições realizadas com suporte de uma radiodifusora local de TV.

O terceiro capítulo, *Alternativa Tecnológica para Canal de Retorno da TV Digital baseada em Redes IEEE802.11b/g*, apresenta três mecanismos distribuídos de alocação de canais em redes IEEE 802.11b/g, sendo seu uso aplicável ao canal de retorno da TV Digital. Os métodos descritos apontam soluções ótimas e subótimas, que minimizam a interferência e otimizam a capacidade de transmissão através dessas redes.

O quarto capítulo, *Métodos Objetivos para Avaliação de Qualidade de Vídeo Digital* apresenta técnicas de avaliação de vídeo digital. Duas novas técnicas de avaliação objetiva são apresentadas: uma técnica de referência reduzida baseada na diferença de atividade dos coeficientes DCT (*Discrete Cosine Transform*) e uma técnica sem referência baseada em um

modelo sigmoidal que incorpora características espaço-temporais relacionadas à componente de luminância do sinal de vídeo. Tais técnicas têm por finalidade quantificar a degradação existente em um vídeo digital, oriunda dos sistemas de compressão e/ou ocorridos na transmissão do sinal até o ponto de visualização.

Este livro busca disseminar o conhecimento gerado no decorrer do projeto *Formação de Pessoal Qualificado em TV Digital no Paraná*, ao mesmo tempo em que pretende contribuir para o desenvolvimento tecnológico no Brasil na área de sistemas televisivos.

Keiko Verônica Ono Fonseca – UTFPR

Alexandre de Almeida Prado Pohl – UTFPR

Sumário

1	Ferramenta de Previsão de Desempenho do Sistema ISDB-T de TV Digital	1
	<i>Christofer Schwartz e Evelio Martín García Fernández</i>	
2	Previsão e Medição da Propagação de Sinais de TV Digital na Região Metropolitana de Curitiba	39
	<i>Célio Lúcio Vasco e Marco Antônio Costa de Borba</i>	
3	Alternativa Tecnológica para Canal de Retorno da TV Digital Baseada em Redes IEEE802.11b/g	89
	<i>Tânia Lúcia Monteiro, Marcelo Eduardo Pellenz, Manoel Camillo Penna, Fabrício Enembreck e Richard Demo Souza</i>	
4	Métodos Objetivos para Avaliação de Qualidade de Vídeo Digital	119
	<i>Wyllian Bezerra da Silva, Keiko Verônica Ono Fonseca e Alexandre de Almeida Prado Pohl</i>	

Ferramenta de Previsão de Desempenho do Sistema ISDB-T de TV Digital

Christofer Schwartz e Evelio Martín García Fernández *

Resumo: Neste capítulo é especificada uma ferramenta de previsão de desempenho para o sistema de transmissão no padrão brasileiro de TV digital. A ferramenta permite avaliar o impacto causado pelas diferentes combinações de parâmetros que o sistema de transmissão pode assumir através de um software de interface amigável. Os resultados obtidos permitem afirmar que esta ferramenta pode ser de grande utilidade para engenheiros de radiodifusão na correta configuração deste sistema de transmissão objetivando um uso mais eficiente dos recursos disponíveis e um aumento na qualidade do serviço oferecido.

Palavras-chave: ISDB-T, TV Digital, Parâmetros, Desempenho.

Abstract: *A performance prediction tool for the transmission system for Brazilian digital TV broadcasting is specified in this chapter. The tool allows to evaluate the impact caused by different combinations of parameters that the transmission system can take over a user-friendly software. The results obtained allow affirming that this tool can be very useful for broadcast engineers in proper configuration of the transmission system aiming at more efficient use of available resources and an increase in quality of service.*

Keywords: *ISDB-T, Digital TV, Parameters, Performance.*

*E-mails: christofer@ieee.org e evelio@ufpr.br

1. Introdução

Avanços tecnológicos nas áreas de microeletrônica e telecomunicações tornaram possível a implementação de sistemas de televisão digital, os quais permitem melhora na qualidade do sinal e adição de novos recursos em relação aos sistemas analógicos. Na atualidade existem quatro padrões de TV Digital usados no mundo.

O sistema norte americano ATSC (*Advanced Television System Committee*), primeiro padrão de TV Digital desenvolvido, teve como principal objetivo a melhoria na qualidade de som e imagem, não se preocupando com a recepção móvel. Esse padrão utiliza a técnica de modulação vestigial de oito níveis 8-VSB (*Eight-Vestigial Side Band*).

O sistema DVB-T (*Digital Video Broadcasting Terrestrial*), iniciado em 1993 na Europa, teve atenção voltada para robustez frente a interferências do canal de transmissão. Assim, o comitê criado para a concepção do padrão optou pela técnica COFDM (*Coded Orthogonal Frequency Division Multiplexing*). Outros recursos foram previstos¹, como a utilização de divisão hierárquica de transmissão e a transmissão através de SFN (*Single Frequency Network*).

O padrão japonês (ISDB-T), desenvolvido pela ARIB (*Association of Radio Industries and Businesses*) em 1998, teve de se preocupar com outro aspecto de transmissão: recepção móvel. Assim, foi utilizado o sistema de modulação BST-OFDM (*Band Segmented Transmission Orthogonal Frequency Division Multiplexing*). Esta variação de configuração OFDM é utilizada visando recepção parcial do sinal por meio de receptores portáteis. O sistema ISDB-T também conta com o entrelaçamento temporal de subportadoras, que visa minimizar as interferências causadas pela recepção móvel e multi-percurso.

Em setembro de 2006, o sistema DMB-T/H (*Digital Multimedia Broadcast - Terrestrial/Handheld*) foi criado devido ao governo chinês ser contra o pagamento de *royalties* aos países desenvolvedores dos padrões já existentes. Tal padrão é fruto da junção de duas tecnologias desenvolvidas nas universidades de Tsinghua e Shanghai Jiaotong ([Estadão, 2006](#)).

No Brasil optou-se por escolher um dos padrões já existentes. Vários testes foram feitos em território brasileiro, onde foi atestado que o padrão japonês teve desempenho superior frente a seus principais concorrentes. É importante citar que pesquisadores brasileiros tiveram participação ativa na versão do ISDB-T adquirida pelo Brasil, denominada ISDB-T_b, sugerindo algumas adaptações no sistema que foram acatadas pela ARIB, como por exemplo a codificação de vídeo H.264/AVC (H.264/*Advanced Video Coding*).

É possível afirmar que a eficiência do sistema está diretamente relacionada com a correta parametrização dos transmissores nas diferentes

¹ Disponível na internet: <http://www.dvb.org>

regiões do país. Deve-se levar em consideração o relevo da região onde se deseja realizar a transmissão e o tipo de mídia que se deseja transmitir, podendo ser SDTV (*Standard Definition Television*), EDTV (*Enhanced Definition Television*) ou HDTV (*High-definition Television*).

A eficiência do sistema de transmissão é dada por uma combinação de várias etapas de tratamento da informação que, por sua vez, possuem diversas variáveis com impacto direto na qualidade final do sinal transmitido. A Tabela 1 mostra os parâmetros que podem ser ajustados no transmissor.

Tabela 1. Parâmetros de transmissão do sistema ISDB-T

Parâmetros	Valores
Modos (subportadoras)	1 (1405); 2 (2809); 3 (5617)
Taxa do Intervalo de Guarda	1/4; 1/8; 1/16; 1/32
Tipos de Modulações	D-QPSK; QPSK; 16-QAM; 64-QAM
Taxa da Codificação Interna	1/2; 2/3; 3/4; 5/6; 7/8
Profundidade de Entrelaçamento	0; 380; 760; 1520 símbolos (modo 1)
	0; 190; 380; 760 símbolos (modo 2)
	0; 95; 190; 380 símbolos (modo 3)

O fato de existirem inúmeras variáveis dificulta o trabalho do operador do transmissor, que necessita saber qual a melhor combinação dos parâmetros a serem escolhidos, visando obter um desempenho otimizado do sistema. Sabe-se que a configuração correta desses parâmetros depende de variáveis externas dadas pelo canal de comunicação que, por sua vez, depende do relevo e do meio em que se deseja transmitir.

Tendo em mente as dificuldades dos profissionais da área em realizar uma parametrização otimizada dos transmissores, bem como possíveis limitações técnicas dos mesmos, a ferramenta descrita em [Schwartz \(2011\)](#) tem como motivação criar um software de consulta visando assessorar tais profissionais, bem como outros profissionais da área de telecomunicações, quando o assunto é ISDB-T_b. Uma descrição superficial dessa ferramenta será feita neste capítulo. A solução proposta está dividida em duas etapas:

- Criação de um simulador para gerar os resultados.
- Criação de um software de consulta de resultados.

O simulador foi desenvolvido inteiramente em MATLAB, onde os resultados obtidos através das simulações são salvos em pastas distintas utilizando uma estrutura de arquivos padronizada. Essa padronização é necessária visando facilitar a integração do simulador e do software de consulta. O software de consulta foi desenvolvido utilizando programação orientada a objetos (Delphi/Pascal) para plataforma Windows, como uma primeira implementação.

Uma simulação que apresente resultados satisfatórios para uma determinada combinação de parâmetros pode levar horas para ser concluída. Surge então um grande diferencial do sistema em análise: a criação de uma base de dados de resultados juntamente com um software de consulta de fácil utilização e interface amigável.

2. Visão Global do Sistema ISDB-T

A norma brasileira que descreve o sistema ISDB-T de TV digital, [ABNT NBR 15601:2007](#), é baseada na norma escrita pela ARIB ([ARIB STD-B31 Versão 1.6](#)). Assim, todo o conteúdo desta seção será baseado nessas referências, tornando necessário citar apenas outras referências complementares ao longo do texto. Uma descrição superficial das etapas de tratamento da informação do transmissor que são relevantes para o entendimento desse capítulo será feita. Como o sistema simulado não opera em tempo real, alguns blocos do sistema como por exemplo, o bloco de ajuste de atrasos, não são implementados, o que torna o detalhamento desses blocos desnecessários.

Os itens abordados fazem referência ao sistema de transmissão, pois não existem normas responsáveis pela construção dos receptores, sendo isso de responsabilidade dos fabricantes. Contudo, a norma prevê estruturas para auxiliar os receptores como, por exemplo, as portadoras de controle (TMCC). Essas são utilizadas pelos receptores para equalização de canal e sincronismo, onde os fabricantes dos receptores poderão utilizar métodos de equalização e estimação que julgarem mais adequados e vantajosos de serem implementados. Uma visão mais detalhada do sistema de transmissão é apresentada na Figura 1.

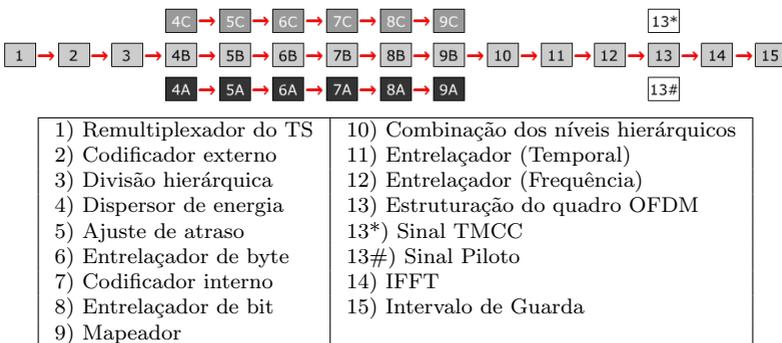


Figura 1. Diagrama de blocos do sistema de transmissão do ISDB-T_b.

2.1 Remultiplexador

A ideia principal da remultiplexação do fluxo de transporte (TS) é criar uma estrutura de informação que será utilizada pelo sistema de transmissão. Com esse intuito, um ou mais TS são divididos em estruturas de pacotes, chamadas de TSP. Assim, uma estrutura de N TSPs é chamada de quadro multiplex. N deve respeitar a quantidade de informação que irá ser transmitida dentro de um quadro OFDM. Os TSPs são compostos de 204 bytes totais, sendo que os últimos 16 bytes são nulos. Sua estrutura pode ser vista na Figura 2. Futuramente, no bloco de codificação externa, os bytes nulos serão substituídos pela redundância do código corretor de erros. A quantidade de TSPs que irá compor um quadro multiplex pode ser vista na Tabela 6 da norma [ABNT NBR 15601:2007](#).

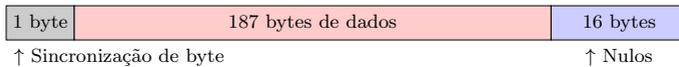


Figura 2. Estrutura de um TSP do sistema ISDB-T.

A sequência de TSPs que é criada pelo remultiplexador leva em consideração a divisão hierárquica que irá acontecer em etapas futuras. Assim, os TSPs de cada uma das três camadas são arranjados para formar uma única sequência de pacotes distribuídos de maneira que o circuito de divisão hierárquica possa separá-los entre as camadas desejadas.

2.2 Codificação de canal

O esquema de codificação de canal utilizado no sistema ISDB-T_b consiste na concatenação de um código externo de Reed Solomon (RS) e um código interno convolucional. O código externo utilizado é o código encurtado RS(204,188) que tem a capacidade de codificar um TSP inteiro, ou seja, $k = 188$ bytes, corrigindo até 8 bytes por cada bloco codificado de $n = 204$ bytes. Este código tem símbolos em $GF(2^8)$ construído a partir do polinômio primitivo $p(x) = x^8 + x^4 + x^3 + x^2 + 1$ e seu polinômio gerador é dado por:

$$g(x) = (x - \lambda^0)(x - \lambda^1)(x - \lambda^2) \cdots (x - \lambda^{15}), \quad (1)$$

onde $\lambda = 02_{\text{HEX}}$ é um elemento primitivo de $GF(2^8)$.

O código interno utilizado no sistema ISDB-T_b é um código convolucional puncionado que pode operar com taxas de codificação de 1/2, 2/3, 3/4, 5/6 ou 7/8. Isto possibilita a configuração do sistema tendo em conta o compromisso entre maior robustez (menor taxa de codificação) e maior taxa de transmissão de dados (maior taxa de codificação).

O algoritmo de decodificação mais utilizado junto à codificação convolucional é o algoritmo de Viterbi. Esse algoritmo pode operar com decisão abrupta ou com decisão suave. Na decisão abrupta, as

seqüências de entrada do decodificador são compostas por “zeros” e “uns” utilizando-se como métrica a distância de Hamming entre seqüências binárias. Na decisão suave são utilizados os valores quantizados vindos dos demoduladores utilizando-se como métrica a distância Euclidiana.

2.3 Mapeadores

Nesta etapa, as seqüências de bits são mapeadas em seqüências de símbolos complexos, sendo as constelações utilizadas: $\pi/4$ - DQPSK, QPSK, 16-QAM e 64-QAM. Os diagrama de espaço de sinais juntamente com o desempenho de erro destas constelações no canal Gaussiano podem ser vistos na Figura 3. Note que apesar da constelação QPSK parecer ser mais robusta que a DQPSK com rotação de $\pi/4$, a última oferece facilidades de sincronismo.

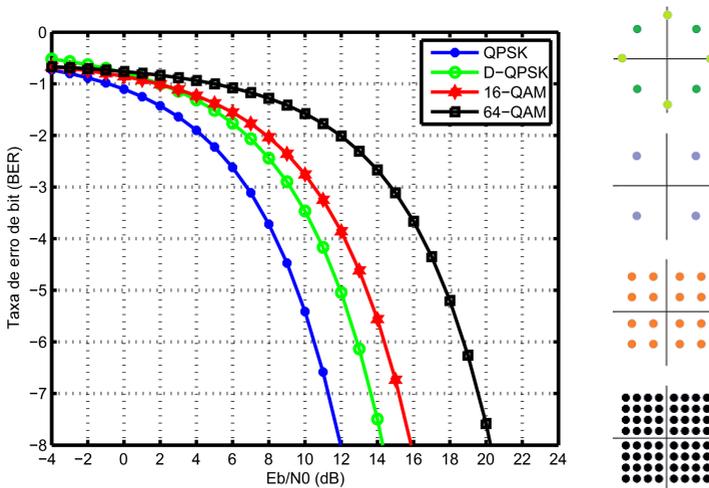


Figura 3. Constelações e desempenho das modulações do sistema ISDB-T em canais AWGN.

Por fim, após o mapeamento os símbolos gerados são normalizados conforme a Tabela 12 da norma [ABNT NBR 15601:2007](#).

2.3.1 Detecção coerente e não-coerente

Pode-se classificar as modulações de acordo com a estratégia de detecção utilizada na demodulação, são elas:

- Detecção coerente: utilizada em receptores que são capazes de estar em sincronismo com o oscilador do transmissor, tanto em frequência

quanto em fase, o qual fornece a portadora utilizada para modular o sinal de mensagem (QPSK, 16-QAM e 64-QAM).

- Detecção não-coerente: o receptor não tem conhecimento da fase da portadora do transmissor (DQPSK).

Na modulação DQPSK, cada símbolo é representado por uma “variação de fase”, diferente da QPSK, cujos símbolos são representados por um ângulo de fase fixo. A principal vantagem da detecção da modulação DQPSK está na simplicidade do receptor, que não necessita de uma referência de fase para realizar a demodulação, tornando-o mais barato e simples.

Ao analisar a Figura 3, é possível ver que o desempenho da QPSK diferencial é um pouco inferior em comparação com a QPSK coerente. Essa queda de desempenho pode ser entendida analisando uma situação hipotética, onde o receptor recebeu um símbolo corrompido com um ruído qualquer. Esse símbolo corrompido será armazenado pelo receptor para utilização na decodificação do próximo símbolo. Se o próximo símbolo vier correto, a diferença entre os dois símbolos será uma variação de fase que poderá representar outro símbolo que, por sua vez, será comparado a outro símbolo, propagando o erro.

2.3.2 Estratégias de decisão dos demoduladores

Nos receptores, as saídas digitais dos demoduladores podem ser obtidas fazendo uso de dois métodos de decisão distintos:

- Decisão abrupta (*Hard Decision*).
- Decisão suave (*Soft Decision*).

Na saída do método de decisão abrupta, o sinal é composto por zeros e uns, pois o decodificador leva em consideração o símbolo mais próximo do símbolo recebido para tomar sua decisão. Já o método de decisão suave entrega, na saída do decisor, valores inteiros quantizados que representam a probabilidade do bit ser “1” ou “0”. Quanto maior for o número inteiro mais certeza o decisor tem de que esse bit é zero. Quanto menor for o número, podendo possuir valores negativos, maior a certeza de que o bit recebido é “1”.

Existem vários modelos matemáticos utilizados para estimação suave: método dos momentos, método dos mínimos quadrados e o método da máxima verossimilhança. Alguns algoritmos de máxima verossimilhança já implementados pelo MATLAB são:

- Algoritmo de máxima verossimilhança logarítmica exata (*Exact Log-likelihood Ratio Algorithm*).
- Algoritmo de máxima verossimilhança logarítmica aproximada (*Approximate Log-Likelihood Ratio Algorithm*).

O *Log-Likelihood Ratio* (LLR), é o logaritmo da probabilidade do bit transmitido ser zero. No LLR exato são levados em conta todos os pontos da constelação para o cálculo desta probabilidade, enquanto no LLR aproximado, o cálculo da probabilidade é feito levando em consideração apenas o ponto da constelação mais próximo do sinal recebido. Sendo este último utilizado pelo simulador descrito neste capítulo.

2.4 O quadro OFDM

Variar a quantidade de subportadoras por símbolo OFDM pode ser interessante do ponto de vista de flexibilidade do sistema perante a interferências do canal de transmissão. Pode-se tornar o sistema mais robusto a interferências causadas por multipercursos aumentando o número de subportadoras, o que irá torná-lo mais frágil a interferências por efeito Doppler. Por essa razão o sistema prevê três modos de operação que torna possível o operador escolher qual a melhor configuração para a região que se deseja transmitir. As quantidades de subportadoras por símbolo OFDM utilizadas são: 1405 subportadoras no modo 1, 2809 subportadoras no modo 2 e 5617 subportadoras no modo 3.

Independente do modo de operação, um quadro OFDM (composto por 204 símbolos OFDM) é dividido em 13 segmentos no domínio da frequência. Esses segmentos são agrupados em até 3 camadas distintas, como mostra a Figura 4. Cada camada possui taxas de codificação e modulações diferentes e independentes entre si.

Em [Furht & Ahson \(2008\)](#) pode-se encontrar alguns exemplos de combinações entre número de segmentos e quantidade de camadas, como mostra a Figura 4. Na configuração 1 é usada apenas uma camada hierárquica, assim os 13 segmentos são usados para transmissão em HDTV ou múltiplas transmissões SDTV. Na segunda opção de configuração é utilizado o segmento central para recepção parcial e os demais segmentos para HDTV ou múltiplos SDTV. Por fim, a configuração 3 mostra um segmento para recepção parcial, três segmentos para recepção móvel em SDTV e nove segmentos para HDTV.



(a) Configuração 1.



(b) Configuração 2.



(c) Configuração 3.

Figura 4. Ilustração do espectro de frequências dos exemplos de distribuição de segmentos às camadas hierárquicas.

2.4.1 Sinais pilotos

Um aspecto interessante do sistema ISDB-T é a função de cada um dos quatro tipos de portadoras pilotos que o sistema possui, são eles:

- Piloto espalhado, SP (*Scattered Pilot*).
- Piloto contínuo, CP (*Continual Pilot*).
- Controle e configuração de transmissão e de multiplexação, TMCC (*Transmission and Multiplexing Configuration Control*).
- Canal auxiliar, AC (*Auxiliary Channel*).

As portadoras SP têm seu valor gerado com o auxílio de um circuito gerador de sequências pseudo-aleatórias que tem como polinômio gerador $g_{sp}(x) = X^{11} + X^9 + 1$. Os valores iniciais dos registradores de deslocamento do circuito de geração variam para cada segmento e seguem os valores dados pela Tabela 23 da norma brasileira [ABNT NBR 15601:2007](#). A saída do circuito de geração é mapeada com $-4/3 + j0$ para “1” e $4/3 + j0$ para “0”.

As portadoras CP também devem ser moduladas com modulação BPSK, e é levada em consideração a posição da portadora dentro do segmento. As CPs possuem a mesma função que as portadoras SPs, a diferença é que as CPs são utilizadas em modulações diferenciais e as SPs em modulações síncronas. Salvo que à direita de cada quadro OFDM é inserida uma portadora CP.

As portadoras TMCC utilizam modulação DBPSK e a informação por elas carregada é protegida por um código corretor de erro próprio, o código de diferença cíclica encurtado (184, 102). Composto por 204 bits, o sinal TMCC carrega a informação das taxas dos códigos corretores de erros para cada uma das camadas, tipos de modulações, profundidade de entrelaçamento, sincronismo, entre outros.

A Figura 5 (Modificada de [Sasaki \(2004\)](#)) ilustra a distribuição destas portadoras de controle para modulações síncronas e diferenciais para um segmento OFDM criado em modo 1.

Por fim, as portadoras de canais auxiliares transportam sinais adicionais para controle de sinal de transmissão, também utilizando modulação DBPSK. São previstos pela norma dois canais auxiliares (AC1 e AC2) para modulação diferencial e apenas um canal (AC1) para modulação síncrona. Para garantir diversidade de aplicações, somente a modulação DBPSK deve ser aplicada ao sinal. Na ausência de informações auxiliares, serão transmitidos apenas bits “uns” para preenchimento dessas portadoras.

A Tabela 2 mostra a distribuição dessas portadoras, cuja quantidade varia com o tipo de modulação utilizada no segmento e com o modo de operação. As posições das portadoras CP, AC e TMCC variam com a numeração do segmento em que estão contidas. No caso de segmentos com modulações diferenciais, as posições das portadoras CP, AC e TMCC nos

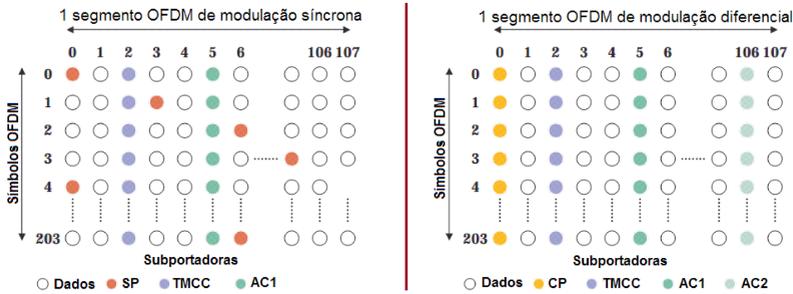


Figura 5. Exemplo de distribuição das portadoras de controle em um segmento OFDM modo 1.

modos 1, 2 e 3 são dadas pelas Tabelas 17, 18 e 19 da norma [ABNT NBR 15601:2007](#). Para modulações síncronas, as posições das portadoras AC e TMCC dentro dos segmentos são dadas pelas Tabelas 20, 21 e 22 da [ABNT NBR 15601:2007](#).

Tabela 2. Quantidade de portadoras de controle para um símbolo OFDM.

Tipo de portadoras	Modo 1	Modo 2	Modo 3
Total	$108 \times S_t + 1 = 1405$	$216 \times S_t + 1 = 2809$	$432 \times S_t + 1 = 5617$
Dados	$96 \times S_t = 1248$	$192 \times S_t = 2496$	$384 \times S_t = 4992$
SP	$9 \times S_s$	$18 \times S_s$	$36 \times S_s$
CP	$S_d + 1$	$S_d + 1$	$S_d + 1$
TMCC	$S_s + 5 \times S_d$	$2 \times S_s + 10 \times S_d$	$4 \times S_s + 20 \times S_d$
AC1	$2 \times S_t = 26$	$4 \times S_t = 52$	$8 \times S_t = 104$
AC2	$4 \times S_d$	$9 \times S_d$	$19 \times S_d$

S_s é o número de segmentos OFDM com modulação síncrona;

S_d é o número de segmentos OFDM com modulação diferencial;

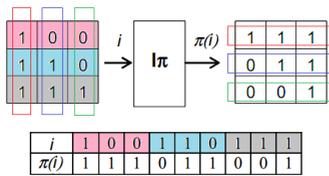
S_t é o número total de segmentos OFDM ($S_s + S_d$).

2.5 Entrelaçadores

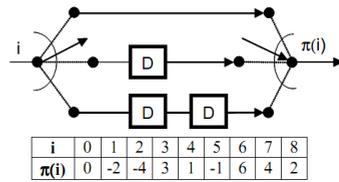
Os entrelaçadores têm o objetivo de embaralhar a informação de um vetor, aumentando o desempenho do esquema de correção de erros. Distribuindo os erros gerados pelo canal entre várias palavras codificadas, os entrelaçadores trazem a quantidade de erros de uma palavra código para dentro da capacidade de correção do código utilizado. Em outras palavras, os entrelaçadores evitam sequências de erros nos dados através da aleatorização dos mesmos. Na literatura, pode-se encontrar dois tipos clássicos de entrelaçadores ([Goldsmith, 2005](#)):

- Entrelaçadores de blocos.
- Entrelaçadores convolucionais.

Nos entrelaçadores de blocos os dados são arranjados em forma de matrizes. Com o objetivo de embaralhá-los, tais dados são lidos no sentido perpendicular ao sentido em que foram distribuídos. Assim, se os dados são arranjados no sentido das linhas, o entrelaçador irá lê-los no sentido das colunas. O entrelaçador de blocos pode ser representado por uma matriz $M_{l,c}$, sendo o período do entrelaçador $T_{ent} = l \times c$. Um diagrama de funcionamento desse tipo de entrelaçador é trazido pela Figura 6(a) (Modificada de Akamine (2004)).



(a) Diagrama ilustrativo de um entrelaçador de bloco.



(b) Ilustração do entrelaçador convolucional de $T_{ent} = 3$.

Figura 6. Exemplos de entrelaçadores de bloco e convolucional.

Já os entrelaçadores convolucionais utilizam registradores de deslocamento para atrasar os bytes ou bits, como mostra a Figura 6(b) (Modificada de Akamine (2004)), que ilustra um entrelaçador convolucional de período $T_{ent} = 3$. Mais detalhes a respeito destes entrelaçadores, juntamente com alguns parâmetros, como: causalidade, latência, memória e espalhamento, podem ser encontrados em Akamine (2004).

No sistema ISDB-T_b existem quatro blocos de entrelaçamento distintos, como mostrado na Figura 1: bloco de entrelaçamento de byte, bloco de entrelaçamento de bit, bloco de entrelaçamento no tempo e bloco de entrelaçamento na frequência. Este último é dividido em 4 etapas:

- Entrelaçamento entre os segmentos.
- Entrelaçamento dentro do segmento:
 - Rotação de portadoras por número de segmentos.
 - Aleatorização das portadoras.

De forma geral, todos têm o mesmo objetivo: aumentar a eficiência dos códigos corretores. Os entrelaçadores listados não serão descritos neste capítulo, porém, pode-se encontrar mais detalhes em Schwartz (2011).

3. Metodologia do Simulador

Os parâmetros que o simulador deve possuir estão inclusos na Tabela 1. A distribuição de segmentos entre as camadas é feita de acordo com

a Figura 4. A implementação de todos os parâmetros é perseguida, porém a modulação $\pi/4$ -DQPSK não é implementada. Essa modulação é desprezada devido ao MATLAB não possuir um método de decisão suave (*soft-decision*) para a mesma². Porém, todas as configurações previstas para modulação diferencial são implementadas (configuração de segmento, entrelaçamentos, portadoras piloto, entre outros) visando uma futura implementação da mesma. No fundo, a modulação diferencial é desejada em sistemas de recepção móvel. Como os resultados desejados *a priori* são para recepção fixa, a falta dessa modulação não acarreta grandes problemas.

Algumas considerações a respeito do transmissor e do receptor são feitas com o objetivo de tornar o simulador mais simples de ser implementado, são elas:

- Sincronização ideal entre início de símbolo OFDM (Dias, 2006).
- Sincronização ideal de frequências de portadoras (Dias, 2006).
- Sincronização de intervalo de amostragem (Dias, 2006).
- Dados transportados pelas TMCCs 100% livre de interferências do canal.
- Byte de sincronismo do início do TSP utilizado para identificação das camadas.
- Blocos de ajuste de atrasos não utilizados (necessários somente se o simulador operasse em tempo real).
- Etapas de *upconversion* (responsável por colocar o sinal na faixa de frequência que se deseja transmitir) e *downconversion* (responsável por colocar o sinal em banda base novamente) não foram implementadas devido a essas não interferirem nos resultados desejados.
- Utilização de arredondamentos de informação em quadros e vetores de dados (ocorre quando se deseja transmitir uma quantidade finita de dados que não seja suficiente para preencher todo um quadro de informação ou quando um vetor não retorna um valor inteiro ao ser fracionado).

Os blocos que são relevantes para a obtenção dos resultados desejados, contidos na Figura 1, são: 1, 2, 3, 6, 7, 8, 9, 10, 11, 12, 13, 13*, 14 e 15. Da mesma forma, pode-se visualizar na Figura 7 os blocos do lado do receptor.

Nota-se que os blocos contidos no transmissor e no receptor são praticamente os mesmos, com funções inversas. Porém no receptor temos

² Sem um método de decisão suave, o desempenho do algoritmo de Viterbi diminui consideravelmente, fazendo com que os resultados simulados se distanciem da realidade.

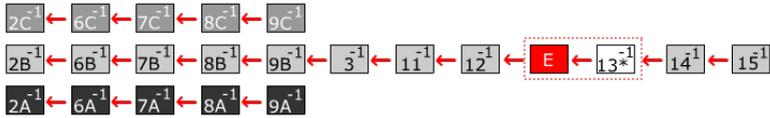


Figura 7. Blocos do receptor implementados em MATLAB.

um bloco de estimação de canal. Esse bloco tem a função de calcular a resposta estimada do canal visando corrigir as distorções causadas por ele.

Os dados utilizados para transmissão podem ser lidos externamente ou gerados pelo próprio simulador. Os dados gerados pelo simulador podem ser binários aleatórios ou provenientes de imagens TIFF e PNG. A opção de ler os dados de fontes externas existe com o intuito de tornar o simulador capaz de transmitir qualquer tipo de dados, que podem ser provenientes de vídeos, imagens, sons, entre outros. Esses dados devem obedecer uma estrutura pré-definida, onde cada linha do arquivo representa um TSP e cada coluna representa um byte do TSP.

A Figura 8 ilustra uma matriz de dados lida a partir de um arquivo externo. Caso os dados externos não possuam TSPs codificados com o código externo, essa matriz terá 188 colunas, como visto na Figura 8(a). Vale lembrar que caso os TSPs não tenham sido codificados, o próprio simulador irá codificá-los com o código RS(204,188), se essa opção estiver habilitada. Caso esses dados já estejam codificados, a matriz importada terá 204 colunas, como ilustrado pela Figura 8(b). A Figura 8(c) apresenta os dados antes de serem importados pelo simulador, visualizados com a ajuda de um editor de textos para arquivos de extensão “.txt”.

1	2	3	...	188	204
1	162	124		118	79
1	155	127		109	247
1	154	134		123	190
1	158	134		96	72
1	171	131	...	82	23
1	129	131		85	245
1	102	134		124	233
1	104	135		58	36
1	109	132		56	178
1	113	132		0	142

1	162	124	135	237	79
1	155	127	128	79	247
1	154	134	132	25	190
1	158	134	125	70	72
1	171	131	126	...	186
1	129	131	126	...	44
1	102	134	117	...	185
1	104	135	121	...	116
1	109	132	158	...	82
1	113	132	153	...	132

(a) TSPs importados sem RS. (b) c/RS (c) Arquivo externo.

Figura 8. Exemplo de estrutura de dados lidos de um arquivo externo.

Lembre-se que a primeira coluna de dados indica a camada que os TSPs pertencem, onde nas Figuras 8(a) e 8(c) todos os TSPs pertencem à camada A, representada pelo número “1”. Note que os bytes são separados uns dos outros através de um espaço simples.

3.1 Estrutura do simulador

O funcionamento do simulador pode ser entendido analisando um fluxograma que está dividido entre as Figuras 9, 11 e 12.

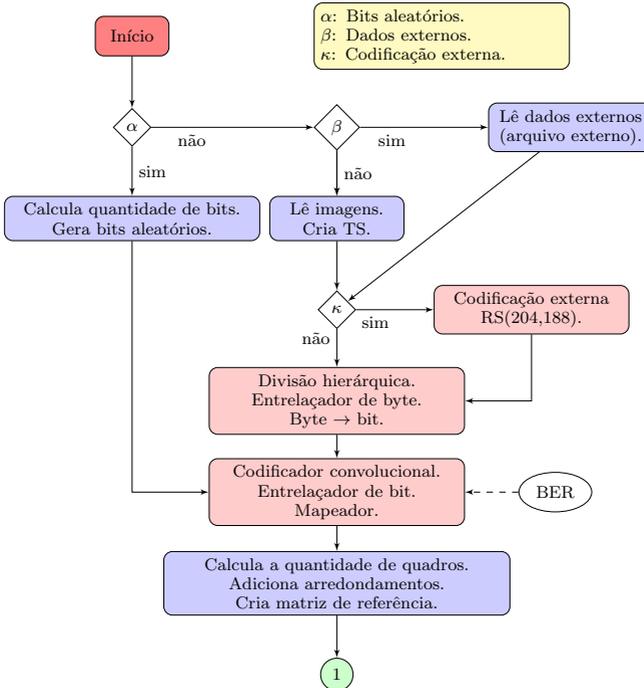


Figura 9. Fluxograma do simulador (parte 1).

Na primeira etapa, representada pela Figura 9, os parâmetros a serem utilizados na simulação já foram escolhidos: configuração, modo de transmissão, modulações, taxas de codificação, modelo de canal, entre outros. Deve-se definir se serão transmitidos dados externos ou dados gerados pelo próprio simulador.

Seguindo o fluxograma, se a opção de codificação externa estiver habilitada, a mensagem será codificada com o código RS(204,188). Em seguida os TSPs sofrem a divisão hierárquica, ou seja, o simulador separa os dados em camadas de acordo com o índice 1, 2 ou 3 localizados na posição de primeiro byte do TSP. Já separados em camadas, os dados passam por bloco de entrelaçamento de byte e posteriormente são convertidos de decimal para binário.

Note que no caso de transmissão de dados aleatórios binários, os passos descritos no parágrafo anterior, bem como a aplicação da codificação

externa, não são realizados. Ou seja, para a geração de dados binários aleatórios, o tratamento da informação acontece a partir deste momento, como ilustra o fluxograma. Assim, todas as etapas de tratamento da informação a nível de byte são poupadas, o que torna a simulação mais rápida. Com isso, o cálculo da BER considerando o esquema de codificação concatenado será feita analiticamente com base na BER calculada após a decodificação do código convolucional. Isso se dá devido a mensagem binária aleatória não passar pelo processo de codificação externa.

O próximo passo consiste em realizar a codificação interna, que irá aumentar o tamanho das sequências de dados de cada camada. Realiza-se também o entrelaçamento e mapeamento dos bits em símbolos para cada uma das camadas hierárquicas.

Em seguida, com os dados já mapeados, é calculada a quantidade de quadros necessária para realizar a transmissão. Se necessário, são inseridos alguns arredondamentos aos dados de cada camada. Nesta etapa do fluxograma, é criada uma matriz que será usada como referência posteriormente. A matriz de referência tem as mesmas dimensões do quadro OFDM e, através de representações numéricas, indica as posições que irão ser ocupadas por subportadoras e por portadoras piloto. Note que o tipo de modulação (diferencial ou síncrona) e o modo de operação do sistema influenciam nas posições das portadoras nos segmentos, que por sua vez, alteram a estrutura final do quadro OFDM. Cada tipo de portadora piloto tem uma representação numérica na matriz de referência. A Figura 10 representa parte de um segmento OFDM, em modo 1, operando com modulação síncrona.

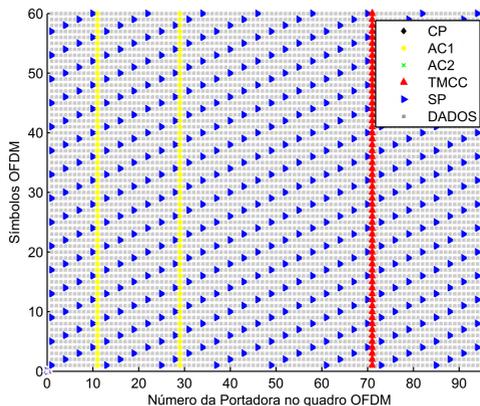


Figura 10. Exemplo de um segmento OFDM (60 símbolos OFDM).

Observando o início do fluxograma da Figura 11, vê-se que os dados que saem do bloco de arredondamento de quadros caem dentro de uma

estrutura de repetição. Tal estrutura varia de um até a quantidade de quadros OFDM necessários para passar toda a informação desejada. Em seguida são entregues aos blocos de entrelaçamento no tempo e na frequência. Após o término do processo de entrelaçamento, são adicionadas as portadoras piloto com base na matriz de referência criada anteriormente.

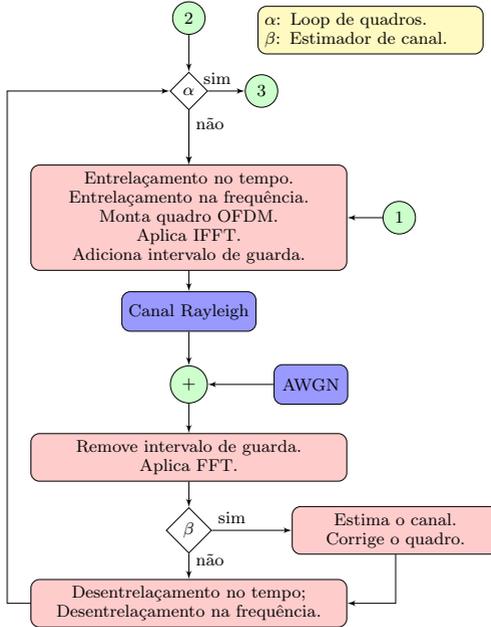


Figura 11. Fluxograma do simulador (parte 2).

Após a inserção de portadoras piloto, é feita a IFFT de cada símbolo OFDM a uma taxa de amostragem de 2048, 4096 e 8192 amostras para os modos 1, 2 e 3, respectivamente. Note que esses valores são potências de 2, o que torna a IFFT mais eficiente do ponto de vista computacional. Um símbolo OFDM irá ter 1405, 2809 e 5617 portadoras para os modos 1, 2 e 3. Para completar o tamanho da IFFT pode-se adicionar zeros no centro de cada símbolo OFDM por meio de *zero-padding*. Após a IFFT, a potência média do símbolo OFDM é normalizada, assim, a potência média das subportadoras do símbolo OFDM, C , segue a relação (Wang, 2002):

$$C = \frac{N_{\text{dados}} \times 1 + N_{\text{pilotos}} \times (4/3)^2}{N_{\text{total}}}, \quad (2)$$

onde N_{dados} representa a quantidade de portadoras de dados do símbolo OFDM, N_{pilotos} a quantidade de portadoras piloto do símbolo OFDM

e N_{total} a quantidade de portadoras do símbolo OFDM. Note que as portadoras piloto tem mais energia que as portadoras de dados.

Por fim, adiciona-se o prefixo cíclico a cada símbolo e em seguida a mensagem é entregue ao canal de transmissão. O tempo de cada amostra, na saída da IFFT, após adição do prefixo cíclico, T_{ss} , é dado por:

$$T_{ss} = \frac{T_s + T_s \times R_{IG}}{FS + FS \times R_{IG}} = \frac{T_s}{FS}, \quad (3)$$

onde T_s pode ser $252 \mu s$, $504 \mu s$ ou $1008 \mu s$, para os modos 1, 2 ou 3 respectivamente, FS é a quantidade de amostras na saída da IFFT (2048, 4096 ou 8192, para os modos 1, 2 ou 3 respectivamente) e R_{IG} representa a duração do intervalo de guarda expressado como uma fração da duração do símbolo OFDM (1/4, 1/8, 1/16 ou 1/32).

Note que ao alterar o valor de T_s , a quantidade de amostras geradas na saída da IFFT/FFT também muda, como mostra a expressão:

$$T_{ss} = \frac{252\mu s}{2048} = \frac{504\mu s}{4096} = \frac{1008\mu s}{8192} = 1,2305 \times 10^{-7} s, \quad (4)$$

isso torna o valor de T_{ss} uma constante. Fazendo o inverso desse valor, obtém-se a frequência de amostragem, F_{ss} , de um símbolo OFDM:

$$F_{ss} = \frac{1}{1,2305 \times 10^{-7}} = 8126984 \text{ Hz}. \quad (5)$$

Como consequência da aplicação da técnica de *zero-padding* no domínio da frequência, para realizar a IFFT com um valor que seja potência de dois (2048, 4096 ou 8192), ocorre um aumento na quantidade de amostras geradas no domínio do tempo. Essas amostras extras devem ser descartadas antes da transmissão para que a largura de banda do canal seja respeitada. Posteriormente as amostras descartadas são recuperadas por interpolação no receptor. Quando o número de amostras não é uma potência de dois, o MATLAB utiliza a IDFT. Nas simulações feitas em MATLAB, não se observou um ganho computacional significativo que justifique a aplicação do *zero-padding*. Desta forma, o comprimento da IFFT utilizado nas simulações é igual a quantidade de subportadoras do símbolo OFDM: 1405, 2809 e 5617 para os modos 1, 2 e 3 respectivamente. Com isso, foi poupada a remoção das amostras extras geradas pela IFFT, bem como a aplicação de um método de interpolação para recuperar tais amostras do lado do receptor.

Através de (3), pode-se calcular um novo tempo de amostragem que será utilizado pelo simulador. Assim, para os modos 1, 2 e 3 obtêm-se os valores $1,7936 \times 10^{-7}$, $1,7942 \times 10^{-7}$ e $1,7946 \times 10^{-7}$ respectivamente. Pode-se calcular também a largura de banda do canal que será ocupada pelo sinal transmitido para cada modo de operação. Como $B_w = 1/T_{ss}$, têm-se como resultados: $1/1,7936 \times 10^{-7} = 5575000 \text{ Hz}$, $1/1,7942 \times 10^{-7} =$

5573000 Hz e $1/1,7946 \times 10^{-7} = 5572000$ Hz respectivamente para cada modo de operação.

Após a mensagem passar pelo canal com desvanecimento Rayleigh, com parâmetros que serão descritos posteriormente, é adicionado ruído branco gaussiano. A quantidade de ruído a ser adicionada ao sinal, pela função *awgn* do MATLAB, é dada por:

$$SNR(\text{dB}) = \frac{E_b}{N_0}(\text{dB}) + 10 \log_{10}(k_{map}R_{cc}) + 10 \log_{10}\left(\frac{N_{total}}{N_{amostras}}\right), \quad (6)$$

onde *SNR* especifica a relação sinal-ruído (por amostras) do sinal, $N_{amostras}$ é a quantidade de amostras do símbolo OFDM, k_{map} é a quantidade de bits por símbolo do mapeador (2, 2, 4 ou 6 para DQPSK, QPSK, 16-QAM e 64-QAM, respectivamente) e R_{cc} é a taxa do codificador interno. A última parcela da equação é responsável por corrigir a quantidade de erro que será introduzida no sinal devido ao aumento da quantidade de amostras que representam um símbolo OFDM³.

Agora pode-se variar E_b/N_0 para obter o nível de ruído que se deseja adicionar na entrada do receptor. Porém, os valores de k_{map} e R_{cc} serão únicos somente se o sistema estiver utilizando apenas uma camada hierárquica (configuração 1). Para resolver esse problema, o simulador faz uma relação entre os valores de E_b/N_0 de cada camada. Para melhor entender o que isso representa, pode-se supor que estão sendo utilizadas as três camadas do sistema com os seguintes valores de k_{map} para as três camadas: 2, 4 e 6 respectivamente para as camadas A, B e C. Suponha também que as taxas dos códigos convolucionais são: 1/2, 2/3 e 3/4, respectivamente para as camadas A, B e C.

Desta forma, respectivamente para as camadas A, B e C, tem-se: $R_{cc}k_{map} = 1$, $R_{cc}k_{map} = 2,66$ e $R_{cc}k_{map} = 4,5$. Convertendo esses valores para dB, obtém-se 0dB, 4,15dB e 6,53dB para as camadas A, B e C, respectivamente. O simulador irá utilizar na Equação 6 o menor valor encontrado, ou seja, 0dB, e irá subtrair esse valor dos outros dois encontrados. Isso significa que ao entrarmos com E_b/N_0 igual a 1 na Equação 6, será obtido um valor de SNR que representa exatamente $E_b/N_0 = 1$ para a camada A, porém para a camada B, esse valor mensurado representa $E_b/N_0 = 1 - 4,15 = -3,15$ dB. Para a camada C, o ruído mensurado pela função irá representar $E_b/N_0 = 1 - 6,53 = -5,53$ dB.

Com isso, encontramos uma relação que permite avaliar o desempenho das três camadas como mostra a Tabela 3.

Seguindo com a descrição do fluxograma, após a mensagem passar pelo canal e sofrer adição de ruído, ela chega ao receptor. A primeira etapa do

³ Causado pela adição do intervalo de guarda, *zero padding* (caso seja utilizado) e devido ao aumento de amostras decorrentes do intervalo de guarda.

Tabela 3. Exemplo de relação de E_b/N_0 para três camadas.

Camadas	E_b/N_0 (dB)							
	1	3	5	7	9	11	13	15
A (-4,15dB)	-3,15	-1,15	0,85	2,85	4,85	6,85	8,85	10,85
C (-6,53dB)	-5,53	-3,53	-1,53	0,47	2,47	4,47	6,47	8,47

processo é a remoção do intervalo de guarda, em seguida é feita a FFT de cada símbolo OFDM, onde as sequências de zeros adicionados pelo processo de *zero padding* são removidas (quando utilizada essa técnica).

Se o estimador de canal não estiver habilitado, a mensagem recebida é reestruturada. Nesse processo são removidas as portadoras piloto, deixando a mensagem pronta para ser desentrelaçada na frequência e no tempo. Caso o estimador de canal esteja habilitado, antes do processo de desentrelaçamento na frequência e no tempo, o quadro OFDM é reestruturado e dividido pela resposta estimada do canal. A resposta estimada do canal é obtida utilizando um estimador de duas dimensões com método de interpolação *Cubic-Spline*. Esse processo será abordado na Seção 3.3.

Com isso, chega-se ao fim da estrutura de repetição que é responsável por repetir o processo até que sejam passados todos os quadros OFDM. Em outras palavras, o fluxograma da Figura 11 será repetido até que todos os quadros necessários sejam transmitidos.

Por fim, pode-se analisar o fluxograma da Figura 12. Nota-se que o início e o fim do segundo fluxograma analisado se encontra interno a esse terceiro. A saída do segundo fluxograma, indicada pelo círculo de número 3, é entregue a uma outra estrutura de repetição que é responsável por repetir o processo duas vezes: uma para tratar a informação do quadro sem o uso do estimador e outra para tratar o quadro estimado (se for desejada uma análise sem o uso de estimador).

Seguindo a estrutura, são removidas as subportadoras de arredondamentos de quadros de cada uma das camadas. Então, inicia-se o processo de demapeamento em *soft-decision*, seguido pelas etapas de desentrelaçamento de bits e de decodificação das mensagens pelo algoritmo de Viterbi. Note que nesse ponto os dados da saída do decodificador são utilizados para calcular a BER na recepção.

Se os bits são gerados de maneira aleatória pelo simulador, o processo termina nesse ponto, desprezando os blocos que se encontram internos ao retângulo cinza da Figura 12. Ao concluir a estrutura de repetição responsável por tratar os dados estimados e não estimados, todo o processo do segundo fluxograma é repetido devido a estrutura de repetição de níveis de ruídos. Assim, todos os quadros são transmitidos novamente para cada valor de ruído que se deseja simular (E_b/N_0). Ao final de cada repetição é calculada a taxa de erro de bits.

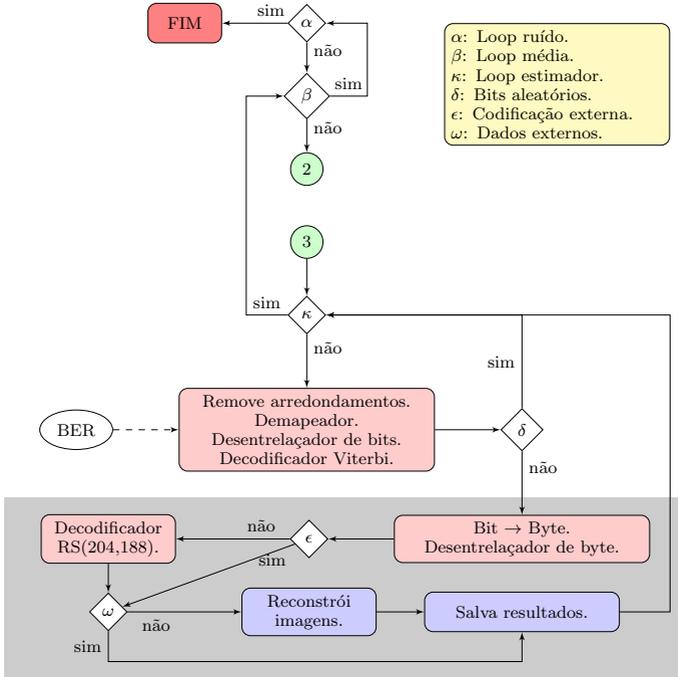


Figura 12. Fluxograma do simulador (parte 3).

Pode-se notar que entre a estrutura de repetição de nível de ruído e o círculo de número 2 (que marca o início do segundo fluxograma), existe uma outra estrutura de repetição. Esta estrutura, “Loop Média”, permite que a BER de cada nível de ruído possa ser calculada mais de uma vez, resultando em uma BER média. Desnecessária no caso de o canal de transmissão ser Gaussiano, essa estrutura de repetição é usada para que a curva de desempenho não varie entre as simulações. Quando se utiliza o canal de desvanecimento Rayleigh, é necessário repetir a simulação para cada nível de ruído mais de uma vez. Esse procedimento simula a transmissão do quadro para várias condições do modelo de canal utilizado, pois para cada repetição são calculados novos ganhos para os caminhos de multipercursos do modelo em uso. A troca dos ganhos dos caminhos (multipercursos) também é feita de quadro para quadro. Porém, em simulações em que a quantidade de quadros OFDM é muito pequena, deve-se utilizar um valor suficiente de repetições para estabilizar a curva de desempenho.

A taxa de erro de bits após a codificação externa é estimada analiticamente, tendo como base de cálculo a BER obtida após a

decodificação interna. Esse procedimento torna o simulador mais simples de ser implementado, poupando outros pontos de cálculo de taxas de erro de bit no sistema. Outra vantagem obtida é a possibilidade de que sejam traçadas curvas de desempenho do sistema como um todo, mesmo trabalhando com a geração de dados binários aleatórios, onde o processo de codificação externa não é realizado pelo simulador. A BER após a codificação externa é dada por:

$$BER = \sum_{e=t+1}^n \binom{n}{e} (BER_{cc})^e (1 - BER_{cc})^{n-e}, \quad (7)$$

onde e representa a quantidade de símbolos da palavra código que foram corrompidos por ruídos. Essa relação pode ser feita devido ao processo de entrelaçamento que garante a aleatorização dos erros.

Se a mensagem não for do tipo binária aleatória, ainda existem alguns passos antes do final das estruturas de repetições de quadro não-estimado/estimado e de níveis de ruído. As mensagens nas saídas dos decodificadores de Viterbi de cada camada são entregues ao bloco de conversão de bit para byte e passam pelo bloco de desentrelaçamento de byte. Se a codificação externa estiver habilitada, as mensagens seguem para o decodificador RS(204,188) e chegam a mais um bloco de decisão. Se a mensagem transmitida não foi gerada pelo simulador, ou seja, se foi obtida através de um arquivo externo, os dados recebidos são salvos na mesma estrutura em que foram lidos. Caso a mensagem tenha sido gerada através de imagens, as mesmas são reconstruídas e salvas.

Cada valor de relação sinal-ruído, pelo qual a mensagem é submetida, gera subpastas que irão conter os resultados obtidos. Essas pastas recebem como nome o número da repetição em que a estrutura de repetição de nível de ruído se encontra. Essas pastas ficam internas a uma pasta criada no início da simulação, fazendo com que cada simulação realizada tenha uma pasta própria.⁴

3.2 Modelos de canais utilizados

Junto ao simulador estão implementados os modelos de canais Brasil A, B, C, D e E. Esses modelos são encontrados em diversos trabalhos como, por exemplo, Baran (2006), Guimarães & Gomes (2006), Trentin (2006) e Ishini & Akamine (2009). Tais modelos descrevem alguns cenários de recepção fixa, porém, alguns trabalhos utilizam esses modelos juntamente com efeito Doppler para representar canais de recepção móvel.

⁴ Alguns *scripts* auxiliares não foram incluídos nos fluxogramas visualizados para não torná-los muito difíceis de serem analisados. Eles são responsáveis por processar e salvar alguns resultados, gerar imagens de segmentos e estrutura de quadros, salvar os passos do simulador (arquivos de *log*), criação de *tag* de identificação da simulação, entre outros.

O modelo Brasil A representa um ambiente de recepção externa que, de maneira geral, simula um sinal com múltiplos percursos de baixa intensidade e com valores de atrasos pequenos. Diferente do modelo Brasil A, o modelo Brasil B busca caracterizar um ambiente de recepção com antena interna, onde para isso são parametrizados atrasos grandes e atenuações de sinal relativamente altas, podendo até representar uma antena externa com defeitos, atingindo atenuações de até 4 dB.

O perfil Brasil C considera atrasos pequenos e de grande potência. Esse modelo pode simular a propagação em ambientes montanhosos sem linha de visada. Bem semelhante ao modelo Brasil C, o modelo Brasil D caracteriza uma recepção em ambiente com antena interna debilitada. Por fim, o modelo Brasil E avalia um cenário de recepção com rede de frequência única. Nesse caso, o receptor estará recebendo sinais de mesma amplitude defasados no tempo entre si.

A Tabela 4 mostra os valores de atraso e atenuações utilizados para parametrizar a função Rayleigh de cada um dos cinco modelos apresentados.

Tabela 4. Parâmetros dos modelos de canais Brasil.

Canais Brasil		1	2	3	4	5	6
A	Atraso (μs)	0	0,15	2,22	3,05	5,86	5,93
	Atenuação (dB)	0	13,8	16,2	14,9	13,6	16,4
B	Atraso (μs)	0	0,3	3,5	4,4	9,5	12,7
	Atenuação (dB)	0	12	4	7	15	22
C	Atraso (μs)	0	0,09	0,42	1,5	2,32	2,79
	Atenuação (dB)	2,8	0	3,8	0,1	2,5	1,3
D	Atraso (μs)	0,1	0,63	2,22	3,05	5,86	5,93
	Atenuação (dB)	0,1	3,8	2,6	1,3	0	2,8
E	Atraso (μs)	0	1	2	-	-	-
	Atenuação (dB)	0	0	0	-	-	-

3.3 Estimação de canal

É possível utilizar dois tipos de estimadores para o sistema ISDB-T:

- Estimação em uma dimensão: 1D (frequência).
- Estimação em duas dimensões: 2D (frequência e tempo).

A diferença entre essas duas formas de estimação está na leitura das portadoras de referência. A Figura 13 mostra a aquisição de portadoras SPs para estimação 1D e 2D que serão utilizadas como referência devido ao receptor conhecer os valores destas portadoras geradas pelo transmissor, ou seja, sem a interferência do canal.

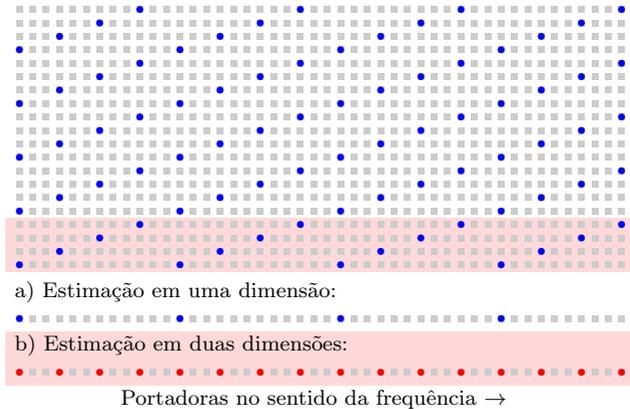


Figura 13. Exemplo de distribuição das portadoras SPs utilizadas para estimativa de canal.

Para o método de estimativa 1D, são utilizadas apenas as portadoras de referência de um único símbolo, Figura 13(a). Após a leitura destas portadoras de referências, os valores lidos são comparados com os valores conhecidos pelo receptor, calculando a resposta estimada do canal para cada SP. A relação entre sinal transmitido (antes da IFFT no transmissor) e o sinal recebido (após a FFT no receptor) para cada subportadora é dada por:

$$Y(n_s, n_p) = X(n_s, n_p)H(n_s, n_p) + N(n_s, n_p), \quad (8)$$

onde $Y(n_s, n_p)$ e $X(n_s, n_p)$ são, respectivamente, o sinal recebido e transmitido para cada subportadora do quadro OFDM, $H(n_s, n_p)$ é a resposta do canal naquele ponto e $N(n_s, n_p)$ é o ruído AWGN adicionado à subportadora. Os índices n_s e n_p representam, respectivamente, o número do símbolo OFDM no quadro e o número da subportadora (no sentido da frequência) no símbolo OFDM.

A resposta estimada do canal para cada subportadora ($\hat{H}(n_s, n_p)$), pode ser calculada através de:

$$\hat{H}(n_s, n_p) = \frac{Y(n_s, n_p)}{X(n_s, n_p)}. \quad (9)$$

Por fim, basta realizar uma interpolação entre os pontos calculados, buscando obter uma resposta estimada do canal para as demais subportadoras que estão localizadas entre as portadoras de referência.

Na Figura 13(b), são utilizadas as portadoras de referências de quatro símbolos OFDM. As portadoras de três símbolos OFDM, localizados em posições superiores no quadro OFDM, são espelhadas no símbolo em

análise. Como especificado por [ABNT NBR 15601:2007](#), as portadoras SPs são espaçadas de 12 subportadoras no sentido da frequência e de 4 portadoras no sentido dos símbolos OFDM. Assim, o espelhamento resulta em um espaçamento significativamente reduzido entre as SPs, como mostrado na [Figura 13](#). Essa diminuição de espaçamento entre duas portadoras piloto torna a interpolação mais precisa, acarretando em um ganho no desempenho do estimador.

O método de interpolação utilizado é o *Cubic-Spline*, que utiliza polinômios de terceira ordem para realizar a interpolação. As interpolações por *splines* são casos particulares da interpolação *piecewise* polinomial. Esse método de interpolação apresentou resultados superiores frente aos demais métodos: interpolação exponencial, *piecewise* e linear, avaliados por [Ishini & Akamine \(2009\)](#), o que justifica a escolha deste método.

4. Resultados e Validações

Visando obter uma validação do simulador, em [Schwartz \(2011\)](#) é feita uma comparação com os resultados apresentados em [Wang \(2002\)](#). São comparadas as curvas de desempenho das diferentes modulações para os seguintes parâmetros: $R_{cc} = 3/4$, profundidade de entrelaçamento temporal, $I = 0$, taxa do $IG = 1/8$ e modo de operação igual a 1. Em ambos os trabalhos, resultados similares são obtidos. As curvas de desempenho obtidas com a utilização do simulador construído são mostradas na [Figura 14](#).

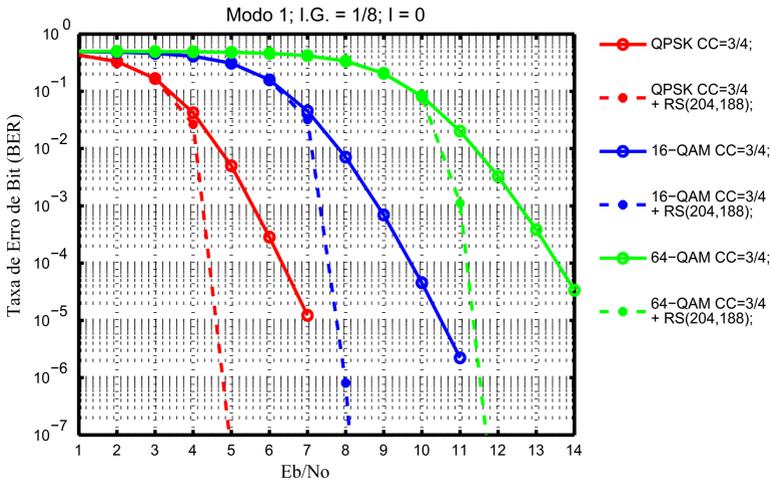
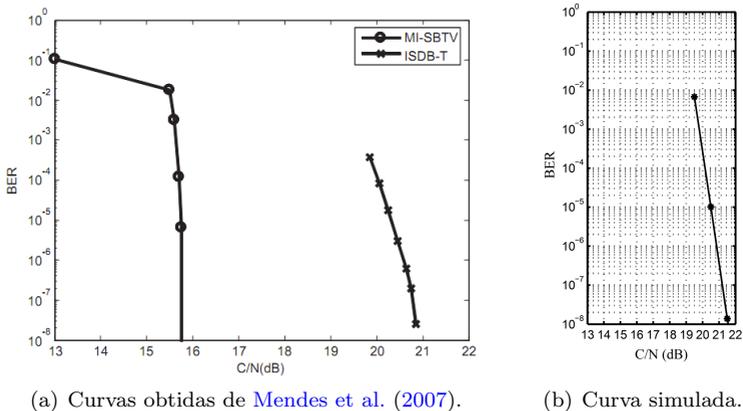


Figura 14. Curvas de desempenho do sistema em canais AWGN.

Os resultados de uma outra validação são mostrados na Figura 15 utilizando o modelo de canal Brasil A. A Figura 15(a), retirada de Mendes et al. (2007), mostra a curva de desempenho do sistema ISDB-T obtida através de medições, enquanto a Figura 15(b) mostra a curva gerada pelo simulador. Cada quadro OFDM é passado 110 vezes pelo canal e, ao final, é calculada uma média entre as taxas de erro de bit obtidas (“Loop Média” igual a 110). Pode-se observar que o resultado simulado está bem próximo do resultado obtido por medições, o que valida a implementação do estimador e do modelo de canal. Mais detalhes com relação à parametrização do simulador podem ser encontrados em Schwartz (2011).



(a) Curvas obtidas de Mendes et al. (2007).

(b) Curva simulada.

Figura 15. Curva de desempenho para modulação 64-QAM com CC de taxa 3/4 concatenado com RS(204,188), para canal Brasil A.

4.1 Bloco de estimação de canal

Para analisar o desempenho do bloco de estimação, foi criado um cenário para transmissão de um quadro OFDM que possui suas portadoras de dados mapeadas com modulação síncrona 16-QAM, modo de operação 1, configuração 1 e taxa de intervalo de guarda igual a 1/8. O modelo de canal utilizado foi o Brasil B (grandes atrasos e atenuações de sinais relativamente altas) e não foi inserido ruído AWGN (o termo $N(n_s, n_p)$ de (8) é zero). Assim, os danos causados à constelação recebida são atribuídos às características de multipercurso do canal. Isso torna possível avaliar o impacto direto do estimador, pois a inserção de ruído irá prejudicar o desempenho deste.

As Figuras 16 e 17 fazem um comparativo entre dois métodos de estimação (estimação 1D ou 2D). Note que é feita uma comparação da

resposta estimada do canal com a resposta real. A resposta real é calculada utilizando (9) para as subportadoras do símbolo OFDM.

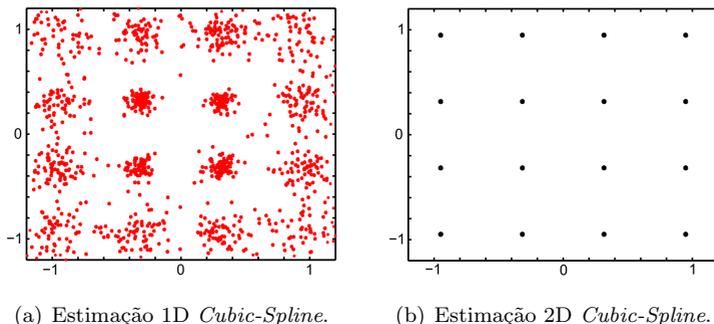


Figura 16. Constelações 16-QAM estimadas das subportadoras de 1 símbolo OFDM em modo 1 (Canal Brasil B).

Nota-se que a distância entre as subportadoras SPs não é suficiente para garantir que a interpolação seja precisa. Por menor que seja o erro de estimação, os símbolos QAM ainda irão se encontrar deslocados de sua posição ideal, como ilustra a Figura 16(a). Analisando a Figura 17(b), visualiza-se que os erros de interpolação foram significativamente reduzidos pela estimação de canal 2D. Essa melhoria é atribuída à redução de espaçamento entre as portadoras de referência. A constelação corrigida com o método de estimação 2D é mostrada na Figura 16(b).

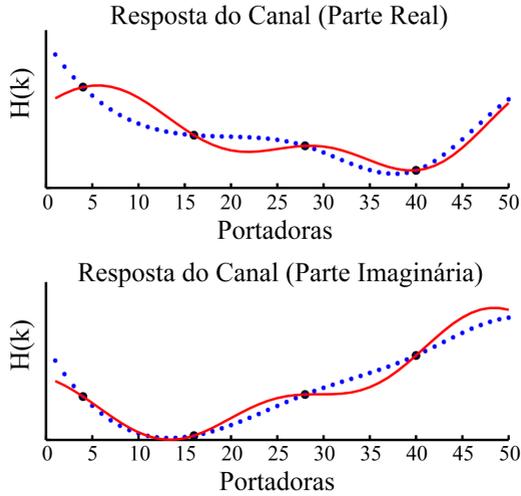
Pode-se notar que a estimação 2D *Cubic Spline* corrigiu todos os danos causados pelo canal ao símbolo OFDM analisado. Porém, na prática, isso não ocorre. Como os efeitos do canal estão acompanhados de ruído, o desempenho do estimador cai consideravelmente devido ao ruído das portadoras de referência. Essa queda de desempenho do estimador provoca, em canais AWGN, uma redução de desempenho do sistema de aproximadamente 2 dB, que pode ser vista analisando a Figura 18.

A análise feita nesta seção pode ser estendida para outras configurações de símbolos OFDM, bem como outros modelos de canais.

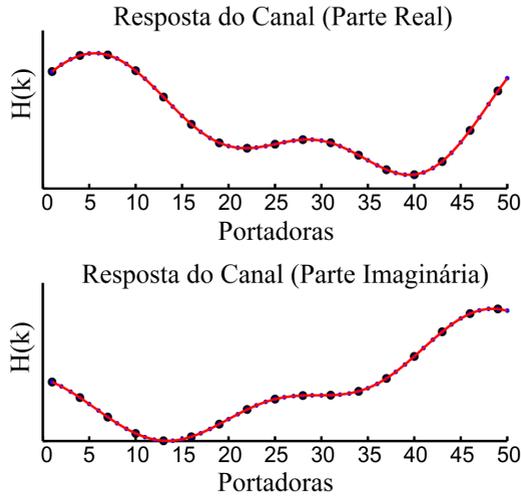
4.2 Curvas de desempenho

As curvas de desempenho ($\text{BER} \times E_b/N_0$), geradas pelo simulador, têm papel fundamental para atingir o objetivo proposto no início deste capítulo. Através destas curvas pode-se observar os impactos das variações de parâmetros do sistema de maneira clara e objetiva, verificando o ganho na robustez da transmissão.

Atualmente no Brasil, a maior parte das transmissoras estão optando por manterem parâmetros fixos no transmissor. A configuração de



(a) Estimação 1D *Cubic-Spline*.



(b) Estimação 2D *Cubic-Spline*.

Figura 17. Resultados dos estimadores de canal 1D e 2D *Cubic-Spline* para as 50 primeiras subportadoras de 1 símbolo OFDM em modo 1 (Canal Brasil B). Os símbolos representam: — Resposta calculada, · Resposta estimada, ○ Portadora piloto.

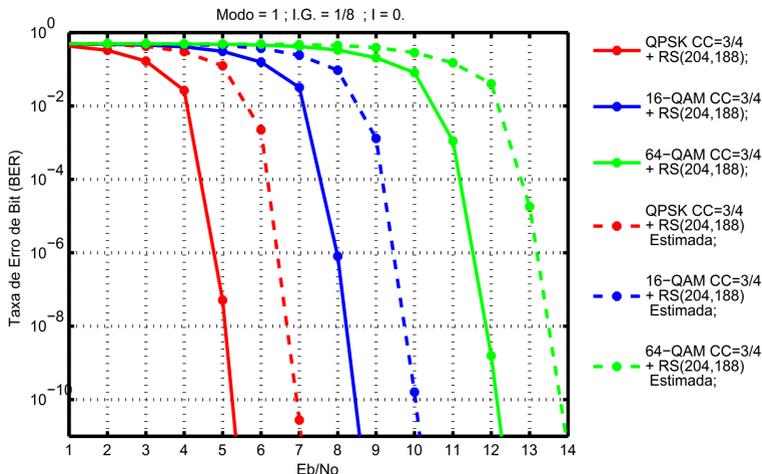


Figura 18. Queda de desempenho decorrentes de erros de estimação em canais com desvanecimento + ruído AWGN.

segmentos utilizada é a configuração 2, em que o segmento central é utilizado para recepção parcial e os demais para transmissão em alta definição. Os parâmetros utilizados são:

- Modulação: 64-QAM.
- Modo de operação: 3.
- Taxa do código convolucional: 3/4.
- Intervalo de guarda: 1/8 ou 1/16.
- Profundidade de entrelaçamento temporal (I): 0.

O entrelaçamento temporal apenas é utilizado em regiões que apresentem forte influência de ruído impulsivo. Evita-se utilizar tal entrelaçador devido ao atraso causado na recepção do sinal que é fortemente percebido quando o telespectador realiza uma troca de canal. Pode-se dizer que em uma avaliação subjetiva, o impacto do atraso na troca de canal será mais negativo do que a presença de alguns artefatos na imagem decorrentes de ruído impulsivo no sinal recebido.

Atualmente, essa configuração é mantida fixa mesmo se o vídeo a ser transmitido não for um vídeo de alta definição. Neste caso, é feita uma conversão de SDTV para HDTV, resultando em um “falso” HD. Para os parâmetros anteriormente listados, podem-se obter as seguintes taxas de transmissão (ABNT NBR 15601:2007):

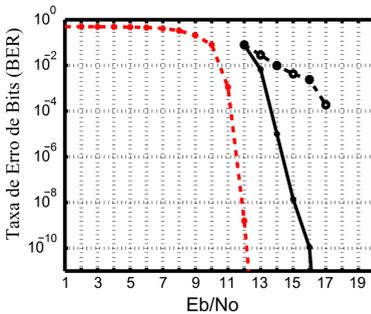
- HDTV (Modulação 64-QAM; $R_{cc} = 3/4$; $IG = 1/8$): 18,25 Mb/s.

- HDTV (Modulação 64-QAM; $R_{cc} = 3/4$; $IG = 1/16$): 19,33 Mb/s.

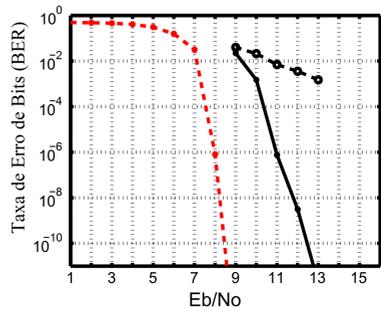
Alterando as modulações, pode-se obter as taxas de transmissão para os demais tipos de vídeos:

- EDTV (Modulação 16-QAM; $R_{cc} = 3/4$; $IG = 1/8$): 12,17 Mb/s.
- EDTV (Modulação 16-QAM; $R_{cc} = 3/4$; $IG = 1/16$): 12,89 Mb/s.
- SDTV (Modulação QPSK; $R_{cc} = 3/4$; $IG = 1/8$): 6,08 Mb/s.
- SDTV (Modulação QPSK; $R_{cc} = 3/4$; $IG = 1/16$): 6,44 Mb/s.

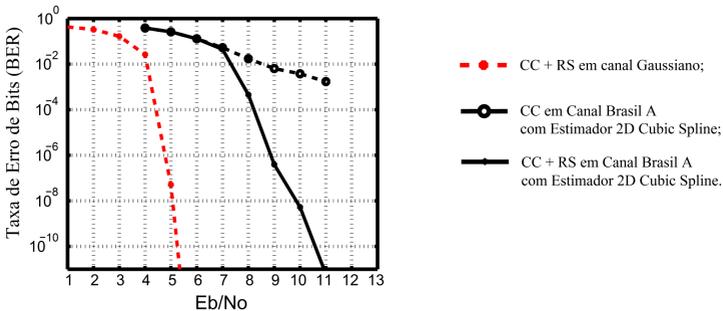
A Figura 19(a) mostra o desempenho de uma transmissão em HDTV com os parâmetros utilizados atualmente pelas transmissoras.



(a) Transmissão em HDTV



(b) Transmissão em EDTV.



(c) Transmissão em SDTV.

Figura 19. Exemplos de desempenho de transmissões em modo 3, $CC=3/4$, I.G. de $1/8$ e $I=0$.

Comparando o desempenho das transmissões, analisando as Figuras 19(a) e 19(c), pode-se notar um ganho significativo de robustez no sinal. Na

prática esse ganho pode ser aplicado no caso de parametrizar o transmissor de forma correta quando o vídeo transmitido não estiver em alta definição. Assim, o processo de “conversão” de SDTV para HDTV seria poupado e o sinal ficaria mais imune às interferências do canal. Ganhos em área de cobertura do sinal transmitido também podem ser obtidos (Schwartz, 2011).

Através de simulações, outras comparações podem ser feitas como, por exemplo, alterações na taxa do intervalo de guarda. Com isso, uma previsão do desempenho da transmissão pode ser feita ao se consultar a base de dados.

Para apresentar os resultados o simulador gera duas figuras distintas. Uma para camadas individuais e outra com os resultados das transmissões de todas as camadas. A primeira figura contém a curva simulada para a mensagem codificada com o código convolucional e a curva de desempenho do código convolucional concatenado com o código RS(204,188). Na segunda figura gerada pelo simulador são apresentados somente os resultados da curva de desempenho do código convolucional concatenado com o código RS(204,188), porém para as três camadas (quando utilizadas).

Com objetivo didático, são gerados algumas figuras que visam facilitar o entendimento de algumas partes do sistema ISDB-T. São elas: figura dos TSPs remultiplexados, figura da distribuição de segmentos e representação gráfica da inserção do intervalo de guarda. Também são gerados resultados que buscam mostrar os impactos sofridos pela transmissão de forma amigável: vídeos das constelações e imagens transmitidas (TIFF ou PNG).

A Figura 20 mostra uma imagem PNG recebida utilizando o modelo de canal Brasil A ($E_b/N_0 = 5\text{dB}$). Foram necessários 3 quadros OFDM, com arredondamento no último, para concluir a transmissão (6% de uma imagem de resolução $1920 \times 1080 \times 3$). Os parâmetros utilizados para realizar a transmissão desta imagem são: modulação QPSK, intervalo de guarda $1/8$, $R_{cc} = 3/4$, $I = 0$ e modo de operação igual a 3.

Note que o sistema ISDB-T_b utiliza codificação de vídeo H.264/AVC, desta forma, a imagem PNG não irá apresentar efetivamente os artefatos que o vídeo transmitido teria na prática. Para ter uma visão mais realista, seria necessário transmitir um vídeo codificado. Porém, a imagem transmitida pode ser utilizada pelo usuário do software de visualização de resultados, com o objetivo de facilitar a interpretação das curvas de desempenho apresentadas. Tais imagens dão uma ideia visual do impacto do ruído ao sinal transmitido.

5. O Software Criado

O software de visualização de resultados, desenvolvido em Delphi, recebeu o nome de “ISDB-T Analyser”. Compilado para plataforma Windows, o software expõe os resultados das simulações da seguinte forma:

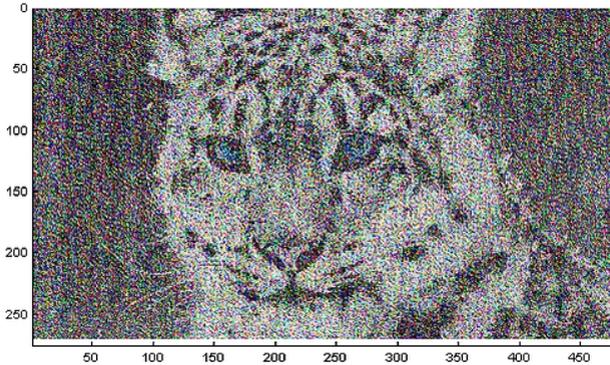


Figura 20. Exemplo de imagem PNG transmitida pelo simulador.

- Curvas de desempenho da simulação ($BER \times E_b/N_0$).
- Imagens TIFF/PNG recebidas para cada nível de ruído.
- Vídeos das constelações para cada nível de ruído.
- Estrutura do TS remultiplexado.
- Configuração do quadro OFDM.
- Representação gráfica da inserção do intervalo de guarda.
- Relatório contendo os passos realizados pelo simulador.

Tal software está preparado para receber a estrutura de arquivos criada pelo simulador. Essa estrutura é criada de maneira automatizada, garantindo que o usuário não precise estruturar cada simulação manualmente. Outro ponto a se destacar é que o software gera os vídeos das constelações através de um vetor de dados salvo dentro das pastas de simulações, isso para cada nível de ruído.

De maneira geral, a estrutura do ISDB-T Analyser é bem simples, e pode ser analisada com o auxílio do fluxograma da Figura 21. Esse fluxograma é simplificado, onde ficaram de fora blocos de verificação de existência de arquivos, blocos de avisos, blocos de *skins*, importação de simulações e blocos de opções extras.

Ao iniciar o programa, o software realiza uma busca por simulações e lista as encontradas. Caso não encontre nenhuma, avisa ao usuário da necessidade de importar as simulações. Caso encontre arquivos de simulações já importados, os parâmetros destas simulações serão incluídos nas listas de parâmetros disponíveis para consulta.

Em seguida, o programa aguarda a seleção dos parâmetros desejados e, quando requisitado, o programa busca por simulações que contenham tais parâmetros. Caso a busca retorne vários resultados, uma lista

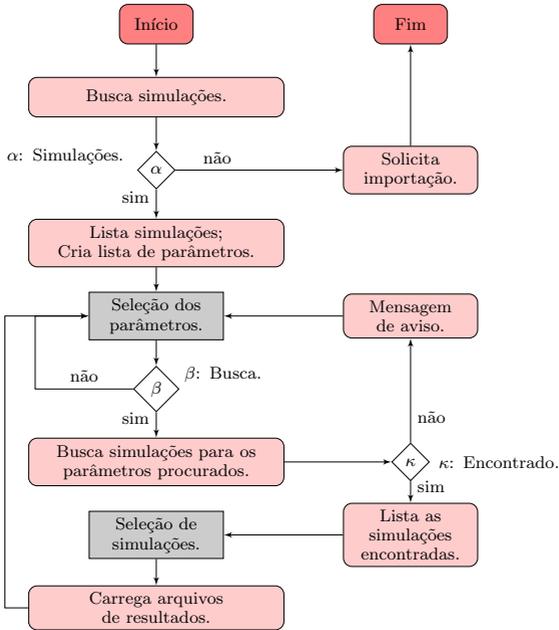


Figura 21. Fluxograma do software de visualização de resultados.

de simulações é criada para que o usuário selecione a desejada. Após selecionada, os dados da simulação são carregados e os botões de visualização de resultados são habilitados. Quando pressionados, os botões mostram as janelas responsáveis por exibir o resultado desejado.

Como o objetivo do software é auxiliar os profissionais que trabalham com os transmissores, eventualmente acadêmicos ou pessoas que não dominam a teoria do sistema por completo, deseja-se que este tenha arquivos de ajuda com comentários, dicas e explicações para auxiliar no entendimento dos resultados consultados. Com esse fim, uma base de dados de informações foi criada de modo que possam ser inseridos comentários e sugestões dos próprios usuários do software.

Essa base de dados de informações possui uma interface independente do software de visualização de resultados e utiliza Access como base de dados (*Data Base*, DB). A escolha desse tipo de base de dados é justificada pelo fato de não ser necessário a instalação de nenhum tipo de software extra para acessar o conteúdo do DB, sendo que o Delphi possui *drivers* destinados para leitura e escrita desse tipo de base de dados.

5.1 O ISDB-T analyser

Ao executar o programa a primeira imagem que o usuário irá visualizar é a mostrada na Figura 22. A primeira barra de *status* indica o andamento da varredura das pastas de simulações importadas anteriormente. Após a varredura ser concluída, o programa irá processar os dados de configuração encontrados na busca anterior. Esse processamento de dados consiste em selecionar os parâmetros de cada simulação encontrada, eliminar os parâmetros repetidos e adicioná-los às caixas de seleção. Desta forma, apenas os parâmetros que se encontram efetivamente simulados e importados irão aparecer como opção dentro do programa.

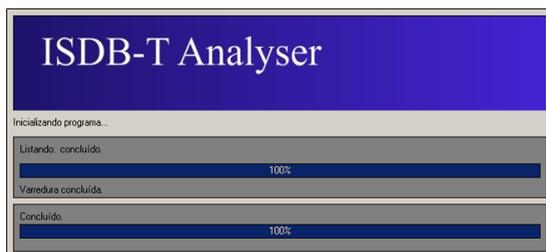


Figura 22. Janela de *splash* do software de visualização de resultados.

Caso o programa não encontre nenhuma simulação, ele sugere a importação das mesmas. Se o usuário optar por importá-las, ele será conduzido a uma janela igual à ilustrada na Figura 23. Após a seleção do caminho onde se encontra a pasta contendo as simulações, deve-se clicar em buscar simulações. Desta forma, o programa irá localizar as pastas que contenham simulações válidas, ou seja, cuja estrutura de arquivos contidos nestas pastas estejam de acordo com a estrutura requerida pelo programa. Se não for a primeira vez que o usuário estiver importando simulações, após localizá-las, o programa irá verificar pastas de simulações duplicadas. Caso já existam simulações com os mesmos parâmetros que as simulações encontradas, o software pergunta ao usuário se deseja remover as simulações anteriormente importadas.

Considerando que o processo de importação de simulações já tenha sido concluído, o usuário será levado à janela principal, onde pode-se selecionar os parâmetros das simulações que deseja localizar. Se o usuário preferir, ele pode selecionar a simulação na barra de ferramentas, selecionando a opção “lista de simulações”. Essa última opção facilita o uso do programa no caso do sistema possuir poucas simulações armazenadas.

Após feita a seleção dos parâmetros, basta clicar no botão “Buscar Resultados” e o programa irá listar as simulações encontradas para os parâmetros inseridos, como mostra a Figura 24 (parâmetros selecionados

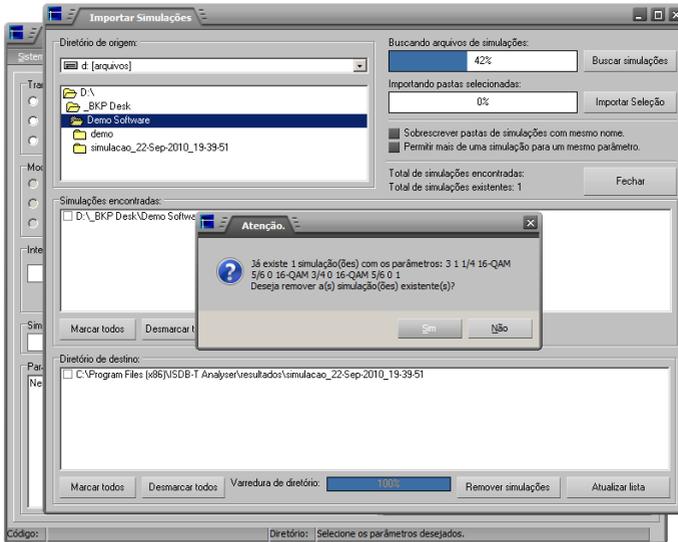


Figura 23. Janela de importação de simulações do ISDB-T Analyser.

nesta ilustração são fictícios e não representam uma configuração comumente usada).

Após selecionar uma simulação dentre as encontradas, os botões de visualização de resultados são liberados. Ao clicar em “constelações recebidas”, uma janela igual a da Figura 25 irá aparecer. Essa janela é responsável por gerar os vídeos das constelações de cada uma das camadas, sendo essas estimadas ou não. Para gerar os vídeos, são necessários converter os dados salvos nos arquivos de textos em coordenadas numéricas do quadro de reprodução. Essa conversão pode ser feita em tempo real ou com um esquema de pré-processamento dos dados. Por padrão, os dados são tratados em tempo real, porém, caso o computador no qual o software se encontra instalado não tenha recursos para realizar o processamento dos dados em tempo real, pode-se selecionar a opção “pré-processamento” para preparar os dados antes da reprodução do vídeo. Outro recurso de reprodução do vídeo é o “zoom”, que pode ser ajustado pelo usuário da maneira que ele preferir. Cada ponto do vetor de símbolos, gerado pelo simulador, é um ponto na constelação. Desta forma, o usuário pode selecionar a quantidade de pontos por quadro que serão plotados e a quantidade de quadros por segundo, alterando a velocidade de reprodução do vídeo. O tamanho do ponto na imagem também pode ser alterado.

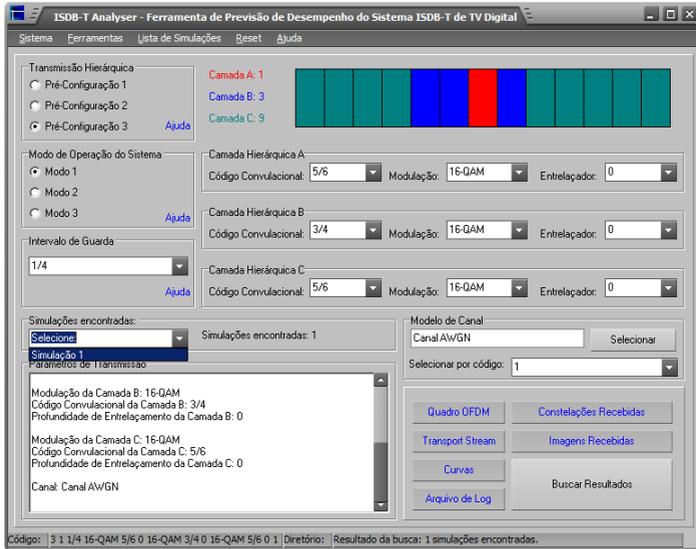


Figura 24. Janela principal do ISDB-T Analyser.

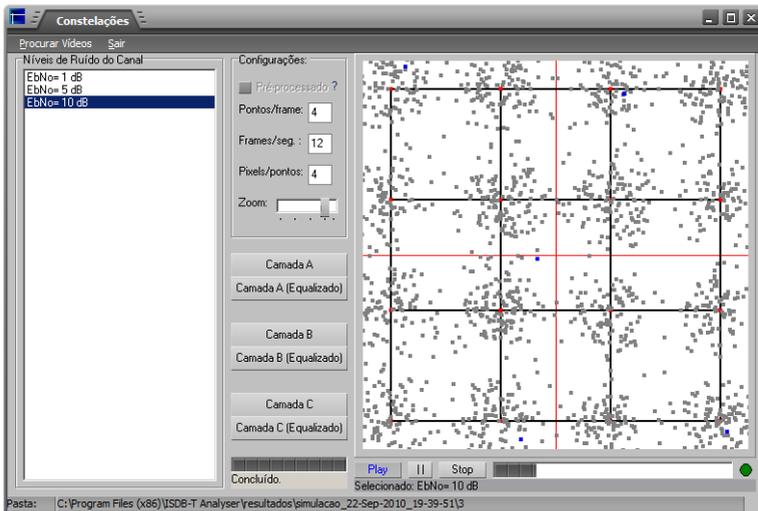


Figura 25. Janela de geração do vídeo das constelações recebidas.

Retornando à janela principal, caso o usuário clique na opção “Quadro OFDM”, uma janela igual à da Figura 26 irá aparecer. Essa janela irá mostrar a imagem da distribuição dos segmentos no quadro OFDM e a adição do intervalo de guarda ao símbolo. A mesma estrutura de janela é usada para mostrar as informações do TS remultiplexado gerado pelo simulador. Basta clicar em “Transport Stream”, na janela principal do programa. No canto direito da janela, o usuário pode consultar as descrições e comentários deixados pelos usuários a respeito destes parâmetros. Essas informações se encontram no banco de dados do programa e podem ser alteradas através do editor de informações a qualquer momento.

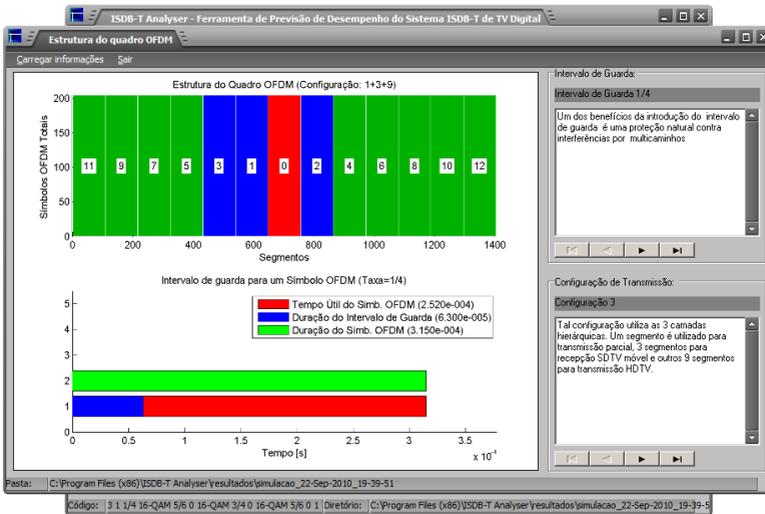


Figura 26. Janela de visualização da distribuição de segmentos do quadro OFDM e adição do intervalo de guarda.

O software conta com algumas opções extras, como por exemplo, tabela de portadoras de controle, *skins* e botões de “ajuda rápida”. Indo em “Ferramentas”, no menu superior, e clicando em “Tipos de portadoras” o usuário poderá ver a quantidade de portadoras totais necessárias para compor um quadro OFDM com os parâmetros selecionados por ele.

6. Considerações Finais

Foi apresentada uma solução relativamente simples de previsão de desempenho do sistema ISDB-T para diferentes configurações do transmissor. Uma abordagem sucinta das várias etapas de tratamento

da informação que o sistema ISDB-T possui foi feita. Nesta abordagem pode-se verificar que esse sistema tem capacidade de operar com várias configurações de transmissão.

As implementações propostas foram apresentadas de maneira superficial, dando uma noção das etapas de tratamento da informação do simulador que foram implementadas em MATLAB. Algumas formas de visualização de resultados, juntamente com alguns exemplos, foram vistas e discutidas. Um exemplo bem prático de como o software proposto pode ser útil para as transmissoras também foi comentado visando dar ao leitor uma ideia da importância da correta parametrização do transmissor, justificando o trabalho aqui apresentado.

O princípio de funcionamento do software de visualização de resultados também foi descrito. Algumas janelas foram apresentadas e descritas. Em [Schwartz \(2011\)](#) algumas ferramentas de auxílio trazidas pelo software (editor de banco de dados, tópicos de ajuda, entre outros) são ilustradas e podem ser consultadas pelo leitor que estiver interessado em uma descrição mais aprofundada.

Um ponto importante a ser considerado é a total independência entre o simulador e o software de visualização de resultados. Isso torna possível a geração de resultados por parte de vários grupos de estudos, os quais podem compartilhar os resultados obtidos e inserir comentários à base de dados através do editor, garantindo a troca de informação e o aperfeiçoamento dos resultados em questão.

Por fim, vale frisar que para maiores detalhes a respeito do trabalho descrito neste capítulo, bem como sugestões de implementações futuras, deve-se consultar [Schwartz \(2011\)](#).

Agradecimentos

Os autores agradecem à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pelo apoio financeiro ao projeto, bem como ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ). Um agradecimento também à Arielle Ribeiro Golin pela análise ortográfica do texto.

Referências

- ABNT NBR 15601:2007. *Televisão Digital Terrestre - Sistema de Transmissão*. Associação Brasileira de Normas Técnicas, 2007. Revisão 2008.
- Akamine, C., *Análise de desempenho do sistema de TV digital DVB-T com Entrelaçador Temporal em Canais com Ruído Impulsivo*. Dissertação de mestrado, Universidade Estadual de Campinas - Pós-Graduação em

- Engenharia Elétrica - Departamento de Telecomunicações, Campinas, SP, 2004.
- ARIB STD-B31 Versão 1.6. *Transmission System For Digital Terrestrial Television Broadcasting*. Association of Radio Industries and Businesses, 2001. Revisão 2005.
- Baran, I.R., *Explorando a Coerência Temporal e a Comunicação Oportunística em Sistemas de Comunicação Móvel Multiusuário*. Dissertação de mestrado, Universidade Federal de Santa Catarina, Florianópolis, SC, 2006.
- Dias, J.T., *Desempenho de Algoritmos para Estimação de Parâmetros de Sincronização em Sistemas OFDM*. Dissertação de mestrado, Instituto Militar de Engenharia, Rio de Janeiro, RJ, 2006.
- Estadão, , TV digital da China vai movimentar US\$ 62,5 bilhões em 2015. <http://www.estadao.com.br>, 2006.
- Furht, B. & Ahson, S., *Handbook of Mobile Broadcasting - DVB-H, DMB, ISDB-T and MEDIAFLO*. Internet and Communications. New York, USA: Auerbach Publications, 2008.
- Goldsmith, A., *Wireless Communication*. New York, USA: Cambridge University Press, 2005.
- Guimarães, D.A. & Gomes, G.G.R., Caracterização do canal para o MI-SBTVD. *Revista Telecomunicações*, 9(1):32–45, 2006.
- Ishini, A.K. & Akamine, C., Técnicas de estimação de canal para o sistema ISDB-T_b. *Revista de Radiodifusão*, 3(3), 2009.
- Mendes, L.L.; Brito, J.M.C.; Cardoso, F.A.; Guimarães, D.A.; Lima, G.C.; Gomes, G.G.R.; Arantes, D.S. & Souza, R.D., MI-SBTVD: a proposal for the brazilian digital television system SBTVD. *Journal of the Brazilian Computer Society*, 12(4), 2007.
- Sasaki, M., Technologies and services of digital broadcasting - terrestrial digital television broadcasting. *Broadcast Technology*, 20:14–19, 2004.
- Schwartz, C., *Ferramenta de Previsão de Desempenho do Sistema ISDB-T de TV Digital*. Dissertação de mestrado, Universidade Federal do Paraná - Departamento de Engenharia Elétrica, Curitiba, PR, 2011.
- Trentin, R.G., *Técnicas de Processamento MIMO-OFDM Aplicadas à Radiodifusão de Televisão Digital Terrestre*. Dissertação de mestrado, Universidade Federal de Santa Catarina, Florianópolis, SC, 2006.
- Wang, X., *Performance Evaluation of the ISDB-T Standard for Multimedia Services*. MSc dissertation, University of British Columbia - Department of Electrical and Computer Engineering, Vancouver, Canada, 2002.

Previsão e Medição da Propagação de Sinais de TV Digital na Região Metropolitana de Curitiba

Célio Lúcio Vasco e Marco Antônio Costa de Borba*

Resumo: É um interesse permanente das entidades de teledifusão conhecer a região geográfica servida por suas emissões. Modelos baseados em métodos de predição normalizados fornecem informações sobre a cobertura, fundamentando as decisões de investimentos nas fases de projeto e de licenciamento. Entretanto, a medição do sinal irradiado é imprescindível para validar as predições e para identificar áreas sem cobertura. Neste capítulo, dois estudos de caso são apresentados, baseados em conjuntos específicos de dados. No primeiro, alguns modelos de propagação foram comparados aos dados coletados, e foi proposto um modelo empírico para ser utilizado na região de Curitiba. No segundo caso, a ênfase foi na medição e no tratamento de um significativo volume de dados em uma estação já instalada de TV Digital, para detalhar a cobertura efetivamente atingida e orientar o seu aperfeiçoamento.

Palavras-chave: Modelos de propagação, Medidas, TV Digital.

Abstract: *Knowing the exact region served by a broadcasting station is a permanent interest. Models based on standardized prediction methods provide maps of the coverage, essential to the decision-making process of the investors in the design phase. However, measurement of irradiated signals is still required to validate the predictions, as well as to investigate situations that may give rise to reception difficulties. Two case-studies are presented in this chapter, based in specific sets of measurements. In the first one, some current propagation models were compared to the data obtained, and a new model has been proposed for planning the coverage in Curitiba region. In the second case, emphasis was moved to measurements and processing a significant amount of data collected after the digital TV station start up, in order to access the details of the true coverage and support improvements.*

Keywords: *Propagation models, Measurement, Digital TV.*

*E-mail: marcoborba2010@gmail.com

1. Introdução

O SBTVD-T *Sistema Brasileiro de Televisão Digital Terrestre* foi instituído pelo Decreto n° 4.901, de 26 de novembro de 2003 (Brasil, 2003), e teve sua implementação determinada pelo Decreto n° 5.820 (Brasil, 2006), de 29/6/2006, que instituiu o sistema japonês ISDB-T (*Integrated Services Digital Broadcasting - Terrestrial*), com pequenas alterações, como padrão de transmissão e definiu um prazo de dez anos como período de transição do padrão analógico para o digital.

O início oficial das transmissões de TV Digital no Brasil ocorreu no dia 2 de dezembro de 2007, quando se iniciaram as transmissões na cidade de São Paulo. Desde então, novas implantações em outras cidades do país vêm seguindo um cronograma definido pelo Ministério das Comunicações.

As entidades de teledifusão, privadas ou públicas, têm um interesse permanente em conhecer a região geográfica servida por suas emissões. Modelos computacionais podem fornecer informações detalhadas sobre a área coberta, ao implementar variados métodos de predição baseados em normas nacionais (Anatel, 2005, 1991) e internacionais (Radiocommunication Study Groups ITU, 2003, 2007b). Enquanto tais métodos computacionais são indispensáveis à definição dos investimentos, e para que os órgãos governamentais reguladores definam frequências, potências e localização das emissoras, *medir* o sinal após sua implementação é atividade imprescindível para validar as predições (Radiocommunication Study Groups ITU, 2005), e para identificar precisamente áreas de recepção difícil, a fim de solucionar as deficiências (Pullen, 1999).

Modelagem computacional e medição direta não são abordagens que competem entre si; são antes complementares. Até porque há métodos computacionais de predição importantes que também empregam modelos empíricos, baseados em um número elevado de medições, como o de Okumura-Hata e o da Recomendação 1546-1 da ITU (Radiocommunication Study Groups ITU, 2003).

Os radiodifusores da capital do Paraná sentiram a necessidade de avaliar a condição da propagação na região de Curitiba nesta nova conjuntura de TV digital, com novos canais de radiofrequência em UHF e com novos limiares de recepção definidos pela norma ABNT NBR 15604 (ABNT, 2007).

Com o intuito de fomentar a pesquisa na área de TV Digital em Curitiba, foi feita uma parceria entre a UTFPR (*Universidade Tecnológica Federal do Paraná*) e a RPC (*Rede Paranaense de Comunicação*), que cedeu a infraestrutura de transmissão e recepção para que pudessem ser feitos diversos testes de propagação de sinais de radiofrequência já transmitidos em TV Digital.

Dois estudos de casos são apresentados neste capítulo, com ênfases na *previsão* e na *medição* respectivamente:

- A análise da propagação de sinais de TV Digital na Região Metropolitana de Curitiba, que foi elaborada a partir da medição dos níveis de recepção em 223 pontos nos arredores de Curitiba, distribuídos entre as distâncias de 0,21km até 57,7km, coletados de junho a setembro de 2008. Nesse primeiro caso, as medições foram feitas com o equipamento receptor *parado* nos diversos pontos de medidas e cobrindo toda a região de cobertura da transmissão. Esse estudo de caso está descrito na Seção 2.
- A medição de cobertura em regiões densamente edificadas de Curitiba, que foi realizada a partir da medição de 1580 pontos em junho de 2010. Nesse segundo caso, as medições foram feitas com o equipamento receptor operando em um carro *em movimento*, permitindo a coleta sequencial dos dados, em diversos trechos de medidas, com alguns quilômetros de comprimento. Esse estudo de caso está descrito na Seção 3.

Descrever métodos apropriados para *prever* e para *medir* a cobertura de um sinal de TV digital no sistema brasileiro ISDB-T_B (*Teledifusão Digital de Serviços Integrados-versão B*) na cidade de Curitiba é o objetivo deste capítulo.

2. Previsão da Propagação

O início dos ensaios da transmissão de TV Digital em Curitiba deu-se a partir da autorização para a UTFPR executar o Serviço Especial para Fins Científicos ou Experimentais conforme o Ato nº 1388 (Anatel, 2008a), de 12/3/2008, emitido pela Anatel (*Agência Nacional de Telecomunicações*). Utilizou-se um transmissor de fabricação nacional da Linear Equipamentos Eletrônicos S.A., modelo IS74K1 de 4,1kW, operando no canal 41 de UHF (635,142857MHz), conectado a uma antena Transtel Conti & Cia Ltda, modelo TTSL8 U0 - 41-10D, instalada à 90m de altura. Para as medições de recepção utilizou-se uma unidade móvel, contendo um mastro telescópico de 8m, instalado com uma antena log-periódica do fabricante Proeletronic, banda total UHF, conectado a um analisador de espectro do fabricante Agilent modelo ESA E4405B para a medição dos níveis de recepção.

2.1 Metodologia

A metodologia aplicada foi realizar um levantamento em campo de medições do nível de recepção dos sinais de radiofrequência em diferentes pontos da região de Curitiba. Foram coletadas medidas em 223 pontos avaliados no período de 13 de junho de 2008 até a data de 25 de setembro de 2008, distribuídos pelo município de Curitiba e Região Metropolitana,

distribuídos entre as distâncias de 0,21km até 57,7km, em diferentes azimutes e em locais com a possibilidade de acesso por automóvel, de modo a cobrir a maior área possível da região, especialmente em áreas com maior aglomeração urbana (Vasco et al., 2009). Analisou-se os dados coletados em campo nos 223 pontos, contendo as medidas de recepção da potência de canal para uma banda passante de 6MHz, a potência média para uma banda de 30kHz e a relação sinal/ruído do canal, para as condições de potência de transmissão em 4,1kW e 1kW, que são relevantes para a análise de propagação dos sinais. Como foram evidenciadas algumas inconsistências entre as medidas realizadas para um mesmo ponto, foi feita uma pré-seleção dos dados coletados para retirar estes pontos anômalos, o que resultou em um conjunto final de 174 pontos válidos. As estimativas de propagação foram feitas utilizando somente a medida da potência de canal para 6MHz na condição de transmissão de 4,1kW. Esta Seção é organizada do seguinte modo: na Seção 2.2 são apresentados os principais modelos aplicados, os resultados obtidos na análise do conjunto de pontos medidos e a proposta de um novo modelo; na Seção 2.3 o modelo proposto é comparado com alguns modelos da literatura; e na Seção 2.4 os modelos aplicados e o proposto são utilizados para estimativa da cobertura do sinal transmitido.

2.2 Estudo dos principais modelos aplicados à previsão da propagação

A partir da comparação de alguns modelos de propagação referenciados na literatura, ou seja, modelo de espaço livre, de Hata, da Recomendação ITU-R P.1546 e log-distância com uma ou múltiplas regressões, foi possível ajustar o modelo log-distância para o comportamento da propagação observada no conjunto de medidas obtidos em campo, e então propor o modelo empírico de propagação log-distância misto que poderá ser utilizado na determinação da cobertura dos sinais de TV Digital na região de Curitiba. Para cada modelo de propagação aplicado foi levantada a curva da potência de recepção estimada, mostrada em um gráfico do nível de recepção, em dBm (*dB em relação a 1 mW*), em função da distância de separação entre o transmissor e o receptor, em km, permitindo uma análise qualitativa da propagação. Para a comparação quantitativa, foram calculados o erro médio da diferença entre o nível medido e o nível estimado pelo modelo para cada ponto, e o desvio padrão σ do conjunto de pontos medidos em relação aos valores estimados para cada modelo. O resumo dos resultados obtidos na aplicação dos diversos modelos de propagação está apresentado na Tabela 1.

2.2.1 Modelo espaço livre

O modelo de espaço livre possui uma estimativa bastante otimista quando comparado com os valores medidos, como pode ser visto na Figura 1, com

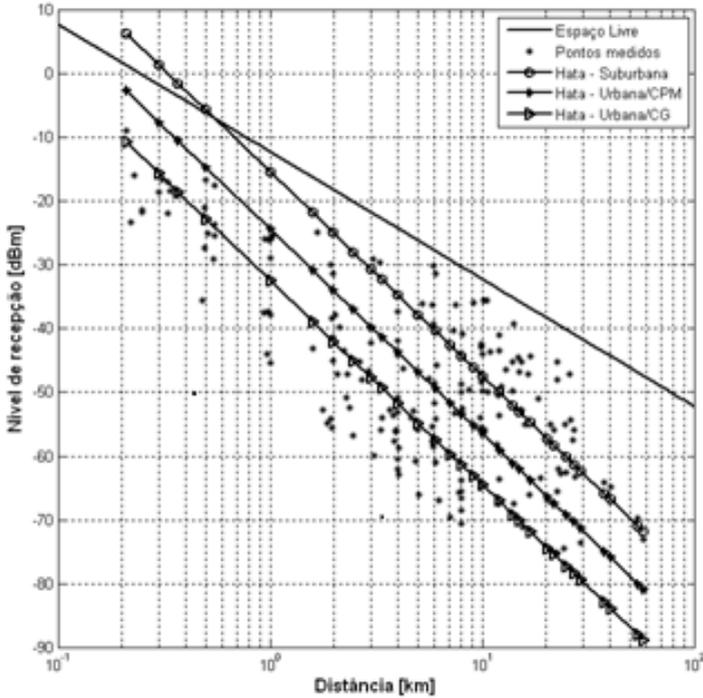


Figura 1. Modelos de espaço livre e de Hata.

uma dispersão alta de 22,9dB conforme dados mostrados na Tabela 1. Isto se deve ao fato que este método corresponde à atenuação devido à irradiação da potência do transmissor no espaço livre, considerando somente a distância até o receptor, em uma condição com linha de visada entre o transmissor e o receptor, que não é o caso estudado, onde há uma aglomeração urbana com prédios e construções, associado à topografia da região não ser plana.

2.2.2 Modelo de Hata

O Modelo de Hata (Hata, 1980) foi implementado para as condições de áreas urbanas em cidades grandes (CG) e em cidades pequenas e médias (CPM), e também para áreas suburbanas, considerando a altura da antena transmissora de 90m e a altura da antena receptora de 8m. Pela análise da Figura 1, a curva da estimativa para áreas urbanas em cidades grandes (Urbana/CG) adapta-se melhor para os pontos medidos até 8km e a estimativa para áreas suburbanas adapta-se melhor para os pontos medidos

além de 9km. A estimativa para áreas urbanas em cidades pequenas e médias (Urbana/CPM) é a que apresenta menor erro médio e variância, conforme dados da Tabela 1, quando comparados os três modelos.

2.2.3 Modelo ITU-R P.1546

O modelo da recomendação ITU-R P.1546 foi implementado conforme definido no Anexo II da Resolução da Anatel n° 398 (Anatel, 2005). Como a estimativa é em função da altura efetiva da antena do transmissor, foram obtidos os dados topográficos da região no sistema SIGAnatel (Anatel, 2008b). Foram feitas duas estimativas utilizando este método. A primeira considerando uma altura efetiva média de 122m que corresponde à média da altura efetiva de todos os pontos medidos considerando os dados topográficos da região e a regra de cálculo definida pela Rec. ITU-R P.1546. A segunda, considerando a altura efetiva calculada para cada um dos pontos e aplicando as correções para percursos curtos urbanos e suburbanos, aplicáveis para distâncias inferiores a 15km, relacionados à obstrução dos prédios, considerando uma altura dos arredores de 30m, e a correção para a altura da antena receptora de 8m em relação à referência de 10m, para uma altura de cobertura do solo de 30m. Verifica-se, pelos dados da Tabela 1, que os resultados da utilização da Rec. ITU-R P.1546 são similares ao resultado da estimativa para áreas urbanas em cidades pequenas e médias (Urbana/CPM).

2.2.4 Modelo log-distância de um segmento

Este trabalho então se concentrou na busca de um modelo de propagação ponto-área, a partir de um modelo log-distância, que é obtido por meio da regressão pelo MMSE (*método do erro médio quadrático mínimo*) (Rappaport, 1996). Esta regressão permite obter uma curva que melhor se ajusta ao conjunto de pontos medidos, onde a soma do quadrado dos erros entre os valores medidos e os valores estimados é minimizada. A atenuação por propagação PL na distância d , que é a distância do transmissor ao ponto medido, é dada pela Equação (1):

$$PL(d) = PL(d_0) + 10n \log_{10}(d/d_0) \quad (1)$$

onde $PL(d_0)$ é a atenuação por propagação na distância d_0 , que é uma distância de referência próxima ao transmissor e n o expoente da atenuação por propagação. A atenuação $PL(d_0)$ foi obtida pela média dos níveis de recepção medidos na distância de $d_0=0,21$ km. Os resultados do modelo log-distância de um segmento são apresentados na Figura 2 e Tabela 1. O valor de n encontrado é de 2,53.

2.2.5 Modelo log-distância de dois segmentos

Na busca de diminuir o desvio padrão, foi aplicado o modelo log-distância em dois segmentos (do inglês *dual slope*), ou modelo de regressão

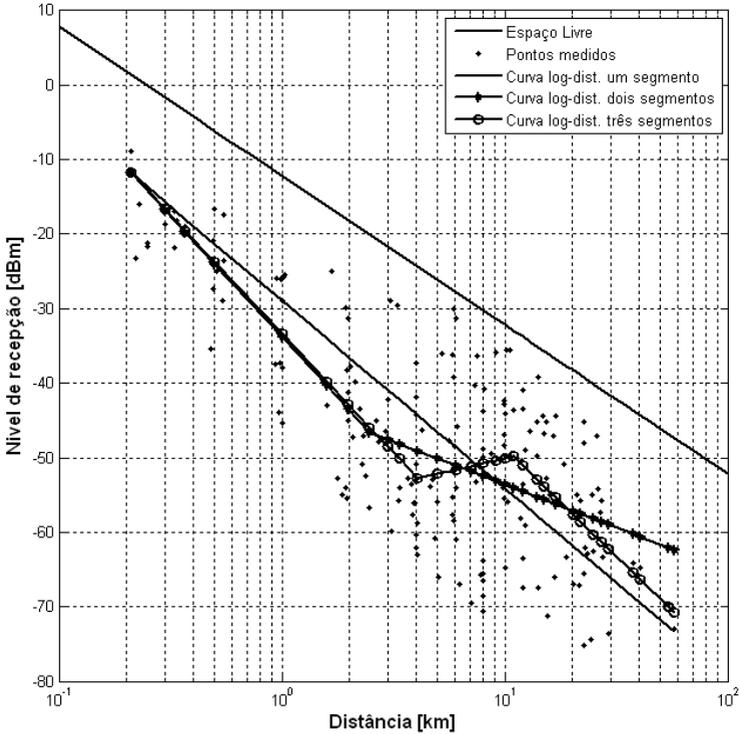


Figura 2. Modelos log-distância de um, dois e três segmentos.

dupla (Feuerstein et al., 1994), que é uma variante do modelo log-distância, quando aplicado ao modelo de reflexão por dois raios, que considera duas regiões de comportamento de propagação, separadas por um ponto de joelho (do inglês *breakpoint*). Este ponto é encontrado pela distância de obstrução da primeira zona de Fresnel, ou seja, a distância do transmissor onde a borda da primeira zona de Fresnel está no nível do solo. A curva foi encontrada pelo processo iterativo de arbitrar o ponto de joelho e calcular pelo método de MMSE cada segmento, sendo adotada a curva que apresentou o menor desvio padrão. Pelos dados da Tabela 1, o desvio padrão é menor que no modelo log-distância de um segmento, mas observa-se que a tendência da curva do modelo de dois segmentos, mostrada na Figura 2, é divergente da tendência das medidas do conjunto de pontos.

2.2.6 Modelo log-distância de três segmentos

Como a tendência da curva do modelo de dois segmentos não é convergente, então, foi utilizado um modelo com regressão tripla, com um segundo ponto

de joelho, mais afastado da antena. Semelhante ao modelo log-distância de dois segmentos, a curva do modelo de três segmentos foi encontrada pelo processo iterativo de arbitrar cada um dos dois pontos de joelho e calcular, pelo método de MMSE, cada segmento, sendo adotada a curva que apresentou o menor desvio padrão. Pelos dados da Tabela 1, o desvio padrão é menor que no modelo log-distância de dois segmentos, mas pela Figura 2, observa-se que o segundo segmento deste modelo é negativo, não possuindo uma explicação física. Portanto este é um modelo puramente teórico.

2.2.7 Modelo log-distância misto – proposto

O modelo proposto, chamado de Modelo Log-distância Misto, considera três segmentos, onde o primeiro e segundo segmentos possuem os índices de propagação n_1 e n_2 iguais ao do modelo log-distância de dois segmentos e o terceiro índice de propagação n_3 igual ao do modelo log-distância de três segmentos. O primeiro joelho j_1 é igual ao joelho do modelo log-distância de dois segmentos e o segundo joelho j_2 , é o ponto de encontro entre as curvas dos modelos log-distância de dois e três segmentos. Pelo processo de regressão MMSE obteve-se a curva apresentada na Figura 3.

Pelos resultados da Tabela 1, o desvio padrão deste modelo não é o menor, mas o modelo log-distância de três segmentos é meramente matemático, sem explicação física. A equação da atenuação por propagação do modelo log-distância misto PLLDM é dada por:

$$P_{LLDM}(d) = \begin{cases} 32,5 \log_{10}(d/0,21) + 87,29 & (0,21 \leq d \leq 2,5 \text{ km}) \\ 11,5 \log_{10}(d/2,5) + 122,25 & (2,5 \leq d \leq 19,0 \text{ km}) \\ 29,0 \log_{10}(d/19) + 132,38 & (d \geq 19,0 \text{ km}) \end{cases}$$

onde d é a distância entre a antena transmissora e o receptor, em quilômetros. Como o levantamento foi feito no canal 41, está aproximadamente no meio da faixa dos canais de UHF de televisão que estão entre 14 e 59.

O resumo dos resultados obtidos na aplicação dos diversos modelos de propagação é apresentado na Tabela 1.

2.3 Comparação entre os modelos

Primeiramente foi feita a comparação do modelo misto com os valores médios da potência recebida nos pontos. A média da potência recebida foi calculada para vários trechos de distâncias de modo a permitir a visualização do comportamento médio da propagação na região.

Pela análise da Figura 4, nota-se que a curva do modelo misto acompanha a média do sinal de recepção medido nos pontos, indicando ser um modelo adequado para as condições da região. O gráfico mostra que há um ponto de cruzamento da curva log-distância de um segmento

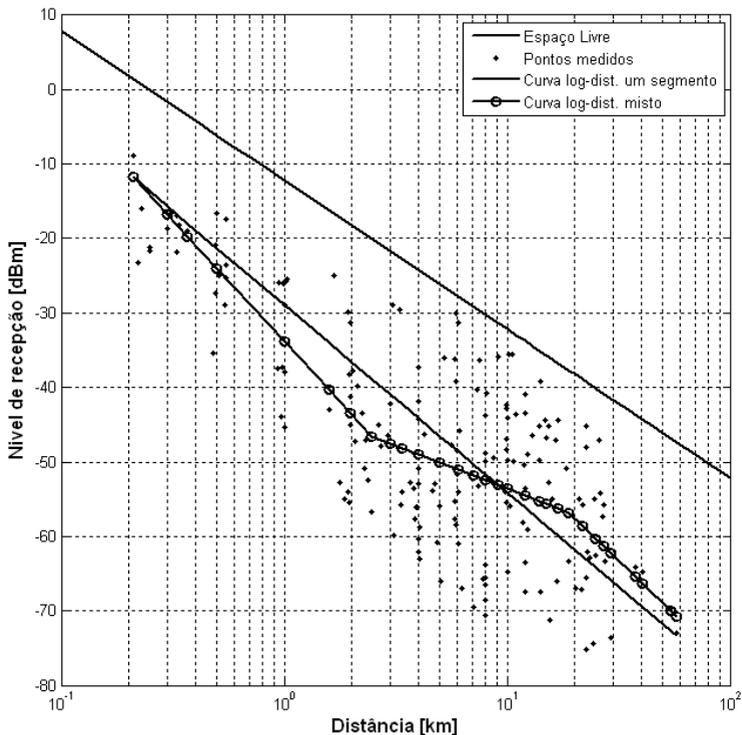


Figura 3. Modelo proposto log-distância misto.

Tabela 1. Resumo dos resultados obtidos.

Método de propagação	Erro médio [dB]	σ [dB]
Espaço livre	-20,8	22,9
Hata - área urbana cidades grandes (CG)	7,2	14,0
Hata - área urbana cidades pequenas e médias (CPM)	-0,8	12,0
Hata - área suburbana	-9,9	15,6
ITU-R P.1546 com altura efetiva de 122m	-5,4	14,1
ITU-R P.1546 com altura efetiva pontual e com correção	1,4	12,0
Log-distância de um segmento	-1,4	10,3
Log-distância de dois segmentos	-0,2	9,2
Log-distância de três segmentos	-0,2	8,9
Log-distância misto	0,2	9,1

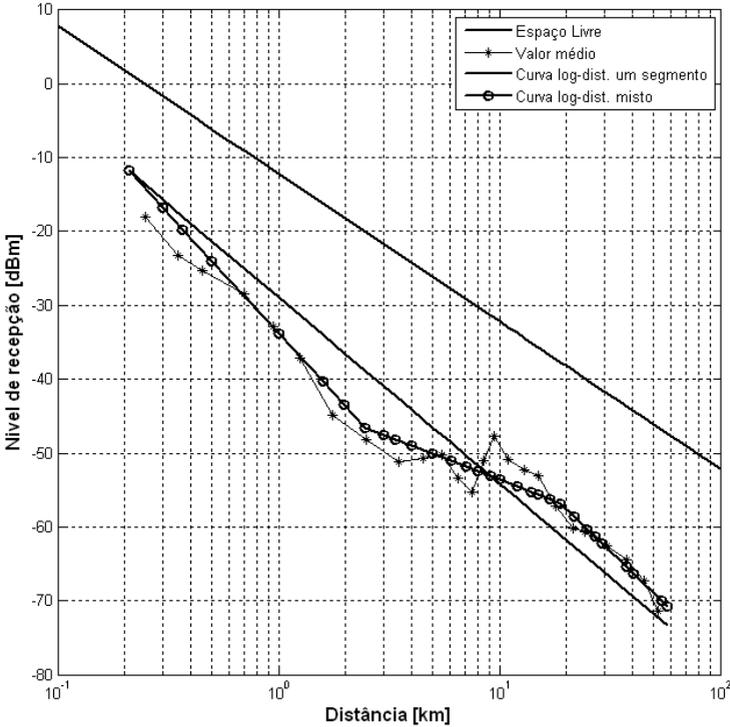


Figura 4. Comparação do nível médio e dos modelos log-distância misto e de um segmento.

e do modelo misto na distância de 9,1km, onde há uma mudança de comportamento do nível médio de recepção medido nos pontos, e consequentemente na estimativa do modelo misto. Antes deste ponto, desde o primeiro ponto medido, o nível médio está abaixo da curva log-distância e, após o ponto até o último ponto medido, o nível médio está acima da curva log-distância.

Então, pode-se considerar que existem dois comportamentos da propagação na região. Provavelmente isto se deve ao fato da região urbana densamente habitada da cidade de Curitiba estar espalhada até uma distância média, aproximada, de 10km.

Na segunda análise, foi realizada uma comparação do modelo misto com os modelos de Hata para área urbana em cidades grandes e em cidades pequenas e médias, além da área suburbana, mostrada na Figura 5.

Pode-se observar que o primeiro segmento do modelo misto adapta-se à estimativa do modelo de Hata para área urbana em cidades grandes, até a distância de aproximadamente 3km, onde o índice de propagação

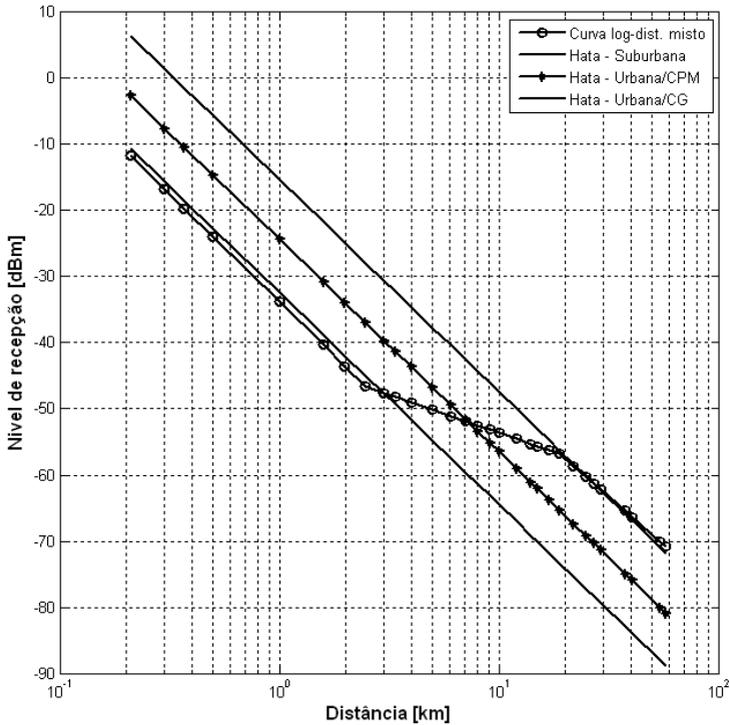


Figura 5. Comparação dos modelos log-distância Misto e Hata.

do modelo misto $n_1=3,25$, bem próximo do valor de $n=3,21$ do modelo de Hata. O terceiro segmento do modelo misto adapta-se ao modelo de Hata para área suburbana, após a distância de 19,0km, onde o índice de propagação do modelo misto $n_3=2,90$ é aproximadamente igual ao $n=3,21$ do modelo de Hata.

Assim foi realizada uma comparação dos valores estimados pelo modelo com os valores medidos, para trechos de distâncias, obtendo-se os resultados mostrados na Tabela 2.

Pela análise dos dados da Tabela 2, pode-se ver que para o trecho de 0,21 a 4,1km, os métodos log-distância misto e de Hata para área urbana em cidades grandes têm resultado aproximados. Para o trecho de 19,0 a 57,7km, após o segundo ponto de joelho do modelo misto, este modelo possui um resultado levemente pior no erro médio que o modelo de Hata, mas o mesmo resultado do desvio padrão. Entretanto o modelo misto pode não ser o melhor modelo em cada trecho, mas certamente é o melhor considerando todo o conjunto de medidas, portanto é indicado para ser utilizado na estimativa da cobertura dos sinais da TV Digital na região.

Tabela 2. Resultados obtidos para trechos.

Método de propagação	Dist. [km]	Erro médio [dB]	σ [dB]
Log-distância misto - segmento 1	0,21 a 4,1	-1,1	7,9
Hata - área urbana cidades grandes	0,21 a 4,1	-1,1	7,8
Rec. ITU-R P.1546 - altura efetiva	19,0 a 57,7	6,0	10,6
Log-distância com um segmento	19,0 a 57,7	3,1	8,6
Hata - área suburbana	19,0 a 57,7	-0,3	8,1
Log-distância misto - segmento 3	19,0 a 57,7	-0,7	8,1

Adicionalmente, na tentativa de interpretar o comportamento da propagação em relação a alguns pontos de joelho analisados, obteve-se, no sistema SIGAnatel ([Anatel, 2008b](#)), a distribuição da aglomeração urbana da região de Curitiba, centralizado no local da antena transmissora, que é apresentada na Figura 6. As áreas escuras representam regiões com concentração urbana, as linhas escuras indicam o limite dos municípios e os círculos indicam as distâncias de 4,1km, 9,1km e 19,0km, referentes à, respectivamente, distância do ponto de joelho j_1 do modelo log-distância de três segmentos, distância onde a curva do modelo misto cruza com a curva do modelo log-distância de um segmento, e distância do ponto de joelho j_2 do modelo misto.

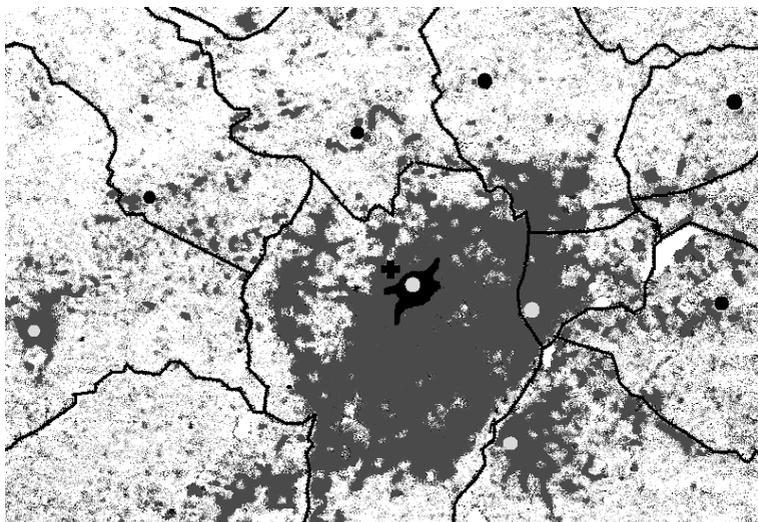


Figura 6. Aglomeração urbana na região de Curitiba, adaptado do sistema SIGAnatel.

Pode-se notar que o centro da cidade está localizado no círculo de 4,1km da antena transmissora, onde há a maior concentração de edifícios, e, conseqüentemente, há a maior atenuação do sinal de radiofrequência devido à obstrução. O ponto de joelho j_1 do modelo misto em 2,5km representa, provavelmente, o ponto onde a densa aglomeração urbana estaria terminando, modificando o comportamento da propagação. A partir de 19,0km, ponto de joelho j_2 do modelo misto, observa-se uma região suburbana da cidade, com baixa aglomeração urbana. A distância de 9,1km, onde há uma mudança do comportamento da atenuação da propagação, está em uma região intermediária entre a grande aglomeração urbana e regiões suburbanas, indicando, provavelmente, uma mudança no perfil das construções a partir desta distância.

2.4 Análise da cobertura utilizando os modelos estudados

Considerando o nível mínimo de -77,0dBm, que é a sensibilidade de recepção definida pela norma ABNT NBR 15604 (ABNT, 2007), e por meio de uma extrapolação das curvas obtidas em cada modelo estudado, foi calculada a distância limite entre o transmissor e o receptor que atende esta condição, apresentada na Tabela 3.

Tabela 3. Resultados obtidos para cobertura.

Modelo de propagação	Distância [km]
Log-distância misto	95
Hata - área suburbana	81
Log-distância de um segmento	81
Hata - área urbana cidades pequenas e médias	43
ITU-R P.1546 - altura efetiva de 122m	40
Hata - área urbana - cidades grandes	24
Log-distância de dois segmentos	>150

O modelo log-distância misto que foi considerado o indicado para a avaliação da cobertura, e por se tratar de um modelo ponto-área, obteve uma distância média de 95km. Os métodos de Hata para área suburbana e log-distância de um segmento obtiveram estimativas de cobertura médias idênticas com uma distância de 81km, sendo uma coincidência porque estes modelos possuem curvas e pontos de origem diferentes. Como era esperado, o modelo da curva log-distância de dois segmentos não convergiu o resultado até a distância máxima de 150km. Os modelos de Hata para área urbana em cidades grandes e em cidades pequenas e médias, e o método da ITU-R P1546 obtiveram estimativas pessimistas na distância média de cobertura entre 24 a 43km.

3. Medição de Cobertura

As entidades de teledifusão, privadas ou públicas, têm um interesse permanente em conhecer a região geográfica servida por suas emissões. Modelos computacionais podem fornecer informações detalhadas sobre a área coberta, ao implementar variados métodos de predição baseados em normas nacionais (Anatel, 2005, 1991) e internacionais (Radiocommunication Study Groups ITU, 2003, 2007b).

Enquanto tais métodos computacionais são indispensáveis para orientar os investimentos de implantação da estação, e para que os órgãos governamentais reguladores definam frequências, potências e localização das emissoras, *medir* o sinal após sua implementação é atividade imprescindível para validar as predições (Radiocommunication Study Groups ITU, 2005), e para identificar precisamente áreas de recepção difícil, a fim de solucionar as deficiências (Pullen, 1999). Analisar as especificidades inerentes, e validar uma metodologia prática para conhecer a cobertura real de um sistema de TV digital é o objetivo desta Seção.

Modelagem computacional e medição direta não são abordagens que competem entre si; são antes complementares. Até porque há métodos computacionais de predição importantes que também empregam modelos empíricos, baseados em um número elevado de medições, como o de Okumura-Hata e o da Recomendação 1546-1 da ITU (Radiocommunication Study Groups ITU, 2003). Assim, um objetivo secundário deste trabalho é apreciar a coerência entre os valores preditos e os medidos. Também será observada a influência das implantações sobre o terreno próximas ao local de recepção (*clutter*), e os valores medidos dos retardos de multipercurso.

3.1 Especificidades inerentes à medição de cobertura

O efeito combinado das múltiplas reflexões e do pequeno comprimento de onda acarreta uma variabilidade muito acentuada no valor da potência recebida no ambiente urbano. Mudanças de alguns centímetros na posição da antena receptora impõem variações de dezenas de dB ao sinal captado. Devido a essa grande variabilidade do mensurando, sua medição direta reveste-se de peculiaridades que influenciam a escolha dos métodos. Algumas dessas particularidades são abordadas nesta Seção, porquanto embasam a metodologia de medição apresentada na Seção subsequente.

3.1.1 Variabilidade do mensurando

A intensidade do sinal recebido varia tanto no tempo quanto em localização (Radiocommunication Study Groups ITU, 2007a). O sinal pode ser formado por componentes diretas, difratadas, refletidas ou refratadas. É de se esperar, portanto, variações significativas no sinal recebido.

Uma percepção qualitativa da variabilidade no valor de potência recebida pode ser formada a partir da observação da Figura 7. Ela

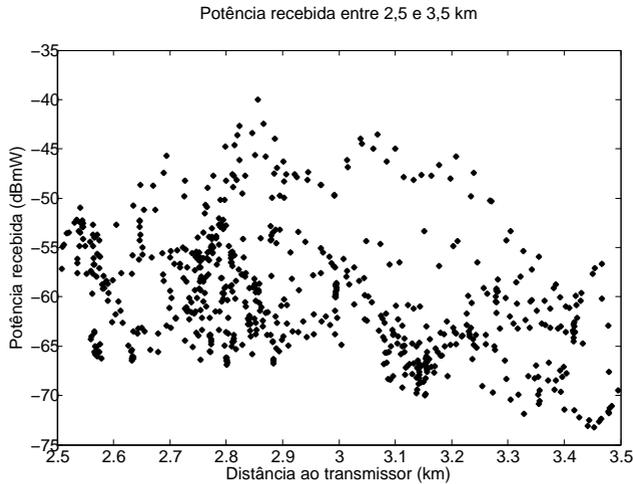


Figura 7. Variabilidade nos valores medidos de potência recebida em função da distância entre emissor e receptor.

apresenta valores medidos de potência recebida em função da separação, em quilômetros, entre transmissor e receptor. Trata-se de um sinal de TV digital, com faixa de 632 a 638MHz, medido em Curitiba. É parte do conjunto de dados que embasou o estudo apresentado nesta Seção.

3.1.2 Os métodos de uso corrente

Durante a fase de dimensionamento da estação de teledifusão, decisões precisam ser tomadas relativamente a alturas de torres, localização, potência dos transmissores e antenas de emissão. Nesse momento, modelos computacionais são utilizados para avaliar a cobertura que pode ser esperada a partir de algumas combinações, ditadas mais por considerações econômicas do que técnicas. É quando *softwares* sofisticados de predição de cobertura, como (EDX Wireless LLC, 2008), embasam o processo decisório do empresário de teledifusão. Por seu turno, os órgãos reguladores do uso do espectro eletromagnético precisam de informações quantitativas sobre as intensidades de campo que podem ser produzidas por determinado emissor, tanto para o enquadramento normativo de sua cobertura quanto para avaliar possíveis interferências. Normas internacionais, como (Radiocommunication Study Groups ITU, 2003), e nacionais, como (Anatel, 1991), foram criadas para orientar o processo.

A Tabela 4 reúne um conjunto abrangente de modelos de propagação recomendados pela ITU.

Cálculos de propagação podem ser enquadrados em duas grandes categorias: métodos de previsão *ponto-a-ponto* e métodos de previsão *ponto-área*. Os primeiros são adequados quando a localização do receptor é perfeitamente conhecida, e não muda. Os segundos são concebidos para o caso de *broadcasting*, onde a localização exata de cada receptor não é conhecida *a priori*, e também para outras situações em que os receptores são móveis.

Nos métodos ponto-a-ponto, são em geral conhecidas a topografia e as implantações sobre o terreno ao longo do percurso onde ocorre a propagação. Isso permite um cálculo mais refinado, que pode considerar em detalhe as eventuais obstruções do primeiro elipsóide de Fresnel, difrações em torno de obstáculos e efeitos da refração troposférica.

3.1.3 Variabilidade das localidades

Nos métodos ponto-área, a região de cobertura presumível é dividida em quadrículas de dimensões variáveis: 200 x 200m (Anatel, 1991), 1 x 1km (Pullen, 1999), quadrículas com lados entre 100 e 200m para VHF e UHF (Radiocommunication Study Groups ITU, 2003), quadrículas com lados de 500m (Radiocommunication Study Groups ITU, 2007b) para citar alguns exemplos. A cada quadrícula é associado um valor de campo elétrico *mediano*, isto é, um valor que será superado por 50% das medições feitas dentro da quadrícula à qual está associado o valor de intensidade de campo elétrico, durante 50% do tempo. Esse valor é geralmente denominado $E(50, 50)$ (campo elétrico excedido em 50% das localidades e em 50% do tempo), e é um dos principais valores fornecidos pelos métodos ponto-área. Se o tamanho da quadrícula for relativamente pequeno, o percurso entre transmissor e receptor tende a não se alterar significativamente com a movimentação do receptor no interior da quadrícula. O fator de maior influência sobre as variações de campo elétrico medido, conhecida como *variabilidade das localidades*, será a alteração da cobertura local do terreno (prédios, árvores, pessoas, veículos, etc.) nas proximidades da antena receptora. Essa variabilidade tem uma distribuição estatística aproximadamente lognormal (Radiocommunication Study Groups ITU, 2003), o que significa dizer que os valores de campo, expressos em dB, têm uma distribuição estatística Gaussiana (densidade de probabilidade normal).

Assim, para uma quadrícula onde o campo elétrico mediano é $E(50, 50)$, o valor de campo elétrico $E(q)$ que será excedido em $q\%$ das medições é avaliado pela Equação (2) (Radiocommunication Study Groups ITU, 2003):

$$E(q) = E(50, 50) + Q_i(q/100)\sigma_L(f) \quad (2)$$

Tabela 4. Modelos de propagação da ITU.

Radio Propagation Models	Applied Phenomena and Parameters	Applied Frequency Ranges	Corresponding ITU Recommendations
Gaseous Attenuation Model	Oxygen and water vapor absorption	> 1.0 GHz	ITU-R P.676,836 [12,17]
Scintillation and Multipath Model for Troposphere	Scintillation and multipath fading	> 1.0 GHz	ITU-R P.618 [18]
Attenuation Frequency Scaling Model	Rain attenuation	7 - 50 GHz	ITU-R P.618 [18]
Cumulative Distribution of Fade Duration Model	Fading duration	> 400 MHz	[1,10]
Diffusely Scattering Wave Model	Satellite to mobile with Doppler shift	> 100 MHz	[1,10]
Empirical Roadside Shadowing Model	Canopy shadow for land mobile	> 100 MHz	[1,10]
Maritime Link: Fading Due to Sea Reflection	Rough sea surface for marine mobile	0.8 - 8.0 GHz	ITU-R P.680 [10]
Altshuler Fog Model	Fog attenuation	5 - 100 GHz	ITU-R P.840 [19]
ITU Cloud Model	Cloud attenuation	5 - 100 GHz	ITU-R P.840 [19]
Radio Noise Model	Galactic and atmospheric noise	100 MHz - 100 GHz	ITU-R P.372 [1,10]
ITU Rain Model	Rain attenuation, rain rate distribution	1 - 400 GHz	ITU-R P.618, 837, and 838 [18,20,21]
Global Rain Model	Rain climate zones	1 - 400 GHz	[1,2]
Wet Antenna Surface Model	Wet factor	20 GHz	[1,11]
Ice and Rain Depolarization Model	Co-polar attenuation (CPA) and XPD	4 - 35 GHz	ITU-R P.618,530 [18,22]
Site Diversity Model	Site separation during rain	10 - 30 GHz	ITU-R P. 618 [18]
Line of Sight with Multipath Model	Space to Earth or Earth to Earth	> 2.0 GHz	ITU-R P. 530, and 619 [22,23]
Terrain Diffraction Model	Knife-edge and rounded surface	30 MHz - 10 GHz	ITU-R P.526, and 530 [16,22]
Transhorizon Propagation Model	Anomalous models: Ducting and scattering	> 0.7 GHz	ITU-R P.452 [15]

onde:

- $Q_i(x)$: distribuição cumulativa normal complementar inversa, em função da probabilidade, detalhada nas equações (4) e (5) a seguir;
- σ_L : desvio-padrão da distribuição Gaussiana das médias locais na área sob estudo.

O valor do desvio-padrão para os sistemas digitais com largura de banda *inferior* a 1MHz e para sistemas analógicos são dados, em dB, pela Equação (3) ([Radiocommunication Study Groups ITU, 2003](#)):

$$\sigma_L(f) = K + 1,6 \log_{10}(f) \quad (3)$$

com:

- $K=2,1$ para sistemas móveis em localidades urbanas;
- $K=3,8$ para sistemas móveis em localidades suburbanas ou entre morros suaves;
- $K=5,1$ para sistemas de radiodifusão analógicos;
- f =frequência em MHz.

Para sistemas digitais com largura de banda *igual* ou *superior* a 1MHz, deve-se usar um desvio-padrão de 5,5dB em todas as frequências, conforme a Recomendação ITU-R P.1546-1 ([Radiocommunication Study Groups ITU, 2003](#)).

Uma aproximação para $Q_i(x)$, válida para $0,01 \leq x \leq 0,99$, é fornecida em (4) se $x \leq 0,5$, e em (5) se $x \geq 0,5$ ([Radiocommunication Study Groups ITU, 2003](#)):

$$Q_i(x) = T(x) - \xi(x) \quad (x \leq 0,5) \quad (4)$$

$$Q_i(x) = -[T(1-x) - \xi(1-x)] \quad (x \geq 0,5) \quad (5)$$

onde:

$$T(x) = \sqrt{[-2 \ln(x)]} \quad (6)$$

$$\xi(x) = \frac{(C_2 T(x) + C_1) T(x) + C_0}{[(D_3 T(x) + D_2) T(x) + D_1] T(x) + 1} \quad (7)$$

com:

- $C_0=2,515517$;
- $C_1=0,802853$;
- $C_2=0,010328$;
- $D_1=1,432788$;
- $D_2=0,189269$;

- $D_3=0,001308$;

Considerado isoladamente, o alcance de uma emissão de RF (*radiofrequência*) seria infinito. O que efetivamente impõe um limite para o alcance de um sinal desejado são outros sinais, *indesejados*, que com ele competem. São formados pelo ruído radioelétrico (natural ou produzido pelo homem), pelo ruído gerado internamente no receptor, e por outras emissões, distantes e no mesmo canal, ou em canais adjacentes. Para prever coberturas, os métodos ponto-área estimam a distância máxima na qual o sinal desejado ainda tem uma amplitude suficiente para concorrer favoravelmente com os sinais indesejados presentes na mesma antena receptora. A avaliação do efeito combinado que vários sinais interferentes têm em um mesmo ponto de recepção pode ser feita por um método devido a Schwartz e Yeh, *apud Pullen (1999)*.

3.1.4 Variabilidade com o tempo

Percursos mais longos acentuam a influência do índice de refração troposférico sobre a amplitude do sinal recebido na faixa de 30 a 3000MHz, usualmente modeladas por meio do raio terrestre equivalente. Em percentagens baixas de tempo, o sinal emitido pode produzir intensidades de campo a longas distâncias que não seriam esperadas, se consideração fosse feita apenas às curvas que fornecem $E(50,50)$. Portanto, para estudar interferências que determinado sinal pode produzir, interessa saber que intensidades ele pode atingir, mesmo que em percentagens menos significativas do tempo de observação. Com esse objetivo, as normas costumam fornecer dados para $E(50,10)$ (*campo elétrico excedido em 50% das localidades e em 10% do tempo*) e $E(50,1)$ (*campo elétrico excedido em 50% das localidades e em 1% do tempo*), valores de campo elétrico superados apenas em 10 e em 1% respectivamente do tempo de observação para uma determinada separação entre emissor e receptor.

Além de outras fontes, esses valores de $E(50,10)$ e $E(50,1)$ podem ser encontrados em ([Radiocommunication Study Groups ITU, 2003](#)) e ([Radiocommunication Study Groups ITU, 2007b](#)). Essas informações embasam a decisão dos órgãos administradores do espectro eletromagnético quanto à separação mínima que deve existir entre emissoras, a fim de que qualquer delas não interfira sobre os receptores localizados na área de cobertura de outra emissora, dentro do chamado contorno protegido.

Para distâncias de 50km, e para percursos terrestres em UHF, a Recomendação ITU-R P.1406 informa um desvio-padrão para variabilidade temporal de apenas 2dB ([Radiocommunication Study Groups ITU, 2007a](#)). Para as distâncias ainda menores encontradas nas medições em Curitiba, as variações temporais podem ser desconsideradas, particularmente quando confrontadas com os altos valores de desvios-padrão devidos às variações de localidade.

3.2 Por que medir

O emprego de métodos computacionais de previsão diminui, em princípio, a necessidade de extensos levantamentos experimentais. Entretanto, o trabalho de campo ainda é necessário para validar as previsões ([Radiocommunication Study Groups ITU, 2005, 2007a](#)). Também é importante para localizar áreas de recepção difícil, a fim de que recebam a atenção necessária ([Pullen, 1999](#); [Plumb et al., 1999](#); [Radiocommunication Study Groups ITU, 2007a](#)). Visando a melhorar a acurácia dos prognósticos de cobertura, softwares como o EDX ([EDX Wireless LLC, 2008](#)) disponibilizam vários modelos de cálculo, à escolha do utilizador. São modelos sofisticados, que incorporam refinamentos, e buscam realizar previsões que se aproximem o mais possível das medições experimentais. À medida em que se comparam os valores preditos pelos modelos com aqueles efetivamente medidos, compreende-se melhor por que há tantas opções de modelos incorporadas ao software. Para que as previsões dos modelos se tornassem mais acuradas, seria preciso incluir no modelo objetos não transparentes às irradiações, como prédios, torres, e outras implantações. Seria necessário obter suas dimensões e localizações precisas, e também características eletromagnéticas de seus materiais constitutivos. Esse procedimento é ainda mais demorado e caro do que a medição direta do sinal efetivamente recebido.

Desde que os custos e prazos permitam, medir é sempre mais preciso que calcular. A acurácia de uma medição, ainda que efetuada com recursos limitados, supera a das previsões dos modelos, como se verá no decorrer deste trabalho.

3.2.1 Medir é viável

As técnicas para medir intensidade de campo recebido estão já estão bem estabelecidas ([Terman & Pettit, 1952](#); [Rohde & Schwarz, 1985](#)). Para o caso de *broadcasting*, importa que o número de medições seja expressivo, para que a área a ser investigada abranja quantidade significativa de receptores ([Radiocommunication Study Groups ITU, 2007a](#)). Significa dizer que o método de medição precisa coletar um número elevado de medidas sem se tornar excessivamente custoso ou demorado. Portanto, automatizar as medições, e tomá-las de um veículo em movimento, é importante para tornar o método viável. Essa foi a abordagem escolhida para os ensaios em Curitiba. Ela difere do método tradicional, com antena receptora montada a 10m do solo, em um mastro telescópico ([Gonzalez Castellanos et al., 2010](#); [Vasco, 2009](#); [Radiocommunication Study Groups ITU, 2003](#)), e permite coletar um volume maior de dados em menos tempo, conforme analisado na Seção 3.6.3. Medir com a antena receptora a 10m do solo aproxima a condição de antena externa em casa térrea, medir com antena a 2m do

solo, sobre um veículo em movimento, aproxima a condição de recepção móvel.

Para ilustrar as ordens de grandeza, um trecho denominado “primeiro afastamento na avenida Canadá” tem seu trajeto detalhado no mapa da Figura 8. O trecho é sobre a avenida Canadá, entre a rua Holanda e o Parque General Iberê de Mattos, no bairro Bacacheri, em Curitiba-PR. Nele foram tomadas 60 medições de potência recebida e de resposta impulsiva. O trecho mede 1914m, e foi percorrido em 4,932 minutos. A velocidade média no percurso foi de 23,28km/h, o que permitiu medir 31,348 pontos por quilômetro. Isto significa uma medida a cada 32m de percurso, a uma taxa média de 12,16 medidas por minuto.

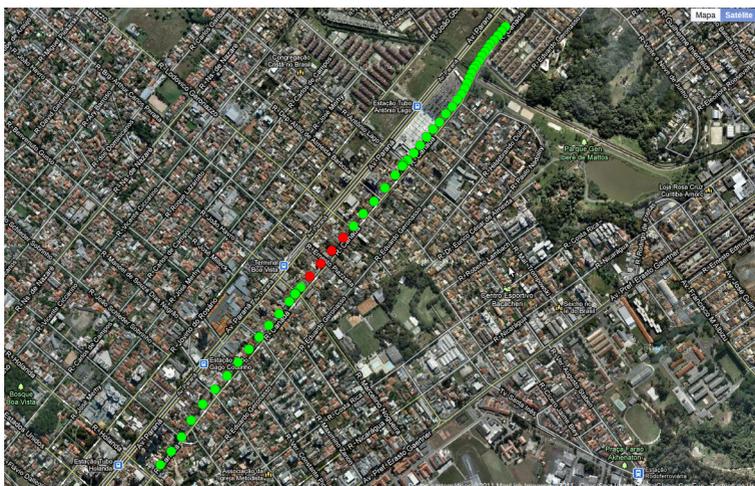


Figura 8. Trajeto do primeiro afastamento na avenida Canadá. Círculos vermelhos assinalam potência recebida abaixo do limiar.

3.2.2 Repetibilidade e reprodutibilidade das medições

Uma medição só é útil se puder ser posteriormente repetida e/ou reproduzida com coerência. Com o objetivo de verificar essa característica, algum tipo de suavização dos dados medidos precisa ser aplicada, para que a grande variabilidade da potência recebida guarde coerência com a região de medição. A Figura 9 apresenta os valores medidos sem qualquer tratamento. A dificuldade de repetir ou de reproduzir a mesma medida em um dado ponto é facilmente percebida.

Visando a suavizar as variações rápidas com a distância na intensidade do sinal recebido, calculou-se o valor médio de várias medidas adjacentes, para associar esse valor médio ao local de medição. Após alguns ensaios,

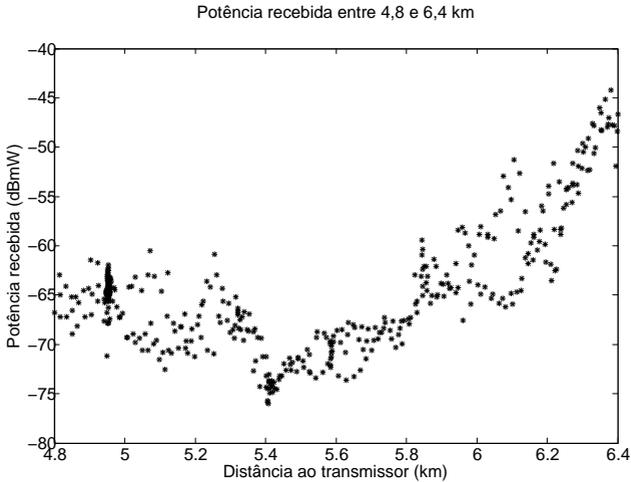


Figura 9. Medidas sem qualquer tratamento, tomadas entre 4,8 e 6,4 km.

elegeu-se o valor 5 para o tamanho da janela deslizante. Isto significa que, a cada valor de distância, associa-se a média de 5 medidas em 5 pontos adjacentes a essa distância. Esse valor de 5 medições representa uma quadrícula com cerca de 150m de lado, valor compatível com a Recomendação ITU-R P.1546-1 ([Radiocommunication Study Groups ITU, 2003](#)). Assim, cada quadrícula passa a ter uma “distância” (média de 5 distâncias) e uma “potência recebida” (média de 5 valores de potência recebida). O resultado obtido encontra-se na Figura 10.

O caminho percorrido é o mesmo já representado na Figura 8. Nela, os círculos verdes assinalam os locais onde a potência recebida encontra-se acima do limiar (o limiar adotado neste trabalho foi uma potência recebida no canal de -72dBmW , porque correspondeu ao nível em que falhou, em campo, a recepção em HD no equipamento empregado nas medições). Os círculos vermelhos indicam os locais onde não é esperado ocorrer recepção nas condições em que a medição foi realizada, com antena receptora posicionada na pista de rodagem, e altura de 2m em relação ao solo.

Não basta o resultado ser suave, é preciso que ele possa ser *repetido* e *reproduzido* com coerência. Para verificar a repetibilidade, compararam-se as potências médias recebidas em duas passagens sucessivas pela mesma avenida, com um intervalo de cerca de 40 minutos entre um e outro percurso. A repetibilidade da medição pode ser avaliada por meio da Figura 11. É útil ressaltar que, embora as medições em ambos os percursos

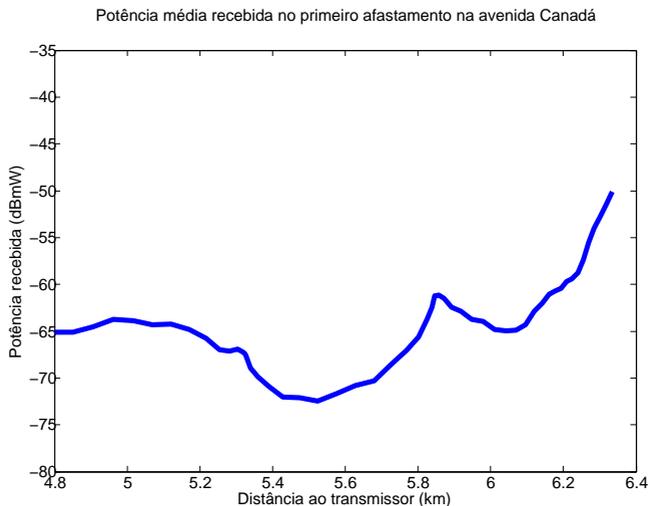


Figura 10. Média de 5 medidas adjacentes de potência recebida na avenida Canadá, primeira passagem.

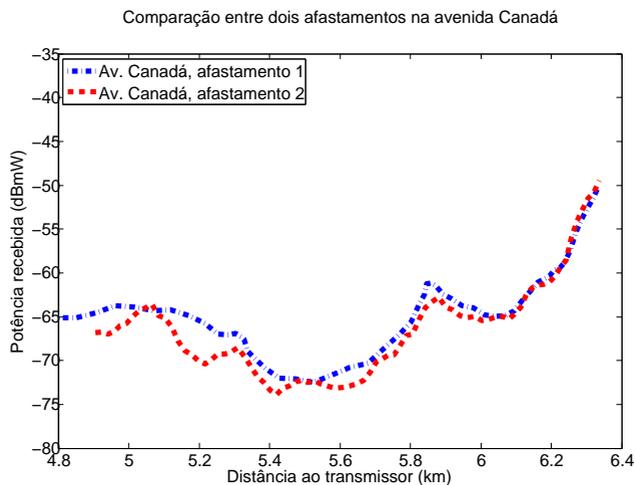


Figura 11. Comparação entre os sinais medidos na avenida Canadá, em duas passagens sucessivas, separadas por 40 minutos, para verificar a repetibilidade do método.

ocorreram a não muito mais de meia hora entre elas, o veículo esteve sujeito às restrições do trânsito. Significa que não trafegou na mesma velocidade, nem que percorreu necessariamente a mesma pista de rodagem. Em face do pequeno comprimento de onda utilizado no ensaio, da ordem de 0,5m, essas modificações podem alterar mesmo os valores médios.

Ainda assim, é perceptível a alta correlação entre os valores medidos nas duas diferentes passagens pelo mesmo trajeto, o que atesta a boa repetibilidade do método de medição utilizado.

O caminho percorrido no segundo afastamento na avenida Canadá está na Figura 12.



Figura 12. Trajeto do segundo afastamento na avenida Canadá. Círculos vermelhos assinalam potência recebida abaixo do limiar.

Com o objetivo de avaliar a reprodutibilidade e a coerência dos valores medidos, foi percorrido um percurso próximo ao já mencionado, representado na Figura 14. Ali estão representadas as potências recebidas em duas avenidas paralelas e adjacentes. O trajeto da primeira aproximação na avenida Paraná está representado na Figura 13. Como anteriormente, os círculos verdes assinalam os locais onde a potência recebida encontra-se acima do limiar. Os círculos vermelhos indicam os locais onde não ocorreu recepção nas condições em que foi realizada a medição.

Note-se que, apesar da pequena variação em distância ao transmissor, os níveis médios do sinal têm variação da ordem de 5dB entre uma avenida e a outra. Estas diferenças provavelmente são causadas pelos prédios

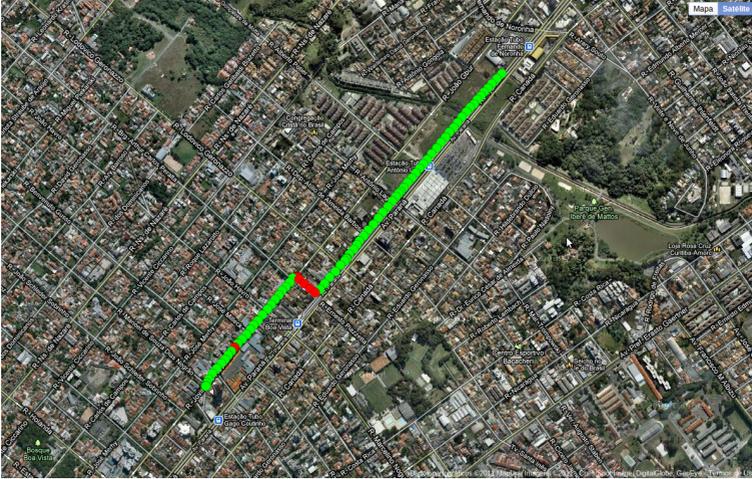


Figura 13. Trajeto da primeira aproximação na avenida Paraná. Círculos vermelhos assinalam potência recebida abaixo do limiar.

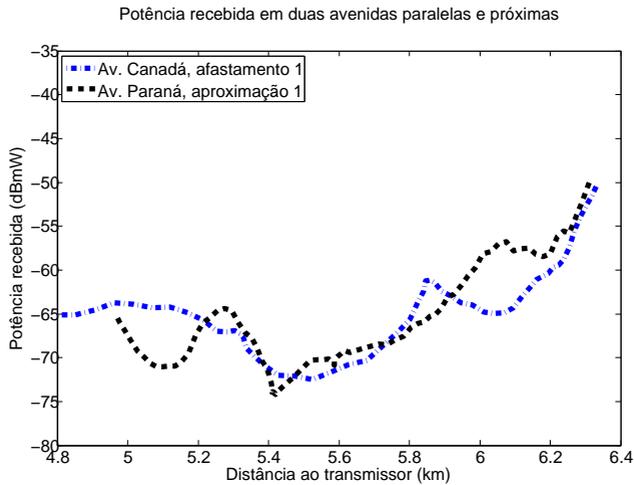


Figura 14. Comparação entre os sinais medidos nas avenidas Canadá e Paraná.

às suas margens. Como os prédios não são considerados nos modelos computacionais, é impossível que qualquer das simulações predissesse as diferenças encontradas na Figura 14, o que reforça a importância de medir.

3.2.3 Escolha do trajeto

O percurso escolhido para o ensaio não teve a ambição de constituir-se em um levantamento de cobertura. Antes, a seleção do trajeto baseou-se nas seguintes considerações:

- Facilidade de deslocamento;
- Alta densidade de telespectadores;
- Facilidade em passar mais de uma vez pela mesma via (para verificar a repetibilidade);
- Coerência (esperava-se sinal mais fraco *atrás* dos prédios);
- Multipercurso (havia dúvidas sobre o quanto ele seria prejudicial).

As especificidades abordadas nesta Seção subsidiaram as definições do método de medição, assunto da Seção a seguir.

3.3 Método

A escolha do método de medir foi determinada tanto pelas especificidades abordadas na Seção anterior quanto pelos equipamentos e recursos efetivamente disponíveis.

A abrangência das medições, e os materiais, instrumentos e configurações estão a seguir descritos.

3.3.1 Escopo dos ensaios de campo

O escopo dos ensaios em campo está limitado pelo equipamento disponível. O objetivo geral é verificar a cobertura do sinal de TV digital sob condições que não são facilmente reprodutíveis em laboratório ou em simulações (Weck, 1999). Nos ensaios que embasaram o presente trabalho, foram registrados, em *cada* medição automatizada, os seguintes valores:

- Número de ordem do ponto de medição;
- Hora em que foi registrada a medição, fornecida por GPS;
- Latitude e longitude do ponto de medição, avaliadas por GPS;
- Potência recebida em toda a faixa do canal em dBmW (*dB em relação a 1 miliwatt*);
- MER (*taxa de erro de modulação*); e
- Resposta impulsiva, constituída pelas dez versões mais intensas do sinal que trafegaram por percursos múltiplos, caracterizadas por suas amplitudes relativas e retardos, também relativos, sempre à versão multipercurso mais intensa registrada nesta medição. Uma amostra típica está na Figura 16.

3.3.2 Critérios para avaliar a recepção

Quando há multipercurso, a taxa de erro de bit pode ser mais alta do que o esperado. Significa dizer que, para atingir-se uma determinada taxa de erro de bit, é necessária uma C/N mais elevada em situação de multipercurso (Pullen, 1999). Na prática, há ainda o ruído impulsivo e as diversas interferências, que tendem a exigir uma C/N ainda maior. O ruído radioelétrico emitido pelos veículos automotores pode degradar sistemas em UHF em 10dB ou mais (Miyoshi & Sanches, 2002).

Embora haja diversas considerações bem elaboradas sobre o limiar a considerar (QEF – *Quasi Error Free* (Weck, 1999; Radiocommunication Study Groups ITU, 2005), CRR – (*Correct Reception Rates*) (Akamine et al., 2010)), o limiar adotado neste trabalho foi uma potência recebida no canal de -72dBmW, porque correspondeu ao nível em que falhou, em campo, a recepção em HD no equipamento empregado nas medições.

3.3.3 Instrumentos e configurações

As medições foram tomadas a bordo de um veículo especialmente destinado a esse fim, equipado com GPS (*Global Positioning System*), inversor de DC (*corrente contínua*) para 60Hz, acomodações apropriadas para operadores, instrumentos e antenas.

3.3.4 Sinal de teste

O sinal de TV digital utilizado nas medições foi aquele irradiado publicamente pela estação que serviu de fonte de sinal para os ensaios, no canal 41 de UHF, com as características a seguir (Anatel, 2006):

- Faixa de frequências de 632 a 638MHz;
- Polarização horizontal;
- Potência do transmissor 8kW;
- Ganho da antena transmissora 10,7dBd (*dB em relação a um dipolo de meia-onda*);
- EIRP 80,42dBmW; e
- Altura do centro de fase sobre o terreno 90m.

O sinal emitido pela estação que serviu de fonte de sinal de teste para os ensaios, no canal 41 de UHF, tem um espectro praticamente plano, composto por 5617 subportadoras COFDM (modo 3). O intervalo de guarda utilizado foi de $63\mu\text{s}$ (1/16 da duração do símbolo de $1008\mu\text{s}$). A Tabela 5 resume a configuração do sinal usado nas medições.

Tabela 5. Parâmetros do sinal usado nas medições.

Camadas	A	B
Largura de faixa (MHz)	0,432	6
Intervalo de guarda	1/16	1/16
Quantidade de segmentos	1	12
Tipo de modulação	QPSK	64-QAM
Código convolucional	2/3	3/4
Entrelaçamento no tempo (ms)	400	200
Taxa de bits (Mbits/s)	0,44	17,842

3.3.5 Configuração dos instrumentos

A via principal de recepção inicia-se por uma antena do tipo monopólo vertical, com comprimento próximo a $\lambda/4$. Mesmo que o sinal seja transmitido com polarização horizontal, as múltiplas reflexões que sofre em ambiente urbano acabam por criar componente vertical na polarização. De acordo com as considerações feitas em [Borba \(2011\)](#), uma perda de 7dB foi adotada como consequência desse descasamento de polarização.

O monopólo vertical foi escolhido como antena receptora por ser essencialmente não diretivo no plano horizontal, de modo a tornar a recepção quase independente da direção do percurso. Além da conveniência mecânica de sua base magnética de fixação à capota do veículo, o monopólo vertical é uma antena receptora típica para a recepção móvel.

O sinal recebido pelo monopólo vertical foi separado em dois caminhos por um divisor de 3,15dB. Um deles alimentou um receptor de TV digital, mantido ligado durante o percurso para avaliar subjetivamente a recepção. O segundo caminho foi conectado a um analisador de espectro do fabricante Rohde&Schwarz, modelo ETL. Esse instrumento mede a potência total na faixa de frequências do canal, BER (*taxa de erro de bit*), MER (*taxa de erro de modulação*), e consegue distinguir os sinais que percorreram caminhos múltiplos, informando suas amplitudes e retardos relativos ([Rohde & Schwarz, 2011](#)). A Tabela 6 resume algumas características do instrumento analisador de espectro que foi empregado nas medições.

Um diagrama simplificado do material usado na recepção encontra-se na Figura 15, e uma tela típica do analisador de espectro alimentado por um sinal real de TV digital ISDB-T_B pode ser vista na Figura 16. Todo o material da Figura 15 esteve montado a bordo de uma *van* especificamente destinada à atividade, equipada com inversor para 60Hz.

3.3.6 Procedimentos de medição

O procedimento de medição consistiu em sintonizar o receptor de TV no canal 41, recebido pela antena monopólo vertical $\lambda/4$, cuja base de fixação magnética permaneceu aderida à capota da *van*, a 2m do solo. O sinal que provém dessa antena, através de um cabo coaxial com impedância

Tabela 6. Especificações do analisador de TV *Rohde und Schwarz* ETL.

Warm-up time	30 minutes	Three hours storage at ambient temperature
Setting range of signal level	-80dBm to +20dBm in steps of 0,1dB	
Level measurement uncertainty	95% confidence level, +20°C to +30°C, S/N>16dB, 0dB to -50dB from reference level, 50MHz < f < 3GHz	< 1,0dB
Attenuator uncertainty	< 0,3dB	
Uncertainty of signal level setting	< 0,1dB (nom.)	

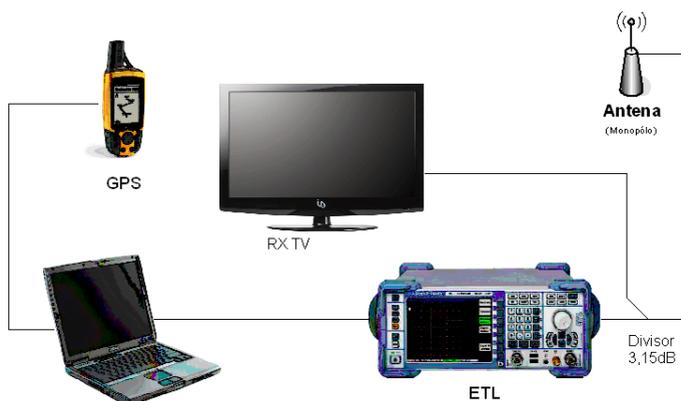


Figura 15. Diagrama em blocos do sistema de recepção.

característica de 75Ω , é dividido em dois caminhos, por um divisor de 3,15dB; uma via alimenta o receptor de TV, e a outra se dirige à entrada do analisador de espectro. Com isto, a potência medida à entrada do analisador é praticamente a mesma que alimenta o receptor de TV. Quando a imagem ou som no receptor de TV deixarem de ser satisfatórios, todas as características do sinal estarão disponíveis ao analisador. Os parâmetros do sinal extraídos pelo analisador ETL são enviados ao *laptop*, onde são

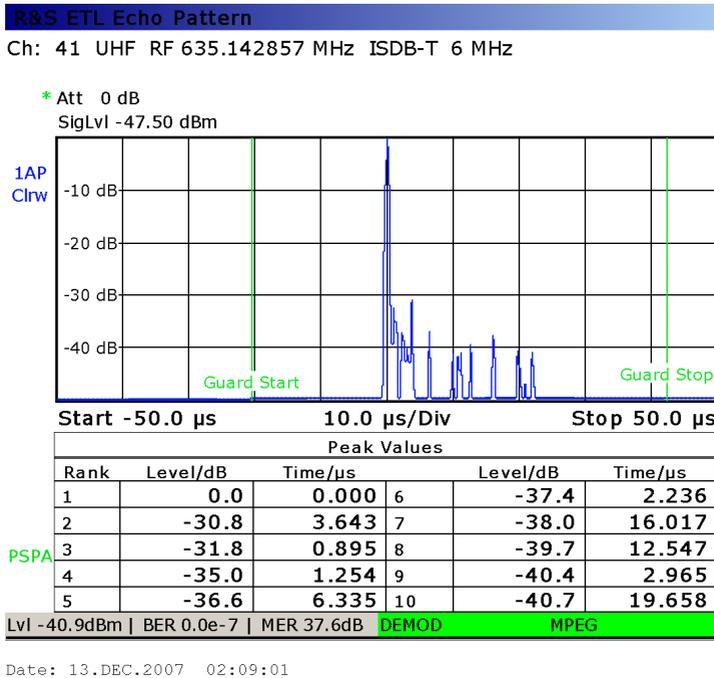


Figura 16. Tela típica do analisador ETL.

gravados em uma planilha, para análise posterior. São gravadas todas as informações anteriormente detalhadas na Seção 3.3.1.

Após o tempo de *warm-up* especificado na Tabela 6, a *van* segue o itinerário previamente definido, e medidas são automática e continuamente realizadas e armazenadas na planilha, a um ritmo determinado pela capacidade do ETL. A velocidade de deslocamento do veículo não deve exceder um limite máximo, sob pena de não serem armazenadas medições relativas a multipercurso. Com a configuração adotada, esse limite de velocidade foi experimentalmente determinado entre 10 e 20km/h. A escolha da velocidade é um compromisso entre o prazo disponível para cobrir determinada área e o número de amostras desejado por quadrícula de medição. O tamanho efetivo da quadrícula dependerá de quantas medições serão agrupadas em uma janela deslizante para produzir as estatísticas de tendência central e de dispersão. Essa escolha pode ser feita *a posteriori*, já que os dados medidos estarão armazenados. Neste trabalho, foi adotado o tamanho 5 para a janela deslizante; significa que 5 amostras serão analisadas em conjunto para produzir as estatísticas, o que corresponde

a uma quadrícula com cerca de 150m de lado, conforme já mencionado na Seção 3.2.2.

3.4 Incertezas na medição

A avaliação da incerteza a respeito dos valores medidos de potência recebida foi feita de acordo com os procedimentos metrológicos usuais (Albertazzi & Sousa, 2008). Como cada amostra tem 5 medições, a incerteza com respeito à média amostral é avaliada com a distribuição t (de Student) (Spiegel, 1977), de acordo com a Equação (8).

$$I = I_i + t_{2\sigma} s \quad (8)$$

onde:

- I (*incerteza metrológica*): diferença entre o valor medido e o valor verdadeiro, que não é superada em 95,45% dos casos;
- I_i (*incerteza metrológica do instrumento de medição*): diferença entre o valor medido pelo instrumento e o valor verdadeiro, que não é superada em 95,45% dos casos;
- $t_{2\sigma}$ (*valor crítico t de Student (Gosset)*): valor crítico da variável t que delimita uma área de 95,45% em torno da média. $t_{2\sigma}=2,869$ para 4 graus de liberdade;
- s (*desvio-padrão da amostra*): desvio-padrão da amostra com 5 medições;

Esses valores de incerteza, calculados com $t=2,869$, foram empregados nas Figuras da Seção 3.5.

Convém notar também que a altura de 2m adotada nos ensaios, além de ser inferior à altura representativa do *clutter*, nos termos da Recomendação ITU-R P.1546-1 (mínimo de 10m) (Radiocommunication Study Groups ITU, 2003), também expõe a antena receptora a um sinal com mais multipercurso e a um nível maior de ruído impulsivo, em decorrência da maior proximidade com os sistemas de ignição dos veículos no tráfego. É de se esperar portanto que o limiar de recepção satisfatória ocorra em uma potência recebida maior do que o nível medido em laboratório, de -77dBm. Em campo, o limiar de recepção satisfatória em HD foi determinado experimentalmente em -72dBm à entrada de antena do receptor de TV digital, com as considerações já apresentadas em 3.3.2.

Colhidos e armazenados os dados, conforme os procedimentos descritos, posteriormente podem ser analisados, assunto abordado na Seção que segue.

3.5 Análise dos dados medidos

Uma análise dos dados coletados segundo o método descrito na Seção anterior chama atenção sobre alguns aspectos. O primeiro deles é a nítida influência que prédios altos têm sobre a propagação urbana.

3.5.1 Intensidade do sinal medido

Medições foram tomadas em duas avenidas paralelas e próximas na cidade de Curitiba. Ambas as vias estão situadas quase que perpendicularmente ao percurso de propagação. Entre elas, existem muitos prédios elevados, construídos em consonância com o Plano Diretor da cidade. Portanto, seria de se esperar que o sinal recebido na sombra eletromagnética dos prédios (na avenida Silva Jardim) fosse mais fraco que o sinal recebido antes das construções (na avenida Visconde de Guarapuava). Essa expectativa claramente se confirma, como se pode perceber pelas Figuras 17 e 18 a seguir. A altura de cada barra representa a fração do total de medidas no trecho que cai dentro da faixa de potência centrada no valor em dBmW sobre o qual está localizada a barra. Em ambas as figuras encontram-se destacados os casos em que houve falha na recepção em HD.

Ao comparar os dois trechos, percebe-se que há uma diferença significativa, da ordem de 20dB, entre os valores médios de potência

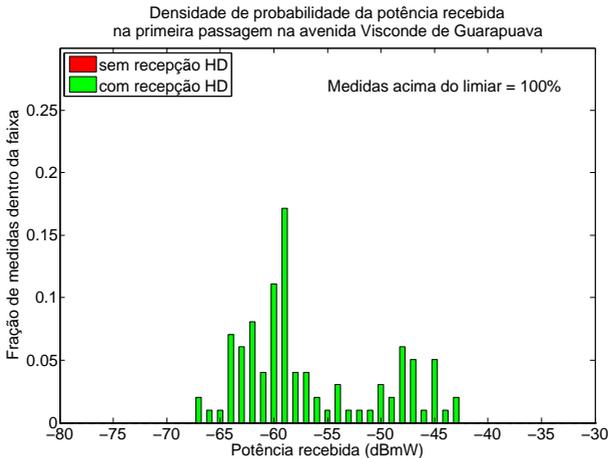


Figura 17. Fração do número total de medidas no trecho que se encontra dentro da faixa de potência especificada pela abcissa de cada barra na avenida Guarapuava, primeira passagem. Barras verdes significam recepção HD; barras vermelhas assinalam falhas na recepção.

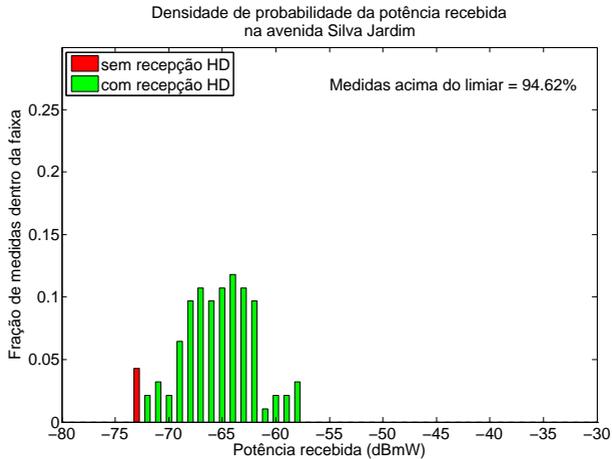


Figura 18. Fração do número total de medidas no trecho que se encontra dentro da faixa de potência especificada pela abscissa de cada barra na avenida Silva Jardim. Barras verdes significam recepção HD; barras vermelhas assinalam falhas na recepção.

recebida nas duas avenidas. É razoável atribuí-la à presença dos prédios altos, implantados entre as duas vias. Como as ruas são aproximadamente perpendiculares ao percurso de propagação, a linha de construções altas funciona como um muro, mais efetivo ainda quando as medições são tomadas no nível do solo. Quando se comparam os histogramas das avenidas Guarapuava (Figura 17) e Silva Jardim (Figura 18), percebe-se que os valores de potência recebida nessa última agrupam-se em torno de valores mais baixos, o que parece confirmar a atenuação causada pelas construções altas permitidas pelo zoneamento da cidade de Curitiba entre as duas vias.

A curva $P(90, 50)$ da Figura 19 é um avaliação teórica, calculada com o modelo Gaussiano da Recomendação ITU-R P.1546-1. A partir da observação apenas dessa curva, é possível concluir que, para assegurar que em 90% das localidades o sinal recebido fique acima do limiar naquele trecho, seria necessário aumentar a potência de transmissão entre 3 e 4 dB. Todavia, os valores efetivamente medidos e representados no histograma da Figura 20 mostram que houve cobertura HD em mais de 93% das localidades, fato que pode ser interpretado como mais uma indicação da importância de medir.

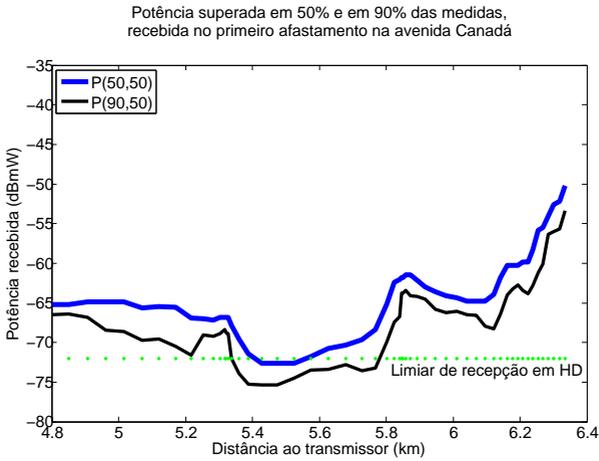


Figura 19. Valor mediano $P(50, 50)$ (linha espessa) e valor superado em 90% das localidades $P(90, 50)$ (linha fina) de potência recebida por quadrícula na avenida Canadá, primeira passagem.

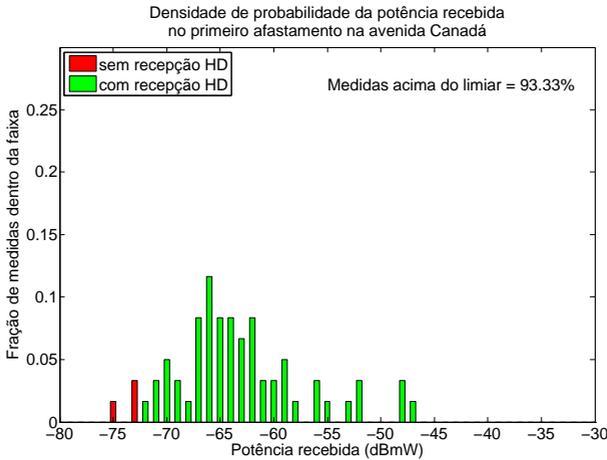


Figura 20. Fração do número total de medidas no trecho que se encontra dentro da faixa de potência especificada pela abcissa de cada barra na avenida Canadá, primeiro afastamento. Barras verdes significam recepção HD; barras vermelhas assinalam falhas na recepção.

3.5.2 Multipercurso

A qualidade da recepção de TV analógica nos grandes centros sempre foi prejudicada pelas imagens secundárias criadas a partir de réplicas do sinal atrasadas no tempo, consequência da multiplicidade de percursos possíveis para a propagação eletromagnética em ambiente urbano. Não é sem fundamento, portanto, considerar os efeitos do multipercurso sobre a TV digital. Para isso, foi aproveitada a capacidade do analisador de espectro empregado nos ensaios de registrar a resposta impulsiva do canal.

Uma ideia intuitiva de como o canal responde a um impulso pode ser formada a partir da Figura 16, da Seção 3.3.5. Cada versão do sinal, propagada por diferentes percursos, fica representada por sua amplitude e por seu retardo, ambos medidos em relação a uma referência. A referência adotada nos ensaios foi a réplica *mais intensa* do sinal, que não é necessariamente a primeira a chegar. Consequentemente poderá haver retardos com tempos negativos, se chegarem versões do sinal *antes* da réplica mais intensa adotada como referência. Esses casos são, às vezes, denominados de *pré-ecos*. É importante notar que a existência de um único retardo negativo é suficiente para confirmar que o sinal recebido em visada direta não foi o mais intenso.

A amplitude de cada réplica, por sua vez, é medida em relação à referência, arbitrariamente considerada como 0dB. Todas as demais terão amplitude negativa em dB, porque são medidas em relação à versão mais intensa que chegou. O conjunto de potências e retardos relativos de N réplicas do sinal propagadas por múltiplos caminhos pode ser denominado perfil de ecos (Rappaport, 2002).

O equipamento utilizado nas medições tem capacidade para registrar até dez réplicas ($N=10$). Portanto, cada uma das 1580 medidas tomadas durante o ensaio tem associada a si vinte valores relativos a ecos. Dez deles são as intensidades, em dB, com relação ao mais intenso. Os outros dez valores são os retardos, em microssegundos, relativamente também ao eco mais intenso.

Com o fim de resumir a informação contida nesses 31.600 números medidos, utilizou-se a formulação da Recomendação ITU-R P.1407 *Multipath Propagation and Parameterization of its Characteristics* (Radiocommunication Study Groups ITU, 1999), os assim chamados “retardo médio” T_D e ‘espalhamento *rms* dos retardos’ S , apresentada a seguir:

$$T_D = \frac{\sum_{i=1}^N P_i t_i}{\sum_{i=1}^N P_i} \quad (9)$$

$$S = \sqrt{\frac{\sum_{i=1}^N P_i t_i^2}{\sum_{i=1}^N P_i} - \frac{(\sum_{i=1}^N P_i t_i)^2}{(\sum_{i=1}^N P_i)^2}} \quad (10)$$

onde:

- N = número de versões multipercurso consideradas;
- P_i = potência da i -ésima versão multipercurso;
- t_i = retardo da i -ésima versão multipercurso.

Observe-se que o retardo médio T_D é uma média ponderada dos N retardos levados em consideração (em nosso experimento foram dez). O peso de cada retardo na formação da média é a sua potência relativa. É uma medida de tendência central, e pode ter valor negativo no caso de os sinais propagados por multipercurso terem intensidade significativa com respeito ao propagado em visada direta. O espalhamento *rms* S é uma medida de dispersão, e avalia o quanto os retardos estão espalhados no tempo.

Durante o ensaio, foi percebido que o analisador de espectro só conseguia registrar o perfil de ecos quando a velocidade de deslocamento ficasse abaixo de um valor máximo. Esse valor é inferior a 20km/h, e foi avaliado em aproximadamente 10km/h com base no velocímetro do veículo. Devido às condições de tráfego, nem sempre foi possível manter velocidade de deslocamento compatível com o registro. Por isso, algumas partes dos trechos não têm registro do perfil de ecos, de modo que as curvas apresentam interrupções.

Para comparar o comportamento em diferentes regiões, a Figura 21 mostra um exemplo de valores de T_D , calculados com a Equação (9) e de S , calculados com a Equação (10) para um trecho de região suburbana, enquanto a Figura 22 apresenta a informação correspondente a um trecho de região urbana.

É possível perceber que os valores de retardo são maiores em região suburbana do que nos trechos urbanos, ao contrário do que se poderia esperar. A causa provável é que, nas regiões urbana e urbana densa, a presença de obstáculos elevados junto ao receptor impede que sinais refletidos de objetos mais distantes atinjam a antena de recepção no nível do solo. Por seu turno, a maior visibilidade permitida pelos prédios mais baixos dos trechos suburbanos permite receber ecos de objetos refletores mais distantes.

Outra observação que pode ser feita na Figura 21, é que os retardos tendem a diminuir com o aumento da distância ao transmissor, porque a cota do terreno cresce com a distância nesse trecho, e permite maior visibilidade. Com ela, aumenta a intensidade do sinal em visada direta, o que acarreta uma diminuição do peso dos sinais retardados no cômputo do tempo médio, causando a progressiva diminuição de T_D , de acordo com a Equação (9).

Os dados experimentais mostram também que os valores de retardo médio encontrados na prática são pequenos (da ordem de $6\mu\text{s}$) quando

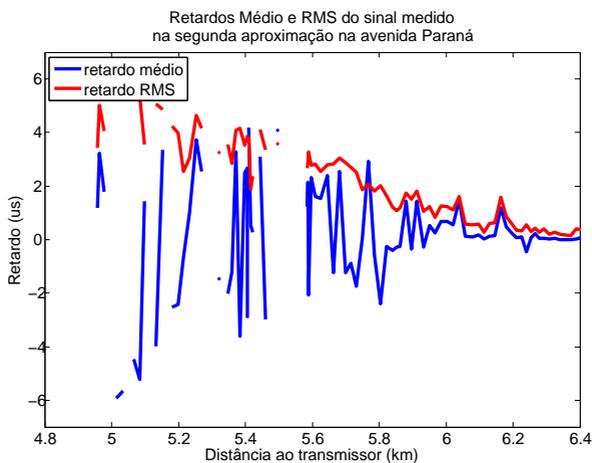


Figura 21. Valores de retardo médio (linha azul) e *rms* (linha vermelha), calculados segundo a Recomendação ITU-R P.1407 para a segunda aproximação na avenida Paraná, em região suburbana.

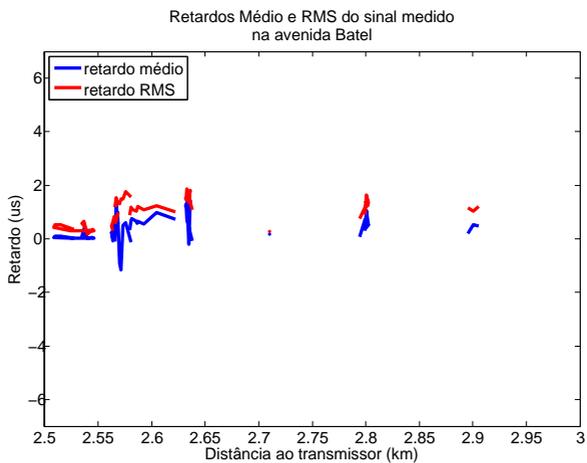


Figura 22. Valores de retardo médio (linha azul) e *rms* (linha vermelha), calculados segundo a Recomendação ITU-R P.1407 para o trecho das avenidas Benjamin Lins e Dr. Pedrosa, em região urbana.

comparados com o intervalo de guarda de $63\mu\text{s}$ utilizado na transmissão. É um indício de que o multipercorso perturba pouco o sinal de TV digital nas regiões investigadas, pelo menos enquanto não for utilizada uma rede de frequência única (SFN).

3.6 Resultados

Coletados e analisados os dados medidos, é oportuno reorganizar as informações de forma mais concisa, propósito perseguido nas seções subsequentes.

3.6.1 Comparação com modelos

Conforme já exposto na Seção 3.1.3, os cálculos de propagação procuram prever o valor *mediano* do sinal. Para verificar a concordância entre os valores calculado e medido, foi elaborada a Figura 23. Ali estão comparados os valores médios das medições por quadrícula com o que prevê a Recomendação ITU-R P.1546-1 ([Radiocommunication Study Groups ITU, 2003](#)). Essa Recomendação foi adotada com *status* de norma pela Anatel ([Anatel, 1991](#)).

Se for considerado que o valor mediano é aquele que é superado por metade das medições, pode-se dizer que a previsão é boa. Sua falha é não apresentar os detalhes da variação do sinal.

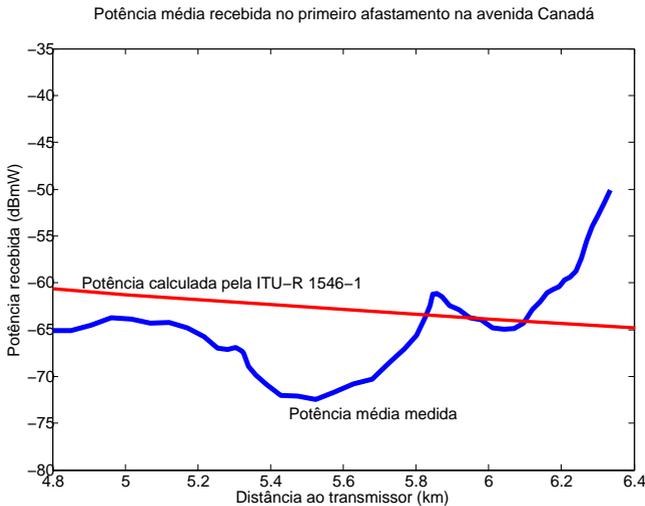


Figura 23. Valor médio medido (linha espessa) e calculado segundo a Recomendação ITU-R P.1546-1 (linha fina) da potência recebida por quadrícula na avenida Canadá, primeiro afastamento.

Após as medições de campo, a equipe técnica da estação que serviu de fonte de sinal de teste para os ensaios utilizou um *software* comercial, denominado EDX (EDX Wireless LLC, 2008), que dispõe de uma variedade de modelos, à escolha do utilizador. Essa ferramenta gerou previsões de potência recebida segundo 32 modelos diferentes, para comparar com as medições. O programa produziu curvas de potência recebida *versus* distância e, por meio de cálculos internos, avaliou parâmetros de erro das previsões. A denominação dos modelos e a acurácia de suas previsões estão na Tabela 7.

Um modelo que apresentou boa concordância com os valores medidos na avenida Canadá foi o popular modelo de Okumura-Hata, comparado com o experimento na Figura 24.

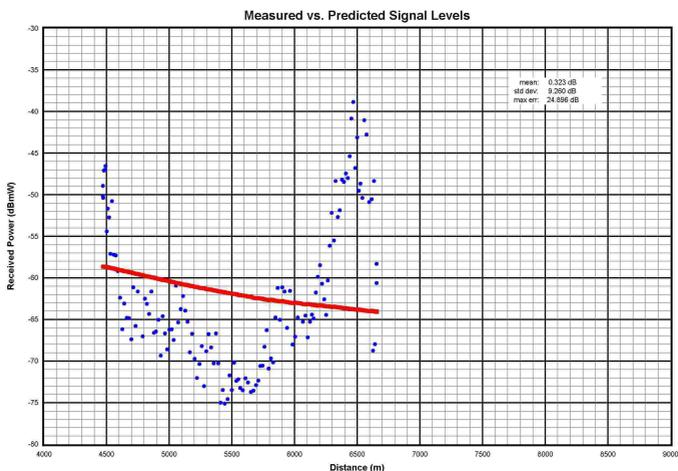


Figura 24. Comparação entre os valores medidos (círculos azuis) e a previsão do *software* EDX (quadrados vermelhos) segundo o modelo Okumura-Hata-suburban (sem difração), na avenida Canadá. Erro médio=0,323dB; desvio-padrão do erro=9,26dB; erro máximo=24,89dB.

Como no caso anterior, a previsão de valores medianos é boa, mas não informa quanto às variações. Com o objetivo de melhorar esse detalhamento, o *software* sofisticou o modelo de Hata, a ele acrescentando as perdas por difração em obstáculos múltiplos, calculadas com base no relevo com o método de Epstein-Peterson. Por essa técnica, cada obstáculo é tratado individualmente, com o obstáculo anterior considerado como transmissor, e o obstáculo seguinte como receptor (EDX Wireless LLC, 2008). Os resultados dessa sofisticação podem ser vistos na Figura 25.

Tabela 7. Modelos de propagação do *software* EDX usados na Av. Canadá.

Modelo de Propagação	Erro Médio [dB]	Desvio-Padrão [dB]
Anderson 2D v1.00	15.54	10.22
FCC - EDX	31.22	9.65
FCC - FCC	31.33	9.69
FCC - Pt22	31.22	9.65
FCC + RMD	21.81	9.60
Free Space + RMD	21.46	9.28
Hata-extended-open	14.59	9.33
Hata-extended-suburban	-7.63	9.42
Hata-extended-urban	-12.68	9.33
ITU-R 1546 cold sea curves with delta H	41.90	9.11
ITU-R 1546 cold sea curves without delta H	41.90	9.11
ITU-R 1546 Land curves with delta H	27.86	9.34
ITU-R 1546 Land curves without delta H	27.86	9.34
ITU-R 1546 Warm sea curves with delta H	41.90	9.11
ITU-R 1546 Warm sea curves without delta H	41.90	9.11
ITU-R 370 Cold sea curves with delta H	41.97	9.12
ITU-R 370 Cold sea curves without delta H	41.97	9.12
ITU-R 370 Land curves with delta H	41.97	9.12
ITU-R 370 Land curves without delta H	41.97	9.12
ITU-R 370 Warm sea curves with delta H	41.97	9.12
ITU-R 370 Warm sea curves without delta H	41.97	9.12
ITU-R 370 + RMD Cold sea curves with delta H	32.14	8.60
ITU-R 370 + RMD Cold sea curves without delta H	32.14	8.60
ITU-R 370 + RMD Land curves with delta H	32.14	8.60
ITU-R 370 + RMD Land curves without delta H	32.14	8.60
ITU-R 370 + RMD Warm sea curves with delta H	32.14	8.60
ITU-R 370 + RMD Warm sea curves without delta H	32.14	8.60
Longley rice v1.2.2	17.64	10.30
Okumura-Hata-open	23.95	9.50
Okumura-Hata-suburban	0.323	9.26
Okumura-Hata-urban	-3.32	9.50
TIREM - EDX	17.13	10.03

Embora o modelo acerte ao prever uma diminuição do sinal devida à perda por difração, erra na localização dessa região de sinal fraco. É possível que esse desvio se deva ao fato de o modelo considerar apenas o relevo do terreno, e não as construções nele implantadas.

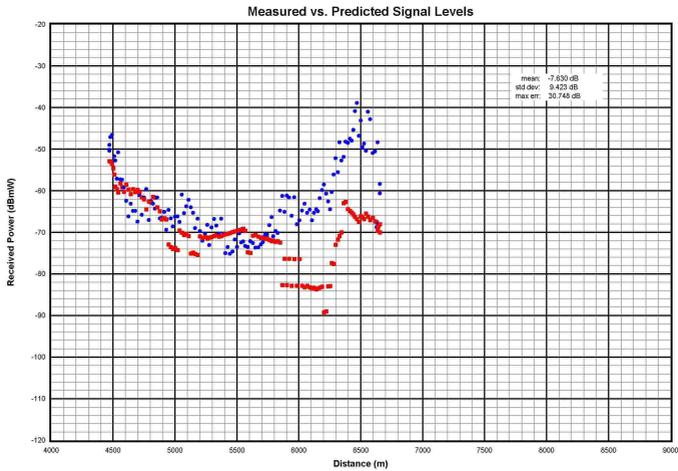


Figura 25. Comparação entre os valores medidos (círculos azuis) e a previsão do *software* EDX (quadrados vermelhos) segundo o modelo Hata-extended-suburban, na avenida Canadá. Erro médio=-7,63dB; desvio-padrão do erro=9,42dB; erro máximo=30,74dB.

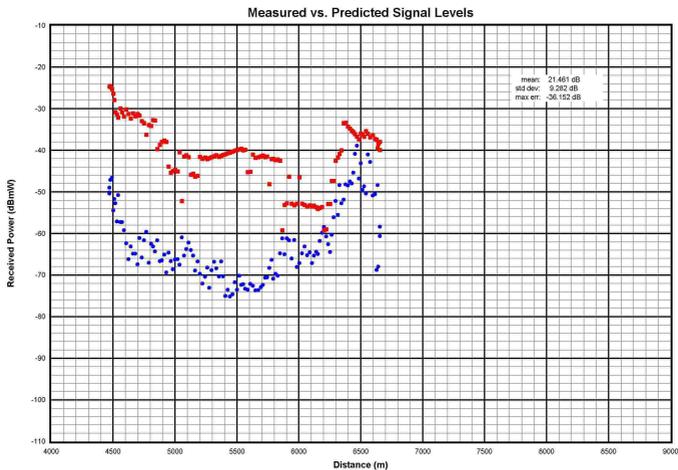


Figura 26. Comparação entre os valores medidos (círculos azuis) e a previsão do *software* EDX (quadrados vermelhos) segundo o modelo Free space+RMD, na avenida Canadá. Erro médio=21,46dB; desvio-padrão do erro=9,28dB; erro máximo=-36,15dB.

Há também modelos que fizeram uma previsão bastante imprecisa. A título de exemplo, tomamos o modelo denominado Free space +RMD (*reflection plus multiple diffraction*), que acrescenta as perdas por difração à perda normal de espaço livre. Suas previsões podem ser comparadas às medições na Figura 26.

3.6.2 Cobertura

A análise em separado de diversos trechos onde foram coletados dados de medição em campo foram analisados sob vários aspectos em (Borba, 2011). Um resumo dos resultados pode ser encontrado na Tabela 8.

Tabela 8. Resultados por região.

Região	Nome do trecho	Desvio-padrão [dB]	% recepção acima do limiar
Suburbana	Canadá afastamento 1	2,87	93,33
Suburbana	Canadá afastamento 2	2,65	80
Suburbana	Paraná aproximação 1	2,76	87,78
Suburbana	Paraná aproximação 2	2,78	86,79
Urbana	Guarapuava passagem 1	2,47	100
Urbana	Guarapuava passagem 2	2,72	100
Urbana	Batel e Dr. Pedrosa	2,82	100
Urbana densa	Silva Jardim início	1,76	95
Urbana densa	Silva Jardim passagem 1	2,55	94,62
Urbana densa	Dulcídio	6,09	100
Urbana densa	Deodoro e Pernetá	2,54	100

3.6.3 Aspectos práticos

A configuração do material necessário às medições já foi descrita na Seção 3.3.5. O item mais caro é o analisador de espectro, encarregado de medir a potência e de registrar o perfil de ecos. Também foi constatado, na Seção 3.5.2, que a importância de medir o perfil de ecos não é preponderante, pelo menos quando não há uso de SFN. Portanto, é concebível substituir o analisador de espectro por um equipamento de menor custo, capaz de medir a potência recebida, mesmo sem registrar o perfil de ecos. Um sintonizador poderia medir a potência a partir de uma curva de calibração de seu AGC (*controle automático de ganho*), solução já adotada em (Akamine et al., 2010). A incerteza na medição de potência efetuada com o AGC de um sintonizador certamente não será tão estreita quanto o 1,0dB que representa a incerteza do analisador de espectro utilizado (Rohde & Schwarz, 2011). Não obstante, em face da

variabilidade do mensurando, os valores obtidos com esse método mais econômico ainda podem atender aos propósitos da medição.

O GPS e o *notebook* são equipamentos comuns no comércio, e também o inversor para a alimentação, de modo que o custo da configuração pode permanecer na faixa de alguns milhares de reais.

Os equipamentos podem ser embarcados em veículo de praticamente qualquer porte, não é necessário o emprego de uma *van* de uso dedicado.

O levantamento de campo pode ser dimensionado a partir do tamanho estipulado para a quadrícula e do número de medidas que se deseja obter dentro de cada uma delas.

Para uma quadrícula com 150m de lado, cada quilômetro de percurso cobre uma área de 0,15 quilômetros quadrados. Uma área urbana com cem quilômetros quadrados pode ser pesquisada com 667km rodados. A uma velocidade média de 20km/h, e com 6h de levantamento por dia, é possível concluir a campanha em menos de seis dias, com uma taxa da ordem de 18 quilômetros quadrados por dia.

As medidas analisadas neste trabalho foram coletadas com tempo ensolarado, temperatura ambiente a 15 graus Celsius. As primeiras análises foram reunidas em (Borba et al., 2010).

Consideradas as particularidades e as complexidades peculiares à medição da cobertura de TV digital, em tudo semelhantes à determinação da área coberta por outros sistemas sem-fio, as seções a seguir embasam a convicção de que o método estudado é válido, mesmo apresentando limitações e possibilidades de aperfeiçoamento futuro.

3.6.4 Viabilidade do método

As considerações feitas na Seção 3.6.3 permitem admitir que o método é viável, em termos de custos e de prazos. A principal vantagem do método experimental porém é a delimitação precisa das áreas de recepção difícil para que as correções possam ser encaminhadas, atributo que os métodos teóricos ainda não conseguiram reproduzir. Adicionalmente, o levantamento de campo pode confirmar os limites da cobertura teórica negociada com os órgãos de administração do espectro, bem como estimular a geração de receitas a partir do interesse dos anunciantes.

3.6.5 Determinação da cobertura real

Mesmo dentro das limitações de custo e de prazo consideradas na Seção 3.6.3, a metodologia aqui empregada fornece aos teledifusores uma informação muito próxima da cobertura real de suas estações, já consideradas as influências do relevo, das construções, do multipercorso, das interferências e do ruído radioelétrico, condições difíceis de reproduzir em simulações, em especial simultaneamente.

3.6.6 Coerência dos modelos

A concordância entre as previsões dos vários modelos e as medições experimentais foi considerada na Seção 3.6.1. A partir dos dados analisados, é possível formar a convicção de que os modelos, em sua maioria corretos ao prever valores medianos do sinal, não apresentam a mesma precisão ao quantificar e ao localizar as áreas de recepção mais difícil. É possível que a acurácia dos modelos se aperfeiçoe, à medida em que os *softwares* incorporem aos seus bancos de dados informações mais detalhadas sobre a localização e as propriedades eletromagnéticas dos obstáculos. Permanece a dúvida se buscar essas informações mais detalhadas não supera, em dificuldade, a medição direta do sinal de interesse.

3.6.7 A configuração do esquema de modulação

A concepção do sinal de TV digital permite ao teledifusor uma ampla gama de opções no que diz respeito ao esquema de modulação. A configuração utilizada nos ensaios, e apresentada na Tabela 5, pode ser alterada, para atingir objetivos específicos de serviço, como recepção móvel, por exemplo. Validar essas escolhas é mais um emprego que pode ser dado ao método aqui utilizado.

3.6.8 Limitações

Embora não tenha sido implementada nas medições realizadas, a capacidade de avaliar a recepção em mobilidade é uma das possibilidades da metodologia apresentada.

Basta que o veículo possa se movimentar nas velocidades exigidas pelo teste, e que o sistema de registro permaneça operacional nessa situação, para que o comportamento em mobilidade possa ser detalhadamente avaliado. Já a recepção *indoor*, que é uma informação importante para o teledifusor, não pode ser aferida pelo método aqui estudado.

3.6.9 Recomendações para trabalhos futuros

O grande diferencial da TV digital terrestre é que o equipamento que a recebe não precisa ter fios. A mobilidade e a portabilidade são vantagens competitivas importantes quando comparada a outras modalidades (IPTV, satélite, cabo). Há indícios fortes de que o sistema brasileiro de TV digital possibilita a recepção móvel, mesmo em alta definição, desde que amparada por diversidade espacial (Akamine et al., 2010).

Nesse contexto, seria interessante esclarecer em trabalhos futuros:

- As antenas internas;
- O comportamento do sinal dentro de prédios, a partir daquele medido no nível da rua;

- A variação do sinal com a altura, no exterior e no interior de prédios urbanos.

Um desdobramento mais ambicioso do método automatizada de medição validado neste trabalho seria a criação do que poderia ser denominado um “mapa-rádio” da região sob estudo. Em vez de medir apenas *um* canal de televisão digital terrestre, informações adicionais poderiam ser coletadas, limitadas apenas pela capacidade do equipamento em uso. Nelas poderiam ser incluídos:

- Outros canais de TV, digitais ou analógicos;
- Emissoras de FM (*frequência modulada*);
- Emissoras de AM (*amplitude modulada*);
- Estações radiobase de telefonia celular;
- Outros sinais cuja intensidade e expectativa de permanência justifiquem seu registro em um banco de informações-rádio.

O conjunto de informações-rádio poderia ser associado a um sistema de dados geográficos, o que criaria um “mapa-rádio” da região. Ele permitiria que cada radiodifusor conhecesse detalhadamente a área de sua cobertura. Propiciaria aos órgãos de administração do espectro um conhecimento muito mais detalhado da utilização desse recurso em sua região de responsabilidade, embasando decisões sobre interferências, ou a respeito de áreas que necessitem do apoio de *gap-fillers*.

O uso mais desafiador de um ‘mapa-rádio’ porém, talvez seja possibilitar um sistema de radiolocalização de baixo custo, independente de satélites, capaz de operar *dentro* de construções e em plataformas móveis, e que tem a parte mais cara de sua infra-estrutura, os transmissores, já implantada.

4. Conclusões

A estimativa da atenuação por propagação é complexa devido à grande quantidade de variáveis envolvidas, como as condições ambientais, as condições do terreno ao longo do percurso da propagação e os obstáculos da aglomeração urbana no trajeto. No primeiro estudo de caso, o levantamento na cidade de Curitiba apresentou desde áreas urbanas densamente habitadas até áreas rurais, e com diversas situações topográficas. Isso dificultou a aplicação de um modelo único de propagação. Nessa abordagem, optou-se pela utilização de métodos ponto-área que, apesar de mais gerais, tornam os cálculos mais simples quando da sua aplicação. Com base no primeiro estudo de caso, com ênfase na Previsão da Propagação em Curitiba, pode-se concluir que:

- O método log-distância misto inovou na aplicação de regressões múltiplas para o método log-distância, na busca de encontrar o ponto ideal da distância de joelho, associando as características físicas da propagação à característica do cálculo da regressão;
- Adicionalmente o método log-distância misto provou ser a união, em um único método, dos métodos de Hata para área urbana em cidades grandes e para área suburbana para a região de Curitiba;
- O método log-distância misto inovou na aplicação de regressões múltiplas para o método log-distância, na busca de encontrar o ponto ideal da distância de joelho, onde existe uma mudança do comportamento da propagação, associando as características físicas da propagação à característica do cálculo da regressão;
- O método log-distância misto proposto neste trabalho, obtido empiricamente, apresentou uma melhor estimativa quando comparado com outros métodos tradicionais, para toda a faixa de pontos medidos, podendo ser aplicado para a região de Curitiba.

Com base no segundo estudo de caso, com ênfase na Medição da Propagação em Curitiba, é possível concluir que:

- As considerações feitas na Seção 3.6.3 permitem admitir que o método é viável, em termos de custos e de prazos. A principal vantagem do método experimental porém é a delimitação precisa das áreas de recepção difícil para que as correções possam ser encaminhadas, atributo que os métodos teóricos ainda não conseguiram reproduzir;
- Mesmo dentro das limitações de custo e de prazo consideradas na Seção 3.6.3, a metodologia de medição descrita na Seção 3 fornece aos teledifusores uma informação muito próxima da cobertura real de suas estações, já consideradas as influências do relevo, das construções, do multipercurso, das interferências e do ruído radioelétrico, condições difíceis de reproduzir em simulações, em especial simultaneamente.
- Embora os modelos apropriados acertem no valor mediano do campo elétrico, são superados pela medição no que diz respeito ao detalhamento da cobertura. As minúcias associadas às implantações sobre o terreno ainda são difíceis de incorporar aos modelos.
- Validar escolhas no que diz respeito ao esquema de modulação (intervalos de guarda, por exemplo) é mais um emprego que pode ser dado ao método de medição aqui utilizado.

Alguns desdobramentos futuros podem ser concebidos a partir do presente trabalho:

- A pesquisa de modelos empíricos, adaptados aos dados coletados em Curitiba, para outras faixas de frequência;
- A viabilidade da *recepção móvel* em LD (*low definition*), e mesmo em HD (*high definition*);
- O desempenho das antenas internas;
- O comportamento do sinal dentro de prédios, a partir daquele medido no nível da rua;
- A variação do sinal com a altura, no exterior e no interior de prédios urbanos.

Referências

- ABNT, , NBR 15604-1:2007 Televisão digital terrestre - Receptores. 2007.
- Akamine, C.; Yamada, F.; Bedicks Jr., G.; Dantas, C.; Sukys, F.; Motoyama, R.; Maroja, R.; Valeira, G. & Horta, E., Field trials for Brazilian DTV using space diversity. In: *Proc. of IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)*. p. 1–11, 2010.
- Albertazzi, A. & Sousa, A.R., *Fundamentos de Metrologia Científica e Industrial*. 1a edição. Barueri, SP: Manole, 2008.
- Anatel, , N - 002/91 - Norma para Determinação da Intensidade de Campo em Canais de VHF e UHF. 1991. Anexo à Portaria N° 053, de 07 de junho de 1991.
- Anatel, , Anexo II à Resolução 398 de 7 de abril de 2005. 2005.
- Anatel, , Descrição do Sistema - Características Técnicas de Operação da Estação 690881088. 2006.
- Anatel, , Ato 1388, de 12 de março de 2008. Diário Oficial da União, 2008a. Brasília, seção 1, p.78-79.
- Anatel, , SIGAnatel - Sistema de Informações Geográficas da Anatel. 2008b. <http://sistemas.anatel.gov.br/siganatel>.
- Borba, M.A., *Medição de Cobertura no Sistema Brasileiro de TV Digital Terrestre*. Dissertação de mestrado, Universidade Tecnológica Federal do Paraná – CPGEI, Curitiba, PR, 2011.
- Borba, M.A.C.; Mayer, C.E.N.; Fonseca, K.V.O.; Pohl, A.A.P.; Good, K.R.; Mildemberg, L.; Wosch, D.L. & Martins, R.A., Investigação experimental de sinais de TV digital com multipercurso na cidade de Curitiba. *Revista de Radiodifusão*, 4(4), 2010.
- Brasil, , Decreto n° 4.901, de 26 de novembro de 2003. Diário Oficial da União, 2003.
- Brasil, , Decreto n° 5.820, de 29 de junho de 2006. Diário Oficial da União, 2006.

- EDX Wireless LLC, . *EDX Signal: Comprehensive Engineering Software for Wireless Communication System Design*. Oregon, USA, 2008. p. 10. Rev. 018. <http://www.edx.com>.
- Feuerstein, M.J.; Blackard, K.L.; Rappaport, T.; Seidel, S.Y. & Xia, H.H., Path Loss, Delay Spread, and Outage Models as Functions of Antenna Height for Microcellular System Design. *IEEE Transactions on Vehicular Technology*, 43(3):487–498, 1994.
- Gonzalez Castellanos, P.V.; Pérez García, N.A.; Paredes, J.L.; Uzcátegui, J.; Peña, J.; Duque, L.; Silva Mello, L.A.R. & Souza, R.S.L., Measurements of Terrestrial Digital TV signals at Two Cities in South America. In: Press, I. (Ed.), *Proc. of IEEE Fourth European Conference on Antennas and Propagation (EuCAP)*. Piscataway, USA, p. 1–4, 2010.
- Hata, M., Empirical Formula for Propagation Loss in Land Mobile Radio Services. *IEEE Transactions on Vehicular Technology*, 29(3):317–325, 1980.
- Miyoshi, E.M. & Sanches, C.A., *Projetos de Sistemas Rádio*. 2a edição. São Paulo: Editora Érica, 2002.
- Plumb, G.D.; Pullen, I.R. & Tait, B., UK DVB-T network planning and coverage verification. In: *Proc. International Broadcasting Convention Conference (IBC)*. p. 393–398, 1999.
- Pullen, I., DTT coverage predictions – how they are made and tested. *Digital News*, 9:14–15, 1999.
- Radiocommunication Study Groups ITU, , *ITU Recommendation ITU-R P.1407: Multipath Propagation and Parameterization of its Characteristics*. Geneva, Switzerland, 1999.
- Radiocommunication Study Groups ITU, , *ITU Recommendation ITU-R P.1546-1: Method for Point-to-area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*. Geneva, Switzerland, 2003.
- Radiocommunication Study Groups ITU, , *ITU Recommendation ITU-R BT.1735: Métodos de Evaluación Objetiva de la Calidad de la Cobertura de las Señales de Radiodifusión de Televisión Digital Terrenal del Sistema B Especificado en la Recomendación UIT-R BT.1306*. Geneva, Switzerland, 2005.
- Radiocommunication Study Groups ITU, , *ITU Recommendation ITU-R P.1406-1: Propagation Effects Relating to Terrestrial Land Mobile and Broadcasting Services in the VHF and UHF Bands*. Geneva, Switzerland, 2007a.

- Radiocommunication Study Groups ITU, , *ITU Recommendation ITU-R P.1546-3: Method for Point-to-area Predictions for Terrestrial Services in the Frequency Range 30 MHz to 3000 MHz*. Geneva, Switzerland, 2007b.
- Rappaport, T.S. (Ed.), *Wireless Communications Principles and Practice*. 1a edição. Communications Engineering and Emerging Technologies. Upper Saddle River, USA: Prentice Hall, 1996.
- Rappaport, T.S. (Ed.), *Wireless Communications Principles and Practice*. 2a edição. Communications Engineering and Emerging Technologies. Upper Saddle River, USA: Prentice Hall, 2002.
- Rohde & Schwarz, , *Signal-strength and Interference Measurements – News from Rohde & Schwarz-Polarad*. N. special 1. Munich, Germany: Rohde & Schwarz-Polarad, 1985. 86 p.
- Rohde & Schwarz, , *ETL TV Analyzer Specifications*. Munich, Germany, 2011. Data Sheet Version 11.00.
- Spiegel, R.M., *Estatística*. São Paulo: McGraw-Hill do Brasil Ltda, 1977.
- Terman, F.E. & Pettit, J.M., *Electronic Measurements*. 2a edição. Tokyo, Japan: Kogakusha Company, Ltd., 1952.
- Vasco, C.L., *Modelo de Propagação Empírico para Sinais de TV Digital em Curitiba*. Dissertação de mestrado, Universidade Tecnológica Federal do Paraná – CPGEI, 2009.
- Vasco, C.L.; F., R.J.; Martins, R.A.; Agostinho, E.; Pohl, A.A.P.; Souza, R.D.; Fonseca, K.V.O. & Pellenz, M.E., Ensaio de transmissão e recepção de TV digital terrestre na região metropolitana de Curitiba. *Revista de Radiodifusão*, 3(3):63–76, 2009.
- Weck, C., VALIDATE Field Trials of Digital Terrestrial Television (DVB-T). Relatório Técnico, Institut für Rundfunktechnik GmbH (Broadcasting Systems Development), Munich, Germany, 1999.

Alternativa Tecnológica Para Canal de Retorno da TV Digital Baseada em Redes IEEE802.11b/g

Tânia Lúcia Monteiro*, Marcelo Eduardo Pellenz, Manoel Camillo Penna,
Fabrício Enembreck e Richard Demo Souza

Resumo: O governo Brasileiro, através da TV Digital Interativa, visa uma tecnologia de comunicação que permita à população o acesso às redes de telecomunicações e também a alfabetização digital. Neste contexto, um ponto importante é a definição do canal de retorno, sendo que as redes sem fio IEEE 802.11b/g são uma possível solução. Este capítulo apresenta três mecanismos distribuídos de alocação de canais em redes IEEE 802.11b/g, sendo portanto seu uso aplicável ao canal de retorno da TV Digital. Os métodos propostos apontam para soluções ótimas e subótimas, minimizam a interferência e otimizam a capacidade destas redes.

Palavras-chave: TV Digital, Canal de Retorno, Redes sem Fio, Alocação de Canais

Abstract: *The Brazilian government, through the Interactive Digital TV, aims at a communication technology that allows the access to the telecommunications networks as well as the digital inclusion. In this regard, an important aspect is the definition of the return channel, in which the IEEE 802.11b/g networks are a possible solution. This chapter presents three distributed channel allocation mechanisms which can be used in the Digital TV return channel. The proposed methods point to suboptimal and optimal solutions, minimize the interference and optimize the capacity of such networks.*

Keywords: *Digital TV, Return Channel, Wireless Networks, Channel Allocation*

*E-mails: tmonteiro@utfpr.edu.br ou taniamonteiro@ppgia.pucpr.br

1. Introdução

Ao surgir como um novo paradigma de comunicação, a TV Digital além de proporcionar inúmeras vantagens sobre o sistema analógico, tais como melhor qualidade de áudio e vídeo e maior número de programas em um único canal, apresenta-se sobretudo como uma nova forma de comunicação quando passa a oferecer aos telespectadores a possibilidade de interagirem com a emissora, permitindo aos mesmos passarem de uma atividade essencialmente passiva para uma atividade participativa. Surge então a TV Digital Interativa onde a ação de interatividade representa a capacidade de um dispositivo interagir ou permitir interação com o seu respectivo usuário. Deve-se observar que a existência de interatividade está estritamente relacionada à existência de um meio eletrônico, intermediando a interação (de Oliveira & Albuquerque, 2005). Para Reisman (2002) e Lemos (1997) o conceito de interatividade pode ser classificado sobretudo quanto ao nível de abrangência: *restrita* ao uso do controle remoto ou por exemplo de telefone ou mensagens de correio eletrônico (e-mail); *plena* onde o telespectador pode controlar o conteúdo do programa que está assistindo, sendo capaz de controlar a programação que gostaria de assistir; e *coativa* onde são acrescentadas funcionalidades que permitirão aos telespectadores obterem informações sobre esportes, meteorologia, notícias, etc., além de poderem acessar detalhes sobre produtos anunciados e até efetuarem a compra dos mesmos. Já Montez & Becker (2005) complementam ao considerarem a possibilidade do telespectador interferir plenamente na programação e não apenas reagir aos programas transmitidos pelas emissoras. Diante de níveis superiores de abrangência, surge a necessidade de um canal de comunicação que seja capaz de prover recursos para a transmissão de conteúdo do telespectador para a emissora.

Paralelo ao gradual acesso a toda esta tecnologia, pode-se observar que os mais diversos serviços eletrônicos, virtuais e interativos vem se ampliando e popularizando, criando demandas e tornando ainda maior a quantidade e a diversidade de usuários que desejam acessar estas redes. Neste sentido, em função das limitações da população brasileira, tanto de acesso às redes de telecomunicações quanto de alfabetização digital, uma das alternativas que se consolida é a utilização do receptor de TV como um terminal eficiente e simples para acesso a estes serviços, independente das redes de telecomunicações a que estejam conectados. A Televisão Digital Interativa se apresenta então como uma alternativa de baixo custo a esta realidade, apresentando-se como solução tecnológica que pode atender à faixa da população excluída digitalmente, possibilitando acesso à Internet, ensino a distância, governo eletrônico, entretenimento, serviços de saúde, consulta a saldo do FGTS, etc. O dispositivo capaz de receber o sinal transmitido pela emissora pode estar embutido em uma televisão digital ou pode ser um dispositivo externo, neste caso, denominado set top

box ou IRD (*Integrated Receiver Decoder*) ou URD (*Unidade Receptora-Decodificadora*). Em se tratando de algum serviço de interatividade, o receptor que possui facilidades de um microcomputador, deverá enviar dados do telespectador para a emissora ou para outro destino qualquer, provendo a comunicação com o canal de comunicação, denominado de canal de retorno ou canal de interatividade.

A tecnologia sobre o qual o canal de comunicação (canal de retorno) é implementado, define suas características de transmissão. Esta solução não se apresenta como única e dependerá de características e disponibilidade de recursos de cada região, cidade, operadora, usuário, etc., assim como da concentração populacional local que poderá viabilizar ou não as mais diversas soluções. Dentre as soluções possíveis, pode-se citar: PLC, telefonia fixa, telefonia celular, ADSL, satélite, WiMax, redes Wi-Fi¹, etc.

- PLC (*Power Line Communication*) – Esta tecnologia permite a utilização da rede de energia elétrica para transmissão de sinais digitais, possibilitando, desta forma, que qualquer ponto de energia possa vir a ser um ponto de rede. Apresenta-se como uma solução interessante devido a sua capilaridade e taxas de transmissão. Contudo, segundo (ANATEL, 2003; Forum PLC, 2006) os estudos para implantação em larga escala não foram conclusivos;
- STFC (Sistema de Telefonia Fixa Comutada) – Embora seja uma alternativa possível, esta tecnologia dispõe de uma taxa de transmissão baixa (56kbps), limitando a banda de transmissão, e apresenta custo de tarifação que acabam por comprometer sua utilização;
- Telefonia Celular – A transmissão de dados através de redes celulares vem evoluindo e hoje, com a tecnologia 3G, suporta taxas de transmissão de cerca de 700kbps, podendo chegar a 1Gbps com a versão 4G. No entanto, o custo destes serviços ainda é proibitivo para a maioria da população brasileira;
- ADSL (*Assymetrical Digital Subscriber Line*) – Esta tecnologia é muito utilizada como solução de banda larga em ambientes comerciais e residências, pode chegar a 8Mbps. Como alternativa tecnológica não dispõe da capilaridade necessária. Sua contratação é considerada desvantajosa em relação ao custo final, pois o acesso a tecnologia está associado à disponibilidade de uma linha telefônica e de um provedor de conteúdo;
- Satélite – Esta tecnologia possibilita um grande alcance a consideráveis taxas de transmissão, entretanto o custo não é acessível para a maioria da população;

¹ Termo licenciado pela Wi-Fi Alliance para se referir a redes sem fio baseadas no padrão IEEE 802.11

- WiMax (*Worldwide Interoperability for Microwave Access*) – O padrão IEEE 802.16, é capaz de prover uma taxa de transmissão aproximada de 70Mbps, com alcance de até 50 km (em condições ideais). Existe expectativa de que a tecnologia WiMax² seja uma alternativa viável para solução de canal de retorno;
- Redes Wi-Fi – Redes locais sem fio baseadas no padrão IEEE 802.11. Operam em faixas de frequências que não necessitam de licença para instalação e/ou operação. Possuem capacidade teórica que varia de 11Mbps a 600Mbps, dependendo do padrão adotado. Possuem tecnologia extremamente flexível o que torna viável sua implantação nos mais variados locais a custos reduzidos.

As redes sem fio padrões 802.11b/g estão se tornando cada vez mais populares, graças a seu baixo custo, capacidade, flexibilidade, cobertura, fácil implantação e operação simplificada. Uma diversidade de produtos com adaptadores de redes sem fio (computadores, telefones, eletrodomésticos, sensores, etc.) tem atendido a diferentes classes e perfis de consumidores, sendo utilizados em ambientes residenciais, edifícios, cidades, etc. Nos últimos anos estas redes têm sido amplamente investigadas para aplicação em dois cenários principais: as redes locais sem fio – WLANs (*Wireless Local Area Network*) e as redes *mesh* sem fio – WMN (*Wireless Mesh Networks*).

Quando se deseja considerar esta tecnologia de rede para prover acesso em banda larga, existem desafios relacionados ao desempenho que devem ser solucionados. O problema de maior impacto enfrentado pelas soluções WLAN e WMN, continua sendo a redução da capacidade total, devido à interferência co-canal (vizinhos configurados no mesmo canal) e de canal adjacente (vizinhos configurados em canais sobrepostos), que ocorrem entre nós da rede ou nós de redes vizinhas. Vale salientar que o nível de interferência depende sobretudo da atribuição dos canais, da distribuição geográfica dos APs e dos usuários, do número de clientes associados a cada ponto de acesso (AP) e também da carga de tráfego. A existência de outros fatores, não mencionados, podem intensificar ou não o nível de interferência detectado e a capacidade útil da rede.

Neste contexto apresenta-se a tecnologia baseada nos padrões 802.11b/g (IEEE802.11std, 2012) como solução de rede para implantação de canal de retorno. Esta tecnologia além de suas características técnicas permite o acesso por todas as camadas sociais da população, levando vantagem no quesito como tecnologia mais difundida entre as anteriormente citadas. Este trabalho foca em garantir a performance de cenários onde os APs podem pertencer a domínios diferentes, uma realidade muito comum na implantação de soluções sem fio em áreas muito densas. O problema de

² WiMAX Forum: <http://www.wimaxforum.org>

alocação de canais é formalizado como um DCOP (*Distributed Constraint Optimization Problem*), e são propostos dois novos algoritmos distribuídos que encontram a solução ótima, denominados DSAA-O (*Distributed Channel Assignment Algorithm – Optimal*) e DOCA (*Distributed Optimal Channel Assignment*), e duas estratégias subótimas, denominadas DSAA-S (*Distributed Channel Assignment Algorithm – Optimal*) e DSCA (*Distributed SubOptimal Channel Assignment*), desenvolvidas para reduzir o número de mensagens trocadas entre APs, mantendo a solução subótima muito próxima da ótima.

2. Fundamentação Teórica

O objetivo de uma WLAN é de prover conectividade sem fio sobre uma certa área, formada por múltiplos APs, onde cada AP possui uma abrangência máxima de cobertura. Já as WMN se destacam como uma das melhores alternativas para implantação de redes de acesso, sendo definidas como um conjunto de roteadores e clientes mesh, onde os roteadores mantêm-se praticamente estáticos e formam o backbone da rede, provendo acesso para clientes nativos *mesh* e de outras tecnologias. [Akyildiz & Wang \(2005\)](#) definem redes *mesh* como redes dinâmicas, auto-organizáveis e auto-configuráveis, cujos nós estabelecem e mantêm conexão entre eles, criando, de fato, uma rede ad hoc (sem infra-estrutura). Para [Bicket et al. \(2005\)](#) as duas arquiteturas representam abordagens comuns para implantação de redes urbanas comunitárias compartilhadas. As redes WMN através de uma solução de múltiplos saltos, utilizam antenas direcionais, capazes de prover enlaces de mais qualidade do que enlaces com antenas omnidirecionais, e informação da localização de seus clientes. Segundo os autores, esta abordagem requer um grupo de trabalho de maior preparo técnico mas apresenta uma rede com alta capacidade de transmissão e conectividade. As redes WLAN ao se basearem na operação individual dos APs para conexão direta com seus clientes, podem atuar de forma independente. Contudo, a solução não é capaz de prover a mesma cobertura que as redes WMN. Na prática uma grande quantidade de APs podem estar operando em uma mesma região. Este ambiente leva a um cenário onde o gerenciamento dos canais das redes sem fio é de extrema importância, uma vez que a interferência co-canal e de canais adjacentes pode reduzir significativamente a performance da rede.

Ao utilizarem tipicamente APs equipados de apenas uma interface de rádio, as redes acabam por sofrer degradação de sua capacidade devido à natureza *half duplex* do meio sem fio ([Gupta & Kumar, 2000](#)). O padrão 802.11 ([IEEE802.11std, 2012](#)) possui três extensões onde são definidas as faixas de frequências que podem ser utilizadas pelos dispositivos sem fio, assim como os diferentes tipos de esquemas de modulação e codificação que podem ser empregados. As extensões

802.11b (IEEE802.11b, 2000) e 802.11g (IEEE802.11g, 2004) do padrão 802.11 coordenam a operação dos dispositivos nas faixas de frequências não licenciadas, ISM (*Industrial Scientific and Medical Band*): 2,4 GHz, fornecendo taxas de transmissão de 11 Mbps (802.11b) e 54 Mbps (802.11g). No entanto estas faixas de frequências não são exclusivas das redes 802.11, podendo ser compartilhadas com outros dispositivos como: redes *Bluetooth*, rádios comunicadores, telefones sem fio, fornos de microondas, etc. Dessa maneira as redes 802.11 b/g acabam por sofrer interferências provindas de seu ambiente local. Embora diferentes técnicas de transmissão digital sejam definidas (DSSS³ ou FHSS⁴) e tenham o objetivo de minimizar o efeito de interferências, em alguns casos não apresentam resultados positivos.

As leis regulatórias da maioria dos países, especificamente na América do Norte e Europa, estabelecem pelo menos 11 canais diferentes, cujas portadoras são alocadas entre 2400 e 2483MHz, com a frequência central do canal começando em 2412MHz, espaçada em intervalos de 5MHz, até 2462MHz. Cada canal possui uma largura de banda de 22MHz. No entanto a especificação da camada física do protocolo IEEE 802.11 (IEEE 802.11 PHY) permite a operação simultânea de múltiplos canais não sobrepostos⁵ (por exemplo, 1, 6 e 11), desde que estejam espaçados 25MHz entre si. Estas faixas de frequências podem ser usadas livremente, embora deva-se respeitar os limites legais para a potência de transmissão. A capacidade de utilizar vários canais, beneficiando-se de todo o espectro disponível, aumenta substancialmente a largura de banda efetiva e a capacidade das redes sem fio (Ramachandran et al., 2006).

As técnicas de atribuição de canais visam melhorar a utilização do espectro sem fio. O crescimento continuado das implantações de soluções WLAN e WMN leva a cenários de rede onde um grande número de APs interferentes podem estar localizados em uma mesma região. Consequentemente a estratégia de atribuição de canais deve ser cuidadosamente planejada, caso contrário, pode haver uma perda de desempenho para os usuários das redes sem fio. Além disso, a utilização de uma técnica centralizada não é viável porque APs podem pertencer a diferentes domínios administrativos. Portanto, uma estratégia escalável e distribuída deve ser aplicada. Uma pergunta natural sobre algoritmos distribuídos é se eles são capazes de atingir a solução global ótima para o

³ DSSS - *Direct Sequence Spread-Spectrum*: esquema de modulação *spread-spectrum* que gera um padrão redundante de bits para cada bit transmitido.

⁴ FHSS *Frequency Hopping Spread*: um esquema de modulação *spread-spectrum* que utiliza uma portadora de banda estreita alterando a frequência segundo um padrão conhecido pelo transmissor e pelo receptor.

⁵ O que ocorre na realidade é que entre os canais 1 e 6 existe uma pequena quantidade de sobreposição, da mesma forma que entre os canais 6 e 11, contudo os demais canais sofrem uma sobreposição de canal adjacente significativa se comparada aos canais 1, 6 e 11.

problema de alocação de canais. Solução ótima significa a soma mínima de todos os fatores de interferência entre os APs interferentes na rede (Mishra et al., 2005).

As estratégias de atribuição de canais podem ser classificadas como estáticas ou dinâmicas (Kysanur & Vaidya, 2004). Na abordagem estática, os canais são atribuídos às interfaces de rádio de forma permanente ou por longos períodos. Em geral nenhum mecanismo de coordenação é necessário entre os nós, e os algoritmos buscam por uma solução topológica que maximize o uso da banda. De acordo com Ramachandran et al. (2006), com o crescimento explosivo das redes WiFi, que operam no mesmo espectro (não licenciado), como as redes WLAN ou as WMN, qualquer atribuição estática de canais provavelmente resultará na operação da rede, em canais que também estejam sendo utilizados por equipamentos vizinhos. Ocorre então um aumento resultante da interferência, o que poderá prejudicar o desempenho da rede. Já a estratégia dinâmica pressupõe que qualquer interface pode assumir ou trocar de canal de comunicação. Neste caso, um mecanismo de coordenação é necessário para garantir uma adequada correspondência de canal entre nós (APs) da rede. A capacidade de mudança dinâmica permite que vários canais possam ser utilizados por um número reduzido de interfaces, mas o desafio está no projeto do mecanismo de coordenação da comutação.

Baseados em arquitetura centralizada ou distribuída, alocação cooperativa ou não cooperativa, inúmeros métodos de acesso dinâmico ao espectro têm sido propostos, incluindo teoria de jogos (Nie & Comaniciu, 2005), mecanismos de leilões (Huang et al., 2006), negociação local (Cao & Zeng, 2005) e coloração de grafos (Zheng, 2005; Peng et al., 2006). Uma estratégia distribuída permite uma auto-reorganização parcial de setores da rede que estão sofrendo maior degradação ou perda momentânea de desempenho. A não dependência de um elemento ou nó de processamento central também é uma grande vantagem. Com uma técnica de otimização distribuída qualquer subconjunto de nós (APs) pode executar um algoritmo de realocação de canais.

3. Metodologia

Os mecanismos distribuídos e cooperativos desenvolvidos para a solução do problema de alocação de canais, foram baseados nos algoritmos ADOPT (*Asynchronous Distributed OPTimization*) (Modi et al., 2005) e DPOP (*Dynamic Programming Optimization Protocol*) (Petcu & Faltings, 2005; Petcu, 2007), mais detalhados a seguir.

3.1 Definição formal DCOP

Em pesquisas na área de inteligência artificial, o processo para encontrar uma solução para atender um conjunto de restrições que impõem

condições sobre um conjunto de variáveis, é classicamente denominado *constraint satisfaction* (Rossi et al., 2006). Formalmente um problema CSP (*Constraint Satisfaction Problem*) é definido como um conjunto de variáveis cujo estado deve satisfazer um número de restrições, que serão solucionadas através de métodos *constraint satisfaction*. Em cenários onde variáveis e restrições são distribuídas entre múltiplos agentes, tem-se um problema distribuído denominado DisCSP (*Distributed Constraint Satisfaction Problem*). A abordagem do DisCSP não é adequada para resolver determinadas situações do mundo real, pois requer que a solução seja caracterizada pela resposta *satisfeito* ou *insatisfeito* (Hirayama & Yokoo, 1997). Algumas soluções práticas podem requerer graus de qualidade ou de custo. Esta classe de problemas é modelada como um “problema de otimização distribuída com restrições” ou DCOP (Meisels, 2008). Um DCOP consiste em uma série de variáveis que são distribuídas a um grupo de agentes colaborativos como restrições de valor, ou seja, restrições que são descritas como funções de valor, as quais retornam valores em um intervalo específico. A meta é otimizar uma função objetivo global, minimizando-se o peso das restrições atendidas (Modi et al., 2003a, 2005).

Um problema de otimização distribuída com restrições é formalmente definido considerando-se um conjunto de N agentes, $A = \{a_1, a_2, \dots, a_N\}$, e um conjunto de N valores, $D = \{d_1, d_2, \dots, d_N\}$, onde cada valor d_j é atribuído a um agente a_j e pertence a um domínio discreto e finito \mathbb{D}_j . Um agente a_j tem controle de seus valores d_j armazenados, e conhecimento de seu domínio \mathbb{D}_j . Matematicamente, um DCOP é modelado através de um grafo de restrições não direcionado, $\mathbf{G} = (V, E)$, onde V representa o grupo de vértices (nós) e $E \subseteq V^2$ o grupo de arestas de conflito (*links*). Cada agente a_j representa um vértice $v_j \in V$ no grafo de restrição. Para cada par de agentes (a_i, a_j) conectado por uma aresta de conflito $e_k = (i, j) \in E$, no grafo de conflito, uma função custo $f_{ij}(x, y) : \mathbb{D}_i \times \mathbb{D}_j \rightarrow \mathbb{R}$ é definida. Os agentes devem se coordenar a fim de encontrarem um conjunto de valores que otimizem uma função global, estabelecendo custos para as restrições. Portanto, a meta de um algoritmo baseado em DCOP é encontrar o melhor conjunto, denominado D^* , cujos valores minimizem uma função custo global, $g(D) = g(d_1, d_2, \dots, d_N) = \sum_{e_k \in E} f_{ij}(x, y)$. No caso de dois agentes conectados (a_i, a_j) , com diferentes domínios a eles associados $\mathbb{D}_i \neq \mathbb{D}_j$, os agentes, além de conhecerem seu próprio domínio, devem conhecer o domínio dos agentes a eles conectados para que possam avaliar a função custo. No entanto, em muitos cenários DCOP os domínios discretos e finitos, \mathbb{D}_j , são os mesmos para todos os agentes.

Existem basicamente duas classes de algoritmos para resolver um DCOP. A primeira classe é executada diretamente sobre o grafo de restrição, sem impor ordem hierárquica no processo de comunicação entre os agentes. Alguns exemplos incluem os algoritmos *NC-BB*, *ADOPT-NB* e *OPT-APO* (Chechetka & Sycara, 2006; Modi et al., 2003b; Mailler &

Lesser, 2004). A segunda classe, onde estão incluídos os algoritmos ADOPT e DPOP, requerem durante suas execuções uma ordem pré-definida para troca de mensagens. Nestes casos, a ordenação dos agentes é geralmente obtida utilizando-se um algoritmo de ST (*Spanning Tree*). Para estes algoritmos, baseados em árvores, tem-se o conceito de agentes pais e filhos conectados a um agente a_j . É importante ressaltar que dois agentes não vizinhos podem estar conectados no grafo de restrições, por uma aresta de conflito, se ambos estiverem ligados a um agente descendente em comum a_k . Estes agentes pertencerão ao conjunto de agentes pais de a_k .

3.2 Algoritmo ADOPT

Modi et al. (2003a) mencionam em seu trabalho que uma grande classe de problemas de coordenação, tais como planejamento distribuído, escalonamento distribuído, alocação de recursos distribuídos, entre outros, podem ser modelados como um DCOP. Os autores propõem um algoritmo de otimização de restrições, distribuído e de espaço-polinomial chamado ADOPT. O algoritmo ADOPT é o primeiro algoritmo para DCOP que pode achar a solução ótima, ou subótima dentro de um limite especificado pelo usuário, a partir da ótima. Utiliza apenas comunicação assíncrona e localizada entre os agentes. As idéias básicas do ADOPT incluem:

- uma nova estratégia de pesquisa assíncrona onde as soluções podem ser abandonadas antes de serem comprovadas como subótimas,
- reconstrução eficiente das soluções abandonadas, e
- detecção de encerramento.

A primeira versão do ADOPT apresentada em Modi et al. (2003a) é denominada *Simple-ADOPT*⁶. Este algoritmo é definido pelos autores como um BSA (*Backtracking Search Algorithm*), executado assincronamente e em paralelo em todos os agentes, atualizando a cada iteração limites inferiores, garantindo assim a qualidade da solução global. O algoritmo requer que os agentes sejam priorizados em uma ordem total, por meio de um algoritmo distribuído de ST (*Spanning Tree*). A ST é obtida através do algoritmo DFS (*Depth-First Search*), uma ordenação para árvore de pesquisa inicialmente proposta por Cheung (1983). Considerando a ordenação adotada, os autores usam o termo pai (*parent*) para se referirem a um agente (imediatamente conectado) de prioridade superior, e filho (*child*) para se referir a um agente (imediatamente conectado) de prioridade inferior. A Figura 1b ilustra a ordem total assumida pelos agentes para o grafo de restrições da Figura 1a, gerada através do algoritmo DFS. Dada esta ordem, como indicado

⁶ As versões mais recentes do algoritmo ADOPT (Modi et al., 2005), não foram consideradas porque requerem mensagens de controle adicionais e mais recursos para armazenamento de informações locais.

na Figura 1c, um agente comunica seu valor d_j para todos os agentes de prioridade mais baixa. Após a execução desta operação, o agente comunica o custo limite inferior, a um único agente de prioridade mais alta, como indicado na Figura 1d. Para alterar o valor local de cada agente e informações de custo mínimo entre agentes, o algoritmo *Simple-ADOPT* utiliza dois tipos de mensagens de controle: mensagens de *VALUE* e de *VIEW*. A direção do fluxo destas mensagens, também pode ser observada nas Figuras 1c e 1d. Estes grafos constituem parâmetros importantes para o algoritmo e são denominados de grafo de *VALUE* e grafo de *VIEW*. Seja P_j^{value} o conjunto de pais e C_j^{value} o conjunto de filhos de um agente a_j no grafo de *VALUE*. Da mesma forma, seja P_j^{view} o conjunto de pais e C_j^{view} o conjunto de filhos de um agente a_j no grafo de *VIEW*. É importante ressaltar que um agente a_j tem apenas um pai no grafo de *VIEW*, como indicado na Figura 1d.

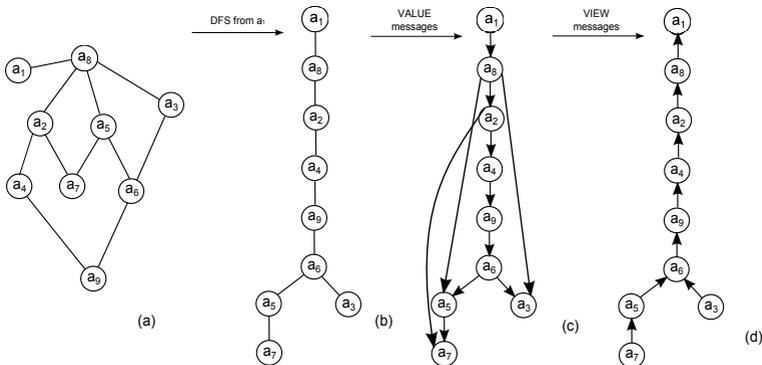


Figura 1. Exemplo DCOP : (a) Grafo de restrições; (b) Grafo de restrições com agentes totalmente ordenados; (c) Grafo para fluxo das mensagens de VALUE; (d) Grafo para fluxo das mensagens de VIEW.

O algoritmo *Simple-ADOPT* inicia com cada agente a_j escolhendo localmente e concorrentemente um valor $d_j \in \mathbb{D}_j$. Este valor é enviado, através de uma mensagem de *VALUE*, para todos os seus agentes descendentes, conectados a ele, $\{a_{k1}, a_{k2}, \dots, a_{kn}\} \in C_j^{value}$. Os agentes filhos assincronamente esperam e respondem pelas mensagens recebidas. Dada uma mensagem de *VALUE* recebida de seus agentes pais $\{a_{i1}, a_{i2}, \dots, a_{il}\} \in P_j^{value}$, o agente a_j escolhe seu valor local d_j de acordo com a Equação 1.

$$d_j = x \mid \min_{x \in \mathbb{D}_j} \sum_{a_i \in P_j^{value}} f_{ij}(d_i, x). \quad (1)$$

Durante a execução do algoritmo, cada agente a_j deve considerar três valores distintos de custo, o custo local (*local cost*), o menor custo corrente (*current lower bound cost*) e o menor custo local estimado (*estimated lower bound cost*). O custo local, $lc_j(d_j)$, é definido por $lc_j(d_j) = \sum_{a_i \in P_j^{value}} f_{ij}(d_i, d_j)$. Este custo é calculado através da soma dos resultados da função custo para todas as mensagens de *VALUE* recebidas dos agentes (pais), utilizando os valores $d_j \in \mathbb{D}_j$, obtidos através da Equação 1. O menor custo local para a subárvore, $sc_j(d_j)$, é atualizado conforme $sc_j(d_j) = \max(sc_j(d_j), z_k^*)$, onde z_k^* representa o menor custo local estimado recebido através do grafo de *VIEW*. Para o nó a_j o custo z_k^* é calculado por $z_j^* = lc_j(d_j) + sc_j(d_j)$, onde d_j passa então a assumir o valor $d_j = x \mid \min_{x \in \mathbb{D}_j} lc_j(x) + sc_j(x)$. Se um agente não tiver nenhum descendente, seu menor custo estimado, z_j^* , será somente seu custo local lc_j .

Sempre que um agente a_j recebe uma mensagem de *VALUE* de um antecessor a ele ligado, $a_i \in P_j^{value}$, este agente armazena o valor corrente recebido (a_i, d_i) em sua variável **Currentvw_j**, representando assim o contexto corrente de a_j . Para este contexto corrente, o agente a_j reporta para seu pai, $a_i \in P_j^{view}$, o menor custo estimado z_j^* , através de uma mensagem de *VIEW*. Por sua vez, o seu agente pai a_i , irá usar o menor custo estimado recebido z_j^* para calcular seu menor custo estimado para com sua respectiva subárvore, $sc_i(d_i)$. O agente pai desconsidera uma mensagem de *VIEW* quando existir uma incompatibilidade de contexto. Esta situação pode acontecer em dois casos: qualquer agente pai a_i tem uma view corrente mais atualizada que os valores armazenados pelo seu filho a_j , ou a_j tem sua view corrente, mais atualizada que a de seu pai a_i . Quando um agente a_j reporta para seu pai P_j^{view} , seu menor custo estimado z_j^* , localmente o agente assume o valor d_j que minimiza z_j^* . Consequentemente os limites inferiores são sempre conservados, o que é necessário a fim de se garantir a integridade da solução. Se isto não ocorrer, o custo total de uma solução pode ser superestimado ou erroneamente descartado por um agente pai.

Finalmente, quando um agente a_j recebe uma mensagem de *VALUE*, este atualiza seu contexto corrente, **Currentvw_j**. O agente então deleta seu custo armazenado z_j^* , uma vez que este agora pode estar inválido. O algoritmo *Simple-ADOPT* armazenando apenas um valor de contexto corrente, possui requisitos de espaço linear em cada agente. O algoritmo alcança a estabilidade quando todos os agentes passam a aguardar por uma mensagem. Neste caso a atribuição completa escolhida pelos agentes é igual ao conjunto ótimo de atribuições, denominado, D^* , cujos valores minimizam a função custo global, $g^* = g(D^*)$.

3.3 Algoritmo DPOP

O algoritmo DPOP (Petcu & Faltings, 2005; Petcu, 2007) emprega um método de propagação de valores resultantes de uma função utilidade, com base em programação dinâmica, inspirado em algoritmos de soma-produto. Considera uma *spanning tree* associada ao grafo de restrições, \mathbf{G} , obtida com a execução do algoritmo DFS (Petcu, 2007). Conforme ilustrado na Figura 2, o grafo de restrições é composto de arestas denominadas *tree-edges* (linhas contínuas), que pertencem a ST, e arestas *back-edges* (linhas tracejadas), que não fazem parte da ST. Um caminho no grafo que é inteiramente representado por arestas *tree-edges* é denominado *tree-path*. É importante ressaltar que a execução de um algoritmo distribuído para a criação de uma *spanning tree* se inicia a partir de um dos nós a_j . Este nó pode ser escolhido através de um algoritmo distribuído de eleição de líder, ou simplesmente adotando-se um critério baseado na identificação dos nós (menor ou maior ID). Conforme Petcu & Faltings (2005) e Petcu (2007), para se gerar uma ordenação através de um algoritmo de DFS é necessário um número linear de $2 \times |E|$ mensagens, onde $|E|$ representa o número de arestas no grafo de restrições. O tamanho destas mensagens também é linear.

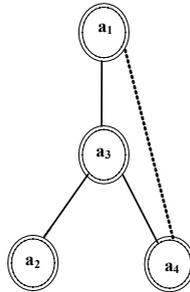


Figura 2. Exemplo de pseudoárvore para o DPOP

O princípio de operação do algoritmo DPOP, considera as seguintes definições: um conjunto de agentes (APs), $\mathbb{A} = \{a_1, a_2, \dots, a_N\}$; um domínio finito composto das variáveis dos agentes (conjunto de canais disponíveis), $\mathbb{D} = \{d_1, d_2, \dots, d_N\}$; um nó único P_j (pai de a_j), superior na hierarquia da árvore DFS, que está diretamente conectado ao nó a_j através de uma aresta *tree-edge* (ex. $P_3 = a_1$); um conjunto de nós inferiores na árvore DFS, C_j (filhos de um nó a_j), que estão diretamente conectados ao nó a_j através de arestas *tree-edges* (ex. $C_3 = \{a_2, a_4\}$); um conjunto de nós superiores na hierarquia da árvore DFS, PP_j (pseudopais de um nó a_j), que estão diretamente conectados ao nó a_j através de arestas *back-edges* (ex. $PP_4 = a_1$); um conjunto de nós inferiores na hierarquia da árvore DFS, PC_j (pseudofilhos de a_j), que estão diretamente conectados

ao nó a_j através de arestas *back-edges* (ex. $PC_1 = a_4$); e finalmente de um conjunto de nós, denominado Sep_j , formado por ancestrais do nó a_j , que estão conectados a a_j ou a descendentes de a_j . Um nó a_j pode facilmente encontrar seu conjunto Sep_j considerando (a) a união dos conjuntos Sep_i recebidos de seus filhos, (b) seus pais e pseudopais, menos ele próprio. Isso é matematicamente definido como $Sep_j \doteq \bigcup_{a_k \in C_j} Sep_j \cup P_j \cup PP_i \setminus a_j$.

O algoritmo DPOP é dividido em 3 fases. Na primeira fase os agentes se organizam logicamente numa ST usando o DFS. Esta estrutura será utilizada nas duas fases seguintes denominadas de propagação de *UTIL* e de *VALUE*.

3.3.1 Fase 1 – propagação de *UTIL*

Na fase de propagação de *UTIL*, as mensagens de *UTIL* se propagam na direção da raiz da árvore. Os agentes folha enviam, para seus pais, as mensagens de *UTIL* somente através das arestas *tree-edges*. Um filho a_k de um nó a_j vai enviar um vetor da utilidade ótima $u_{a_k}^*(d_j^n)$. O vetor com apenas uma dimensão contém o custo da melhor solução que a_k (para todas as possíveis combinações de valores) pode obter dadas as possíveis escolhas em \mathbb{D} de um nó a_j . Estas mensagens dependem somente da subárvore roteada no respectivo nó (não existem arestas para outras partes da árvore) e das restrições entre o nó a_k e seu pai a_j . Considerando a Figura 2 como um exemplo, a mensagem de *UTIL* ($a_2 \rightarrow a_3$) depende apenas da variável de a_3 . Se o grafo contém ciclos, uma mensagem enviada de um nó para seu pai também pode depender de variáveis provenientes de nós superiores na árvore DFS. Isto pode ocorrer quando existe uma aresta *back-edge* conectando o nó emissor a_k com um nó a_b (pseudopai) cuja variável é d_b . Na Figura 2 a mensagem *UTIL* ($a_4 \rightarrow a_3$) necessita ser formada considerando esta situação, onde as utilidades que a subárvore roteada em a_4 podem assumir não dependem exclusivamente de seu pai a_3 . O agente a_4 está conectado com a_1 através da aresta *back-edge* $a_1 \rightarrow a_4$, então a_4 tem que levar em consideração esta dependência quando for enviar uma mensagem para a_3 .

Este é o momento onde a abordagem de programação dinâmica é considerada, e um nó a_k (ex. a_4) vai computar o valor ótimo que sua subtree pode determinar, para cada combinação de valor de seu pai P_j (ex. a_3) e seu pseudopai a_b (ex. a_1) (para cada combinação de valor da tupla $\langle a_3, a_1 \rangle$). Em seguida o nó vai montar uma mensagem *UTIL* denominada hiper-cubo de 2 dimensões, levando em consideração as possíveis escolhas de variáveis de seu pai P_j e também as possíveis escolhas de variáveis de seus pseudopais (uma para a variável em a_3 e outra para a_1 , enviando a mensagem para a_3 , seu pai, como representado na Tabela 1). Para os nós folhas, os mesmos imediatamente calculam estes valores inspecionando as restrições que os mesmos tem com seus vizinhos, ou seja eles iniciam

todo o processo. Apenas após ter recebido a mensagem *UTIL* de todos os seus filhos, um nó a_j pode computar os valores ótimos que podem ser alcançados por toda subárvore roteada através dele. Uma vez que todas as subárvores de a_j são disjuntas, somando seus resultados, é possível se determinar quanto de cada valor de seus filhos representa para toda a subarvore roteada em a_j . Desta forma a_j pode enviar sua mensagem *UTIL* ao seu pai P_j e armazenar seus valores ótimos, correspondendo aos valores recebidos de seus filhos C_j . Cada nó transmite sua mensagem conforme a fase de propagação de *UTIL* do algoritmo.

Tabela 1. Mensagem *UTIL* enviada de a_4 para a_3

$a_4 \rightarrow a_3$	$a_3 = d_3^1$	$a_3 = d_3^2$...	$a_3 = d_3^n$
$a_1 = d_1^1$	$u_{a_4}^*(d_1^1)$	$u_{a_4}^*(d_1^1)$...	$u_{a_4}^*(d_1^1)$
...
$a_1 = d_1^n$	$u_{a_4}^*(d_1^n)$	$u_{a_4}^*(d_1^n)$...	$u_{a_4}^*(d_1^n)$

3.3.2 Fase 2 – propagação de *VALUE*

Na fase de propagação de *VALUE* as mensagens são propagadas na direção dos nós folha da árvore DFS. Esta fase começa após o nó raiz ter recebido as mensagens *UTIL* de todos os seus filhos. O nó raiz em seguida escolhe o valor que leva a solução ótima e informa seus filhos, C_j , enviando uma mensagem de *VALUE* a todos eles. Após receber a mensagem de *VALUE* de seu pai P_j , cada nó é capaz de reconhecer o valor ótimo d_j^* para ele próprio, e passar este valor para seus filhos. É importante ressaltar que a mensagem $VALUE_{P_j}^{a_j}$ que um nó a_j recebe de seu pai P_j , também contém os valores de todas as variáveis que estavam presentes no contexto da mensagem *UTIL* de a_j para seu pai P_j . Conforme o exemplo da Figura 2, o nó raiz a_1 envia para seu filho a_3 a mensagem de $VALUE_1^3$ com ($a_1 \leftarrow d_1^*$), então a_3 envia para seu filho a_4 a mensagem de $VALUE_3^4$ com ($a_1 \leftarrow d_1^*$, $a_3 \leftarrow d_3^*$). Neste momento, o algoritmo é encerrado para os nós a_1 , a_3 e a_4 . De acordo com o algoritmo DPOP, o número de mensagens de *UTIL* e de *VALUE* está diretamente relacionado ao número de agentes, n , no problema. A fase de propagação de *UTIL* requer $n - 1$ mensagens dos nós folha para o nó raiz, e a fase de propagação de *VALUE* requer $n - 1$ mensagens do nó raiz para os nós folha.

3.4 Modelo de interferência

O modelo de interferência define o conjunto de conexões que podem interferir em qualquer outra conexão da rede sem fio. Devido à natureza de difusão do sinal de rádio, transmissões através de um enlace de comunicação (entre um par de nós da rede sem fio) podem interferir em transmissões que estejam ocorrendo em outros enlaces da rede. Dois enlaces

interferentes entre si podem proporcionar transmissões mal sucedidas se transmitirem simultaneamente no mesmo canal. O gerenciamento dos canais das redes sem fio é de extrema importância pois as interferências co-canal (*co-channel*) e de canais adjacentes (*adjacent-channel*) podem reduzir significativamente o desempenho da rede (Rappaport, 2002). A métrica de qualidade considerada para recepção do sinal sob interferência é a relação sinal-ruído+interferência, SINR (*signal-to-interference-plus-noise ratio*). No contexto das redes sem fio, o SINR percebido por um AP ou cliente é função de todas as transmissões simultâneas que ocorrem no mesmo canal e em canais adjacentes parcialmente sobrepostos. Considere um usuário u_k associado a um AP a_i , ($u_k \in \text{Clientes}(a_i)$), operando em um canal c_n . O SINR percebido por a_i é definido como $\gamma_{a_i}^{c_n} = \frac{S_{a_i \leftarrow u_k}}{N_{a_i} + I_{a_i}}$ onde $S_{a_i \leftarrow u_k}$ indica a potência média recebida do cliente u_k e N_{a_i} indica ruídos e interferências externas. A potência da interferência I_{a_i} é calculada por $I_{a_i} = I_{a_i}^a + I_{a_i}^u = \sum_{j \neq i} S_{a_i \leftarrow a_j} + \sum_{u_l \notin a_i} S_{a_i \leftarrow u_l}$, onde $I_{a_i}^a$ e $I_{a_i}^u$ indicam a potência da interferência proveniente de todas as transmissões de outros APs e clientes, respectivamente. Os nós interferentes podem estar operando no mesmo canal ou em canais parcialmente sobrepostos. O número de clientes associados a cada AP e o perfil de tráfego dos clientes também afetam o cálculo de I_{a_i} (Mishra et al., 2005; Chen et al., 2009). Considere uma rede de acesso sem fio composta por diversos APs, como arquitetura outdoor para o canal de retorno. Qualquer estratégia de atribuição de canais tenta minimizar a interferência entre áreas de cobertura de APs vizinhos. O valor do SINR depende basicamente do fator de sobreposição espectral entre canais interferentes. Portanto, o fator de sobreposição é adotado como métrica da função custo aplicada na solução do problema de alocação de canais usando DCOP. Seguindo a abordagem proposta em Briggs & Tijmes (2009); Monteiro et al. (2011), a sobreposição espectral entre canais adjacentes pode ser derivada matematicamente para redes que utilizam interfaces sem fio do padrão IEEE802.11b/g. O sinal modulado não filtrado possui densidade espectral de potência dada pela Equação 2:

$$\mathcal{S}(n, f) = \begin{cases} \left| \frac{\sin[2\pi\mathcal{X}(n, f)]}{2\pi\mathcal{X}(n, f)} \right|^2, & \text{for } \mathcal{X}(n, f) \neq 0 \\ 1, & \text{for } \mathcal{X}(n, f) = 0 \end{cases} \quad (2)$$

onde $\mathcal{X}(n, f) = \frac{f - 2412 - 5(n-1)}{BW}$, BW é a aproximação de largura de banda do canal segundo o critério *null-to-null* e n é o número do canal ($n = 1, 2, \dots, 11$). As interfaces WLAN empregam um filtro de IF em ambos os caminhos, transmissão e recepção, a fim de reduzirem a potência interferente de bandas adjacentes. Este filtro possui largura de banda (3dB bandwidth) de 17MHz e atenuação mínima de 50dB na banda de corte (22MHz). A resposta em frequência aproximada deste filtro de IF é dada por:

$$\mathcal{F}(n, f) = \frac{1}{[1 + 2.6 \mathcal{X}(n, f)]^6}. \quad (3)$$

A partir das equações (2) e (3) podemos calcular o fator de sobreposição entre os canais de número n e m como:

$$\mathcal{SO}(n, m) = \int_{2200}^{2700} \mathcal{S}(n, f) \cdot \mathcal{F}(n, f) \cdot \mathcal{S}(m, f) \cdot \mathcal{F}(m, f) df. \quad (4)$$

Os valores normalizados do fator de sobreposição $\mathcal{SO}(n, m)$ são apresentados na Tabela 2 como função do espaçamento entre canais $|n - m|$. Os valores selecionados d_j de um agente a_j representam o número do canal selecionado de um AP_j , logo o domínio de d_j é representado por $\mathbb{D}_j = \{1, 2, \dots, 11\}$. Para um par de agentes conectados (par de APs interferentes) no grafo de restrições, a função custo para um par de valores (d_i, d_j) é definida como $f_{ij}(x, y) : \mathbb{D}_i \times \mathbb{D}_j \rightarrow \mathbb{R}$ onde os valores em \mathbb{R} representam os fatores de interferência. É importante ressaltar que os resultados numéricos apresentados na Tabela 2 podem variar para diferentes configurações de *chipset* e filtros das interfaces IEEE802.11b/g.

Os APs devem se coordenar para encontrar um conjunto de valores de canais alocados, D^* , que minimize a função custo global e então reduza a quantidade total de interferência na rede. O inverso da função custo representa o ganho ou a função utilidade (UTIL) para os canais alocados entre dois nós no grafo de restrições.

Tabela 2. Fator de interferência normalizado.

Espaçamento entre Canais	Fator de Interferência
0	1
1	0,7272
2	0,2714
3	0,0375
4	0,0054
5	0,0008
6	0,0002
7 – 10	0

4. Algoritmos Propostos

Os algoritmos desenvolvidos com o objetivo de otimizar a capacidade das redes IEEE 802.11 b/g, tratando o problema de alocação de canais como um DCOP, foram primeiramente derivados do algoritmo Simple-ADOPT e denominados de DCAA (*Distributed Channel Assignment Algorithms*).

Também foram propostos algoritmos baseados no DPOP, denominados de DOCA (*Distributed Optimal Channel Assignment*) e DSCA (*Distributed SubOptimal Channel Assignment*).

4.1 Algoritmos DCAA

Os algoritmos DCAA foram desenvolvidos em duas versões. A primeira versão apresentada através do algoritmo DCAA-O (*Distributed Channel Assignment Algorithm - Optimal*) e a segunda versão através do algoritmo DCAA-S (*Distributed Channel Assignment Algorithm - SubOptimal*). Ambos são executados de forma síncrona e em paralelo em todos os APs, buscando respectivamente, de maneira cooperativa, a solução ótima e subótima para o problema de atribuição de canais. O objetivo destes algoritmos não é o de apenas obter a melhor atribuição de canal, mas também minimizar o número de mensagens de controle trocadas entre APs.

O primeiro algoritmo desenvolvido, DCAA-O, garante a convergência através da atribuição ótima de canais. Este algoritmo também possui duas versões: a primeira versão proposta em Monteiro et al. (2011), opera de maneira próxima de um método exaustivo, com critérios de poda para valores em \mathbb{D}_j , apenas no AP raiz e nos APs folha da árvore de restrições. A segunda versão proposta em Monteiro et al. (2012a), opera de maneira otimizada e considera em seu procedimento de poda apenas valores em \mathbb{D}_j que apontem para uma solução de menor custo local, sem perder a otimalidade do resultado. O segundo algoritmo, DCAA-S, é baseado em uma heurística (Monteiro et al., 2012a), que é uma versão modificada do algoritmo DCAA-O, onde a heurística considerada contribui para a eficiência ao invés da otimalidade. Como consequência, o número de mensagens trocadas é significativamente reduzido, ao custo de uma alocação de canal sub-ótima. Os algoritmos DCAA-O e DCAA-S empregam a mesma notação básica para nós pais e nós filhos adotadas pelo algoritmo ADOPT, bem como a definição dos grafos de VALUE e de VIEW.

A escolha do valor d_j é um ponto importante da execução dos algoritmos. Um dado AP a_j irá executar iterações com sua subárvore, através de seus filhos de VALUE e de VIEW, para diferentes escolhas temporários $d_j \in \mathbb{D}_j$. Durante a iteração, o AP a_j constrói suas variáveis locais baseado em restrições existentes com seus vizinhos. Após testar os valores $d_j \in \mathbb{D}_j$, o AP a_j é capaz de identificar, em seu contexto, o valor d_j que minimiza o custo da subárvore, representado por $z_j(d_j)$. A partir deste momento o AP a_j envia uma mensagem de VIEW para seu pai $a_i \in P_j^{view}$, que poderá continuar com o procedimento para escolha de seu canal local d_i . Na primeira versão do algoritmo DCAA-O, cada AP é capaz de identificar o valor d_j , somente após testar todos os valores $d_j \in \mathbb{D}_j$. Na segunda versão, cada nó passa a considerar apenas os valores d_j que são candidatos a minimizar o custo da subárvore no contexto.

Esta versão permite que cada AP a_j , exceto o AP raiz e os APs folha, após receber as mensagens de VIEW de seus filhos C_j^{view} , exclua valores não recomendados de $d_j \in \mathbb{D}_j$ para cálculo de seu custo mínimo. Esta escolha se baseia nos valores de custo local, $lc_j(d_j)$ do AP a_j com seus pais no grafo de VALUE e valores do custo de sua subárvore, $sc_j(d_j)$, tendo como base informações obtidas durante iterações anteriores entre APs pais e filhos. Os valores d_j excluídos do domínio \mathbb{D}_j serão aqueles que possam resultar em valores estimados locais, $z_j(d_j)$, superiores ao valor mínimo computado e armazenado em $z_{ja}(d_j)$. Elimina-se desta forma processamento desnecessário e a quantidade de mensagens trocadas entre os APs.

O algoritmos DCAA-O e DCAA-S podem iniciar o processo de realocação de canal, em uma rede composta de múltiplos APs, quando um AP identifica interferências co-canais ou de canais adjacentes. Em uma etapa preliminar, um AP a_j inicia um algoritmo distribuído para formação de uma *spanning tree* (Cheung, 1983), o que permite a definição dos grafos de VALUE e de VIEW em todos os nós. Após a execução do algoritmo, o nó raiz envia uma mensagem identificada como *TERMINATE*, contendo a solução de alocação ótima, para todos os seus filhos C_1^{view} , através do grafo de VIEW. O AP, a_j , que inicia a execução do algoritmo é identificado como o AP raiz, a_1 , que deve escolher um novo valor d_1 para sua variável local e enviar este valor, via uma mensagem de VALUE para todos os descendentes a ele ligados em P_1^{value} . Considerou-se como o primeiro valor assumido pelo AP raiz, o menor valor $d_1 \in \mathbb{D}_1$. O AP raiz espera sincronamente pelas mensagens de VIEW, de seus filhos de VIEW em C_1^{view} , para então computar suas variáveis locais. É importante chamar atenção para a soma dos custos $z_k(d_k)$ recebidos de todos os filhos. Após ter recebido as mensagens de VIEW de seus filhos, o valor $sc_1(d_1)$ deve ser checado. Se $sc_1(d_1) \neq 0$ então o AP raiz envia o próximo valor d_1 para seus descendentes em C_1^{view} e repete o procedimento. Este procedimento se repetirá para todos os valores em \mathbb{D}_1 ou até que a_1 receba $sc_1(d_1) = 0$ de seus filhos de VIEW. Quando o AP raiz tiver checado todas as possibilidades d_1 com seus descendentes de VALUE, o nó seleciona o menor valor custo recebido de suas subárvores $sc_1(d_1)$ e envia uma mensagem *TERMINATE* contendo a solução de alocação ótima, para todos os APs em C_1^{view} através do grafo de VIEW. A partir desta etapa, cada AP a_j toma conhecimento de seu valor d_j que minimiza a função custo global, e envia uma mensagem *TERMINATE* para todos os APs em C_j^{view} , finalizando a execução do algoritmo.

O algoritmo DCAA-S teve seu desenvolvimento baseado no algoritmo DCAA-O (versão otimizada), e considera que cada AP a_j , exceto o AP raiz, somente selecione valores d_j em \mathbb{D}_j que resultem no menor valor $lc_j(d_j)$, a quantidade de interferência em a_j proveniente de seus pais em P_j^{value} . Esta estratégia, simples, modifica o domínio \mathbb{D}_j para um domínio reduzido \mathbb{H}_j . O domínio reduzido \mathbb{H}_j é obtido por cada AP, através da execução

de um procedimento adicional conforme [Monteiro et al. \(2012a\)](#). O AP a_j deve considerar o domínio redefinido \mathbb{H}_j para calcular seu custo local com suas subárvores, $sc_j(d_j)$. O algoritmo DCAA-S atinge um estado estável sem penalizar a performance da rede em termos de mensagens trocadas, obtém um valor de custo global em geral próximo da solução ótima (podendo apresentar uma solução ótima). Durante as iterações, o algoritmo deverá considerar todo o novo domínio \mathbb{H}_j , mas poderá ter sua execução interrompida localmente, no AP a_j , quando o valor de custo com a subárvore $sc_j(h_j) = 0$. Dado que no domínio \mathbb{H}_j só existem valores h_j que proporcionam o menor custo local, $lc_j(h_j)$, do AP a_j com seus pais de VALOR P_j^{value} , estes valores, $lc_j(h_j)$, se somados a um valor de custo com a subárvore $sc_j(h_j) = 0$ gerarão valores mínimos idênticos. Portanto, não faz sentido o prosseguimento da execução do algoritmo para os outros valores de $h_j \in \mathbb{H}_j$.

Os algoritmos DCAA-O e DCAA-S foram implementados e simulados de forma distribuída, utilizando-se uma biblioteca para simulação de eventos discretos denominada *SIMPATICA* ([Maziero, 2009](#)), baseada no paradigma atores/mensagens. Conforme este paradigma, um modelo de simulação é composto por um conjunto de atores (tarefas) que se comunicam entre si através de mensagens. Esta biblioteca proporcionou a simulação de um ambiente distribuído síncrono através de três tipos de entidades: tarefa, fila e mensagem. Cada AP é implementado como uma tarefa e possui três filas para recebimento de mensagens, identificadas como *VIEW_QUEUE*, *VALUE_QUEUE*, e *TERMINATE_QUEUE*, respectivamente. Os pseudocódigos dos algoritmos DCAA-O e DCAA-S foram omitidos por questões de brevidade, maiores detalhes podem ser obtidos em [Monteiro et al. \(2011, 2012a\)](#).

4.2 Algoritmos DOCA e DSCA

Com o objetivo de melhorar o desempenho das estratégias baseadas no algoritmo ADOPT, apresenta-se uma nova estratégia distribuída e cooperativa entre APs, baseada no algoritmo DPOP ([Petcu & Faltings, 2005](#); [Petcu, 2007](#)). O objetivo é o de reduzir o número de mensagens de controle trocadas entre APs, assim como avançar em termos de escalabilidade para a solução do problema de alocação de canais para redes IEEE 802.11 b/g. Os algoritmos são denominados DOCA (*Distributed Optimal Channel Assignment*) ([Monteiro et al., 2012b,c](#)) e DSCA (*Distributed SubOptimal Channel Assignment*). Como a estratégia DPOP, os algoritmos DOCA e DSCA são baseados em programação dinâmica. Operam sobre a ordenação de suas variáveis, dada através de uma árvore DFS, construída a partir do grafo de restrições. Uma das vantagens ao se utilizar como estratégia o algoritmo DPOP, é que os algoritmos DOCA e DSCA requerem um número linear de troca de mensagens, introduzindo menor sobrecarga na rede quando aplicados a

uma solução distribuída. Entretanto, a complexidade, que não deve deixar de ser considerada, está relacionada ao tamanho das mensagens *UTIL* que são delimitadas exponencialmente pela largura induzida, criada pelo ordenamento DFS escolhido. Os algoritmos empregam a mesma notação básica para nós pais e nós filhos adotadas pelo algoritmo DPOP, bem como a execução das três fases do algoritmo: criação da árvore DFS, propagação de *UTIL* e propagação de *VALUE*.

Para se aplicar o algoritmo DPOP como um método para atribuição de canais em redes sem fio, um grafo de restrições $\mathbf{G} = (V, E)$ é construído a partir de uma topologia física da rede. O conjunto de vértices V do grafo de restrições \mathbf{G} representa os APs, e o conjunto de arestas E representa a conectividade entre agentes interferentes, significando que seus respectivos canais alocados irão afetar o desempenho de cada um devido a interferência co-canal ou de canal-adjacente. Utilizando-se o indicador de intensidade de sinal recebido RSSI (*Received Signal Strength Intensity*), um AP pode medir a força de um sinal recebido e identificar a interferência com seus vizinhos. Os cenários considerados se referem a uma ou mais redes sem fio, compostas por APs, arbitrariamente distribuídas sobre uma área. Assume-se que cada AP emprega uma interface de rádio (IEEE802.11b, 2000) para troca de mensagens pelos algoritmos. Estas interfaces de controle dedicadas operam em um mesmo canal pré-definido c , que é previamente selecionado entre um dos 11 canais disponíveis nos padrões IEEE.802.11b/g. Numa primeira abordagem assumiu-se que este canal não pertence a nenhum domínio ($c \notin \mathbb{D}$). Mensagens *UTIL* e *VALUE* são consideradas como de controle que propagam informações na direção do nó raiz e na direção dos nós folha, respectivamente.

O algoritmo DOCA garante a convergência através da atribuição ótima de canais, contudo sua complexidade é exponencial para redes com maior número de APs e maior grau de conectividade entre APs, inviabilizando sua aplicação para redes mais complexas. O algoritmo DSCA foi proposto a partir do comprometimento da escalabilidade apresentada pelo algoritmo DOCA. O algoritmo DSCA teve seu desenvolvimento baseado no algoritmo DOCA, tendo como premissa garantir a linearidade do tamanho das mensagens trocadas entre APs. Esta linearidade é garantida através de critérios que limitam a dimensão do hiper-cubo a ser enviado por um nó a_j a seu pai P_j , através da propagação de *UTIL*. Os critérios adotados foram: adotar um tamanho máximo (parametrizável) para o hiper-cubo e incluir no hiper-cubo apenas combinações que proporcionem custos locais inferiores ou iguais a um limite de custo local pré-estabelecido (também parametrizável). As combinações que não forem incluídas no hiper-cubo ficam localmente armazenadas em cada AP para que sejam, caso necessário, posteriormente utilizadas no decorrer da execução das fases de propagação de *UTIL* ou de *VALUE*. Como consequência, o tamanho das mensagens trocadas se

mantém linear evitando-se então uma complexidade exponencial, ao custo de uma alocação de canal subótima.

5. Resultados

5.1 Cenários avaliados

As simulações foram divididas em duas etapas. Na primeira etapa avaliou-se o desempenho dos algoritmos DCAA-O e DCAA-S. Foram consideradas topologias (grafos de restrições) com $|V| = 9$, $|V| = 16$ e $|V| = 25$ nós. Adicionalmente para cada topologia de rede foram consideradas densidades de $\{1, 2, 3, 4\}$ para topologias de 9 nós, e $\{1, 2, 3, 4, 5\}$ para 16 e 25 nós. Um grafo com densidade (*link density*) l_d tem $(l_d \cdot N)$ ligações (*links*), onde N é o número de nós no grafo e $l_d \leq (N - 1)/2$ (Modi et al., 2003a). Para cada grupo de nós, e para cada densidade, foram geradas 100 topologias randômicas. Como primeiro cenário de estudo, permitiu-se que cada AP selecionasse o canal correspondente a partir de um conjunto de apenas 3 canais: $\{1, 6, 11\}$. No segundo cenário os APs puderam selecionar um entre 11 canais disponíveis. Na segunda etapa avaliou-se o desempenho dos algoritmos LO-A, DCAA-S e DOCA através de topologias com $|V| = 9$ e $|V| = 10$. Também foram realizadas simulações sobre um conjunto de 100 topologias. Cada AP possuindo grau médio $AD = 4$, $AD = 6$ e $AD = 8$. Os APs puderam selecionar um dentre os 11 canais disponíveis.

5.2 Desempenho

As métricas de comparação adotadas para análise dos resultados foram o número e o tamanho de mensagens trocadas para atingir a convergência e os respectivos valores da função custo global. O valor final da função custo global reflete a degradação global devido a interferências e consequente degradação proporcional da capacidade da rede (Monteiro et al., 2012b,c). A principal vantagem em se utilizar os resultados de uma função custo na avaliação dos algoritmos, é que esta estratégia permite que indiretamente se infira no desempenho da rede sem a necessidade de extensivas simulações no nível de pacote. Como os resultados das simulações representam valores médios das topologias, foram incluídos nos gráficos o intervalo de confiança de 90% para os resultados.

O algoritmo LO-A foi usado como algoritmo de referência e implementado como em Chen et al. (2009). LO-A é um algoritmo distribuído simples e escalável, que apresenta equilíbrio entre qualidade da solução e o número de mensagens trocadas entre os APs. No entanto, em Chen et al. (2009), um AP precisa localmente se coordenar com outros APs através de uma rede cabeada (*backbone network*), a fim de efetuar a troca de canais. Para fins de comparação, as mensagens do algoritmo LO-A são trocadas entre APs via um canal de controle pré-definido, como para o algoritmos DCAA, DOCA e DSCA. O valor computado, como ponto

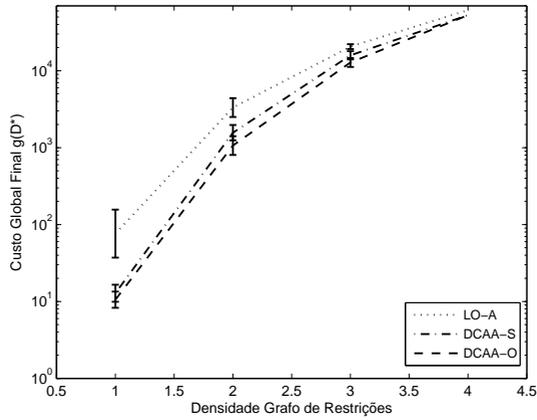
de convergência do algoritmo LO-A, representa um número mínimo de iterações em que passa a não ocorrer nenhuma melhora ou nova alocação de canal entre os APs. Considerou-se o limite inferior de 50 iterações para LO-A. Este valor foi determinado através de simulação como um critério adequado para o encerramento do algoritmo. As simulações do algoritmo DOCA foram realizadas usando-se a plataforma FRODO (Léauté et al., 2009) integrada com o software MATLAB®. FRODO é uma ferramenta, *open-source*, multitarefa para otimização distribuída. As topologias de rede sem fio e os grafos de restrições correspondentes foram gerados através do MATLAB.

O desempenho dos algoritmos DCAA-O, DCAA-S e LO-A em termos de valor de custo global final para topologias formadas por 9 APs, com 11 canais disponíveis e densidades de 1 a 4, é apresentado na Figura 3 (a). Como pode-se observar através desta figura, o algoritmo DCAA-O atinge o menor custo global final, significando ter atingido a melhor solução global. O desempenho do algoritmo DCAA-S se apresenta muito próximo do desempenho do algoritmo DCAA-O, enquanto o algoritmo LO-A apresenta o pior desempenho entre os três algoritmos. A diferença de desempenho entre os algoritmos diminui com o aumento da densidade (ou da conectividade da rede).

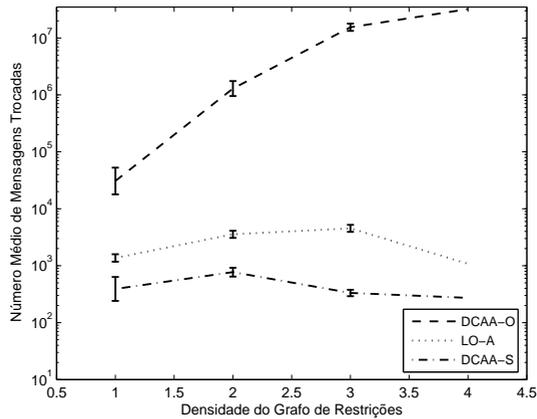
A Figura 3(b) considera o mesmo cenário, entretanto a comparação se dá em termos do número médio de mensagens trocadas entre APs. O algoritmo DCAA-O exige um número exponencial de mensagens trocadas, enquanto o algoritmo DCAA-S exige uma quantidade bem inferior. Esta diferença é representada por algumas ordens de grandeza. O algoritmo DCAA-S também demanda menos troca de mensagens do que o algoritmo LO-A. É importante ressaltar que o algoritmo DCAA-S, claramente, demanda o menor número de mensagens trocadas e também atinge um custo global final, muito próximo do resultado ótimo apresentado pelo algoritmo DCAA-O. Portanto, esta redução na troca de mensagens não acarreta uma diminuição de desempenho que possa ser considerada como relevante. Para maiores detalhes sobre os testes e avaliação de desempenho dos algoritmos DCAA-O, DCAA-S e LO-A, deve-se consultar Monteiro et al. (2011, 2012a).

A Tabela 3 apresenta a média final do custo global para os algoritmos DOCA, DCAA-S e LO-A. É possível observar que os algoritmos LO-A e DCAA-S convergem algumas vezes para um valor mínimo local. Apenas o algoritmo DOCA converge para a solução ótima em 100% dos casos. As soluções ótimas foram checadas através de pesquisa exaustiva. O algoritmo DCAA-S mostrou bom desempenho em termos de convergência, quando comparado ao LO-A, mas nem sempre converge para a solução ótima.

A diferença entre os valores do custo global final apresentados pelos algoritmos LO-A e DCAA-S, em relação ao algoritmo DOCA, é maior para grafos mais conectados (9 e 10 APs com $AD=6$ e $AD=8$). Nestes



(a) Custo global final.



(b) Média de mensagens trocadas.

Figura 3. Métricas DCAA-O, LO-A e DCAA-S (9 APs, 11 canais).

casos a complexidade do problema de alocação de canais aumenta. Deve-se chamar a atenção para o fato de que ao apresentar a solução ótima, o algoritmo DOCA apresenta um número de mensagens cerca de uma ordem de magnitude menor que os algoritmos LO-A e DCAA-S.

A Figura 4 apresenta o número de mensagens trocadas, pelo algoritmo DOCA, para diferentes tamanhos e graus de topologias. Deve-se notar que,

Tabela 3. Média do custo global final.

	9 APs			10 APs		
	AD = 4	AD = 6	AD = 8	AD = 4	AD = 6	AD = 8
LO-A	0,339	2,074	6,166	0,335	1,956	5,967
DCAA-S	0,199	1,510	5,550	0,179	1,407	4,345
DOCA	0,038	1,282	5,291	0,039	0,369	3,408
Optimal	0,038	1,282	5,291	0,039	0,369	3,408

para mensagens de controle, o algoritmo apresenta resultados linearmente escalares em função do número de APs, não dependendo do grau médio das topologias. Entretanto, para a construção das árvores DFS, o algoritmo não produz o mesmo número de mensagens, mostrando que esta fase depende do número de arestas, e que por sua vez está relacionado com o número de nós e o grau médio da topologia.

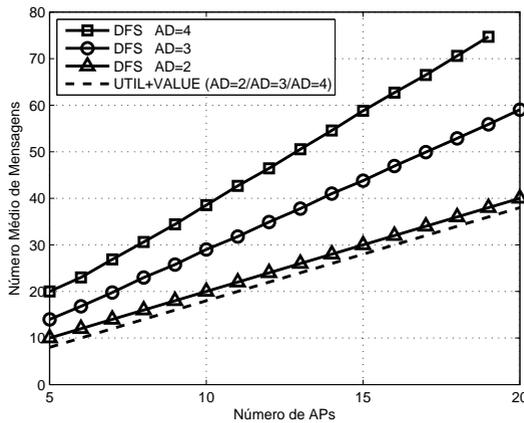
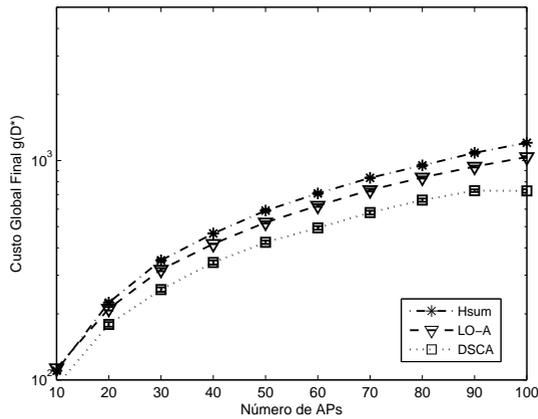


Figura 4. Número de mensagens trocadas – DOCA.

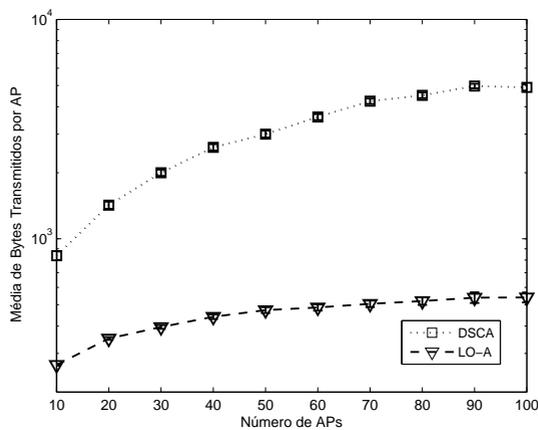
Conforme citado na Seção 4 o desenvolvimento do algoritmo DSCA deu-se devido à necessidade de aumentar a escalabilidade do algoritmo DOCA. Neste sentido a Figura 5 apresenta resultados para topologias mais complexas, até 100 APs e $AD = 5$. Na Figura 5(a) pode-se comparar os resultados da média do custo global final em relação aos algoritmos LO-A e Hsum (Mishra et al., 2005), este último foi considerado como referência por utilizar cenários realistas de interferência e de reutilização de canal, além de alcançar redução de interferência para topologias esparsas e densas. Nota-se que o algoritmo DSCA além de ter escalabilidade aumentada, apresentou melhores resultados mesmo para redes mais complexas. Em relação ao

algoritmo LO-A apresentou melhores resultados em cerca de 30%, e em relação ao algoritmo Hsum em cerca de 40%.

Através da Figura 5(b) observa-se que mesmo ao tratar redes mais complexas, o algoritmo DSCA se manteve linear. Para o caso de 100 nós e $AD = 5$ são requeridos aproximadamente 5Kbytes. O resultado é ainda escalável em termos de implementação do protocolo, uma vez que o tamanho das mensagens é inferior ao máximo MTU da unidade de dados da camada MAC. Para os outros casos o número de bytes é negligenciável.



(a) Custo global final.



(b) Bytes transmitidos por AP.

Figura 5. Métricas LO-A e DSCA.

6. Comentários Finais

A solução para o canal de retorno da TV digital interativa, baseada nas redes sem fio padrão 802.11b/g, possui recursos suficientes para disponibilizar a inclusão digital além de ser uma tecnologia de comunicação que se destaca por sua capacidade de penetração. Tal solução se consolida como opção tecnológica capaz de atingir as mais variadas parcelas da população brasileira.

Este trabalho utiliza a combinação de métodos de pesquisa em IAD (*Inteligência Artificial Distribuída*) com otimização de problemas em redes sem fio. Trata a alocação de canais em redes sem fio padrão 802.11b/g como um *Distributed Constraint Optimization Problem* (DCOP) onde os APs podem pertencer a diferentes domínios administrativos. O primeiro algoritmo proposto (DCAA-O) alcança a solução ótima para o problema de atribuição de canais, portanto, reduz a interferência de canal adjacente e maximiza o desempenho da rede. Entretanto, o número médio de mensagens trocadas entre APs se mostra excessivo para alguns cenários específicos. Para reduzir a quantidade de mensagens trocadas pelo algoritmo DCAA-O, foi proposto um algoritmo baseado em uma heurística (DCAA-S), que atinge uma solução subótima para o problema da atribuição de canais e reduz consideravelmente o número de mensagens trocadas. Esta redução na troca de mensagens não acarreta uma diminuição de desempenho que possa ser considerada como relevante. Na sequência foram desenvolvidos os algoritmos DOCA e DSCA com o objetivo de reduzir o número de mensagens de controle encontradas nos algoritmos anteriores. O algoritmo DOCA alcança a solução ótima mas apresenta problemas de escalabilidade devido ao tamanho das mensagens trocadas entre APs. A heurística desenvolvida através do algoritmo DSCA estabelece uma linearidade quanto ao tamanho das mensagens trocadas, garante escalabilidade da solução e resolve eficientemente o problema de alocação de canais. Em relação ao número de mensagens, os algoritmos apresentam resultados linearmente escaláveis em função do número de APs, não dependendo do grau médio das topologias.

Agradecimentos

Os autores agradecem à CAPES e ao CNPq pelo apoio financeiro.

Referências

- Akyildiz, I.F. & Wang, X., A survey on wireless mesh networks. *IEEE Communications Magazine*, 43(9):S23–S30, 2005.
- ANATEL, , GT PowerLine – power line communication ou broadband over power lines. 2003. Disponível em <http://www.anatel.gov.br/>.

- Bicket, J.; Aguayo, D.; Biswas, S. & Morris, R., Architecture and evaluation of an unplanned 802.11b mesh network. 2005. Disponível em <http://pdos.csail.mit.edu/roofnet/>.
- Briggs, K. & Tijmes, M., Optimal channel allocation for wireless cities. In: *Proc. of the IEEE 69th Vehicular Technology Conference*. p. 1–5, 2009.
- Cao, L. & Zeng, H., Distributed spectrum allocation via local bargaining. In: *Proc. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. p. 475–486, 2005.
- Chechetka, A. & Sycara, K., No-commitment branch and bound search for distributed constraint optimization. In: *Proc. of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*. New York, USA: ACM, p. 1427–1429, 2006.
- Chen, J.K.; de Veciana, G. & Rappaport, T.S., Site-specific knowledge and interference measurement for improving frequency allocations in wireless networks. *IEEE Transactions on Vehicular Technology*, 58(5):2366–2377, 2009.
- Cheung, T.Y., Graph traversal techniques and the maximum flow problem in distributed computation. *IEEE Transactions on Software Engineering*, 9(4):504–512, 1983.
- Forum PLC, , PLC forum. 2006. Disponível em <http://www.plcforum.org/>.
- Gupta, P.P. & Kumar, P.R., The capacity of wireless networks. *IEEE Transactions on Information Theory*, 46:388–404, 2000.
- Hirayama, K. & Yokoo, M., Distributed partial constraint satisfaction problem. In: *Principles and Practice of Constraint Programming*. p. 222–236, 1997.
- Huang, J.; Berry, R.A. & Honig, M.L., Auction-based spectrum sharing. *Journal of Mobile Network Applications*, 11(3):405–418, 2006.
- IEEE802.11b, , Supplement to IEEE Standard for Information Technology-Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area Networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band. *IEEE Std 80211b-1999*, 2000.
- IEEE802.11g, , Draft IEEE Standard for Information Technology Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Amendment 6: Medium Access Control (MAC) Security Enhancements (Amendment to IEEE Std 802.11, 1999 Edition as amended by IEEE Std 802.11g-2003 and IEEE Std 802.11h-2003). *IEEE Std P80211i/D100*, 2004.

- IEEE802.11std, , IEEE Standard for Information Technology Telecommunications and Information Exchange Between Systems - Local and Metropolitan Area networks - Specific Requirements - Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. *IEEE Std 80211-2012 (Revision of IEEE Std 80211-2007)*, 2012.
- Kyasanur, P. & Vaidya, N.H., Routing in multi-channel multi-interface ad hoc wireless networks. Technical Report, University of Illinois at Urbana-Champaign, 2004. 12 p.
- Léauté, T.; Ottens, B. & Szymanek, R., FRODO 2.0: An open-source framework for distributed constraint optimization. In: Hirayama, K.; Yeoh, W. & Zivan, R. (Eds.), *Proc. of the IJCAI'09 - Distributed Constraint Reasoning Workshop*. Pasadena, USA, p. 160–164, 2009, disponível em <http://frodo2.sourceforge.net>.
- Lemos, A., Anjos interativos e retribalização do mundo. Sobre interatividade e interfaces digitais. 1997. <http://www.facom.ufba.br/ciberpesquisa/lemos/interativo.pdf>.
- Mailler, R. & Lesser, V., Solving distributed constraint optimization problems using cooperative mediation. In: *Proc. of the Third International Joint Conference on Autonomous Agents and Multiagent Systems*. Washington, USA: IEEE Computer Society, v. 1, p. 438–445, 2004.
- Maziero, C.A., Library SIMPATICA. 2009. Disponível em <http://www.ppgia.pucpr.br/~maziero/doku.php/software:simulation>.
- Meisels, A., *Distributed Search by Constrained Agents*. London, UK: Springer-Verlag, 2008.
- Mishra, A.; Banerjee, S. & Arbaugh, W., Weighted coloring based channel assignment for WLANs. *ACM SIGMOBILE Mobile Computing and Communications Review*, 9(3):19–31, 2005.
- Modi, P.J.; Shen, W.M.; Tambe, M. & Yokoo, M., An asynchronous complete method for distributed constraint optimization. In: *Proc. of the Second International Joint conference on Agents and Multiagent Systems - AAMAS*. p. 161–168, 2003a.
- Modi, P.J.; Shen, W.M.; Tambe, M. & Yokoo, M., Solving distributed constraint optimization problems optimally, efficiently an asynchronously. 2003b. Extension of paper: An Asynchronous, Complete Method for General Distributed Constraint Optimization.
- Modi, P.J.; Shen, W.M.; Tambe, M. & Yokoo, M., Adopt: asynchronous distributed constraint optimization with quality guarantees. *Artificial Intelligence*, 161(1-2):149–180, 2005.

- Monteiro, T.L.; Pellenz, M.E.; Penna, M.C.; Enembreck, F. & Souza, R.D., On optimal distributed channel allocation for access points in WLANs. In: Casares-Giner, V.; Manzoni, P. & Pont, A. (Eds.), *Proc. of International IFIP TC 6 Workshops*. Heidelberg: Springer-Verlag, v. 6827 de *Lecture Notes in Computer Science*, p. 73–84, 2011.
- Monteiro, T.L.; Pellenz, M.E.; Penna, M.C.; Enembreck, F.; Souza, R.D. & Pujolle, G., Channel allocation algorithms for WLANs using distributed optimization. *AEU - International Journal of Electronics and Communications*, 66(6):480–490, 2012a.
- Monteiro, T.L.; Pujolle, G.; Pellenz, M.E.; Penna, M.C. & Souza, R.D., A multi-agent approach to optimal channel assignment in w lans. In: *Proc. of IEEE Wireless Communications and Networking Conference*. p. 2637–2642, 2012b.
- Monteiro, T.L.; Pujolle, G.; Pellenz, M.E.; Penna, M.C. & Souza, R.D., An optimal channel assignment strategy for WLANs using distributed optimization. In: *Proc. of IEEE Network Operations and Management Symposium*. p. 286–292, 2012c.
- Montez, C. & Becker, V., *TV digital interativa: conceitos, desafios e perspectivas para o Brasil*. Florianópolis, SC: Editora da UFSC, 2005.
- Nie, N. & Comaniciu, C., Adaptive channel allocation spectrum etiquette for cognitive radio networks. In: *Proc. First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks*. p. 269–278, 2005.
- de Oliveira, E.C.R. & Albuquerque, C.V.N., TV digital interativa: Padrões para uma nova era. In: *Anais da V Escola Regional de Informática RJ/ES*. Sociedade Brasileira de Computação, p. 1–27, 2005.
- Peng, C.; Zheng, H. & Zhao, B.Y., Utilization and fairness in spectrum assignment for opportunistic spectrum access. *Journal of Mobile Network Applications*, 11(4):555–576, 2006.
- Petcu, A., *A Class of Algorithms for Distributed Constraint Optimization*. PhD. Thesis No. 3942, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, 2007.
- Petcu, A. & Faltings, B., A scalable method for multiagent constraint optimization. In: *Proc. of the 19th International Joint Conference on Artificial Intelligence*. San Francisco, USA: Morgan Kaufmann, p. 266–271, 2005.
- Ramachandran, K.; Belding, E.M.; Almeroth, K.C. & Buddhikot, M.M., Interference-aware channel assignment in multi-radio wireless mesh networks. In: *Proc. of 25th IEEE International Conference on Computer Communications*. p. 1–12, 2006.
- Rappaport, T.S., *Wireless Communications: Principles and Practice*. 2a edição. Upper Saddle River, USA: Prentice Hall, 2002.

- Reisman, R.R., Rethinking Interactive TV – I want my Coactive TV. 2002.
<http://www.teleshuttle.com/cotv/CoTVIntroWtPaper.htm>.
- Rossi, F.; van Beek, P. & Walsh, T., *Handbook of Constraint Programming*.
Amsterdam, The Netherlands: Elsevier Science, 2006.
- Zheng, H., Collaboration and fairness in opportunistic spectrum access. In:
Proc. IEEE International Conference on Communications. p. 3132–
3136, 2005.

Métodos Objetivos para Avaliação de Qualidade de Vídeo Digital

Wyllian Bezerra da Silva, Keiko Verônica Ono Fonseca e
Alexandre de Almeida Prado Pohl*

Resumo: Sinais de vídeo digital estão sujeitos a inúmeras degradações devido a processos de compressão, transmissão em canais ruidosos e processamento durante sua decodificação. Em consequência, a avaliação da qualidade de vídeo tornou-se uma necessidade para os radiodifusores e provedores de conteúdo interessados em oferecer um sinal de alta qualidade aos usuários. Este capítulo apresenta uma introdução às técnicas de avaliação de vídeo digital ao mesmo tempo em que apresenta duas novas técnicas de avaliação objetiva: uma técnica de referência reduzida baseada na diferença de atividade dos coeficientes DCT e uma técnica sem referência baseada em um modelo sigmoidal que incorpora características espaço-temporais relacionadas à componente de luminância do sinal de vídeo.

Palavras-chave: Avaliação de qualidade de vídeo digital, Métrica referência reduzida, Métrica sem referência.

Abstract: *Digital video signals are subject to several degradations due to compression processes, transmission over noisy channels and processing during the decoding. Therefore, the video quality assessment has become a necessity for broadcasters and content providers interested in offering a high quality signal to the customers. This chapter presents an introduction to the techniques of digital video quality assessment and describe two new objective assessment techniques: a reduced-reference technique based on the activity-difference of the DCT coefficients and a no-reference technique based on a sigmoidal model that incorporates spatio-temporal features related to the luminance component of the video signal.*

Keywords: *Digital video quality assessment, Reduced reference metric, No-reference metric*

*E-mail: pohl@utfpr.edu.br

1. Introdução

O avanço da tecnologia voltada para a produção e distribuição de conteúdo visual tem sido expressivo durante os últimos anos. Neste contexto, merece destaque o vídeo digital, transmitido pela Internet ou via radiodifusão para os mais diversos tipos de dispositivos. Contudo, vídeos digitais estão sujeitos a inúmeras degradações e distorções devido a processos de compressão, transmissão em canais ruidosos e processamento durante sua decodificação. Assim, a avaliação da qualidade de vídeo tornou-se uma necessidade, cujo resultado tem por objetivo auxiliar o radiodifusor ou o provedor de vídeo a melhorar a qualidade do conteúdo visual transmitido.

O estudo e desenvolvimento de técnicas de avaliação de vídeo tornaram-se um importante e desafiador tema de pesquisa. Tal avaliação pode ser realizada de forma subjetiva ou objetiva. A avaliação subjetiva estima a qualidade visual através de observadores humanos e é considerada a mais apropriada pelo fato de os próprios seres humanos serem o alvo do conteúdo transmitido. Entretanto, tal técnica requer uma grande quantidade de avaliadores, o que a torna inconveniente e custosa do ponto de vista prático. Por outro lado, a técnica objetiva faz uso de algoritmos para estimar a qualidade do conteúdo visual e é baseada em modelos matemáticos, cujo intuito é obter uma estimativa de qualidade mais próxima possível daquela percebida pelo Sistema Visual Humano (SVH). As métricas de avaliação objetiva apresentam vantagens, pois são imparciais, reproduzíveis, confiáveis e apresentam baixo custo. Sobre tudo nos últimos anos tem se intensificado o desenvolvimento de métodos objetivos de avaliação de qualidade de imagem e vídeo que apresentam maior grau de correlação com os resultados obtidos através do sistema visual humano (VQEG, 2000; Wang et al., 2002a, 2003a, 2004b,a; Engelke & Zepernick, 2007).

As métricas objetivas de avaliação de qualidade de vídeo podem ser subdivididas em três categorias: métrica com referência completa, reduzida e sem referência. A métrica com referência completa necessita de um vídeo de referência, idealmente sem degradação, para sua execução e está restrita a ambientes onde o vídeo de referência é facilmente acessível. A técnica com referência reduzida possui a vantagem de requerer apenas alguns parâmetros do vídeo de referência, que devem ser enviados simultaneamente com o vídeo no canal de transmissão. Finalmente, a técnica sem referência não requer a existência do vídeo original no ponto de avaliação, o que a torna bastante flexível, pois pode ser aplicada em qualquer ponto onde se tenha acesso ao sinal recebido. De forma geral, as métricas objetivas de avaliação de qualidade de vídeo podem ainda ser subdivididas em três categorias: métricas com referência completa (FR, *Full Reference*), com referência reduzida (RR, *Reduced Reference*) e sem referência (NR, *No Reference*).

Este capítulo aborda os fundamentos das técnicas de avaliação de vídeo digital e duas novas métricas desenvolvidas no âmbito do projeto “Formação de Pessoal Qualificado em Sistemas de Transmissão de TV Digital no Paraná”. O capítulo foi dividido em 5 Seções, incluindo esta introdução. A Seção 2 apresenta uma visão geral sobre métodos subjetivos e objetivos de avaliação de qualidade vídeo, a qual inclui a técnica de referência reduzida baseada na diferença de atividade dos coeficientes DCT. A Seção 3 descreve em detalhes a técnica sem referência baseada em um modelo sigmoidal que incorpora características espaço-temporais relacionadas à componente de luminância do sinal de vídeo e a metodologia de ensaio necessária para se testar a métrica. A Seção 4 apresenta os resultados experimentais obtidos, seguida pela conclusão deste capítulo na Seção 5.

2. Avaliação de Qualidade de Vídeo

A avaliação de qualidade de vídeo pode ser realizada de forma subjetiva ou objetiva. A primeira estima a qualidade visual através de avaliadores humanos e requer uma grande quantidade de pessoas, tornando-se impraticável em ambientes de radiodifusores e provedores de conteúdo (VQEG, 2000; Wang et al., 2002a, 2003a, 2004b,a; Engelke & Zepernick, 2007), não apenas pela inconveniência de se dispor constantemente de um grande número de avaliadores como também pelo custo envolvido (Eskicioglu & Fisher, 1995). A segunda faz uso de algoritmos para estimar a qualidade do conteúdo visual e é baseada em modelos matemáticos, cujo intuito é obter uma estimativa de qualidade mais próxima possível da oferecida pela percepção visual humana. A seguir serão abordadas as principais características dessas métricas.

2.1 Avaliação subjetiva de qualidade de vídeo

A qualidade percebida pelo usuário tem papel fundamental na orientação e validação de projetos de sistemas de conteúdo visual (Pappas & Safranek, 2000). A medida de qualidade obtida através da avaliação humana é representada através de um escore, denominado média de escores das opiniões (MOS, *Mean Opinion Score*) e a diferença média de escores das opiniões (DMOS, *Difference Mean Opinion Score*) (ITU-R, 2004). A MOS é indicada para validação de métricas NR, enquanto que a DMOS é uma medida subjetiva indicada para validar métricas FR e RR (VQEG, 2008, 2009, 2010). As recomendações do VQEG (*Video Quality Expert Group*) definem a DMOS conforme a Equação (1) a seguir.

$$DMOS = MOS(PVS) - MOS(ref) + 5, \quad (1)$$

onde PVS (*Processed Video Sequences*) é a sequência de vídeo processada e $MOS(ref)$ é a MOS dos vídeos de referência. Segundo as recomendações

do VQEG, quanto maior o valor da DMOS, melhor é a qualidade do vídeo avaliado.

A avaliação subjetiva de qualidade de vídeo emprega uma metodologia que permite avaliar o grau de satisfação do usuário frente ao conteúdo visual a fim de quantificar a degradação existente no sinal de vídeo. Tal avaliação pode ainda ser dividida em dois métodos: o método que se propõe avaliar o desempenho de sistemas de vídeo em condições ideais (em termos qualitativos) e o método que se compromete a manter a qualidade sob condições ótimas, designado como avaliação de imparidade. A recomendação ITU-R:BT-500.11 (ITU-R, 2004) descreve os diversos ensaios que podem ser aplicados para a realização dos testes subjetivos, tais como: o método SIS (*Double Stimulus Impairment Scale*, escala de imparidades com duplo estímulo e Medidas de robustez do sistema, ou seja, características de falha), DSCQS (*Double Stimulus Continuous Quality scale*, escala contínua de qualidade com duplo estímulo e medidas de qualidade de sistemas referenciados para medição da qualidade de codificação de imagens), SSCQE (*Single Stimulus Continuous Quality Evaluation*, avaliação contínua de qualidade com estímulo simples e medidas de qualidade de sistemas sem referência) e SDSCE (*Simultaneous Double Stimulus for Continuous Evaluation*, avaliação contínua com duplo estímulo simultâneo e Medidas de fidelidade entre duas sequências de vídeos previamente avaliados, comparação entre ferramentas distintas de recuperação de erros).

2.2 Avaliação objetiva de qualidade de vídeo

A avaliação objetiva é compreendida como uma modelagem matemática que possibilita avaliar o grau de degradação do vídeo, após algum processo de distorção. A degradação é perceptível quando surgem artefatos em seu conteúdo, tais como blocagem, borramento, distorção localizada em áreas pouco nítidas do quadro e travamentos (Wu et al., 2007), que interferem diretamente com sua qualidade. A métrica de avaliação mais apropriada é aquela que apresenta maior correlação com os resultados obtidos pelas métricas subjetivas. Neste sentido, o VQEG se apresenta como o fórum internacional no qual a padronização e validação de métricas objetivas para conteúdos visuais são discutidos (Brunnstrom et al., 2009).

2.3 Métricas de referência completa

As métricas FR necessitam de um vídeo de referência para sua implementação. São extensamente utilizadas no processo de validação de propostas RR ou NR, por meio da comparação de desempenho entre os escores objetivos. A seguir são apresentados algumas métricas de referência completa, i.e., PSNR (*Peak Signal-to-Noise Ratio*), SSIM (*Structural SIMilarity index*) e MS-SSIM (*Multi-Scale Structural SIMilarity index*),

amplamente utilizadas na literatura para a avaliação de qualidade de vídeos processados ou transmitidos (Wang et al., 2003b, 2004a; Wu et al., 2007).

O PSNR é uma métrica FR clássica e de fácil implementação, cuja unidade é representada em dB. Entretanto, seus resultados não apresentam boa correlação com as observações realizadas por observadores humanos em alguns casos. Por exemplo, artefatos como a blocagem e o borramento, diferentemente dos artefatos presentes em vídeos analógicos, são efeitos de natureza altamente estruturada oriundos do processo de compressão (codificação através de transformada DCT ou wavelet (Wu et al., 2007; Salomon, 2007)) ao qual um vídeo é submetido.

Sua medida é processada sobre o erro quadro a quadro, i.e., entre o quadro original x e o quadro degradado y . A Equação (2) a seguir determina o PSNR para um conjunto de quadros F .

$$PSNR = \frac{1}{F} \sum_{f=1}^F 20 \log_{10} \left(\frac{v_f}{\sqrt{MSE_f}} \right), \quad (2)$$

onde,

$$MSE_f = \frac{1}{NM} \sum_{i=0}^N \sum_{j=0}^M [x(f, i, j) - y(f, i, j)]^2, \quad (3)$$

onde o termo v_f deriva de $(2^k - 1)$, cujo k é o número de *bits* por *pixel* (componente de luminância) do quadro f . Os termos $x(f, i, j)$ e $y(f, i, j)$ são os valores de luminância do quadro de origem e degradado, respectivamente. As componentes bidimensionais são representadas pelo número de colunas N e pelo número de linhas M . O termo MSE_f é definido como erro quadrático médio ou MSE (*Mean Square Error*) do quadro f que pode assumir valores no intervalo $[0, \infty)$. Assim, a qualidade do vídeo melhora à medida que o MSE se aproxima de zero, por conseguinte, o valor do PSNR tende a aumentar.

A métrica SSIM tem sua origem no fato de o SVH ser fortemente adaptado para extrair informações acerca das características estruturais de um quadro ou imagem. Assim, uma medida de similaridade estrutural (ou distorção) pode prover boa aproximação para a qualidade perceptual de um quadro de vídeo (Wang & Bovik, 2002; Wang et al., 2004a; Wang & Bovik, 2006; Shi et al., 2009).

Sejam x e y dois sinais não-negativos, onde y é o sinal degradado e x é o sinal original (sem perda de qualidade), a medida de similaridade é utilizada como medida quantitativa da qualidade do sinal degradado, ou seja, x é tomado como referência para medir a qualidade de y . As componentes de luminância, contraste e estrutura, que compõe a medida, são definidas conforme as Equações (4-6), respectivamente (Wang et al., 2004a).

$$l(f, x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (4)$$

$$c(f, x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (5)$$

$$s(f, x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}, \quad (6)$$

onde μ_x , μ_y , σ_x^2 , σ_y^2 e σ_{xy} são a média de x e y , a variância de x e y e a covariância de x e y , respectivamente. Os termos C_1 , C_2 e C_3 são constantes de baixa magnitude. Wang et al. (2004a) definem $C_1 = (K_1L)^2$, onde L é um valor dinâmico de *pixel* (255 para quadros em tom cinza de 8 *bits*) e $K_1 \ll 1$; $C_2 = (K_2L)^2$ e $K_2 \ll 1$; $C_3 = \frac{C_2}{2}$.

A Equação (7) do índice SSIM relaciona x e y com as componentes de luminância, contraste e estrutura para um conjunto de quadros F , conforme (Wang et al., 2004a):

$$SSIM = \frac{1}{F} \sum_{f=1}^F [l(f, x, y)]^\alpha [c(f, x, y)]^\beta [s(f, x, y)]^\gamma, \quad (7)$$

onde os parâmetros α , β e γ são usados para ajustar as três componentes do SSIM. Wang et al. (2004a) simplificaram a Equação (7), tornando os parâmetros α , β e γ positivos e iguais a 1.

A métrica SSIM apresenta diversas extensões, dentre as quais a versão MS-SSIM, proposta por Wang et al. (2003b) que assume uma abordagem multiescalar, através de um método iterativo (Φ iterações) em que se emprega um filtro passa-baixa e uma subamostragem (*downsampling*) no quadro ou imagem. A Equação (8) a seguir resume o método.

$$MS-SSIM = \frac{1}{F} \sum_{f=1}^F [l_\Phi(f, x, y)]^{\alpha_\Phi} \cdot \prod_{k=1}^{\Phi} [c_k(f, x, y)]^{\beta_k} [s_k(f, x, y)]^{\gamma_k}, \quad (8)$$

onde $\sum_{k=1}^{\Phi} \alpha_k + \beta_k + \gamma_k = 1$, $\Phi = 5$ e $\forall k \in [1, \dots, \Phi]$; $0 \leq \alpha_k \leq 1$; $0 \leq \beta_k \leq 1$; $0 \leq \gamma_k \leq 1$. Analogamente ao que foi abordado em (7), os expoentes α_Φ , β_k e γ_k são utilizados para ajustar as diferentes componentes da Equação (8). A componente de luminância do quadro f , denotada por $[l_\Phi(f, x, y)]^{\alpha_\Phi}$ é calculada somente na escala Φ .

2.4 Métricas de referência reduzida

Métricas de referência reduzida extraem um certo número de características do vídeo original, em função do movimento ou em detalhes espaciais. Este método é largamente utilizado no monitoramento de transmissões em rede (Silva et al., 2013). Neste cenário, o vídeo é transmitido com uma sequência de informação codificada (*overhead*) e no lado receptor ocorre a sua decodificação, seguido pelo cálculo do índice de qualidade, o qual é obtido através da comparação entre a representação reduzida de informação nos pares fonte e destino (Le Callet et al., 2006). Tipicamente, métricas RR são implementadas através de funções, divididas em duas etapas (Le Callet & Barba, 2001):

1. Cálculo do erro entre os vídeos original e degradado, através da diferença de suas características para compor a representação reduzida da informação;
2. Função que agrupa os erros ou diferenças para obter um índice de qualidade global.

Algumas categorias de métricas RR exploram propriedades dos artefatos, através da extração de características e um modelo de parametrização (Miyahara et al., 1998), cujo foco é concentrado em alguns tipos particulares de artefatos.

Esta abordagem pode ser encontrada em (Silva et al., 2013) com a proposta do método RRVQA (*Reduced-Reference Video Quality Assessment*) para avaliação de qualidade de vídeo baseado na diferença de atividade dos coeficientes DCT (*Discrete Cosine Transform*) (Rao & Yip, 1990). As componentes de alta frequência, i.e., os coeficientes AC (*Alternating Coefficient*), são responsáveis pelos detalhes em um quadro e são menos perceptíveis pelo SVH. Por esse motivo, os sistemas de compressão (i.e., MPEG-2 e H.264) reduzem a informação contida nesses coeficientes, tendo como consequência a produção de artefatos típicos do processo de compressão, tais como blocagem e borramento. Assim, o método RRVQA opera sobre a diferença entre os coeficientes AC e DC (*Direct Coefficient*) em um macrobloco com resolução $\tau \times \tau$, com $\tau = 16$, tipicamente. Os coeficientes em cada macrobloco são representados por $coef_p$ (coeficientes AC e DC) e a média dos coeficientes DC representada por \overline{DC} . Logo, a diferença absoluta de atividade para um macrobloco j com resolução $\tau \times \tau$ é definida pela Equação (9) a seguir.

$$Act_{f_j} = \frac{1}{\tau \times \tau} \sum_{p=1}^{\tau \times \tau} |coef_p - \overline{DC}|, \quad (9)$$

onde \overline{DC} é definido como

$$\overline{DC} = \frac{64}{\tau \times \tau} \sum_{k=1}^{\frac{\tau \times \tau}{64}} |DC_k|. \quad (10)$$

O parâmetro $Actf_j$ é calculado sobre cada macrobloco dos quadros do vídeo de origem ou de referência e encapsulado em um TS (*Transport Stream*), o qual é transmitido sobre o canal de comunicação, i.e., TV digital ou a rede IP (*Internet Protocol*). Este método gera um *overhead* entre 19 e 21 *bits* para cada macrobloco. No lado receptor, calcula-se a variável $Actf_j$ (9) e o erro quadrático entre as diferenças absolutas de atividades nos lados emissor ($ActfS_j$) e receptor ($ActfR_j$), conforme a Equação (11) a seguir.

$$SEf_j = (ActfS_j - ActfR_j)^2. \quad (11)$$

Esta abordagem requer o cálculo do MSE no domínio da frequência, i.e., a média da Equação (11) que considera todos os macroblocos (M) de um quadro i .

$$MSEf_i = \frac{1}{M} \sum_{j=1}^M SEf_j, \quad (12)$$

Logo, o método RRVQA tem uma abordagem similar à métrica PSNR, contudo, no domínio da transformada DCT, cuja definição é dada pela Equação (13) a seguir.

$$RRVQA = \frac{1}{N} \sum_{i=1}^N 10 \times \log_{10} \frac{[\max(cs_i, cr_i)]^2}{MSEf_i}, \quad (13)$$

onde N é o número de quadros em um vídeo, cs_i e cr_i são os coeficientes DCT do lado emissor e receptor, respectivamente. Somente, as variáveis $Actf_j$ calculada na Equação (9) e cs_i são encapsulados no TS.

2.5 Métricas sem referência

Diferentemente das métricas FR ou RR, métodos NR tentam realizar suposições acerca do conteúdo ou das degradações de um determinado vídeo, baseado na relação entre suas características e no SVH (Wu et al., 2007). Basicamente há duas abordagens no desenvolvimento de métricas objetivas sem referência para avaliação de qualidade de vídeos:

- i. extração de parâmetros sobre o fluxo de *bits* do vídeo, tais como taxa de compressão, informações relativas ao GoP (*Group of Pictures*) e ao parâmetro de quantização (QP – *Quantization Parameter*) (Slanina et al., 2007; Sugimoto et al., 2009; Oelbaum et al., 2009; Staelens et al., 2010; Yang et al., 2010);

- ii. extração de características espaço-temporais, a fim de que sejam detectados distorções ou artefatos nos quadros que impactam diretamente sobre a qualidade do vídeo (Kawayoke & Horita, 2008; Keimel et al., 2009; Yao et al., 2009; Wang et al., 2009).

Métodos de avaliação de qualidade de imagens NRIQA (No-Reference Image Quality Assessment) podem recorrer à extração de características espaciais associadas a artefatos de blocagem e borramento através de uma modelagem não-linear que utiliza o método iterativo de Levenberg-Marquardt para minimização de mínimos quadrados (Wang et al., 2002b). Métricas NRIQA também podem ser baseadas em modelagem sigmoidal que incorpora outras características espaciais, tais como a distorção dos *pixels* e a medida de inarfação de borda (Parvez Sazzad et al., 2008) ou em características baseadas no SVH, i.e., brilho e a complexidade espacial em imagens com artefatos de blocagem (Song & Yang, 2009). A avaliação objetiva de qualidade sem referência apresenta maior complexidade quando aplicada em sequências de vídeo, devido à variação temporal dos quadros, conteúdo das cenas e à dificuldade de compreensão das características do SVH (Wang et al., 2003a). Algumas técnicas exploram características espaciais e temporais na estimação de qualidade de vídeo (Yang et al., 2005; Eden, 2008; Kawano et al., 2010). Além desta abordagem, também é possível utilizar as informações contidas no fluxo de *bits*, tais como tamanho e tipo de GoP, taxa de compressão, parâmetro de quantização, tipo de *codec*, informações acerca dos estimadores de movimento, perfil e nível de codificação (Hemami & Reibman, 2010; Keimel et al., 2011a,b; Yamagishi et al., 2012).

A seguir é apresentada a métrica JPEG-NR proposta por Wang et al. (2002b) que emprega a detecção de artefatos de blocagem e borramento em imagens, cuja componente de luminância é representado por $y(f, i, j)$ com $i \in [1, M]$ e $j \in [1, N]$, onde M equivale ao número de linhas e N ao número de colunas em uma imagem ou quadro f . As diferenças ao longo das direções horizontal (colunas) e vertical (linhas) são determinadas conforme as Equações (15-15):

$$d_h(f, i, j) = y(f, i, j + 1) - y(f, i, j), j \in [1, N - 1], \quad (14)$$

$$d_v(f, i, j) = y(f, i + 1, j) - y(f, i, j), i \in [1, M - 1]. \quad (15)$$

O efeito de blocagem pode ser estimado pela média das diferenças entre as bordas dos blocos DCT nas direções horizontal e vertical, conforme as Equações (16-17), com resolução típica de 8×8 *pixels* no bloco DCT, i.e., $\tau = 8$.

$$B_h = \frac{1}{M \left(\lfloor \frac{N}{\tau} \rfloor - 1 \right)} \sum_{i=1}^M \sum_{j=1}^{\lfloor \frac{N}{\tau} \rfloor - 1} |d_h(f, i, \tau j)|, \quad (16)$$

$$B_v = \frac{1}{N \left(\lfloor \frac{M}{\tau} \rfloor - 1 \right)} \sum_{i=1}^{\lfloor \frac{M}{\tau} \rfloor - 1} \sum_{j=1}^N |d_v(f, \tau i, j)|. \quad (17)$$

A combinação entre B_h e B_v resulta no descritor de blocagem B_f :

$$B_f = \frac{B_h + B_v}{2}. \quad (18)$$

A medida do artefato de borramento é obtida através da redução de atividade espacial nas direções horizontal e vertical, devido ao processo de quantização no domínio da DCT. As Equações (19-20) expressam as médias das diferenças da redução de atividade espacial nas direções horizontal e vertical, respectivamente.

$$A_h = \frac{1}{\tau - 1} \left[\frac{\tau}{M(N - 1)} \sum_{i=1}^M \sum_{j=1}^{N-1} |d_h(f, i, j)| - B_h \right], \quad (19)$$

$$A_v = \frac{1}{\tau - 1} \left[\frac{\tau}{N(M - 1)} \sum_{i=1}^{M-1} \sum_{j=1}^N |d_v(f, i, j)| - B_v \right]. \quad (20)$$

A combinação entre A_h e A_v gera o descritor A_f para detecção de artefatos de borramento:

$$A_f = \frac{A_h + A_v}{2}. \quad (21)$$

O segundo fator que contribui para a detecção de artefatos de borramento é a taxa *zero-crossing* (ZC) nas direções horizontal e vertical.

$$Z_h = \frac{1}{M(N - 2)} \sum_{i=1}^M \sum_{j=1}^{N-2} z_h(f, i, j), \quad (22)$$

$$Z_v = \frac{1}{N(M - 2)} \sum_{i=1}^{M-2} \sum_{j=1}^N z_v(f, i, j), \quad (23)$$

onde z_h e z_v são expressos como:

$$z_h(f, i, j) = \begin{cases} 1, & ZC \text{ na direção horizontal, com } d_h(f, i, j) \\ 0, & \text{caso contrário} \end{cases} \quad (24)$$

$$z_v(f, i, j) = \begin{cases} 1, & ZC \text{ na direção vertical, com } d_v(f, i, j) \\ 0, & \text{caso contrário} \end{cases} \quad (25)$$

A combinação entre Z_k e Z_v gera o descritor de borrimento Z_f :

$$Z_f = \frac{Z_h + Z_v}{2}. \quad (26)$$

Finalmente, o modelo NR proposto por Wang et al. (2002b) combina os descritores A_f , B_f e Z_f para um dado número de quadros F , conforme a Equação (27):

$$JPEG-NR = \frac{1}{F} \sum_{f=1}^F \alpha + \beta B_f^{\gamma_1} A_f^{\gamma_2} Z_f^{\gamma_3}, \quad (27)$$

onde $\alpha = -245,9$, $\beta = 261,9$, $\gamma_1 = -0,0240$, $\gamma_2 = 0,0160$ e $\gamma_3 = 0,0064$ são parâmetros estimados através de uma regressão não-linear com o método iterativo de Levenberg-Marquardt (Levenberg, 1944; Marquardt, 1963; Moré, 1977), descrito no apêndice deste capítulo.

Além das características espaciais, as características temporais devem também ser consideradas. O SVH é mais sensível à luminância do que à crominância (Wu et al., 2007), por isso, nesta abordagem são usadas apenas as características espaço-temporais relacionadas à componente de luminância. Como visto acima, o modelo de Wang et al. (2002b) emprega três características espaciais associadas à detecção de artefatos de bloqueio com o descritor B e de borrimento através dos descritores A e Z (Wang et al., 2002b). A eles deve ainda ser acrescentado três descritores relacionados à variação temporal entre quadros: informação perceptual temporal TI , média da diferença absoluta MAD (*Mean Absolute Difference*) e média da diferença absoluta ponderada $MADw$ (*Mean Absolute Difference weight*).

O descritor de bloqueio B para um conjunto de quadros F é obtido a partir da média da Equação (18):

$$B = \frac{1}{F} \sum_{f=1}^F B_f. \quad (28)$$

Os descritores de borrimento A e Z para um sequência de vídeo com F quadros podem ser reescritos a partir das médias das Equações (21) e (26):

$$A = \frac{1}{F} \sum_{f=1}^F A_f, \quad (29)$$

$$Z = \frac{1}{F} \sum_{f=1}^F Z_f. \quad (30)$$

O descritor *TI* (ITU-T P.910, 1999) considera a diferença de movimento de quadros adjacentes f e $f - 1$, representada por $m(f, i, j)$ com os valores de luminância de *pixels* localizados em uma mesma região espacial $y(f, i, j)$, com i representando a linha e j a coluna do *pixel* y no quadro f . O descritor *TI* é definido como:

$$TI = \frac{1}{F-1} \sum_{f=2}^F \sigma[m(f, i, j)], \quad (31)$$

onde F é o número total de quadros de um vídeo e $\sigma[m(f, i, j)]$ é o desvio-padrão da diferença de movimento entre quadros sucessivos, i.e., $m(f, i, j) = y(f, i, j) - y(f-1, i, j)$ com $f > 1$.

O descritor *MAD* (Ding et al., 2008) representa a diferença temporal absoluta entre quadros sucessivos $m(f, i, j)$, cuja definição é:

$$MAD_k = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |m(f, i, j)|, \quad (32)$$

onde MAD_k tem $F-1$ amostras, i.e., $k = 1, 2, \dots, F-1$. Logo, o descritor MAD_k tem a média definida como:

$$\overline{MAD} = \frac{1}{F-1} \sum_{k=1}^{F-1} MAD_k. \quad (33)$$

O descritor *MADw* (Ding et al., 2008) considera a razão entre o valor de *MAD* do quadro atual k e do anterior $k-1$, conforme a definição a seguir.

$$MADw = \frac{1}{F-1} \sum_{k=1}^{F-1} \frac{MAD_k}{MAD_{k-1}}. \quad (34)$$

A Figura 1 exibe o comportamento dos seis descritores A , B , Z , TI , \overline{MAD} e $MADw$ que caracterizam trinta vídeos codificados em MPEG-2. Em algumas amostras de vídeo da Figura 1, i.e., no intervalo [5, 10] e [25, 30], observa-se a mesma tendência desses descritores tanto nas curvas espaciais (A , B , Z) quanto nas temporais (TI , \overline{MAD}).

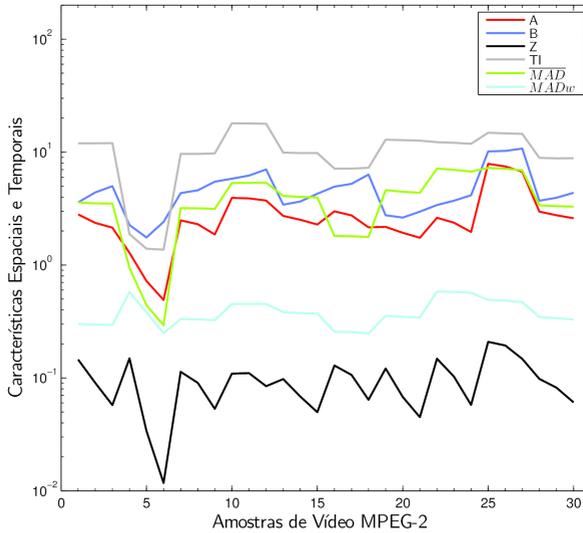


Figura 1. Características espaço-temporais para vídeos em MPEG-2.

3. Método NRVQA-LM

O método NRVQA-LM (*No-Reference Video Quality Assessment using LM method*) proposto tem origem em estudos empíricos envolvendo manipulações matemáticas entre as características espaciais e temporais descritas na seção anterior e os escores subjetivos das bases de dados estudadas. A métrica NRVQA-LM é determinada por uma abordagem analítica baseada em um modelo sigmoidal que incorpora três características espaciais e três temporais descritas na Seção 2.5. Além disto, a função sigmoidal proposta apresenta um comportamento monotônico reportado em diversos trabalhos relacionados à problemas de avaliação de qualidade de imagem e vídeo (Parvez Sazzad et al., 2008; Kawano et al., 2010; Keimel et al., 2011a,b). A fase de treinamento compreende o mapeamento entre as entradas ($B, Z, A, TI, MAD, MADw$) e os escores subjetivos, a partir de uma solução de mínimos quadrados. As Equações (35), (36) e (37) definem os métodos espacial, temporal e espaço-temporal proposto, respectivamente.

$$NRVQA-LMe = \frac{1}{1 + e^{(\beta_1 B + \beta_2 Z + \beta_3 A + \beta_4)}}, \quad (35)$$

$$NRVQA-LMt = \frac{1}{1 + e^{(\beta_4 TI + \beta_5 MAD + \beta_6 MADw + \beta_7)}}, \quad (36)$$

$$NRVQA-LM = \frac{1}{1 + e^{(\beta_1 B + \beta_2 Z + \beta_3 A + \beta_4 TI + \beta_5 MAD + \beta_6 MADw + \beta_7)}}, \quad (37)$$

onde os parâmetros β_1 a β_7 são otimizados pelo método iterativo LM (Levenberg, 1944; Marquardt, 1963; Moré, 1977) descrito no apêndice deste capítulo e reportado pela literatura em diversos trabalhos relacionados à avaliação de qualidade de imagem e vídeo que o empregam na solução de problemas de mínimos quadrados (Wang et al., 2002b; Ries et al., 2006; Engelke & Zepernick; Keimel et al., 2009; Brandão & Queluz, 2010; Shahid et al., 2011; Calyam et al., 2012; Kipli et al., 2012).

A Figura 2 apresenta o diagrama de blocos do método proposto pela Equação (37). A fase de treinamento compreende a seleção de um determinado conjunto de vídeo, i.e., baseado no tipo de artefato ou codificação, bem como os seus escores subjetivos normalizados no intervalo $[0, 1]$, os parâmetros β' s iniciais, i.e., na primeira iteração $\beta_{1,\dots,7} = 0$, e as seis características espaço-temporais ($A, B, Z, TI, MAD, MADw$). A fase de teste compreende a atualização do modelo NRVQA-LM por meio da otimização dos parâmetros β' s através do método iterativo LM e a obtenção do escore de qualidade, a partir da seleção do vídeo de teste.

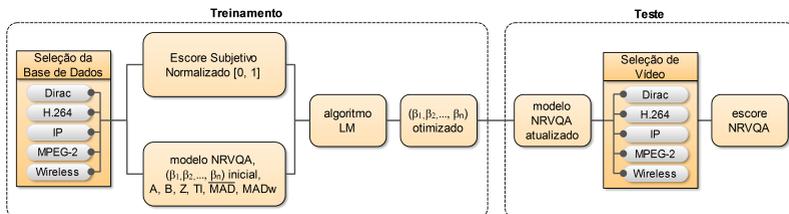


Figura 2. Diagrama de blocos do método NRVQA-LM com o treinamento dos parâmetros β' s pelo método iterativo LM (Silva & Pohl, 2012).

A Figura 3 compara o desempenho entre as componentes espacial (35), temporal (36) e espaço-temporal (37) do método proposto e a DMOS normalizada (DMOSn) no intervalo $[0, 1]$ para o conteúdo em MPEG-2 da base de dados LIVE. Em algumas amostras de vídeo, conforme inspeção visual da Figura 3 observa-se que o método proposto em (37), composto tanto pelas características espaciais quanto temporais possui maior correlação com a medida subjetiva. A acurácia do método NRVQA-LM é igual a 0,9473, enquanto que as componentes espacial e temporal apresentam uma acurácia de 0,8906 e 0,3544, respectivamente.

Além do modelo sigmoidal proposto na Equação (37), outro fator que tem forte influência no desempenho da métrica proposta é o treinamento dos parâmetros β' s que estão intrinsecamente relacionados com o conteúdo da base de dados escolhida.

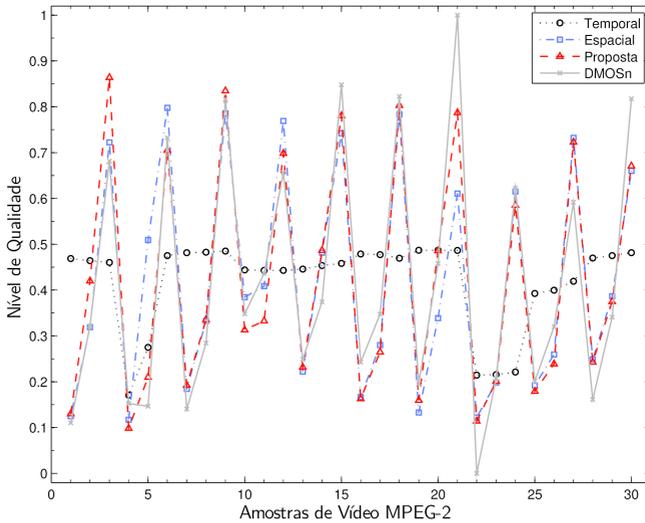


Figura 3. Comparação entre os modelos sigmoidal temporal, espacial, espaço-temporal e a DMOSn para vídeos em MPEG-2 (base de dados LIVE).

3.1 Metodologia de treinamento e ensaio

Nesta seção é descrita a metodologia empregada para ensaio do método. Além de gráficos bidimensionais, também foi utilizada uma forma de representação que sintetiza estatisticamente o comportamento de uma série de dados através de um diagrama de caixa ou *box-plot* (Tukey, 1977). A Figura 4 apresenta um exemplo com escores de qualidade distribuídos segundo o diagrama de caixa à direita. Os valores extremos ou discrepantes são representados pelo símbolo $+$, enquanto que os valores não discrepantes inferiores e superiores estão confinados entre o símbolo \perp e o primeiro quartil (Q_1) e entre o terceiro quartil (Q_3) e o símbolo \top , respectivamente. A mediana (segundo quartil ou Q_2), média e desvio-padrão são representados pelos símbolos $—$, \bullet e \blacksquare , respectivamente.

O desvio-quartil (amplitude interquartílica) e a mediana (Q_2) são medidas de dispersão que não são influenciadas pelos valores extremos, como ocorre com a média e desvio-padrão. A amplitude interquartílica corresponde à diferença entre o terceiro e o primeiro quartil ($Q_3 - Q_1$) e contém 50% dos dados da distribuição. O primeiro (Q_1), segundo (Q_2) e terceiro (Q_3) quartis indicam que os valores da distribuição são menores que 25%, 50% e 75%, respectivamente.

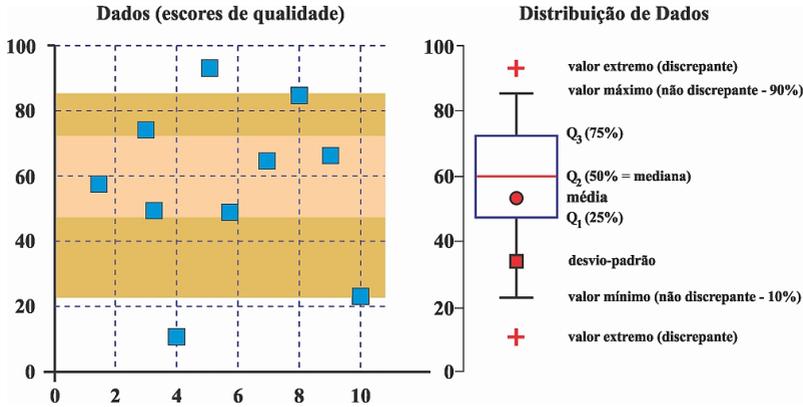


Figura 4. Diagrama de caixa (*box-plot*).

3.2 Bases de dados de vídeo

Para ensaio das métricas NR foi utilizada o banco de vídeos LIVE (Seshadrinathan et al., 2010), acompanhada dos descritores de informação temporal (TI) e espacial (SI – *Spatial perceptual Information*). Os gráficos TI vs. SI foram obtidos conforme a recomendação ITU-T P.910 (ITU-T P.910, 1999) que considera os valores de pico destas variáveis. Esta base de dados inclui 150 amostras de vídeos com dez conteúdos que contemplam distorções por compressão MPEG-2 e H.264, transmissão em canal ruidoso sem fio e rede IP. Nesta base de dados há sete seqüências de vídeo com 25 fps e as três restantes com 50 fps. Os arquivos não contêm cabeçalho, possuem o formato 4:2:0 progressivo e resolução de 768×432 pixels, cujos escores subjetivos estão em escala DMOS. A Figura 5 descreve o comportamento espacial (Figura 5-a), temporal (Figura 5-b) e temporal vs. espacial (Figura 5-c) para os vídeos de referência dessa base de dados.

Na validação cruzada a base de dados foi dividida em três grupos, G1 (Grupo 1), G2 (Grupo 2) e S (superconjunto) em cada uma das cinco categorias H.264, IP, MPEG-2, *Wireless* e toda base de dados (Todos) com 40, 40, 30, 40 e 150 amostras, respectivamente. Para a categoria H.264, por exemplo, G1 e G2 possuem 20 amostras aleatórias e distintas para cada grupo e S possui todas as amostras (40). Este experimento foi repetido mil vezes ($k = 1, \dots, 1000$) com seqüências diferentes tanto para $\{G1_k\}$, quanto para $\{G2_k\}$, cuja formação do par treinamento-teste obedeceu à alternância entre todos os grupos: $\{G1_k\}$ - $\{G1_k\}$, $\{G1_k\}$ - $\{G2_k\}$, $\{G1_k\}$ - $\{S\}$, $\{G2_k\}$ - $\{G1_k\}$, $\{G2_k\}$ - $\{G2_k\}$, $\{G2_k\}$ - $\{S\}$, $\{S\}$ - $\{G1_k\}$, $\{S\}$ - $\{G2_k\}$, $\{S\}$ - $\{S\}$. As métricas PSNR, SSIM, MS-SSIM, JPEG-NR e o método proposto NRVQA-LM foram comparados no processo de validação. A Equação (38) mostra a

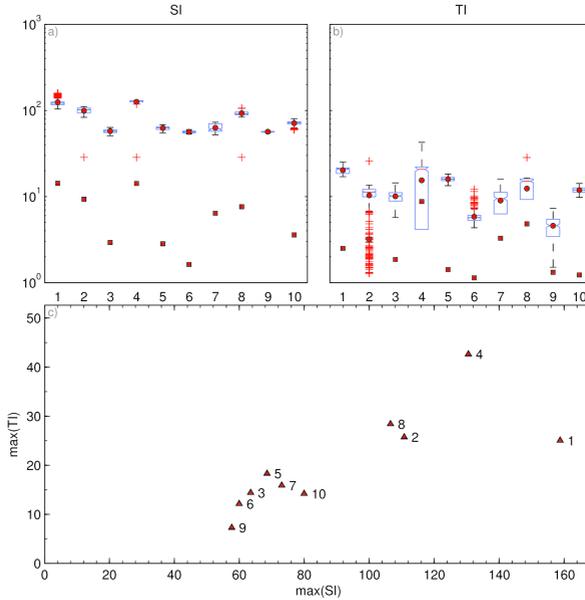


Figura 5. Diagrama TI vs. SI para a base de dados LIVE. (a) *box-plot* para SI (b) *box-plot* para TI (c) $\max(TI)$ vs. $\max(SI)$.

relação entre os conjuntos $\{G1_k\}$, $\{G2_k\}$ e $\{S\}$. Logo, a união entre $\{G1_k\}$ e $\{G2_k\}$ para qualquer k resulta em $\{S\}$, embora não exista interseção entre $\{G1_k\}$ e $\{G2_k\}$ ambos são subconjuntos de $\{S\}$.

$$\begin{aligned}
 \{G1_k\} \cup \{G2_k\} &= \{S\}, \\
 \{G1_k\} &\subset \{S\} \text{ e } \{G2_k\} \subset \{S\}, \\
 \{G1_k\} \cap \{G2_k\} &= \emptyset.
 \end{aligned}
 \tag{38}$$

3.3 Processamento dos escores objetivos e subjetivos

O mapeamento entre os escores objetivos (disponíveis no banco de vídeos) e subjetivos e o processo de validação dos métodos propostos neste trabalho foram seguidos conforme as recomendações disponíveis nas versões mais atuais do Grupo de Especialistas em Qualidade de Vídeo (VQEG, 2008, 2009, 2010). A Figura 6 resume as etapas de mapeamento entre os escores objetivos (VQAS – *Video Quality Assessment Scores*) e subjetivos (DMOS), bem como o processo de validação de métodos objetivos de qualidade de vídeo.

Neste trabalho, utilizou-se uma função polinomial cúbica no mapeamento não-linear entre VQAS e a DMOS, conforme recomendações

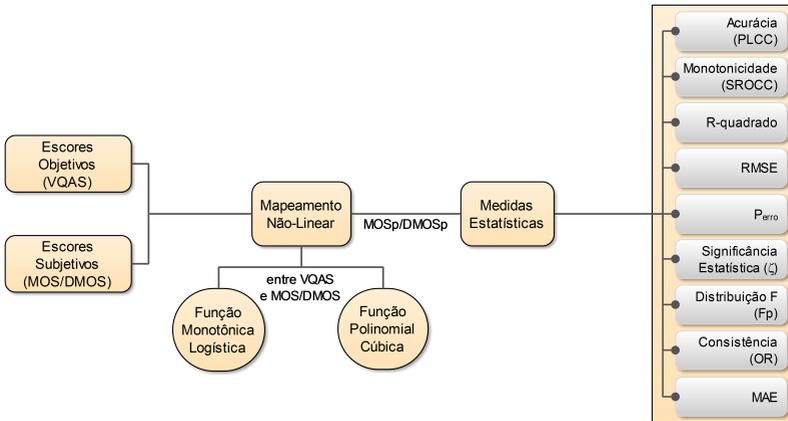


Figura 6. Pós-processamento dos escores objetivos e medidas estatísticas de desempenho.

mais atuais do VQEG. (VQEG, 2008, 2009, 2010) que além de ser uma predição mais simples, não causa sobreposição de dados como ocorre com a função monotônica logística (Engelke et al., 2009). A função polinomial cúbica é definida como

$$DMOSp = ax^3 + bx^2 + cx + d, \quad (39)$$

onde os coeficientes a , b , c e d são obtidos através de um ajuste (*fitting*) cúbico entre os escores subjetivos e objetivos.

Após o mapeamento não-linear, os escores objetivos (VQAS) são mapeados para a escala DMOS e recebem a denominação de DMOSp, conforme observado na Figura 6. Procedimento análogo ao da Equação (39) também pode ser aplicado à DMOSp. Em seguida, ocorre o processo de validação entre DMOS e DMOSp através de medidas estatísticas de desempenho, dentre as quais o coeficiente de correlação linear de Pearson ou PLCC (*Pearson Linear Correlation Coefficient*) e o coeficiente de correlação de postos de Spearman ou SROCC (*Spearman Rank Order Correlation Coefficient*), R-quadrado (R^2), RMSE, P_{erro} (erro entre os escores subjetivos e objetivos), significância estatística (ζ), distribuição F percentual (F_p) com quatro graus de liberdade, OR (*Outlier Ratio*) ou consistência e a média do erro absoluto ou MAE (*Mean Absolute Error*).

4. Resultados Experimentais

Esta seção apresenta os resultados experimentais dos métodos propostos com a base de dados LIVE (768×432p) que contém escores subjetivos (DMOS) de 150 vídeos com distorções em H.264, IP, MPEG-2 e *Wireless*.

A Figura 7 compara o desempenho entre a métrica PSNR e o método RRVQA, em termos da acurácia (PLCC) e da monotonicidade (SROCC). O método RRVQA apresenta uma acurácia de 0,7039 e monotonicidade de 0,6756, enquanto que a métrica PSNR tem acurácia e monotonicidade igual a 0,5533 e 0,5222, respectivamente.

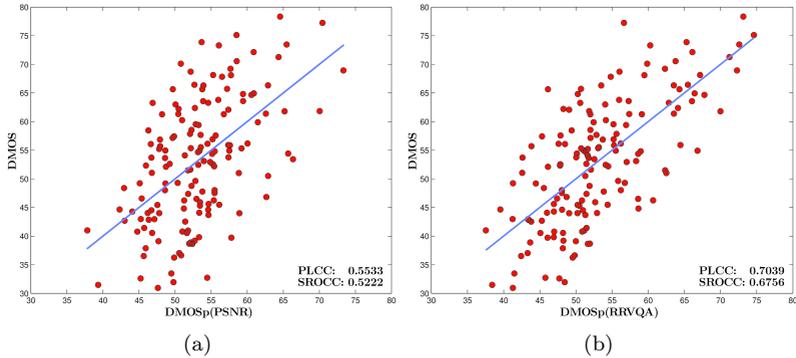


Figura 7. Comparação da acurácia (PLCC) entre a métrica PSNR e o método RRVQA para todos os vídeos da base de dados LIVE (Seshadrinathan et al., 2010). (a) SSIM vs. DMOS; (b) RRVQA vs. DMOS.

A Figura 8 compara a distribuição estatística da acurácia (PLCC) entre o método NRVQA-LM e as métricas PSNR, SSIM, MS-SSIM e JPEG-NR para os conteúdos da base de dados LIVE. Pela inspeção visual desta figura, observa-se que o método NRVQA-LM apresenta melhor desempenho para os conteúdos H.264 e MPEG-2. Os artefatos de compressão H.264 e MPEG-2 causam efeitos monotônicos e não-contínuos na qualidade das amostras de vídeo (Wu et al., 2007). Esta pode ser a razão pela qual o método proposto tenha apresentado melhor desempenho. Entretanto, o mesmo apresenta menor acurácia quando os conteúdos estão embaralhados (categoria Todos), i.e., Figura 8-y, devido ao treinamento dos coeficientes β^i s que estão associados com as características de um conteúdo específico de uma base de dados. Além disso, os artefatos para os conteúdos IP e *Wireless* foram gerados artificialmente e de uma maneira não-monotônica Seshadrinathan et al. (2010). Logo, estas condições podem reduzir a acurácia na previsão de qualidade do método proposto.

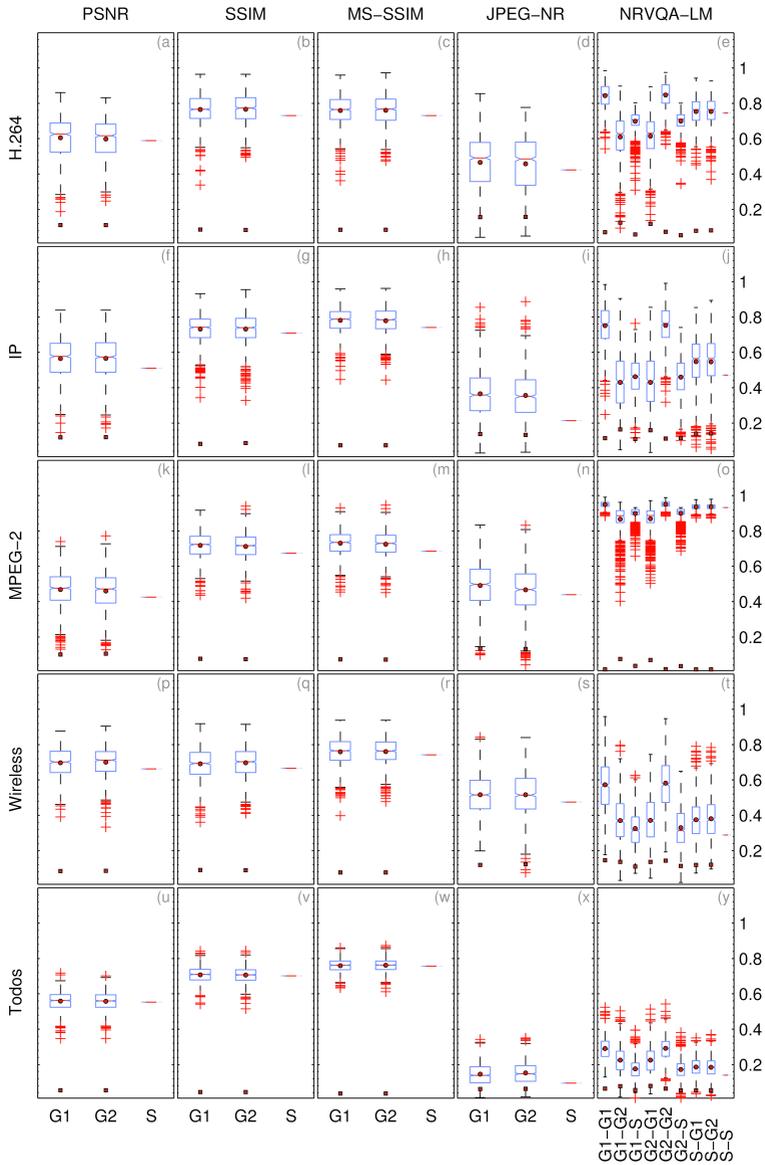


Figura 8. O eixo das abscissas compreende os grupos de teste das métricas comparadas e os pares treinamento-teste para o método NRVQA-LM. O eixo das ordenadas correspondem aos valores de acurácia (PLCC) da base de dados LIVE.

5. Conclusão

Este capítulo apresentou dois métodos para avaliação objetiva de vídeo: um método com referência reduzida (RRVQA) e um método sem referência com abordagem em modelagem sigmoïdal (NRVQA-LM). O método RRVQA é baseado na diferença de atividade dos coeficientes DCT e, embora o SVH tenha menor sensibilidade às componentes de alta frequência, estas componentes são mais susceptíveis a erros em uma transmissão sob condições ruidosas. Estes erros são detectados pelo método RRVQA através da diferença de atividade no domínio da frequência, envolvendo os coeficientes AC e DC. Este método apresenta maior acurácia do que a métrica FR PSNR, conforme os resultados observados na Figura 7. O método NRVQA-LM explora a extração de características espaço-temporais relacionadas à componente de luminância. Os resultados experimentais comparam a distribuição estatística da acurácia (PLCC) com a validação cruzada entre o método sem referência apresentado e as métricas de referência completa PSNR, SSIM e MS-SSIM e sem referência JPEG-NR. O método NRVQA-LM apresenta melhor desempenho para os conteúdos H.264 e MPEG-2, conforme observa-se na Figura 8. A comparação entre os resultados propostos e os reportados na literatura é limitada, devido ao uso de diferentes funções de mapeamento. A Função de mapeamento não-linear polinomial cúbica é sugerida pelas mais recentes recomendações do VQEG, embora a maioria dos trabalhos encontrados na literatura utilizam recomendações já ultrapassadas, as quais sugerem o mapeamento entre os escores objetivos e subjetivos através de uma função não-linear monotônica logística. Os métodos propostos são adequados para o monitoramento de qualidade de vídeo em sistemas de *broadcasting*, tais como TV Digital e IPTV com a recepção do sinal de vídeo em dispositivos móveis, i.e., *smartphones*, *tablets*, PCs móveis ou em aparelhos com a tecnologia WiDi (*Wireless Display*).

Apêndice

O método iterativo LM (Levenberg, 1944; Marquardt, 1963; Moré, 1977) combina a vantagem de aceleração no processo de localização do mínimo de uma função não-linear baseada no método do gradiente descendente (Widrow & Hoff Jr., 1960) e na aceleração de convergência baseada na localização de mínimos adjacentes através do método de Gauss-Newton (Hoffman, 2001).

As Equações (40) a (48) detalham o método LM, cuja finalidade é minimizar o vetor de erro ($\mathbf{d} = \mathbf{s} - \hat{\mathbf{s}}$) que considera o vetor de escores de qualidade \mathbf{s} e a saída desejada, i.e., o vetor de escores subjetivos $\hat{\mathbf{s}}$.

$$\arg \min_{(\beta_1, \beta_2, \dots, \beta_7)} \mathbf{d}^T \mathbf{d}, \quad (40)$$

onde

$$\mathbf{d} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_\phi] \text{ e}$$

$$d_\phi = [Subj_\phi - f(A_\phi, B_\phi, Z_\phi, TI_\phi, \overline{MAD}_\phi, MADw_\phi; \beta_1, \beta_2, \dots, \beta_7)]$$

com $Subj_\phi$ denotando o escore subjetivo da amostra de vídeo ϕ de um determinado subconjunto de uma base de dados. A função f que descreve o produto $\mathbf{d}^T \mathbf{d}$ em (41) é determinada pela relação entre as características espaço-temporais descritas na Seção 2.5 e os parâmetros β 's. Assim, f é definida pelo modelo não-linear proposto, cujos resultados são os escores objetivos, representados pelo vetor \mathbf{s} .

$$\mathbf{d}^T \mathbf{d} = f(A_\phi, B_\phi, Z_\phi, TI_\phi, \overline{MAD}_\phi, MADw_\phi; \beta_1, \beta_2, \dots, \beta_7), \quad (41)$$

O método de Gauss-Newton é uma extensão do método de Newton. Este requer que a matriz Hessiana (\mathbf{H}) seja definida positivamente e inversível (Moré, 1977). Todavia, nem sempre há garantias de que \mathbf{H} obedeça a estas condições, ao passo que o método de Gauss-Newton requer apenas a matriz Jacobiana do vetor de erros. O método de Gauss-Newton é utilizado para minimizar a função f através de uma fórmula recorrente que considera os termos \mathbf{H} e o gradiente de uma função (∇f).

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \mathbf{H}^{-1} \nabla f, \quad (42)$$

onde k é o número de iteração, \mathbf{H} é a matriz quadrada das derivadas parciais de segunda ordem da função f que pode ser aproximada conforme a Equação (43) a seguir.

$$\mathbf{H} = \mathbf{J}^T \mathbf{J}, \quad (43)$$

e o gradiente ∇f pode ser expresso conforme

$$\nabla f = \mathbf{J}^T \mathbf{d}, \quad (44)$$

onde \mathbf{J} , conforme (45), é a matriz Jacobiana que contém as primeiras derivadas do vetor de erros \mathbf{d} .

$$\mathbf{J} = \begin{pmatrix} \frac{\partial d_1}{\partial \beta_1} & \dots & \frac{\partial d_1}{\partial \beta_7} \\ \vdots & \ddots & \vdots \\ \frac{\partial d_\phi}{\partial \beta_1} & \dots & \frac{\partial d_\phi}{\partial \beta_7} \end{pmatrix}. \quad (45)$$

O método de Gauss-Newton apesar de rápido, apresenta problemas de estabilidade quanto à invertibilidade da matriz Hessiana aproximada, i.e., $\mathbf{H} = \mathbf{J}^T \mathbf{J}$. Levenberg (1944) propôs uma alteração no método de Gauss-Newton com a introdução de uma solução iterativa, envolvendo a matriz identidade \mathbf{I} e o coeficiente de combinação λ (fator de ajuste), cujo valor é alterado a cada iteração.

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} + \lambda_k \mathbf{I}, \quad (46)$$

onde k é o número de iteração e λ deve ser sempre positivo para que torne \mathbf{H} definida positivamente e inversível.

O método de Levenberg altera o método de Gauss-Newton através de uma solução aproximada de \mathbf{H} . Logo, a Equação (42) pode ser reescrita em termos da Jacobiana do vetor de erros.

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \left(\mathbf{J}^T \mathbf{J} + \lambda_k \mathbf{I} \right)^{-1} \mathbf{J}^T \mathbf{d}. \quad (47)$$

Caso λ seja muito pequeno ($\lambda \rightarrow 0$), emprega-se o método de Gauss-Newton. Entretanto, se λ for grande ($\lambda \rightarrow \infty$) utiliza-se o método do gradiente descendente com um pequeno tamanho de passo (Levenberg, 1944). Inicialmente, neste trabalho adotou-se $\lambda = 10^{-4}$, mas a cada iteração seu valor é alterado.

O método de Levenberg apresenta instabilidades quando λ cresce muito ($\lambda \rightarrow \infty$). Marquardt (1963) propôs um aperfeiçoamento no método de Levenberg (47) com a substituição da matriz identidade \mathbf{I} pela matriz diagonal dos elementos de $\mathbf{J}^T \mathbf{J}$.

$$\mathbf{s}_{k+1} = \mathbf{s}_k - \left[\mathbf{J}^T \mathbf{J} + \text{diag} \left(\mathbf{J}^T \mathbf{J} \right) \lambda_k \right]^{-1} \mathbf{J}^T \mathbf{d}, \quad (48)$$

Logo, solução da Equação (48) é definida como método LM (Levenberg, 1944; Marquardt, 1963; Moré, 1977).

Referências

- Brandão, T. & Queluz, M.P., No-reference quality assessment of H.264/AVC encoded video. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11):1437–1447, 2010.
- Brunnstrom, K.; Hands, D.; Speranza, F. & Webster, A., VQEG validation and ITU standardization of objective perceptual video quality metrics [standards in a nutshell]. *IEEE Signal Processing Magazine*, 26(3):96–101, 2009.
- Calyam, P.; Chandrasekaran, P.; Trueb, G.; Howes, N.; Ramnath, R.; Yu, D.; Liu, Y.; Xiong, L. & Yang, D., Multi-resolution multimedia QoE models for IPTV applications. *International Journal of Digital Multimedia Broadcasting*, 2012(ID 904072):1–13, 2012.
- Ding, W.; Tong, Y.; Zhang, Q. & Yang, D., Image and video quality assessment using neural network and SVM. *Tsinghua Science and Technology Journal*, 13(1):112–116, 2008.
- Eden, A., No-reference image quality analysis for compressed video sequences. *IEEE Transactions on Broadcasting*, 54(3):691–697, 2008.

- Engelke, U.; Kusuma, M.; Zepernick, H.J. & Caldera, M., Reduced-reference metric design for objective perceptual quality assessment in wireless imaging. *Signal Processing: Image Communication*, 24(7):525–547, 2009.
- Engelke, U. & Zepernick, H.J., An artificial neural network for quality assessment in wireless imaging based on extraction of structural information. In: *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing*.
- Engelke, U. & Zepernick, H.J., Perceptual-based quality metrics for image and video services: A survey. In: *Proc. EuroNGI Conference Next Generation Internet Networks*. p. 190–197, 2007.
- Eskicioglu, A.M. & Fisher, P.S., Image quality measures and their performance. *IEEE Transactions on Communications*, 43(12):2959–2965, 1995.
- Hemami, S.S. & Reibman, A.R., No-reference image and video quality estimation: applications and human-motivated design. *Signal Processing: Image Communication*, 25(7):469–481, 2010.
- Hoffman, J.D., *Numerical Methods for Engineers and Scientists*. 2a edição. New York, USA: Marcel Dekker, 2001.
- ITU-R, , *Methodology for the Subjective Assessment of the Quality for Television Pictures, Rev. 11*. Relatório Técnico, Geneva, Switzerland, 2004.
- ITU-T P.910, , *Subjective Video Quality Assessment Methods for Multimedia Applications*. Recommendation P.910, International Telecommunication Union, Geneva, Switzerland, 1999.
- Kawano, T.; Yamagishi, K.; Watanabe, K. & Okamoto, J., No reference video-quality-assessment model for video streaming services. In: *Proc. 18th International Packet Video Workshop (PV'10)*. p. 158–164, 2010.
- Kawayoke, Y. & Horita, Y., NR objective continuous video quality assessment model based on frame quality measure. In: *Proc. of 15th IEEE International Conference on Image Processing (ICIP'08)*. 2008.
- Keimel, C.; Habigt, J.; Klimpke, M. & Diepold, K., Design of no-reference video quality metrics with multiway partial least squares regression. In: *Proc. of the Third International Workshop on Quality of Multimedia Experience (QoMEX)*. p. 49–54, 2011a.
- Keimel, C.; Klimpke, M.; Habigt, J. & Diepold, K., No-reference video quality metric for HDTV based on H.264/AVC bitstream features. In: *Proc. of the 18th IEEE International Conference on Image Processing (ICIP'11)*. p. 3325–3328, 2011b.

- Keimel, C.; Oelbaum, T. & Diepold, K., No-reference video quality evaluation for high-definition video. In: *Proc. of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'09)*. p. 1145–1148, 2009.
- Kipli, K.; Muhammad, M.S.; Masra, S.M.W.; Zamhari, N.; Lias, K. & Mat, D.A.A., Performance of Levenberg-Marquardt backpropagation for full reference hybrid image quality metrics. In: *Proc. International Multiconference of Engineers and Computer Scientists*. v. I, p. 704–707, 2012.
- Le Callet, P. & Barba, D., Image quality assessment: From sites errors to a global appreciation of quality. In: *Proc. of the Picture Coding Symposium*. p. 105–108, 2001.
- Le Callet, P.; Viard-Gaudin, C. & Barba, D., A convolutional neural network approach for objective video quality assessment. *IEEE Transactions on Neural Networks*, 17(5):1316–1327, 2006.
- Levenberg, K., A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics*, 2:164–168, 1944.
- Marquardt, D.W., An algorithm for least-squares estimation of nonlinear parameters. *SIAM Journal on Applied Mathematics*, 11(2):431–441, 1963.
- Miyahara, M.; Kotani, K. & Algazi, V.R., Objective picture quality scale (PQS) for image coding. *IEEE Transactions on Communications*, 46(9):1215–1226, 1998.
- Moré, J., The Levenberg-Marquardt algorithm: Implementation and theory. In: Watson, G.A. (Ed.), *Numerical Analysis*. Berlin, Germany: Springer, v. 630 de *Lecture Notes in Mathematics*, p. 105–116, 1977.
- Oelbaum, T.; Keimel, C. & Diepold, K., Rule-based no-reference video quality evaluation using additionally coded videos. *IEEE Journal of Selected Topics in Signal Processing*, 3(2):294–303, 2009.
- Pappas, T.N. & Safranek, R.J., Perceptual criteria for image quality evaluation. In: *Handbook of Image and Video Processing*. New York, USA: Academic Press, p. 669–684, 2000.
- Parvez Sazzad, Z.M.; Kawayoke, Y. & Horita, Y., No reference image quality assessment for JPEG2000 based on spatial features. *Signal Processing: Image Communication*, 23(4):257–268, 2008.
- Rao, K.R. & Yip, P., *Discrete Cosine Transform: Algorithms, Advantages, Applications*. San Diego, USA: Academic Press Professional, 1990.
- Ries, M.; Kubanek, J. & Rupp, M., Video quality estimation for mobile streaming applications with neuronal networks. In: *Proc. 5th International Conference on Measurement of Audio and Video Quality in Networks (MESAQIN'06)*. Prague, Czech Republic, n. 5164913, 2006.

- Salomon, D., *Data Compression: The Complete Reference*. 4a edição. Northridge, CA: Springer, 2007.
- Seshadrinathan, K.; Soundararajan, R.; Bovik, A.C. & Cormack, L.K., Study of subjective and objective quality assessment of video. *IEEE Transactions on Image Processing*, 19:1427–1441, 2010.
- Shahid, M.; Rossholm, A. & Lovstrom, B., A reduced complexity no-reference artificial neural network based video quality predictor. In: *Proc. 4th International Congress on Image and Signal Processing (CISP'11)*. v. 1, p. 517–521, 2011.
- Shi, Y.; Ding, Y.; Zhang, R. & Li, J., Structure and hue similarity for color image quality assessment. In: *Proc. International Conference on Electronic Computer Technology*. p. 329–333, 2009.
- Silva, W.B. & Pohl, A.A.P., No-reference video quality assessment method based on the Levenberg-Marquardt minimization. In: *Proc. of XXX Brazilian Symposium on Telecommunications (SBTr'12)*. Brasília, Brazil, 2012.
- Silva, W.B.; Pohl, A.A.P. & Fonseca, K.V.O., A reduced-reference video quality assessment method based on the activity-difference of DCT coefficients. *IEICE Transactions on Information and Systems*, E96-D(3):708–718, 2013.
- Slanina, M.; Ricny, V. & Forchheimer, R., A novel metric for H.264/AVC no-reference quality assessment. In: *Proc. of 14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference Focused on Speech and Image Processing, Multimedia Communications and Services*. p. 114–117, 2007.
- Song, X. & Yang, Y., A new no-reference assessment metric of blocking artifacts based on HVS masking effect. In: *Proc. of the 2nd International Congress on Image and Signal Processing (CISP'09)*. p. 1–6, 2009.
- Staelens, N.; Vercammen, N.; Dhondt, Y.; Vermeulen, B.; Lambert, P.; Van de Walle, R. & Demeester, P., ViQID: a no-reference bit stream-based visual quality impairment detector. In: *Proc. of the Second International Quality of Multimedia Experience Workshop (QoMEX)*. p. 206–211, 2010.
- Sugimoto, O.; Naito, S.; Sakazawa, S. & Koike, A., Objective perceptual video quality measurement method based on hybrid no reference framework. In: *Proc. 16th IEEE International Conference on Image Processing (ICIP'09)*. Piscataway, USA: IEEE Press, p. 2237–2240, 2009.
- Tukey, J.W., *Exploratory Data Analysis*. New York, USA: Addison-Wesley, 1977.

- VQEG, , *Validation of Objective Models of Video Quality Assessment*. Final Report, Video Quality Experts Group (VQEG), 2000.
- VQEG, , *Validation of Objective Models of Multimedia Quality Assessment, Phase I*. Final Report, Video Quality Experts Group (VQEG), 2008.
- VQEG, , *Validation of Reduced-reference and No-reference Objective Models for Standard Definition Television, Phase I*. Final Report, Video Quality Experts Group (VQEG), 2009.
- VQEG, , *Validation of Video Quality Models for High Definition Video Content*. Technical Report, Version 2.0, Video Quality Experts Group (VQEG), 2010. <http://www.its.bldrdoc.gov/vqeg/projects/hdtv/>.
- Wang, A.; Jiang, G.; Wang, X. & Yu, M., New no-reference blocking artifacts metric based on human visual system. In: *Proc. of the International Conference on Wireless Communications Signal Processing (WCSP'09)*. p. 1–5, 2009.
- Wang, Z. & Bovik, A.C., A universal image quality index. *IEEE Signal Processing Letters*, 9(3):81–84, 2002.
- Wang, Z. & Bovik, A.C., *Modern Image Quality Assessment*. Synthesis Lectures on Image, Video, and Multimedia Processing. San Rafael, USA: Morgan & Claypool Publishers, 2006.
- Wang, Z.; Bovik, A.C. & Lu, L., Why is image quality assessment so difficult? In: *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'02)*. Piscataway, USA: IEEE Press, v. 4, p. IV–3313–IV–3316, 2002a.
- Wang, Z.; Bovik, A.C.; Sheikh, H.R. & Simoncelli, E.P., Image quality assessment: from error visibility to structural similarity. *IEEE Signal Processing Letters*, 13(4):600–612, 2004a.
- Wang, Z.; Lu, L. & Bovik, A.C., Video quality assessment based on structural distortion measurement. *Signal Processing: Image Communication*, 19(2):121–132, 2004b.
- Wang, Z.; Sheikh, H.R. & Bovik, A.C., No-reference perceptual quality assessment of JPEG compressed images. In: *Proc. of the IEEE International Conference on Image Processing (ICIP'02)*. Piscataway, USA: IEEE Press, v. 1, p. I–477–I–480, 2002b.
- Wang, Z.; Sheikh, H.R. & Bovik, A.C., Objective video quality assessment. In: Furht, B. & Marques, O. (Eds.), *The Handbook of Video Databases: Design and Applications*. Boca Raton, USA: CRC Press, p. 1041–1078, 2003a.
- Wang, Z.; Simoncelli, E.P. & Bovik, A.C., Multiscale structural similarity for image quality assessment. In: *Proc. of the 37th IEEE Asilomar Conference on Signals, Systems, and Computers*. Pacific Grove, USA: IEEE Computer Society, v. 2, p. 1398–1402, 2003b.

- Widrow, B. & Hoff Jr., M.E., Adaptive switching circuits. *IRE WESCON Convention Record*, 4:96–104, 1960.
- Wu, H.R.; Rao, K.R. & Kassim, A.A., Digital video image quality and perceptual coding. *Journal of Electronic Imaging*, 16(3):039901, 2007.
- Yamagishi, K.; Okamoto, J.; Hayashi, T. & Takahashi, A., No reference video-quality-assessment model for monitoring video quality of IPTV services. *IEICE Transactions*, 95-B(2):435–448, 2012.
- Yang, F.; Wan, S.; Chang, Y. & Wu, H.R., A novel objective no-reference metric for digital video quality assessment. *IEEE Signal Processing Letters*, 12(10):685–688, 2005.
- Yang, F.; Wan, S.; Xie, Q. & Wu, H.R., No-reference quality assessment for networked video via primary analysis of bit stream. *IEEE Transactions on Circuits and Systems for Video Technology*, 20(11):1544–1554, 2010.
- Yao, J.; Zhang, Y.; Xu, G. & Jin, M., No-reference objective quality assessment for video communication services based on feature extraction. In: *Proc. of the 2nd International Congress on Image and Signal Processing (CISP'09)*. p. 1–6, 2009.

Novas Técnicas e Ferramentas Aplicadas ao Sistema Brasileiro de TV Digital

A transmissão de sinais de vídeo e áudio por meio de radiodifusão apresenta um rico e longo histórico de desenvolvimento. O Brasil vive um momento de transição do sistema analógico para o sistema digital de transmissão de TV. Até 2016 a nova tecnologia deve estar implantada em todo o território nacional com o subsequente desligamento da transmissão analógica. Dentro desse contexto, inúmeros projetos e técnicas têm sido desenvolvidos e implantados. Este livro apresenta uma contribuição ao desenvolvimento tecnológico do sistema brasileiro de televisão digital, em que uma ferramenta de previsão de desempenho para o sistema de transmissão, um método de previsão e medição da propagação de sinais de TV digital, mecanismos de alocação de canais em redes IEEE 802.11b/g para aplicação no canal de retorno e métricas para avaliação de qualidade de vídeo digital são apresentados e discutidos.



Agência Brasileira do ISBN
ISBN 978-85-64619-14-2

