



**UNIVERSIDADE DO ESTADO DO RIO DE JANEIRO**

Centro de Tecnologia e Ciências

Faculdade de Engenharia

Renata Caminha Coelho de Souza

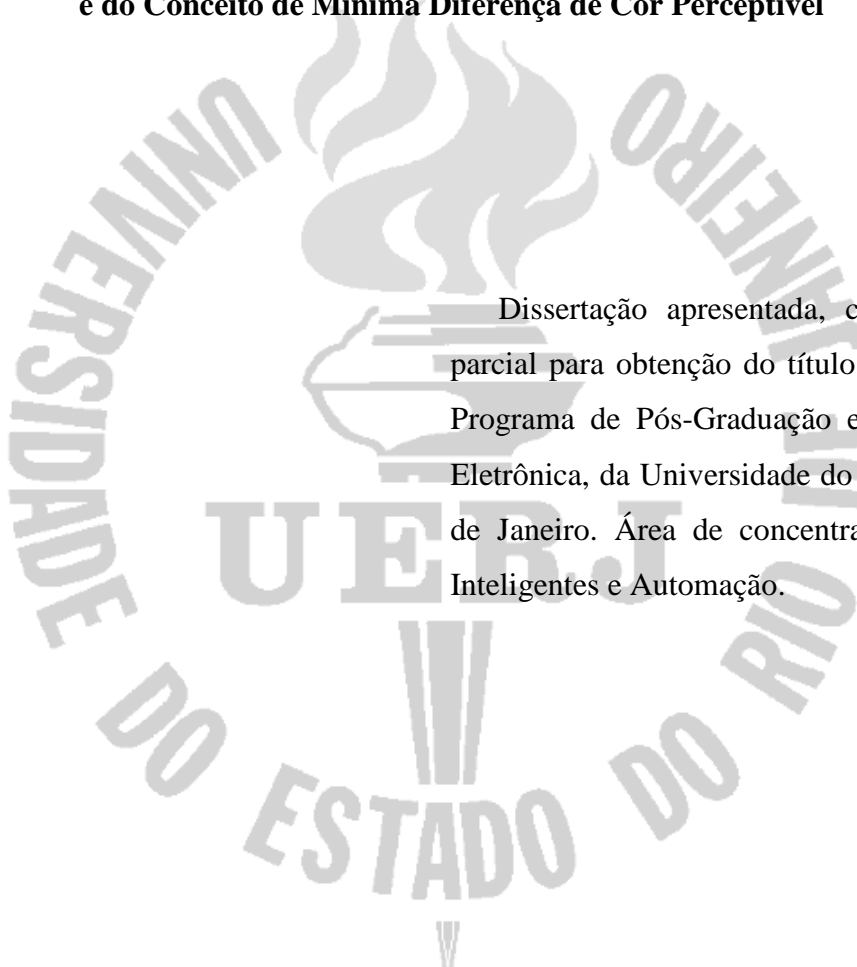
**Avaliação de Imagens Através de Similaridade Estrutural  
e do Conceito de Mínima Diferença de Cor Perceptível**

Rio de Janeiro

2009

Renata Caminha Coelho de Souza

**Avaliação de Imagens Através de Similaridade Estrutural  
e do Conceito de Mínima Diferença de Cor Perceptível**



Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

Orientador: Prof. D.Sc. Lisandro Lovisolo

Rio de Janeiro

2009

CATALOGAÇÃO NA FONTE  
UERJ/REDE SIRIUS/CTC/B

S719 Souza, Renata Caminha Coelho de.  
Avaliação de Imagens Através de Similaridade  
Estrutural e do Conceito de Mínima Diferença de Cor  
Perceptível / Renata Caminha Coelho de Souza. – 2009.  
117 f.: il.

Orientador: Lisandro Lovisolo.

Dissertação (Mestrado) – Universidade do Estado do Rio de  
Janeiro, Faculdade de Engenharia.

Bibliografia: f.112-115

1. Imagem. 2. Transformada de imagem. 3. Imagens coloridas. 4.  
Qualidade de imagem. I. Lovisolo, Lisandro. II. Universidade do  
Estado do Rio de Janeiro. Faculdade de Engenharia. III. Título.

CDU 621.397

Autorizo, apenas para fins acadêmicos e científicos, a reprodução total ou parcial desta tese.

Renata Souza

Assinatura

20/ outubro / 2009

Data

Renata Caminha Coelho de Souza

**Avaliação de Imagens Através de Similaridade Estrutural  
e do Conceito de Mínima Diferença de Cor Perceptível**

Dissertação apresentada, como requisito parcial para obtenção do título de Mestre, ao Programa de Pós-Graduação em Engenharia Eletrônica, da Universidade do Estado do Rio de Janeiro. Área de concentração: Sistemas Inteligentes e Automação.

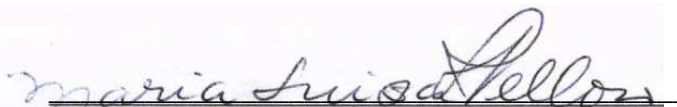
Aprovado em 20 de outubro de 2009.

Banca Examinadora:



---

Prof. D.Sc. Lisandro Lovisolo (Orientador)  
Faculdade de Engenharia da UERJ



---

Prof.ª D.Sc. Maria Luiza Fernandes Velloso  
Faculdade de Engenharia da UERJ



---

Prof. Ph.D. Eduardo Antônio Barros da Silva  
Faculdade de Engenharia da UFRJ

Rio de Janeiro  
2009

## **DEDICATÓRIA**

À minha família,  
com a qual eu aprendo valores como  
união, força, perseverança, amor e fé.

## AGRADECIMENTOS

A Deus, o Senhor, que me dotou de inteligência, perseverança e fé, e que permitiu que todo o contexto da minha vida fosse favorável para iniciar, desenvolver e concluir o mestrado.

Ao meu caríssimo professor orientador, Lisandro Lovisolo, por ser uma fonte inesgotável de idéias, por sua disposição e por sua confiança. Sempre foi um prazer estar em sua companhia durante as muitas horas de conversas, telefonemas, almoços, cafés, e principalmente de muito trabalho.

Ao meu namorado, Stefano Bertuola, pelo seu incentivo, seu apoio, sua cobrança e sua paciência. Também sou grata por suas opiniões e críticas, que certamente fizeram este trabalho muito mais colorido.

À minha mãe, Elza Caminha, pelo apoio e pelo amor incondicionais.

Aos queridos companheiros de trabalho no mestrado, Felipe Henriques e Pedro Arlindo Barroso Hardman Vianna que, compartilhando dos mesmos desafios, brilhavam para mim como faróis de portos seguros e serviram como inspiração quando o desânimo e o cansaço pareciam grandes demais.

A todos aqueles que apoiaram minha empreitada no mestrado e que agora podem se alegrar comigo por essa conquista.

Disciplina é liberdade.

*Kant*

## RESUMO

SOUZA, Renata Caminha Coelho de. Avaliação de imagens através de Similaridade Estrutural e do conceito de Mínima Diferença de Cor Perceptível. 2009. 115f. Dissertação (Mestrado em Engenharia Eletrônica) – Faculdade de Engenharia, Universidade do Estado do Rio de Janeiro, Rio de Janeiro, 2009.

A avaliação objetiva da qualidade de imagens é de especial importância em diversas aplicações, por exemplo na compressão de imagens, onde pode ser utilizada para regular a taxa que deve ser empregada para que haja a máxima compressão (permitindo perda de dados) sem comprometer a qualidade final; outro exemplo é na inserção de marcas d'água, isto é, introdução de informações descritivas utilizadas para atestar a autenticidade de uma imagem, que devem ser 'invisíveis' para o observador. O SSIM (*Structural SIMilarity*) é uma métrica de avaliação objetiva da qualidade de imagens de referência completa projetada para imagens em tons de cinza. Esta dissertação investiga sua aplicação na avaliação de imagens coloridas. Para tanto, inicialmente é feito um estudo do SSIM utilizando quatro diferentes espaços de cores – RGB, YCbCr,  $L\alpha\beta$  e CIELAB. O SSIM é primeiramente calculado nos canais individuais desses espaços de cores. Em seguida, com inspiração no trabalho desenvolvido em (1) são testadas formas de se combinar os valores SSIM obtidos para cada canal em um valor único – os chamados SSIM Compostos. Finalmente, a fim de buscar melhores correlações entre SSIM e avaliação subjetiva, propomos a utilização da mínima diferença de cor perceptível, calculada utilizando o espaço de cores CIELAB, conjuntamente com o SSIM. Para os testes são utilizados três bancos de dados de imagens coloridas, LIVE, IVC e TID, a fim de se conferir consistência aos resultados. A avaliação dos resultados é feita utilizando as métricas empregadas pelo VQEG (*Video Quality Experts Group*) para a avaliação da qualidade de vídeos, com uma adaptação. As conclusões do trabalho sugerem que os melhores resultados para avaliação da qualidade de imagens coloridas usando o SSIM são obtidas usando os canais de luminância dos espaços de cores YCbCr,  $L\alpha\beta$  e especialmente o CIELAB. Também se concluiu que a utilização da mínima diferença de cor perceptível contribui para o melhoramento dos resultados da avaliação objetiva.

Palavras-chaves: Avaliação da qualidade de imagens. SSIM. Mínima Diferença de Cor Perceptível (JND). CIELAB.



## ABSTRACT

Objective image quality evaluation is of special interest in many image applications, for example for image compression, where it can be used to control the rate in order to keep a tradeoff between lost of data and image quality; another example is in the application of watermarks, i.e., introduction of descriptive information used to guarantee the authenticity of an image, that must be invisible to the observer who looks at the image. SSIM (*Structural SIMilarity*) index is a full-reference image quality assessment metric developed to evaluate gray images. This work investigates the application of SSIM in the evaluation of color images. Therefore, four different color spaces are tested – RGB, YCbCr, L $\alpha\beta$  e CIELAB. Initially SSIM is calculated individually for each one of color spaces channels. Then, inspired in (1), the results of the SSIM in the individual channels are combined in a unique result – the so called Composite SSIM. Finally, in order to improve the correlations between, calculated using CIELAB color space, together with SSIM. Three color image databases, LIVE, IVC and TID, were employed in the tests in order to confer solidity to the results. The evaluation of the results is made using VQEG (*Video Quality Experts Group*) methodology, developed for video quality evaluation with an adaptation regarding the time “dimension” that does not exist in the image domain. The conclusions from the work were that SSIM performs better in the evaluation of color images when applied to luminance channel of YCbCr, L $\alpha\beta$  and especially to CIELAB color spaces. It was also concluded that the use of just noticeable difference concept improve objective assessment results.

Keywords: Image quality assessment. SSIM. Just Noticeable Difference (JND). CIELAB.

## LISTA DE FIGURAS

Figura 1 – Exemplos de aplicação de avaliação objetiva da qualidade de imagens.....	17
Figura 2 – Imagens de referência usadas no banco de dados LIVE.....	27
Figura 3 – Exemplo de imagem comprimida com JPEG2000.....	28
Figura 4 – Exemplo de imagem comprimida com JPEG.....	29
Figura 5 – Exemplo de imagem adicionada de ruído branco nos componentes RGB.....	29
Figura 6 – Exemplo de imagem com embaçamento Gaussiano.....	30
Figura 7 – Exemplo de imagem com erro de bit devido à transmissão por um canal sem-fio.	30
Figura 8 – Imagens de referência usadas no banco de dados IVC.....	31
Figura 9 – Imagens de referência usadas no banco de dados TID.....	32
Figura 10 – Percepção da sensibilidade ao contraste de acordo com a frequência espacial. ...	37
Figura 11 – Diagrama esquemático da função de dispersão de ponto que atua na imagem adquirida pelo sistema visual humano. Inspirado em (31). .....	39
Figura 12 – Estrutura da computação do SSIM.....	39
Figura 13 – Separação das distorções de luminância, de contraste e estrutural da imagem de referência I. Inspirado em (26). .....	40
Figura 14 – Filtro Gaussiano com desvio padrão = 1,5 e janela = 11, e sua resposta em frequência. ....	45
Figura 15 – Comparação entre SSIM e MSE para imagens degradadas. Inspirado em (26). ..	47
Figura 16 – Distribuição dos fotorreceptores na retina em relação à fóvea. Fonte: (35). .....	49
Figura 17 – Curvas normalizadas das sensibilidades dos fotorreceptores. Fonte: (35).....	51
Figura 18 – Diagrama de cromaticidade xy do espaço CIE XYZ. Fonte: (42). .....	52
Figura 19 – Espaço de cores sRGB. Fonte: (14). .....	53
Figura 20 – Cubo sRGB representado nos eixos YCbCr. ....	54

Figura 21 – Elipses de MacAdam representadas sobre o diagrama de cromaticidade xy do espaço de cores CIE XYZ. Fonte: (48).....	57
Figura 22 – Cubo sRGB representado nos eixos CIELAB em diferentes perspectivas. Fonte: (14).....	58
Figura 23 – Diagrama de blocos da aplicação do $\Delta E$ juntamente ao SSIM na avaliação da qualidade de imagens.....	59
Figura 24 – Mapeamento dos dados objetivos para a escala subjetiva. ....	60
Figura 25 – Diagrama de blocos do SSIM aplicado aos canais individuais.....	66
Figura 26 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores RGB, usando o banco de dados LIVE. ....	66
Figura 27 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores YCbCr.....	67
Figura 28 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores $L\alpha\beta$ .....	68
Figura 29 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais dos espaços de cores RGB, YCbCr, $L\alpha\beta$ e CIELAB.....	69
Figura 30 – Curvas de ajuste do SSIM aplicado aos canais individuais nos bancos de dados IVC, TID e LIVE.....	70
Figura 31 – Correlações entre MOS/DMOS e SSIM com ajuste aplicado aos canais individuais dos espaços de cores. ....	71
Figura 32 – Gráficos de dispersão do banco de dados LIVE identificados por distorção.....	74
Figura 33 – Curva de ajuste por distorção no canal L do CIELAB.....	75
Figura 34 – Métricas de análise para a aplicação das três componentes do SSIM individualmente e combinadas duas a duas.....	78
Figura 35 – Diagrama de blocos do cálculo do SSIM Composto. ....	80
Figura 36 – Curvas de ajuste dos SSIM Compostos para cada espaço de cores usando o banco de dados LIVE.....	83

Figura 37 – Correlações entre DMOS e SSIM com ajuste no banco de dados LIVE para os cálculos dos SSIM Compostos e SSIM dos canais individuais com maior correlação.....	84
Figura 38 – Comparação entre as melhores curvas de ajuste utilizando os canais individuais e SSIM Compostos .....	86
Figura 39 – Diagrama de blocos da utilização do $\Delta E$ conjuntamente ao SSIM filtrado para a primeira abordagem, cálculo do MSSIM como média dos pixels onde $\Delta E > JND$ . .....	89
Figura 40 – Curvas de correlação entre SSIM e DMOS por JND, no banco de dados LIVE..	90
Figura 41 – RMSE, OR e Percentis dos SSIM Compostos e SSIML após ajuste.....	90
Figura 42 – Diagrama de blocos da utilização do $\Delta E$ filtrado conjuntamente ao SSIM filtrado para a primeira abordagem. ....	92
Figura 43 – Diagrama de blocos da implementação do SSIM sem filtragem para a primeira abordagem.....	93
Figura 44 – Curvas de correlação dos testes da aplicação da filtragem no mapa SSIM e no mapa $\Delta E$ para a primeira abordagem da aplicação do $\Delta E$ conjuntamente ao SSIM. ....	94
Figura 45 – Métricas de avaliação da primeira abordagem para o banco de dados LIVE. ....	95
Figura 46 – Métricas de avaliação da primeira abordagem para o banco de dados IVC. ....	95
Figura 47 – Métricas de avaliação da primeira abordagem para o banco de dados TID.....	96
Figura 48 – Diagrama de blocos da utilização do $\Delta E$ conjuntamente ao SSIM filtrado para a segunda abordagem, com substituição das regiões da imagem distorcida menores que $\Delta E$ pelas equivalentes da imagem de referência.....	98
Figura 49 – Diagrama de blocos da utilização do $\Delta E$ filtrado conjuntamente ao SSIM filtrado para a segunda abordagem.....	99
Figura 50 – Diagrama de blocos da implementação do SSIM sem filtragem para a segunda abordagem.....	100
Figura 51 – Curvas de correlação dos testes da aplicação da filtragem no mapa SSIM e no mapa $\Delta E$ para a segunda abordagem da aplicação do $\Delta E$ conjuntamente ao SSIM. ....	102

Figura 52 – Métricas de avaliação da segunda abordagem para o banco de dados LIVE.....	102
Figura 53 – Métricas de avaliação da segunda abordagem para o banco de dados IVC.....	103
Figura 54 – Métricas de avaliação da segunda abordagem para o banco de dados TID.....	103
Figura 55 – Comparação entre os valores de pico das correlações e correlações em JND = 0 para todos os testes das duas abordagens de aplicação do $\Delta E$ conjuntamente ao SSIM. ....	105
Figura 56 – Melhoria percentual com a utilização da distância entre cores para o melhor caso de cada experimento .....	106

## LISTA DE TABELAS

Tabela 1– Conceito para avaliação utilizando o método DSIS .....	22
Tabela 2 – Conceito para avaliação utilizando o método SS .....	23
Tabela 3 – Conceito para avaliação utilizando o método de Comparação de Estímulos .....	24
Tabela 4– Medidas de avaliação do SSIM aplicado aos canais dos espaços de cores .....	72
Tabela 5– Métricas de análise por tipo de distorção .....	76
Tabela 6 – Valores obtidos pelas métricas de análise para componentes e combinações SSIM .....	79
Tabela 7– Medidas de avaliação dos SSIM Compostos nos espaços CIELAB e $L\alpha\beta$ .....	84
Tabela 8– Medidas de avaliação dos SSIM Compostos nos espaços RGB e YCbCr .....	85
Tabela 9 – Melhores correlações com JND correspondente .....	91
Tabela 10 – Comparação entre os resultados obtidos pelo JND que apresentou maior correlação em cada experimento e os resultados sem utilização do $\Delta E$ para a primeira abordagem .....	97
Tabela 11 – Comparação entre os resultados obtidos pelo JND que apresentou maior correlação em cada experimento e os resultados sem utilização do $\Delta E$ com a segunda abordagem .....	104
Tabela 12 – Valores JND dos picos de correlação .....	107

## SUMÁRIO

<b>INTRODUÇÃO</b> .....	16
<b>Métricas subjetivas</b> .....	17
<b>Métricas objetivas</b> .....	17
<b>Objetivos desta dissertação</b> .....	18
<b>Organização</b> .....	19
<b>1 AVALIAÇÃO DA QUALIDADE DE IMAGENS</b> .....	21
1.1 <b>Métodos subjetivos</b> .....	21
1.1.1 <u>Método DSIS (Double Stimulus Impairment Scale)</u> .....	22
1.1.2 <u>Método DSCQS (Double Stimulus Continuous Quality Scale)</u> .....	22
1.1.3 <u>Método SS (Single Stimulus)</u> .....	23
1.1.4 <u>Método de Comparação de Estímulos (Stimulus Comparison Method)</u> .....	23
1.2 <b>Cálculo das médias</b> .....	24
1.2.1 <u>Cálculo do MOS</u> .....	25
1.2.2 <u>Cálculo do DMOS</u> .....	25
1.3 <b>Bancos de Dados</b> .....	26
1.3.1 <u>Banco de dados LIVE</u> .....	27
1.3.2 <u>Banco de dados IVC</u> .....	31
1.3.3 <u>Banco de dados TID</u> .....	32
1.4 <b>Métodos objetivos</b> .....	33
<b>2 O ÍNDICE SSIM</b> .....	36
2.1 <b>Percepção no sistema visual humano</b> .....	36
2.1.1 <u>Sensibilidade ao contraste</u> .....	36
2.1.2 <u>Percepção da luz e luminância</u> .....	37
2.1.3 <u>Estrutura</u> .....	38

2.1.4	<u>Dispersão de ponto</u> .....	38
2.2	<b>O SSIM</b> .....	39
2.2.1	<u>Interpretação geométrica</u> .....	40
2.2.2	<u>Formulação</u> .....	41
2.2.3	<u>Média SSIM</u> .....	44
2.2.4	<u>Filtro suavizante</u> .....	45
2.3	<b>MSE versus SSIM</b> .....	46
3	<b>ESPAÇOS DE CORES</b> .....	48
3.1	<b>Percepção da cor</b> .....	48
3.2	<b>Imagens digitais e espaços de cores</b> .....	50
3.3	<b>Espaço RGB</b> .....	53
3.4	<b>Espaço YCbCr</b> .....	53
3.5	<b>Espaço <math>L\alpha\beta</math></b> .....	55
3.6	<b>Espaço CIELAB</b> .....	56
3.7	<b>Utilização do <math>\Delta E</math> na avaliação da qualidade de imagens</b> .....	58
4	<b>METODOLOGIA DE ANÁLISE</b> .....	60
4.1	<b>Mapeamento para a escala subjetiva</b> .....	61
4.2	<b>Correlação de Pearson</b> .....	62
4.3	<b>RMSE</b> .....	62
4.4	<b>Razão Outlier</b> .....	63
4.5	<b>Percentis 95 e 99</b> .....	64
5	<b>COMPORTAMENTO DO SSIM EM DIFERENTES ESPAÇOS DE CORES</b> ..	65
5.1	<b>SSIM aplicado aos canais individuais</b> .....	65
5.1.1	<u>Espaço RGB</u> .....	66
5.1.2	<u>Espaço YCbCr</u> .....	67
5.1.3	<u>Espaço <math>L\alpha\beta</math></u> .....	68
5.1.4	<u>Espaço CIELAB</u> .....	69



5.1.5	<u>Comparação entre os espaços</u> .....	69
5.2	<b>Análise do SSIM por tipo de distorção</b> .....	73
5.2.1	<u>Curvas de ajuste por distorções</u> .....	75
5.3	<b>Combinações alternativas das componentes do SSIM</b> .....	77
6	<b>SSIM COMPOSTOS</b> .....	80
6.1	<b>Formulação dos SSIM Compostos</b> .....	81
6.2	<b>Análise dos resultados dos SSIM Compostos</b> .....	82
7	<b>APLICAÇÃO DA DISTÂNCIA ENTRE CORES NO CIELAB AO SSIM</b> .....	87
7.1	<b>SSIM como média dos pixels para os quais <math>\Delta E &gt; JND</math></b> .....	88
7.1.1	<u>Filtragem SSIM</u> .....	89
7.1.2	<u>Filtragem SSIM e <math>\Delta E</math></u> .....	91
7.1.3	<u>Sem filtragem</u> .....	92
7.1.4	<u>Resultados da primeira abordagem</u> .....	93
7.2	<b>SSIM como média de todos os pixels substituindo <math>\Delta E &lt; JND</math> pela referência</b> .....	98
7.2.1	<u>Filtragem SSIM</u> .....	98
7.2.2	<u>Filtragem SSIM e <math>\Delta E</math></u> .....	99
7.2.3	<u>Sem filtragem</u> .....	100
7.2.4	<u>Resultados da segunda abordagem</u> .....	100
7.3	<b>Comparação entre os resultados das abordagens</b> .....	105
8	<b>CONCLUSÃO</b> .....	109
8.1	<b>Observações e sugestões para trabalhos futuros</b> .....	111
	<b>REFERÊNCIAS</b> .....	112

## INTRODUÇÃO

Imagens digitais que são trocadas e distribuídas através de redes de comunicações digitais estão sujeitas a diversos tipos de distorções durante a aquisição, processamento, compressão, transmissão e reprodução das mesmas. Essas distorções podem reduzir a qualidade no momento de exibição. As causas de distorção podem se originar em diversas etapas do processo; como exemplos, a câmera que captura a imagem pode introduzir distorções devido à sua ótica, a ruídos, à calibração das cores, ao controle de exposição, ao movimento da câmera, etc. (2)

Após a captura, as imagens em geral são processadas por um algoritmo de compressão com perdas para fins de armazenamento ou transmissão. Tais algoritmos geralmente são projetados para produzir compressões poupando espaço de armazenamento com o custo de permitir que ocorram distorções controladas no sinal imagem. De forma similar, erros de transmissão de dados que ocorrem devido a canais inerentemente falhos, como canais sem-fio, ou mais raramente em seu armazenamento, também tendem a introduzir distorções (3). Além dessas formas de distorção, o dispositivo utilizado para exibição ao usuário final pode introduzir suas próprias distorções, como baixa resolução de reprodução, calibração ruim, (2) etc. A quantidade de distorção que cada um desses estágios pode acarretar depende em maior parte da economia de banda e /ou da limitação física dos dispositivos.

É óbvio o interesse em se medir a qualidade de imagens e avaliar a distorção após os diferentes estágios. Uma métrica objetiva da qualidade da imagem pode desempenhar um importante papel em uma ampla gama de aplicações como na aquisição de imagens, compressão, comunicação, exibição, impressão, restauração, melhoramentos, análise e aplicação de marcas d'água (2).

A avaliação objetiva de imagens, de forma sucinta, pode ser usada em dois contextos: para *monitorar* dinamicamente e ajustar a qualidade da imagem e para *otimizar* algoritmos e ajustar parâmetros do sistema de processamento de imagens. A Figura 1 traz uma representação esquemática da avaliação objetiva aplicada à compressão de imagens e à inserção de marcas d'água, inserindo informações descritivas não perceptivas em uma imagem, que são utilizadas para atestar sua autenticidade proporcionando, por exemplo, proteção de direitos autorais.

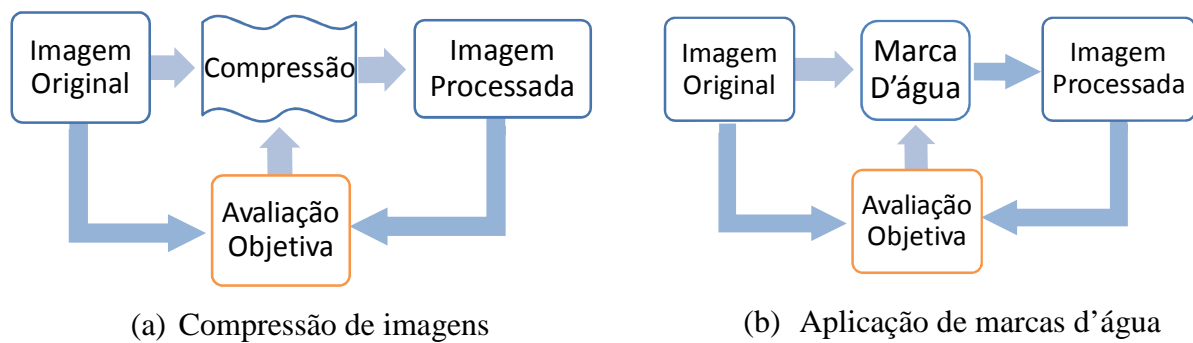


Figura 1 – Exemplos de aplicação de avaliação objetiva da qualidade de imagens.

### Métricas subjetivas

A maneira mais confiável de avaliar a qualidade de imagens é por meio de avaliação subjetiva, uma vez que observadores humanos são os receptores finais para a maioria das aplicações. O MOS (*Mean Opinion Score*), uma medida de qualidade produzida com a participação de um grupo de observadores, é um método amplamente usado para avaliação subjetiva da qualidade (2). Seu cálculo consiste na média das notas dadas para uma determinada imagem por um grupo de observadores em condições controladas. Uma variação do MOS é o DMOS, *Differential Mean Opinion Score*, que calcula o grau de qualidade de uma imagem como a diferença entre a avaliação da imagem distorcida e a avaliação da imagem de referência (4). Para o DMOS é necessário que sejam avaliadas tanto a imagem distorcida quanto a correspondente imagem de referência.

Entretanto, para a maior parte das aplicações o método MOS é inconveniente, pois é lento e dispendioso, uma vez que emprega um grupo de pessoas no processo de avaliação. Além disso, tal método não é factível quando se deseja que técnicas de medição da qualidade de imagens estejam embarcadas.

### Métricas objetivas

Com o intuito de mitigar o problema da necessidade de pessoas no processo de avaliação da qualidade, foram propostas métricas para avaliação *objetiva* da qualidade como em (3), (5) e (1). Estas métricas, diferentemente das métricas subjetivas, são calculadas de forma automática por algoritmos, não necessitando da participação humana. Geralmente as métricas

objetivas são classificadas em três categorias com base na quantidade de informação da imagem original requerida para a avaliação da qualidade (3): de Referência Completa, Sem Referência e de Referência Reduzida.

Métricas de Referência Completa (3) (cujas siglas, FR, vem de seu nome em inglês *Full-Reference*) realizam uma comparação pixel-a-pixel entre a imagem de referência (a original) e a imagem sendo testada (a processada), requerendo que a referência seja completamente disponível. Métodos Sem-Referência (sigla NR de *No-Reference*) analisam a imagem processada sem a necessidade de uma referência explícita – estes métodos sempre necessitam fazer suposições sobre o conteúdo da imagem e/ou sobre as distorções de interesse (5). Métricas de Referência Reduzida (RR – *Reduced Reference*) são projetadas para ser um compromisso entre métricas FR e NR. As métricas RR utilizam certa quantidade de características da imagem de referência, como por exemplo detalhes espaciais, para executar a comparação com a imagem de teste com base nestas características (3).

Inicialmente, métodos estatísticos de avaliação de imagens FR como MSE (Erro Médio Quadrático) e PSNR (Razão Sinal-Ruído de Pico) foram muito usados (6) e o são ainda hoje, devido a suas simplicidades comportamentais: simetria, diferenciabilidade, não-negatividade e aditividade para fontes independentes (6). Entretanto, apesar de suas amplas presenças na literatura, não são bons métodos para avaliação da qualidade de imagens, pois seus resultados não possuem altas correlações com a percepção humana (6).

Métricas de avaliação da qualidade baseadas na percepção humana começaram a ser desenvolvidas, como em (7) e (8). Porém estas, em sua maior parte, se concentram na avaliação de imagens em escala de cinza, assumindo que a maior parte da informação visual se concentra no chamado canal de luminância (explicado no Capítulo 3), enquanto que os canais de cromaticidade, que carregam informação de cor, concentram menos informação. Apenas poucos estudos foram propostos para avaliar a influência da cor na qualidade de imagens (1).

### **Objetivos desta dissertação**

Este trabalho se propõe a analisar o desempenho do SSIM (*Structural SIMilarity*) (9), uma métrica FR inspirada em características da percepção humana, originalmente proposto para a avaliação objetiva da qualidade de imagens em escala de cinza, aplicado na avaliação de imagens coloridas.

Para avaliar imagens coloridas deve-se empregar um espaço de cores. É interessante que o espaço de cores empregado na avaliação produza resultados correlacionados àqueles obtidos pela percepção humana. Assim, de forma a investigar o espaço que fornece resultados mais correlacionados aos obtidos pela avaliação subjetiva, foram estudados os comportamentos de quatro diferentes espaços de cores: RGB (10), pois é comumente usado em *displays* coloridos e câmeras; YCbCr (11), uma vez que é o espaço usado pelo estado-da-arte em compressão de vídeo e imagens;  $L\alpha\beta$  (12), devido ao fato de que este espaço de cores foi utilizado em (1) com o SSIM na avaliação de imagens coloridas; e CIELAB (13), um espaço de cores perceptual desenvolvido pelo CIE (Comissão Internacional em Iluminação), projetado para se aproximar da forma como as cores são percebidas pela visão humana.

O espaço de cores CIELAB possui uma importante característica: a distância medida entre cores nesse espaço é uma boa aproximação da distância visual percebida (14). Desta forma, o limiar de percepção da cor conhecido como JND (*just noticeable difference*) (15) ou mínima diferença de cor perceptível, pode ser utilizado de forma a aumentar a correlação da avaliação objetiva com a avaliação subjetiva. Assim, a aplicação da distância entre cores conjuntamente com a métrica objetiva pode produzir resultados melhores, uma vez que se incorpora mais uma característica da percepção humana.

Para os testes, foram utilizados três bancos de dados produzidos por laboratórios envolvidos no estudo de imagens digitais. Esses bancos de dados contêm avaliações subjetivas de imagens: o banco de dados IVC (16) possui 130 imagens das quais 10 imagens são de referência e 120 são versões degradadas dessas 10 por meio de compressão JPEG, compressão JPEG2000 e embaçamento; o banco de dados LIVE (17) conta com um total de 808 imagens, das quais 29 são imagens de referência e 779 são versões degradadas por compressão JPEG2000, compressão JPEG, adição de ruído Gaussiano branco nos componentes RGB, embaçamento Gaussiano nos componentes RGB e erros de transmissão em imagens codificadas via JPEG2000; e o banco de dados TID2008 (18), que contém 25 imagens de referência, (24 das quais contidas também no banco de dados LIVE), porém contém 1700 imagens degradadas por 17 diferentes tipos de distorções.

## **Organização**

Esta dissertação está dividida em oito capítulos. O primeiro Capítulo trata da avaliação da qualidade de imagens, explicando alguns métodos subjetivos utilizados e investigando os

métodos objetivos mais comuns – MSE e PSNR, explicando por que estes não são bons métodos para avaliação objetiva da qualidade de imagens. Neste capítulo também é feita uma descrição dos bancos de dados de avaliações subjetivas usados nos testes. O segundo Capítulo introduz o índice SSIM, sua formulação e suas características. O terceiro apresenta os quatro espaços de cores propostos para avaliação de imagens coloridas – RGB, YCbCr,  $L\alpha\beta$  e CIELAB e suas características. O Capítulo 4 trata da metodologia de análise utilizada para avaliação dos resultados inspirada na metodologia usada pelo VQEG (*Video Quality Experts Group*) (4), para avaliação da qualidade de vídeo. O Capítulo 5 estuda o comportamento do SSIM para avaliação de imagens coloridas aplicado nos canais individuais dos espaços de cores sugeridos no Capítulo 3. No Capítulo 6 são propostas três formas de cálculo do SSIM considerando os três canais dos espaços de cores de forma conjunta – os chamados SSIM Compostos. O Capítulo 7 estuda a aplicação da distância entre cores calculada no espaço CIELAB conjuntamente ao SSIM a fim de melhorar os resultados da avaliação subjetiva via SSIM. O Capítulo 8 apresenta as conclusões bem como sugestões para futuros trabalhos.

## 1 AVALIAÇÃO DA QUALIDADE DE IMAGENS

A avaliação da qualidade de imagens é importante para dimensionar distorções que podem reduzir a qualidade no momento de exibição. Como descrito no capítulo anterior, existem duas formas de se obter a avaliação da qualidade de uma imagem: através de métodos subjetivos, isto é, por meio de notas produzidas por observadores humanos, ou através de métodos objetivos, onde algoritmos simulam o comportamento do sistema visual humano produzindo a avaliação da qualidade.

Este capítulo trata das metodologias de avaliação da qualidade de imagens e está dividido em quatro seções: métodos subjetivos, cálculo das médias, bancos de dados e métodos objetivos. Na primeira é apresentada uma descrição dos métodos de avaliação subjetiva empregados na construção dos bancos de dados utilizados neste trabalho. Na segunda são descritas as formas de cálculo das médias a partir das notas subjetivas usando MOS (*Mean Squared Error*) ou DMOS (*Difference Mean Squared Error*). Na terceira são apresentados três bancos de dados de avaliação subjetiva de imagens utilizados para comparação com a avaliação objetiva. Na quarta, duas métricas de avaliação objetivas – MSE e PSNR – são descritas, e são apresentados alguns motivos que sugerem que estas não são boas métricas de avaliação da qualidade de imagens.

### 1.1 Métodos subjetivos

A medida da qualidade percebida em imagens requer o uso de métodos subjetivos. A condição para tal medida apresentar significado é haver uma relação entre as características físicas do estímulo, nesse caso a imagem distorcida apresentada aos observadores em um teste, e a magnitude e a natureza da sensação causada pelo estímulo.

Diversos métodos experimentais para avaliação subjetiva foram criados voltados para diferentes propósitos (19). A escolha de um método para uma aplicação particular depende de fatores como o contexto, o propósito e onde, no processamento da imagem, o teste deve ser executado (20).

Os métodos usados na confecção ou citados nos bancos de dados deste trabalho – métodos DSIS (*Double Stimulus Impairment Scale*), DSCQS (*Double Stimulus Continuous Quality Scale*), SS (*Single Stimulus*) e o Método de Comparação de Estímulos, são descritos a seguir.

### 1.1.1 Método DSIS (Double Stimulus Impairment Scale)

Neste método as imagens sob teste são mostradas em seqüência. O observador é primeiramente apresentado à imagem de referência e, em seguida, à imagem processada. É solicitado ao observador que avalie o grau de distorção da imagem processada com relação à imagem de referência. A avaliação deve ser dada em termos de conceitos (ou notas) a respeito da degradação da imagem, seguindo a escala da Tabela 1 (20).

Tabela 1– Conceito para avaliação utilizando o método DSIS

<b>Conceito</b>	<b>Distorção</b>
<b>5</b>	Imperceptível
<b>4</b>	Perceptível, não incomodando
<b>3</b>	Ligeiramente irritante
<b>2</b>	Irritante
<b>1</b>	Muito irritante

### 1.1.2 Método DSCQS (Double Stimulus Continuous Quality Scale)

Nesse método o observador assiste as imagem de referência e distorcida. A ordem temporal das imagens durante as seqüências é alterada de forma pseudo-aleatória, e não é dado o conhecimento aos observadores sobre qual imagem é a referência. Ambas as imagens de cada seqüência devem ser avaliadas (20).

O número de repetições depende da duração do teste. O avaliador assiste a cada uma das imagens por aproximadamente dez segundos, de duas a três vezes, memoriza a percepção e então dá a nota correspondente a cada uma delas, entre os limites de 0 a 100.

Não é conveniente associar os números da pontuação dada neste método com os resultados e adjetivos do método anterior. Os resultados obtidos devem ser tratados apenas como a diferença de classificação numérica entre a referência e a imagem sendo avaliada (20). Desta forma, um número pequeno na escala DSCQS indica que a diferença de qualidade entre a imagem de referência e a imagem sendo avaliada é pequena, ou seja, há pouca degradação percebida. Porém, não há uma indicação da qualidade final da imagem. Para isso, é necessário um processamento das notas calculando-se o DMOS, descrito adiante.



### 1.1.3 Método SS (Single Stimulus)

No método de estímulo simples as imagens sob teste são apresentadas uma por vez e avaliadas independentemente com uma escala de categorias. Este método especifica que após cada apresentação o observador deve avaliar a qualidade da imagem exibida. Uma escala de cinco níveis de qualidade é em geral usada. A Tabela 2 apresenta os conceitos referentes a cada nível de qualidade que pode ser atribuído à imagem sendo testada (20). Caso haja necessidade de uma quantidade maior de conceitos de discriminação, uma escala de nove níveis pode ser usada (20).

Tabela 2 – Conceito para avaliação utilizando o método SS

<b>Conceito</b>	<b>Qualidade</b>
<b>5</b>	Excelente
<b>4</b>	Boa
<b>3</b>	Razoável
<b>2</b>	Pobre
<b>1</b>	Ruim

A diferença fundamental entre os métodos de estímulo simples (SS) e estímulo duplo (DSIS ou DSCQS) é que no método de estímulo simples o observador é apresentado à imagem sendo avaliada uma única vez enquanto que no método de estímulo duplo, a imagem pode ser apresentada duas ou mais vezes.

### 1.1.4 Método de Comparação de Estímulos (Stimulus Comparison Method)

Este é um método alternativo aos mais comuns DSIS e DSCQS. Neste método, duas seqüências de imagens de mesmo conteúdo – uma de referência e outra sob teste – são mostradas simultaneamente, e o observador deve dar uma nota correspondente à relação entre as imagens. Nos testes para televisão, há três metodologias de avaliação (20):

- Julgamento por adjetivos: a relação entre as imagens é medida com uso de palavras que expressam a existência de diferenças perceptíveis, de acordo com a Tabela 3.

Tabela 3 – Conceito para avaliação utilizando o método de Comparação de Estímulos

Conceito	Imagem sob teste comparada à imagem de referência
-3	Muito pior
-2	Pior
-1	Levemente pior
0	Igual
1	Levemente melhor
2	Melhor
3	Muito Melhor

- Julgamento sem categoria: a relação da qualidade observada é informada por um número, entre limites abertos ou pré-definidos, ou ainda pelo desenho de um ponto ou risco em uma escala vertical com dois limites rotulados nas pontas desta escala (exemplo: igual, diferente).
- Método do desempenho: cada imagem do par tem graus diferentes de um determinado sintoma, ou uma delas não o contém. O avaliador decide qual das imagens contém o problema, ou em qual delas ele está presente em maior grau. O acerto e velocidade na resposta são usados como índices da relação entre as duas imagens.

## 1.2 Cálculo das médias

Obtidas as avaliações subjetivas para uma imagem por um grupo de observadores, é necessário que seja feita uma média das notas a fim de se ter um único valor correspondente à qualidade da imagem. O MOS (*Mean Opinion Score*) e o DMOS (*Difference Mean Opinion Score*) são dois cálculos de médias utilizados nos bancos de dados usados nesse trabalho. Tanto o MOS quanto o DMOS podem ser usados para o cálculo da média. A diferença básica entre eles é que para o cálculo do DMOS é necessário que se tenha a avaliação subjetiva da imagem de referência explicitamente, enquanto que para o cálculo do MOS a nota explícita da imagem de referência não é necessária. Seus cálculos são descritos a seguir.

### 1.2.1 Cálculo do MOS

O MOS (*Mean Opinion Score*) fornece a média das notas obtidas dos observadores para uma imagem. Para seu cálculo é necessário que a imagem seja avaliada por um grupo de observadores através de qualquer dos métodos descritos anteriormente. O  $MOS_j$  fornece a média das notas de  $i$  observadores para a imagem  $j$  via:

$$MOS_j = \frac{1}{N} \sum_{i=1}^N m_{ij}, \quad (1)$$

onde  $m_i$  é a nota dada pelo observador  $i$  para a imagem  $j$  sendo avaliada e  $N$  é o número de observadores que avaliaram a imagem em questão. A escala do MOS segue a escala do método utilizado para a obtenção das notas.

O cálculo do MOS não abarca a relação entre a imagem de referência e a imagem distorcida. Esta é somente considerada na obtenção das notas no momento da avaliação subjetiva se o método empregado assim o fizer. Para considerar explicitamente a relação entre uma imagem distorcida e sua correspondente original, deve-se utilizar o DMOS conforme descrito a seguir.

### 1.2.2 Cálculo do DMOS

O DMOS (*Difference Mean Opinion Score*) é calculado pela média da diferença entre as notas da imagem de referência e da imagem modificada (4). Para sua computação é necessário que sejam fornecidas as notas da imagem de referência e da imagem distorcida. Quando o método DSCQS é utilizado, por exemplo, as notas de diferença podem ser computadas. Para o cálculo do DMOS as notas dos observadores são primeiramente convertidas em notas de diferença de qualidade a partir da seguinte fórmula:

$$d_{ij} = r_{ij} - r_{idis(j)}, \quad (2)$$

onde  $r_{ij}$  é a nota do  $i$ -ésimo observador para a  $j$ -ésima imagem original, e  $r_{idis(j)}$  é a nota de qualidade assumida pelo  $i$ -ésimo observador para a  $j$ -ésima imagem distorcida.

Como nem todos os observadores usam toda a faixa possível de valores na escala numérica, o que poderia gerar variabilidade, as notas cruas de diferença  $d_{ij}$  para o  $i$ -ésimo

observador e para a  $j$ -ésima imagem são convertidos em “Z-scores” (21), após a remoção dos *outliers* e dos observadores rejeitados, através de:

$$z_{ij} = \frac{d_{ij} - \bar{d}_i}{\sigma_i}, \quad (3)$$

onde  $\bar{d}_i = \frac{1}{N} \sum_{i=1}^N d_i$  é a média das notas de diferença de todas as imagens avaliadas pelo observador  $i$ ,  $\sigma_i$  é o desvio padrão das notas dado por  $\sigma_i = \frac{1}{(N-1)} \sum_{i=1}^N (d_i - \bar{d}_i)^2$  e  $d_i$  são as notas dadas pelo avaliador  $i$  para todas as imagens, onde  $i = 1, \dots, N$ . É então feita uma média dos observadores para se ter o  $\bar{z}_j$  para a  $j$ -ésima imagem.

Para obter os valores DMOS para todo o banco de dados, é assumido um mapeamento linear entre Z-scores e DMOS fazendo-se  $\text{DMOS}(\bar{z}) = p_1 \bar{z} + p_2$ . Os valores de  $p_1$  e  $p_2$  são obtidos minimizando o erro médio quadrático entre  $\text{DMOS}(\bar{z}_j)$  e  $\text{DMOS}_j$  (21). Os Z-scores podem então ser ou não reescalados para abranger toda a faixa de valores de 1 a 100.

### 1.3 Bancos de Dados

Para avaliar o desempenho das métricas e a contribuição da cor é necessária a utilização de bancos de dados de imagens coloridas com suas respectivas avaliações subjetivas. Esta seção descreve os três bancos de dados de imagens coloridas com avaliação subjetiva usados para o estudo das métricas.

Para avaliar o desempenho de métricas objetivas de avaliação da qualidade, a avaliação produzida pela métrica objetiva deve ser comparada com a avaliação subjetiva da qualidade produzida por observadores humanos. Para este propósito, um banco de dados de imagens com avaliação subjetiva deve ser usado.

O valor médio das notas de qualidade subjetivas para uma dada imagem é geralmente chamado de MOS – *Mean Opinion Score*. Alguns autores processam o valor do MOS transformando-o em DMOS a fim de obter uma avaliação subjetiva mais confiável, conforme descrito na Seção 1.2.

Neste trabalho, três bancos de dados com imagens coloridas distorcidas e de referência bem como suas respectivas avaliações subjetivas são usados: o banco de dados LIVE (17),

composto por 808 imagens, o banco de dados IVC (16), com 120 imagens e o banco de dados TID (18), com 1700 imagens.

O banco de dados LIVE é o mais utilizado na literatura. Assim, foi escolhido como o banco de dados de referência para os experimentos deste trabalho. Os bancos de dados IVC e TID são utilizados de forma a validar os resultados obtidos com o banco de dados LIVE.

Os detalhes da composição e produção das imagens de cada um dos bancos de dados são descritos a seguir.

### 1.3.1 Banco de dados LIVE

Este banco de dados de avaliações subjetivas de imagens foi produzido pelo LIVE (*Laboratory for Image and Video Engineering*) da Universidade do Texas (17).

O banco de dados LIVE conta com um total de 808 imagens, dentre as quais 29 são imagens de referência e 779 são versões degradadas das referências. Ele reflete uma boa diversidade de conteúdo. As imagens incluem rostos, pessoas, animais, cenas naturais, objetos ou nenhum objeto específico de interesse. As imagens de referência são apresentadas na Figura 2. Todas as imagens desse banco de dados foram usadas nos testes aqui apresentados.



Figura 2 – Imagens de referência usadas no banco de dados LIVE.

O método de avaliação subjetiva utilizado na geração do banco de dados é o estímulo simples (*Single Stimulus*) descrito na Seção 1.1.3, onde as imagens de referência são avaliadas na mesma sessão que as imagens distorcidas. A metodologia de um único estímulo foi escolhida no lugar da de duplo estímulo (*Double Stimulus Continuous Quality Scale*) devido ao grande número de imagens a ser avaliado (um total de 779), uma vez que o procedimento de duplo estímulo leva de 3 a 4 vezes mais tempo por imagem que o procedimento de estímulo simples (21). Entretanto, uma vez que as imagens de referência também foram avaliadas, uma nota da diferença de qualidade DMOS foi calculada para todas as imagens distorcidas com notas variando entre 0 e 100. As imagens foram avaliadas por grupos de 20 a 29 observadores a uma distância de 2 a 2,5 vezes a altura da tela.

A produção das imagens degradadas empregou as seguintes distorções: compressão JPEG2000, compressão JPEG, inserção de ruído Gaussiano branco nos componentes RGB, embaçamento Gaussiano nos componentes RGB e erros de bit em fluxos codificados em JPEG2000 quando transmitidos através de um canal de desvanecimento rápido *Rayleigh* (17). As especificações para sua elaboração bem como um exemplo de cada tipo são apresentadas a seguir:

- Compressão JPEG2000: As imagens distorcidas foram produzidas comprimindo as imagens de referência coloridas no formato JPEG2000 com taxas de bits variando entre 0,028 bits por pixel (bpp) e 3,15 bpp. O codificador Kakadu versão 2.2 (22) foi usado para gerar essas imagens. Um exemplo de imagem comprimida com JPEG2000 é visto na Figura 3.



(a) Imagem de referência “dancers”



(b) Imagem comprimida com JPEG2000  
à taxa de 0,25079 bpp

Figura 3 – Exemplo de imagem comprimida com JPEG2000.

- Compressão JPEG: As imagens distorcidas foram geradas comprimindo as imagens coloridas de referência no formato JPEG com taxas de bits variando entre 0,15 bpp e 3,34 bpp. A implementação usada foi a função *imwrite* do Matlab (23). A Figura 4 mostra um exemplo de imagem comprimida com JPEG.



(a) Imagem de referência “church and capitol”



(b) Imagem comprimida com JPEG à taxa de 0,29125 bpp

Figura 4 – Exemplo de imagem comprimida com JPEG.

- Adição de ruído branco: Ruído branco foi adicionado a cada uma das componentes RGB das imagens de referência. O mesmo desvio padrão  $\sigma_N$  para o ruído foi usado nas componentes R, G e B. Os valores usados para o  $\sigma_N$  ficaram entre 0,012 e 2,0. Um exemplo de imagem adicionada de ruído branco Gaussiano está na Figura 5.



(a) Imagem de referência “monarch”



(b) Imagem adicionada de ruído branco com desvio padrão  $\sigma_N = 1$

Figura 5 – Exemplo de imagem adicionada de ruído branco nos componentes RGB.

- **Embaçamento Gaussiano:** As componentes R, G, e B foram filtradas usando um filtro Gaussiano simétrico circular 2D de desvio padrão  $\sigma_B$  entre 0,42 e 15 pixels. A Figura 6 apresenta um exemplo de imagem modificada com embaçamento Gaussiano.



(a) Imagem de referência “parrots”

(b) Imagem com embaçamento Gaussiano de desvio padrão  $\sigma_B = 7,66665$ 

Figura 6 – Exemplo de imagem com embaçamento Gaussiano.

- **Perda de bits devido à transmissão por um canal sem-fio simulado, com desvanecimento rápido *Rayleigh*:** Imagens codificadas em JPEG2000 foram distorcidas por erros de bits durante a transmissão do fluxo JPEG2000 através de um canal sem-fio. A Figura 7 mostra um exemplo de imagem modificada com erro de bit devido à transmissão por um canal sem-fio. O SNR no receptor foi variado



(a) Imagem de referência “plane”

(b) Imagem com erro de bit para potência de ruído igual a 20,3 dB

Figura 7 – Exemplo de imagem com erro de bit devido à transmissão por um canal sem-fio.



para gerar fluxos de bits corrompidos com diferentes proporções de erros de bits. O fluxo fonte JPEG2000 foi gerado usando o codificador já mencionado, mas com o recurso de robustez ao erro ativado com limites de 64 x 64 (21). A taxa da fonte foi fixada em 2,5 bits por pixel para todas as imagens e nenhum algoritmo de supressão de erros foi utilizado. O SNR no receptor usado para variar a potência da distorção variou entre 15,5 e 26,1 dB.

### 1.3.2 Banco de dados IVC

A equipe IVC (*Image Vidéo Communication*) do Instituto de Pesquisa e Comunicações em Cibernética de Nantes IRCCyN (*Institut de Recherche en Communications et en Cybernétique de Nantes*), na França, produziu o banco de dados IVC (16).

As avaliações subjetivas das imagens foram feitas a uma distância de observação de 6 vezes a altura da tela, usando o método DSIS (*Double Stimulus Impairment Scale*) descrito na Seção 1.1.1, com 15 observadores. O MOS foi então calculado como a média das notas dadas para cada imagem.

As intensidades das distorções para cada tipo de imagem foram otimizadas de forma a cobrir uniformemente toda a escala subjetiva na faixa de 1 a 5 (24).

O banco de dados contém 10 imagens originais e 185 versões dessas imagens distorcidas por 4 diferentes tipos de processamento: compressão JPEG, compressão JPEG2000, codificação LAR e aplicação de embaçamento. A Figura 8 mostra as imagens de referência do banco de dados.

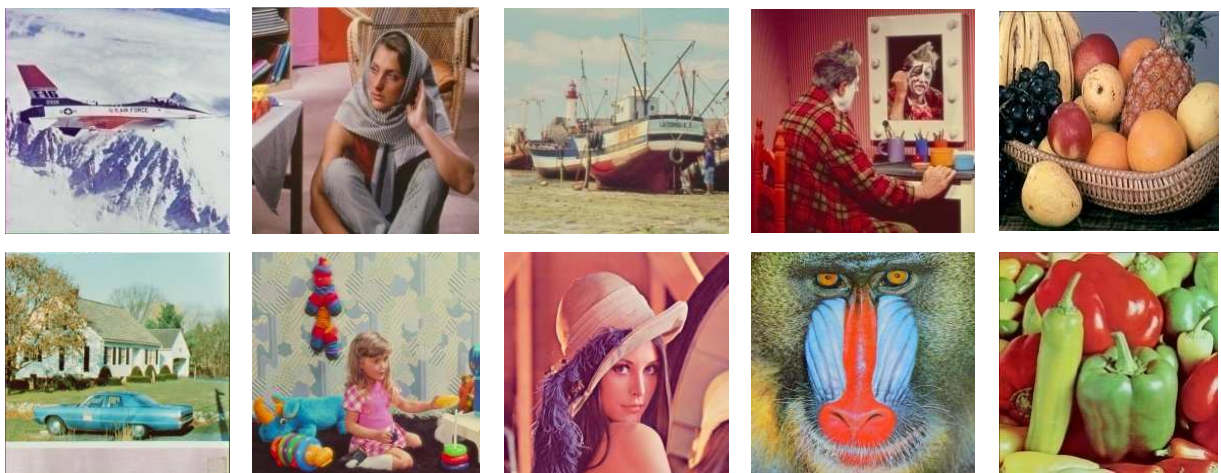


Figura 8 – Imagens de referência usadas no banco de dados IVC.

Para os testes deste trabalho foram selecionadas as imagens degradadas por meio de compressão JPEG, compressão JPEG2000 e embaçamento, contando com um total de 10 imagens de referência e 120 imagens distorcidas. A descrição das distorções pode ser encontrada em (24).

### 1.3.3 Banco de dados TID

O banco de dados TID – *Tampere Image Database 2008* (18), elaborado em laboratórios de universidades de três países – na Universidade Aeroespacial Nacional, na Ucrânia, na Universidade de Tecnologia de Tampere, na Finlândia e na Universidade de Roma, na Itália, é o maior banco de dados de imagens distorcidas (25).

Este banco de dados contém 25 imagens de referência coloridas com diferentes características, com regiões homogêneas, bordas e detalhes e 1700 imagens modificadas por meio de 17 diferentes tipos de distorções. A Figura 9 apresenta as imagens de referência utilizadas nos testes.

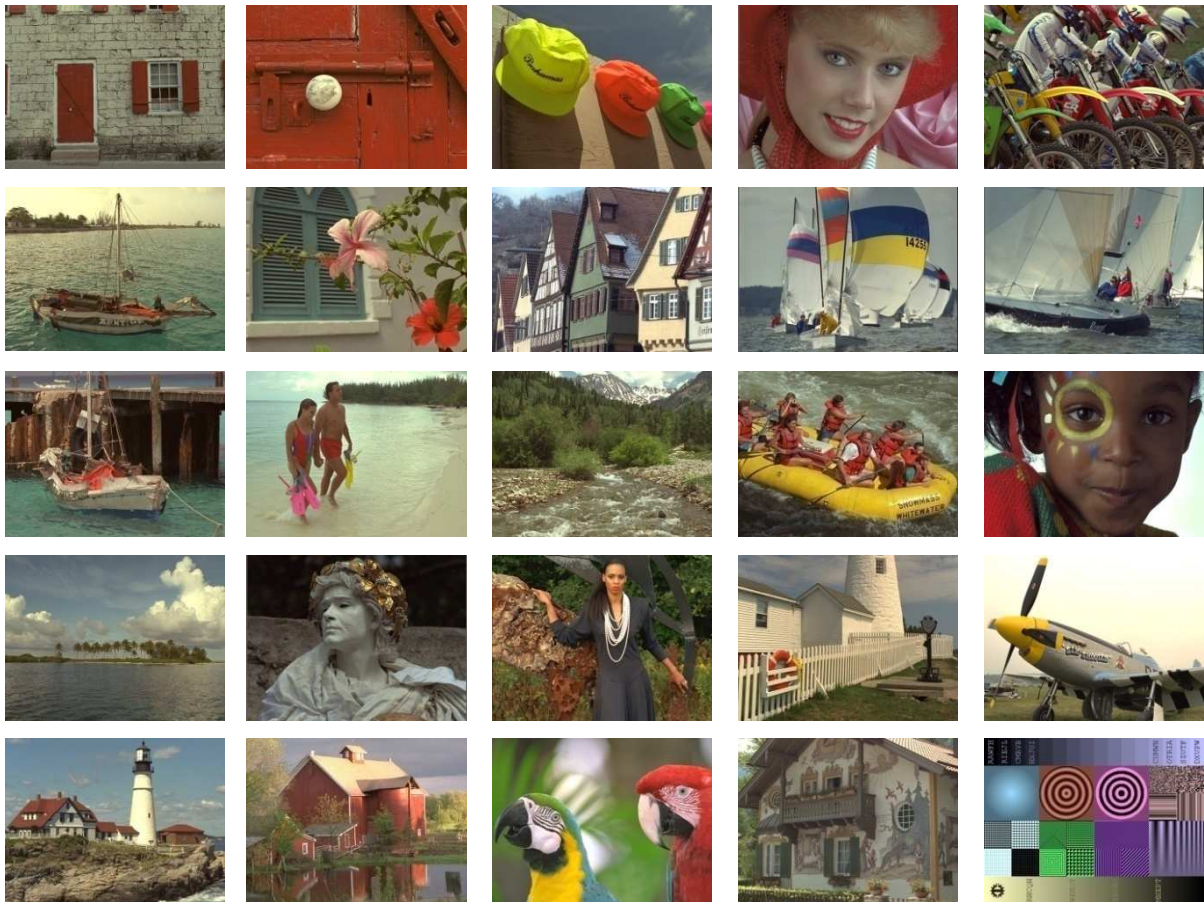


Figura 9 – Imagens de referência usadas no banco de dados TID.

A avaliação subjetiva foi realizada de forma comparativa utilizando o Método de Comparação de Estímulos descrito na Seção 1.1.4. O MOS foi obtido através dos resultados de 830 experimentos realizados por observadores da Finlândia, Itália e Ucrânia. No total, 838 observadores realizaram 256428 comparações de qualidade visual de imagens distorcidas, ou usando outra interpretação, 512856 avaliações de qualidade visual relativa em pares de imagens. O MOS abrange uma faixa de valores de 0 (mínimo) a 9 (máximo), onde um maior MOS corresponde a uma melhor qualidade da imagem.

Para a criação das imagens com defeitos foram aplicados os seguintes tipos de distorções nas imagens de referência (25): ruído Gaussiano aditivo, ruído aditivo nas componentes de cores, ruído espacialmente correlatado, ruído mascarado, ruído de alta frequência, ruído impulsivo, ruído de quantização, embaçamento Gaussiano, redução de ruído da imagem, compressão JPEG, compressão JPEG2000, erros de transmissão em codificação JPEG, erros de transmissão em codificação JPEG2000, ruído de padrão não-excêntrico, distorção por blocos locais de diferentes intensidades, deslocamento da intensidade e mudança de contraste. Cada tipo de distorção foi aplicada com 4 níveis de intensidade às imagens de referência. Detalhes sobre a produção das imagens distorcidas podem ser encontrados em (25).

#### **1.4 Métodos objetivos**

A medida objetiva da qualidade, em oposição à medida subjetiva produzida por observadores humanos, procura determinar a qualidade de imagens de forma algorítmica. Assim, quer-se fornecer uma nota quantitativa computacionalmente que descreva o grau de similaridade/fidelidade ou, de modo oposto, o nível de erro/distorção entre imagens.

Existem três classificações para métricas objetivas de avaliação da qualidade. Elas podem ser métricas de Referência Completa, métricas de Referência Reduzida ou métricas Sem Referência.

Em métricas com referência, assume-se que uma das imagens é a original e possui qualidade perfeita enquanto a outra é distorcida ou contaminada por erros. Desta forma, a avaliação da qualidade usando o método de Referência Completa se torna um problema de avaliação da fidelidade de uma imagem distorcida em relação à imagem original.

Diversas métricas objetivas foram sugeridas na literatura para avaliar a qualidade de imagens digitais. Nesta seção são apresentadas duas métricas que são historicamente as mais usadas e de simples implementação: o MSE e o PSNR.

O MSE e o PSNR são as métricas de avaliação objetiva da qualidade que são utilizadas há mais tempo. O MSE tem sido a métrica quantitativa de desempenho dominante no campo de processamento de sinais, enquanto que para imagens o PSNR, em geral, é mais usado (6).

Imagens digitais são representadas por uma ou mais matrizes de números. Imagens em escala de cinza são compostas por uma única matriz e o valor de cada elemento da matriz está diretamente relacionado à intensidade do tom de cinza. Por exemplo, para codificações usando 8 bits/pixel, 0 representa preto, 255 representa branco, e os demais valores representam tons intermediários de cinza.

Sejam  $\mathbf{x} = \{x_i | i = 1, 2, \dots, N\}$  e  $\mathbf{y} = \{y_i | i = 1, 2, \dots, N\}$  os valores dos pixels da imagem de referência e da imagem modificada que se deseja comparar. Desta forma (6), o MSE é definido via:

$$MSE = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2. \quad (4)$$

A partir do MSE pode-se definir o PSNR (razão sinal-ruído de pico) (6). O PSNR é dado por:

$$PSNR = 10 \log_{10} \frac{L^2}{MSE}, \quad (5)$$

onde  $L$  é a faixa dinâmica da intensidade dos pixels na imagem. Para imagens codificadas em 8 bits/pixel e escala de cinza,  $L = 2^8 - 1 = 255$ . O PSNR é útil quando se deseja comparar imagens com diferentes faixas dinâmicas, ou no caso de imagens coloridas codificadas em espaços de cores cujas faixas dinâmicas dos canais difiram entre si (26), caso contrário não acrescenta novas informações além das já fornecidas pelo MSE. Mas tornou-se mais empregado que o MSE, provavelmente por que ele aumenta com a diminuição da distorção.

Ambas as medidas, MSE e PSNR, são fáceis de calcular. O MSE é geralmente a mais conveniente para fins de otimização de algoritmos uma vez que é diferenciável e, quando combinado com ferramentas da álgebra linear, soluções previamente formuladas podem ser encontradas para problemas reais (6). Além disso, o MSE tem um claro significado físico – a energia do sinal de erro. Tal medida de energia é preservada após transformações lineares ou ortogonais como, por exemplo, a transformada de Fourier. Estas são as grandes razões pelas quais o MSE (e o PSNR) são extensivamente usados na literatura de processamento de sinais, imagens, e comunicações. Não obstante, o MSE vem sendo criticado por sua fraca correlação com a qualidade da imagem percebida (6).

É possível notar que o arranjo de indexação aqui aplicado não leva em conta as informações sobre as posições espaciais dos pixels ou sobre as relações entre eles, mas ao invés disso, os ordena como um vetor unidimensional. Uma vez que o MSE pode ser definido usando essa representação unidimensional, é visível que o MSE não faz uso de nenhuma informação sobre a posição dos valores dos pixels na imagem, o que pode ser de grande valor quando se mede a qualidade de imagens.

No capítulo seguinte é estudado o SSIM, uma métrica de avaliação objetiva da qualidade de imagens inspirada em características do sistema visual humano, como alternativa ao uso das métricas PSNR e MSE.

## 2 O ÍNDICE SSIM

O índice SSIM (*Structural SIMilarity*) é uma métrica objetiva de avaliação da qualidade de imagens em tons de cinza, desenvolvida considerando características da percepção humana em sua elaboração (27).

O SSIM foi introduzido em (27) e é mais formalmente descrito em (9) e (26). O argumento principal sobre o qual se baseia o SSIM é o de que o sistema visual humano é altamente adaptado para extrair informação estrutural de cenas visuais (6).

Neste capítulo o SSIM é explicado, iniciando com a apresentação das principais características da percepção visual consideradas em seu projeto. Na seção seguinte, é feita uma descrição da formulação do SSIM. Por fim, a última seção apresenta um exemplo comparando o desempenho do SSIM e do MSE na avaliação de um conjunto de imagens coloridas.

### 2.1 Percepção no sistema visual humano

Na literatura de neurociência computacional, há muito tempo assume-se a hipótese de que o sistema visual humano é altamente adaptado ao ambiente visual natural (28), não somente em termos da habilidade de adaptação no curto prazo (por exemplo, adaptação às condições luminosas do ambiente), mas também através da evolução e do desenvolvimento neural de longo prazo.

Como resultado, o estudo das características do sistema visual humano (SVH) oferece uma boa base de inspiração para a criação de métricas objetivas de avaliação da qualidade. Nesta seção são descritas algumas das características relacionadas à percepção visual consideradas na elaboração do SSIM, como a sensibilidade ao contraste, a percepção de luz e luminância, a noção de estrutura e a dispersão de ponto.

#### 2.1.1 Sensibilidade ao contraste

A função de sensibilidade ao contraste (CSF) modela a sensibilidade do SVH em função da frequência espacial presente nos estímulos visuais (26). Em geral, a CSF tem uma natureza passa-banda (29). A frequência na qual o pico ocorre varia com o observador e geralmente se encontra entre 3 a 10 ciclos por grau de ângulo visual e decai significativamente com o aumento ou com a diminuição da frequência (29).

Uma medida direta da CSF é possível considerando uma gradação senoidal da variação de contraste (razão entre a intensidade máxima e mínima) e frequência espacial, como mostra a Figura 10. Supondo que a percepção do contraste seja determinada somente pelo contraste, as faixas claras e escuras alternadas devem aparentar a mesma altura por toda direção horizontal ao longo de uma imagem. Entretanto, observa-se que as barras verticais parecem significativamente mais altas no centro da imagem, seguindo a forma da CSF. De fato, a posição do pico muda com a distância de visão, e este efeito observado é uma propriedade do SVH e não da imagem de teste (5).

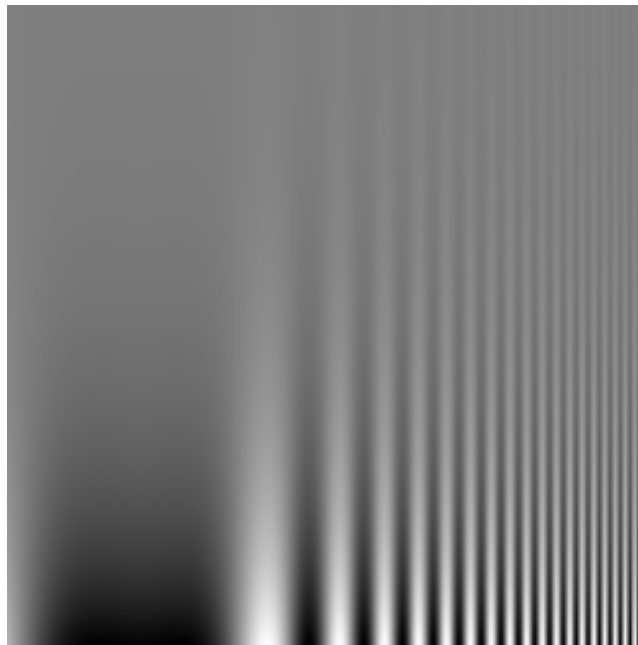


Figura 10 – Percepção da sensibilidade ao contraste de acordo com a frequência espacial.

### 2.1.2 Percepção da luz e luminância

A luminância de um objeto é definida como a quantidade de luz incidente que é refletida pelo mesmo (30). Em uma imagem, a luminância correspondente à superfície de um objeto é o produto da iluminação e das reflexões, mas a estrutura do objeto na cena é independente da iluminação. A sensibilidade à luminância percebida pelo SVH obedece à lei de Weber (9), que estabelece a relação entre estímulo e sensação, permitindo a sua mensuração. Sua definição é dada por:

$$\frac{\Delta I}{I} = K, \quad (6)$$

onde  $I$  é a luminância de fundo,  $\Delta I$  é o mínimo incremento de luminância percebido neste fundo e  $K$  é uma constante chamada de fração de Weber.

Segundo a lei de Weber, por exemplo, a diferença percebida ao se aumentar a intensidade de iluminação de uma lâmpada de 100 para 110 watts será a mesma sentida aumentando-se a intensidade de iluminação de 1000 para 1100 watts. Isto significa que a sensação de luminância está relacionada não com a iluminação em si, mas com a razão entre o incremento de luminância e a luminância de fundo.

Além disso, o sistema visual humano é apto para determinar as cores de um objeto independentemente do iluminante (14). Esta característica é conhecida como *constância da cor* (14). Isto indica que a mudança na iluminação exerce pouco impacto nas estruturas dos objetos na cena visual (26).

### 2.1.3 Estrutura

Distorções não-estruturais são causadas por condições ambientais ou pelas condições instrumentais durante a aquisição e a exibição de imagens (6). Exemplos dessas distorções são mudanças na luminância ou no brilho, mudança de contraste e distorções de Gamut. Essas distorções não alteram a estrutura das imagens (6). Entretanto outras distorções, como por exemplo, adição de ruído, embaçamento e compressão com perdas, distorcem significativamente as estruturas das imagens. Se o sistema visual humano for interpretado como um extrator de informação que busca identificar e reconhecer objetos na cena visual, então ele precisa ser altamente sensível a distorções estruturais e compensar automaticamente distorções não estruturais. Conseqüentemente, uma medida objetiva da qualidade de imagens deve incorporar tal característica.

### 2.1.4 Dispersão de ponto

O sistema ótico visual focaliza o estímulo visual na retina, porém este processo gera um efeito passa-baixa na imagem devido às inerentes imperfeições e limitações da lente do olho, produzindo um efeito visual de embaçamento. Este efeito é tipicamente modelado por um sistema linear invariante no espaço caracterizado por uma função de dispersão de ponto ou *Point Spread Function (PSF)* (2).

A Figura 11 ilustra a função de dispersão de ponto através de um diagrama esquemático. Analisando um ponto qualquer de um objeto, a luz refletida nele chega até o olho. Quando a



imagem é formada na retina, diz-se que a imagem original do objeto foi convoluída com a função de dispersão de ponto, ou PSF.

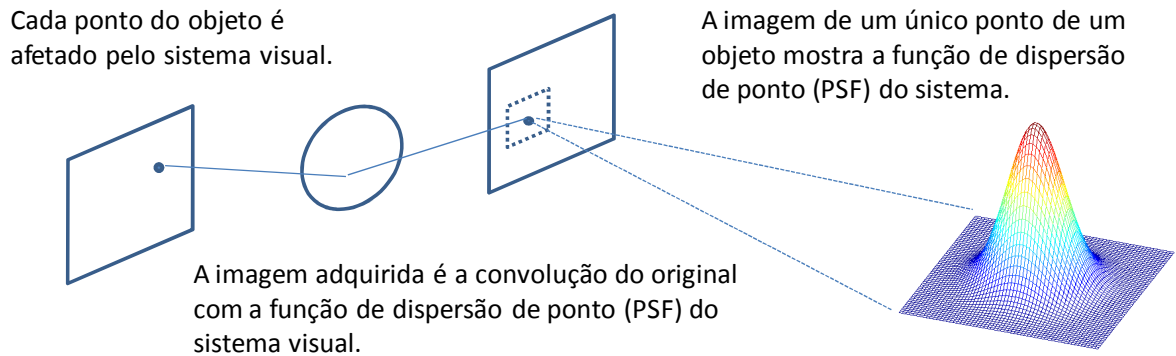


Figura 11 – Diagrama esquemático da função de dispersão de ponto que atua na imagem adquirida pelo sistema visual humano. Inspirado em (31).

## 2.2 O SSIM

O SSIM foi originalmente motivado pela observação de que imagens naturais são altamente estruturadas, o que significa que os pixels de imagens deste tipo tem fortes dependências com os vizinhos e que essas dependências carregam importantes informações sobre as estruturas dos objetos na cena visual (26). A principal assunção em que se baseia o SSIM é que o sistema visual humano é altamente adaptado para extrair informação estrutural das cenas visuais. Desta forma, para a medida da fidelidade de imagens, a retenção da estrutura deve ser um importante ingrediente.

A estrutura da computação do SSIM é mostrada na Figura 12.

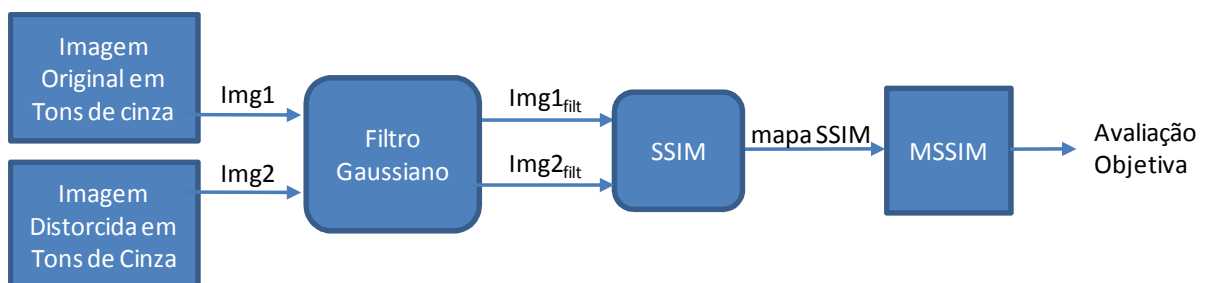


Figura 12 – Estrutura da computação do SSIM.

### 2.2.1 Interpretação geométrica

A métrica SSIM pode ser entendida representando uma imagem como um vetor no qual o valor das coordenadas corresponde à escala de cinza dos pixels da imagem. Qualquer distorção na imagem pode ser interpretada como a adição de um vetor de distorção ao vetor que representa a imagem original.

Considerando uma imagem no  $\mathbb{R}^3$ , isto é, uma imagem composta por três pixels, somente para interpretação, o comprimento de um vetor de distorção não representa uma boa medida da qualidade da imagem, mas as direções desses vetores podem ter importantes significados perceptivos. Imagens com alterações somente na luminância ou no contraste geralmente retêm uma alta qualidade perceptiva. Isto está de acordo com o princípio da similaridade estrutural. A luminância (da superfície de um objeto observado ou cuja imagem é captada) é o produto da iluminação e das reflexões, mas a estrutura do objeto na cena é independente da iluminação. O maior impacto em uma imagem de mudanças na iluminação são variações nos contrastes e luminâncias locais médios. Conseqüentemente, é desejável separar as medidas das distorções causadas pela luminância e contraste das demais distorções estruturais que podem afetar a imagem (26).

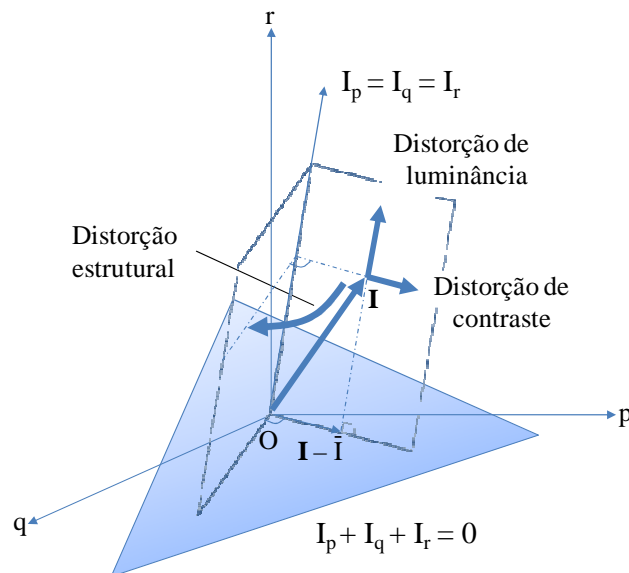


Figura 13 – Separação das distorções de luminância, de contraste e estrutural da imagem de referência  $\mathbf{I}$ . Inspirado em (26).

A Figura 13 mostra como as distorções de luminância e de contraste são separadas da distorção estrutural no espaço da imagem  $\mathbf{I}$ . Sendo  $p$ ,  $q$  e  $r$  três eixos perpendiculares entre si, as distorções de luminância se situam na direção  $I_p = I_q = I_r$  que é perpendicular ao plano

$I_p + I_q + I_r = 0$ . Mudanças no contraste são determinadas pela direção de  $\mathbf{I} - \bar{\mathbf{I}}$ , que é a projeção do vetor imagem  $\mathbf{I}$  no plano  $I_p + I_q + I_r = 0$ . No espaço da imagem, os dois vetores que definem as direções das distorções de luminância e de contraste se estendem em um subespaço determinado pelo vetor da imagem de referência  $\mathbf{I}$ . Assim vemos que a distorção de estrutura está relacionada ao ângulo entre os vetores representativos das imagens retirando as médias. Qualquer outra distorção corresponde à rotação deste plano de certo ângulo pode ser interpretada como uma distorção estrutural, como ilustrado na Figura 13.

O índice SSIM é uma função de duas imagens. Assumindo-se que uma das imagens tem qualidade perfeita (imagem de referência), o SSIM é interpretado como a medida de qualidade da outra imagem. O algoritmo SSIM separa a tarefa da medida da similaridade em três comparações: de luminância, de contraste e de estrutura.

### 2.2.2 Formulação

Supondo que  $\mathbf{x}$  e  $\mathbf{y}$  são janelas tiradas na mesma posição de duas imagens  $\mathbf{X}$  e  $\mathbf{Y}$  que estão sendo comparadas, o SSIM local mede a similaridade de três elementos destas janelas: a similaridade –  $l(\mathbf{x}, \mathbf{y})$  – entre as luminâncias das janelas (valores de brilho), a similaridade –  $c(\mathbf{x}, \mathbf{y})$  – entre os contrastes das janelas, e a similaridade –  $s(\mathbf{x}, \mathbf{y})$  – entre as estruturas das janelas.

Primeiramente, a luminância local de cada janela é estimada pela intensidade média

$$\mu_x = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i. \quad (7)$$

A função de comparação da luminância  $l(\mathbf{x}, \mathbf{y})$  é então expressa em função de  $\mu_x$  e  $\mu_y$ :

$$l(\mathbf{x}, \mathbf{y}) = \text{luma}(\mu_x, \mu_y), \quad (8)$$

onde  $\mu_x$  e  $\mu_y$  são as médias das amostras de  $\mathbf{x}$  e  $\mathbf{y}$  respectivamente.

A seguir, é removida a intensidade média do sinal. O sinal resultante  $\mathbf{x} - \mu_x$  é a projeção do vetor  $\mathbf{x}$  no plano  $\sum_{i=1}^N x_i = 0$ , como representado na Figura 13. O desvio padrão é, então, usado como uma estimativa aproximada do contraste da imagem. O desvio padrão é dado por:

$$\sigma_x = \left( \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2 \right)^{1/2}. \quad (9)$$

A comparação de contraste  $c(\mathbf{x}, \mathbf{y})$  é uma função dos desvios padrão  $\sigma_x$  e  $\sigma_y$ :

$$c(\mathbf{x}, \mathbf{y}) = cont(\sigma_x, \sigma_y). \quad (10)$$

Faz-se então a transformação das imagens usando suas médias e desvios padrões. A comparação de estrutura  $s(\mathbf{x}, \mathbf{y})$  é conduzida nas imagens normalizadas resultantes:

$$s(\mathbf{x}, \mathbf{y}) = estrut \left( \frac{\mathbf{x} - \mu_x}{\sigma_x}, \frac{\mathbf{y} - \mu_y}{\sigma_y} \right). \quad (11)$$

Finalmente, os três componentes são combinados de forma a produzir um SSIM local:

$$SSIM(\mathbf{x}, \mathbf{y}) = f[l(\mathbf{x}, \mathbf{y}) \cdot c(\mathbf{x}, \mathbf{y}) \cdot s(\mathbf{x}, \mathbf{y})]. \quad (12)$$

É válido notar que os três componentes são relativamente independentes, o que é fisicamente sensato uma vez que a mudança na luminância e/ou no contraste exerce pouco impacto nas estruturas dos objetos na cena visual.

Para completar a definição, é necessário definir as funções  $l(\mathbf{x}, \mathbf{y})$ ,  $c(\mathbf{x}, \mathbf{y})$  e  $s(\mathbf{x}, \mathbf{y})$  e como combiná-las. Além disso, deseja-se que a medida de similaridade satisfaça as seguintes condições (26):

1. Simetria:  $SSIM(\mathbf{x}, \mathbf{y}) = SSIM(\mathbf{y}, \mathbf{x})$ . Quando quantificando a similaridade entre duas imagens, a alteração da ordem das imagens não deve afetar a medida resultante.
2. Delimitação:  $SSIM(\mathbf{x}, \mathbf{y}) \leq 1$ . Um limite superior pode servir como um indicador do quão próximas duas imagens estão de serem perfeitamente idênticas.
3. Máximo único:  $SSIM(\mathbf{x}, \mathbf{y}) = 1$  se e somente se  $\mathbf{x} = \mathbf{y}$ . A avaliação perfeita é alcançada somente quando as imagens sendo comparadas são idênticas. Em outras palavras, a medida de similaridade deve quantificar quaisquer variações que possam existir entre as imagens.

Para a comparação de luminância define-se:

$$l(\mathbf{x}, \mathbf{y}) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}, \quad (13)$$

onde a constante  $C_1$  é incluída para evitar instabilidade quando  $\mu_x^2 + \mu_y^2$  é muito próximo de zero. Especificamente em (26), foram escolhidos  $C_1 = (K_1L)^2$  onde  $L$  é a faixa dinâmica de valores do pixel (255 para imagens de 8 bits), e  $K_1 \ll 1$ . Considerações semelhantes também são aplicadas à comparação de contraste e à comparação de estrutura. Vê-se que a equação (13) obedece às três propriedades listadas anteriormente. Essa equação é também consistente com o princípio de uma sensibilidade à mudança relativa da luminância, e não à mudança absoluta da luminância.

A função de comparação de contraste assume uma forma semelhante:

$$c(\mathbf{x}, \mathbf{y}) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}, \quad (14)$$

onde  $C_2 = (K_2L)^2$  e  $K_2 \ll 1$ . Esta definição também satisfaz às três propriedades listadas acima. Uma importante e desejável característica desta função é que dada uma mudança específica no contraste  $\Delta\sigma = \sigma_y - \sigma_x$ , a função é menos sensível para um alto contraste  $\sigma_x$  do que para baixos contrastes.

A comparação de estrutura é conduzida após a subtração da luminância e normalização do contraste. Especificamente os dois vetores unitários  $(\mathbf{x} - \mu_x)/\sigma_x$  e  $(\mathbf{y} - \mu_y)/\sigma_y$ , ambos sobre o plano  $\sum_{i=1}^N x_i = 0$ , são associados com a estrutura das duas imagens. A correlação (produto interno) entre esses vetores é uma simples e eficaz medida para quantificar a similaridade estrutural. Notando que a correlação entre  $(\mathbf{x} - \mu_x)/\sigma_x$  e  $(\mathbf{y} - \mu_y)/\sigma_y$  é equivalente ao coeficiente de correlação entre  $\mathbf{x}$  e  $\mathbf{y}$ , a função de comparação da estrutura é definida por:

$$s(\mathbf{x}, \mathbf{y}) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}. \quad (15)$$

Em uma implementação discreta o coeficiente de correlação  $\sigma_{xy}$  pode ser estimado como

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y). \quad (16)$$

Assim, geometricamente, o coeficiente de correlação é o cosseno do ângulo entre os vetores  $(\mathbf{x} - \mu_x)$  e  $(\mathbf{y} - \mu_y)$ . Como nas medidas da luminância e do contraste, é introduzida uma pequena constante no numerador e no denominador.

Os elementos luminância, contraste e estrutura das equações (13), (14) e (15) são então combinados em um índice SSIM entre  $\mathbf{x}$  e  $\mathbf{y}$ , através de:

$$\begin{aligned} SSIM(\mathbf{x}, \mathbf{y}) &= [l(\mathbf{x}, \mathbf{y})]^\alpha \cdot [c(\mathbf{x}, \mathbf{y})]^\beta \cdot [s(\mathbf{x}, \mathbf{y})]^\gamma \\ &= \left( \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1} \right)^\alpha \cdot \left( \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2} \right)^\beta \cdot \left( \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3} \right)^\gamma, \end{aligned} \quad (17)$$

onde  $\alpha > 0$ ,  $\beta > 0$  e  $\gamma > 0$  são parâmetros que ajustam a importância relativa dos três componentes. É fácil verificar que esta definição satisfaz aos três princípios anteriormente descritos.

Para simplificar a expressão e reduzir o número de parâmetros usou-se  $\alpha = \beta = \gamma = 1$  e  $C_3 = C_2/2$ , conforme a implementação original desenvolvida em (32), o que resulta em uma forma específica do SSIM (9):

$$SSIM(\mathbf{x}, \mathbf{y}) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \quad (18)$$

### 2.2.3 Média SSIM

O índice SSIM é, em geral, computado localmente dentro de uma janela deslizante que se move pixel a pixel por toda a imagem, resultando em um mapa SSIM (26). Porém, na prática, requer-se um único valor que meça a qualidade da imagem inteira. Para esta finalidade a média dos índices do mapa SSIM (MSSIM) é usada para avaliar a qualidade da imagem completa. O MSSIM é dado por:

$$\text{MSSIM}(\mathbf{X}, \mathbf{Y}) = \frac{1}{M} \sum_{j=1}^M \text{SSIM}(\mathbf{x}_j, \mathbf{y}_j), \quad (19)$$

onde  $\mathbf{X}$  e  $\mathbf{Y}$  são respectivamente as imagens de referência e modificada;  $\mathbf{x}_j$  e  $\mathbf{y}_j$  são os conteúdos das imagens na  $j$ -ésima janela local; e  $M$  é o número de janelas locais na imagem. Para uma terminologia mais simples, ao longo deste trabalho o MSSIM será referido simplesmente como SSIM e em alguns casos específicos quando SSIM se refere ao mapa SSIM faz-se a identificação.

#### 2.2.4 Filtro suavizante

Conforme visto no início do capítulo, o olho humano tem melhor percepção para baixas frequências, devido à função de dispersão de ponto da ótica do olho. De forma a considerar esse efeito, um filtro passa-baixa suavizante é aplicado à imagem na janela local antes do cálculo do SSIM, de forma a simular a função de dispersão de ponto do olho humano.

Um filtro Gaussiano com desvio padrão = 1,5 e tamanho de janela = 11 é utilizado (32). O filtro e o módulo de sua resposta em frequência são mostrados na Figura 14.

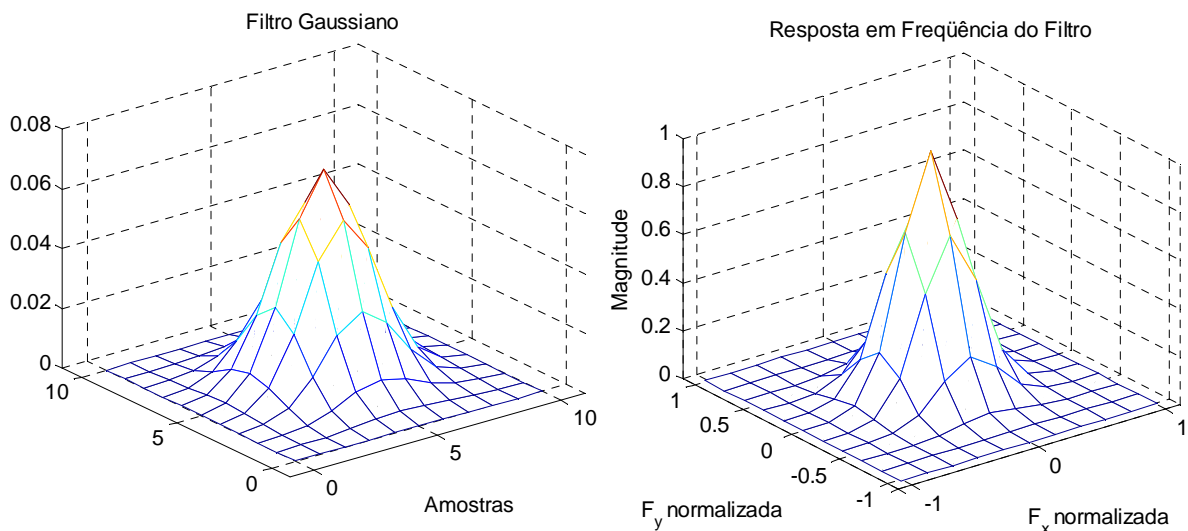


Figura 14 – Filtro Gaussiano com desvio padrão = 1,5 e janela = 11, e sua resposta em frequência.

Neste trabalho realizaram-se experimentos variando a aplicação do filtro suavizante e a não aplicação deste. Os valores das constantes  $K_1 = 0,01$  e  $K_2 = 0,03$  foram usados, e  $L = 1$ ,

pois as imagens de entrada foram convertidas para a faixa de valores entre 0 e 1. Esses experimentos são apresentados nos Capítulos 5 e 7.

### 2.3 MSE versus SSIM

Conforme apresentado na Seção 1.4, o MSE não é uma boa métrica para a avaliação da qualidade de imagens. Para ilustrar o comportamento do MSE para a avaliação da qualidade de imagens e compará-lo ao SSIM, ambas as métricas são aplicadas a imagens distorcidas.

A Figura 15 apresenta a imagem original Lena em (a) modificada de três formas perceptivelmente diferentes: em (b) com aumento do contraste, em (c) com a adição de ruído Gaussiano nas componentes RGB e em (d) com a adição de embaçamento Gaussiano.

É possível notar que a imagem com ruído (c) apresenta uma qualidade visual pior do que a imagem com contraste (b). A imagem (b), inclusive, parece ter uma qualidade visual melhor que a imagem original (a). Por sua vez, a imagem com embaçamento (d) possui uma qualidade muito inferior às imagens (b) e (c). Entretanto, para as três imagens, o MSE calculado no canal L da imagem convertida para CIELAB (a ser discutido na Seção 3.6) foi o mesmo. Por outro lado, o resultado do SSIM calculado para as mesmas imagens se assemelha muito mais à percepção humana do que aqueles produzidos pelo MSE.

A partir deste exemplo é possível perceber que o MSE não representa uma boa métrica para avaliação da qualidade de imagens. Esta métrica não apresenta uma boa correlação com a avaliação subjetiva, uma vez que imagens avaliadas pela percepção humana de maneiras muito distintas são avaliadas pela métrica com o mesmo grau.

No capítulo seguinte são estudados alguns espaços de cores e características relativas à percepção da cor, a fim de fundamentar o estudo sobre a aplicação do SSIM para avaliação de imagens coloridas.





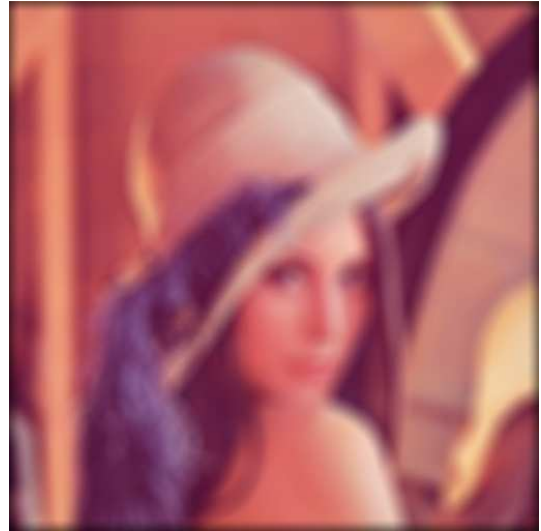
(a) Lena original. MSE = 0 e SSIM = 1.



(b) Lena modificada com aumento do contraste. MSE = 71 e SSIM = 0.8631.



(c) Lena modificada com ruído Gaussiano nos canais RGB. MSE = 71 e SSIM = 0.2254.



(d) Lena modificada com embaçamento. MSE = 71 e SSIM = 0.1693.

Figura 15 – Comparação entre SSIM e MSE para imagens degradadas. Inspirado em (26).

### 3 ESPAÇOS DE CORES

Imagens coloridas para serem representadas necessitam que seja definido um espaço de cores. Este capítulo descreve a forma como as imagens são representadas digitalmente e apresenta alguns espaços de cores estudados para aplicação na avaliação de imagens coloridas utilizando o SSIM.

Para melhor compreender os fundamentos das imagens digitais e dos espaços de cores é necessário conhecer algumas características do olho humano e da percepção da cor que motivaram a forma como as imagens digitais são representadas e como os espaços de cores foram projetados.

Assim, as duas primeiras seções deste capítulo fazem uma introdução ao assunto, explicando sobre a estrutura do olho humano, suas células fotorreceptoras e sobre como se dá a percepção da cor pelo sistema visual humano, e como imagens digitais coloridas são representadas. As quatro seções seguintes apresentam os principais aspectos dos quatro espaços de cores – RGB, YCbCr,  $L\alpha\beta$  e CIELAB, utilizados para os experimentos desta dissertação. Na última seção é feita uma descrição da medição da distância entre cores, calculada utilizando o espaço CIELAB, e é explicado como ela pode ser aplicada na avaliação de imagens coloridas.

#### 3.1 Percepção da cor

O processamento da informação pelo sistema visual humano começa pelo olho. A luz atravessa a lente e atinge a retina, que está localizada no fundo do globo ocular. A retina consiste basicamente em um conjunto de receptores que medem a quantidade de luz incidente. Nela existem em torno de 126 milhões de fotorreceptores, onde aproximadamente 120 milhões são bastonetes e os 6 milhões restantes são cones (14).

Os bastonetes, aproximadamente 25 vezes mais sensíveis ao estímulo luminoso que os cones (33), são usados quando há pouca luz (visão *escotópica*), enquanto os cones são usados para visão colorida com muita iluminação (visão *fotópica*).

Existem três tipos de cones, cada um contendo um fotopigmento distinto. Suas sensibilidades de pico se encontram ou na região azul do espectro visível (S-cones em comprimento de onda curto), ou na verde (M-cones em comprimento de onda médio) ou na vermelha (L-cones em comprimento de onda longo) (34).

A maioria dos cones está localizada dentro de uma pequena área no centro da retina chamada fóvea, enquanto os bastonetes estão distribuídos por toda a retina, exceto na fóvea. A Figura 16 mostra a distribuição dos fotorreceptores na retina em relação à fóvea, seu ponto central.

Os bastonetes são dominantes fora da fóvea, o que explica porque é mais fácil ver objetos muito brilhantes e difusos (como por exemplo estrelas) quando estão no campo periférico da visão do que quando se olha diretamente para eles (5). O ponto cego indicado na Figura 16 é o ponto da retina de onde sai o nervo óptico. Esta região é completamente desprovida de fotorreceptores. O tamanho e o espaçamento dos fotorreceptores determinam a resolução espacial do sistema visual humano.

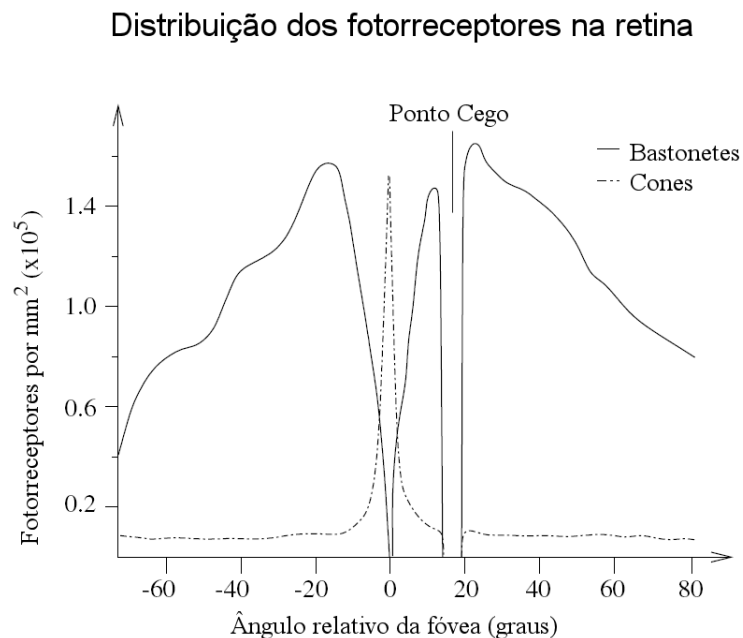


Figura 16 – Distribuição dos fotorreceptores na retina em relação à fóvea. Fonte: (35).

A característica de maior quantidade de fotorreceptores relacionados à luminância e menor quantidade de fotorreceptores relacionados às cores é um importante fundamento considerado no projeto de espaços de cores como YCbCr,  $L\alpha\beta$  e CIELAB.

O principal escopo da visão em cores é auxiliar na identificação de objetos e materiais com base no espectro refletido pelas superfícies. A informação cromática é crucial para diversas tarefas, como por exemplo detecção de contorno, segregação de texturas, agrupamento, busca visual, etc. (36)

A luz absorvida pelos três tipos de cones é quase imediatamente reorganizada em canais de cores opostas (37). Hering (38) foi o primeiro a relatar que alguns pares de tonalidades

podem coexistir em uma única sensação de cor (por exemplo, um amarelo avermelhado é percebido como laranja) enquanto que outros não podem (não se percebe nunca um verde avermelhado, por exemplo) (39). Isto o levou à conclusão de que as sensações de vermelho e verde como também azul e amarelo são codificadas em diferentes caminhos visuais, o que é comumente conhecido como a teoria das cores opostas (37). Os principais componentes do espaço de cores opostas são o preto-branco, vermelho-verde e azul-amarelo.

Experimentos fisiológicos revelaram a existência dos sinais opostos nos caminhos visuais (40). Eles mostraram que, dependendo do tipo de cone, certos comportamentos excitantes ou inibitórios ocorrem muito mais frequentemente que outros: neurônios excitados por L-cones ‘vermelhos’ são geralmente inibidos por M-cones ‘verdes’ e neurônios excitados por S-cones ‘azuis’ são geralmente inibidos por uma combinação de L- e M-cones (40). Desta forma, os campos receptivos destes neurônios sugerem uma conexão entre sinais neurais e cores opostas percebidas.

### **3.2 Imagens digitais e espaços de cores**

Imagens digitais em tons de cinza são representadas por uma única matriz enquanto imagens coloridas são compostas por três ou quatro matrizes (11), dependendo do espaço de cores utilizado para sua representação. Cada uma das matrizes está associada a um canal do espaço de cores utilizado. Por exemplo, no espaço RGB a primeira matriz indica a intensidade do vermelho, a segunda indica a do verde e a terceira a do azul. A cor de um pixel é dada pela combinação das cores de todas as matrizes nos elementos correspondentes ao pixel.

As três matrizes do espaço RGB foram inspiradas no modelo tricromático da visão humana criado por Young-Helmholtz (41), onde cada componente R, G e B representaria células fotorreceptoras do olho humano (células cone) sensíveis em distintas faixas de frequência correspondentes às cores vermelho, verde e azul. O canal R está relacionado aos cones L e à cor vermelha, o canal G representa os cones M e a cor verde e o canal B está para os cones S e para a cor azul. A Figura 17 ilustra as curvas normalizadas das sensibilidades dos fotorreceptores em função do comprimento de onda correspondente a cada uma das cores RGB, captada por cada tipo de cone quando existe alto nível de iluminação, e à luminância, captada pelos bastonetes quando há baixo nível de iluminação e os cones não estão sendo usados.

Existem diversas métricas de avaliação da qualidade que se concentram na avaliação de imagens em escala de cinza, o que é bastante coerente com o que é compreendido sobre o sistema visual humano hoje em dia. Entretanto, há efeitos relevantes da cor que influenciam na qualidade? Este é um dos aspectos investigados neste trabalho.

A princípio, métricas de fidelidade de imagens em tons de cinza poderiam ser aplicadas a imagens coloridas usando-se a simples abordagem de aplicá-las a cada um dos canais de cores individualmente e, então, ponderar/combinar os resultados dos diversos canais em um resultado único. As métricas também poderiam ser aplicadas em somente um dos canais, nos casos em que um dos canais porte mais informações da imagem em detrimento dos demais, como ocorre nos canais de luminância, presente em alguns espaços de cores. Entretanto, em todo caso, a escolha de um espaço de cores adequado é crucial para que sejam produzidos resultados bem correlacionados com a percepção humana.

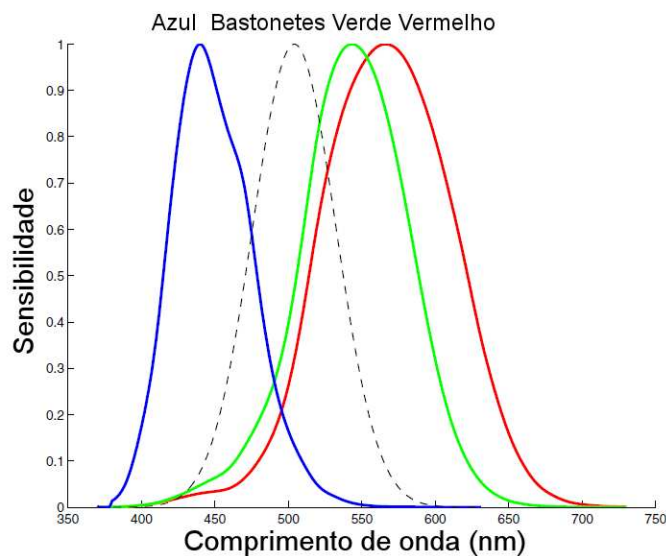


Figura 17 – Curvas normalizadas das sensibilidades dos fotorreceptores. Fonte: (35).

Existe um grande número de diferentes espaços de cores. A maioria dos espaços de cores corresponde a transformações lineares de um mesmo sistema de coordenadas. Cada espaço de cores foi criado por uma variedade de razões. Em (15) pode ser encontrada uma descrição dos espaços de cores CIE e suas principais aplicações.

As fontes espectrais primárias CIE não levam à totalidade das cores reproduzíveis, de fato, não foi encontrado nenhum conjunto de três primárias que possam reproduzir todas as cores. Este motivo levou ao desenvolvimento do sistema CIE XYZ, onde Y representa a luminância da cor, com fontes primárias hipotéticas, de forma que todos os valores triestímulos espectrais

sejam positivos (29). O espaço de cores  $xyY$  é uma especificação do CIE baseada no espaço XYZ que proporciona valores de cromaticidade normalizados. O Y maiúsculo representa a luminância e é o mesmo do espaço de cores XYZ.

O diagrama de cromaticidade é uma forma gráfica de representar um espaço de cores no plano de cromaticidade (29). Os espaços de cores podem ser interpretados como sólidos, uma vez que possuem geralmente três dimensões, e então o diagrama de cromaticidade é a projeção do sólido tridimensional de cores em um plano. A Figura 18 mostra o diagrama de cromaticidade  $xy$  para o sistema espectral primário CIE XYZ (15). Na figura, os pontos na curva de contorno do diagrama se referem ao valor do comprimento de onda correspondente à cor representada no ponto (42). Os eixos  $x$  e  $y$  representam as componentes de cromaticidade da cor normalizadas. Todo o espaço de cores pode ser representado pelas coordenadas  $(x, y, Y)$ , onde  $Y = \text{constante}$  é um plano de cromaticidade.

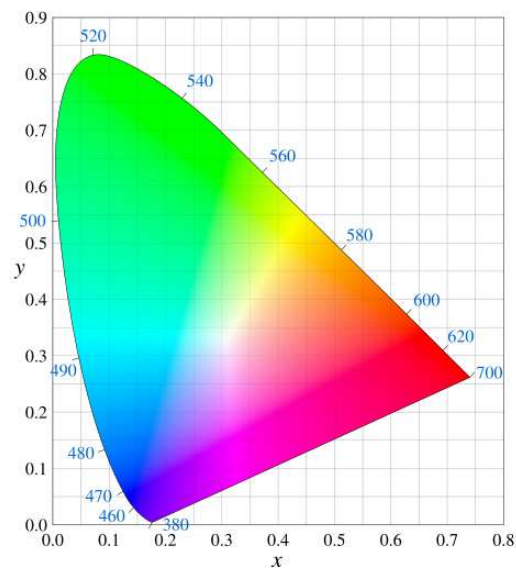


Figura 18 – Diagrama de cromaticidade  $xy$  do espaço CIE XYZ. Fonte: (42).

Dados os conceitos gerais das imagens digitais, as seções seguintes discorrem sobre quatro espaços de cores, RGB, YCbCr,  $L\alpha\beta$  e CIELAB, apresentando suas formulações, seus usos e as principais características que motivam sua utilização para avaliação da qualidade de imagens coloridas.

### 3.3 Espaço RGB

O espaço RGB é o espaço de cores mais comum. O processo de aquisição de imagens normalmente utiliza este espaço de cores (11). Também é comum a sua utilização em displays coloridos como no tubo de raios catódicos (CRT), display de cristal líquido (LCD), display de plasma ou LED (43). Nessas aplicações as componentes R, G e B podem ser representadas com 8 ou 16 bits. No espaço RGB, cada cor é definida pela sua posição em um cubo de lado unitário onde os eixos representam gradações de vermelho, verde e azul.

O espaço RGB é dependente do dispositivo no qual está sendo usado. Isto quer dizer que a exibição de uma imagem RGB depende de parâmetros do dispositivo, como por exemplo o fator de correção gamma. O padrão sRGB foi criado como um espaço de cores de referência independente de dispositivos. É baseado em um espaço de cores RGB calorimétrico calibrado e considera o valor da correção gamma de 2.2 (14). A Figura 19 mostra o cubo sRGB.

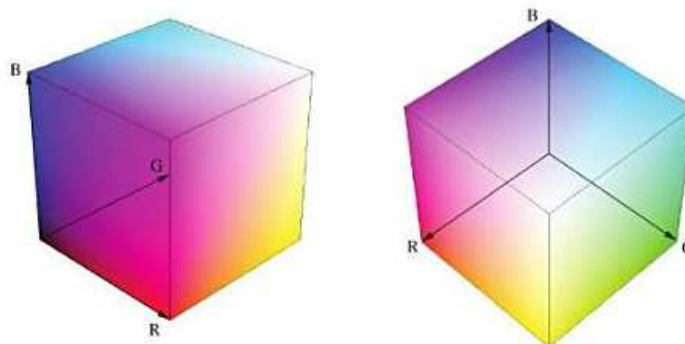


Figura 19 – Espaço de cores sRGB. Fonte: (14).

A avaliação da qualidade de imagens utilizando o espaço de cores RGB pode ser vantajosa em aplicações de captura e exibição, uma vez que não é necessário converter o sinal imagem para outro espaço de cores, reduzindo assim o custo computacional que se teria realizando a conversão.

### 3.4 Espaço YCbCr

O YCbCr não é um espaço de cores absoluto e sim uma forma de codificação das informações digitais RGB. Esta codificação está definida na recomendação ITU-T 601-4 (44). Tal qual o RGB, o YCbCr é dependente do dispositivo. A compressão MPEG usada em DVDs, blu-rays, projetores LCD, televisores digitais de alta definição e câmeras digitais

(MiniDV, DV, Digital Betacam, etc.) produzem vídeo codificado usando YCbCr (10). Um espaço equivalente ao YCbCr utilizado em aplicações analógicas é o YPbPr (10).

A forma YCbCr definida para a televisão padrão para componentes de vídeo digital (45) é derivada do espaço RGB por meio de:

$$\begin{bmatrix} Y \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 0 \\ 128 \\ 128 \end{bmatrix} + \begin{bmatrix} 65,481 & 128,553 & 24,966 \\ -37,797 & -74,203 & 112,00 \\ 112,00 & -93,786 & -18,214 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}, \quad (20)$$

onde Y é o componente de luma (luminância) e Cb e Cr são os componentes de croma (cor) da diferença de azul e da diferença de vermelho, respectivamente.

A Figura 20 exhibe o cubo sRGB representado nos eixos YCbCr.

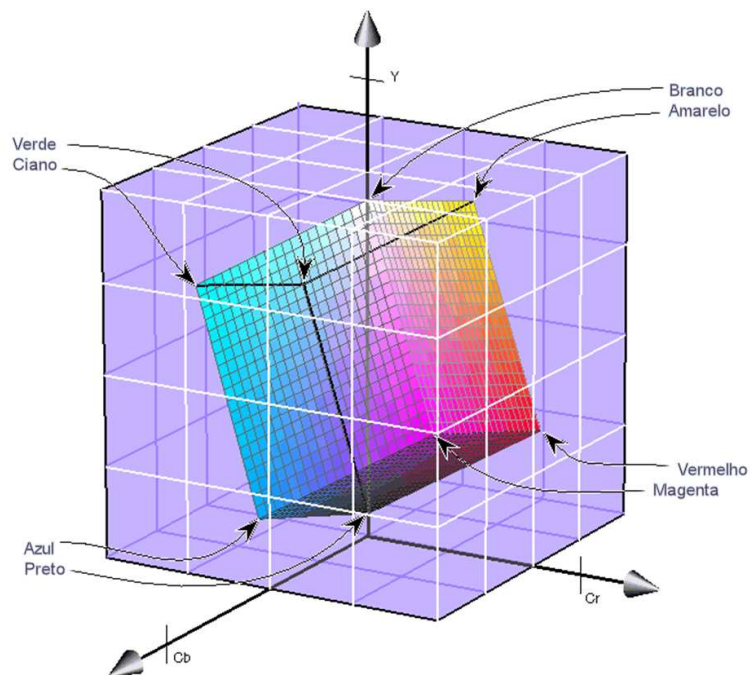


Figura 20 – Cubo sRGB representado nos eixos YCbCr.

A recomendação ITU-R BT.709 define uma forma para o YCbCr ligeiramente diferente, primeiramente para o uso em HDTV. A mais nova forma é também usada em algumas aplicações orientadas para a exibição em computadores (10).

A avaliação da qualidade de imagens utilizando este espaço de cores pode ser vantajosa quando aplicada diretamente nos dispositivos que utilizam esta codificação, reduzindo o custo computacional necessário para converter o sinal para um espaço de cores diferente.



### 3.5 Espaço $L\alpha\beta$

Em (12) é apresentado um estudo utilizando imagens de folhagens naturais para caracterizar a redundância apresentada pelos fotorreceptores atuando na captura de imagens naturais. Este estudo concluiu que existem três direções específicas na resposta logarítmica dos cones ao longo das quais a resposta às imagens naturais são robustamente descorrelacionadas. Das conclusões desse estudo foi proposto um espaço de cores que explora a descorrelação entre a luminância e as cores opostas azul-amarelo e vermelho-verde. O espaço de cores foi denominado  $L\alpha\beta$ .

O espaço de cores  $L\alpha\beta$  é calculado a partir do espaço RGB através das equações (21), (22) e (23):

$$\begin{bmatrix} l' \\ m' \\ s' \end{bmatrix} = \begin{bmatrix} 0,3811 & 0,5783 & 0,0402 \\ 0,1967 & 0,7244 & 0,0782 \\ 0,0241 & 0,1288 & 0,8444 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (21)$$

$$l = \log(l') \quad m = \log(m') \quad s = \log(s') \quad (22)$$

$$\begin{bmatrix} L \\ \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & \frac{1}{\sqrt{6}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & -2 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} l \\ m \\ s \end{bmatrix}. \quad (23)$$

Nas equações,  $l'm's'$  é um espaço de cores intermediário usado para a conversão. O canal L é acromático. Esse é o tipo de informação que domina em imagens naturais e que tem a maior variância (12). Nos canais cromáticos,  $\alpha$  é a direção que tem a maior variação e representa o mecanismo de cores opostas azul-amarelo enquanto a direção  $\beta$  atua como um canal vermelho-verde (12).

Uma métrica de avaliação da qualidade de imagem derivada a partir da aplicação do SSIM aos canais individuais unindo os resultados em uma média vetorial ponderada baseada neste espaço de cores é apresentada em (1).

### 3.6 Espaço CIELAB

O espaço de cores CIELAB (13), também conhecido como  $L^*a^*b^*$  ou simplesmente Lab, foi definido objetivando uma medida da diferença entre as cores que fosse perceptivamente uniforme (13).

Em 1976 a CIE (Comissão Internacional de Iluminação) recomendou internacionalmente dois sistemas de especificação de cores, o CIELAB e o CIELUV, com suas correspondentes fórmulas de diferença de cor associadas (46). Não foi possível para o CIE recomendar naquele momento um único sistema e uma única fórmula de diferença de cores com base em critérios de mérito científico. Adotou-se, então, esta dupla recomendação pensando que assim se estabeleceria certa uniformidade de uso entre a indústria, como de fato ocorreu. Em uma pesquisa realizada em 1990, 92% das indústrias americanas utilizavam o sistema CIELAB (47). O sistema CIELUV, ainda que recomendado pelo CIE ao mesmo nível que o CIELAB, e ainda o sendo, é muito menos empregado que o CIELAB pelas indústrias que necessitam processar alguma informação de cor. Em todo caso, o CIELUV é usado em determinados setores ou aplicações industriais como as referentes a lâmpadas, monitores CRT e outros tipos de telas de visualização. Neste trabalho o espaço CIELAB foi escolhido, pois é o mais utilizado em numerosos estudos científicos (47) e (35).

Ambos os espaços CIELAB e CIELUV foram desenvolvidos baseados nas elipses de MacAdam (15) ilustradas na Figura 21. A área dentro de uma elipse de MacAdam definida no diagrama de cromaticidade XYZ inclui cores visualmente idênticas àquela presente no centro da elipse (48).

O limiar das elipses de MacAdam é conhecido como *Mínima Diferença Perceptível* ou JND (*Just Noticeable Difference*) (15). O conceito de JND foi traduzido para o espaço de cores CIELAB, de forma tal que a distância euclidiana entre as coordenadas neste espaço forneça uma aproximação da diferença percebida pela visão humana entre duas cores distintas. Essa distância é também conhecida como  $\Delta E$  (49).

Uma análise matemática das elipses revelou que não existe nenhuma transformação geométrica que modifique ou transforme simultaneamente todas as elipses em círculos, o que é pré-requisito para um espaço de cores uniforme. Entretanto, muitas tentativas foram feitas a fim de aproximar as elipses por círculos e desenvolver um espaço de cores aproximadamente uniforme. As transformações mais proeminentes e amplamente usadas foram definidas em 1978 como espaços de cores CIELAB e CIELUV (2).

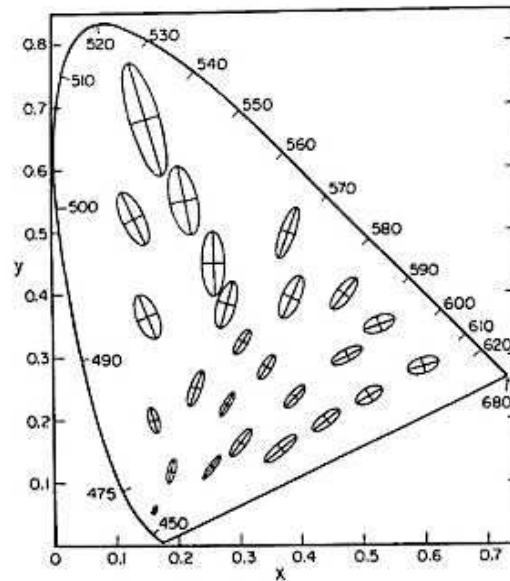


Figura 21 – Elipses de MacAdam representadas sobre o diagrama de cromaticidade xy do espaço de cores CIE XYZ. Fonte: (48).

As três coordenadas CIELAB ( $L, a, b$ ) representam a luminância da cor (onde  $L = 0$  representa o preto e  $L = 100$  indica um branco difuso), a posição entre um vermelho/magenta e verde ( $a$ , onde  $a < 0$  indica verde e  $a > 0$  indica magenta) e a posição entre amarelo e azul ( $b$ , onde  $b < 0$  representa azul e  $b > 0$  indica amarelo). Esta configuração está de acordo com a característica de percepção das cores opostas do SVH descrita na Seção 3.1.

O espaço CIELAB é independente do dispositivo. Para calcular as coordenadas Lab a partir do espaço de cores sRGB faz-se (14):

$$\begin{aligned} L^* &= 116f(Y/Y_n) - 16 \\ a^* &= 500[f(X/X_n) - f(Y/Y_n)] \\ b^* &= 200[f(Y/Y_n) - f(Z/Z_n)] \end{aligned} \quad (24)$$

onde  $X_n, Y_n$  e  $Z_n$  descrevem um estímulo de cor branco de referência D50. A função  $f(\cdot)$  na equação (24) é definida por:

$$f(t) = \begin{cases} t^{1/3} & \text{para } t > (6/29)^2 \\ \frac{1}{3} \left(\frac{29}{6}\right)^2 t + 4/29 & \text{nos demais casos.} \end{cases} \quad (25)$$

A conversão entre o espaço XYZ e sRGB (14) é dada por:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,4124 & 0,3576 & 0,1805 \\ 0,2126 & 0,7152 & 0,0722 \\ 0,0193 & 0,1192 & 0,9505 \end{bmatrix} \begin{bmatrix} R_{sRGB} \\ G_{sRGB} \\ B_{sRGB} \end{bmatrix}. \quad (26)$$

Pode-se notar que há uma raiz cúbica na equação (25). Ela foi introduzida a fim de obter um sólido colorido isotrópico e homogêneo (14). Uma transformação do cubo sRGB no espaço de cores Lab é mostrado na Figura 22.

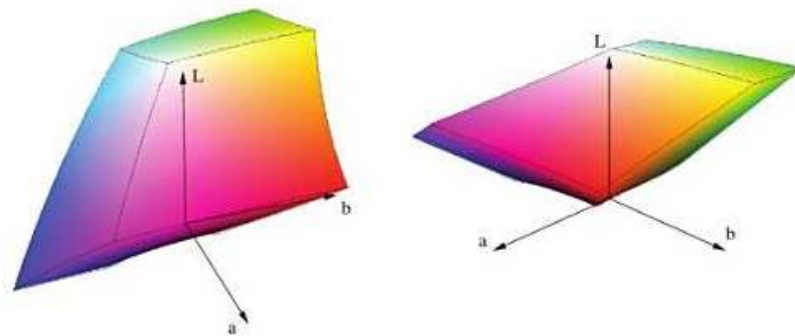


Figura 22 – Cubo sRGB representado nos eixos CIELAB em diferentes perspectivas. Fonte: (14).

A diferença entre as cores medida no espaço CIELAB, também conhecida como  $\Delta E$  (29), é dada por:

$$\Delta E = \sqrt{(L_1 - L_2)^2 + (a_1 - a_2)^2 + (b_1 - b_2)^2}, \quad (27)$$

onde  $(L_1, a_1, b_1)$  e  $(L_2, a_2, b_2)$  são duas cores diferentes no espaço CIELAB.

Outras fórmulas para se calcular a diferença entre as cores baseadas no espaço CIELAB foram desenvolvidas (47). Por exemplo, podem-se citar as fórmulas CIE94 e CIEDE2000, também chamadas de fórmulas de diferença de cor avançadas, que são as duas últimas fórmulas recomendadas pelo CIE (49). Para este trabalho o  $\Delta E$  foi escolhido porque é a forma mais utilizada entre os pesquisadores e também na indústria (47).

### 3.7 Utilização do $\Delta E$ na avaliação da qualidade de imagens

A utilização do espaço de cores CIELAB junto com o SSIM na avaliação de imagens coloridas é ilustrada pelo diagrama de blocos da Figura 23. As imagens RGB de referência e

distorcida são convertidas para CIELAB. Em seguida faz-se o cálculo da distância entre cores, o  $\Delta E$ , onde é gerado um mapa. Neste mapa é realizada uma busca pelos valores que são maiores que o limiar JND. Um  $\Delta E$  menor que um determinado valor JND indica que a diferença entre as cores não é perceptível pelo olho humano (49). As imagens convertidas para CIELAB também passam por um filtro Gaussiano suavizador, que simula o comportamento passa baixas do olho humano conforme explicado nas Seções 2.1.4 e 2.2.4, para em seguida ser calculado o mapa SSIM. Os mapas SSIM e  $\Delta E$  resultantes são comparados. A utilização do  $\Delta E$  não interfere na percepção da mudança de contraste, luminância e estrutura calculada pelo SSIM. Por fim, o MSSIM é calculado produzindo a avaliação objetiva.

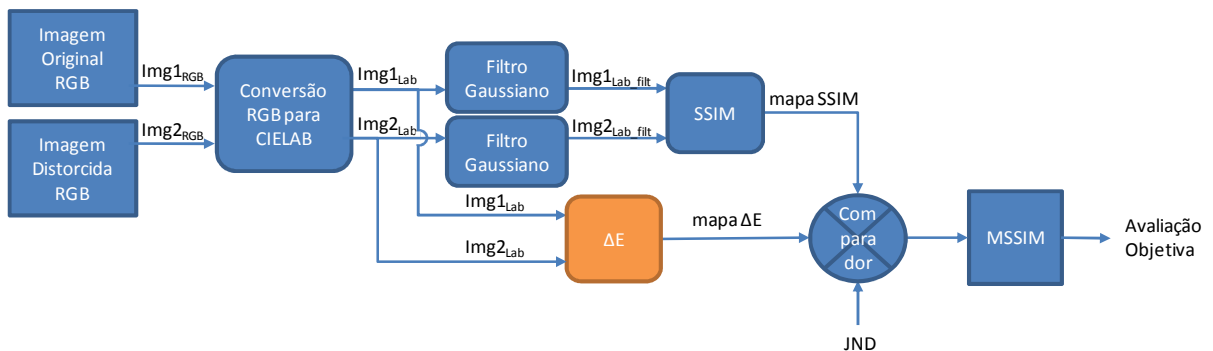


Figura 23 – Diagrama de blocos da aplicação do  $\Delta E$  juntamente ao SSIM na avaliação da qualidade de imagens.

O capítulo seguinte explica a metodologia de análise utilizada para a avaliação dos resultados dos testes objetivos apresentados, quando comparados aos resultados dos testes subjetivos disponibilizados pelos diferentes bancos de dados mencionados.

#### 4 METODOLOGIA DE ANÁLISE

A forma mais natural de avaliação do desempenho de uma métrica objetiva de avaliação da qualidade de imagens é por meio da comparação dos resultados obtidos pela avaliação objetiva com os dados de avaliações subjetivas como MOS ou DMOS. O gráfico de dispersão, que mostra as notas objetivas versus notas subjetivas, costuma ser feito. Entretanto, para a interpretação dos gráficos de dispersão, é necessária uma conversão destes em valores numéricos que permitam sua avaliação.

Para fazer a redução dos gráficos de dispersão para números, no âmbito da avaliação da qualidade de vídeo, inspirou-se na metodologia adotada pelo VQEG (*Video Quality Experts Group*) para a avaliação da qualidade de vídeos, que são seqüências de imagens, isto é, imagens que variam em função do tempo. Para a utilização deste método no âmbito da avaliação da qualidade de imagens, é necessário fazer uma adaptação, uma vez que a “dimensão” tempo só existe no vídeo e não em imagens.

O VQEG definiu uma abordagem de avaliação composta de três métricas estatísticas (50): o cálculo da correlação de Pearson, que mede a *monotonicidade*; o cálculo do RMSE, a raiz quadrada do erro médio quadrático, que caracteriza a *precisão*; e por fim computa-se a razão *outlier*, obtendo, assim, a *consistência* da métrica subjetiva. O VQEG também definiu que seja realizado um mapeamento dos dados objetivos para a escala subjetiva, passo também conhecido como ajuste, antes de se computar as métricas estatísticas. A Figura 24 mostra o esquema de como é feito este mapeamento.

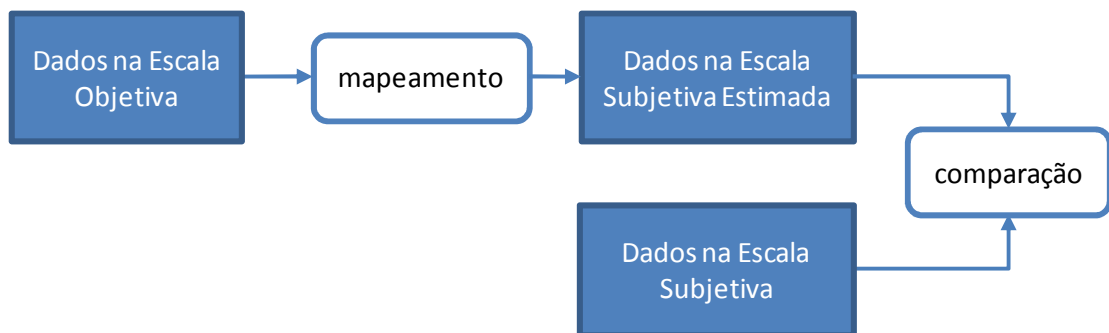


Figura 24 – Mapeamento dos dados objetivos para a escala subjetiva.

Para facilitar a comparação de diferentes bancos de dados, foi realizada uma normalização das métricas subjetivas adequando-as à escala entre 0 e 1. O banco de dados LIVE, cujas notas da avaliação subjetiva DMOS variam na faixa entre 0 e 100 e aumentam conforme a qualidade da imagem sendo avaliada diminui, foi normalizado fazendo-se

$$DMOS\_LIVE_N = \frac{100 - DMOS\_LIVE}{100}. \quad (28)$$

As notas subjetivas MOS do banco de dados IVC, que variam na faixa entre 1 e 5 e aumentam conforme a qualidade da imagem sendo avaliada aumenta, foram normalizadas fazendo-se

$$MOS\_IVC_N = \frac{MOS\_IVC - 1}{4}. \quad (29)$$

Para o banco de dados TID, de forma similar ao que ocorre com o banco de dados IVC, as notas subjetivas MOS que variam na faixa entre 0 e 9 foram normalizadas fazendo-se

$$MOS\_TID_N = \frac{MOS\_TID}{9}. \quad (30)$$

A formulação e a descrição de cada uma das métricas estatísticas utilizadas para a avaliação do desempenho das métricas estão a seguir.

#### 4.1 Mapeamento para a escala subjetiva

Antes do cálculo das métricas de desempenho é realizado um mapeamento dos dados objetivos para a escala subjetiva. Uma função de ajuste não-linear, sugerida pelo VQEG, que desempenha bem empiricamente (50), é o polinômio cúbico:

$$subj_p = a * (obj)^3 + b * (obj)^2 + c * obj + d, \quad (31)$$

onde  $subj_p$  são os dados subjetivos preditos, isto é, os dados objetivos SSIM depois de mapeados para a escala subjetiva e  $obj$  são os dados objetivos SSIM. Os pesos  $a, b, c$  e a constante  $d$  são obtidos pelo ajuste da função aos dados subjetivos MOS/DMOS. A função de mapeamento minimiza o erro médio quadrático entre  $subj_p$  e  $subj$ , sendo  $subj$  os dados subjetivos MOS/DMOS, conforme mostra a equação (32):

$$\min(E[(subj_p - subj)^2]). \quad (32)$$

Em um bom mapeamento de notas subjetivas em objetivas a função de ajuste deve ser monótona dentro da faixa de possíveis valores, de modo que dois valores subjetivos preditos não sejam mapeados em um mesmo valor subjetivo real. Assim, é desejável que a curva de mapeamento apresente derivada sempre de mesmo sinal.

## 4.2 Correlação de Pearson

A correlação de Pearson mede a relação linear entre o desempenho da métrica objetiva e os dados subjetivos. Sua fórmula é dada por:

$$R = \frac{\sum_{i=1}^N (subj_i - \overline{subj})(subj_{i_p} - \overline{subj_p})}{\sqrt{\sum (subj_i - \overline{subj})^2} * \sqrt{\sum (subj_{i_p} - \overline{subj_p})^2}}, \quad (33)$$

onde  $subj_i$  e  $subj_{i_p}$  são respectivamente os dados subjetivos MOS/DMOS e os dados subjetivos preditos para cada imagem;  $\overline{subj}$  e  $\overline{subj_p}$  são as respectivas médias e  $N$  é a quantidade de imagens sendo avaliadas.

A correlação de Pearson pode produzir resultados entre -1 e 1. Quanto mais próximo de 1 for o módulo do resultado mais correlacionados estão os dados objetivos e os dados subjetivos. Valores positivos indicam que os dados sendo comparados são diretamente relacionados enquanto valores negativos indicam que os dados são inversamente relacionados.

## 4.3 RMSE

O RMSE (*Root Mean Square Error*) é uma forma de medir a precisão da métrica objetiva. Este permite quantificar o erro entre os dados subjetivos e subjetivos preditos. O erro entre a medida subjetiva  $subj_i$  e a medida subjetiva predita  $subj_{i_p}$  é dado por:

$$erro_i = subj_i - subj_{i_p}, \quad (34)$$



onde  $subj_i$  são os dados subjetivos MOS/DMOS e  $subj_{i_p}$  são os dados subjetivos preditos. O RMSE é a raiz do erro médio quadrático, dado por:

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N erro_i^2}, \quad (35)$$

onde N denota a quantidade de imagens sendo avaliadas. Quanto menor for o RMSE menor é o erro entre as métricas subjetiva e objetiva.

#### 4.4 Razão Outlier

A consistência da métrica objetiva é avaliada pela razão *outlier*, ou OR. Esta fornece uma medida dos pontos que estão muito distantes do ajuste  $subj_p$  ou do comportamento médio do total das N imagens sendo avaliadas. O OR é medido em porcentagem e sua fórmula é dada por:

$$OR(\%) = \frac{\text{número total de outliers}}{N}, \quad (36)$$

onde um *outlier* é um ponto para o qual

$$erro_i > K * \sigma_{erro}, \quad (37)$$

onde  $\sigma_{erro}$  é o desvio padrão do erro entre a medida subjetiva  $subj_i$  e a medida subjetiva predita  $subj_{i_p}$  para todas as imagens e para um intervalo de confiança de 95% e supondo uma distribuição Gaussiana,  $K = 1,96$  (50).

O OR indica o quão aplicável é a nota subjetiva predita: quanto menor for o OR, menos pontos estão fora do comportamento médio.

#### 4.5 Percentis 95 e 99

Além do cálculo das métricas estatísticas sugeridas pelo VQEG, foram também calculados os percentis 95 e 99 do erro entre o MOS/DMOS mapeados para a escala subjetiva e os dados subjetivos, que dão uma dimensão do tamanho máximo assumido por 95 e 99 por cento da diferença entre os dados.

O percentil é definido como o ponto abaixo do qual a porcentagem especificada de observações ocorre. Neste caso, os percentis 95 e 99 do erro são os valores da diferença entre a medida subjetiva  $subj_i$  e a medida subjetiva predita  $subj_{i,p}$  abaixo da qual se encontram 95 e 99 por cento das amostras. O percentil 95 é definido por:

$$P(95) = arg \left( \int_{-\infty}^x f_{erro}(erro) dx = 0,95 \right), \quad (38)$$

onde  $f_{erro}(erro)$  é a função densidade de probabilidade do erro, definido na equação (34). De modo análogo é definido o percentil 99.

O percentil fornece informação sobre o espalhamento dos dados, sendo possível obter a dimensão da diferença máxima assumida pelo erro. Quanto menor os valores dos percentis, menos distantes estão as medidas subjetivas  $subj_i$  das medidas subjetivas preditas  $subj_{i,p}$ .

O capítulo seguinte investiga o comportamento do SSIM aplicado aos canais de cores, usando a metodologia descrita neste capítulo para análise dos resultados.

## 5 COMPORTAMENTO DO SSIM EM DIFERENTES ESPAÇOS DE CORES

A fim de verificar o comportamento do SSIM para avaliação de imagens coloridas, neste capítulo são apresentados os resultados obtidos para a aplicação do SSIM nos diferentes canais individuais dos espaços de cores RGB, YCbCr,  $L\alpha\beta$  e CIELAB, descritos no Capítulo 3.

Os testes foram divididos em três partes e cada parte é apresentada em uma seção. Na primeira, o SSIM é aplicado a cada um dos canais dos espaços de cores individualmente. Na segunda parte, são estudadas combinações alternativas dos componentes do SSIM – luminância, contraste e estrutura. Por fim, é realizada uma breve análise do comportamento do SSIM por tipo de distorção.

A avaliação dos resultados é feita seguindo a metodologia definida no Capítulo 4: primeiramente é feita uma análise dos gráficos de dispersão, das curvas de ajuste e da correlação de Pearson para os resultados de cada espaço de cores. Em seguida são analisados o RMSE, a razão *outlier* e os percentis 95 e 99.

### 5.1 SSIM aplicado aos canais individuais

O objetivo deste experimento é avaliar o comportamento do SSIM aplicado individualmente em cada canal dos espaços de cores RGB, YCbCr,  $L\alpha\beta$  e CIELAB. A Figura 25 apresenta um esquema do processo: a primeira etapa é a conversão das imagens original e distorcida (em RGB) para o espaço de cores que se deseja trabalhar – YCbCr,  $L\alpha\beta$  ou CIELAB. No caso da avaliação do espaço RGB a etapa de conversão não é necessária. Em seguida, as componentes do espaço de cores passam por um filtro Gaussiano, cujas propriedades estão descritas na Seção 2.2.4. O SSIM é calculado entre o canal filtrado da imagem distorcida que se deseja avaliar e o correspondente da imagem de referência.

Espera-se que nos espaços projetados com um canal correspondente à luminância – YCbCr,  $L\alpha\beta$  e CIELAB, estes canais (respectivamente Y, L e L) concentrem informação relativamente aos canais do espaço RGB. Assim, espera-se que estes canais apresentem melhores desempenhos que os canais RGB. O mesmo se espera desses canais relativamente aos canais de cores (Cb,Cr,  $\alpha,\beta$  e a,b), uma vez que o SSIM foi desenvolvido para ser utilizado na luminância (26).

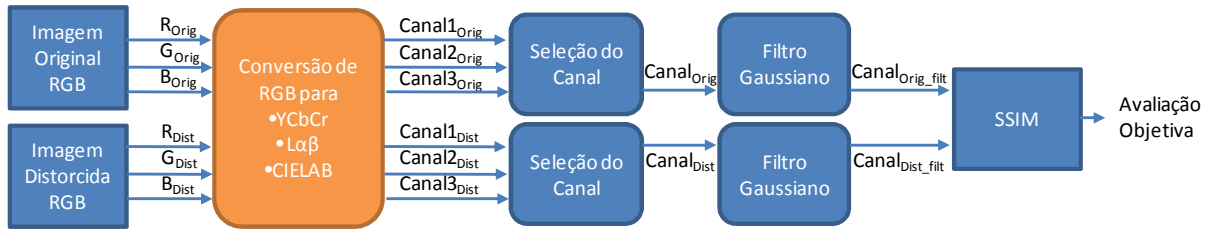


Figura 25 – Diagrama de blocos do SSIM aplicado aos canais individuais.

Com foco na investigação do espaço de cores mais adequado para ser utilizado com o SSIM na avaliação de imagens coloridas, a análise dos dados é feita separadamente por espaço de cores. Em seguida é apresentada uma comparação dos resultados obtidos.

### 5.1.1 Espaço RGB

Em uma imagem representada em RGB não existe concentração de informação em nenhum dos canais em especial. Geralmente há uma distribuição proporcional entre os canais R, G e B. De fato, é possível perceber que nos gráficos de dispersão, onde as coordenadas dos pontos correspondem aos valores obtidos pela métrica objetiva e aos valores subjetivos, tanto para o banco de dados LIVE da Figura 26 quanto nas curvas de ajuste dos bancos de dados IVC e TID da Figura 30 a serem discutidos mais a frente, os canais R, G e B são bastante semelhantes. Desta forma, não há vantagem significativa na utilização de qualquer um dos canais em especial.

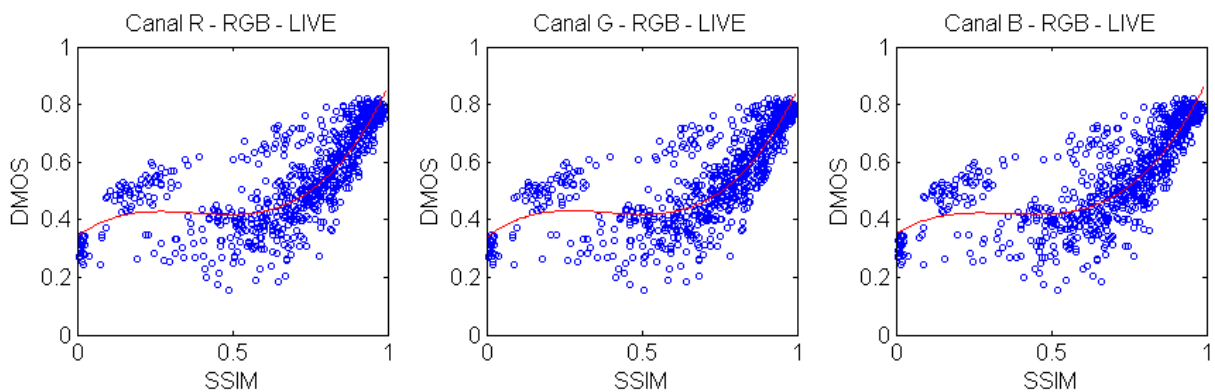


Figura 26 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores RGB, usando o banco de dados LIVE.

Nos gráficos de dispersão da Figura 26 nota-se maior similaridade entre SSIM e DMOS para valores mais altos do SSIM e do DMOS, porém para valores mais baixos a dispersão é grande.

Nota-se que as curvas de ajuste do RGB para o LIVE não são monótonas crescentes, isto é, não apresentam derivada positiva em todos os pontos. A mesma característica é ainda mais evidente nas curvas de ajuste para o IVC e TID na Figura 30. Isto indica que este espaço não proporciona um bom mapeamento das notas objetivas para a escala subjetiva.

As correlações entre as notas subjetivas preditas e as notas subjetivas para os canais R, G e B no banco de dados LIVE ficaram entre 0,8289 e 0,8305 e são mostradas na Figura 31.

### 5.1.2 Espaço YCbCr

No espaço YCbCr a informação é concentrada no canal Y, que representa a luminância, em relação aos canais Cb e Cr, que carregam informação de cor (10).

Muitos dispositivos possuem entradas no formato YCbCr ou no seu equivalente YPbPr. Neste contexto, caso a aplicação do SSIM no canal Y produzisse um bom resultado, poderia ser interessante usar esse espaço de cores diretamente na avaliação da qualidade ao invés de converter o sinal para outro espaço de cores, reduzindo o custo computacional.

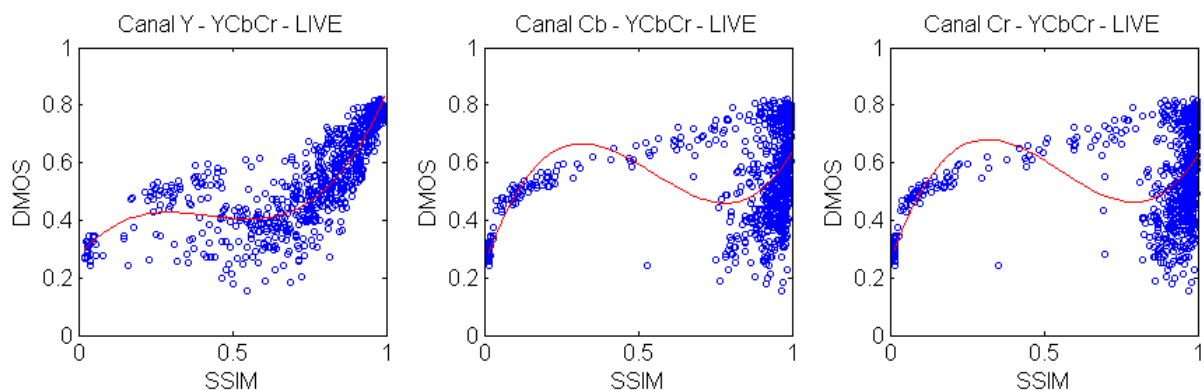


Figura 27 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores YCbCr.

Na Figura 27 verifica-se que o canal Y não apresenta uma “boa distribuição” dos pontos, pois há uma concordância maior para valores mais altos de SSIM e uma dispersão maior para valores de SSIM mais baixos. A curva de ajuste do canal Y não é monótona, e o mesmo ocorre com as curvas de ajuste referentes aos bancos de dados IVC e TID, mostrados na Figura 30. Os canais Cb e Cr apresentaram resultados ainda piores, pois os pontos estão em sua maior parte concentrados para valores muito altos do SSIM e apenas uma pequena parcela dos pontos está distribuída ao longo de toda a faixa de valores SSIM. As curvas de ajuste dos canais Cb e Cr são ruins, pois não apresentam monotonicidade, com mapeamento pior que o das curvas de ajuste do canal Y ou dos canais RGB. Quando o gráfico de dispersão apresenta

uma concentração dos pontos nas extremidades ao invés de um alinhamento na diagonal indica que a métrica objetiva não fornece resultados semelhantes aos obtidos subjetivamente. Isto indica que o canal do espaço de cores usado não representa uma boa escolha para avaliação objetiva.

A correlação alcançada pelo canal Y no banco de dados LIVE, mostrada na Figura 31, foi de 0,8685, pouco mais alta que as correlações alcançadas pelos canais R, G e B anteriormente estudados.

### 5.1.3 Espaço $L\alpha\beta$

No espaço  $L\alpha\beta$ , assim como no espaço YCbCr, uma quantidade maior de informação, referente à luminância, é concentrada no canal L, enquanto os canais  $\alpha$  e  $\beta$  carregam a informação da cor.

Utilizando o espaço de cores  $L\alpha\beta$  e o banco de dados LIVE, nos gráficos de dispersão da Figura 28 nota-se que os canais  $\alpha$  e  $\beta$  não mostram bons resultados: o canal  $\alpha$ , apesar de estar distribuído por toda a faixa de valores SSIM, apresenta uma concentração dos pontos muito grande para valores mais baixos e muito pequena para valores mais altos, enquanto que no canal  $\beta$  ocorre o oposto apresentando uma concentração grande para valores altos de SSIM e pequena para valores baixos de SSIM. Já o canal L apresenta concentração dos pontos em diagonal com boa distribuição por toda extensão de valores SSIM.

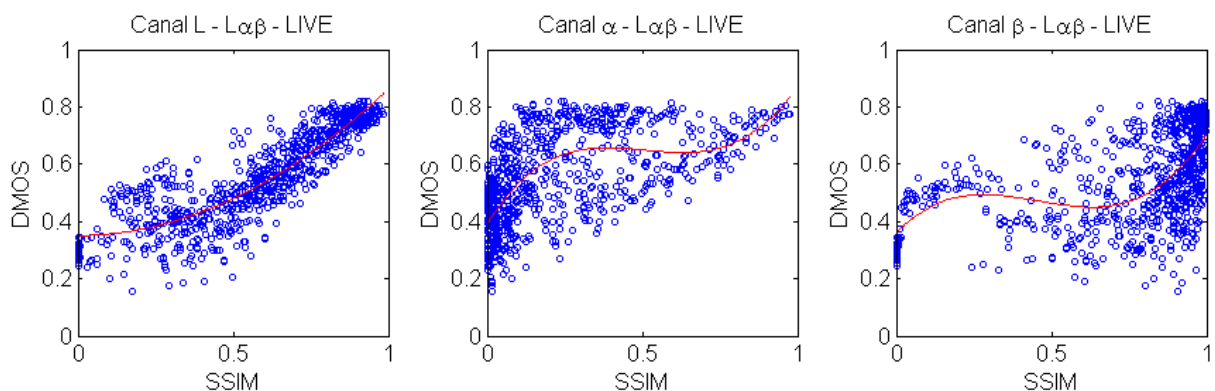


Figura 28 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais do espaço de cores  $L\alpha\beta$ .

A curva de ajuste do canal L tem uma curvatura muito mais suave que a curva de ajuste do canal Y ou dos canais RGB. Para o banco de dados LIVE a curva é monótona, o que indica um bom mapeamento das notas objetivas para a escala subjetiva, porém nos bancos de dados

IVC e TID há uma mudança no sinal da derivada nas extremidades, o que indica a não monotonicidade das curvas de ajuste do canal L para estes bancos de dados.

O canal L do espaço  $L\alpha\beta$  apresentou a maior correlação entre todos os espaços estudados para o banco de dados LIVE, próxima da correlação obtida para o canal Y, valendo 0,8797. Para os bancos de dados IVC e TID a correlação para o canal L do espaço  $L\alpha\beta$  foi ligeiramente menor que a correlação alcançadas para o canal Y do espaço YCbCr.

#### 5.1.4 Espaço CIELAB

O espaço de cores CIELAB é um espaço que pretende ser perceptivo, isto é, aproximar a forma como o SVH percebe as cores. De forma semelhante aos espaços YCbCr e  $L\alpha\beta$ , o espaço CIELAB também tem um canal de luminância, L, que porta a maior parte da informação, e dois canais de crominância, a e b, que portam menor quantidade de informação.

Nos gráficos de dispersão entre DMOS e SSIM para o LIVE, Figura 29, os canais a e b não apresentam um bom resultado, pois concentram a maior parte dos pontos para valores mais baixos de SSIM. Entretanto, no canal L os pontos aparecem bem distribuídos por toda a faixa de valores SSIM em diagonal. A curva de ajuste desse canal também mostra um bom mapeamento, pois é monótona e bem alinhada.

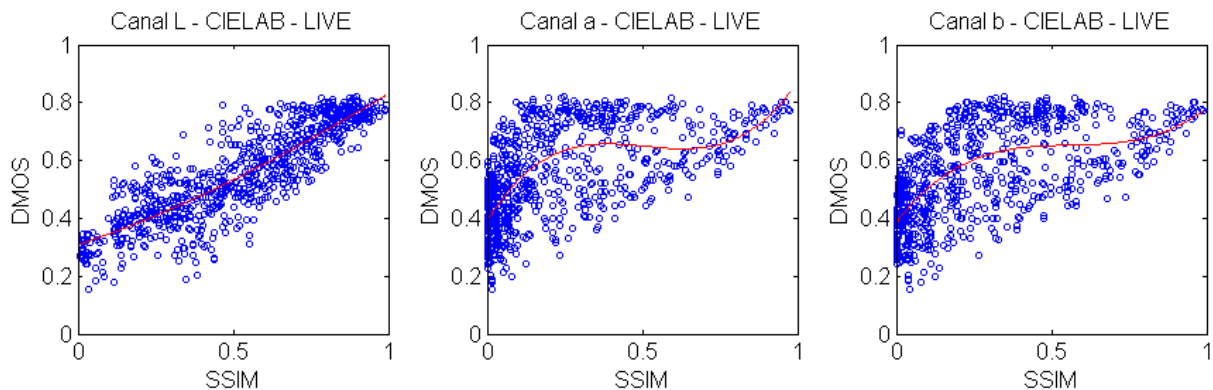


Figura 29 – Gráficos de dispersão do SSIM pela avaliação subjetiva DMOS dos canais individuais dos espaços de cores RGB, YCbCr,  $L\alpha\beta$  e CIELAB.

#### 5.1.5 Comparação entre os espaços

A Figura 30 apresenta as curvas de ajuste entre as notas subjetivas e o SSIM para cada canal dos espaços de cores estudados para os bancos de dados IVC, TID e LIVE para fins de comparação. Para o banco de dados LIVE a figura mostra uma síntese das curvas mostradas nas Figuras 26 a 29.

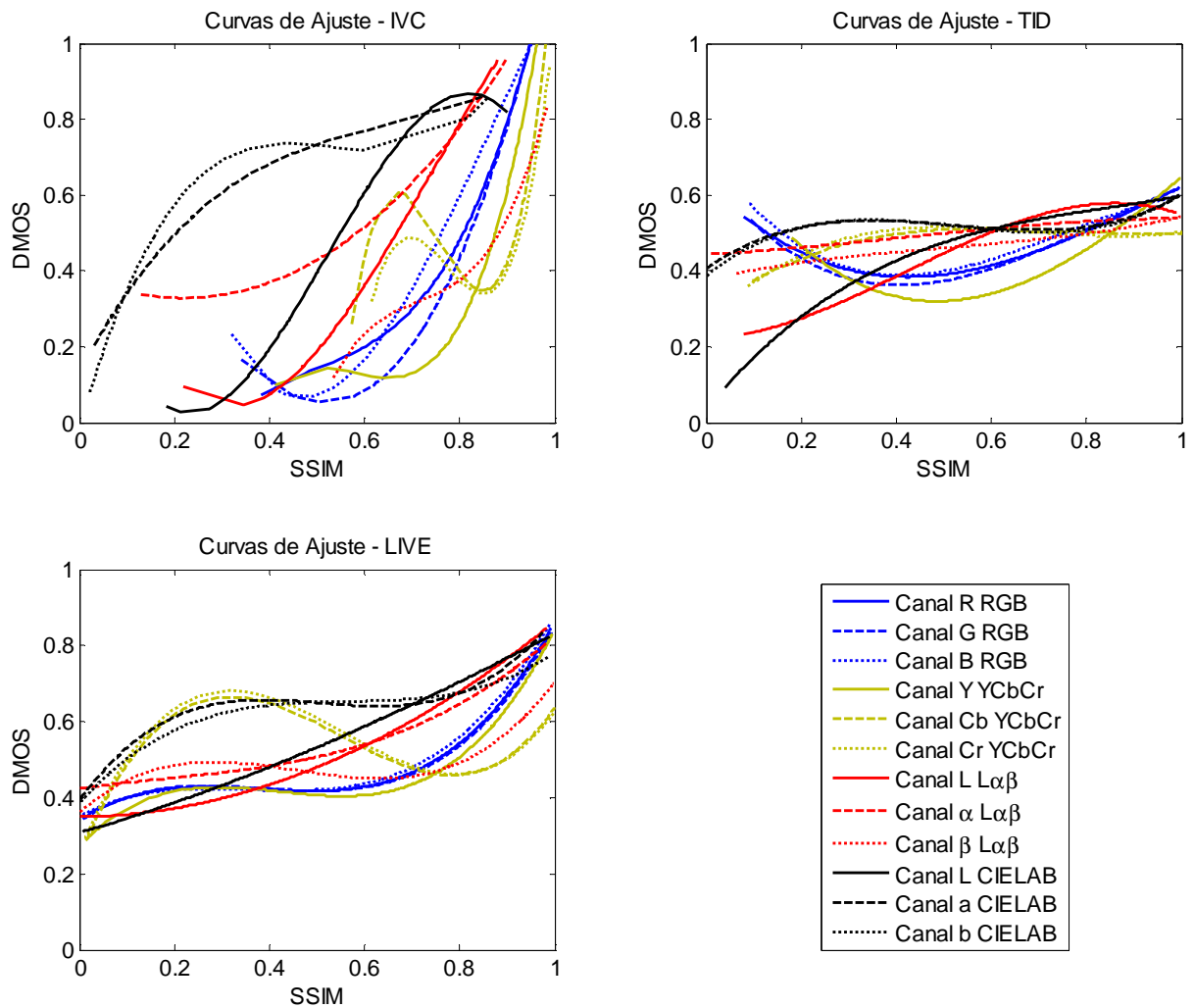


Figura 30 – Curvas de ajuste do SSIM aplicado aos canais individuais nos bancos de dados IVC, TID e LIVE.

Para o banco de dados IVC, a curva de ajuste do canal L do espaço CIELAB tem as extremidades curvas, não apresentando monotonicidade. Este efeito pode ser devido ao fato de que este banco de dados tem poucas imagens e a dispersão de alguns pontos nas extremidades provoca este efeito. Já a curva de ajuste do mesmo canal L para o banco de dados TID é monótona. A curva de ajuste do canal L do espaço  $L\alpha\beta$  para o banco de dados IVC apresenta um bom alinhamento, porém não apresenta monotonicidade devido a uma curva na extremidade inferior. A curva de ajuste correspondente ao mesmo canal para o banco de dados TID apresenta um alinhamento semelhante ao da curva de ajuste do canal L do CIELAB, porém não é monótona devido a uma curvatura na extremidade superior.

Os resultados obtidos para os bancos de dados IVC e TID, de que os canais L dos espaços CIELAB e  $L\alpha\beta$  produzem melhores curvas de ajuste, corroboram com os resultados obtidos



para o banco de dados LIVE. Para todos os bancos de dados a curva do canal L do espaço CIELAB possui uma boa faixa de variação, seguido do canal L do espaço  $L\alpha\beta$ . Os demais canais apresentaram falta de monotonicidade ou maior desalinhamento com a diagonal, e se mostraram menos adequados por oferecerem um mapeamento ruim.

A Figura 31 mostra as correlações alcançadas entre as notas subjetivas e subjetivas previstas com a utilização de cada canal dos espaços de cores. Para o banco de dados LIVE as maiores correlações foram alcançadas usando os canais L dos espaços  $L\alpha\beta$  e CIELAB e o canal Y do espaço YCbCr. De forma semelhante para os bancos de dados IVC e TID os canais L dos espaços  $L\alpha\beta$  e CIELAB e canal Y do espaço YCbCr obtiveram as maiores correlações.

Com relação aos canais de cores, pode-se verificar na Figura 31 que os canais a e b apresentaram correlações mais altas que os canais  $\alpha$  e  $\beta$  e que os canais Cb e Cr.

As correlações alcançadas usando os canais R, G e B usando o banco de dados LIVE foram altas, porém menores que as obtidas usando os canais L dos espaços CIELAB e  $L\alpha\beta$  e Y do YCbCr.

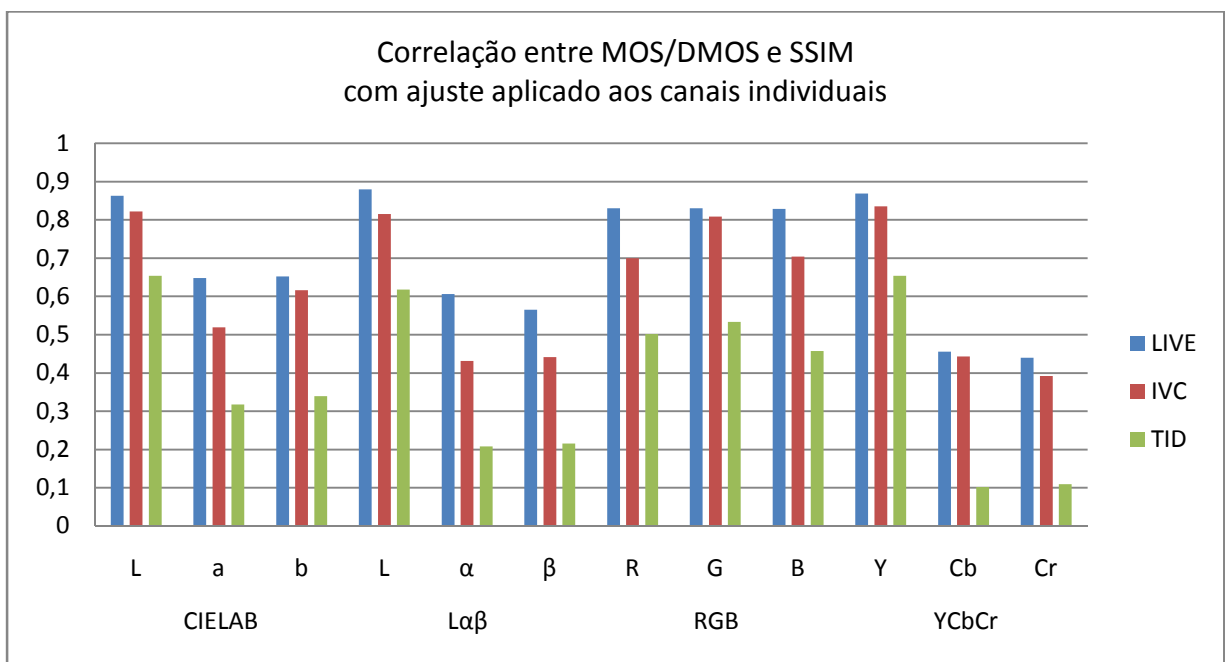


Figura 31 – Correlações entre MOS/DMOS e SSIM com ajuste aplicado aos canais individuais dos espaços de cores.

Após a observação dos gráficos de dispersão e das curvas de ajuste entre DMOS/MOS e SSIM com ajuste, a verificação dos dados RMSE, razão *outlier* e percentis 95 e 99 fornecem importantes parâmetros que complementam a análise dos espaços de cores. A Tabela 4 mostra

os dados das métricas obtidos pelo SSIM aplicado individualmente a cada canal dos espaços de cores dos três bancos de dados. Os três melhores valores de cada métrica estão destacados.

Pode-se notar para os três bancos de dados que os três canais que apresentaram os melhores resultados, isto é, altas correlações e baixos RMSE, OR e percentis, foram o canal Y do espaço YCbCr e os canais L dos espaços CIELAB e  $L\alpha\beta$ .

Os canais do espaço RGB, apesar de terem obtido boas correlações para o banco de dados LIVE, entre 0,8289 a 0,8305, não estão no grupo dos melhores resultados para nenhum dos bancos de dados testados. Além disso, as curvas de ajuste não apresentaram um comportamento monótono. Desta forma conclui-se que os canais Y do YCbCr e os canais L do CIELAB e L do  $L\alpha\beta$  são melhores opções que os canais RGB individualmente, ainda que “conjuntamente” eles pareçam promissores.

Para o banco de dados LIVE o canal Y apresentou alta correlação, 0,8685, baixo RMSE, 0,0032, e baixos percentis e *outlier ratio*. Comportamento semelhante foi observado para os bancos de dados IVC e TID. A curva de ajuste analisada anteriormente também revelou não monotonicidade, o que caracteriza um mapeamento ruim, além dos gráficos de dispersão não terem apresentado boa distribuição por toda a faixa de possíveis valores de SSIM.

Tabela 4– Medidas de avaliação do SSIM aplicado aos canais dos espaços de cores

Espaço	CIELAB	CIELAB	CIELAB	$L\alpha\beta$	$L\alpha\beta$	$L\alpha\beta$	RGB	RGB	RGB	YCbCr	YCbCr	YCbCr
Canal	L	a	b	L	$\alpha$	$\beta$	R	G	B	Y	Cb	Cr
<b>LIVE</b>												
Correlação	<b>0,8633</b>	0,6483	0,6519	<b>0,8797</b>	0,6064	0,5656	0,8305	0,8300	0,8289	<b>0,8685</b>	0,4560	0,4395
RMSE	<b>0,0033</b>	0,0075	0,0075	<b>0,0029</b>	0,0082	0,0088	0,0040	0,0040	0,0041	<b>0,0032</b>	0,0103	0,0105
OR	<b>0,00%</b>	0,13%	0,26%	<b>0,00%</b>	0,51%	0,90%	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	1,03%	1,41%
Perc. 95	<b>0,1689</b>	0,2292	0,2207	<b>0,1595</b>	0,2446	0,2518	0,1952	0,1960	0,1896	<b>0,1663</b>	0,2641	0,2687
Perc. 99	<b>0,2272</b>	0,2882	0,2929	<b>0,2121</b>	0,3072	0,3152	0,2516	0,2524	0,2524	<b>0,2118</b>	0,3164	0,3222
<b>IVC</b>												
Correlação	<b>0,8222</b>	0,5195	0,6161	<b>0,8151</b>	0,4312	0,4412	0,6994	0,8087	0,7040	<b>0,8355</b>	0,4430	0,3922
RMSE	<b>0,0156</b>	0,0352	0,0299	<b>0,0162</b>	0,0392	0,0388	0,0246	0,0167	0,0243	<b>0,0146</b>	0,0387	0,0408
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	0,83%	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	<b>0,3432</b>	0,4487	0,4633	<b>0,3623</b>	0,4950	0,4917	0,4642	0,4081	0,4257	<b>0,3863</b>	0,4559	0,4511
Perc. 99	<b>0,4617</b>	0,5382	0,5136	<b>0,4677</b>	0,5533	0,5948	0,5205	<b>0,4690</b>	0,5262	0,5107	0,5050	0,5174
<b>TID</b>												
Correlação	<b>0,6540</b>	0,3180	0,3398	<b>0,6175</b>	0,2085	0,2156	0,5015	0,5333	0,4574	<b>0,6540</b>	0,1016	0,1099
RMSE	<b>0,0064</b>	0,0100	0,0098	<b>0,0069</b>	0,0106	0,0106	0,0083	0,0080	0,0088	<b>0,0064</b>	0,0110	0,0110
OR	<b>0,76%</b>	2,76%	2,59%	<b>1,12%</b>	3,12%	3,29%	2,12%	2,24%	2,41%	<b>1,53%</b>	3,65%	3,35%
Perc. 95	<b>0,2153</b>	0,2606	0,2608	<b>0,2271</b>	0,2647	0,2685	0,2450	0,2431	0,2533	<b>0,2263</b>	0,2741	0,2723
Perc. 99	<b>0,2810</b>	0,3847	0,3770	<b>0,2956</b>	0,4171	0,4140	0,3556	0,3413	0,3643	<b>0,3092</b>	0,4428	0,4400

O canal L do espaço  $L\alpha\beta$  apresentou a correlação mais alta para o banco de dados LIVE, 0,8797, os menores RMSE, 0,0029, percentil 95, 0,1595, e OR de 0%. Para o banco de dados IVC e TID este canal também apresenta baixos RMSE, OR e percentis.

Para o banco de dados LIVE o canal L do CIELAB apresentou alta correlação de 0,8633, pouco abaixo da apresentada pelo canal L do  $L\alpha\beta$ , e baixos valores de RMSE, 0,0033, percentis 95 e 99 iguais a 0,1689 e 0,2272 respectivamente e OR de 0%. Comportamento semelhante foi observado para os outros bancos de dados, com os menores OR e percentis.

Além desses fatos, o CIELAB possui uma importante característica: a possibilidade de calcular a distância entre as cores. Esta característica pode ser usada para identificar regiões da imagem distorcida cujas cores não são perceptivelmente distintas das existentes na imagem original, e aplicado no sentido de melhorar a métrica objetiva, desconsiderando no cálculo da métrica objetiva as regiões da imagem cuja diferença de cores seja imperceptível. Este princípio é explorado no Capítulo 7.

A seção seguinte aprofunda a análise dos canais dos espaços de cores que apresentaram os melhores desempenhos – Y do espaço YCbCr, L do espaço  $L\alpha\beta$  e L do espaço CIELAB, utilizando o SSIM completo com foco nas distorções.

## 5.2 Análise do SSIM por tipo de distorção

Fazendo a identificação dos tipos de distorção das imagens, é possível verificar os diferentes comportamentos do SSIM para cada tipo de distorção. Assim, identificando o padrão dos pontos referentes aos diferentes tipos de distorção no gráfico de espalhamento é possível ajustar a curva de modo a melhorar a correlação da métrica objetiva com a avaliação subjetiva.

Neste experimento, os canais individuais que apresentaram os melhores resultados na Seção 5.1 – Y do espaço YCbCr, L do espaço CIELAB e L do espaço  $L\alpha\beta$ , são analisados.

A Figura 32 apresenta os gráficos de dispersão do SSIM e da métrica subjetiva para o banco de dados LIVE para esses canais, trazendo marcadores distintos para distorções distintas – compressão JPEG, compressão JPEG2000, adição de ruído Gaussiano, embaçamento e “erro de bits”.

Para o canal Y do espaço YCbCr os pontos referentes ao embaçamento, à compressão JPEG, à compressão JPEG2000 e um pouco menos ao erro de bit, ficaram concentrados em valores altos do SSIM. Já os pontos referentes ao ruído Gaussiano ficaram distribuídos por

toda a faixa de valores SSIM, mas ficaram destacados do restante dos pontos, afastando-se da curva de ajuste para valores intermediários do SSIM.

Para o canal L do espaço  $L\alpha\beta$  os pontos em função das distorções ficaram melhor comportados que para o canal Y. Os pontos referentes ao embaçamento, às compressões JPEG e JPEG2000 e ao erro de bit ficaram mais concentrados especialmente na faixa de valores mais altos do SSIM. Também neste canal, os pontos correspondentes ao ruído Gaussiano ficaram mais distantes do grupo principal, com valores SSIM geralmente menores que o DMOS.

Para o canal L do espaço CIELAB os pontos correspondentes a todos os tipos de distorções ficam bem distribuídos ao longo da diagonal, especialmente os pontos correspondentes ao embaçamento e ao erro de bit. Os pontos que se afastam um pouco do grupo principal, onde o SSIM produz valores mais baixos que o DMOS, são correspondentes ao ruído Gaussiano e alguns poucos pontos são correspondentes à compressão JPEG2000. De forma contrária, alguns pontos cujo SSIM é maior que o DMOS correspondem às imagens comprimidas com JPEG e a algumas imagens comprimidas com JPEG2000.

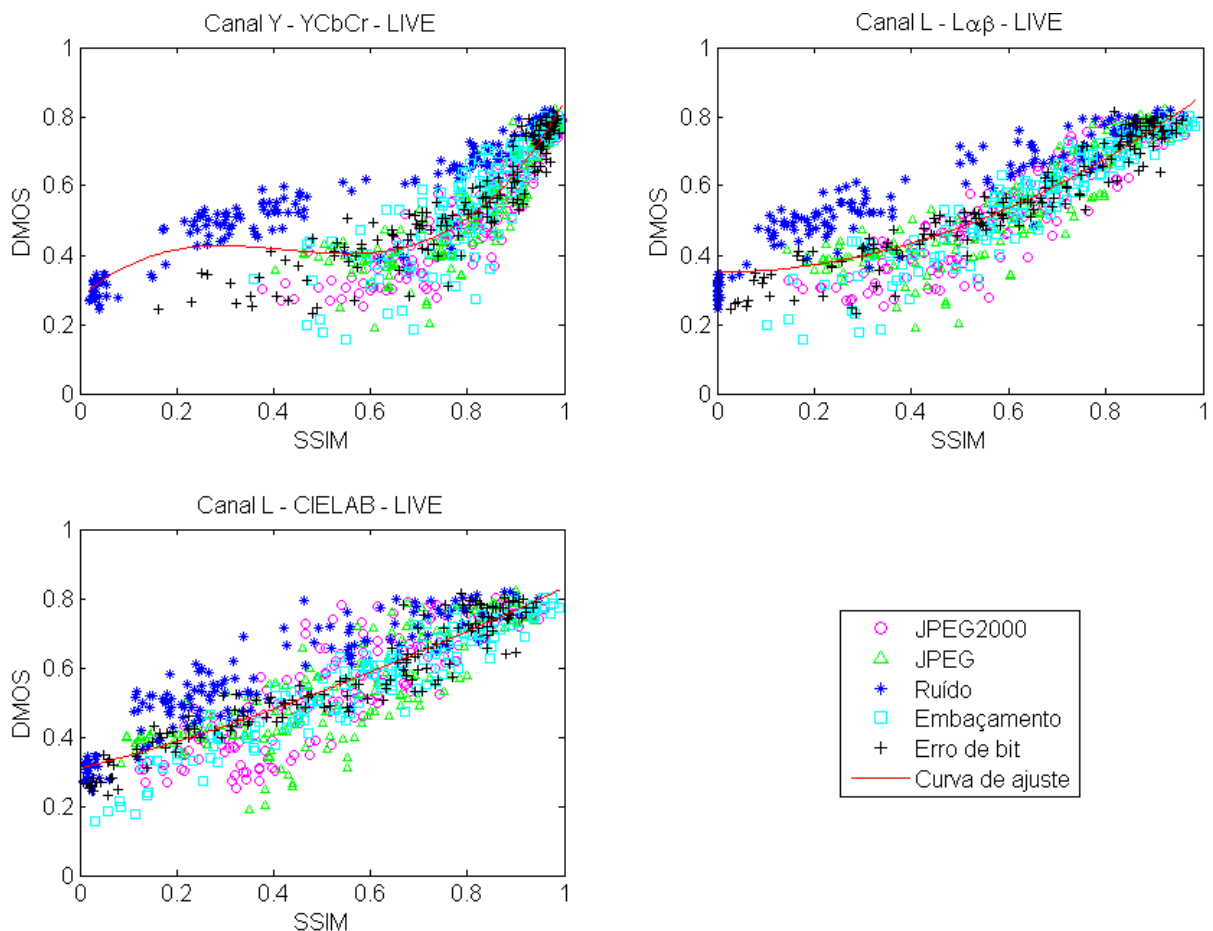


Figura 32 – Gráficos de dispersão do banco de dados LIVE identificados por distorção.

### 5.2.1 Curvas de ajuste por distorções

Caso as métricas de análise apresentem melhores resultados quando calculadas por distorção, um sistema especialista que identifique o tipo de distorção presente na imagem e aplique parâmetros específicos para esta informação pode obter melhores resultados objetivos.

Para analisar o desempenho do SSIM por distorção são plotadas as curvas de ajuste e calculadas as métricas de análise para cada distorção. É utilizado o banco de dados LIVE e o canal L do CIELAB, que apresentou a melhor distribuição dos pontos por toda a faixa de possíveis valores SSIM.

A Figura 33 apresenta os gráficos de dispersão para o SSIM aplicado ao grupo de imagens referentes a cada tipo de distorção e DMOS com suas respectivas curvas de ajuste. A Tabela 5 apresenta os dados de correlação, RMSE, razão *outlier* e percentis 95 e 99 calculados para cada tipo de distorção.

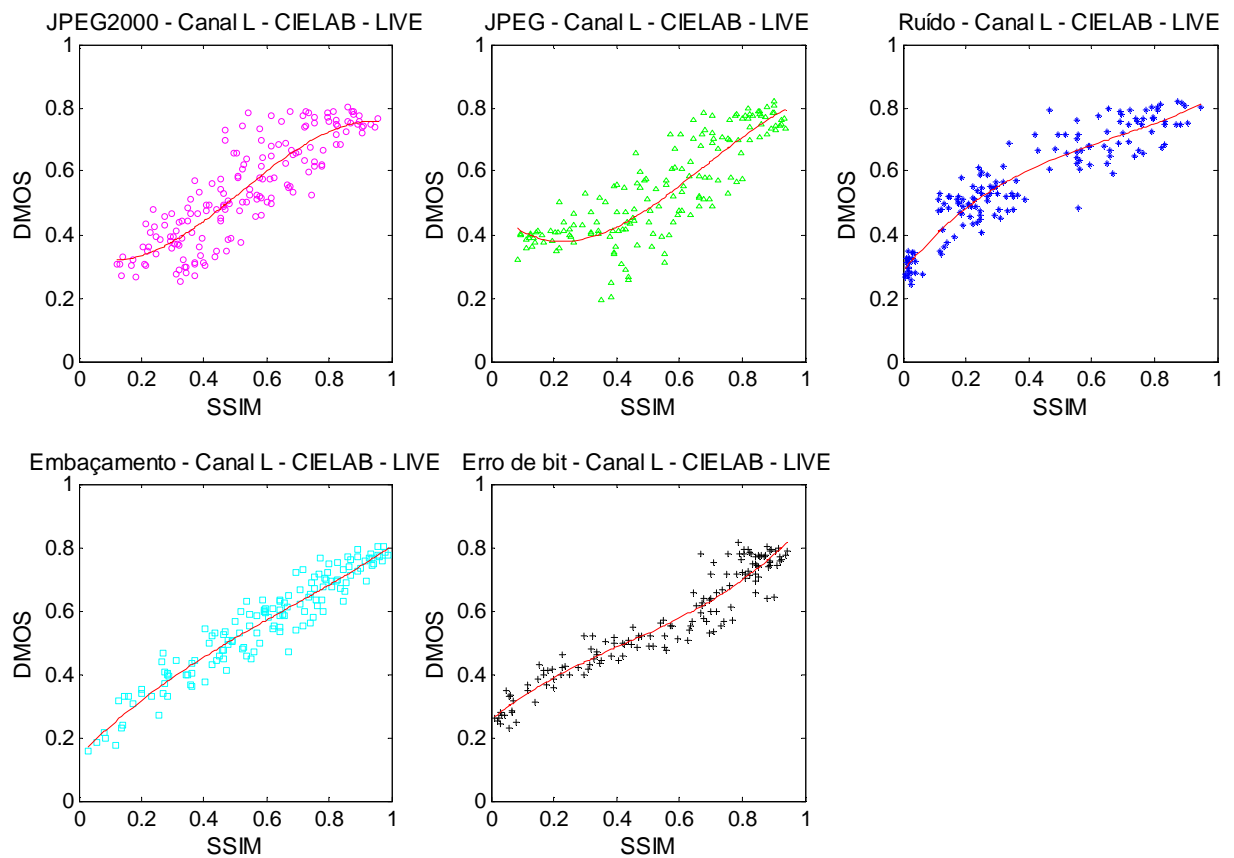


Figura 33 – Curva de ajuste por distorção no canal L do CIELAB.

Verifica-se que para imagens distorcidas por ruído Gaussiano, por embaçamento e por erros de bit, o SSIM apresenta um desempenho muito superior. Os pontos no gráfico de dispersão se mostram mais alinhados e são obtidas maiores correlações e menores RMSE e percentis 95 e 99 do que quando são consideradas todas as distorções.

As maiores correlações alcançadas foram correspondentes ao SSIM aplicado às imagens distorcidas por embaçamento, ruído Gaussiano e erro de bit, valendo respectivamente 0,9548, 0,9568 e 0,9359. Estes três casos também produziram os menores RMSE e percentis 95 e 99. Para todos os casos testados o OR foi 0% indicando que no gráfico de dispersão não há nenhum ponto muito distante do grupo principal.

A correlação alcançada pelo SSIM aplicado em imagens distorcidas com JPEG contrabalança a média produzindo um valor mais baixo que a correlação entre o SSIM e DMOS de todas as imagens juntas. Para esta distorção a correlação do SSIM com DMOS foi de 0,8527. Os valores do RMSE e percentil 95 também foram maiores neste caso do que os valores obtidos por todas as distorções juntas.

O SSIM aplicado em imagens distorcidas pela compressão JPEG2000 apresentou correlação bem próxima e RMSE igual aos obtidos pelo SSIM aplicado em todas as imagens. O valor da correlação para esta distorção foi de 0,8673 e o valor do RMSE foi 0,0033.

Os resultados mostram que o SSIM aplicado ao canal L do espaço CIELAB é uma boa métrica de avaliação da qualidade de imagens distorcidas por erro de bit, embaçamento e ruído Gaussiano, fornecendo resultados altamente correlacionados com a métrica subjetiva.

Tabela 5– Métricas de análise por tipo de distorção

Canal L – CIELAB – Banco de dados LIVE						
Distorção	JPEG2000	JPEG	Ruído	Embaçamento	Erro bit	Todas
Correlação	0,8673	0,8527	<b>0,9359</b>	<b>0,9568</b>	<b>0,9548</b>	0,8633
RMSE	0,0033	0,0035	<b>0,0016</b>	<b>0,0010</b>	<b>0,0012</b>	0,0033
OR	<b>0%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>	<b>0%</b>	0%
Perc. 95	0,1499	0,1764	<b>0,1036</b>	<b>0,0876</b>	<b>0,1027</b>	0,1689
Perc. 99	0,2265	0,2097	<b>0,1592</b>	<b>0,1044</b>	<b>0,1355</b>	0,2272

Na seção seguinte é realizado um estudo baseado nas componentes que formam o SSIM, a fim de verificar se uma combinação diferente destas proporciona resultados melhores para algum dos canais do espaço CIELAB.

### 5.3 Combinações alternativas das componentes do SSIM

O SSIM é composto por três componentes – luminância, contraste e estrutura, como visto no Capítulo 2, cada uma correspondendo a uma característica da percepção do SVH.

Na Seção 5.1 foi verificado que o SSIM calculado nos canais de cromaticidade dos espaços de cores com esta característica não fornece bons resultados. Isto ocorre porque esses canais não carregam informação de estrutura, e o SSIM foi desenvolvido originalmente para trabalhar com o canal de luminância. Nesta seção é realizado um estudo para avaliar se somente o contraste, somente a luminância ou uma combinação diferente das componentes do SSIM aplicada aos canais de cromaticidade apresentaria melhores resultados.

Os testes foram realizados usando o banco de dados LIVE e nos canais L, a e b do espaço de cores CIELAB. A eles aplica-se o SSIM completo e suas combinações (luminância e contraste, contraste e estrutura e luminância e estrutura) a cada uma de suas componentes individuais.

Como explicado no Capítulo 2, cada uma das componentes do SSIM independentemente satisfaz às condições de simetria, limitação e máximo único, o que significa que tanto o SSIM completo como suas componentes e as combinações das componentes são limitadas a 1 e este valor é alcançado se e somente se a imagem sendo avaliada for idêntica à imagem de referência. A Figura 34 apresenta os gráficos e a Tabela 6 mostra os valores obtidos para a correlação, RMSE, OR e percentis 95 e 99 deste experimento.

A partir dos resultados observa-se que o SSIM completo produz melhores resultados do que suas componentes independentes ou suas combinações para o canal de luminância L, com correlação igual a 0,8633, RMSE igual a 0,0033, OR igual a 0%, e percentis 95 e 99 iguais a 0,1689 e 0,2272 respectivamente. Os resultados para os canais de cromaticidade a e b alcançados pelo SSIM completo também foram os melhores, com correlações iguais a 0,6482 e 0,6519 para os canais a e b respectivamente, RMSE igual a 0,0075 para ambos os canais, OR igual a 0,13% e 0,26% para os canais a e b respectivamente, percentis 95 e 99 iguais a 0,2293 e 0,2882 para o canal a e iguais a 0,2207 e 0,2929 para o canal b. Observa-se também que para os canais de cromaticidade os piores resultados foram obtidos utilizando somente a estrutura, apresentando a menor correlação e os maiores RMSE. Isto era esperado, pois corrobora com a idéia de que esses canais portam pequena quantidade de informação estrutural. Entretanto, quando a estrutura foi combinada com a luminância e com o contraste para os canais de cores, os resultados foram melhores.

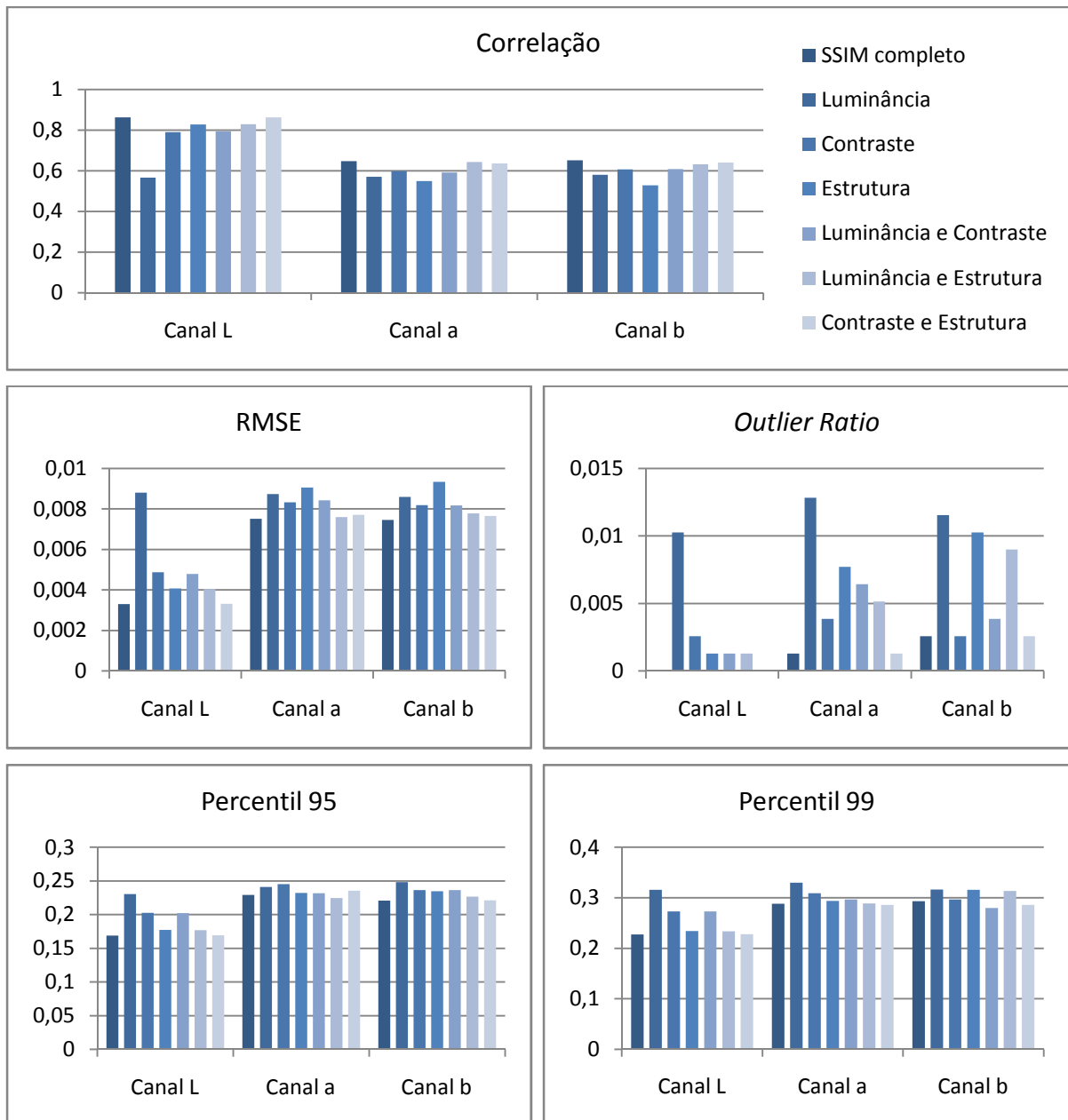


Figura 34 – Métricas de análise para a aplicação das três componentes do SSIM individualmente e combinadas duas a duas.

Visto que em geral o SSIM completo produz melhores resultados que suas componentes independentes ou outras combinações destas, o capítulo seguinte prossegue o estudo do SSIM completo aplicado aos canais dos espaços de cores. A fim de verificar se uma combinação não linear das informações dos três canais do espaço de cores produz melhores resultados do que somente um dos canais individuais, no capítulo seguinte são propostas três formas de combinações não-lineares em valores únicos – os chamados SSIM Compostos. Três diferentes fórmulas de cálculo são apresentadas e os resultados obtidos são analisados.



Tabela 6 – Valores obtidos pelas métricas de análise para componentes e combinações SSIM

Canal	SSIM completo	Luminância	Contraste	Estrutura	Luminância e Contraste	Luminância e Estrutura	Contraste e Estrutura
<b>Correlação</b>							
Canal L	<b>0,8633</b>	0,5669	0,7899	0,8279	0,7943	<b>0,8293</b>	<b>0,8628</b>
Canal a	<b>0,6482</b>	0,5708	0,5983	0,5495	0,5920	<b>0,6432</b>	<b>0,6369</b>
Canal b	<b>0,6519</b>	0,5804	0,6072	0,5288	0,6079	<b>0,6323</b>	<b>0,6402</b>
<b>RMSE</b>							
Canal L	<b>0,0033</b>	0,0088	0,0049	0,0041	0,0048	<b>0,0040</b>	<b>0,0033</b>
Canal a	<b>0,0075</b>	0,0087	0,0083	0,0091	0,0084	<b>0,0076</b>	<b>0,0077</b>
Canal b	<b>0,0075</b>	0,0086	0,0082	0,0093	0,0082	<b>0,0078</b>	<b>0,0077</b>
<b>Outlier Ratio</b>							
Canal L	<b>0,00%</b>	1,03%	0,26%	<b>0,13%</b>	<b>0,13%</b>	<b>0,13%</b>	<b>0,00%</b>
Canal a	<b>0,13%</b>	1,28%	<b>0,39%</b>	0,77%	0,64%	0,51%	<b>0,13%</b>
Canal b	<b>0,26%</b>	1,16%	<b>0,26%</b>	1,03%	0,39%	0,90%	<b>0,26%</b>
<b>Percentil 95</b>							
Canal L	<b>0,1689</b>	0,2304	0,2027	0,1771	0,2020	<b>0,1769</b>	<b>0,1692</b>
Canal a	<b>0,2293</b>	0,2408	0,2452	0,2322	<b>0,2315</b>	<b>0,2243</b>	0,2353
Canal b	<b>0,2207</b>	0,2479	0,2361	0,2346	0,2364	<b>0,2266</b>	<b>0,2213</b>
<b>Percentil 99</b>							
Canal L	<b>0,2272</b>	0,3158	0,2727	<b>0,2340</b>	0,2727	<b>0,2337</b>	0,2277
Canal a	<b>0,2882</b>	0,3299	0,3089	0,2937	0,2963	<b>0,2884</b>	<b>0,2856</b>
Canal b	<b>0,2929</b>	0,3161	0,2966	0,3158	<b>0,2794</b>	0,3136	<b>0,2858</b>

## 6 SSIM COMPOSTOS

Uma vez estudada a aplicação do SSIM em cada um dos canais individuais dos espaços de cores, o trabalho desenvolvido por Toet e Lucassen em (1) serviu como inspiração para verificarmos se melhores resultados poderiam ser obtidos através de uma combinação linear ou não-linear entre os canais dos espaços de cores, em relação à aplicação do SSIM em um dos canais individualmente.

Assim, este capítulo investiga a combinação dos canais em um único valor final SSIM para imagens coloridas – aqui referido como SSIM Composto. O propósito deste experimento é verificar se a contribuição dos três canais resulta em uma curva de ajuste monótona e de melhores resultados gerais. A Figura 35 apresenta a estrutura de cálculo dos SSIM Compostos.

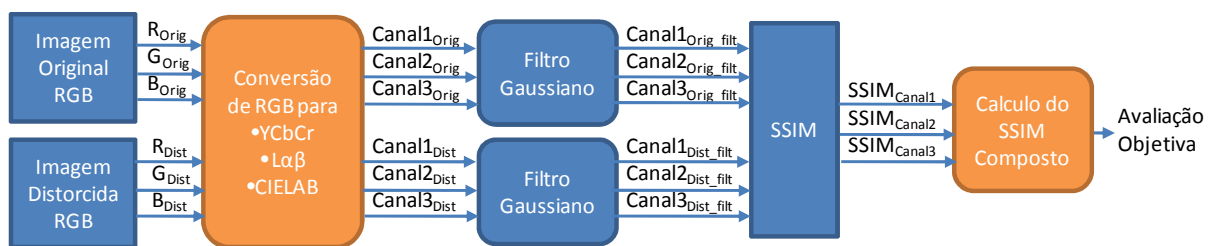


Figura 35 – Diagrama de blocos do cálculo do SSIM Composto.

O mesmo processo do cálculo do SSIM para os canais individuais dos espaços de cores é usado e então os resultados dos três canais são agrupados de forma a se obter um único valor final, o SSIM Composto.

Três propostas de SSIM Compostos foram testadas com duas variações para cada uma: sem ponderação (onde foram atribuídos pesos iguais a cada canal) e com ponderação (atribuindo-se peso 2 para o canal de luminância e peso  $\frac{1}{2}$  para cada um dos canais de cores). Referimo-nos a cada uma destas propostas de SSIM Composto como SSIMC0, SSIMC1, SSIMC2 e às suas variantes ponderadas como SSIMCP0, SSIMCP1 e SSIMCP2.

O espaço de cores RGB possui concentração de luminância em nenhum dos canais e, portanto, os SSIM Compostos ponderados não foram calculados para esse espaço.

Para o cálculo do SSIM Composto primeiramente é necessário que se compute para cada canal de cada espaço de cores o mapa SSIM (uma matriz bi-dimensional), conforme explicado na Seção 2.2.3, para cada canal de cada um dos espaços de cores. Esses mapas

oferecem uma avaliação espacial da qualidade da imagem e um valor único de avaliação da imagem completa é obtido através da média de um mapa.

Referimo-nos aos três canais de qualquer dos espaços de cores como  $I, J, K$ , e aos valores das médias do SSIM, para cada um dos canais, como  $MSSIM(I), MSSIM(J), MSSIM(K)$  e o mapa SSIM para cada canal como  $\mathbf{MAP}_I, \mathbf{MAP}_J, \mathbf{MAP}_K$ , os significados físicos e as fórmulas para os SSIM Compostos são apresentados a seguir.

## 6.1 Formulação dos SSIM Compostos

- SSIMC0 e SSIMCP0

A primeira proposta de SSIM Composto considera conjuntamente os valores dos pixels nos três canais. Esta abordagem considera o pixel como colorido e formado por suas três componentes, ao invés de considerá-las como partes separadas (ver SSIMC1). Assim, o SSIMC0 computa dos valores a partir dos mapas e não do MSSIM e é dado por:

$$SSIMC0(I, J, K) = \sqrt{\frac{1}{3M} \sum_1^M w_I \mathbf{MAP}_I^2 + w_J \mathbf{MAP}_J^2 + w_K \mathbf{MAP}_K^2}, \quad (39)$$

onde  $M$  é a quantidade de pixels dos mapas SSIM dos canais.

Para o cálculo do SSIMC0 cada um dos pesos  $w_I, w_J$  e  $w_K$  assume valor 1 e sua soma vale 3. O cálculo do SSIMCP0 é similar ao do SSIMC0, mas os pesos  $w_I, w_J$  e  $w_K$  assumem respectivamente os valores 2,  $\frac{1}{2}$  e  $\frac{1}{2}$ , com peso maior para I do que para J e K. Desta forma, o canal I correspondente à luminância nos espaços de cores YCbCr,  $L\alpha\beta$  e CIELAB que contém mais informação visual, e assim possui um peso maior que os canais de crominância, que carregam menos informação.

- SSIMC1 e SSIMCP1

A segunda possibilidade foi proposta em (1). As médias dos índices SSIM de cada um dos canais são empregadas para se calcular o valor quadrático médio ponderado entre os canais, que produz o resultado SSIMC1 total:

$$SSIMC1(I, J, K) = \sqrt{\frac{w_I MSSIM(I)^2 + w_J MSSIM(J)^2 + w_K MSSIM(K)^2}{3}}, \quad (40)$$

onde  $w_I$ ,  $w_J$  e  $w_K$  são os pesos atribuídos a cada canal.

Para o cálculo do SSIMC1 os pesos  $w_I$ ,  $w_J$  e  $w_K$  valem 1 e sua soma vale 3. O cálculo do SSIMCP1 é feito de forma similar, porém os pesos  $w_I$ ,  $w_J$  e  $w_K$  valem respectivamente 2,  $\frac{1}{2}$  e  $\frac{1}{2}$ .

- SSIMC2 e SSIMCP2

A terceira proposta é uma média comum dos mapas dos três canais. Esta média é calculada a partir de:

$$SSIMC2(I, J, K) = \frac{1}{3M} \sum_1^M w_I \mathbf{MAP}_I + w_J \mathbf{MAP}_J + w_K \mathbf{MAP}_K. \quad (41)$$

Para o cálculo do SSIMC2 cada um dos pesos  $w_I$ ,  $w_J$  e  $w_K$  assume valor 1 e sua soma vale 3. De forma similar, o cálculo do SSIMCP2 considera os pesos  $w_I$ ,  $w_J$  e  $w_K$  assumindo respectivamente os valores 2,  $\frac{1}{2}$  e  $\frac{1}{2}$ . Diferentemente das propostas anteriores, aqui é feita uma média linear, logo calcular a média utilizando o MSSIM ao invés do mapa SSIM produziria o mesmo resultado.

## 6.2 Análise dos resultados dos SSIM Compostos

Os SSIM Compostos foram aplicados aos bancos de dados LIVE, IVC e TID. As correlações, RMSE, OR e percentis 95 e 99 obtidos por espaço de cores são mostrados na Tabela 7 para os espaços de cores CIELAB e  $L\alpha\beta$  e na Tabela 8 para os espaços de cores RGB e YCbCr. As curvas de ajuste do banco de dados LIVE para os quatro espaços de cores são apresentadas na Figura 36.

De uma maneira geral, os SSIM Compostos ponderados – SSIMCP0, SSIMCP1 e SSIMCP2, apresentaram maiores correlações em comparação aos não-ponderados. Isso pode ser verificado através das curvas de ajuste – mais alinhadas com a diagonal, e pelos valores das correlações com o DMOS mais altos. Entretanto, nenhum dos SSIM Compostos obteve

correlação maior que os melhores resultados alcançados pelos canais individuais, como mostra a Figura 37 em um gráfico comparativo entre as correlações alcançadas pelos SSIM Compostos e SSIM calculados nos canais individuais.

Especificamente sobre os espaços de cores, no espaço RGB é interessante notar que o SSIMC1 e SSIMC2 apresentaram a mesma correlação para o LIVE. Também produziram curvas de ajuste praticamente idênticas, de forma a quase não ser possível distingui-las no gráfico das curvas de ajuste. Nenhuma das curvas de ajuste para o espaço RGB foi monótona.

Para o espaço YCbCr a maior correlação alcançada usando o banco de dados LIVE foi de 0,8326 para o SSIMCP2, ligeiramente menor que a alcançada pelo SSIMCP2 calculado nos espaços  $L\alpha\beta$  e CIELAB. Porém, as curvas de ajuste para o espaço YCbCr não são monótonas e assim não apresentam bom mapeamento.

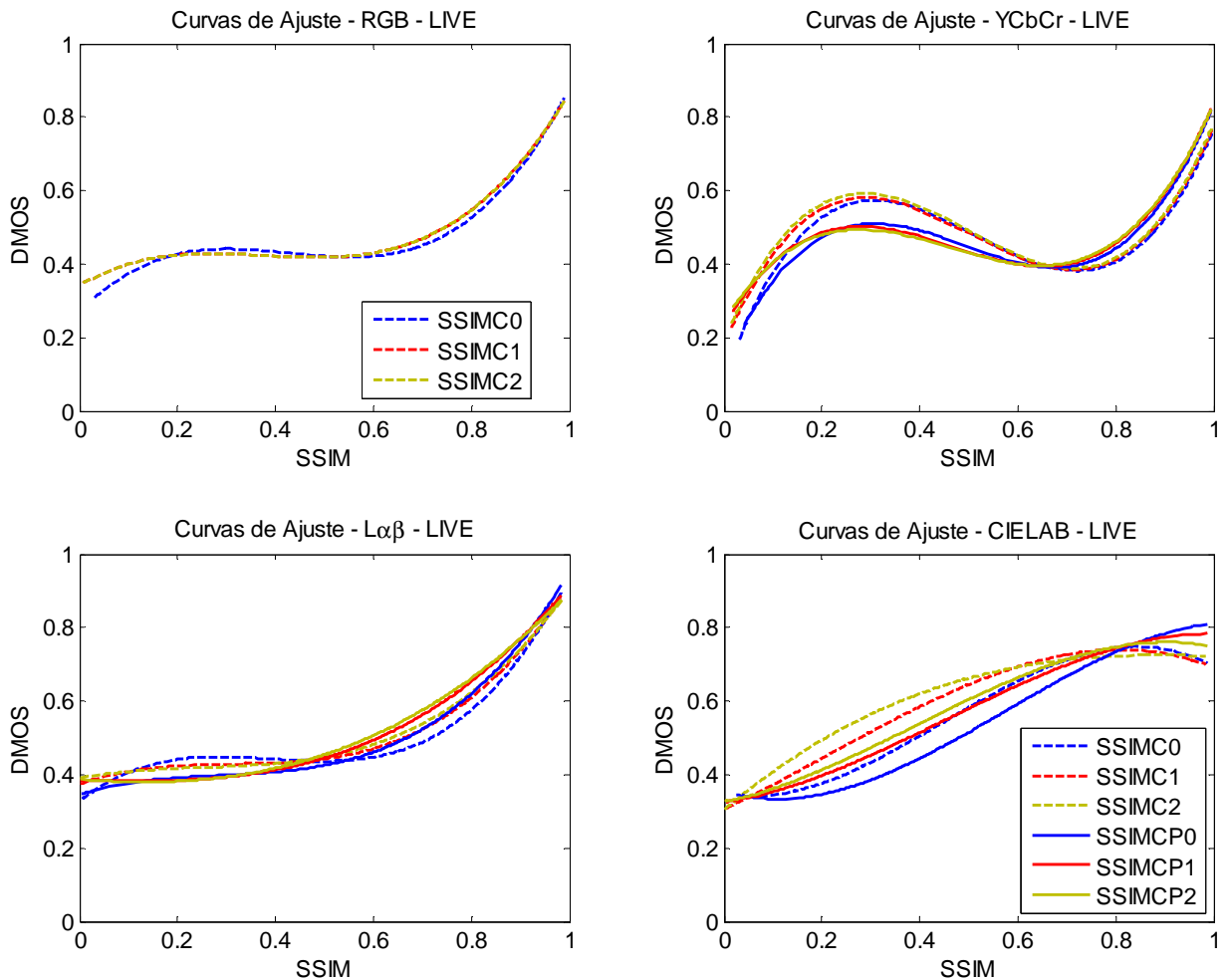


Figura 36 – Curvas de ajuste dos SSIM Compostos para cada espaço de cores usando o banco de dados LIVE.

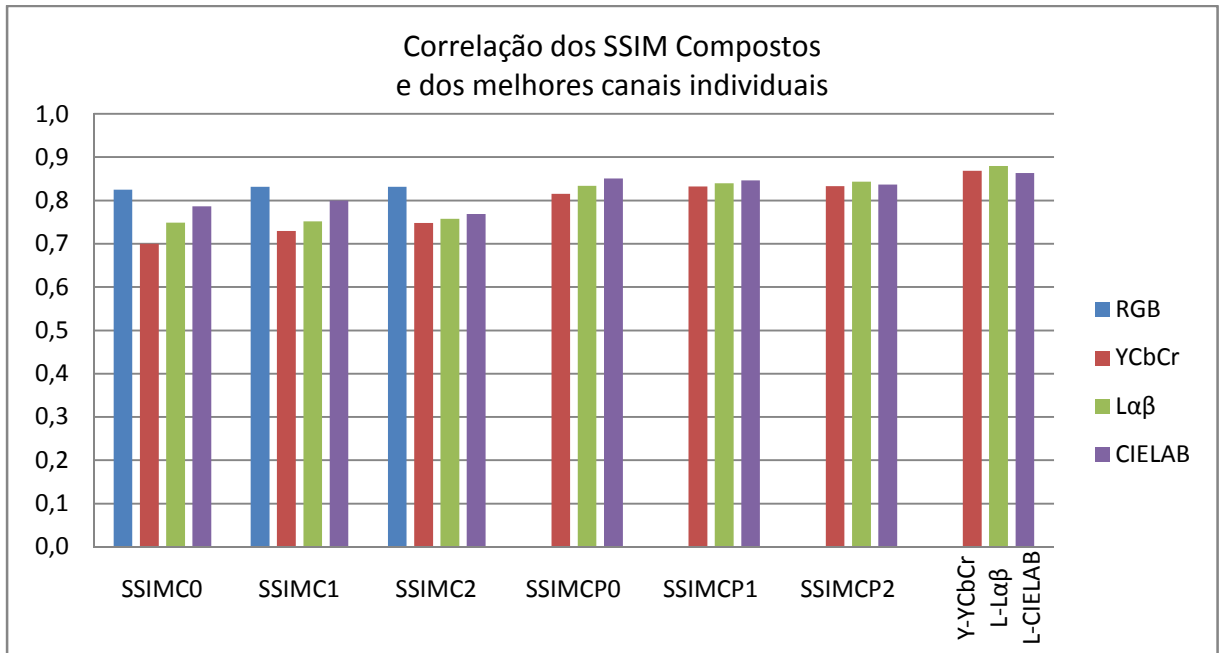


Figura 37 – Correlações entre DMOS e SSIM com ajuste no banco de dados LIVE para os cálculos dos SSIM Compostos e SSIM dos canais individuais com maior correlação.

Tabela 7– Medidas de avaliação dos SSIM Compostos nos espaços CIELAB e Laβ

Espaço	CIELAB	CIELAB	CIELAB	CIELAB	CIELAB	CIELAB	Laβ	Laβ	Laβ	Laβ	Laβ	Laβ
SSIM	SSIMC0	SSIMC1	SSIMC2	SSIMCPO	SSIMCPC1	SSIMCPC2	SSIMC0	SSIMC1	SSIMC2	SSIMCPO	SSIMCPC1	SSIMCPC2
<b>LIVE</b>												
Correlação	0,7861	0,7996	0,7683	<b>0,8505</b>	<b>0,8463</b>	0,8368	0,7485	0,7515	0,7578	0,8339	0,8395	<b>0,8432</b>
RMSE	0,0050	0,0047	0,0053	<b>0,0036</b>	<b>0,0037</b>	0,0039	0,0057	0,0056	0,0055	0,0039	0,0038	<b>0,0037</b>
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	0,26%	0,26%	0,13%	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	0,1859	0,1784	0,1901	<b>0,1692</b>	<b>0,1654</b>	0,1722	0,2176	0,2095	0,2097	<b>0,1766</b>	0,1721	0,1765
Perc. 99	<b>0,2192</b>	0,2201	0,2242	<b>0,2173</b>	0,2269	<b>0,2190</b>	0,2791	0,2698	0,2714	0,2378	0,2393	0,2401
<b>IVC</b>												
Correlação	0,7921	0,8016	0,7641	<b>0,8379</b>	<b>0,8261</b>	<b>0,8268</b>	0,6154	0,6309	0,6190	0,7536	0,7667	0,7568
RMSE	0,0180	0,0172	0,0201	<b>0,0144</b>	<b>0,0153</b>	<b>0,0153</b>	0,0299	0,0290	0,0297	0,0208	0,0199	0,0206
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	0,3471	0,3325	0,3761	<b>0,3270</b>	<b>0,3322</b>	<b>0,3220</b>	0,4757	0,4661	0,4674	0,4089	0,3848	0,3921
Perc. 99	<b>0,4384</b>	0,4659	0,4751	<b>0,4068</b>	<b>0,4448</b>	0,4620	0,5465	0,5235	0,5335	0,4651	0,4489	0,4496
<b>TID</b>												
Correlação	0,4840	0,5739	0,5312	0,5990	<b>0,6264</b>	0,6129	0,3589	0,3842	0,3912	0,5481	0,5474	0,5555
RMSE	0,0085	0,0075	0,0080	<b>0,0071</b>	<b>0,0068</b>	<b>0,0069</b>	0,0097	0,0095	0,0094	0,0078	0,0078	0,0077
OR	1,41%	<b>0,76%</b>	<b>0,82%</b>	<b>0,82%</b>	<b>0,76%</b>	<b>0,71%</b>	2,35%	2,53%	2,53%	1,82%	2,18%	1,94%
Perc. 95	0,2384	0,2267	0,2332	<b>0,2241</b>	<b>0,2169</b>	<b>0,2189</b>	0,2526	0,2516	0,2522	0,2308	0,2291	0,2298
Perc. 99	0,2969	<b>0,2835</b>	0,2859	<b>0,2841</b>	<b>0,2845</b>	0,2839	0,3823	0,3715	0,3713	0,3148	0,3268	0,3234

Tabela 8– Medidas de avaliação dos SSIM Compostos nos espaços RGB e YCbCr

Espaço	RGB	RGB	RGB	YCbCr	YCbCr	YCbCr	YCbCr	YCbCr	YCbCr
SSIM Total	SSIMC0	SSIMC1	SSIMC2	SSIMC0	SSIMC1	SSIMC2	SSIMC0	SSIMC1	SSIMC2
<b>LIVE</b>									
Correlação	0,8249	0,8318	0,8318	0,7002	0,7292	0,7482	0,8156	0,8321	0,8326
RMSE	0,0041	0,0040	0,0040	0,0066	0,0061	0,0057	0,0043	0,0040	0,0040
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	1,54%	1,16%	1,28%	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	0,1964	0,1922	0,1922	0,2546	0,2442	0,2415	0,2224	0,2103	0,2072
Perc. 99	0,2609	0,2540	0,2539	0,3258	0,3201	0,3175	0,2695	0,2598	0,2581
<b>IVC</b>									
Correlação	0,7387	0,7554	0,7540	0,6522	0,6659	0,6704	0,7543	0,7648	0,7644
RMSE	0,0219	0,0207	0,0208	0,0277	0,0268	0,0265	0,0208	0,0200	0,0200
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	1,67%	1,67%	1,67%
Perc. 95	0,4350	0,4225	0,4228	0,4333	0,4400	0,4409	0,3660	0,3631	0,3615
Perc. 99	0,5320	0,5096	0,5105	0,5103	0,5255	0,5260	0,6325	0,6314	0,6330
<b>TID</b>									
Correlação	0,4699	0,4986	0,4956	0,2996	0,3250	0,3460	0,4683	0,5076	0,5071
RMSE	0,0087	0,0084	0,0084	0,0101	0,0099	0,0098	0,0087	0,0083	0,0083
OR	2,29%	2,18%	2,18%	2,41%	2,24%	2,12%	2,53%	2,06%	2,12%
Perc. 95	0,2442	0,2404	0,2410	0,2538	0,2536	0,2553	0,2453	0,2420	0,2399
Perc. 99	0,3640	0,3513	0,3529	0,4017	0,3961	0,3894	0,3648	0,3511	0,3508

Os espaços  $L\alpha\beta$  e CIELAB foram os dois únicos espaços que apresentaram curvas de ajuste monótonas. As versões ponderadas dos SSIM Compostos produziram curvas monótonas no espaço  $L\alpha\beta$  enquanto no espaço CIELAB os SSIMC0 e SSIMC1 também produziram curvas monótonas. As correlações alcançadas também foram maiores para esses dois espaços de cores, comparadas aos outros para todos os bancos de dados. Para o banco de dados LIVE o SSIMC0 no espaço CIELAB alcançou a maior correlação igual a 0,8505 e o espaço  $L\alpha\beta$  alcançou a maior correlação para este banco de dados com o SSIMC2, valendo 0,8432.

A Figura 38 apresenta uma comparação entre as curvas de ajuste dos melhores casos do SSIM aplicado aos canais individuais, os canais L dos espaços de cores CIELAB e  $L\alpha\beta$ , e o canal Y do espaço YCbCr, e dos SSIM Compostos ponderados dos mesmos espaços de cores, utilizando o banco de dados LIVE. Na figura pode-se observar que as curvas dos canais individuais L são mais alinhadas que as curvas dos SSIM Compostos, em especial a curva do canal L do espaço CIELAB.

Além disso, como visto na Figura 37, as correlações da avaliação subjetiva com o SSIM calculado nos canais individuais L dos espaços CIELAB e  $L\alpha\beta$  e Y do espaço YCbCr foram mais altas que as correlações das propostas de SSIM Compostos. Entretanto, como visto na

Figura 38, a curva de ajuste do canal Y não é monotônica crescente, o que torna esta opção menos interessante que o canal L dos espaços CIELAB e  $L\alpha\beta$ .

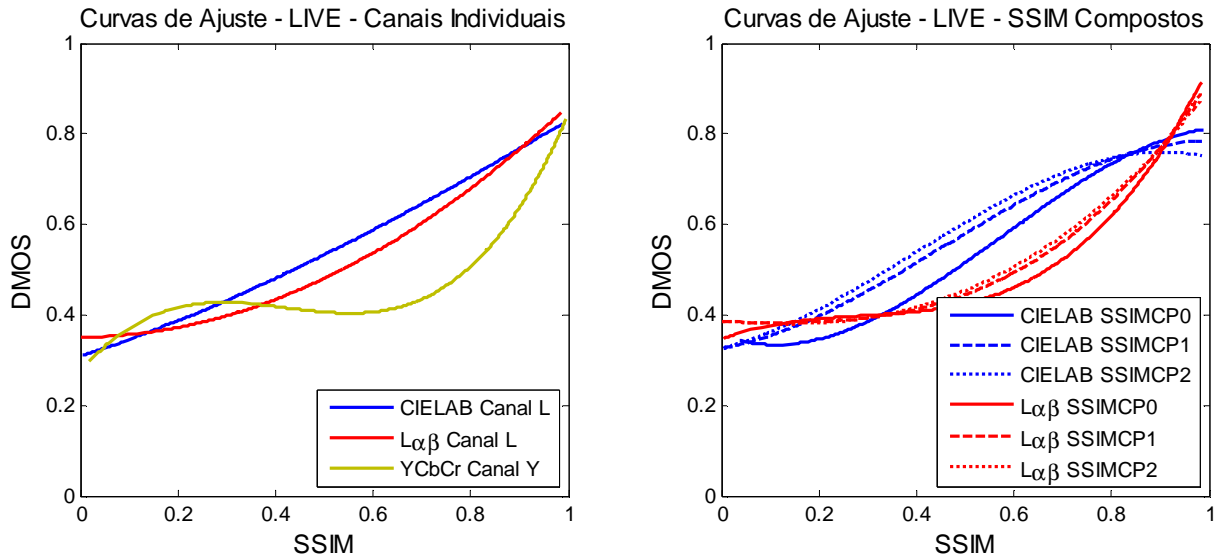


Figura 38 – Comparação entre as melhores curvas de ajuste utilizando os canais individuais e SSIM Compostos

No capítulo seguinte, o espaço de cores CIELAB continua a ser explorado. A característica deste espaço de permitir medir a distância entre cores é investigada, de modo a verificar se a utilização dessa distância permite obter uma melhora do casamento entre a métrica objetiva e a subjetiva.



## 7 APLICAÇÃO DA DISTÂNCIA ENTRE CORES NO CIELAB AO SSIM

No capítulo anterior foi verificado que a idéia de unir o resultado do SSIM calculado nos três canais dos espaços de cores não produziu melhorias frente ao cálculo do SSIM em um dos canais individualmente. Nos Capítulos 5 e 6 verificou-se que a aplicação do SSIM no canal L do espaço CIELAB produz resultados mais bem correlacionados com a avaliação subjetiva do que quando aplicado aos outros canais (a e b) deste espaço ou também quando aplicado aos outros espaços de cores testados. Assim uma nova idéia, de empregar a distância entre cores  $\Delta E$ , discutida na Seção 3.7, conjuntamente ao cálculo do SSIM, surgiu como possibilidade de melhorar o resultado do SSIM.

A proposta consiste em identificar por meio do cálculo da distância entre cores regiões entre duas imagens sendo comparadas nas quais as cores sejam imperceptivelmente diferentes. Em seguida, esta informação é utilizada no cálculo do SSIM.

No cálculo da distância entre cores, a *Mínima Diferença Perceptível* entre duas cores, ou JND, é o limiar a partir do qual duas cores passam a ser percebidas como cores distintas pelo olho humano. Não há um consenso sobre qual o valor JND preciso. Referências às faixas de valores que o JND pode assumir são encontradas em (47), onde são apresentados dois cenários: um genérico e outro voltado para aplicações industriais. As faixas de valores JND para cenários genéricos são definidas como seguem:

- i) entre 0,38 a 0,73 unidades CIELAB é dita como mínima diferença perceptível ou JND;
- ii) em torno de 1,75 unidades CIELAB é dita como uma diferença supra JND;
- iii) acima de 5,0 unidades CIELAB são ditas como grandes diferenças de cor.

Porém, a partir de uma perspectiva industrial (por exemplo, para a reprodução da cor em produtos manufaturados), indica-se que para diferenças de cor:

- i) entre 1,1 e 2,8 unidades CIELAB tem-se uma tolerância de cor rigorosa;
- ii) entre 2,8 e 5,6 unidades CIELAB diz-se de uma tolerância normal e;
- iii) acima de 5,6 unidades CIELAB encontra-se uma tolerância folgada.

Desta forma, para os testes que se seguem são investigados valores para o JND variando entre 0 – o mesmo que não aplicar a distância entre cores – e 7, com passo de 0,2. (Foram simulados um total de 36 diferentes valores de JND).

A proposta é verificar se a aplicação da distância entre cores é capaz de produzir melhores resultados na aplicação do SSIM na avaliação de imagens coloridas do que os resultados obtidos para  $JND = 0$ .

Duas abordagens de utilização da distância entre cores conjuntamente com o SSIM foram propostas: Na primeira, o MSSIM é calculado como a média dos SSIMs dos pixels para os quais  $\Delta E$  é maior que JND. Na segunda, os pixels da imagem distorcida cujo  $\Delta E$  é menor que JND são substituídos pela imagem de referência e então o MSSIM é calculado como a média dos SSIMs de todos os pixels.

Além disto, a idéia da dispersão de ponto, descrita na Seção 2.2.4 e implementada pela aplicação do filtro suavizante, foi testada de três formas: na primeira, o filtro é aplicado somente ao mapa SSIM; na segunda, o filtro é aplicado ao mapa SSIM e ao mapa  $\Delta E$ ; e na terceira, não há aplicação do filtro. Os testes utilizaram os três bancos de dados.

Nas seções seguintes os testes realizados em cada abordagem são explicados detalhadamente e os resultados obtidos são apresentados. Na última seção é realizada uma comparação entre todos os resultados e analisa-se qual proposta produziu resultados mais consistentes.

Em todos os resultados a curva de ajuste foi levada em consideração. Esta foi calculada e as métricas de avaliação (correlação, RMSE, OR e percentis 95 e 99) foram calculadas entre os dados subjetivos preditos (a curva de ajuste) e os dados subjetivos.

### 7.1 SSIM como média dos pixels para os quais $\Delta E > JND$

Nesta abordagem, primeiramente é calculado o mapa  $\Delta E$ . Os pixels neste mapa identificados como menores que o limiar JND são desconsiderados no cálculo do MSSIM. Assim, retêm-se para a avaliação da qualidade somente os pixels cuja distorção é perceptível à visão humana. O mapa SSIM é calculado para todos os pixels, mas somente os pixels para os quais o  $\Delta E$  é maior que JND são considerados no cálculo do MSSIM, conforme a seguinte equação:

$$\text{seja } I = \{i \mid \Delta E_i > JND\} \text{ então } MSSIM = \sum_{i \in I} \frac{SSIM_i}{\#I}, \quad (42)$$

onde  $\#I$  é a quantidade de pixels para os quais  $\Delta E_i > JND$ .

### 7.1.1 Filtragem SSIM

Neste teste o filtro Gaussiano é aplicado durante o cálculo do mapa SSIM, mas não no cálculo do mapa  $\Delta E$ . A Figura 39, já apresentada na Seção 3.7, ilustra a filtragem do SSIM e a utilização do  $\Delta E$  no processo de avaliação objetiva da qualidade de imagens coloridas.

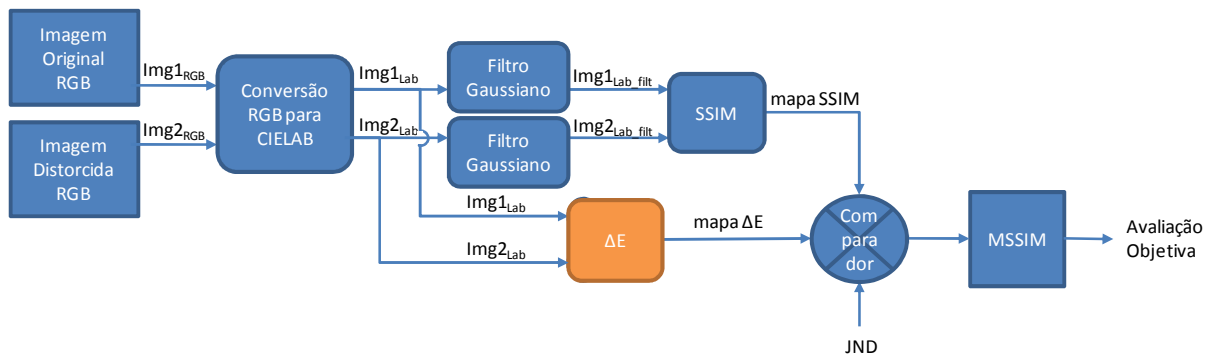


Figura 39 – Diagrama de blocos da utilização do  $\Delta E$  conjuntamente ao SSIM filtrado para a primeira abordagem, cálculo do MSSIM como média dos pixels onde  $\Delta E > JND$ .

As imagens original e distorcida em RGB são convertidas para o espaço de cores CIELAB. A partir das imagens convertidas é calculado o mapa  $\Delta E$ , a distância entre as cores para cada pixel. As imagens convertidas também seguem por outro caminho, onde passam por um filtro Gaussiano e em seguida é calculado o SSIM entre as imagens produzindo o mapa SSIM. Neste ponto, as informações dos dois mapas, SSIM e  $\Delta E$  e o valor JND são agrupadas: o mapa  $\Delta E$  é utilizado junto com a informação do JND para identificar os pixels do mapa SSIM cujo  $\Delta E$  é maior que JND. Os pixels que satisfazem esta condição são usados no cálculo do MSSIM, isto é, calcula-se a média dos SSIMs dos pixels cujo  $\Delta E$  é maior que o valor JND, produzindo o resultado da avaliação objetiva.

Neste experimento, além do cálculo do SSIM no canal L do espaço CIELAB, também é testado o desempenho dos SSIM Compostos a fim de verificar se a utilização da distância entre as cores nos SSIM Compostos revela uma melhora com relação à utilização do canal L do CIELAB.

A Figura 40 apresenta as curvas de correlação entre os SSIM mencionados acima e o DMOS para a faixa de valores JND entre 0 e 7 usando o banco de dados LIVE, e a Figura 41 apresenta os gráficos contendo as curvas para o RMSE, OR e percentis 95 e 99.

As curvas de correlação do SSIM para o canal L e dos SSIM Compostos com o DMOS revelam que o SSIM no canal L obteve maior correlação que todos os SSIM Compostos, atingindo o pico de 0,8809 para  $JND = 2,6$ . Para este JND, o RMSE, OR e percentil 95 do

SSIML também mostram valores bem pequenos, mas não os mínimos, o que é coerente com um bom mapeamento. Já o percentil 99 neste JND apresenta valor mais alto que o de alguns SSIM Compostos.

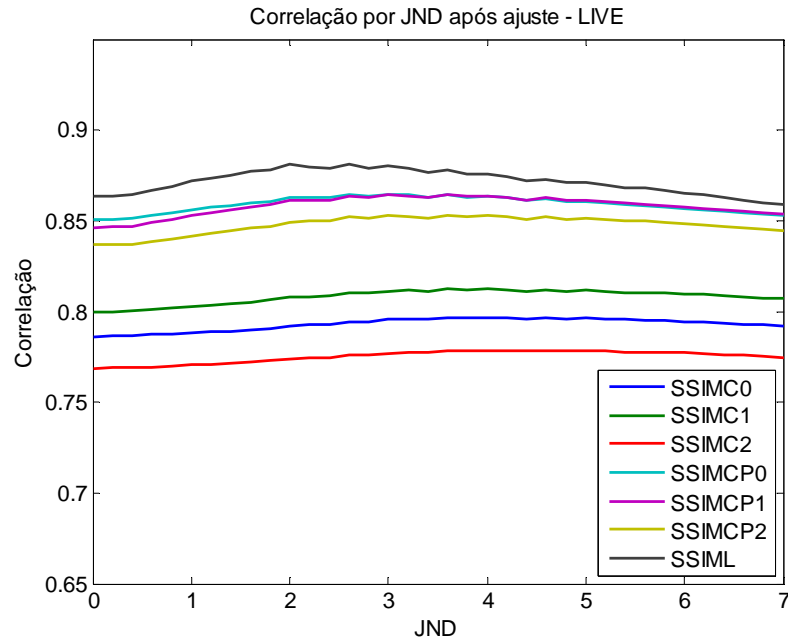


Figura 40 – Curvas de correlação entre SSIM e DMOS por JND, no banco de dados LIVE.

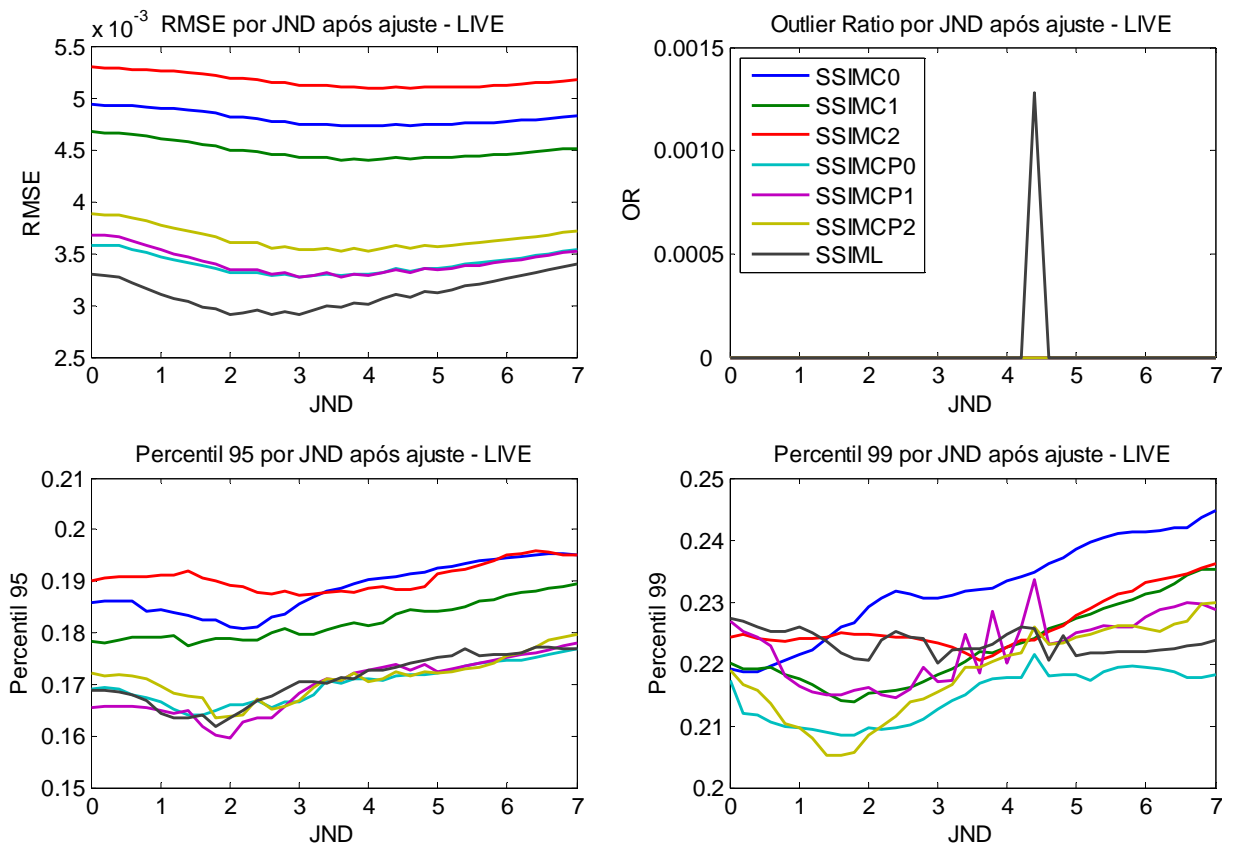


Figura 41 – RMSE, OR e Percentis dos SSIM Compostos e SSIML após ajuste.

Na Tabela 9 são apresentados os valores das maiores correlações obtidas para cada SSIM Composto e SSIM no canal L, com o JND para o qual foi alcançado. Em todos os casos, a melhor correlação foi alcançada para um JND diferente de zero, isto é, para todos os casos a utilização do JND melhora a correlação do SSIM com a métrica subjetiva.

Tabela 9 – Melhores correlações com JND correspondente

SSIM	SSIMC0	SSIMC1	SSIMC2	SSIMCP0	SSIMCP1	SSIMCP2	SSIML
<b>LIVE</b>							
JND	4,0	4,0	4,6	3,0	3,0	3,6	2,6
Correlação	0,7968	0,8124	0,7787	<b>0,8646</b>	<b>0,8645</b>	0,8532	<b>0,8809</b>
RMSE	0,0047	0,0044	0,0051	<b>0,0033</b>	<b>0,0033</b>	0,0035	<b>0,0029</b>
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	0,1904	0,1815	0,1883	<b>0,1666</b>	<b>0,1684</b>	0,1708	<b>0,1677</b>
Perc. 99	0,2334	0,2226	0,2252	<b>0,2127</b>	<b>0,2170</b>	<b>0,2194</b>	0,2243
<b>IVC</b>							
JND	5,8	6,4	5,8	3,2	3,4	3,8	3,0
Correlação	0,8590	0,8821	0,8294	<b>0,8868</b>	<b>0,8903</b>	<b>0,8899</b>	0,8801
RMSE	0,0126	0,0107	0,0150	<b>0,0103</b>	<b>0,0100</b>	<b>0,0100</b>	0,0109
OR	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>	<b>0,00%</b>
Perc. 95	0,3560	0,3268	0,3679	<b>0,2950</b>	<b>0,2718</b>	<b>0,2907</b>	0,3080
Perc. 99	0,4350	<b>0,3848</b>	0,4590	<b>0,3746</b>	<b>0,4035</b>	0,4221	0,4561
<b>TID</b>							
JND	4,4	4,0	1,4	4,4	4,0	4,2	2,0
Correlação	0,5482	0,5916	0,5718	<b>0,6433</b>	<b>0,6466</b>	0,6354	<b>0,6578</b>
RMSE	0,0078	0,0072	0,0075	<b>0,0065</b>	<b>0,0065</b>	0,0066	<b>0,0063</b>
OR	1,71%	<b>1,65%</b>	<b>1,53%</b>	<b>1,65%</b>	<b>1,65%</b>	1,71%	1,71%
Perc. 95	0,2279	0,2196	0,2260	<b>0,2191</b>	<b>0,2184</b>	<b>0,2184</b>	0,2248
Perc. 99	0,3504	<b>0,3213</b>	<b>0,3034</b>	<b>0,3261</b>	0,3277	0,3265	0,3298

Neste teste, utilizando o banco de dados LIVE, nenhuma proposta de SSIM Composto produziu correlação com a medida subjetiva mais alta que o SSIM calculado no canal L. Assim, para os demais testes neste capítulo os cálculos foram feitos somente com o canal L do espaço CIELAB.

#### 7.1.2 Filtragem SSIM e $\Delta E$

Para este teste um filtro Gaussiano com as mesmas características do filtro aplicado ao SSIM para o cálculo do mapa SSIM é aplicado ao  $\Delta E$  antes do cálculo do mapa  $\Delta E$ . Este

experimento é motivado pela característica de dispersão de ponto do SVH, como apresentado na Seção 2.1.4.

Na Figura 42 é mostrado o diagrama esquemático deste experimento. Antes do cálculo do mapa  $\Delta E$  as imagens original e distorcida são filtradas. Em seguida, os mapas SSIM e  $\Delta E$ , juntamente com a informação JND são unidos de forma que o MSSIM é calculado como a média dos SSIM dos pixels cujo  $\Delta E$  é maior que JND.

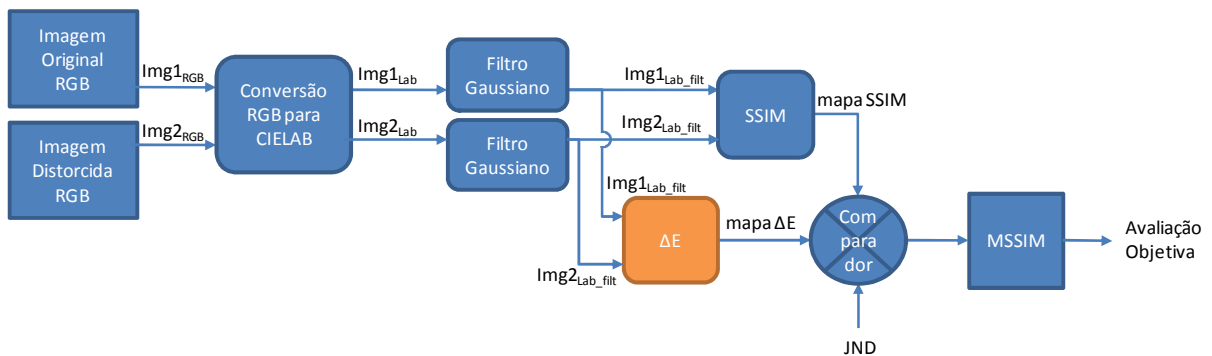


Figura 42 – Diagrama de blocos da utilização do  $\Delta E$  filtrado conjuntamente ao SSIM filtrado para a primeira abordagem.

Foi verificado nos testes com o banco de dados IVC que para  $JND > 2$  pelo menos uma imagem com distorção não apresentou nenhum pixel cujo  $\Delta E$  fosse maior que JND. Concluimos que o problema seria o filtro, que suavizou demais a imagem, uma vez que o mesmo procedimento aplicado à imagem sem a aplicação do filtro não demonstrou problemas. Para contornar o problema, nos casos em que não houve qualquer  $\Delta E > JND$  foi feito  $SSIM=1$ , já que na teoria não existe diferença entre a imagem sendo avaliada e a imagem original.

As curvas de correlação e os resultados das métricas de análise, RMSE, OR e percentis 95 e 99, para o JND que proporcionou a maior correlação são apresentados respectivamente na Figura 44 e Tabela 10.

### 7.1.3 Sem filtragem

Para este teste o filtro Gaussiano foi removido de modo a verificar se sua utilização produz de fato melhores resultados. A Figura 43 mostra um diagrama de blocos desta implementação sem a utilização da filtragem das imagens antes do cálculo do SSIM.

As imagens RGB de referência e distorcida são convertidas para CIELAB. As imagens convertidas são então usadas diretamente no cálculo dos mapas  $\Delta E$  e SSIM. Os mapas são

encaminhados a um comparador que identifica os pixels maiores que JND no mapa SSIM, para se calcular o MSSIM produzindo assim a avaliação objetiva.

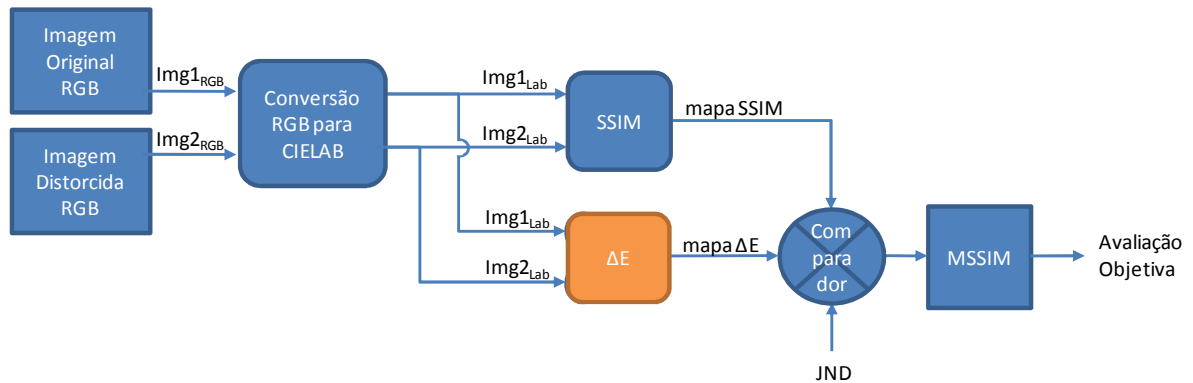


Figura 43 – Diagrama de blocos da implementação do SSIM sem filtragem para a primeira abordagem.

#### 7.1.4 Resultados da primeira abordagem

A correlação entre SSIM e DMOS/MOS para os três bancos de dados nos três testes da primeira abordagem de aplicação da distância de cores juntamente ao SSIM são mostrados na Figura 44. As curvas para as métricas RMSE, OR e percentis 95 e 99 são mostradas nas Figuras 45, 46 e 47, para os bancos de dados LIVE, IVC e TID respectivamente. Os valores das melhores correlações, os JND nos quais ocorreram e sua comparação com o valor da correlação para  $JND = 0$ , são mostrados na Tabela 10.

De uma maneira geral, pode-se observar que os resultados das correlações obtidas para cada um dos testes variaram dependendo do banco de dados avaliado.

Para o banco de dados LIVE é possível verificar que os dois testes com o uso do filtro produziram curvas de correlação bastante parecidas. O teste sem a utilização do filtro em comparação com os demais mostrou uma correlação um pouco menor e aproximadamente constante para toda a faixa de valores JND. Para esse banco de dados é possível notar que houve uma melhora da correlação, ainda que sutil, para um JND diferente de 0.

Para o banco de dados IVC percebe-se que a curva de correlação entre SSIM e MOS correspondente ao teste com filtragem nos mapas  $\Delta E$  e SSIM apresentou a maior correlação. A curva seguiu o comportamento esperado, ascendendo com valores crescentes de JND para em seguida decrescer. Porém, o pico da curva não ocorreu na região esperada, em JND entre 2 e 3 como observado nas as curvas dos outros dois testes, mas sim em  $JND = 5,4$ . Para este

banco de dados as curvas de correlação referente aos testes com filtragem somente no SSIM e sem filtragem ficaram bem semelhantes.

Para o banco de dados TID as curvas de correlação dos testes apresentaram comportamentos bastante distintos. A curva do SSIM filtrado apresentou a mais alta correlação dentre os testes realizados, porém a utilização do  $\Delta E$  não provocou melhora significativa na correlação. O comportamento esperado é uma curva convexa, com o pico dentro da faixa de valores JND proposta, seguida de uma queda, entretanto a curva de correlação correspondente ao teste do SSIM e  $\Delta E$  filtrados apresentou um comportamento indesejado, pois teve seu pico em certo valor JND, decresceu e em um determinado ponto voltou a subir. A curva de correlação correspondente ao teste sem filtragem apresentou valores bem mais baixos que as demais.

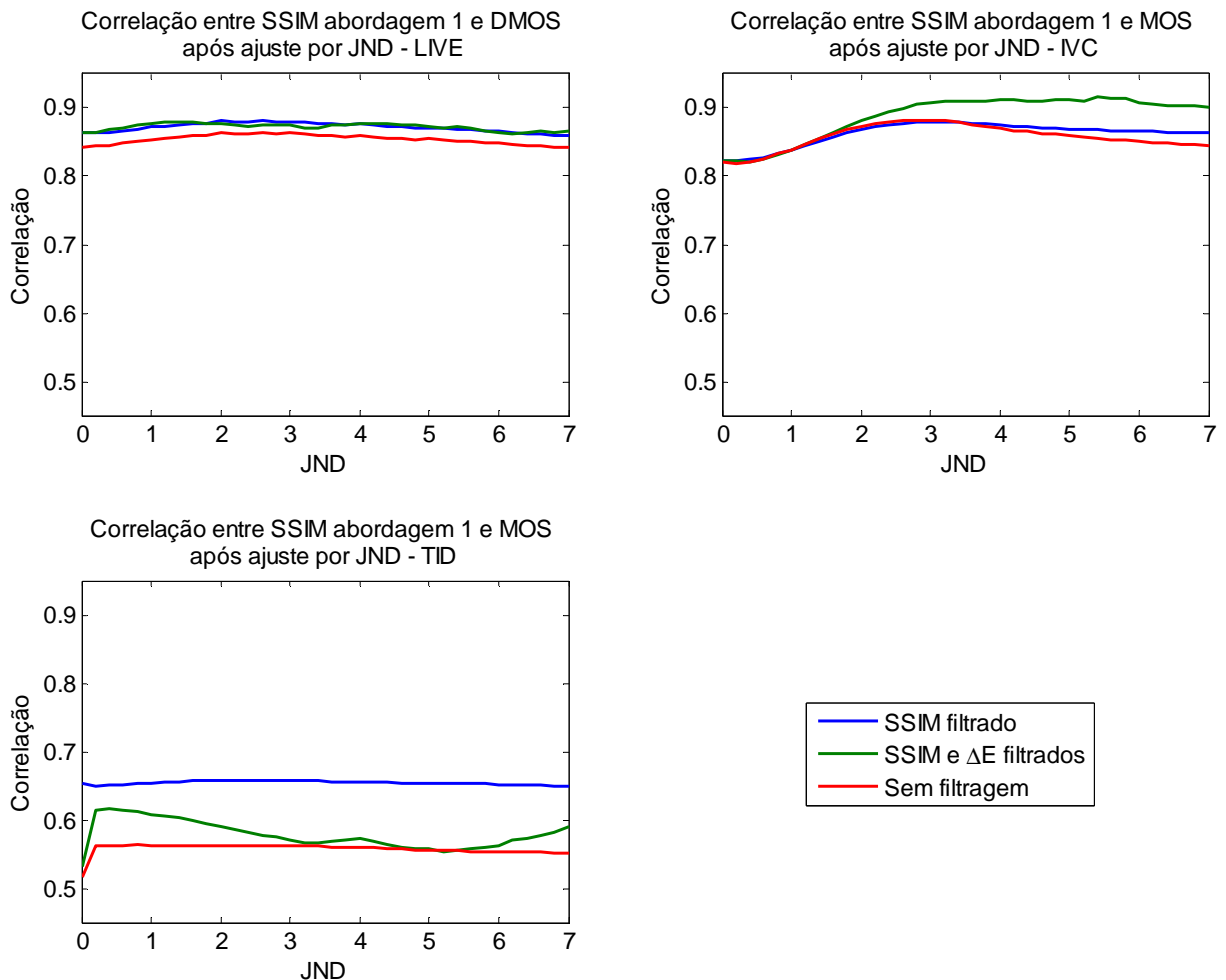


Figura 44 – Curvas de correlação dos testes da aplicação da filtragem no mapa SSIM e no mapa  $\Delta E$  para a primeira abordagem da aplicação do  $\Delta E$  conjuntamente ao SSIM.



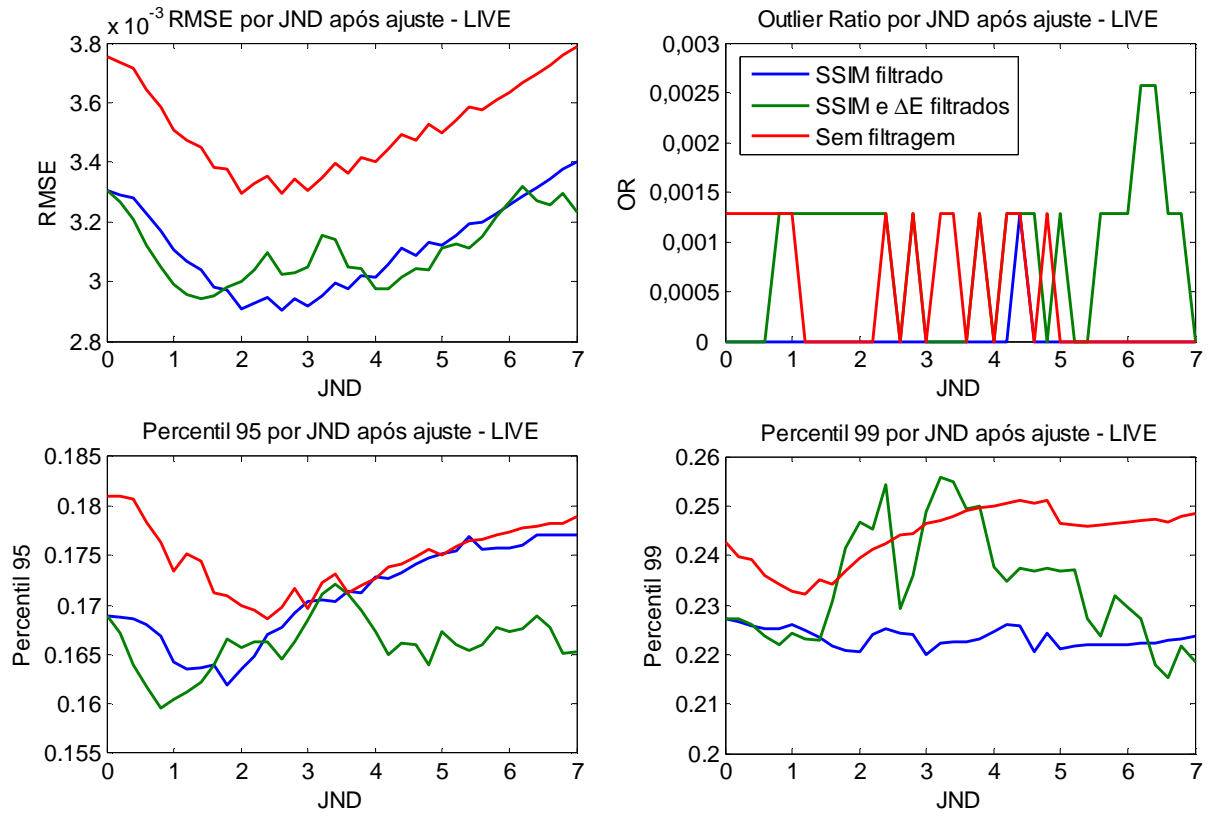


Figura 45 – Métricas de avaliação da primeira abordagem para o banco de dados LIVE.

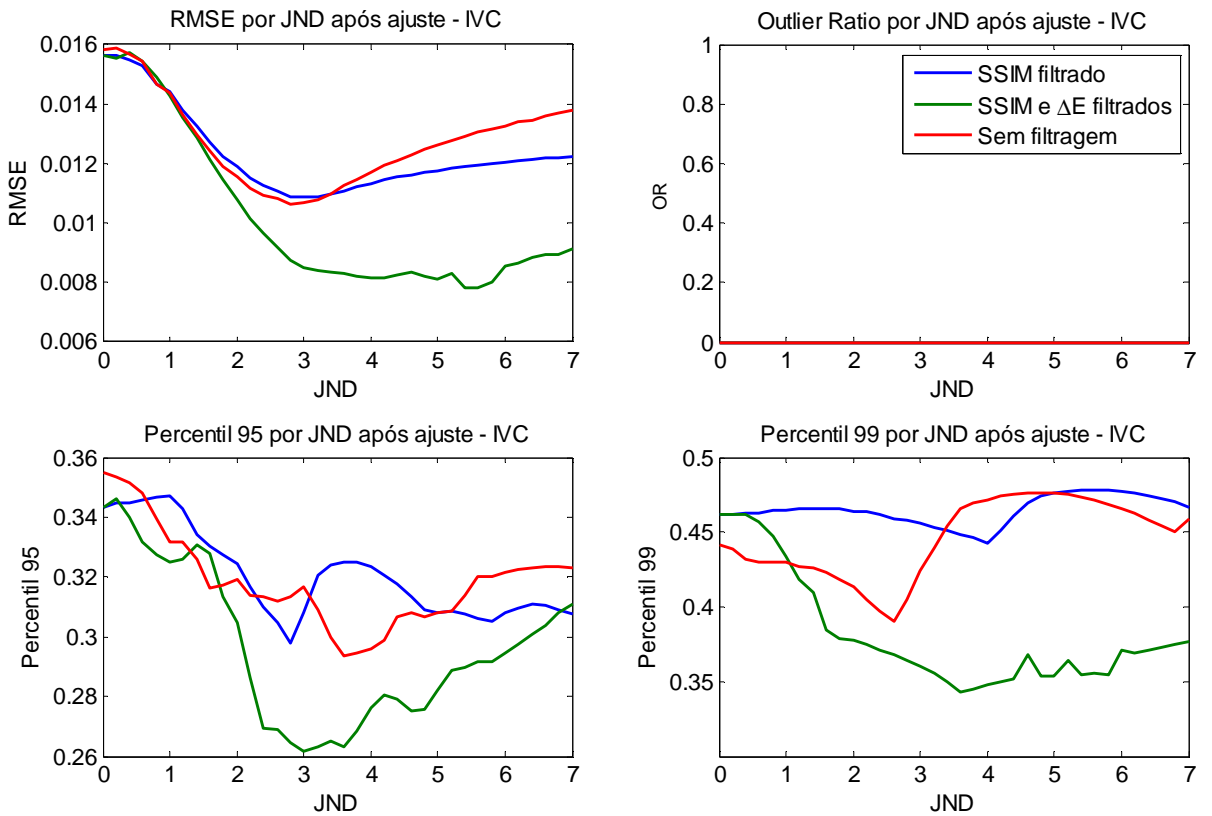


Figura 46 – Métricas de avaliação da primeira abordagem para o banco de dados IVC.

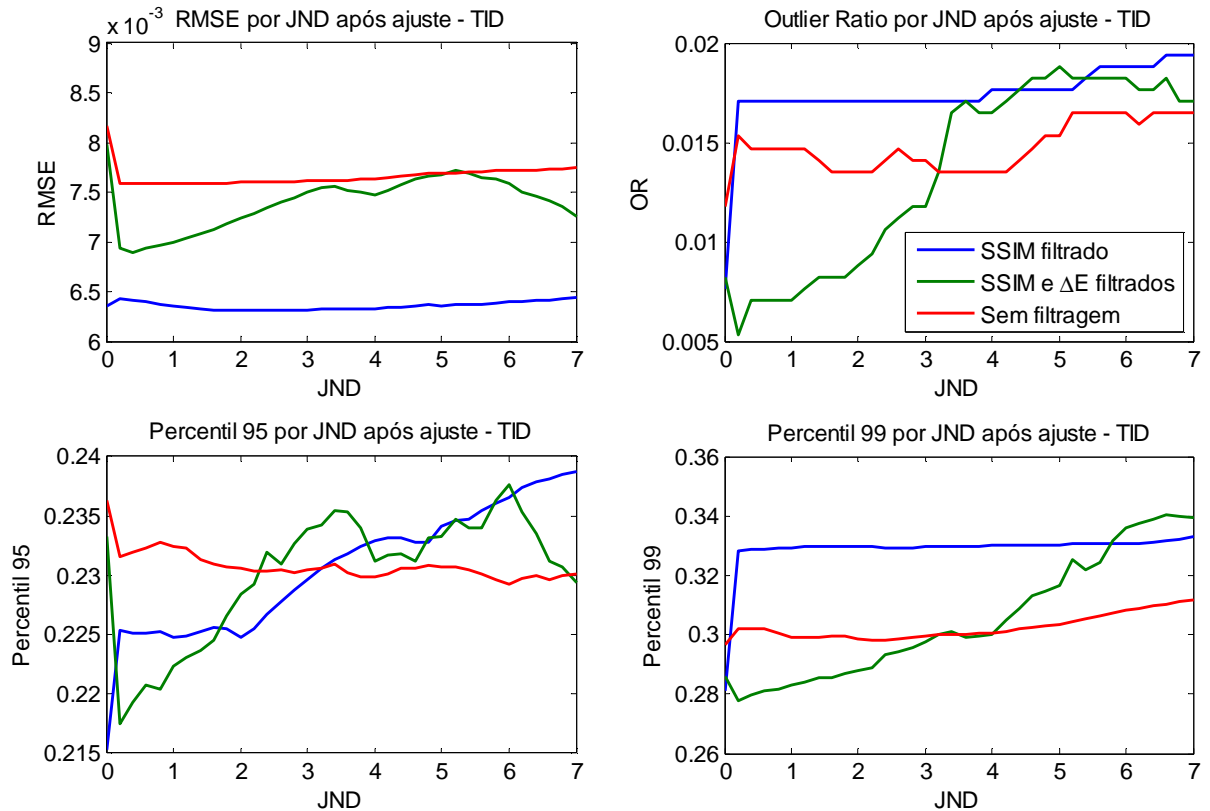


Figura 47 – Métricas de avaliação da primeira abordagem para o banco de dados TID.

A maior correlação alcançada para o banco de dados LIVE foi de 0,8809 em JND = 2,6 para o teste do SSIM filtrado, o que representa uma melhora de 2,04% em relação a não-utilização do  $\Delta E$ . O RMSE e OR para esse valor de JND também foram os menores. O teste do SSIM e  $\Delta E$  filtrados produziu a segunda melhor correlação, valendo 0,8792 em JND = 1,4, o que representou uma melhora de 1,85% em relação ao valor da curva em JND = 0. Também aqui os valores RMSE e OR para esse valor JND foram os menores.

Para o banco de dados IVC a maior correlação foi alcançada para o teste com filtragem do  $\Delta E$  e do SSIM. O valor da correlação foi de 0,9154 para JND = 5,4, um valor de JND alto fora da faixa esperada, representando uma melhora de 11,34% em relação à correlação em JND = 0. As curvas de correlação dos testes SSIM filtrado e sem filtragem foram semelhantes, com picos valendo respectivamente 0,8801 em JND = 3 e 0,8825 em JND = 2,8.

Para o banco de dados TID a maior correlação alcançada foi de 0,6578 pelo SSIM filtrado para JND = 2, uma melhora muito pouco expressiva de 0,59% com relação a não utilização da distância entre cores. Para os três testes as curvas do RMSE, OR e Percentil 95 corroboram o pico apresentado nas curvas de correlação, apresentando os menores valores nos mesmos JNDs.

Dos resultados desta abordagem, pode-se concluir que a utilização da filtragem, simulando o comportamento passa-baixa presente no olho humano, de fato produz melhores resultados. Entretanto, em uma primeira vista, não é possível definir qual dos dois testes empregando a filtragem é mais vantajoso, pois os resultados variaram dependendo do banco de dados utilizado.

Também se conclui que a aplicação da distância entre cores é vantajosa, pois melhora a correlação entre o SSIM e a métrica objetiva, sendo mais óbvia a percepção dessa melhora em alguns casos, como por exemplo, nos experimentos com o banco de dados IVC.

Tabela 10 – Comparação entre os resultados obtidos pelo JND que apresentou maior correlação em cada experimento e os resultados sem utilização do  $\Delta E$  para a primeira abordagem

Abordagem 1 - MSSIM como média dos pixels onde $\Delta E > JND$									
Teste	SSIM filtrado			SSIM e $\Delta E$ filtrados			Sem filtragem		
SSIM	SSIML			SSIML			SSIML		
<b>LIVE</b>									
JND	0,0	2,6	Melhoria usando $\Delta E$	0,0	1,4	Melhoria usando $\Delta E$	0,0	2,6	Melhoria usando $\Delta E$
Correlação	0,8633	0,8809	2,04%	0,8633	0,8792	1,85%	0,8429	0,8637	2,47%
RMSE	0,0033	0,0029	12,06%	0,0033	0,0029	10,93%	0,0038	0,0033	12,26%
OR	0,00%	0,00%	0,00%	0,00%	0,13%	-0,13%	0,13%	0,00%	0,13%
Perc. 95	0,1689	0,1677	0,72%	0,1689	0,1622	3,97%	0,1811	0,1698	6,19%
Perc. 99	0,2272	0,2243	1,28%	0,2272	0,2229	1,89%	0,2426	0,2441	-0,59%
<b>IVC</b>									
JND	0,0	3,0	Melhoria usando $\Delta E$	0,0	5,4	Melhoria usando $\Delta E$	0,0	2,8	Melhoria usando $\Delta E$
Correlação	0,8222	0,8801	7,05%	0,8222	0,9154	11,34%	0,8197	0,8825	7,66%
RMSE	0,0156	0,0109	30,44%	0,0156	0,0078	49,99%	0,0158	0,0107	32,58%
OR	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Perc. 95	0,3432	0,3080	10,23%	0,3432	0,2897	15,59%	0,3549	0,3135	11,66%
Perc. 99	0,4617	0,4561	1,21%	0,4617	0,3549	23,13%	0,4416	0,4050	8,27%
<b>TID</b>									
JND	0,0	2,0	Melhoria usando $\Delta E$	0,0	0,4	Melhoria usando $\Delta E$	0,0	0,8	Melhoria usando $\Delta E$
Correlação	0,6540	0,6578	0,59%	0,5312	0,6163	16,02%	0,5153	0,5632	9,29%
RMSE	0,0064	0,0063	0,88%	0,0080	0,0069	13,61%	0,0082	0,0076	7,03%
OR	0,76%	1,71%	-0,94%	0,82%	0,71%	0,12%	1,18%	1,47%	-0,29%
Perc. 95	0,2153	0,2248	-4,43%	0,2332	0,2193	5,97%	0,2363	0,2328	1,48%
Perc. 99	0,2810	0,3298	-17,36%	0,2859	0,2798	2,13%	0,2968	0,3007	-1,31%

## 7.2 SSIM como média de todos os pixels substituindo $\Delta E < JND$ pela referência

Nesta abordagem de aplicação do  $\Delta E$  juntamente ao SSIM propõe-se que os pixels da imagem distorcida para os quais  $\Delta E < JND$  sejam substituídos pelos da imagem de referência. Desta forma assume-se que os pixels da imagem distorcida para os quais  $\Delta E < JND$  não possuem distorção.

De forma semelhante a na abordagem anterior, são realizados três experimentos com as filtragens dos mapas  $\Delta E$  e SSIM nesta abordagem, sempre utilizando o canal L do espaço CIELAB.

### 7.2.1 Filtragem SSIM

Neste experimento faz-se a filtragem das imagens original e distorcida modificadas antes do cálculo do mapa SSIM. A Figura 48 ilustra a implementação desta abordagem.

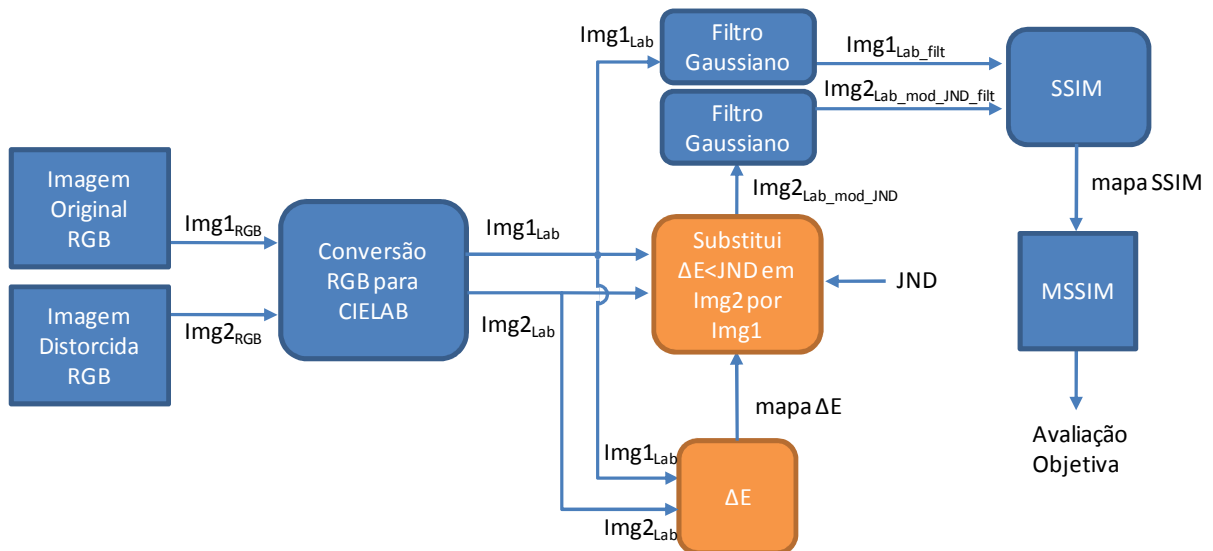


Figura 48 – Diagrama de blocos da utilização do  $\Delta E$  conjuntamente ao SSIM filtrado para a segunda abordagem, com substituição das regiões da imagem distorcida menores que  $\Delta E$  pelas equivalentes da imagem de referência.

Primeiro as imagens original e distorcida são convertidas para o espaço de cores CIELAB. Ambas as imagens são utilizadas para calcular o mapa  $\Delta E$ . A imagem distorcida convertida é modificada tendo algumas regiões substituídas pelas regiões correspondentes da imagem original convertida. Essas regiões são definidas pelos pontos onde o mapa  $\Delta E$  é menor que o JND, conforme a seguinte equação:

$$\text{se } \Delta E_i < JND \text{ então } \text{Img2}_{\text{Lab\_mod\_JND}_i} = \text{Img1}_{\text{Lab}_i}. \quad (43)$$

O resultado da modificação na imagem distorcida é, em seguida, encaminhado ao filtro Gaussiano juntamente com a imagem original convertida. Neste ponto, o mapa SSIM é calculado e uma média do mapa SSIM completo é feita resultando na métrica de avaliação objetiva.

A curva de correlação dos testes para cada banco de dados é apresentada na Figura 51 e as curvas das métricas RMSE, OR e percentis 95 e 99 são mostrados, por banco de dados, para o LIVE na Figura 52, para o IVC na Figura 53 e para o TID na Figura 54.

Os resultados das métricas de análise para o JND que proporcionou a maior correlação de todos os testes são apresentados na Tabela 11.

### 7.2.2 Filtragem SSIM e $\Delta E$

De forma semelhante a da abordagem anterior, um filtro Gaussiano com as mesmas características do filtro aplicado ao SSIM para o cálculo do mapa SSIM é aplicado ao  $\Delta E$  antes do cálculo do mapa  $\Delta E$ . A Figura 49 mostra o diagrama esquemático deste experimento.

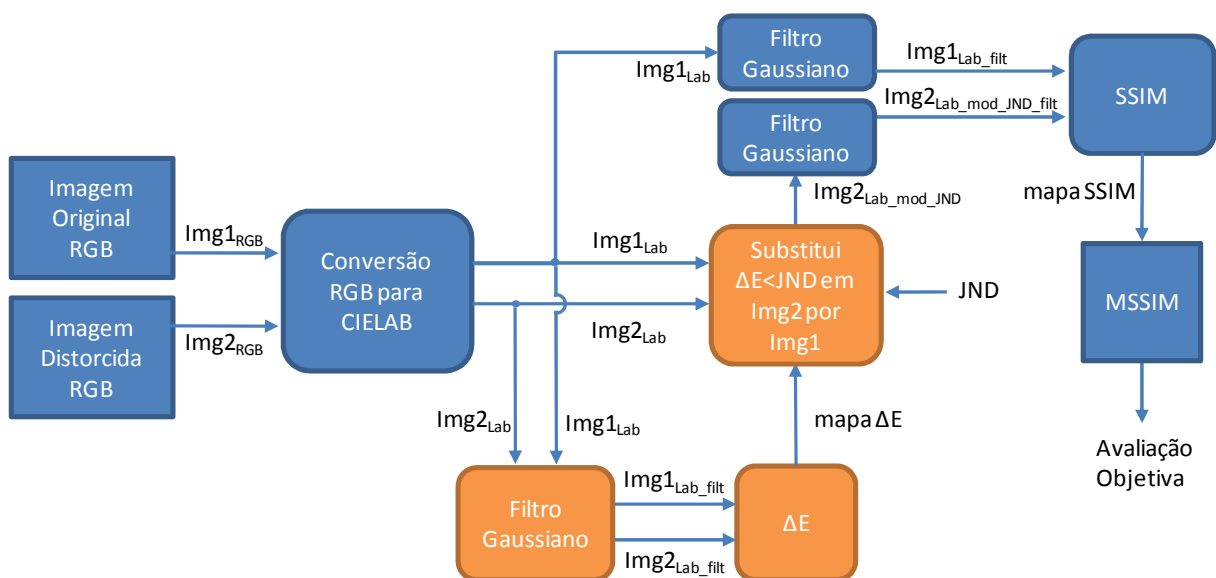


Figura 49 – Diagrama de blocos da utilização do  $\Delta E$  filtrado conjuntamente ao SSIM filtrado para a segunda abordagem.

Após convertidas, as imagens no espaço CIELAB são filtradas para ser calculado o mapa  $\Delta E$ . As imagens convertidas também são dirigidas a um comparador juntamente com o mapa  $\Delta E$  e o valor JND. Este comparador modifica a imagem distorcida substituindo nela os pixels

para os quais o mapa  $\Delta E$  apresenta valores menores que JND pela imagem de referência. A imagem modificada que sai do comparador é então filtrada juntamente com a imagem de referência e é calculado o mapa SSIM entre ambas. O MSSIM é então calculado como a média do mapa SSIM completo para produzir a avaliação objetiva.

### 7.2.3 Sem filtragem

Para este teste, a segunda abordagem da aplicação conjunta do  $\Delta E$  com o SSIM é feita sem a aplicação do filtro Gaussiano no mapa  $\Delta E$  ou no mapa SSIM. A Figura 50 mostra o diagrama esquemático deste experimento.

A partir das imagens convertidas é calculado o mapa  $\Delta E$ . As imagens convertidas também vão para um comparador que, juntamente com o mapa  $\Delta E$  e o valor JND, substitui na imagem distorcida os pixels para os quais o mapa  $\Delta E$  é menor que JND pela imagem de referência. O resultado do comparador juntamente com a imagem original convertida para o espaço CIELAB são utilizados no cálculo do mapa SSIM. O MSSIM é calculado como a média do mapa SSIM completo produzindo a avaliação objetiva.

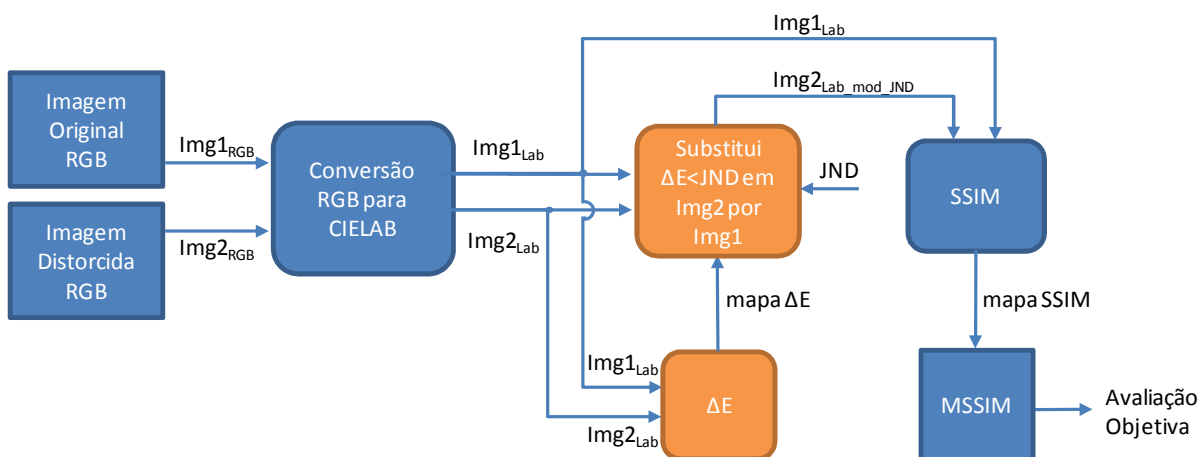


Figura 50 – Diagrama de blocos da implementação do SSIM sem filtragem para a segunda abordagem.

### 7.2.4 Resultados da segunda abordagem

Para esta abordagem do cálculo do SSIM, aplicação do  $\Delta E$  para a substituição dos pixels pela imagem de referência, nota-se que as curvas de correlação, apresentadas na Figura 51, para os testes do SSIM filtrado e sem filtragem são mais bem comportadas quando comparadas às curvas da primeira abordagem. Para os três bancos de dados as curvas seguem o formato esperado, com o valor de correlação atingindo o pico para valores de JND maiores

que zero para em seguida decrescer. Para os três bancos de dados o valor de pico da correlação ocorreu para JND entre 1,2 e 2,8.

Para o teste com filtragem do SSIM e  $\Delta E$ , para os três bancos de dados, a curva de correlação apresentou um comportamento bem diferente das curvas dos outros testes, pois o valor de pico ocorreu em  $JND = 0$ , para em seguida decrescer. Este mesmo comportamento foi verificado nos três bancos de dados, o que leva a concluir que para a segunda abordagem, no teste com filtragem do SSIM e  $\Delta E$  a utilização da distância entre cores não proporciona melhorias com relação à sua não utilização.

Para o banco de dados LIVE o teste com SSIM filtrado apresentou a correlação mais alta com pico de 0,8684 para  $JND = 1,2$ . A melhora com relação ao  $JND = 0$ , porém, foi marginal, de 0,6%. A correlação de pico obtida pelo teste sem filtragem foi de 0,8612 em  $JND = 2$ , representando uma melhoria de 2,17%. Os valores de RMSE e OR, em ambos os testes, para o JND de maior correlação foram os menores, corroborando o resultado.

Para o banco de dados IVC, de forma distinta ao que foi observado nos outros testes, a maior correlação foi alcançada pelo teste sem filtragem. O valor de correlação de pico foi de 0,8892 para  $JND = 2,8$ , representando melhoria de 8,48% com relação à correlação em  $JND = 0$ . Entretanto, o valor da correlação de pico atingido pelo teste com filtragem do SSIM foi bem semelhante ao teste sem filtragem, valendo 0,8815 para o mesmo  $JND = 2,8$  e representando melhoria de 7,22% com relação ao  $JND = 0$ . Os valores de RMSE e OR para ambos os testes foram os menores para esse JND, e os percentis apresentam os seus mínimos nas proximidades desse valor JND.

Para o banco de dados TID, nesta abordagem, a utilização da distância de cores praticamente não produziu melhorias. A correlação de pico para o teste com SSIM filtrado foi de 0,6546 em  $JND = 1,4$  e para o teste sem filtragem de 0,6554 em  $JND = 2,2$ . As melhorias relativamente às correlações para  $JND = 0$  foram respectivamente de 0,09% e 0,23%. Porém, as métricas RMSE, OR e percentis diminuíram com a utilização do  $\Delta E$  nos testes sem filtragem e com filtragem do SSIM. O RMSE atingiu o mínimo em  $JND = 1,4$  para o teste SSIM filtrado e  $JND = 2,2$  para o teste sem filtragem, o OR apresentou os menores valores para JND entre 2,6 e 2,8 para o teste com SSIM filtrado e entre 2,4 e 3,2 para o teste sem filtragem. O percentil 95 só apresentou melhora com o uso do  $\Delta E$  para o teste sem filtragem enquanto o percentil 99 para o teste com SSIM filtrado teve seu mínimo em  $JND = 1,2$  e para o teste sem filtragem o mínimo ficou em  $JND = 3,4$ .

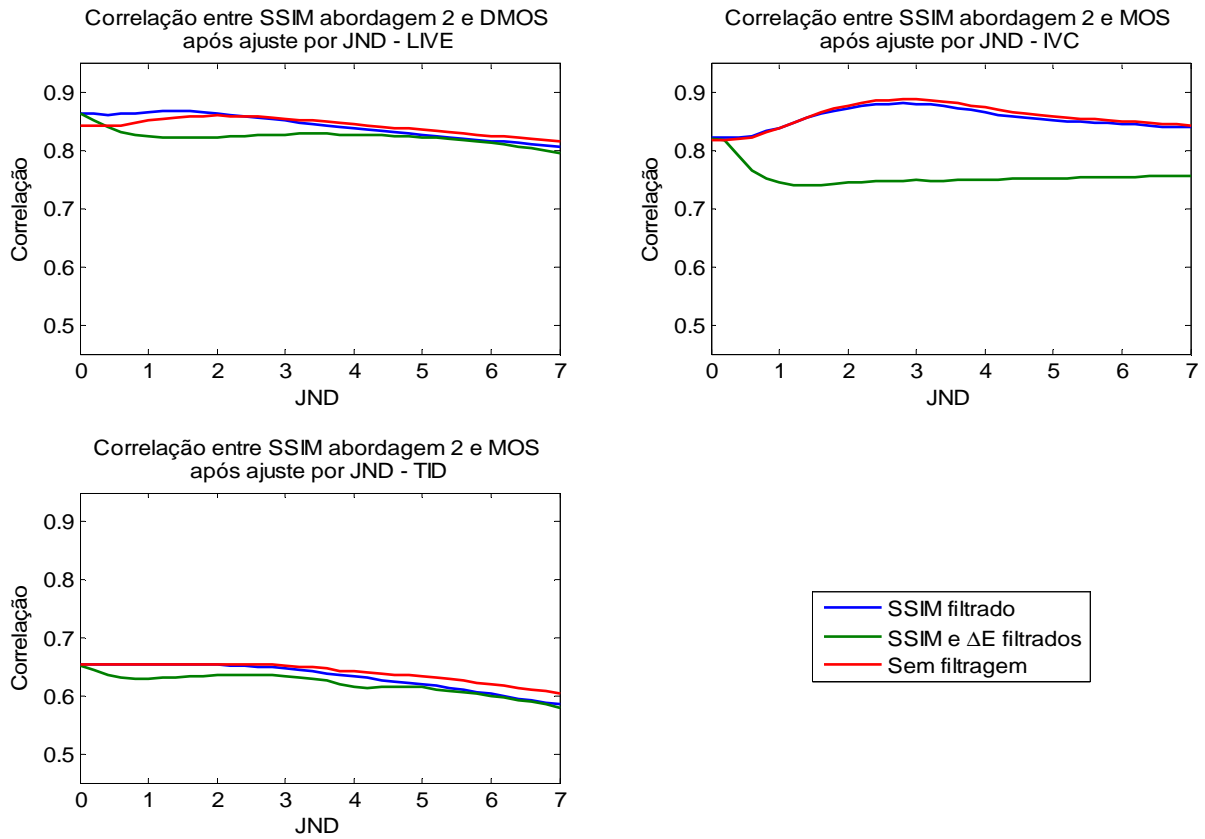


Figura 51 – Curvas de correlação dos testes da aplicação da filtragem no mapa SSIM e no mapa  $\Delta E$  para a segunda abordagem da aplicação do  $\Delta E$  conjuntamente ao SSIM.

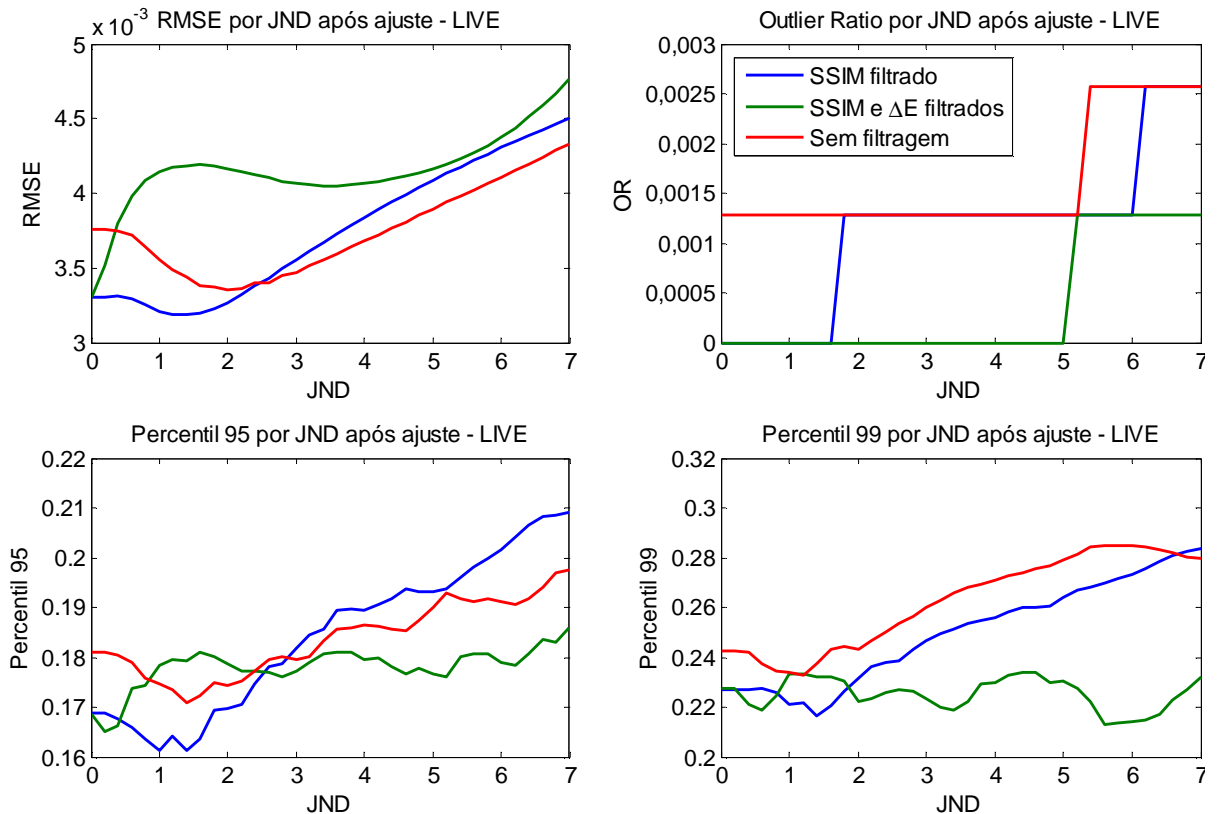


Figura 52 – Métricas de avaliação da segunda abordagem para o banco de dados LIVE.



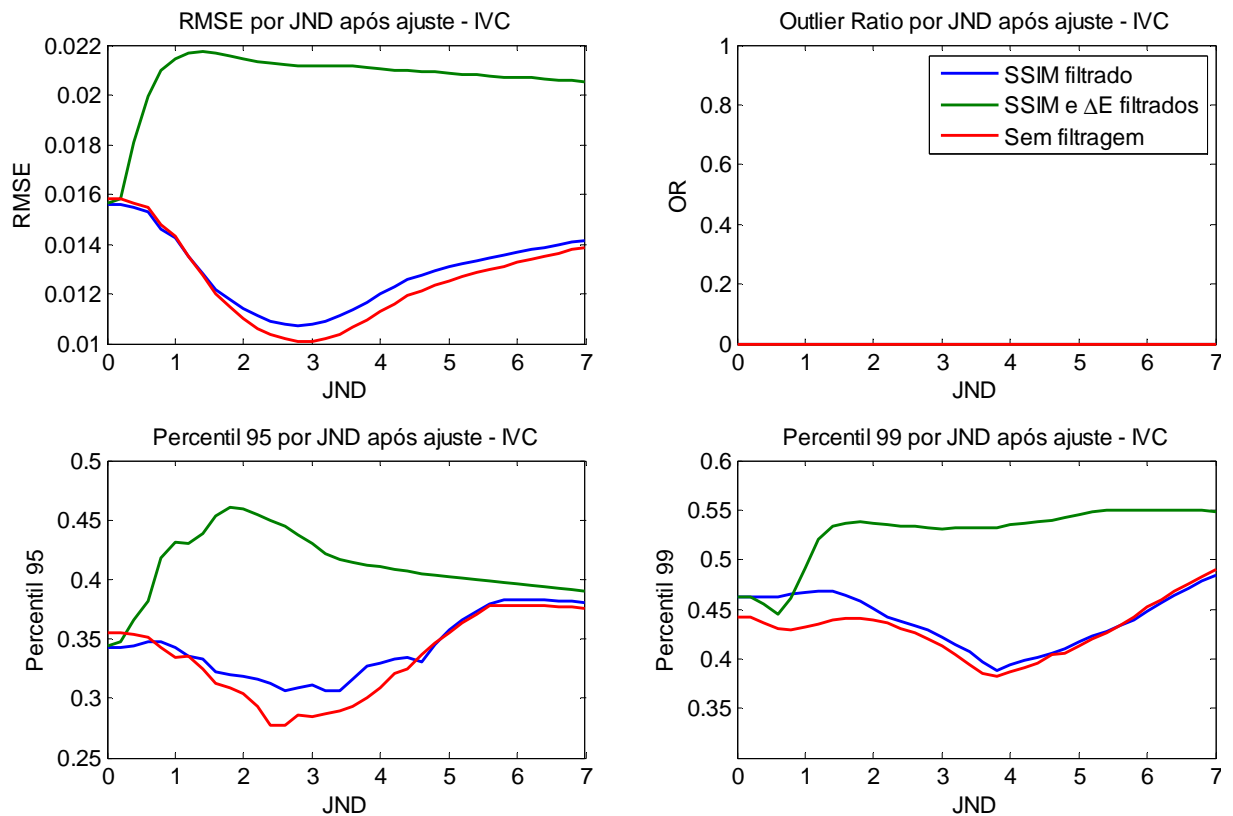


Figura 53 – Métricas de avaliação da segunda abordagem para o banco de dados IVC.

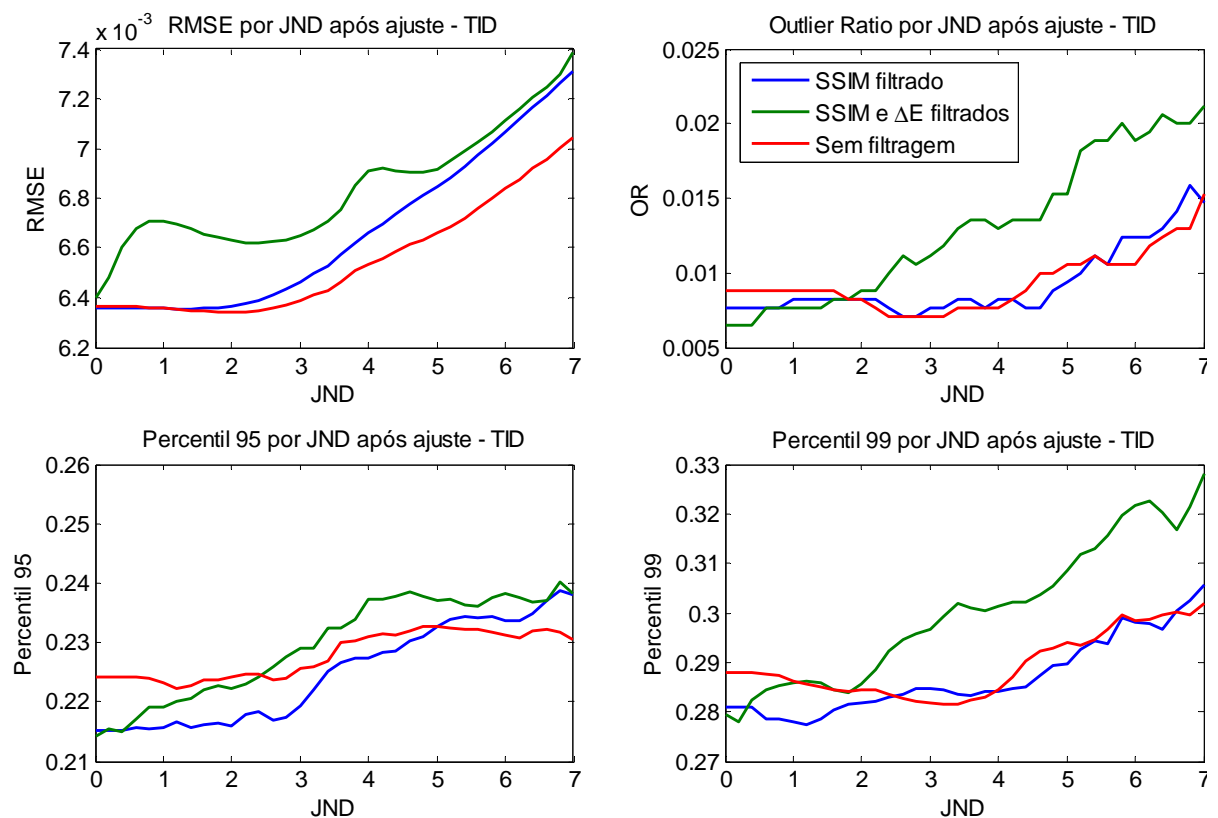


Figura 54 – Métricas de avaliação da segunda abordagem para o banco de dados TID.

Para esta abordagem, pode-se concluir que o valor JND ótimo está entre 1,2 e 2,8, intervalo no qual foram verificadas as maiores correlações dos testes SSIM filtrado e sem filtragem, corroborado pelos resultados mínimos dos RMSE, OR e percentis. Este resultado é coerente com a faixa de valores citada no início do capítulo.

Entretanto, a utilização da distância entre cores com esta abordagem provou ser realmente vantajosa somente para o banco de dados IVC, produzindo melhoria na correlação de 8,48% no teste sem filtragem e de 7,22% no teste com filtragem do SSIM. Para os bancos de dados LIVE e TID a melhoria produzida foi marginal, de no máximo 2,17% no caso do banco de dados LIVE para o teste sem filtragem e de no mínimo 0,23% para o banco de dados TID no teste sem filtragem. Nessa abordagem, a utilização da filtragem no  $\Delta E$  e no SSIM não produziu melhorias na correlação em relação ao  $JND = 0$ .

Tabela 11 – Comparação entre os resultados obtidos pelo JND que apresentou maior correlação em cada experimento e os resultados sem utilização do  $\Delta E$  com a segunda abordagem

Abordagem 2 - Substituição da imagem distorcida onde $\Delta E < JND$ pela imagem de referência									
Teste	SSIM filtrado			SSIM e $\Delta E$ filtrados			Sem filtragem		
SSIM	SSIML			SSIML			SSIML		
<b>LIVE</b>									
JND	0,0	1,2	Melhoria usando $\Delta E$	0,0	0,0	Melhoria usando $\Delta E$	0,0	2,0	Melhoria usando $\Delta E$
Correlação	0,8633	0,8684	0,60%	0,8634	0,8634	0,00%	0,8429	0,8612	2,17%
RMSE	0,0033	0,0032	3,52%	0,0033	0,0033	0,00%	0,0038	0,0033	10,77%
OR	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,13%	0,13%	0,00%
Perc. 95	0,1689	0,1642	2,81%	0,1685	0,1685	0,00%	0,1811	0,1744	3,69%
Perc. 99	0,2272	0,2220	2,29%	0,2276	0,2276	0,00%	0,2426	0,2433	-0,29%
<b>IVC</b>									
JND	0,0	2,8	Melhoria usando $\Delta E$	0,0	0,0	Melhoria usando $\Delta E$	0,0	2,8	Melhoria usando $\Delta E$
Correlação	0,8222	0,8815	7,22%	0,8221	0,8221	0,00%	0,8197	0,8892	8,48%
RMSE	0,0156	0,0107	31,21%	0,0156	0,0156	0,00%	0,0158	0,0101	36,22%
OR	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%	0,00%
Perc. 95	0,3432	0,3090	9,95%	0,3437	0,3437	0,00%	0,3549	0,2859	19,45%
Perc. 99	0,4617	0,4290	7,07%	0,4617	0,4617	0,00%	0,4416	0,4201	4,86%
<b>TID</b>									
JND	0,0	1,4	Melhoria usando $\Delta E$	0,0	0,0	Melhoria usando $\Delta E$	0,0	2,2	Melhoria usando $\Delta E$
Correlação	0,6540	0,6546	0,09%	0,6513	0,6513	0,00%	0,6539	0,6554	0,23%
RMSE	0,0064	0,0064	0,14%	0,0064	0,0064	0,00%	0,0064	0,0063	0,35%
OR	0,76%	0,82%	-0,06%	0,65%	0,65%	0,00%	0,88%	0,76%	0,12%
Perc. 95	0,2153	0,2158	-0,24%	0,2144	0,2144	0,00%	0,2242	0,2246	-0,20%
Perc. 99	0,2810	0,2787	0,84%	0,2795	0,2795	0,00%	0,2878	0,2844	1,21%

### 7.3 Comparação entre os resultados das abordagens

A Figura 55 apresenta uma comparação das duas abordagens de aplicação do  $\Delta E$  conjuntamente ao SSIM para os três bancos de dados. Na figura são mostrados três gráficos, um para cada banco de dados, com as correlações máximas alcançadas por cada um dos três testes realizados com cada abordagem, além da correlação alcançada sem a utilização do  $\Delta E$ , isto é, para  $JND = 0$ .

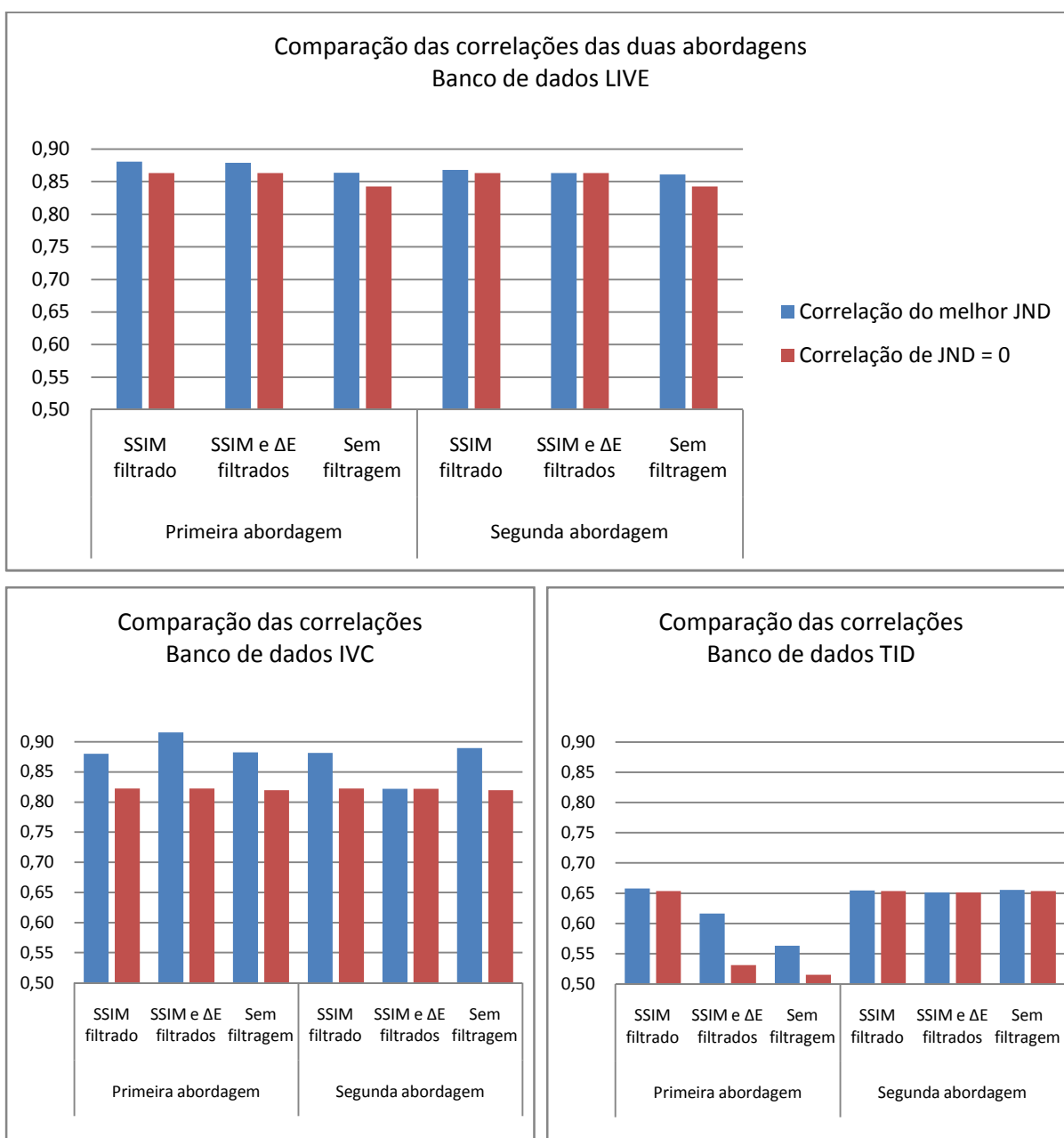


Figura 55 – Comparação entre os valores de pico das correlações e correlações em  $JND = 0$  para todos os testes das duas abordagens de aplicação do  $\Delta E$  conjuntamente ao SSIM.

Na primeira abordagem, calculando o MSSIM somente para os pixels para os quais o mapa  $\Delta E$  é maior que JND, o teste que apresentou a maior correlação para os bancos de dados LIVE e TID foi o do SSIM filtrado. No caso do banco de dados LIVE a diferença entre o resultado desse teste e os demais foi muito pequena. Para o banco de dados IVC, para a mesma abordagem, o teste que proporcionou a maior correlação foi com SSIM e  $\Delta E$  filtrados, apresentando uma diferença razoável para o resultado dos outros testes.

Na segunda abordagem, usando o banco de dados LIVE o teste que produziu maiores correlações foi o do SSIM filtrado. Já para os bancos de dados IVC e TID o teste sem filtragem proporcionou as maiores correlações, porém para o banco de dados TID a diferença para os demais testes foi quase inexistente.

A Figura 56 mostra a melhoria percentual ou ganho, isto é, a diferença entre correlação de pico e correlação para  $JND = 0$ , obtida entre o SSIM e a avaliação subjetiva para cada experimento para cada banco de dados. Os dados das correlações foram os melhores casos de cada experimento, isto é, o valor da correlação no pico da curva.

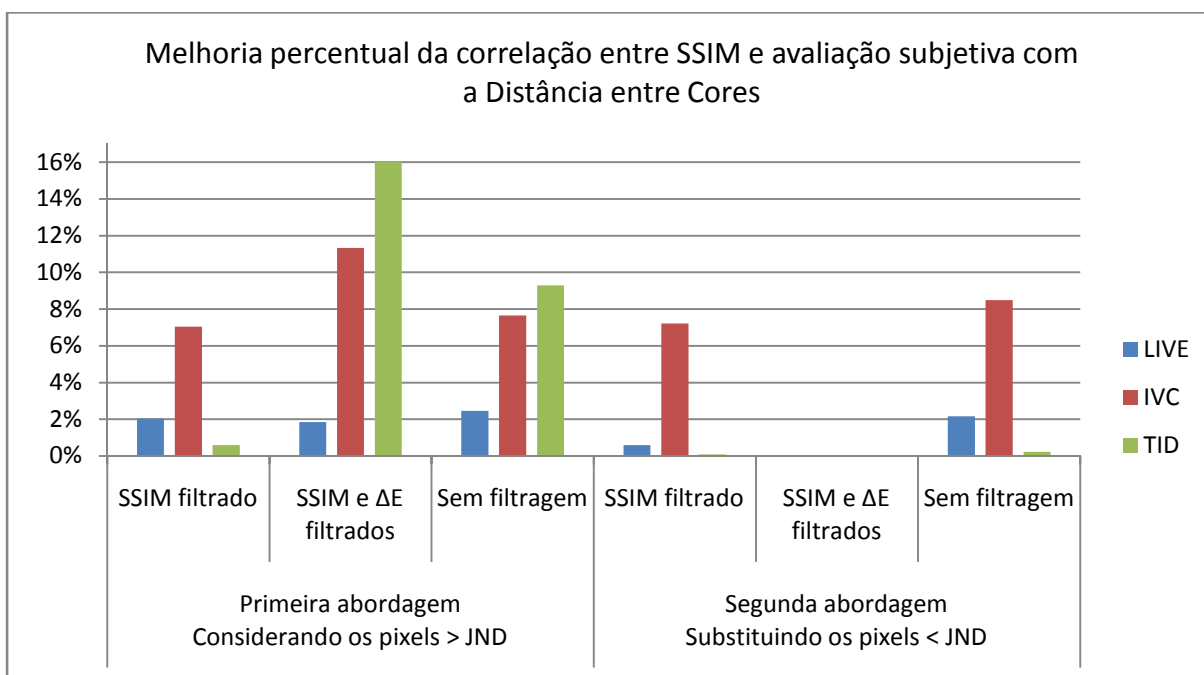


Figura 56 – Melhoria percentual com a utilização da distância entre cores para o melhor caso de cada experimento

Os ganhos obtidos no teste com o SSIM filtrado variaram entre 0,59% e 7,22%, obtidos respectivamente para o banco de dados TID na primeira abordagem e para o banco de dados IVC na segunda abordagem. No teste sem filtragem os ganhos variaram entre 0,23% e 9,29%, alcançados respectivamente utilizando o banco de dados TID na segunda abordagem e na

primeira abordagem. No teste com SSIM e  $\Delta E$  filtrados o ganho foi de 1,85% a 16,02%, utilizando respectivamente os bancos de dados LIVE e TID para a primeira abordagem. A segunda abordagem deste teste não produziu nenhuma melhoria na correlação para valores JND maiores que 0.

Com relação ao valor JND ideal, a Tabela 12 apresenta, em resumo, os valores JND para os quais foram alcançadas as maiores correlações de cada teste em cada abordagem e para cada banco de dados.

Tabela 12 – Valores JND dos picos de correlação

JND do pico de correlação				
Abordagem	Teste	LIVE	IVC	TID
Abordagem 1	SSIM filtrado	2,6	3,0	2,0
	SSIM e $\Delta E$ filtrados	1,4	5,4	0,4
	Sem filtragem	2,6	2,8	0,8
Abordagem 2	SSIM filtrado	1,2	2,8	1,4
	SSIM e $\Delta E$ filtrados	0,0	0,0	0,0
	Sem filtragem	2,0	2,8	2,2

Nota-se que não há um consenso de um valor JND para o pico de correlação para todos os testes, mas é possível identificar a faixa de valores JND entre 0,8 e 3 no qual ocorreram as correlações de pico em 4 dos 6 testes realizados (excluindo os testes com SSIM e  $\Delta E$  filtrados com as duas abordagens).

Os valores JND por teste foram os seguintes: No teste com SSIM filtrado da primeira abordagem os picos de correlação foram alcançados dentro de uma faixa de JND entre 2 e 3. No mesmo teste da segunda abordagem a faixa de JND foi entre 1,2 e 2,8. No teste com SSIM e  $\Delta E$  filtrados há uma total discordância para o JND do pico de correlação. Na primeira abordagem a faixa de valores JND foi de 0,4 a 5,4, enquanto que na segunda abordagem o pico de correlação ocorreu em  $JND = 0$ , isto é, para este teste a utilização da distância entre cores não representou qualquer melhoria. No teste sem filtragem, para a primeira abordagem, os picos de correlação ocorreram em valores JND entre 0,8 e 2,8, enquanto que para a segunda abordagem a faixa de valores JND ficou entre 2,0 e 2,8.

De forma geral, o teste que apresentou os resultados mais consistentes foi com o SSIM filtrado utilizando a primeira abordagem, fazendo o cálculo do SSIM somente para pixels  $>$  JND. Para esse teste, ainda que com diferenças pequenas com relação aos outros testes, as correlações dos bancos de dados LIVE e IVC apresentaram os maiores valores. Além disso,

os valores de pico da correlação para todos os bancos de dados ficaram dentro de uma faixa de valores JND relativamente pequena, entre 2 e 2,6.

## 8 CONCLUSÃO

Este trabalho se propôs a analisar o comportamento do SSIM na avaliação da qualidade de imagens coloridas, testando diferentes espaços de cores e propondo a utilização da distância entre cores calculada no espaço CIELAB.

Para tanto, os testes foram divididos em três etapas: na primeira foram executados testes com os canais individuais dos espaços de cores; na segunda foi proposto um agrupamento dos canais; na terceira foi investigada a aplicação da mínima diferença perceptível – o  $\Delta E$ .

Na primeira etapa, foram testados quatro diferentes espaços de cores, RGB, YCbCr,  $L\alpha\beta$  e CIELAB, aplicados na avaliação da qualidade de imagens coloridas. Esta etapa envolveu três testes: a aplicação do SSIM aos canais individuais dos espaços de cores, combinações alternativas das componentes do SSIM e uma análise do SSIM por tipo de distorção.

No primeiro teste dessa etapa foi observado que os espaços de cores modelados com um canal de luminância – YCbCr,  $L\alpha\beta$  e CIELAB, apresentaram resultados mais consistentes do que o espaço RGB para a aplicação do SSIM em um dos canais individuais.

No segundo teste realizou-se um estudo para avaliar se somente o contraste, somente a luminância ou uma combinação diferente das componentes do SSIM aplicado aos canais do espaço CIELAB apresentaria melhores resultados que o SSIM completo. O resultado foi que o SSIM completo produziu melhores resultados que as outras combinações de suas componentes.

O terceiro teste da primeira etapa aplicou o SSIM nos canais de luminância dos espaços YCbCr,  $L\alpha\beta$  e CIELAB na avaliação de subgrupos de imagens do banco de dados LIVE produzidas com o mesmo tipo de distorção. Dos três espaços o CIELAB apresentou o melhor resultado. Foram feitas então análises por tipo de distorção, plotando suas curvas de ajuste e calculando suas correlações. Como resultado foram observadas maiores correlações por certos tipos de distorção como embaçamento, ruído Gaussiano e erro de bit, do que quando a análise é realizada para todas as distorções. As imagens distorcidas por JPEG obtiveram avaliação pelo SSIM menos correlacionada com a avaliação subjetiva que as imagens com outras distorções.

Na segunda etapa, com a motivação de investigar se uma combinação não-linear entre os canais do espaço de cores produziria melhores resultados, foram propostos 6 diferentes cálculos para o SSIM – os chamados SSIM Compostos. Os resultados destes testes mostraram que o cálculo do SSIM no canal de luminância do espaço de cores é mais vantajoso que a utilização dos SSIM Compostos.

A terceira etapa de testes aplicou o conceito da mínima diferença de cor perceptível ao SSIM na avaliação objetiva de imagens a fim de melhorar a correlação com a métrica subjetiva. Para este teste foi utilizado somente o espaço de cores CIELAB, pois esse espaço permite que seja calculada a distância entre cores  $\Delta E$  necessária para a aplicação do conceito da mínima diferença de cor perceptível.

Nesta etapa foram feitos diferentes testes considerando duas abordagens distintas para emprego da distância de cores conjuntamente com o SSIM – uma calculando o  $\Delta E$  como a média dos pixels maiores que JND e outra realizando o cálculo do SSIM de todos os pixels, mas substituindo os pixels onde  $\Delta E$  é menor que JND pela imagem de referência – a fim de avaliar qual deles configura a melhor opção. Para cada abordagem foram realizados três testes: com o SSIM filtrado, com o SSIM e  $\Delta E$  filtrado e sem nenhuma filtragem.

Os resultados dos testes da terceira etapa não concordaram muito, entretanto é possível tirar algumas conclusões dos resultados. Verifica-se que para todos os casos, exceto em um, a aplicação da distância entre cores aumentou a correlação do SSIM com a métrica subjetiva, porém dependendo do banco de dados usado a melhora foi mais ou menos evidente. Com o banco de dados LIVE a máxima melhoria obtida na correlação com o uso do  $\Delta E$  foi de 2,47% e a menor de 0,6% (excluindo o teste com filtragem do  $\Delta E$  e SSIM da segunda abordagem, caso em que a melhora foi de 0%). Entretanto, com o banco de dados IVC, a melhoria máxima na correlação foi de 11,34% e a menor de 7,05% (excluindo o mesmo caso citado anteriormente), o que já representa uma melhoria considerável. O banco de dados TID apresentou melhoria de até 35% na correlação usando o  $\Delta E$ , porém esse banco de dados apresentou resultados de correlação muito menores que os demais em todos os experimentos realizados. Também se verificou que a utilização da filtragem realizada nos testes com SSIM filtrado, simulando o comportamento passa-baixa presente no olho humano, produziu resultados melhores quando comparados a não utilização da filtragem.

Outra característica avaliada nestes testes foi o JND para o qual a maior correlação seria obtida. Havia a expectativa de se obter valores na faixa entre 1 a 3, pois segundo referências o valor do JND, isto é, o limiar de percepção da diferença entre cores, se situa nessa faixa. Não houve consenso sobre o valor JND ideal nos testes, mas foi possível identificar que na faixa de valores JND entre 0,8 e 3 ocorreram as correlações de pico em 4 dos 6 testes realizados.

Dos testes realizados nesta etapa, o que apresentou os resultados mais consistentes foi com a utilização do SSIM a partir da imagem filtrada para a primeira abordagem, isto é, fazendo a média dos pixels onde  $\Delta E > \text{JND}$ . Este teste apresentou as maiores correlações obtidas utilizando os bancos de dados LIVE e IVC e obteve curvas de correlação e das demais



métricas de análise semelhantes para os três bancos de dados analisados. As correlações de pico ocorreram na faixa de JND entre 2 e 3, obedecendo ao comportamento esperado com uma ascensão, com o pico dentro da faixa de valores JND proposta, seguida de um decaimento.

### **8.1 Observações e sugestões para trabalhos futuros**

Foi observado que os resultados obtidos utilizando o banco de dados TID foram, no geral, bem piores do que os resultados obtidos utilizando os bancos de dados LIVE ou IVC. O motivo dessa discrepância pode ser devido à grande quantidade de imagens do banco de dados e a seus diferentes tipos de distorções. A faixa dinâmica da avaliação subjetiva pode ser diferente para diferentes tipos de distorções, o que atrapalha o cômputo de resultados quando todas as imagens do banco de dados são usadas de uma vez. Uma sugestão é estudar esse banco de dados por partes, focando em cada tipo de distorção independentemente.

Na formulação do SSIM a idéia de contraste é representada pelos desvios padrão entre os valores das janelas. Esta representação de contraste é bem simples e poderia ser melhorada se fossem considerados, por exemplo, a similaridade dos histogramas dos valores dos pixels utilizando a Kullback-Leibler (51).

Recentemente uma nova métrica de avaliação da qualidade usando a resposta Harris (52) foi proposta. O conceito da mínima diferença perceptível calculado no canal L do espaço de cores CIELAB poderia ser aplicado a essa e outras métricas de avaliação objetiva da qualidade a fim de verificar uma melhoria dos resultados objetivos.

Existem outros cálculos de diferença entre cores, as chamadas de fórmulas de diferença de cor avançadas, como o CIE94 e o CIEDE2000. Uma investigação sobre seu emprego juntamente ao SSIM e/ou a outras métricas poderia ser feito a fim de verificar se produzem resultados melhores que a fórmula tradicional empregada neste trabalho.

## REFERÊNCIAS

1. **A. Toet, M. P. Lucassen.** A New Universal Color Image Fidelity Metric. *Displays*. Dec. 2003, Vol. 24, no. 4-5, pp. 197–207.
2. **Z. Wang, H. R. Sheikh, A. C. Bovik.** Objective Video Quality Assessment. [A. do livro] O. Marqure, B. Furht. *The Handbook of Video Databases: Design and Applications*. 2nd ed. Boca Raton, FL, United States of America : CRC Press, 2003, 41, pp. 1041-1078.
3. **H. R. Wu, K. R. Rao.** *Digital Video Image Quality and Perceptual Coding*. Boca Raton, FL, United States of America : CRC Press, 2006.
4. **VQEG.** Final Report from the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment, Phase II. [Online] Aug. 2003. <http://www.vqeg.org/>.
5. **S. Winkler.** *Digital Video Quality: Vision Models and Metrics*. Chichester, England : John Wiley & Sons, 2005.
6. **Z. Wang, A. C. Bovik.** Mean Squared Error: Love It or Leave It? A new look at signal fidelity metrics. *IEEE Signal Processing Magazine*. Jan. 2009, Vol. 26, no. 1, pp. 98–117.
7. **Y. -K. Lai, J. Guo, C. -C. J. Kuo.** Perceptual Fidelity Measure of Digital Color Images. *Proc. SPIE. Human Vision and Electronic Imaging III*, 1988, Vol. 3299, pp. 221-231.
8. **A. B. Watson.** Towards a Perceptual Video Quality Metric. *Human Vision, Visual Processing, and Digital Display VIII*. 1998, Vol. 3299, pp. 139-147.
9. **Z. Wang, A. C. Bovik, H. R. Sheikh, E. P. Simoncelli.** Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Trans. Image Processing*. Apr. 2004, Vol. 13, no. 4, pp. 600–612.
10. **D. Salomon, G. Motta, D. Bryant.** *Data Compression: The Complete Reference*. s.l. : Springer, 2007.
11. **R. L. Myers.** *Display Interfaces Fundamentals and Standards*. Chichester, England : John Wiley and Sons, 2002.
12. **D. L. Ruderman, T. W. Cronin and C.-C. Chiao.** Statistics of Cone Responses to Natural Images: Implications for Visual Coding. *Journal of the Optical Society of America*. Aug. 1998, Vol. 15, no. 8, pp. 2036–2045.
13. **J. Schanda.** *Colorimetry: Understanding the CIE System*. s.l. : Wiley-Interscience, 2007.
14. **M. Ebner.** *Color Constancy*. England : Wiley-IS&T Series in Imaging Science & Technology, 2007.

15. **G. Wyszecki, W. S. Stiles.** *Color Science: Concepts and Methods, Quantitative Data and Formulae*. 2nd. ed. New York, United States of America : John Wiley & Sons, 1982.
16. **P. Le Callet, F. Atrousseau.** Subjective Quality Assessment IRCCyN/IVC Database. [Online] 2005. <http://www.irccyn.ec-nantes.fr/ivcdb/>.
17. **H. R. Sheikh, Z. Wang, L. Cormack, A. C. Bovik.** LIVE Image Quality Assessment Database Release 2. [Online] 2009. <http://live.ece.utexas.edu/research/quality>.
18. **N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola, F. Battisti.** Tampere Image Database 2008 TID2008, version 1.0. [Online] 2009. <http://www.ponomarenko.info/tid2008.htm>.
19. **International Telecommunication Union.** *Subjective Video Quality Assessment Methods for Multimedia Applications Recommendation P.910*. Geneva, Switzerland : s.n., 1999.
20. —. *Recommendation ITU-T P.910. Subjective Video Quality Assessment Methods for Multimedia Applications*. Geneva, Switzerland : s.n., 1999.
21. **H. R. Sheikh, M. F. Sabir, A.C. Bovik.** A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms. *IEEE Transactions on Image Processing*. Nov. 2006, Vol. 15, no. 11, pp. 3440-3451.
22. **D. S. Taubman, M. W. Marcellin.** *JPEG2000: Image Compression Fundamentals, Standards and Practice*. s.l. : Kluwer Academic Publishers, 2001.
23. Matlab. [www.mathworks.com](http://www.mathworks.com). [Online]
24. **A. Ninassi, P. Le Callet, F. Atrousseau.** Pseudo No Reference Image Quality Metric Using Perceptual Data Hiding. *SPIE Human Vision and Electronic Imaging*. Jan. 2006, Vol. 6057, no. 8.
25. **N. Ponomarenko, M. Carli, V. Lukin, K. Egiazarian, J. Astola and F. Battisti.** Color Image Database for Evaluation of Image Quality Metrics. *Proc. of International Workshop on Multimedia Signal Processing*. Oct. 2008, pp. 403-408.
26. **Z. Wang, A. C. Bovik.** *Modern Image Quality Assessment*. San Rafael, CA, United States of America : Morgan & Claypool, 2006.
27. —. A Universal Image Quality Index. *IEEE Signal Processing Lett.* Mar. 2002, Vol. 9, no. 3, pp. 81–84.
28. **E. P. Simoncelli, B. Olshausen.** Natural Images Statistics and Neural Representation. *Annu. Rev. Neuroscience*. May, 2001, Vol. 24, pp. 1193-1216.
29. **A. K. Jain.** *Fundamentals of Digital Image Processing*. Englewood Cliffs, NJ, United States of America : Prentice-Hall, 1989.

30. **J. J. McCann, S. P. McKee, T. H. Taylor.** Quantitative Studies in Retinex Theory. *Vision Research*. 1976, 16, pp. 445-458.
31. **J. C. Russ.** *The Image Processing Handbook*. s.l. : CRC Press, 2007.
32. **Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli.** The SSIM Index for Image Quality Assessment – A Matlab Implementation. [Online] 2009. <http://www.ece.uwaterloo.ca/~z70wang/research/ssim/>.
33. **G. L. Fain, J. E. Dowling.** Intracellular Recordings from Single Rods and Cones in the Mudpuppy Retina. *Science*. 1973, 180, pp. 1178-1181.
34. **H. Wässle.** A Patchwork of Cones. *Nature*. Feb. 1999, Vol. 397.
35. **M. Nadenau.** *Integration of Human Color Vision Models into High Quality Image Compression*. Lausanne, France : Tese de Doutorado - École Polytechnique Fédérale de Lausanne, 2000. p. 185. no. 2296.
36. **K. K. De Valois.** *Seeing*. Handbook of Perception and Cognition Series. s.l. : Academic Press, 2000.
37. **S. K. Shevell, Optical Society of America.** *The Science of Color*. 2nd Edition. s.l. : Elsevier, 2003.
38. **E. Hering.** *Zur Lehre vom Lichtsinne*. Vienna, Austria : Carl Gerolds & Sohn, 1878.
39. **S. Winkler.** Issues in Vision Modeling for Perceptual Video Quality Assessment. *Signal Processing*. Oct. 1999, Vol. 78, no. 2, pp. 231–252.
40. **G. Svaetichin.** Spectral Response Curves From Single Cones. *Acta Physiologica Scandinavica*. 1956, no. 134, pp. 17-46.
41. **H. von Helmholtz.** *Treatise on Physiological Optics*. Dover, NY, United States of America : Southall J. P., 1962.
42. **A. M. Marcos, P. M. Martín.** *Colorimetría*. Madrid, Spain : Editorial Ciencia 3, 1996.
43. **M. Vrhel, E. Saber, H. J. Trussell.** Color Image Generation and Display Technologies. *IEEE Signal Processing Magazine*. Jan. 2005.
44. **International Telecommunication Union.** *Recommendation ITU-R BT.601-4. Encoding Parameters of Digital Television for Studios*. 1994.
45. **T. Acharya, A. K. Ray.** *Image Processing: Principles and Applications*. s.l. : John Wiley and Sons, 2005.
46. **Commission Internationale de L'Eclairage.** *Colorimetry. CIE Publication No. 15.2*. 1986.

47. **M. Melgosa, M. M. Pérez, A. Yebra, R. Huertas, E. Hita.** Some Reflections and Recent International Recommendations on Color Difference Evaluation. *Óptica Pura y Aplicada*. 2001, Vol. 34.
48. **D. L. MacAdam.** Specification of Small Chromaticity Differences in Daylight. *Journal of the Optical Society of America*. Jan. 1943, Vol. 33, no. 1.
49. **M. Melgosa.** Color-Difference Formulas. *Balkan Light 2008*. 2008.
50. **A. M. Rohaly, P. Corriveau, J. Libert et al.** Video Quality Experts Group: Current Results and Future Directions. *Proceedings of SPIE Visual Communications and Image Processing*. 2000, Vol. 4067, no. 3, pp. 742-753.
51. **Y. Rubner, J. Puzicha, C. Tomasi, J. M. Buhmann.** Empirical Evaluation of Dissimilarity Measures for Color and Texture. [ed.] Elsevier Science. *Computer Vision and Image Understanding*. 2001, no. 84, pp. 25-43.
52. **D. -O. Kim, R. -H. Park.** New Image Quality Metric Using the Harris Response. *IEEE Signal Processing Letters*. Jul, 2009, Vol. 16, no. 7.
53. **M. Carnec, P. Le Callet, D. Barba.** Objective Quality Assessment of Color Images Based on a Generic Perceptual Reduced Reference. *Signal Processing Image Communication*. April 2008, Vol. 23, no. 4, pp. 239-256.
54. **H. Wässle.** Parallel Processing in the Mammalian Retina. *Nature Reviews. Neuroscience*, Oct. 2004, Vol. 5.
55. **International Telecommunication Union.** *Recommendation ITU-T BT.601-6. Studio Encoding Parameters of Digital Television for Standard 4:3 and Wide Screen 16:9 Aspect Ratios*. 2007.
56. **R. A. Young.** Oh Say, Can You See? The Physiology of Vision. *Proc. SPIE Human Vision, Visual Processing and Digital Display*. 1991, Vol. 1453, pp. 92-123.