UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

COORDENAÇÃO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

DISSERTAÇÃO DE MESTRADO

RENDERIZAÇÃO NÃO FOTORREALÍSTICA: UMA ABORDAGEM PARA ABSTRAÇÃO DE IMAGENS COLORIDAS DIGITAIS

JOÃO BOSCO DE OLIVEIRA ALENCAR JÚNIOR

ORIENTADORES

HERMAN MARTINS GOMES

JOSÉ EUSTÁQUIO RANGEL DE QUEIROZ

CAMPINA GRANDE, PARAÍBA

NOVEMBRO – 2010

UNIVERSIDADE FEDERAL DE CAMPINA GRANDE

CENTRO DE ENGENHARIA ELÉTRICA E INFORMÁTICA

COORDENAÇÃO DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

RENDERIZAÇÃO NÃO FOTORREALÍSTICA: UMA ABORDAGEM PARA ABSTRAÇÃO DE IMAGENS COLORIDAS DIGITAIS

JOÃO BOSCO DE OLIVEIRA ALENCAR JÚNIOR

Dissertação submetida à Coordenação do Curso de Pós-Graduação em Ciência da Computação da Universidade Federal de Campina Grande como parte dos requisitos necessários para obtenção do grau de Mestre em Ciência da Computação.

Área de concentração: Ciência da Computação

Linha de pesquisa: Modelos Computacionais e Cognitivos

José Eustáquio Rangel de Queiroz

Herman Martins Gomes

Orientadores

Campina Grande, Paraíba

Novembro – 2010

FICHA CATALOGRÁFICA ELABORADA PELA BIBLIOTECA CENTRAL DA UFCG

A368r Alencar Júnior, João Bosco de Oliveira

Renderização não fotorrealística: uma abordagem para abstração de imagens coloridas digitais/João Bosco de Oliveira Alencar Júnior. ─ Campina Grande, 2010.

121 f.: il. col.

Dissertação (Mestrado em Ciência da Computação) – Universidade Federal de Campina Grande, Centro de Engenharia Elétrica e Informática**.**

Orientadores: Prof. Dr. José Eustáquio Rangel de Queiroz, Profº. Ph.D. Herman Martins Gomes.

Referências.

1. Processamento Digital de Imagens. 2. Representações Não Fotorrealísticas. 3. Abstração de Imagens. I. Título.

CDU 004.932(043)

"RENDERIZAÇÃO NÃO FOTO-REALÍSTICA: UMA ABORDAGEM PARA ABSTRAÇÃO DE IMAGENS COLORIDAS DIGITAIS"

JOÃO BOSCO DE OLIVEIRA ALENCAR JUNIOR

DISSERTAÇÃO APROVADA EM 30.11.2010

HERMAN MARTINS GOMES, Ph.D
Orientador(a)

JOSÉ EUSTÁQUIO RANGEL DE QUEIROZ, D.Sc
Orientador(a)

JOSEANA MACÊDO FECHINE, D.Sc
Examinador(a)

JOÃO MARQUES DE CARVALHO, Ph.D
Examinador(a)

BRUNO MOTTA DE CARVALHO, Ph.D
Examinador(a)

CAMPINA GRANDE - PB

AGRADECIMENTOS

À Deus.

A todos os membros da minha família, por terem me ajudado sempre.

Aos professores Herman e Eustáquio, pela paciência, ajuda e pelos ensinamentos que me proporcionaram durante o período do mestrado.

Ao professor Almerindo, por ter me ensinado o caminho para ingressar no mestrado.

Aos membros da banca examinadora, pelas contribuições e sugestões de melhorias para este trabalho.

Aos amigos que me receberam em Campina Grande e me ajudaram desde antes do início do mestrado: Flávio, Gilson, Fernando, Felipe, Saulo, Bruno Alexandre, Carlos Augusto, Marcos, Welflen, Guilherme, e demais membros da QI.

Aos amigos que fizeram parte do Laboratório de Visão Computacional (LVC), entre eles: Cláudio, Eanes, Vinícius, Luciana, Odlon, Paulo, Eduardo, Bruno Brito, Caio, Alexandre, Einstein, Harry Potter, Kimmi Raikkonen.

À Xhycko, também integrante do LVC, pelas contribuições e importante ajuda durante as pesquisas na área de NPR.

Aos estudantes que participaram dos experimentos realizados neste trabalho.

Aos professores e funcionários da UFCG, especialmente aos membros da COPIN Aninha e Vera, pela paciência e ajuda.

À CAPES e à HP, pelo apoio financeiro.

RESUMO

A renderização não fotorrealística é uma área de estudo que engloba uma ampla gama de técnicas compartilhadas pelas áreas da Computação Gráfica e do Processamento Digital de Imagens que visa à geração de imagens artísticas a partir de cenários, modelos geométricos ou imagens fotográficas. Neste contexto, a abstração de imagens digitais constitui-se em um paradigma que objetiva transformar imagens fotográficas complexas em imagens artísticas simplificadas, a fim de facilitar a comunicação visual, enquanto preserva seu conteúdo principal. Nesta dissertação, apresenta-se o arcabouço concebido para a geração de imagens abstraídas ao estilo de cartuns, a partir de algoritmos para suavização e segmentação de regiões, detecção de bordas, ajuste de brilho/saturação e harmonização de cores em imagens, os quais foram investigados ao longo desta pesquisa. No intuito de validar os resultados obtidos, foi realizado um estudo experimental envolvendo avaliadores humanos, no qual se comparou a abordagem proposta com outras cinco abordagens existentes para a cartunização de imagens. A análise estatística realizada sobre os dados experimentais, a partir de testes ANOVA e Tukey-Kramer, evidenciou a existência de diferenças significativas entre os resultados gerados pela abordagem proposta e aqueles gerados pelas demais abordagens comparadas.

ABSTRACT

Non-photorealistic rendering is common to both Computer Graphics and Digital Image Processing, which encompasses a broad class of techniques for creating artistic images from scenes, geometric models or photographic images. In this context, digital image abstraction constitutes a paradigm that aims to transform complex photographic images into simplified artistic images in order to facilitate visual communication, while preserving its main content. In this dissertation it is presented a framework designed to generate cartoon-like abstracted images from algorithms of region smoothing and segmentation, edge detection, adjustment of brightness/saturation and color harmonization, which were investigated throughout this research. In order to validate these results, an experimental study involving human evaluators was conducted, in which it was compared the proposed approach with five other existing approaches for image cartoonization. Statistical analysis performed on experimental data from one-way ANOVA and Tukey-Kramer tests revealed significant differences in preference between results created by the proposed approach and those created by the other approaches considered in this research.

SUMÁRIO

1 INTRODUÇÃO 1

1.1 DESCRIÇÃO DO PROBLEMA 3

1.2 MOTIVAÇÃO 4

1.3 OBJETIVOS 6

1.4 ESTRUTURA DA DISSERTAÇÃO 6

2 TRABALHOS RELACIONADOS 8

2.1 BREVE HISTÓRICO DE NPR 8

2.2 ABSTRAÇÃO DE IMAGENS 13

2.2.1 SUAVIZAÇÃO COM PRESERVAÇÃO DE BORDAS 25

2.2.2 SEGMENTAÇÃO DE IMAGENS 35

2.2.3 AJUSTE DE CORES 41

2.3 CONSIDERAÇÕES FINAIS DO CAPÍTULO 47

3 ABORDAGEM PROPOSTA 49

3.1 VISÃO GERAL DA ARQUITETURA PROPOSTA 49

3.2 ABSTRAÇÃO DA IMAGEM 51

3.2.1 SIMPLIFICAÇÃO POR SUAVIZAÇÃO 51

3.2.2 SIMPLIFICAÇÃO POR SEGMENTAÇÃO 58

3.2.3 EXTRAÇÃO DE BORDAS 59

3.3 AJUSTE DE CORES 61

3.3.1 HARMONIZAÇÃO DE CORES 61

3.3.2 AJUSTE DE BRILHO E SATURAÇÃO 63

3.4 DETALHES DE PROJETO E IMPLEMENTAÇÃO 64

3.5 CONSIDERAÇÕES FINAIS DO CAPÍTULO 66

4 AVALIAÇÃO EXPERIMENTAL 67

4.1 COMPARAÇÃO VISUAL DOS RESULTADOS 68

4.2 EXPERIMENTO SUBJETIVO 72

4.2.1 SOFTWARES UTILIZADOS 72
4.2.2 CALIBRAÇÃO DE PARÂMETROS 74
4.2.3 COMPARAÇÃO DE RESULTADOS 80
4.3 CONSIDERAÇÕES FINAIS DO CAPÍTULO 85
5 CONSIDERAÇÕES FINAIS 86
5.1 SÍNTESE DA PESQUISA 86
5.2 PRINCIPAIS CONTRIBUIÇÕES 87
5.3 TRABALHOS FUTUROS 88
REFERÊNCIAS 90
APÊNDICE A AMOSTRA DE RESULTADOS 99
APÊNDICE B VARIAÇÕES DO FILTRO BILATERAL 103
APÊNDICE C CONVERSÕES ENTRE ESPAÇOS DE CORES 108
APÊNDICE D TESTES DE TUKEY - CALIBRAÇÃO DE PARÂMETROS 111
APÊNDICE E TESTES DE HIPÓTESES 118

LISTA DE FIGURAS

Figura 1.1 Exemplos de utilização de NPR na mídia. 5
Figura 2.1 Exemplos de resultados de NPR 11
Figura 2.2 Contextualização da pesquisa perante os principais domínios relacionados à área de NPR 12
Figura 2.3 Avenidas de Londres. 13
Figura 2.4 Abstração automatizada de imagem 14
Figura 2.5 Operadores de *Sobel* 16
Figura 2.6 Quantização de cores (redução para 16 cores). 17
Figura 2.7 Sequência de passos do método proposto por Wen et al. (2006) 19
Figura 2.8 Abordagem proposta por Winnemöller, Olsen e Gooch (2006) 21
Figura 2.9 Detecção de contornos 23
Figura 2.10 Curva de *Gauss* 27
Figura 2.11 Filtragem por mediana aplicado a diferentes vizinhanças 28
Figura 2.12 Iterações do filtro *Bilateral*. 29
Figura 2.13 Agrupamentos do filtro *Kuwahara* para um pixel localizado na posição (x,y) da imagem. 30
Figura 2.14 Exemplo da Filtragem *Mean Shift* com indicação das coordenadas (x,y), dos centros e das janelas resultantes 32
Figura 2.15 Exemplos de resultados de filtros de suavização. 33
Figura 2.16 Exemplos de resultados NPR de filtros de suavização. 34
Figura 2.17 Redução de cores. 35
Figura 2.18 Componentes Conectados. 36
Figura 2.19 Exemplo de corte no grafo. 38
Figura 2.20 Comparação entre algoritmos de segmentação realizada por Estrada e Jepson (2005). 39
Figura 2.21 Baixa taxa de detecção. 40
Figura 2.22 Espaço cilíndrico HSV. 42
Figura 2.23 Diferenças entre histogramas. 43
Figura 2.24 Transformações *gama* 44

Figura 2.25 Deslocamento de cores. 45
Figura 2.26 Templates Harmônicos. 46
Figura 3.1 Arquitetura da abordagem proposta. 50
Figura 3.2 Filtro *Bilateral Baseado em Fluxo*..... 52
Figura 3.3 Filtro *Bilateral Adaptativo*. 53
Figura 3.4 Ilustração do Filtro FABL.. 54
Figura 3.5 Funções aproximadas às propostas por Zhang e Allebach (2008). 56
Figura 3.6 Redução de cores.. 57
Figura 3.7 Segmentação Mean Shift. 58
Figura 3.8 Suavização das bordas. 59
Figura 3.9 Sobreposição dos contornos. 60
Figura 3.10 Harmonização de cores. 62
Figura 3.11 Ajuste de brilho e saturação. 63
Figura 3.12 Diagrama de classes da aplicação de *software* desenvolvida.. 65
Figura 4.1 Comparações de abstração por suavização. 68
Figura 4.2 Comparações de abstração por suavização. 69
Figura 4.3 Comparações de abstração por segmentação. 70
Figura 4.4 Comparações de abstração por segmentação. 71
Figura 4.5 Softwares utilizados para comparação. 73
Figura 4.6 Excerto da interface com o usuário utilizada na etapa de calibração de parâmetros. 77
Figura 4.7 Excerto da interface com o usuário utilizada na etapa de comparação dos resultados das aplicações de *software* utilizadas. 81
Figura B.1 Direção tangente (Tx) e perpendicular (Gx) ao fluxo..... de tangente de bordas no ponto X. 105
Figura B.2 Valores treinados a partir do LoG. 107

LISTA DE TABELAS

Tabela 4.1 Teste F ANOVA fator único para a calibração dos softwares.....79
Tabela 4.2 Média e desvio padrão da quantidade relativa de votos obtidos por cada aplicação de software. 82
Tabela 4.3 Teste F ANOVA fator único para a comparação dos softwares....82
Tabela 4.4 Teste de Tukey para a comparação dos softwares. 83
Tabela D.1 Configurações de parâmetros do BeFunky. 111
Tabela D.2 Grupos de similaridades obtidos pelo teste de Tukey para o software BeFunky. 112
Tabela D.3 Configurações de parâmetros do CaricatureSoft. 112
Tabela D.4 Grupos de similaridades obtidos pelo teste de Tukey para o software CaricatureSoft. 113
Tabela D.5 Configurações de parâmetros do LiangZhu. 113
Tabela D.6 Grupos de similaridades obtidos pelo teste de Tukey para o software LiangZhu. 114
Tabela D.7 Configurações de parâmetros do MSU Cartoonizer. 114
Tabela D.8 Grupos de similaridades obtidos pelo teste de Tukey para o software MSU Cartoonizer. 115
Tabela D.9 Configurações de parâmetros do KusoCartoon. 115
Tabela D.10 Grupos de similaridades obtidos pelo teste de Tukey para o software MSU KusoCartoon. 116
Tabela D.11 Configurações de parâmetros da abordagem proposta. 117
Tabela D.12 Grupos de similaridades obtidos pelo teste de Tukey para a abordagem proposta. 117

LISTA DE QUADROS

Quadro 2.1 Métodos para abstração de imagens. 48

Quadro 4.1 – Parâmetros de controle dos softwares. 75
Quadro 4.2 – Critérios de avaliação. 76
Quadro 4.3 – Grupos de similaridade obtidos por meio do teste de Tukey. 84
Quadro A.1 Exemplos de resultados de imagens abstraídas por abordagens citadas na revisão bibliográfica. 99

LISTA DE SIGLAS E ABREVIATURAS

ANOVA	***AN****alysis* ***O****f* ***VA****riance* (Análise de Variância)
CG	**C**omputação **G**ráfica
CIE	***C****ommission* ***I****nternationale d'****E****clairage* (Comissão Internacional de Iluminação)
DoG	***D****ifference* ***o****f* ***G****aussians* (Diferença de Gaussianas)
ETF	***E****dge* ***T****angent* ***F****low* (Fluxo de Tangente de Bordas)
FABL	***F****low-Based* ***A****daptive* ***Bil****ateral* (Bilateral Adaptativo Baseado em Fluxo)
FBL	***F****low-Based* ***Bil****ateral* (Bilateral baseado em Fluxo)
FDoG	***F****low-Based* ***D****ifference* ***o****f* ***G****aussians* (Diferença de Gaussianas baseada em Fluxo)
f.d.p.	**f**unção **d**ensidade de **p**robabilidade
gl	**g**raus de **l**iberdade
HSV	***H****ue,* ***S****aturation,* ***V****alue* (Matiz, Saturação, Valor)
MQ	**M**édia dos **Q**uadrados
NPR	***N****on-****P****hotorealistic* ***R****endering* (Renderização Não Fotorrealística)
OpenCV	***Open*** *Source* ***C****omputer* ***V****ision Library* (Biblioteca de Visão Computacional de Código Aberto)
PDI	**P**rocessamento **D**igital de **I**magens
RGB	***R****ed,* ***G****reen,* ***B****lue* (Vermelho, Verde, Azul)
SQ	**S**oma dos **Q**uadrados

Capítulo 1

Introdução

Artistas norte-americanos denominaram de *fotorrealismo* o estilo de pintura à mão que simula o efeito de uma câmera fotográfica, cujas imagens produzidas são dificilmente distinguíveis de fotografias reais (Strothotte e Schlechtweg, 2002). Analogamente, o termo fotorrealismo foi incorporado à Ciência da Computação, como uma forma de reproduzir o mundo real por meio da geração automatizada de imagens indistinguíveis de fotografias.

Porém, projetistas gráficos admitem que nem sempre as imagens fotográficas são a melhor opção para transmitir informação (Markosian et al., 1997). Às vezes, produzir uma imagem cujo conteúdo seja apenas o suficiente para transmitir a informação desejada, pode torná-la mais facilmente compreensível. Quando o propósito da imagem não é ilustrar uma cena real, ou seja, quando o propósito não é a simulação do fenômeno físico da interação entre a luz e os elementos da foto, mas sim a transmissão da essência da cena em um contexto mais genérico ou abstrato (e.g., a ilustração de uma situação ou a exemplificação de um fato ou conceito), então o fotorrealismo pode tornar a informação mais confusa ou menos compreensível (Winnemöeller, 2006).

Existem muitos termos associados ao Processamento Digital de Imagens (PDI) relacionados à geração de imagens que criam a ilusão de realidade, ao invés de simulá-la, dentre os quais se podem citar *renderização artística*, *renderização estilizada*, *renderização simbólica*,

renderização instrutiva e *renderização expressiva*. Porém, o termo consensual que melhor abrange esta área é *renderização não fotorrealística*, usualmente referido como NPR[1] (Winnemöeller, 2006; Thoma, 2002).

Alguns autores afirmam que as pesquisas relacionadas ao *não fotorrealismo* surgiram no início dos anos 90, como resultado de uma contraposição à corrente de expressão fotorrealística (Thoma, 2002). Os resultados de tais pesquisas destinam-se não apenas a técnicas de abstração para a exibição de informação visual, mas também para imitar estilos visuais tradicionais, e.g., pinturas artísticas, desenhos a lápis, revistas em quadrinhos, mosaicos (Santella, 2005).

Segundo Gooch (2001), enquanto o fotorrealismo baseia-se em estudos de fenômenos físicos, tais como efeitos de luz, propriedades de superfícies e materiais, a área do não fotorrealismo envolve a estilização e a comunicação, combinando ciência e arte, em situações em que o processo para a geração da imagem é menos importante do que seu conteúdo comunicativo. Assim, diferentemente do fotorrealismo, em que a eficácia do processo é medida a partir da proximidade entre a imagem obtida e a cena original, trabalhos relacionados à NPR normalmente possuem caráter subjetivo, não existindo uma forma direta e precisa de mensuração de sua eficácia (Strothotte e Schlechtweg, 2002; Gooch, 2001).

Todavia, tanto para o fotorrealismo quanto para o não fotorrealismo, a cor é um atributo importante da imagem. Enquanto no primeiro caso as cores objetivam simular efeitos de luz da natureza (Winnemöeller, 2006), no segundo as cores da imagem são utilizadas como uma ferramenta para a criação da abstração (Santella, 2005).

Inserida neste contexto, a presente dissertação objetiva o estudo de transformações de cores de imagens digitais para o tratamento do problema da geração de imagens abstratas não fotorrealísticas.

---

[1] Sigla do termo da língua inglesa *Non-Photorealistic Rendering.*

1.1 Descrição do Problema

Com o intuito de facilitar a comunicação visual, o processo de abstração de imagens é um paradigma da área de NPR, a partir do qual a maior ênfase na redução de informação serve de base para a transformação de imagens digitais. Tal tarefa de reduzir a quantidade de informação na cena e evidenciar a informação visual principal deve ser feita de uma forma estilizada, tal que o resultado obtido por meio da abstração sirva ao propósito de comunicar a mensagem da cena, assim como seja agradável aos olhos do observador (Kang e Lee, 2008).

Este processo envolve dois principais subproblemas, a saber: (i) a *extração de linhas e contornos* e (ii) a *estilização de regiões*. A combinação das soluções para estes dois problemas geralmente simplifica a imagem original, produzindo imagens ao estilo de desenhos animados e cartuns (Kang, Lee e Chui, 2009).

O problema da *extração de linhas e contornos* tem por objetivo a obtenção de bordas significativas na imagem, resultando em traçados que retratam os elementos da imagem por meio de seus contornos. Estes traçados, sobrepostos à imagem, auxiliam a percepção de detalhes importantes, aumentando o contraste local entre as diferentes regiões que compõem a cena (Winnemöeller, 2006).

Por outro lado, o problema da *estilização de regiões* tem por objetivo a redução de descontinuidades entre pixels pertencentes a uma mesma região da imagem. A partir daí, podem ser obtidas imagens com regiões suavizadas ou homogêneas, o que implica a redução da quantidade de cores que compõem a imagem original. O subproblema de *estilização de regiões* é visto como o principal problema no processo abstração de imagens, responsável pela diminuição da complexidade da cena, enquanto preserva as estruturas importantes (Kang e Lee, 2008).

Além destes dois subproblemas, alguns trabalhos recentes (e.g., Wen et al., 2006; Medhi, Patabandhige e Toyama, 2007) propuseram em suas abordagens ajustes nas cores da imagem. Para tanto, utilizam técnicas

semi-automáticas com o propósito de modificar a saturação e o brilho, visando evidenciar partes de interesse na imagem.

O escopo deste trabalho envolve os subproblemas de estilização de regiões e de ajuste de cores, a partir de estudos de transformação de cores em imagens, por meio de técnicas de PDI relacionadas a filtros de suavização, técnicas de segmentação e ajuste de histogramas. Assim sendo, a abordagem ora descrita foi concebida com o intuito de possibilitar a geração de imagens abstraídas a partir de fotografias digitais.

Não faz parte do escopo deste trabalho um estudo aprofundado do subproblema de extração de linhas. Por conseguinte, a pesquisa ora documentada limitou-se, neste sentido, ao estudo dos principais métodos para a extração de linhas, descritos no estado da arte de abstração de imagens.

1.2 Motivação

O processamento de imagens digitais, com fins à NPR, vem despertando interesse na indústria de *software*, fato que pode ser constatado a partir dos financiamentos atuais de pesquisas por grandes empresas ao longo desta década, e.g., pela Microsoft, HP e IBM. Dentre os projetos recentes que atestam tal interesse pode ser citado o *PhotoToon*, desenvolvido no âmbito de uma cooperação técnico-científica firmada entre a HP e a Universidade Federal de Campina Grande, destinado ao desenvolvimento de pesquisas (as quais deram início ao trabalho apresentado nesta dissertação) relacionadas à criação de cartuns a partir de imagens fotográficas.

O interesse por NPR também é evidenciado na indústria cinematográfica, em filmes tais como *Waking Life* (2001) e *A Scanner Darkly* (2006), assim como em diversas produções televisivas (vinhetas de programas e telenovelas, assim como mensagens publicitárias) e outros meios de divulgação (e.g., revistas, *web sites*). Contudo, a

produção de tais animações de alta qualidade ainda exige muito tempo e esforço, as quais normalmente lançam mão de técnicas de rotoscopia digital (Rzeszutek, El-Maraghi e Androutsos, 2009).

Na Figura 1.1, são exibidas imagens não fotorrealísticas retiradas do filme *Waking Life* (Figura 1.1A), do filme *A Scanner Darkly* (Figura 1.1B), de um programa de TV (Figura 1.1C) e de um anúncio comercial (Figura 1.1D).

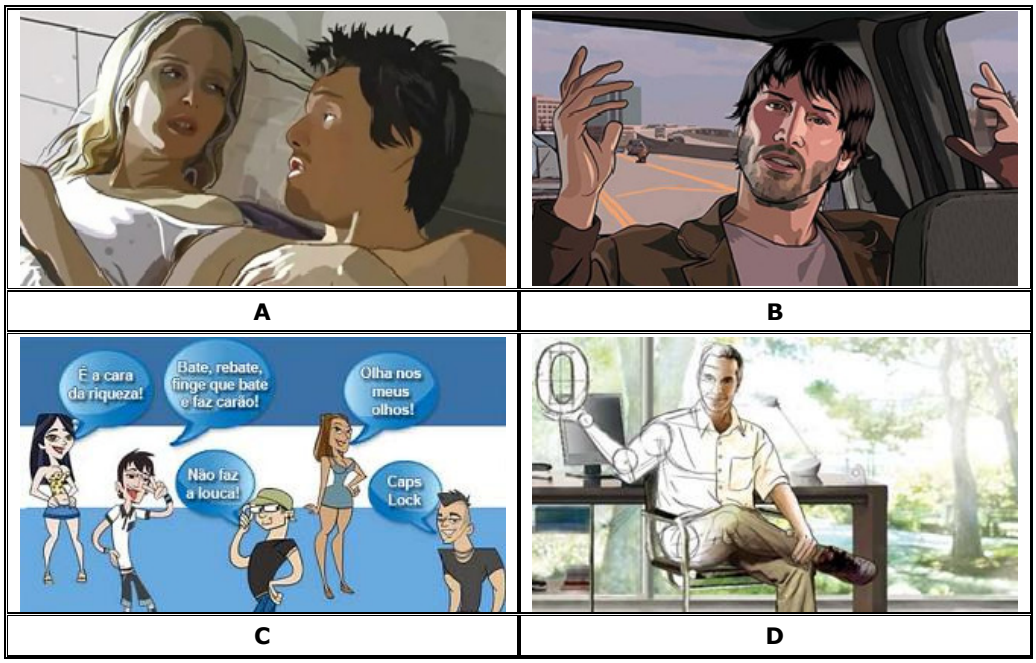


Figura 1.1 - Exemplos de utilização de NPR na mídia: (A) Imagem do filme *Waking Life*; (B) Imagem do filme *A Scanner Darkly*; (C) Animação do *Big Brother Brasil 10*; (D) Propaganda comercial do *Unibanco*.

Além disto, observa-se que o estudo e a concepção de métodos de NPR também resultam em produtos destinados a outras áreas da computação, e.g., aplicações computacionais para o desenvolvimento sócio-econômico (Medhi, Patabandhige e Toyama, 2007), envolvendo conhecimentos interdisciplinares, de áreas tais como a Filosofia, a Psicologia, as Neurociências, a Ciência Cognitiva e a História da Arte, dentre outras (Halper et al., 2003; Hertzmann, 2010).

Apesar de existir um crescente interesse em processos de abstração de imagens digitais, os estudos e as abordagens relacionadas são recentes

e estão em constante transformação. Desta forma, as soluções existentes ainda carecem de pesquisas, para que seja possível alcançar um bom nível de abstração das imagens, mantendo um alto nível de qualidade visual e exigindo o mínimo possível de intervenção do usuário.

1.3 Objetivos

O objetivo geral desta dissertação é propor e validar uma abordagem para abstração de imagens digitais que possibilite a geração de imagens ao estilo de cartuns, tendo como referência o estado da arte da área.

Esta abordagem consiste em uma seqüência de etapas, capazes de serem descritas por algoritmos computacionais que recebem como informação de entrada a imagem a ser processada, o resultado gerado na etapa anterior do processo e os parâmetros de ajuste dos algoritmos envolvidos. O resultado da última etapa da seqüência é a imagem abstraída.

Para atingir o objetivo geral, o presente trabalho exigiu a formulação dos seguintes objetivos específicos:

- Realizar pesquisa bibliográfica relacionada à abstração de imagens;
- Definir a arquitetura do sistema e os algoritmos a serem utilizados na abordagem adotada;
- Codificar os módulos do sistema proposto; e
- Validar os resultados produzidos pelo sistema desenvolvido.

1.4 Estrutura da Dissertação

Este documento é composto por mais quatro capítulos. No Capítulo 2, é apresentada uma síntese da revisão bibliográfica realizada ao longo da pesquisa, a fim de fundamentar a abordagem proposta nesta dissertação. No referido capítulo, a literatura mais relevante e recente relacionada à abstração de imagens é apresentada, assim como os principais trabalhos

relacionados a áreas que fazem parte da solução do problema de abstração de imagens.

No Capítulo 3, descreve-se a abordagem proposta para o problema de abstração de imagens, sendo apresentado o fluxo de processamento, desde a imagem de entrada até a geração do resultado final. Adicionalmente, são apresentados neste capítulo os métodos propostos que compõem os módulos do sistema, incluindo alguns detalhes da codificação e de organização interna do sistema desenvolvido com o propósito de validar a abordagem proposta. O funcionamento de cada módulo é ilustrado, incluindo detalhamentos de passos intermediários.

No Capítulo 4, são apresentados os experimentos realizados e a discussão dos resultados obtidos. São incluídos neste capítulo a descrição e os resultados do experimento de validação realizado.

Finalmente, no Capítulo 5, são apresentadas as considerações finais, as contribuições e recomendações de trabalhos derivados do tema abordado passíveis de pesquisas futuras.

Capítulo 2

Trabalhos Relacionados

Neste capítulo, é apresentado um breve histórico sobre NPR, contextualizando um resumo sobre os diferentes trabalhos que, em geral, vêm sendo desenvolvidos nesta área. Em seguida, são descritos os trabalhos mais específicos e de maior relevância encontrados na literatura, reunidos durante a pesquisa bibliográfica sobre o tema de abstração de imagens, nos quais são apresentadas as principais técnicas utilizadas para obtenção da imagem abstraída. Os trabalhos relacionados à abstração de imagens apresentados neste capítulo serviram como base para a concepção da abordagem para a geração de imagens abstratas, ao estilo de cartuns, proposto nesta dissertação. Uma breve contextualização sobre técnicas de suavização de imagens, segmentação e ajuste de cores também é apresentada neste capítulo, uma vez que estas técnicas são comumente utilizadas no processo de abstração de imagens.

2.1 Breve Histórico de NPR

Historicamente, o principal objetivo dos processos de renderização de imagens é produzir imagens que sejam indistinguíveis de fotografias (Lansdown e Schofield, 1995; Markosian, 97; Winnemöeller, 2006). Até a década de 90, havia um maior interesse no desenvolvimento das técnicas de renderização fotorrealística comparativamente às técnicas não fotorrealistas. A partir da década de 90, a área de NPR começou a

despertar interesse da comunidade científica (Thoma, 2002), estimulando pesquisas com diversos objetivos distintos, tais como a simulação de mídias artísticas tradicionais, o entendimento do sistema visual humano, a abstração de imagens e o aperfeiçoamento da aprendizagem, dentre outros (Gooch, 2003).

Os primeiros trabalhos nesta área objetivavam reproduzir formas tradicionais de arte, tais como pintura a óleo, aquarela e desenho a nanquim (Gooch, 2003). Em 1978, Smith (1978a) já apresentava à comunidade científica resultados de pesquisas na área de NPR, descrevendo uma ferramenta computacional que permitia a criação de pinturas artísticas digitais. Todavia, o autor menciona em seu trabalho a existência de pesquisas e programas mais antigos, destinados à criação de pinturas artísticas digitais, citando as ferramentas desenvolvidas nos *Bell Labs*, em 1969, assim como no *Xerox Palo Alto Research Center*, em 1972.

Inspiradas no trabalho de Smith (1978a), outras pesquisas desta época descreveram técnicas para aperfeiçoar a textura (e.g., Lewis, 1984) e a qualidade das pinceladas geradas digitalmente (e.g., Whitted, 1983; Strassmann, 1986), tendo como meta produzir pinturas semelhantes àquelas produzidas na mídia tradicional.

Também utilizando conceitos descritos por Smith (1978a), Haeberli (1990) desenvolveu uma técnica para a criação de pinturas impressionistas, a partir de uma imagem de entrada. Por meio de um programa interativo, o artista selecionava a direção, tamanho, formato e a posição que a pincelada a ser preenchida na tela da pintura, obtendo automaticamente a cor a partir das cores dos pixels da imagem de entrada. Haeberli (1990) mencionou que as imagens abstratas geradas por sua técnica filtravam e refinavam a informação visual, sendo uma alternativa para resultados gerados por fotorrealismo. Além disto, Haeberli (1990) ressaltou que designers gráficos geralmente utilizam imagens mais abstratas como desenhos ou pinturas, balanceando realismo e eficácia

para a comunicação visual, ressaltando que, algumas vezes, imagens fotográficas podem ser menos eficazes do que imagens estilizadas. É importante observar que o trabalho proposto por Haeberli (1990) tinha como objetivo não apenas gerar pinturas utilizando sistemas computacionais, mas transformar imagens fotorrealísticas (naturais ou sintéticas) em pinturas, a fim de auxiliar a comunicação visual.

Consoantes com Haeberli (1990), outros autores (e.g., Saito e Takahashi, 1990; Seligmann e Feiner, 1991; Winkenbach e Salesin, 1994) afirmaram que técnicas alternativas ao fotorrealismo podem ajudar a compreensão visual, possuindo várias aplicações em desenho industrial e ilustrações anatômicas.

Além da geração de pinturas artísticas, outras vertentes da área de NPR podem ser encontradas na literatura antes dos anos 90, tais como a geração de imagens ao estilo de ilustrações produzidas à mão (e.g., Mezei, Puzin e Conroy, 1974 apud. Landsdown e Schofield, 1995), a geração de caricaturas digitais (e.g., Brennan, 1985), a transformação de imagens em desenhos de linhas (e.g., Parker, 1988), técnicas de *dithering*[2] (e.g., Ulichney, 1988), dentre outras.

Landsdown e Schofield (1995) apresentaram um estudo sobre sistemas de NPR, mostrando a amplitude e diversidade de pesquisas e soluções. Além disto, os autores apontaram que os problemas significativos associados à NPR são principalmente estéticos. O artigo de Landsdown e Schofield (1995) foi um dos pioneiros no uso do termo NPR, apresentando uma visão geral da área e os principais trabalhos realizados pela comunidade científica atuante no âmbito da NPR.

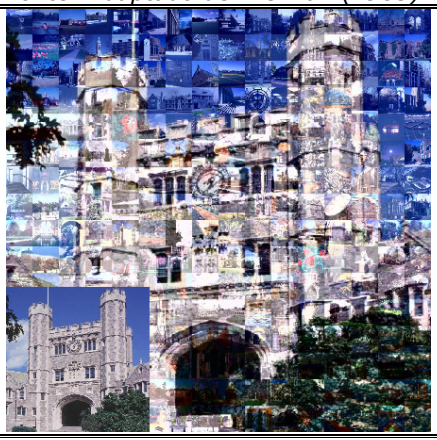
Desde então, constata-se um crescimento na quantidade de artigos publicados, assim como na variedade de problemas estudados, a saber: a geração de imagens que simulam pinturas impressionistas (e.g., Meier, 1996), mosaico (e.g., Finkelstein e Range, 1998), aquarela (e.g., Curtis et

[2] Técnica para criar imagens que causam a "ilusão" de possuírem certas tonalidades de cores ou tons de cinza, para serem utilizadas em dispositivos que não são capazes de exibir estas tonalidades.

al., 1997), quadrinhos (e.g., Kurlander, Skelly e Salesin, 1996), cartuns (e.g., Decaudine, 1996), desenho à mão (e.g., Chen, 2004), dentre outras expressões artísticas culturais (eg., Britto Neto e Carvalho, 2007). O crescente interesse em pesquisas científicas na área de NPR motivou o surgimento de outros marcos históricos: a criação do simpósio *Non-Photorealistic Animation and Rendering* (NPAR) em 2000, realizado em *Annecy* na França, e a publicação de dois livros-texto que abordam o tema, de autoria de Gooch e Gooch (2001) e Strothotte e Schlechtweg (2002). Na Figura 2.1 são exibidos exemplos de resultados obtidos por técnicas NPR.

Geração de caricaturas.
Fonte: Adaptado de Brennan (1985)

Geração de pinturas impressionistas.
Fonte: Adaptado de Haeberli (1990)



Geração de mosaicos.
Fonte: Adaptado de Finkelstein e Range (1998)

Geração de pinturas à óleo.
Fonte: Adaptado de Zeng et al. (2009)

Figura 2.1 – Exemplos de resultados de NPR. As imagens de entrada estão representadas como uma versão reduzida no canto inferior esquerdo de cada imagem.

Além disto, encontram-se também na literatura pesquisas interdisciplinares relacionadas à área de NPR, envolvendo estudos relacionados com aspectos psicológicos (e.g., Halper et al., 2003), atenção visual (e.g., DeCarlo e Santella, 2002), interface homem-máquina (e.g., Medhi, Patabandhige e Toyama, 2007) e teoria da arte (e.g., Hertzmann, 2010).

Pesquisas em NPR produzem imagens estilizadas a partir de ambientes tridimensionais (e.g., Saito e Takahashi, 1990; Seligmann e Feiner, 1991; Decaudine, 1996) e/ou bidimensionais (e.g., Haeberli, 1990; Hertzmann, 1998), utilizando técnicas providas pelas áreas de Computação Gráfica (CG) e Processamento Digital de Imagens (PDI).

O processo de abstração de imagens abordado neste trabalho encontra-se no âmbito de NPR, possuindo como meta a geração de imagens abstratas a partir de imagens bidimensionais e se inserindo no contexto do PDI, conforme ilustrado na Figura 2.2.

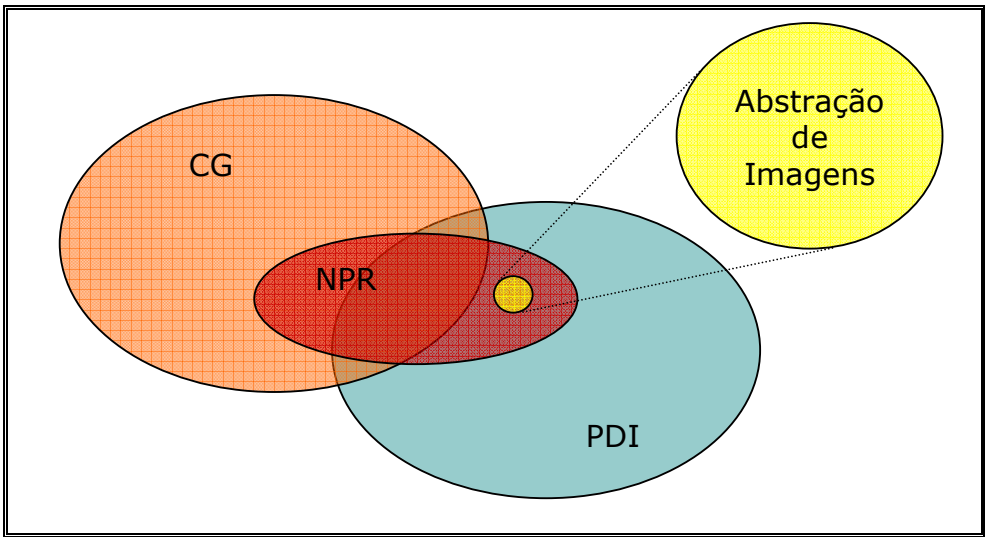


Figura 2.2 – Contextualização da pesquisa perante os principais domínios relacionados à área de NPR.

Dentre os principais problemas existentes na área de NPR, encontram-se a simulação de estilos artísticos e a facilitação da comunicação visual. O tema de abstração de imagens tem sido um tópico

de pesquisa ativo na área de NPR (Kang e Lee, 2008), o qual objetiva a facilitação da comunicação visual, mediante a remoção das informações não relevantes ao entendimento da mensagem, enquanto retém ou enfatiza os elementos mais relevantes.

A seguir, é apresentada uma revisão bibliográfica mais detalhada, contendo as principais abordagens de abstração de imagens digitais.

2.2 Abstração de Imagens

Conforme mencionado anteriormente, uma imagem fotorrealística pode tornar a informação confusa ou até mesmo incompreensível. Por exemplo, na Figura 2.3B, é exibido um excerto do mapa de ruas de Londres, resultante da estilização das feições reais da área. Tal representação se afigura de grande utilidade para o propósito de localização de pontos turísticos, o qual pode não ser atingido a partir da imagem fotorrealística correspondente (Figura 2.3A).

Figura 2.3 – Avenidas de Londres: (A) Fotografia aérea de Londres; (B) Mapa de rodovias e pontos de localização de Londres.

Desta mesma forma, os algoritmos de abstração de imagens possuem como meta remover informações e detalhes menos relevantes ao entendimento da informação, enquanto mantém e reforça os elementos mais relevantes, tornando a mensagem principal mais evidente

e mais facilmente reconhecível. Diferentes métodos automatizados para abstração de imagens têm sido propostos na literatura, dentre os quais podem ser mencionados:

(i) a remoção ou redução de detalhes de textura, iluminação e sombreamento presentes na cena (e.g., Olmos e Kingdom, 2005);

(ii) a simplificação dos formatos dos elementos da cena (e.g., Kang e Lee, 2008);

(iii) a redução do contraste em regiões de baixo contraste e aumento do contraste em regiões de alto contraste (e.g., Winnemöller, Olsen e Gooch, 2006; Kyprianidis e Döllner, 2008); e

(iv) a remoção de elementos não salientes por meio de mapas de saliência (e.g., DeCarlo e Santella, 2002; Zhao et al., 2009).

Na Figura 2.4, ilustra-se um exemplo de abstração de imagem fotorrealística, no qual foram removidos detalhes de textura e grande parte do sombreamento.

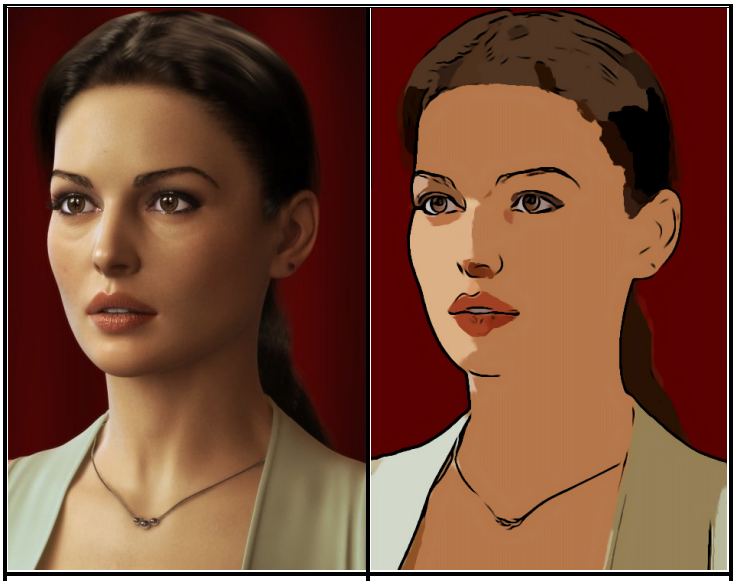


Figura 2.4 – Abstração automatizada de imagem: (A) Imagem original[3]; (B) Imagem abstraída utilizando o método proposto nesta dissertação.

[3] Imagem disponível online em:http://www.anghelescu.net/monica.html

Além disto, algoritmos de NPR (também chamada de renderização artística) objetivam produzir resultados artísticos que sejam visualmente agradáveis ao olho humano (Kang e Lee, 2008; Wen et al., 2006).

O processo de abstração de imagens envolve dois principais subproblemas: extração de linhas (com vistas à obtenção dos contornos dos elementos da imagem); e estilização de regiões (com o propósito de simplificação).

Segundo Gonzalez e Woods (2000), a idéia que fundamenta a maioria das técnicas de detecção de borda é a computação de um operador diferencial local. As técnicas existentes baseiam-se na obtenção de gradientes[4] que determinam o nível de variação da intensidade entre pixels vizinhos (Neoh e Hazanchuk, 2004). A derivada em um ponto qualquer da imagem é obtida a partir da magnitude do gradiente, sendo utilizada para determinar se o pixel correspondente é ou não um pixel de borda.

O cálculo das magnitudes dos vetores de gradiente é feito por meio da aplicação de máscaras (matrizes de pesos) espaciais. Estas máscaras computam, para cada pixel da imagem, o valor da magnitude do gradiente, a partir da soma dos valores dos pixels vizinhos, ponderados pelos valores correspondentes encontrados na máscara. Os operadores de *Roberts*, de *Prewitt*, e de *Sobel* são exemplos de técnicas de detecção de bordas baseadas em gradientes (Neoh e Hazanchuk, 2004). Apesar de serem bastante populares, as técnicas de detecção de bordas baseadas em gradientes são muito sensíveis ao ruído presente nas imagens, além de operarem a partir de uma análise local dos pixels, favorecendo a descontinuidade de bordas (Neoh e Hazanchuk, 2004). Na Figura 2.5A e Figura 2.5B, são exibidas as máscaras do gradiente horizontal (G_x) e vertical (G_y) respectivamente, chamadas de operadores de *Sobel* (Gonzalez e Woods, 2008).

[4] *Gradiente* pode ser definido como um vetor associado a uma coordenada da imagem, o qual aponta para a direção em que os pixels em torno da coordenada sofrem maior variação.

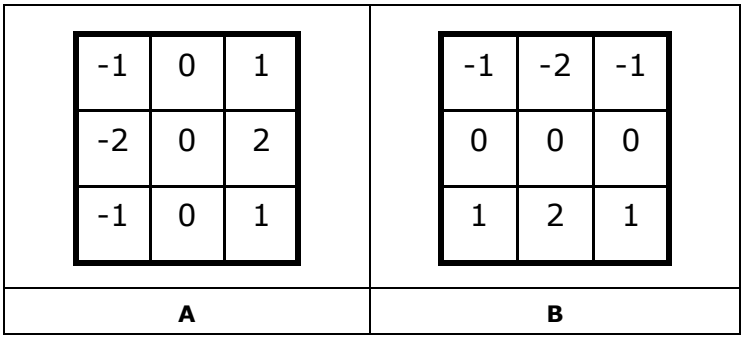


Figura 2.5 – Operadores de *Sobel*: (A) Gradiente X; (B) Gradiente Y.

Embora este fato tenha sido apontado como um dos principais subproblemas do processo de abstração de imagens (Kang, Lee e Chui, 2009), alguns autores preferem não utilizar a extração de linhas como parte de suas abordagens (e.g., Kang e Lee, 2008; Kyprianidis, Kang e Döllner, 2009). Nestes casos, a imagem é abstraída utilizando apenas a estilização das regiões. Desta forma, a estilização das regiões passa a ser apontada por aqueles autores como o principal subproblema no processo abstração de imagens (Kang e Lee, 2008). A partir da estilização de regiões, as cores da imagem são simplificadas, o que geralmente implica a redução da quantidade de cores utilizadas na composição da imagem.

Conforme a literatura da área, os métodos tradicionais destinados à quantização de cores em imagens possuem como objetivo gerar imagens com paletas de cores inferiores às paletas das imagens originais, visando à minimização do erro (diferença entre valores dos pixels da imagem resultante e da imagem original) ocasionado pelo processo de redução de cores (Hautamäki, 2005). Desta forma, estes métodos, em geral, preservam as cores mais freqüentes da imagem, podendo adicionalmente preservar detalhes irrelevantes, tais como o sombreamento de grandes regiões, ao mesmo tempo em que podem remover informações relevantes, tais como o contraste entre regiões menores da imagem.

Como exemplo, considere-se, respectivamente, nas Figura2.6B a 2.6D, os resultados da aplicação dos seguintes algoritmos de redução de cores à imagem da Figura 2.6A: (i) K-Médias (Mikolov, 2007); (ii) mapas auto-organizáveis de Kohonen (Kohonen, 1990); e gás neural

crescente (Holmström, 2002).

A	B	C	D

Figura 2.6 – Quantização de cores (redução para 16 cores): (A) Imagem Original; (B) K-Médias; (C) SOM; (D) GNG.

Percebe-se a perda de detalhes importantes contidos na imagem original (Figura 2.6A), tais como a cor do batom, do cabelo e da pulseira, enquanto são mantidas cores na área corresponde ao segundo plano da imagem, menos relevante ao contexto.

Conforme anteriormente explicitado, o presente trabalho de dissertação tem por objetivo propor uma abordagem para a abstração de imagens digitais, com ênfase no estudo sobre a estilização das regiões que compõem a imagem. Neste sentido, DeCarlo e Santella (2002) descreveram uma estratégia computacional para estilizar e abstrair fotografias, evidenciando estruturas significantes da imagem por meio de desenhos de linhas grossas e grandes regiões de cores constantes. Para produzir regiões uniformemente coloridas, a imagem é submetida a um processo de segmentação, a partir da utilização da filtragem *Mean-Shift* (Comaniciu e Meer, 2002), técnica baseada em um procedimento de reconhecimento de padrões.

Na abordagem proposta por DeCarlo e Santella (2002), os elementos significativos da imagem são identificados a partir de um modelo de percepção humana e de um registro dos movimentos oculares do usuário,

capturados por meio de uma técnica de rastreamento de olhos. Esta técnica baseia-se na hipótese de que os movimentos oculares humanos revelam uma forte evidência da localização do conteúdo significativo de uma imagem. Uma vez identificados os locais considerados mais significativos da imagem, é realizado um ajuste na resolução das regiões (quanto menos importante a região, menor será a sua resolução). Para a obtenção dos contornos da imagem, foi utilizado o detector de bordas *Canny* (Meer e Georgescu, 2001), capaz de contornar problemas associados à descontinuidade de bordas, preservar bordas fracas e suprimir ruído (Neoh e Hazanchuk, 2004).

Contudo, para a localização do conteúdo significativo da imagem, a estratégia de DeCarlo e Santella (2002) necessita de intervenção humana no registro dos movimentos oculares do usuário, o que a caracteriza, portanto, como uma abordagem semi-automática.

Olmos e Kingdom (2005) apresentaram um algoritmo automático para renderização de imagens não fotorrealísticas que se fundamenta na remoção de sombras com base na arquitetura do sistema visual humano. O processo inicia-se por uma conversão do espaço de cores da imagem (RGB para LMS), após o que são computados três canais do sistema visual humano, a saber: (i) iluminação (LUM, do termo inglês *luminance*), (ii) vermelho-verde (RG, do termo inglês *red-green*) e (iii) azul-amarelo (BY, do termo inglês *blue-yellow*), sendo os canais RG e BY correspondentes aos canais cromáticos. Os canais RG e BY são usados para obter um mapa de variações de refletância das superfícies, o qual é posteriormente utilizado para classificar as derivadas dos canais R, G e B da imagem. Por fim, as derivadas classificadas são, então, reintegradas, gerando uma imagem colorida sem sombra, que é sobreposta por contornos de bordas obtidos utilizando operadores de *Sobel*.

Segundo Olmos e Kingdom (2005), os resultados obtidos pelo algoritmo por eles apresentado ressaltam problemas na abordagem proposta, que podem surgir quando ocorrem mudanças significativas no

canal de iluminação da imagem.

Wen et al. (2006) desenvolveram uma nova técnica para a produção de esboços coloridos de imagens imitando o estilo de desenho artístico à mão livre. Nesta técnica, um processo de segmentação interativo é realizado em uma etapa de pré-processamento, na qual o algoritmo *Mean-Shift* (Comaniciu e Meer, 2002) é utilizado para a segmentação da imagem, que é refinada por meio da intervenção do usuário. Após a segmentação, um algoritmo baseado na distribuição da luminosidade nas regiões é utilizado para criar um esboço colorido estilizado, no qual são adicionadas áreas brancas, simulando o brilho das regiões. Por fim, é aplicado um deslocamento nas cores para enfatizar o conteúdo principal e criar uma combinação de cores visualmente agradável. Esta seqüência de etapas é ilustrada na Figura 2.7.

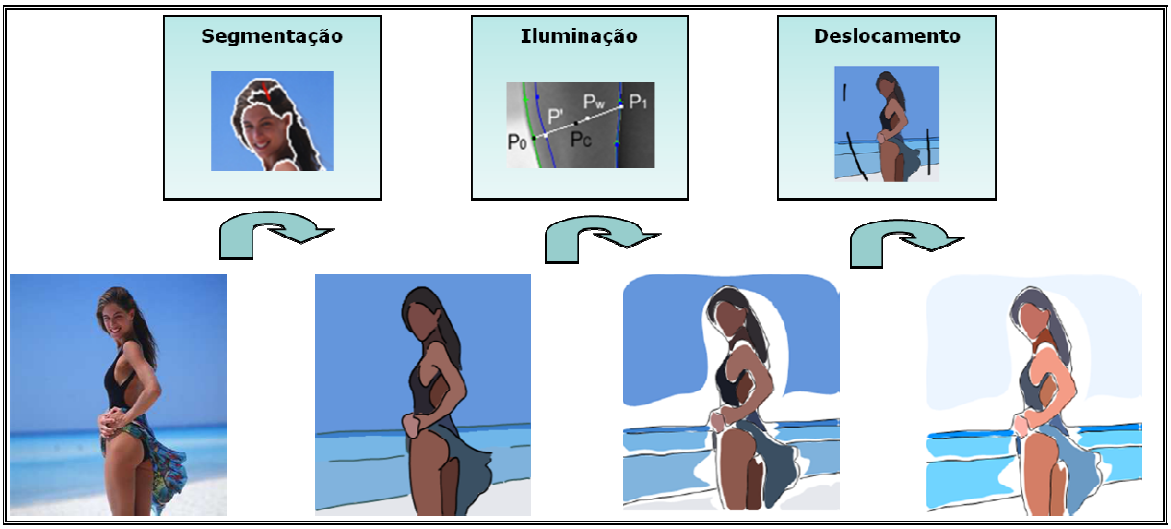


Figura 2.7 – Sequência de passos do método proposto por Wen et al. (2006).

Apesar de demonstrar bons resultados, a técnica proposta por Wen et al. (2006) exige a intervenção do usuário, cujo esforço depende da complexidade da imagem, i.e., quanto mais complexa a imagem, maior a intervenção manual necessária para a obtenção de resultados satisfatórios.

Medhi, Patabandhige e Toyama (2007) apresentaram uma ferramenta

que permite aos usuários a criação de cartuns a partir de imagens fotográficas. Semelhante à técnica proposta por Wen et al. (2006), foi desenvolvida uma interface com o usuário que possibilita refinamentos no processo para produção de melhores resultados, o que inclui correções na etapa de segmentação. Todavia, esta ferramenta disponibiliza ao usuário um modo automático que permite a geração de cartuns sem intervenção humana.

O processo proposto por Wen et al. (2006) inicia-se com a aplicação de um algoritmo de suavização Gaussiana (*Gaussian blurring*), com o objetivo de reduzir a quantidade de componentes obtidos a partir do processo de segmentação. Posteriormente, são extraídas as bordas da imagem, por meio do operador de *Sobel* (Gonzalez e Woods, 2008), utilizando uma máscara alternativa, a qual permite a detecção de bordas a partir de pixels afastados do pixel central. Em seguida, a imagem obtida do processo de suavização é segmentada, utilizando-se uma estratégia de análise de componentes conectados (segmentação baseada na continuidade dos pixels adjacentes). Na etapa seguinte, as regiões resultantes do processo de segmentação são repintadas, sendo identificada uma cor média para cada região, a qual é atribuída aos pixels que a compõem. Por fim, as bordas obtidas no início do processo são sobrepostas à imagem final.

O trabalho de Medhi, Patabandhige e Toyama (2007) teve como alvo serviços de informação (como informações climáticas, oportunidade de emprego, etc.) oferecidos a usuários com baixo nível de alfabetização (analfabetos ou semi-analfabetos). Nota-se que a utilização de um processo automatizado para a abstração a partir da geração de regiões homogêneas na imagem é uma tarefa desejável e alvo de pesquisas não apenas para a área de processamento digital de imagens, mas também para outras áreas da computação (neste caso, a construção de aplicações para o desenvolvimento sócio-econômico de comunidades carentes).

Winnemöller, Olsen e Gooch (2006) desenvolveram um arcabouço

destinado à abstração automática de imagens. Esta ferramenta atua reduzindo o contraste de regiões de baixo contraste a partir da aplicação de um filtro *Bilateral* (Tomasi e Manduchi, 1998), técnica semelhante à *difusão anisotrópica* (Perona e Malik, 1990), ao mesmo tempo em que eleva o contraste nas regiões de mais alto contraste, por meio um processo de extração de bordas, utilizando *Diferença de Gaussianas* (Marr e Hildreth, 1980). A quantização de cores é sugerida como uma etapa adicional, possibilitando a transformação do resultado obtido em imagens ao estilo de cartuns. Neste sentido, foi proposta uma equação que gera uma imagem pseudoquantizada, reduzindo os níveis de luminosidade da imagem.

É sugerido, por fim, a aplicação de uma técnica de pós-processamento para a correção de pequenas falhas nas bordas e para o aguçamento do resultado final. De forma geral, os principais passos da abordagem proposta por Winnemöller, Olsen e Gooch (2006) são ilustrados na Figura 2.8.

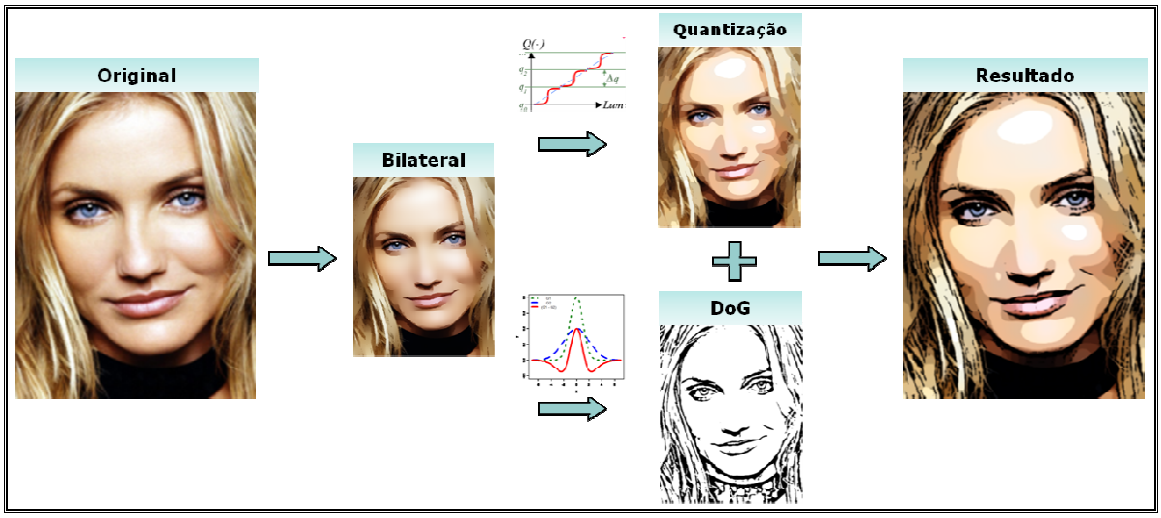


Figura 2.8 – Abordagem proposta por Winnemöller, Olsen e Gooch (2006).

Papari, Petkov e Campisi (2007) apresentaram uma técnica para remoção de detalhes de textura utilizando um filtro de suavização. Nesta técnica, a imagem é simplificada, a partir da utilização de um filtro

Kuwahara, muito similar ao filtro *Bilateral* utilizado por Winnemöller, Olsen e Gooch (2006). Redmond e Dingliana (2007) também propuseram o uso do filtro *Kuwahara*, em conjunto os contornos obtidos pela técnica *DoG*, para a renderização abstrata de cenas 3D. Na abordagem de Redmond e Dingliana (2007), foi utilizada uma segmentação baseada em informações do modelo de dados 3D da cena, a fim de destacar o conteúdo visual importante da imagem.

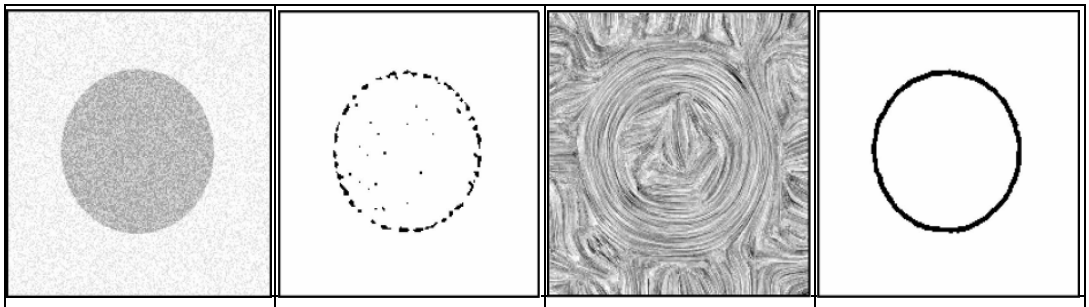
Para Kang e Lee (2008), a utilização de um algoritmo de segmentação de imagem é uma escolha natural para a tarefa de abstração de imagem, reduzindo agressivamente a complexidade da cena, enquanto preserva estruturas importantes. Porém, foi ressaltado que o uso de segmentação normalmente resulta em uma abstração incompleta, visto que as regiões segmentadas continuam apresentando bordas complexas, o que requer iniciativas adicionais de pós-processamento. Além disto, foi argumentado que abordagens como aquela proposta por Winnemöller, Olsen e Gooch (2006) também não simplificam diretamente o formato dos elementos da imagem.

Como solução, Kang e Lee (2008) propuseram a utilização de um campo de direção baseado no fluxo de curvatura médio, juntamente com um filtro de aguçamento de bordas, que objetiva a simplificação das cores das regiões simultaneamente às bordas das regiões. Por fim, foi ressaltado que esta abordagem possui limitações, pois contrai as regiões de formato circular na imagem (por exemplo, os olhos). Para esta situação, foi proposta a incorporação de um detector automático de olhos à abordagem, a fim de contornar o problema.

Kyprianidis e Döllner (2008) e Kang, Lee e Chui (2009) propuseram extensões ao arcabouço de abstração proposto por Winnemöller, Olsen e Gooch (2006). Nestas extensões, os passos de estilização de regiões utilizando o filtro *Bilateral* e de extração de contornos utilizando a Diferença de Gaussianas são melhorados por meio do uso de informações relacionadas ao formato dos elementos que compõem a imagem.

Nos resultados experimentais, observa-se uma melhora significativa na qualidade da detecção dos contornos da imagem, a partir do uso da *Diferença de Gaussianas Baseada em Fluxo* (FDoG, do inglês *Flow-Based DoG*), a qual favorece a geração de contornos bem definidos. Na Figura 2.9, apresenta-se um exemplo comparativo da obtenção de contornos utilizando o DoG (Figura 2.9B) e o FDoG (Figura 2.9D).

As informações relacionadas ao formato dos elementos da imagem são obtidas por meio do cômputo do fluxo dos contornos. Para a representação deste fluxo, Kang, Lee e Chui (2009) geram um campo de vetores (Figura 2.9C) chamado de *Fluxo de Tangente das Bordas* (ETF, do inglês *Edge Tangent Flow*). Em contrapartida, Kyprianidis e Döllner (2008) apresentaram um método para gerar um campo de vetores com características similares ao ETF, que exige menor custo computacional.



A	B	C	D

Fonte: *Adaptado de Kang, Lee e Chui (2009), Flow-Based Image Abstraction*

Figura 2.9 – Detecção de contornos: (A) Imagem Original; (B) DoG; (C) ETF; (D) FDoG.

Zhao et al. (2009) adicionaram ao conjunto de etapas da abordagem proposta por Kang, Lee e Chui (2009) uma técnica destinada à obtenção de um mapa de saliências, o qual é utilizado para identificar as regiões de interesse da imagem. Desta forma, Zhao et al. (2009) controlaram o nível de abstração e de detalhes dos contornos, enfatizando o conteúdo visual nas regiões salientes e removendo detalhes nas regiões não salientes.

Também utilizando informações sobre o formato dos elementos que compõem a imagem, Kyprianidis, Kang e Döllner (2009) propuseram uma

técnica baseada no filtro *Kuwahara*, descrito por Papari, Petkov e Campisi (2007), adaptando-o ao campo de vetores de orientação das bordas proposto por Kyprianidis e Döllner (2008).

Diante das abordagens apresentadas nesta revisão bibliográfica e de resultados apresentados por algumas destas abordagens (vide Apêndice A), nota-se que é possível obter um bom nível de abstração do conteúdo de imagens digitais por meio de transformações em suas paletas de cores. Além disso, percebe-se que as técnicas automáticas apresentadas geram imagens utilizando filtros de suavização.

Em contrapartida, as técnicas que utilizam segmentação sugerem a intervenção do usuário para ajuste do resultado. Segundo Wen et al. (2006), não existe um algoritmo de segmentação automática que produza resultados perfeitos para todas as categorias de imagens.

Desta forma, a presente pesquisa focalizou estudos sobre transformação de cores em imagens, por meio de técnicas de PDI relacionadas a filtros de suavização, técnicas de segmentação e ajuste de cores, visando à proposição de uma abordagem que permita a geração automatizada de imagens abstratas simplificadas a partir de fotografias digitais, que sejam consideradas por observadores humanos mais atrativas do que outras técnicas existentes.

Neste ponto, vale salientar que o objetivo deste trabalho não é a obtenção de uma técnica de segmentação ótima ou do melhor algoritmo para a suavização de imagens, mas sim a obtenção de um arcabouço alternativo para a geração de imagens artísticas simplificadas.

Na próxima subseção, são apresentados estudos relacionados a técnicas de suavização, segmentação e ajuste de cores, com ênfase nos algoritmos utilizados pelas técnicas de abstração apresentadas neste capítulo. O objetivo da apresentação destes estudos foi analisar as características, aplicações e limitações destes algoritmos, com o propósito de auxiliar na concepção da abordagem proposta.

2.2.1 Suavização com Preservação de Bordas

A filtragem é considerada como uma das operações mais fundamentais em processamento de imagens e visão computacional (Tomasi e Manduchi, 1998). Em processamento digital de imagens, as técnicas de filtragem destinam-se, em essência, à suavização de regiões da imagem, ao aguçamento/extração das bordas das regiões que a compõem, ou à extração de características presentes na imagem, normalmente com o propósito de restaurar ou melhorar a qualidade da imagem para a visualização ou para a segmentação e classificação de padrões.

Estas técnicas podem ser classificadas com relação ao domínio em que atuam (filtragem no domínio da freqüência ou no domínio espacial), com relação à natureza da operação aplicada ao sinal (linear ou não linear) e/ou com relação ao propósito com que são aplicadas (suavização, aguçamento/extração de bordas).

No domínio da freqüência, os filtros de suavização (denominados passa-baixas) agem atenuando ou eliminando componentes de alta freqüência que caracterizam bordas e outros detalhes finos nas imagens. A utilização de filtros no domínio da frequência é uma excelente alternativa para tratamento de problemas com propriedades ressaltadas numa representação espectral, como a remoção de ruído periódico. Em outras situações, como em imagens afetadas apenas por ruído aleatório, a utilização de uma técnica de filtragem espacial pode ser uma melhor escolha, consumir menos memória e menor tempo computacional (Inoué e Spring, 1997; Gonzalez e Woods, 2008).

A abordagem básica de um filtro linear espacial, independente do tipo do filtro usado é atribuir, como resposta a um determinado pixel, o somatório de produtos entre as intensidades dos pixels vizinhos e os coeficientes de uma máscara (Gonzalez e Woods, 2008). Apesar dos filtros espaciais não lineares também operarem em vizinhanças, a resposta destes filtros dependem dos valores dos pixels na vizinhança considerada, não utilizando os mesmos coeficientes aplicados a todos os pixels da

imagem.

Filtros de suavização são normalmente utilizados em operações de pré-processamento, para a remoção de pequenos detalhes nas imagens (Gonzalez e Woods, 2008). A estrutura mais simples de um filtro de suavização espacial é a atribuição do valor 1/N para todos os coeficientes da máscara (onde N é o número de elementos da máscara), que tem como resposta a operação de média dos valores dos pixels da vizinhança (Gonzalez e Woods, 2008).

De forma similar ao filtro de média supramencionado, pode-se utilizar uma média ponderada para reduzir o ofuscamento no processo de suavização (Gonzalez e Woods, 2008), como é o caso do filtro Gaussiano. Neste filtro, os valores dos coeficientes da máscara decrescem conforme aumenta a distância (espacial) ao pixel central, partindo do princípio de que as intensidades dos pixels tipicamente variam lentamente no domínio espacial, onde o pixel central possui intensidade similar à de seus vizinhos mais próximos (Tomasi e Manduchi, 1998).

No filtro Gaussiano, os pesos decrescem em uma proporção similar ao comportamento da curva de *Gauss* descrita pela Equação 2.1, na qual d representa a distância ao ponto médio da curva. A intensidade da suavização obtida por este filtro pode ser determinada a partir do desvio padrão (σ) (vide Equação 2.1).

$$G(d) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left(-\frac{d^2}{2\sigma^2}\right)} \qquad (2.1)$$

A distribuição da curva de *Gauss* com média na posição (0,0) e $\sigma = 1$ é ilustrada graficamente na Figura 2.10.

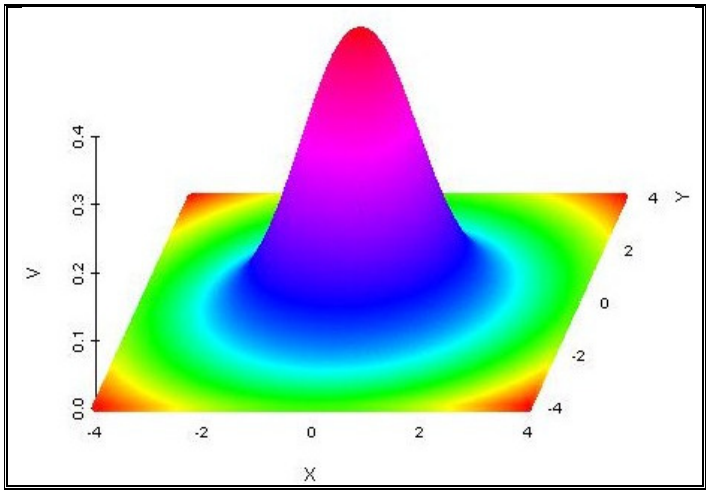


Figura 2.10 – Curva de *Gauss*.

Uma das principais dificuldades de métodos destinados à suavização de imagens é remover ruído ou pequenos detalhes preservando outros detalhes relevantes, e.g., as bordas que estruturam os elementos da imagem. Segundo Tomasi e Manduchi (1998), a premissa de que a intensidade dos pixels varia lentamente falha nas regiões das bordas, as quais são igualmente suavizadas.

Além do filtro Gaussiano, os processos de filtragem da mediana e da moda são outras técnicas utilizadas para a suavização de imagens. Nestes filtros, a resposta para um determinado pixel é obtido, respectivamente, mediante a obtenção da mediana ou da moda dos níveis de cinza dos pixels da vizinhança, o que os caracteriza como filtros não-lineares. Segundo Gonzalez e Woods (2008), o filtro da mediana é bastante popular por prover uma excelente capacidade de redução de ruído, suavizando consideravelmente menos as bordas da imagem do que os filtros de suavização linear que utilizam máscaras de tamanho similar.

Porém, devido a sua natureza, o filtro da mediana degrada detalhes finos (como linhas finas e curvas agudas) na imagem quando se aumenta as dimensões da máscara de filtragem, conforme ilustrado na Figura 2.11. Além disto, a extensão do conceito do filtro da mediana para imagens coloridas não é trivial, por não existir uma forma natural de ordenação para os dados tridimensionais (Pitas e Tsakalides, 1991; Dinet e Robert-

Inacio, 2007).

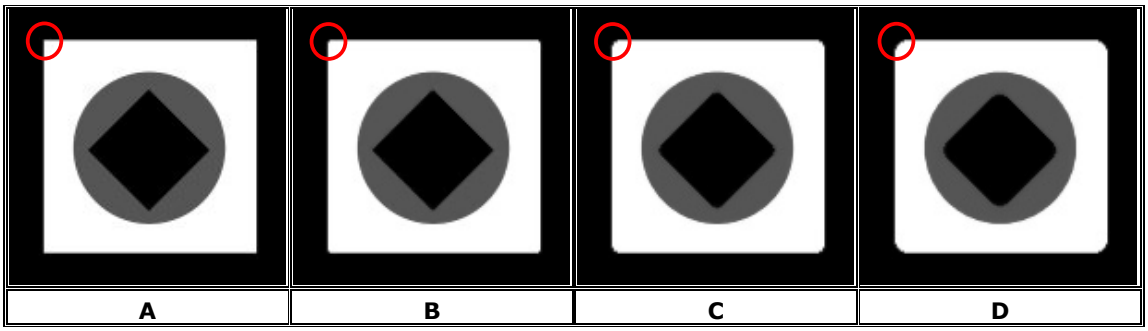


Figura 2.11 – Filtragem por mediana aplicado a diferentes vizinhanças: (A) Imagem original; (B) Vizinhança 3x3; (C) Vizinhança 7x7 (D) Vizinhança 11x11.

Smith e Brady (1997), assim como Tomasi e Manduchi (1998), propuseram um novo conceito de filtragem para a suavização com preservação de bordas. Apesar de terem sido desenvolvidos independentemente e nomeados como filtro *SUSAN* (Smith e Brady ,1997) e filtro *Bilateral* (Tomasi e Manduchi, 1998), em essência fundamentam-se na mesma idéia (Zhang e Allebach, 2008). A idéia da filtragem *Bilateral* é utilizar a distância espacial e radial entre os pixels, combinadas através de duas funções de ponderação (f e g), para determinar o peso atribuído a cada pixel da vizinhança. A função de ponderação tipicamente utilizada tanto para f quanto para g é a função gaussiana (Weiss, 2006).

Seja s um pixel qualquer da imagem de entrada I e Ω o conjunto dos pixels da sua vizinhança (incluindo s), o resultado do filtro *Bilateral* para s é dado pela equação:

$$J_S = k_S^{-1} \sum_{p \in \Omega} f(p-s) g(I_p - I_s) I_p \qquad (2.2)$$

na qual J_S é o resultado do filtro na imagem de saída J para o pixel s, $p-s$ é a distância espacial entre os pixels p e s, I_p e I_s os valores correspondentes dos pixels p e s na imagem I, e k_S é a constante de normalização definida pela Equação (2.3).

$$k_S = \sum_{p \in \Omega} f(p - s) g(I_p - I_s) \qquad (2.3)$$

A atribuição de pesos de baixa magnitude aos pixels cujos valores de intensidades são muito diferentes ao valor do pixel central evita, durante a filtragem nas regiões de borda, que um pixel seja fortemente influenciado por outros pixels vizinhos que pertencem ao outro lado da borda. Diferente do filtro da mediana, o conceito do filtro *Bilateral* pode ser facilmente aplicado a imagens coloridas, utilizando a distância euclidiana para o cômputo de distância entre valores multidimensionais.

Além disto, Tomasi e Manduchi (1998) observaram que a aplicação de múltiplas iterações do filtro *Bilateral* pode ser útil em circunstâncias que a imagem resultante deva apresentar a aparência de cartum (ver Figura 2.12). Algumas técnicas de abstração de imagens se baseiam no uso do filtro *Bilateral*, tal como a abordagem proposta por Winnemöller, Olsen e Gooch (2006).

Adaptado de Tomasi e Manduchi (1998)

Figura 2.12 – Iterações do filtro *Bilateral*: (A)Imagem original; (B) Imagem filtrada com 1 iteração; (C) Imagem filtrada com 5 iterações.

Recentemente, vários estudos têm sido feitos para propor melhorias ao filtro *Bilateral*, tendo sido propostas soluções para (i) reduzir a

complexidade computacional e o tempo de execução do algoritmo (e.g., Pham e Vliet, 2005; Porikli, 2008); (ii) melhorar a eficácia na remoção do ruído (e.g., Jain, 2009); (iii) gerar imagens mais nítidas (e.g., Zhang e Allebach, 2008); e também (iv) melhorar o processo de geração de cartuns (e.g., Kang, Lee e Chui, 2009; Kyprianidis e Döllner, 2008).

Além do filtro *Bilateral*, o filtro de suavização *Kuwahara* é um filtro que preserva as bordas da imagem e também vem sendo estudado e utilizado na área de NPR, por ser capaz de gerar imagens semelhantes a pinturas artísticas (Papari, Petkov e Campisi, 2007). Neste filtro, a resposta para cada pixel na imagem é obtida por meio do cálculo da média e do desvio padrão.

O método utilizado consiste em agrupar os vizinhos de um determinado pixel de acordo com a posição espacial que ocupam e computar, para cada grupo, o desvio padrão das intensidades dos pixels que o compõe. Em sua forma original, o filtro *Kuwahara* agrupa os vizinhos de um pixel em quatro grupos (Redmond e Dingliana, 2007), conforme ilustrado na Figura 2.13.

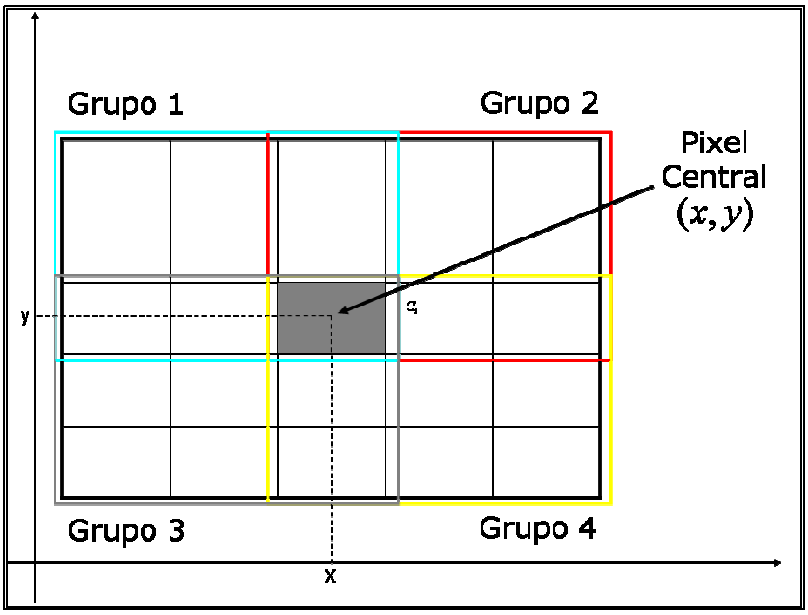


Figura 2.13 – Agrupamentos do filtro *Kuwahara* para um pixel localizado na posição (x,y) da imagem. Os retângulos de diferentes cores representam os diferentes grupos de pixels utilizados pelo filtro Kuwahara.

Então, é atribuído como resultado à imagem de saída o valor da média das intensidades dos pixels do grupo com menor desvio padrão. Em Papari, Petkov e Campisi (2007), é sugerida a utilização dos pixels localizados em uma região circular ao pixel central, a divisão desta área em N grupos diferentes (em contraste ao numero quatro fixo da técnica original) e a substituição do cálculo das médias das regiões por médias ponderadas – o autor utiliza uma função gaussiana – relativas à distância ao pixel central.

Comaniciu e Meer (2002) propuseram um método para a suavização de imagens com preservação de bordas que utiliza uma técnica computacional de reconhecimento de padrões denominada *Mean Shift*, destinada à detecção de máximos locais (modas) de uma função densidade de probabilidade (f.d.p.).

Semelhante ao filtro *Bilateral*, o procedimento adotado utiliza o cálculo de pesos obtidos por meio da distância espacial e radial. Entretanto, ao invés de obter como resposta a média ponderada dos valores da vizinhança, os pesos obtidos neste procedimento são utilizados para obtenção da localização espacial de um novo centro. Em seguida, a janela (matriz cujo centro é inicialmente alinhado ao pixel central) da vizinhança é deslocada em direção ao novo centro, sendo calculados os pesos relacionados à nova vizinhança. Este processo se repete em várias iterações, até convergir (quando o deslocamento da janela é nulo). O resultado do filtro é o valor da média ponderada obtida na janela da última iteração.

Desta forma, este método permite que pixels localizados além da vizinhança do pixel central exerçam influência durante o processo de suavização. Em contrapartida, no filtro *Bilateral* a janela é estática, fixada sobre o pixel central, não permitindo a influência de demais pixels da imagem (Comaniciu e Meer, 2002). O comportamento do filtro baseado no *Mean Shift* é ilustrado na Figura 2.14.

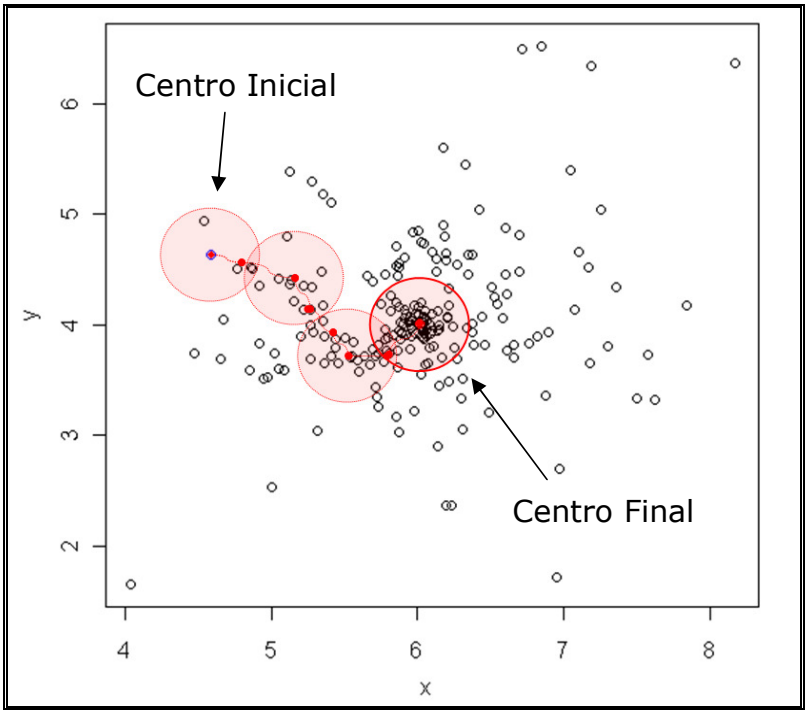


Figura 2.14 – Exemplo da Filtragem *Mean Shift* com indicação das coordenadas (x,y), dos centros e das janelas resultantes (em vermelho).

Vários outros métodos foram propostos na literatura para a suavização com preservação de bordas (Perona e Malik, 1990; Saint-Marc, Chen, e Medioni, 1991; Ramponi, 1995; Donitson, 1996), dentre os quais destaca-se, por sua popularidade, a difusão anisotrópica (Perona e Malik, 1990). A técnica de difusão anisotrópica produz resultados semelhantes aqueles produzidos pelo filtro *Bilateral*. Todavia, trata-se de um método iterativo que pode apresentar problemas relacionados à estabilidade, eficiência e critério de parada (Tomasi e Manduchi, 1998; Comaniciu e Meer, 2002; Barash, 2000).

Na Figura 2.15, são exibidos os resultados obtidos na aplicação de alguns dos filtros de suavização relatados nesta seção: filtro *Gaussiano* (Figura 2.15B), filtro *da mediana* (Figura 2.15C), filtro *Bilateral* (Figura 2.15D), filtro *Kuwahara* (Figura 2.15E) e filtro *Mean Shift* (Figura 2.15F).

Figura 2.15 – Exemplos de resultados de filtros de suavização: (A) Imagem original; (B) Filtro *Gaussiano* (com janela 7x7 e sigma 1.5); (C) Filtro da mediana (com janela 5x5); (D) Filtragem *Bilateral* (com janela 7x7, sigma espacial 5 e sigma radial 12); (E) Filtro *Kuwahara* (com grupos de tamanho 4x4); (F) Filtro *Mean Shift* (com sigma espacial 10 e sigma radial 8).

É possível perceber, a partir dos resultados apresentados na Figura 2.15, a deficiência do filtro Gaussiano para conter a suavização excessiva das bordas mesmo quando empregado um sigma baixo. Além disto, também pode ser observada a geração de artefatos nas cores da imagem

resultante da aplicação do filtro da mediana (notar alteração nas cores dos livros localizados na parte superior direita da estante, na Figura 2.15C), o qual é aplicado separadamente a cada canal de cor da imagem.

De forma geral, os filtros de suavização são abordagens propostas para a remoção de ruído e a obtenção da imagem fotorrealística sem ruído. Desta forma, para obtenção de imagens não fotorrealísticas, normalmente necessita-se a aplicação de várias iterações e/ou a utilização de grandes máscaras. Na Figura 2.16 são exibidos exemplos de resultados utilizando um maior grau de suavização para os filtros *Bilateral* (Figura 2.16B), *Kuwahara* (Figura 2.16C) e *Mean Shift* (Figura 2.16D), tendo por objetivo a obtenção do efeito de NPR.

Figura 2.16 – Exemplos de resultados NPR de filtros de suavização: (A) Imagem original; (B) 5 iterações da filtragem bilateral (com janela 5x5, sigma espacial 10 e sigma radial 12); (C) 2 iterações do filtro *Kuwahara* (com grupos de tamanho 3x3); (D) 1 iterações do filtro *Mean Shift* (com sigma espacial 10 e sigma radial 20).

O Filtro *Kuwahara* perde suas características de preservação de bordas quando utiliza grupos de tamanho iguais ou superiores a 5x5 pixels (fato observado por Redmond e Dingliana, 2007). Isto também ocorre ao ser aplicado em várias iterações (Figura 2.16C). Recentes alterações feitas na estratégia do filtro *Kuwahara* (Papari, Petkov e Campisi, 2007; Kyprianidis, Kang e Döllner, 2009) mostraram excelentes resultados para a renderização ao estilo de pinturas artísticas. O filtro *Bilateral* e o filtro *Mean Shift* apresentam bons resultados na preservação das bordas durante a suavização. Porém, apesar de possuírem estratégias que se

assemelham quanto ao domínio em que atuam (ponderando valores de pixels através da distância espacial e radial), a técnica *Mean Shift* favorece o aparecimento de transições de cores abruptas nas regiões, quando o deslocamento da janela de pixels vizinhos segue em direção a diferentes centros.

Com base nestas observações, assim como em resultados documentados na literatura (Winnemöller, Olsen e Gooch ,2006; Kyprianidis e Döllner, 2008; Kang, Lee e Chui ,2009; Zhao et al., 2009), foi conduzido um estudo mais aprofundado sobre o filtro *Bilateral*, para incorporá-lo à abordagem proposta.

2.2.2 Segmentação de Imagens

Para gerar regiões homogêneas em um processo de quantização de cores de imagens, com fins à renderização não fotorrealística, usualmente são empregadas técnicas de segmentação. A partir da obtenção de regiões bem definidas na imagem, pode-se repintá-las, utilizando cores representativas para a região, a fim de gerar uma representação NPR, conforme ilustrado na Figura 2.17.

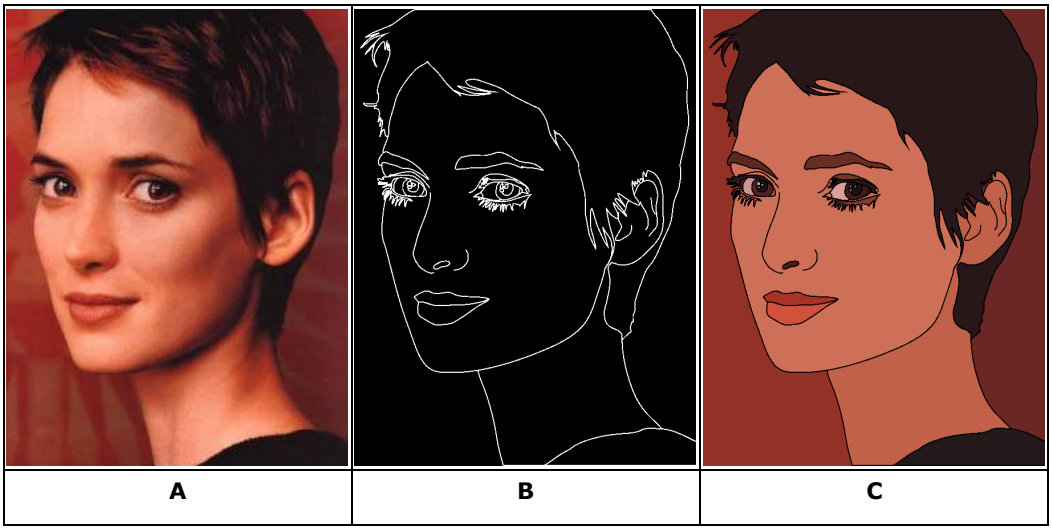


Figura 2.17 – Redução de cores: (A) Imagem original; (B) Bordas de regiões obtidas a partir de um processo de segmentação manual (cortesia do professor José Eustáquio Rangel de Queiroz); (C) Regiões repintadas com cores representativas.

A segmentação de imagens é uma tarefa que possibilita a identificação de regiões de interesse e de objetos em uma imagem (Deng, Manjunath e Shin, 1999). Uma vez identificadas as regiões de uma imagem, a homogeneização das cores destas regiões possibilita a remoção de ruídos e detalhes e, por conseguinte, a abstração de seu conteúdo. Apesar do contínuo crescimento desta área, a tarefa de se obter uma segmentação precisa para uma imagem qualquer ainda não está solucionada (Estrada e Jepson, 2005).

A abordagem proposta por Medhi, Patabandhige e Toyama (2007) utiliza para a abstração das imagens uma segmentação baseada na rotulagem dos componentes conectados, conceito bastante utilizado por aplicações de PDI (Gonzalez e Woods, 2008). Esta técnica consiste em agrupar pixels de forma que para qualquer par de elementos deste grupo exista um caminho ligando estes dois elementos, formado por pixels também pertencentes a este grupo, e que a diferença entre dois pixels adjacentes neste caminho não ultrapassa um determinado limiar. Na Figura 2.18, é ilustrado um exemplo de segmentação utilizando a rotulagem por componentes conectados, obtido a partir de um limiar igual a cinco.

1	2	2	3	29	31	30
1	1	2	3	26	30	33
1	2	2	7	25	27	28
1	3	5	6	13	14	15
3	3	6	4	14	13	17
10	6	16	15	15	16	17
8	16	13	12	14	15	15

A

1	1	1	1	2	2	2
1	1	1	1	2	2	2
1	1	1	1	2	2	2
1	1	1	1	3	3	3
1	1	1	1	3	3	3
1	1	3	3	3	3	3
1	3	3	3	3	3	3

B

Figura 2.18 – Componentes Conectados: (A) Matrix de valores dos pixels; (B) Resultado da rotulagem dos componentes conectados.

Medhi, Patabandhige e Toyama (2007) mencionam que, para a

maioria das imagens, o resultado obtido neste processo de segmentação deve ser corrigido manualmente, problema que se agrava quando as imagens contêm muito detalhes de textura.

Alguns algoritmos de segmentação apresentam resultados satisfatórios para uma grande variedade de imagens, dentre os quais se destacam pela popularidade, assim como pelo desempenho, precisão e pela disponibilidade de implementações públicas, os algoritmos *Mean Shift* (Comaniciu e Meer, 2002) e *Normalized Cuts* (Shi e Malik, 2000) (Estrada e Jepson, 2005; Arbeláez et al., 2009).

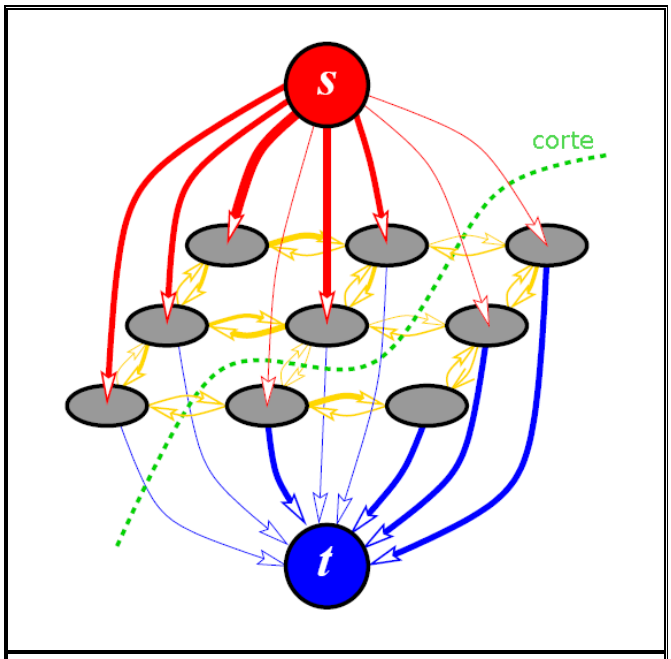
A técnica *Mean Shift* é uma abordagem de agrupamento (*clusterização*) que pode ser utilizada no desenvolvimento de algoritmos de visão computacional para uma grande variedade de tarefas, dentre as quais a suavização e a segmentação (Comaniciu e Meer, 2002). O procedimento de segmentação de imagens baseado na técnica *Mean Shift* é uma extensão direta do filtro de suavização *Mean Shift*, mencionado anteriormente neste trabalho.

Durante a aplicação do filtro de suavização são armazenadas informações sobre as modas (região densas), as quais são agrupadas em clusters de acordo com suas proximidades no domínio espacial e radial. Por fim, cada pixel da imagem é associado ao *cluster* que contém a moda à qual o pixel convergiu durante o processo de suavização.

Outro algoritmo popular na literatura, fundamentado em uma técnica de corte de grafos para garantir uma solução ótima global de forma rápida (Boykov e Kolmogorov, 2004), foi proposto por Boykov e Jolly (2001). Todavia, tal técnica é semi-automática, além de segmentar a imagem somente em duas regiões.

Assim como Boykov e Jolly (2001), ambos os algoritmos *Normalized Cuts* (Shi e Malik, 2000) e *SE-MinCut* (Estrada e Jepson, 2005) tratam o problema de agrupamento dos pixels da imagem como um problema de particionamento de grafos. Nestes algoritmos, a imagem é descrita como um grafo, no qual cada pixel é representado por um vértice que é

conectado aos vértices que representam seus vizinhos por arestas. A estas arestas são atribuídos pesos proporcionais à similaridade entre os pixels correspondentes aos vértices. O objetivo destes dois algoritmos é separar os elementos da imagem utilizando cortes no grafo, removendo as arestas que conectam as partes do grafo, de forma que a soma dos pesos das arestas removidas seja a mínima. Na Figura 2.19, é ilustrado um corte em um grafo separando-o em dois grupos (*s* e *t*) distintos.



Fonte: Adaptado de *Boykov e Kolmogorov (2004).*

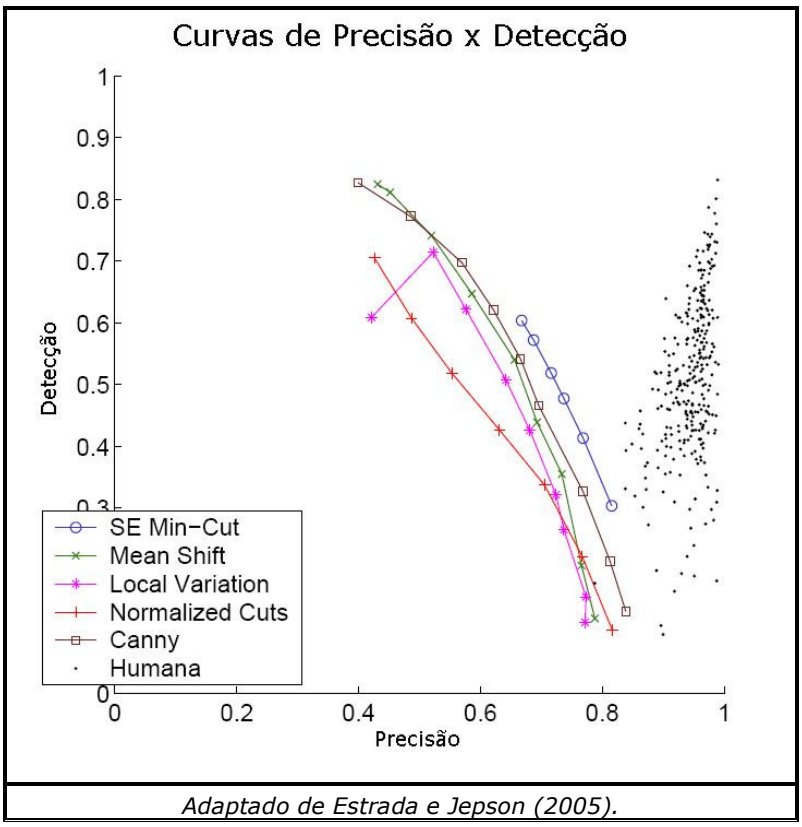
Figura 2.19 – Exemplo de corte no grafo.

Estrada e Jepson (2005) apresentaram resultados comparando algoritmos de segmentação, que incluem *Normalized Cuts* (Shi e Malik, 2000), *SE-MinCut* (Estrada e Jepson, 2005), *Mean Shift* (Comaniciu e Meer, 2002) e *Local Variation* (Felzenszwalb e Huttenlocher, 1998). Nesta comparação foram utilizadas as imagens presentes na base de imagens BSD[5]. Para cada imagem desta base, são obtidos os resultados da segmentação de cada algoritmo e comparados individualmente com o

[5] *Berkeley Segmentation Dataset* (*BSD*) - base de imagens, disponível em: http://www.eecs.berkeley.edu/Research/Projects/CS/vision/grouping/

resultado da segmentação humana (imagens também disponíveis na base BSD). Tendo como as imagens de segmentação humana como a resposta correta e desejada dos algoritmos de segmentação, são obtidos dados relacionados à taxa de precisão (*precision*) e à taxa de detecção (*recall)* das fronteiras entre as regiões detectadas, procedimento utilizado para avaliação do desempenho de algoritmos de segmentação (Estrada e Jepson, 2005).

A taxa de detecção é uma medida que indica a relação entre as bordas corretamente detectadas e as bordas que deveriam ter sido detectadas, enquanto a taxa de precisão indica a relação entre as bordas corretamente detectadas e todas as bordas detectadas pelo algoritmo. O resultado obtido por Estrada e Jepson (2005) é ilustrado na Figura 2.20.



Adaptado de Estrada e Jepson (2005).

Figura 2.20 – Comparação entre algoritmos de segmentação realizada por Estrada e Jepson (2005).

Apesar de obter excelentes taxas de precisão, Estrada e Jepson

(2005) observaram que o algoritmo proposto por eles possui limitações quanto à obtenção de muitos segmentos na imagem, não obtendo taxas altas de detecção, o que provoca a remoção de muitos detalhes da imagem.

Como um dos principais objetivos almejados pelos algoritmos de abstração de imagens é a facilitação da comunicação visual, a obtenção de poucos elementos da imagem pode ocasionar perdas de detalhes importantes, comprometendo consequentemente o propósito de comunicação. Desta forma, a impossibilidade de obter uma maior taxa de detecção (mesmo que ocasione a diminuição da taxa de acerto) pode implicar um efeito inverso àquele desejado (vide Figura 2.21).

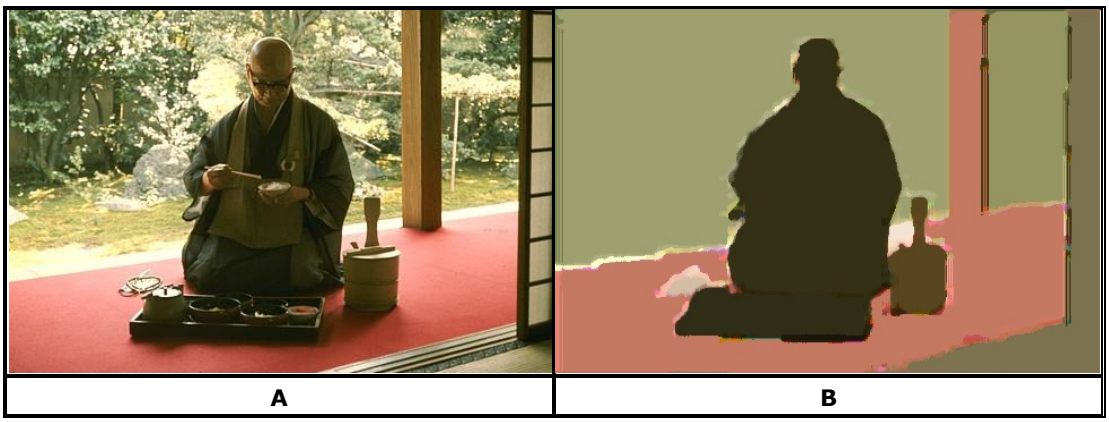


Figura 2.21 – Baixa taxa de detecção: (A) Imagem original (imagem da base BSD); (B) Imagem segmentada (baixo número de segmentos identificados).

Na comparação entre os outros algoritmos, destacam-se os resultados obtidos pelos algoritmos *Canny* (Meer e Georgescu, 2001) e *Mean Shift* (Comaniciu e Meer, 2002), que obtiveram melhores valores de precisão para valores mais altos de detecção (≥ 0.3). Porém, Estrada e Jepson (2005) ressaltaram que a comparação com o algoritmo de *Canny* é injusta, pois este algoritmo é um detector de bordas que não produz uma segmentação da imagem.

O algoritmo de segmentação *Mean Shift* foi utilizado por algumas abordagens de NPR (DeCarlo e Santella, 2002; Wen et al., 2006; Wang et

al., 2004), tendo sido destacada sua eficácia para o processo da abstração de imagens (DeCarlo e Santella, 2002).

2.2.3 Ajuste de Cores

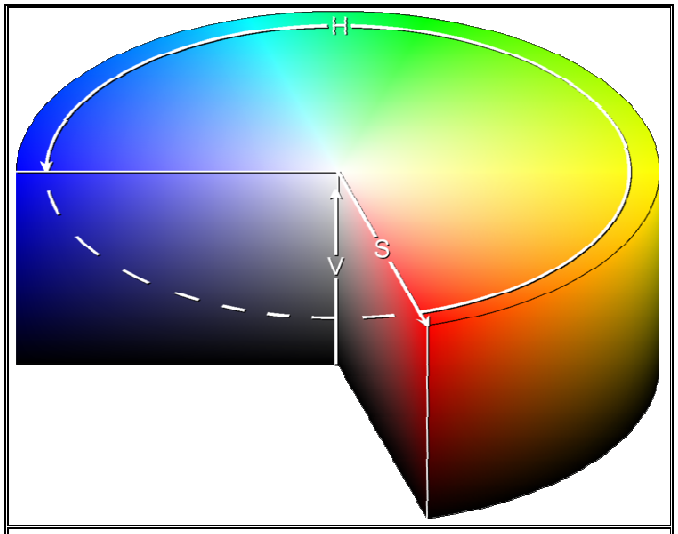
Normalmente, as informações de cores das imagens digitais são descritas por meio do espaço de cores RGB. Neste espaço cromático, as cores são formadas a partir da combinação de diferentes intensidades de três cores primárias: vermelho (R), verde (G) e azul (B). Além deste espaço de cores ser ideal para implementações em *hardware*, corresponde perfeitamente ao fato de que o sistema visual humano é fortemente perceptivo às cores vermelho, verde e azul (Gonzalez e Woods, 2008).

Contudo, o espaço RGB não é adequado para descrever as cores em termos práticos para a interpretação humana (Gonzalez e Woods, 2008). Em contrapartida, com o intuito de capturar noções comuns à descrição de cores humana, Smith[6] (1978b) propôs o modelo de cores HSV, no qual as cores são descritas por três componentes - matiz, saturação e valor, destinadas a capturar conceitos artísticos, tais como tonalidade e sombreamento.

Resumidamente, o matiz (normalmente associado a nomes como verde, azul, amarelo) é o atributo que descreve a cor pura (Gonzalez e Woods, 2008), a saturação descreve o grau em que a cor pura é diluída pela luz branca (Gonzalez e Woods, 2008) e o valor descreve a luminosidade da cor (Schwarz, Cowan, e Beatty, 1987).

As cores no espaço HSV são representadas em um espaço de coordenadas cilíndrico, onde o matiz (H, do termo inglês *hue*) corresponde ao ângulo, a saturação (S, do termo inglês *saturation*) corresponde à distância ao eixo central, e o valor (V, do termo inglês *value*) à altura. Esta representação é ilustrada na Figura 2.22.

[6] Mesmo autor da ferramenta computacional *Paint* (Smith, 1978a), mencionado anteriormente na Seção 2.1.



Fonte: http://daileygraphicdesign.com/Color.html

Figura 2.22 – Espaço cilíndrico HSV.

Segundo Ianeva (2003), dentre as características que distinguem um cartum de uma imagem fotográfica estão a redução de cores e o exagero na saturação. Por meio de análise de histogramas de canais do espaço de cores HSV, Ianeva (2003) mostrou que os cartuns possuem médias de saturação (S) e luminosidade (V) maiores do que aquelas exibidas por imagens fotográficas.

Este comportamento (ilustrado na Figura 2.23) observado por Ianeva (2003) sugere a aplicação de funções de transformação de intensidade aos canais de saturação e valor para obtenção de imagens ao estilo de cartuns a partir de imagens fotográficas.

Existem três tipos básicos de funções, utilizadas freqüentemente para correção de imagens, a saber: lineares, logarítmicas e de potenciação (Gonzalez e Woods, 2008). As funções logarítmicas e de potenciação são funções que mapeiam os níveis de intensidade em um canal de cor para expandir ou comprimir estes valores. Porém, as transformações obtidas por funções de potenciação (conhecidas como transformações *gama*) são mais versáteis a este propósito (Gonzalez e Woods, 2008).

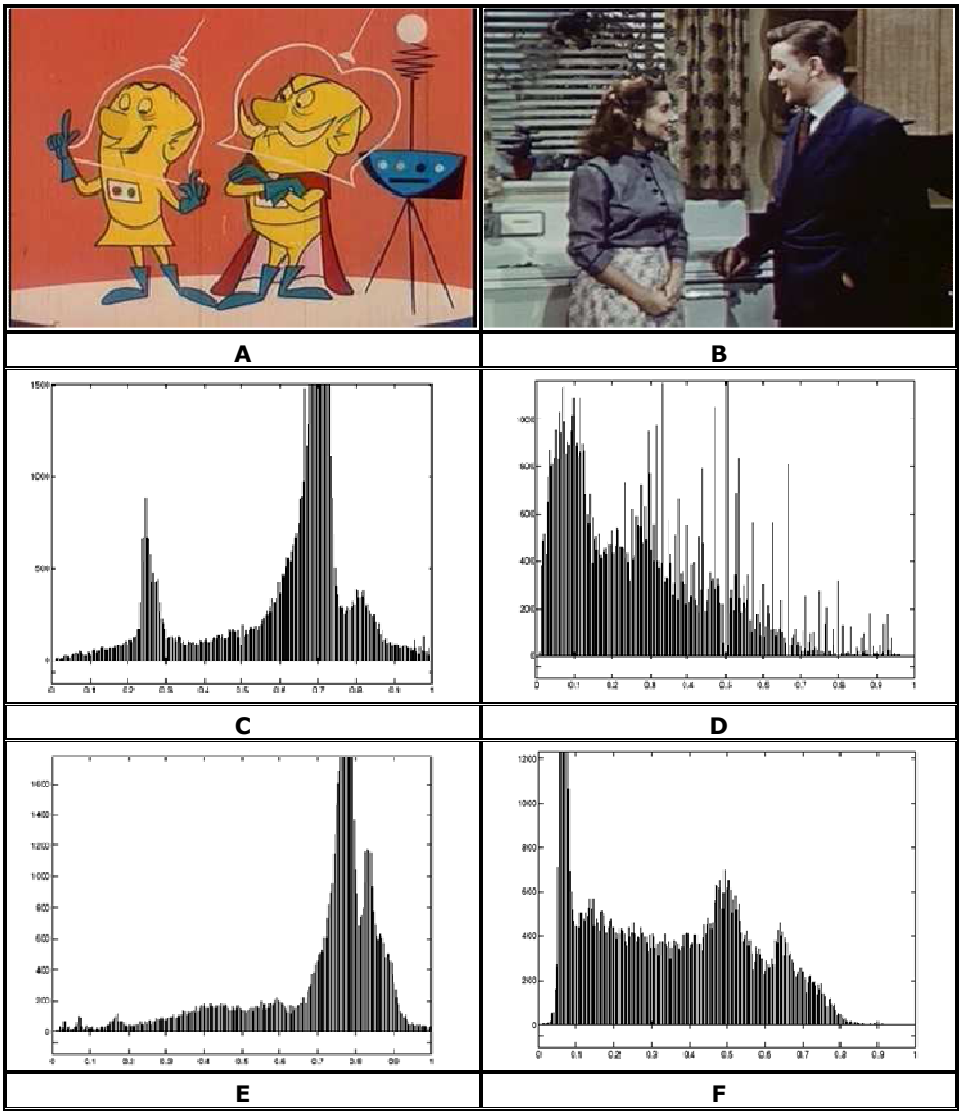


Figura 2.23 – Diferenças entre histogramas (imagens adaptadas de Ianeva, 2003): (A) Cartum; (B) Fotografia; (C e D) Histogramas da saturação do cartum e da foto respectivamente; (E e F) Histogramas da luminosidade do cartum e da foto respectivamente.

Seja r a intensidade de entrada a ser mapeada, s a intensidade de saída da transformação, c e *gama* valores constantes, as funções de transformações *gama* são descritas pela Equação (2.4) (Gonzalez e Woods, 2008).

$$s = cr^{\gamma} \tag{2.4}$$

Para valores de *gama* menores do que 1, a função de transformação

expande os valores de entrada (r) da função; para valores de *gama* maiores do que 1, os valores de entrada são comprimidos (ver Figura 2.24).

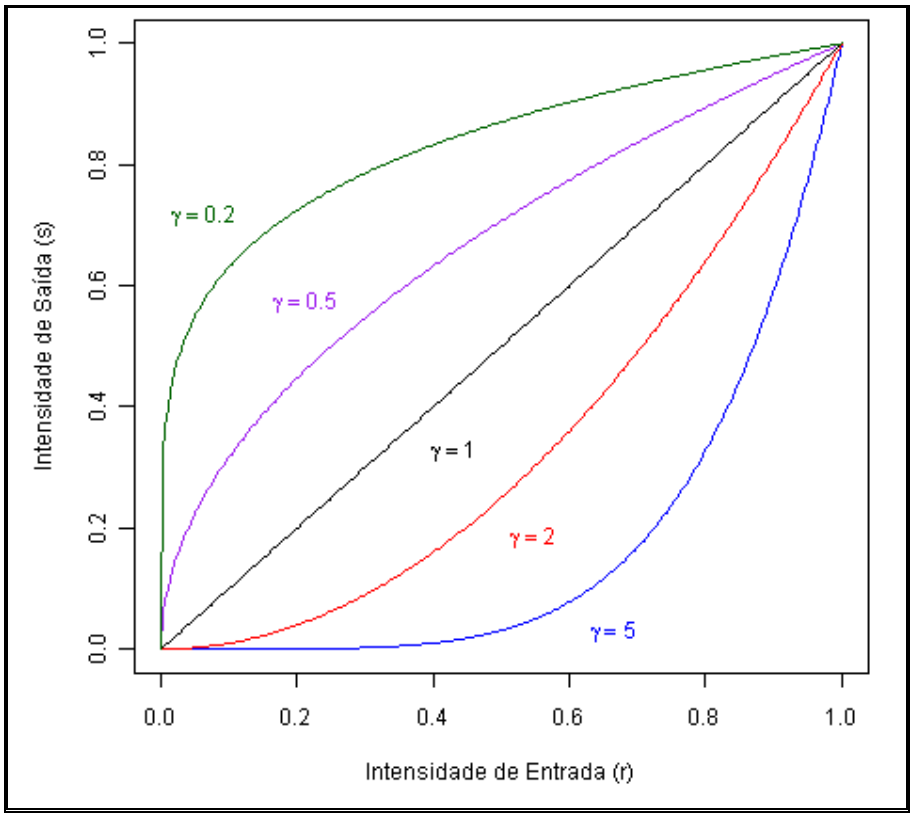


Figura 2.24 - Transformações *gama*.

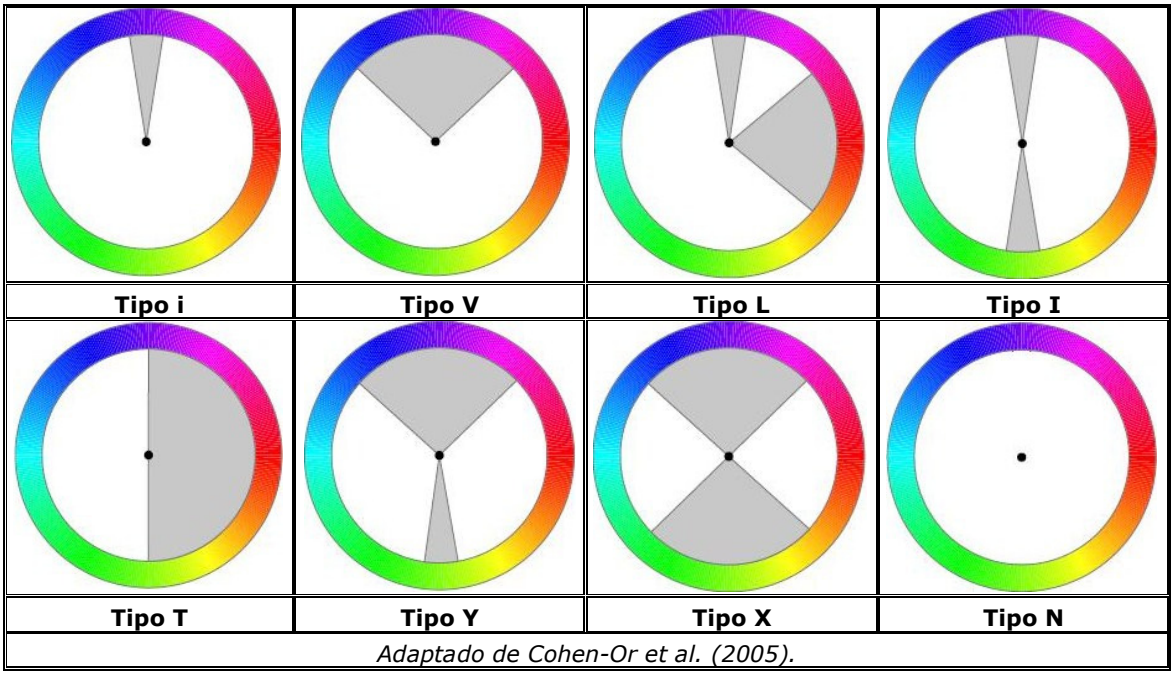
Além dos atributos saturação e valor do espaço HSV, o matiz também pode ser utilizado para imitar estilos artísticos em NPR, além de melhorar o aspecto visual da imagem (Zeng et al., 2009).

Zeng et al. (2009) observaram diferenças relacionadas à temperatura das cores das imagens. Os autores observaram que as cores das pinturas a óleo são, em geral, mais quentes (de acordo com a percepção humana) do que aquelas presentes em imagens naturais. A sensação psicológica de temperatura das cores está diretamente relacionada ao matiz, sendo a cor laranja considerada quente e a cor azul considerada fria (Zeng et al., 2009). O deslocamento dos valores dos matizes dos pixels proporciona mudanças na temperatura das cores, conforme ilustrado na Figura 2.25.

Figura 2.25 – Deslocamento de cores: (A) Imagem original; (B, C e D) Matizes deslocados, respectivamente, de 30º, 190° e 330°.

Tokumaru et al. (2002) desenvolveram um sistema que avalia a harmonia entre as cores da imagem. Com base neste sistema, Cohen-Or et al. (2005) propuseram um método (denominado de harmonização de cores) automático para recolorir uma imagem utilizando um conjunto de cores harmônicas, com o intuito de melhorar o aspecto visual. Neste método, os valores de matizes que compõem a imagem são deslocados para se encaixarem ao melhor esquema harmônico encontrado.

Um esquema harmônico utilizado por Cohen-Or et al. (2005) é baseado em oito *templates* harmônicos (Figura 2.26) definidos para o canal de matizes (H) no espaço de cores HSV.



Adaptado de Cohen-Or et al. (2005).

Figura 2.26 – Templates Harmônicos. Um conjunto de cores que se localizam nos setores (áreas escuras) de um template é considerado harmônico. Os templates podem ser rotacionados para qualquer ângulo.

Um esquema harmônico é definido por um *template* T_m associado a um ângulo α. A função $F(X,(m,\alpha))$ que mede a harmonia de uma imagem X em relação ao esquema (m,α) é definida pela Equação (2.5),

$$F(X,(m,\alpha)) = \sum_{p \in X} \left\| H(p) - E_{T_m(\alpha)}(p) \right\| \cdot S(p) \qquad (2.5)$$

na qual p é um pixel pertencente à imagem X, $H(p)$e $S(p)$ são os valores do matiz e da saturação do pixel p, e $E_{T_m(\alpha)}(p)$ é a borda do setor mais próximo à $H(p)$, referente ao *template* T_m posicionado no ângulo α. O melhor esquema harmônico para uma imagem é definido pelo *template* harmônico associado ao ângulo que minimiza a função F.

Após identificar o melhor *template* e ângulo, o valor do matiz de cada pixel na imagem X é deslocado para um dos setores do *template* de acordo com sua proximidade.

Estes ajustes aplicados aos canais do espaço de cores HSV têm por

objetivo melhorar o aspecto visual do cartum, tendo sido incorporados à abordagem proposta como etapas opcionais de pós-processamento. É conveniente ressaltar que a aplicação destas técnicas altera, por outro lado, as informações das cores das imagens, podendo alterar, por conseguinte, elementos importantes para a comunicação e entendimento da informação. Desta forma, torna-se necessária a intervenção humana no processo.

2.3 Considerações Finais do Capítulo

Neste capítulo, destacou-se que as pesquisas científicas na área de NPR se intensificaram após a década de 90, tendo como um dos principais propósitos a geração de imagens abstratas, com fins a facilitar a comunicação visual. Os trabalhos mais recentes de abstração de imagens baseiam-se na abordagem proposta por Winnemöller, Olsen e Gooch (2006), demonstrando, por meio de seus contínuos avanços, que o tema de abstração de imagens permanece como um tópico ativo na área de NPR.

O arcabouço proposto por Winnemöller, Olsen e Gooch (2006), possui como passos principais a remoção de detalhes da imagem utilizando filtro de suavização, a redução de cores através de uma pseudo-quantização, e a obtenção de contornos dos elementos da imagem. Estes passos são seguidos também por Kang, Lee e Chui (2009), cuja abordagem possui como base o arcabouço de Winnemöller, Olsen e Gooch (2006).

Por outro lado, nas abordagens de DeCarlo e Santella (2002) e Medhi, Patabandhige e Toyama (2007) a imagem é simplificada utilizando algoritmo de segmentação, porém, estas abordagens utilizam como passo inicial algoritmos de suavização para auxiliar o processo de segmentação. Apesar de não estar explícito o filtro de suavização na abordagem de DeCarlo e Santella (2002), a abstração é obtida por meio da segmentação *Mean Shift*, que por sua vez utiliza a suavização *Mean Shift* (explicado

anteriormente). Também faz parte destas abordagens a obtenção dos contornos dos elementos da imagem.

Desta forma, estas abordagens seguem mesmos passos para abstração das imagens, tendo como principal diferença a utilização de diferentes técnicas de PDI para implementação destes passos. Outras abordagens encontradas na literatura (e.g. Kyprianidis e Döllner, 2008; Zhao et al., 2009) seguem estes passos, sumarizados no Quadro 2.1.

Quadro 2.1 – Métodos para abstração de imagens.

Abordagem	Suavização	Redução de Cores	Extração de Contornos
DeCarlo e Santella (2002)	*Mean Shift*	*Mean Shift*	*Canny*
Winnemöller, Olsen e Gooch (2006)	*Bilateral*	Quantização	*DoG*
Medhi, Patabandhige e Toyama (2007)	*Gaussiana*	Rotulagem dos Componentes Conectados	*Sobel*
Kyprianidis e Döllner (2008)	*Bilateral* utilizando fluxo de tangente das bordas	Quantização	*DoG* utilizando fluxo de tangente das bordas
Kang, Lee e Chui (2009)	*Bilateral* utilizando fluxo de tangente das bordas	Quantização	*DoG* utilizando fluxo de tangente das bordas
Zhao et al. (2009)	*Bilateral* utilizando fluxo de tangente das bordas e mapa de saliência	Quantização	*DoG* utilizando fluxo de tangente das bordas e mapa de saliência

Além destes passos, DeCarlo e Santella (2002) finalizam o processo de abstração corrigindo as regiões de bordas geradas. Esta correção é importante para técnicas que utilizam segmentação, para reduzir o nível de detalhes e artefatos presentes nas fronteiras entre regiões (DeCarlo e Santella, 2002; Kang e Lee, 2008; Kang, Lee e Chui, 2009).

A revisão bibliográfica apresentada neste capítulo pôde fornecer informações relevantes para o entendimento do funcionamento de sistemas de abstração de imagens, e do funcionamento das principais técnicas utilizadas (suavização, segmentação e ajuste de cores), necessários para a elaboração do método proposto.

Capítulo 3

Abordagem Proposta

Considerando a revisão realizada no capítulo anterior, definiram-se como núcleo da abordagem para abstração de imagens proposto nesta dissertação dois módulos, responsáveis pela: (i) *simplificação da imagem* (suavização e redução de cores); e (ii) *extração de contornos*. Adicionalmente, como pós-processamento, a fim de melhorar o aspecto visual do cartum gerado, também foram incorporados à abordagem os módulos acessórios de *ajuste de brilho e saturação* e *harmonização de cores*.

Desta forma, o propósito deste capítulo é apresentar a abordagem proposta para a geração de imagens não fotorrealísticas ao estilo de cartuns, a partir de imagens fotográficas digitais. Inicialmente, é fornecida uma visão geral da arquitetura da abordagem, seguida do detalhamento de cada fase do processo, e dos módulos que as compõem.

3.1 Visão Geral da Arquitetura Proposta

A arquitetura da abordagem proposta foi inspirada em abordagens encontradas na revisão bibliográfica, em especial as abordagens propostas por DeCarlo e Santella (2002), Winnemöller, Olsen e Gooch (2006), Medhi, Patabandhige e Toyama (2007) e Kang, Lee e Chui (2009). Na arquitetura descrita neste documento, há duas fases principais:

1. *Abstração da imagem*: nesta fase, a imagem de entrada é

submetida aos módulos de simplificação e extração dos contornos, visando à obtenção da imagem abstrata; e

2. *Ajuste de cores*: nesta fase, a imagem abstrata é submetida aos módulos de *harmonização de cores* e de *ajuste de brilho e saturação*.

O diagrama apresentado na Figura 3.1 contém a visão geral da arquitetura da abordagem proposta.

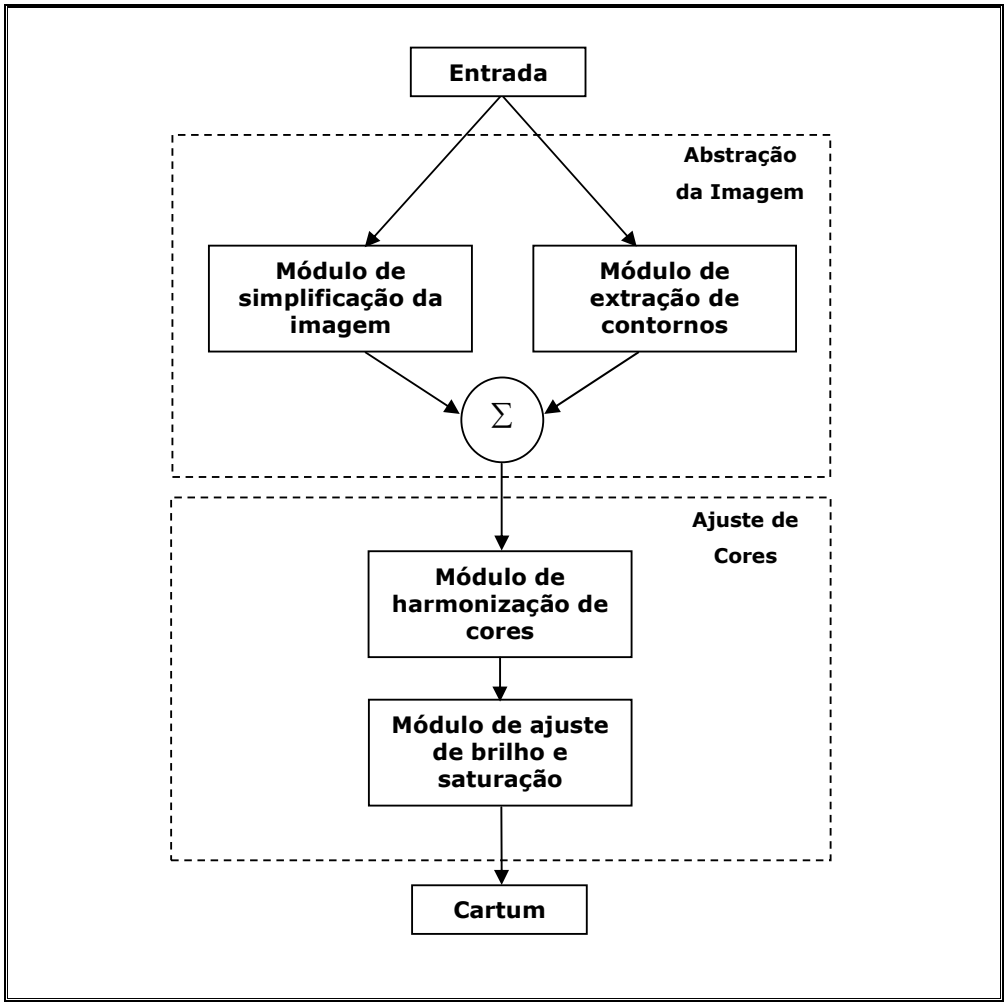


Figura 3.1 – Arquitetura da abordagem proposta.

Para a validação desta abordagem, foi implementada uma aplicação de *software* utilizando e estendendo técnicas presentes no estado da arte de abstração de imagens, detalhadas nas próximas seções deste capítulo.

3.2 Abstração da Imagem

O objetivo da primeira fase da abordagem proposta, também considerado o principal objetivo deste trabalho, é a obtenção da imagem abstrata semelhante a um cartum. Esta fase possui dois módulos, destinados à: (i) *simplificação da imagem*; e (ii) *extração de contornos*.

A fim de simplificar a imagem, os principais trabalhos encontrados na literatura utilizam algoritmos de suavização ou segmentação. A escolha do método a ser utilizado depende do nível de abstração que se deseja alcançar. Os algoritmos de segmentação reduzem a complexidade da cena, de forma mais efetiva do que os filtros de suavização (Kang e Lee, 2008). Contudo, necessitam normalmente de correções não automáticas, devido à dificuldade do processo de segmentação.

Técnicas para a extração de bordas são utilizadas para extrair os contornos dos elementos significantes da imagem, fundamentando-se usualmente na detecção de transições significativas nos valores das intensidades de pixels vizinhos.

Serão apresentados nesta seção, dois métodos diferentes para a simplificação de uma imagem. Um deles utiliza um algoritmo de suavização, enquanto o outro utiliza um algoritmo de segmentação. Por fim, será apresentado um método para a extração dos contornos dos elementos da imagem.

3.2.1 Simplificação por Suavização

Dentre os filtros de suavização pesquisados e apresentados no capítulo anterior, o filtro escolhido para ser codificado e incorporado ao módulo de simplificação da imagem por suavização foi o filtro *Bilateral*.

Devido a sua eficácia e facilidade de uso, o filtro *Bilateral* ganhou muita atenção da comunidade científica nos últimos anos (Kang, Lee e Chui, 2009). Uma variante deste filtro foi originalmente empregada na abstração de imagens por Winnemöller, Olsen e Gooch (2006), sendo adotado em outras pesquisas mais recentes (Kyprianidis e Döllner, 2008;

Kang, Lee e Chui, 2009; Zhao et al., 2009) que utilizam a informação da direção das bordas para reforçar a estilização e a preservação das bordas na imagem abstrata. Nos trabalhos supramencionados, mais especificamente, foi utilizado o filtro FBL. A Figura 3.2 contém um exemplo de aplicação deste filtro com 3 ou 5 iterações.

A	B	C

Adaptado de Kang, Lee e Chui (2009)

Figura 3.2 – Filtro *Bilateral Baseado em Fluxo*: (A) Imagem Original; (B) Filtro *Bilateral* (3 iterações); (C) Filtro *Bilateral Baseado em Fluxo* (5 iterações).

A abordagem proposta para a simplificação da imagem por suavização é uma extensão inédita do FBL, proposta nesta dissertação, na qual se utiliza o caráter adaptativo da filtragem radial proposta por Zhang e Allebach (2008) para manter a nitidez das bordas da imagem após sucessivas iterações. Os funcionamentos do filtro FBL e do filtro proposto por Zhang e Allebach (2008) são descritos no Apêndice B.

Após a suavização, as cores da imagem são reduzidas utilizando o método originalmente proposto por Winnemöller, Olsen e Gooch (2006). Nos próximos parágrafos, são apresentados em maiores detalhes os métodos para a suavização da imagem e a redução da paleta de cores da imagem.

I Filtro Bilateral Adaptativo Baseado em Fluxo

Na abordagem de Zhang e Allebach (2008), o valor do sigma e do centro (média) da gaussiana no domínio radial é adaptado para cada pixel, de

acordo com a resposta do operador *Laplaciano da Gaussiana* (Marr e Hildreth, 1980). Esta alteração tem por objetivo reduzir a influência dos pixels que se encontram na região de transição das bordas, evitando a suavização destas transições, e favorecendo o aumento do contraste entre as bordas. Na Figura 3.3 é exibida uma comparação entre os resultados obtidos pelo filtro *Bilateral* convencional e pelo filtro *Bilateral Adaptativo*.

Figura 3.3 – Filtro *Bilateral Adaptativo*: (A e D) Imagem Original; (B e E) Filtragem *Bilateral* (vizinhança = 9x9; sigma espacial = 20; sigma radial = 10); (C e F) Filtragem *Bilateral Adaptativa* (vizinhança = 9x9; sigma espacial = 20; LoG(sigma) = 1,2; LoG(vizinhança) = 7x7). Dimensões das imagens: (A, B e C) 800x600 pixels; (D, E e F) 270x180 pixels.

Visando adicionar esta característica do filtro *Bilateral Adaptativo* ao

filtro FBL, foi incorporado o cálculo do LoG e substituída a função gaussiana de ponderação no domínio radial pela função proposta por Zhang e Allebach (2008). O resultado do filtro *Bilateral Adaptativo Baseado em Fluxo* (FABL) proposto é ilustrado na Figura 3.4.

Figura 3.4 – Ilustração do Filtro FABL: (A e D) Imagem Original (800x600 px); (B e E) Imagem filtrada pelo filtro FBL (direção da tangente: sigma espacial = 2.0; sigma radial = 10; direção do gradiente: sigma espacial = 0.6; sigma radial = 0.04); (C e F) Imagem filtrada pelo filtro FABL (direção da tangente: sigma espacial = 2.0; sigma radial = 10; direção do gradiente: sigma espacial = 0.6; LoG(sigma) = 1,2; LoG(vizinhança) = 7x7). Dimensões das imagens: (A, B e C) 800x600 pixels; (D, E e F) 270x180 pixels.

A alteração no filtro FBL foi efetuada apenas na filtragem na direção

do gradiente das bordas, visto que é nesta direção que ocorre a transição dos valores entre as duas regiões da borda. Vale salientar que tal alteração também visou evitar a inibição da ação de estilização dos formatos das bordas, a qual necessita de valores altos do sigma da função radial.

Por fim, as funções propostas por Zhang e Allebach (2008) obtidas por treinamento foram substituídas por funções similares e de fácil ajuste, para se obter maior controle do filtro, possibilitando aumentar ou diminuir os valores dos parâmetros sigma, assim como regular a intensidade do deslocamento do centro da gaussiana, aumentando ou diminuindo o aguçamento das bordas.

Desta forma, as funções $\zeta(p)$ e $\sigma^2(p)$ da gaussiana (ver Apêndice B) foram substituídas pelas Equações (3.1) e (3.2) respectivamente.

$$\zeta(p) = m \cdot LoG(p) \tag{3.1}$$

$$\sigma^2(p) = \left(e^{\left(-\frac{(LoG(p))^2}{2\sigma^2 s} \right)} \left(MAX_{sigma} - MIN_{sigma} \right) \right) + MIN_{sigma} \tag{3.2}$$

A partir da Equação (3.1), obtêm-se os valores do deslocamento proporcionalmente aos valores do LoG. Este comportamento foi observado por Zhang e Allebach (2008), na análise das funções obtidas no treinamento. A Equação (3.2) proposta é uma função gaussiana deslocada para uma faixa de valores entre $[MIN_{sigma}, MAX_{sigma}]$ e controlada por $\sigma^2 s$. Embora a função gaussiana não possua o mesmo comportamento da função obtida por Zhang e Allebach (2008), esta foi escolhida com o objetivo de permitir maior suavização nos pixels que não pertencem a regiões de bordas, os quais apresentarão valores mais baixos na resposta do LoG.

Na Figura 3.5 é ilustrado um exemplo de resposta obtida pelas

funções propostas para $m = 0.42$, $MIN_{sigma} = 5$, $MAX_{sigma} = 17$ e $\sigma_s = 16$.

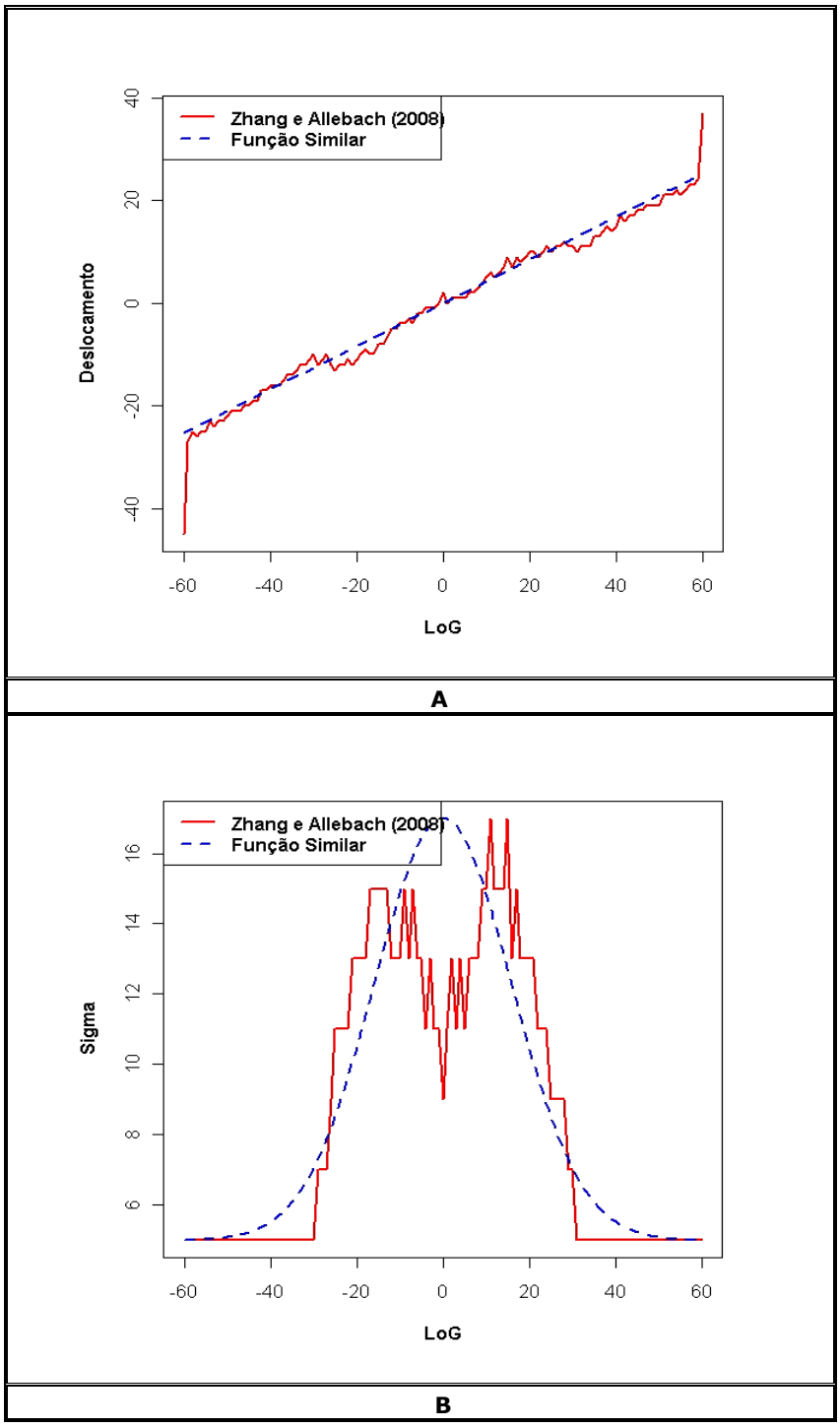


Figura 3.5 – Funções similares às propostas por Zhang e Allebach (2008): (A) Função $\zeta(p)$ para m=0.42; (B) Função $\sigma^2(p)$ para MIN_{sigma}=5, MAX_{sigma}=17 e σ_s=16.

II. Pseudo-Quantização da Iluminação

Após a obtenção da imagem suavizada, obtém-se a imagem simplificada ao estilo de cartuns por meio de um processo de redução de cores. A técnica utilizada por Kyprianidis e Döllner (2008), Kang, Lee e Chui (2009) e Zhao et al. (2009) é aquela originalmente proposta por Winnemöller, Olsen e Gooch (2006): a obtenção da paleta reduzida de cores se dá a partir da aplicação de um processo de pseudo-quantização no canal da iluminação. Um exemplo do resultado deste processo é mostrado na Figura 3.6.

Figura 3.6 – Redução de cores: (A) Imagem Suavizada; (B) Imagem com cores reduzidas.

Este processo de quantização é aplicado ao canal *L** do espaço de cor CIE L*a*b* (CIELAB), no qual *L** corresponde à iluminação e *a** e *b** às componentes cromáticas (as conversões destes espaços cromáticos são descritas no Apêndice C). A função de quantização aplicada ao canal *L** é descrita pela Equação (3.3).

$$Q(x) = q_{próximo} + \frac{\Delta q}{2} \tanh\left(\varphi.\left(f(x) - q_{próximo}\right)\right) \qquad (3.3)$$

Esta função divide o canal *L** em n níveis equidistantes (distância igual a Δq), na qual $f(x)$ é a intensidade do canal *L** para o pixel x e $q_{próximo}$ é o nível mais próximo de $f(x)$. A suavidade das transições entre estes níveis é controlada pelo parâmetro φ.

3.2.2 Simplificação por Segmentação

As principais abordagens de abstração por segmentação utilizam a técnica do *Mean Shift* para a simplificação da imagem. Um exemplo de resultado obtido a partir desta técnica é ilustrado na Figura 3.7.

Figura 3.7 – Segmentação Mean Shift: (A) Imagem original; (B) Imagem segmentada.

É necessário ressaltar que, ao simplificar a imagem utilizando segmentação, produz-se bordas com muitos detalhes entre as regiões obtidas. Para corrigir estas bordas, DeCarlo e Santella (2002) propuseram a suavização das curvas das bordas.

Considerando este fato, na abordagem proposta nesta dissertação foi utilizado o filtro *Bilateral Adaptativo Baseado em Fluxo*, apresentado na seção anterior, com valores altos para os sigmas do filtro aplicado na direção da tangente das bordas. Um exemplo de resultado desta suavização é exibido na Figura 3.8.

Figura 3.8 – Suavização das bordas: (A) Imagem segmentada; (B) Resultado após suavização das bordas.

Apesar de terem sido apresentadas duas abordagens diferentes para simplificação da imagem, apenas uma foi utilizada para a validação do abordagem proposta. Para efeito de comparação com as aplicações de *software* utilizadas na validação, os quais permitem obter diferentes níveis de abstração, a abordagem escolhida foi a de abstração por segmentação.

3.2.3 Extração de Bordas

Após obter a imagem simplificada, as estruturas dos elementos importantes na imagem são evidenciadas sobrepondo as linhas escuras dos contornos destes elementos na imagem simplificada, favorecendo o contraste das bordas.

Para a extração de tais contornos, foi utilizada neste trabalho a técnica FDoG, que favorece a geração de contornos bem definidos a partir das informações do formato dos elementos presentes no ETF. Esta técnica vem sendo comumente adotada em trabalhos mais recentes encontrados na revisão bibliográfica (e.g., Kyprianidis e Döllner, 2008; Kang, Lee e Chui, 2009; Zhao et al., 2009).

Os contornos resultantes da técnica DoG são obtidos pela subtração

do resultado da suavização de duas funções gaussianas com parâmetros sigma diferentes. Este método possui como resultado uma aproximação do método de detecção de bordas de Marr e Hildreth (1980), a partir do qual se obtêm as bordas por meio da detecção do cruzamento do zero no resultado do LoG. No FDoG, os contornos são detectados pela diferença entre os filtros gaussianos aplicados na direção do gradiente (direção que ocorre a transição entre as regiões da bordas). Por fim, os contornos obtidos são suavizados por um filtro gaussiano na direção da tangente das bordas.

Após a obtenção dos contornos, estes são sobrepostos na imagem simplificada multiplicando as intensidades dos pixels da imagem pelo mapa de contornos, conforme ilustrado na Figura 3.9.

Figura 3.9 – Sobreposição dos contornos: (A) Imagem original; (B) Contornos extraídos; (C) Imagem simplificada; (D) Sobreposição dos contornos.

Para reduzir a quantidade de detalhes dos contornos da imagem extraídos nesta etapa, a imagem de entrada é submetida a um pré-processamento de suavização. O processo de abstração da imagem encerra-se ao simplificar a imagem e sobrepor os contornos.

Após a obtenção da imagem abstrata, esta é submetida à fase de ajuste de cores da imagem, a qual foi inserida na abordagem visando agrupar módulos opcionais de pós-processamento, os quais têm por objetivo melhorar o aspecto visual do cartum gerado.

3.3 Ajuste de Cores

A fase de ajuste de cores proposta neste trabalho, é composta por módulos que realizam transformações no brilho, na saturação e na combinação dos matizes das cores da imagem. Por conseguinte, são aplicados ao espaço de cores HSV. As conversões entre os espaços cromáticos HSV e RGB encontram-se no Apêndice C.

A fase de ajuste de cores é composta por dois módulos, responsáveis pela(o): (i) *harmonização de cores*; e (ii) *ajuste de brilho e saturação*. Estes módulos modificam separadamente os histogramas do matiz, saturação e valor. Todavia, o módulo de harmonização de cores necessita do canal da saturação precedendo, portanto, o módulo de ajuste de brilho e saturação.

3.3.1 Harmonização de Cores

O processo de harmonização de cores tem por objetivo utilizar um conjunto de cores considerado harmônico para renderizar a imagem e, assim, melhorar seu aspecto visual. A harmonia de um conjunto de cores da imagem é descrita pela relação de posição que estas ocupam no espaço de cores, a qual provê uma percepção visual agradável ao olho humano (Cohen-Or et al., 2005). Na abordagem proposta, foi utilizada a técnica de harmonização de cores descrita por Cohen-Or et al. (2005), que estabelece a harmonia entre as cores por meio de *templates* de relações espaciais entre os matizes das cores.

Este processo pode auxiliar a redução da quantidade de cores da imagem (Cohen-Or et al., 2005), porquanto agrupa os matizes nos intervalos correspondentes aos setores de um dos *templates* harmônicos.

O módulo de harmonização de cores fundamenta-se na abordagem de Cohen-Or et al. (2005), a qual busca o melhor esquema harmônico para uma imagem, i.e., o esquema harmônico que minimiza a função $F(X,(m,\alpha))$ (vide Equação 2.5, no capítulo anterior). Este esquema produz

uma imagem harmonizada com as cores mais próximas possíveis à imagem original.

Assim como na abordagem proposta por Cohen-Or et al. (2005), neste trabalho não é utilizado o padrão (*template*) do tipo N para a harmonização das imagens, visto que tal padrão corresponde a imagens em escala de cinza. Na Figura 3.10, é exibido um exemplo de resultado da harmonização das cores de um cartum gerado pelo processo de abstração por segmentação descrito neste capítulo.

Figura 3.10 – Harmonização de cores: (A) Imagem original; (B) Imagem abstrata; (C) Imagem harmonizada (esquema: *template* I, ângulo 22° sentido horário); (D, E, e F) Histogramas dos matizes das imagens A, B e C, respectivamente.

Ao deslocar os matizes das cores da imagem, o processo de harmonização de cores estudado nesta dissertação pode gerar resultados não naturais, e.g. mudança na cor da pele, do céu, conforme anteriormente descrito por Cohen-Or et al. (2005). Por isto, o processo necessita da intervenção humana para identificar tais problemas e, caso necessário, alterar manualmente o *template* ou o ângulo do esquema

harmônico.

3.3.2 Ajuste de Brilho e Saturação

O ajuste de brilho e saturação da imagem fundamenta-se na característica dos histogramas de saturação (S) e luminosidade (V) dos cartuns, observada por Ianeva (2003). Para reproduzir esta característica nos cartuns gerados, são aplicadas transformações *gama* aos canais S e V da imagem (ver Figura 3.11).

Figura 3.11 - Ajuste de brilho e saturação: (A) Imagem original; (D) Imagem abstrata harmonizada; (G) Transformações *gama* (*Gama* = 0.5) aos canais S e V da imagem D; (B, E e H) Canais de saturação das imagens A, D e G respectivamente; (C, F e I) Canais de luminosidade (valor) das imagens A, D e G respectivamente.

Segundo Ianeva (2003), as médias de saturação e luminosidade nos cartuns são maiores do que aquelas exibidas por imagens fotográficas. Os valores dos parâmetros gama foram objeto de avaliação experimental discutida no próximo capítulo. O resultado obtido neste módulo é o resultado final da abordagem proposta.

3.4 Detalhes de Projeto e Implementação

Para a validação da abordagem proposta, foi desenvolvida uma aplicação de *software* para a geração de imagens abstratas ao estilo de cartuns. Os principais módulos descritos neste capítulo foram incorporados a esta aplicação.

Por diversas razões, a aplicação de *software* foi desenvolvida na linguagem C++, dentre as quais podem ser citadas: (i) a capacidade de gerar programas de maior desempenho computacional em relação a outras linguagens (e.g. linguagens interpretadas ou semi-interpretadas); (ii) a possibilidade de empregar recursos de orientação a objetos que incorporam aspectos favoráveis à engenharia do *software*, e.g., reusabilidade e manutenibilidade, dentre outros; e (iii) a facilidade de integração a bibliotecas avançadas de processamento de imagens, tais como a biblioteca OpenCV (utilizada neste trabalho). Para a construção, compilação e depuração do programa foi utilizada a ferramenta Microsoft Visual Studio 2005.

A aplicação de *software* desenvolvida foi estruturada em três grupos de classes, a saber: (i) classes relacionadas ao controle do processo, responsáveis por operações de leitura e escrita em disco e controle do fluxo de execução do processo; (ii) classes relacionadas à fase de *abstração da imagem*; e (iii) classes relacionadas à fase de *ajuste de cores*. As classes que fazem parte da aplicação desenvolvida são exibidas na Figura 3.12 e a separação entre os grupos destas classes é ilustrada pelas cores verde (*controle*), azul (*abstração da imagem*) e amarela (*ajuste de cores*).

Para a codificação do módulo do algoritmo de segmentação *Mean Shift*, foi utilizado o código fonte do EDISON[7], disponível em um conjunto de classes codificadas na linguagem C++.

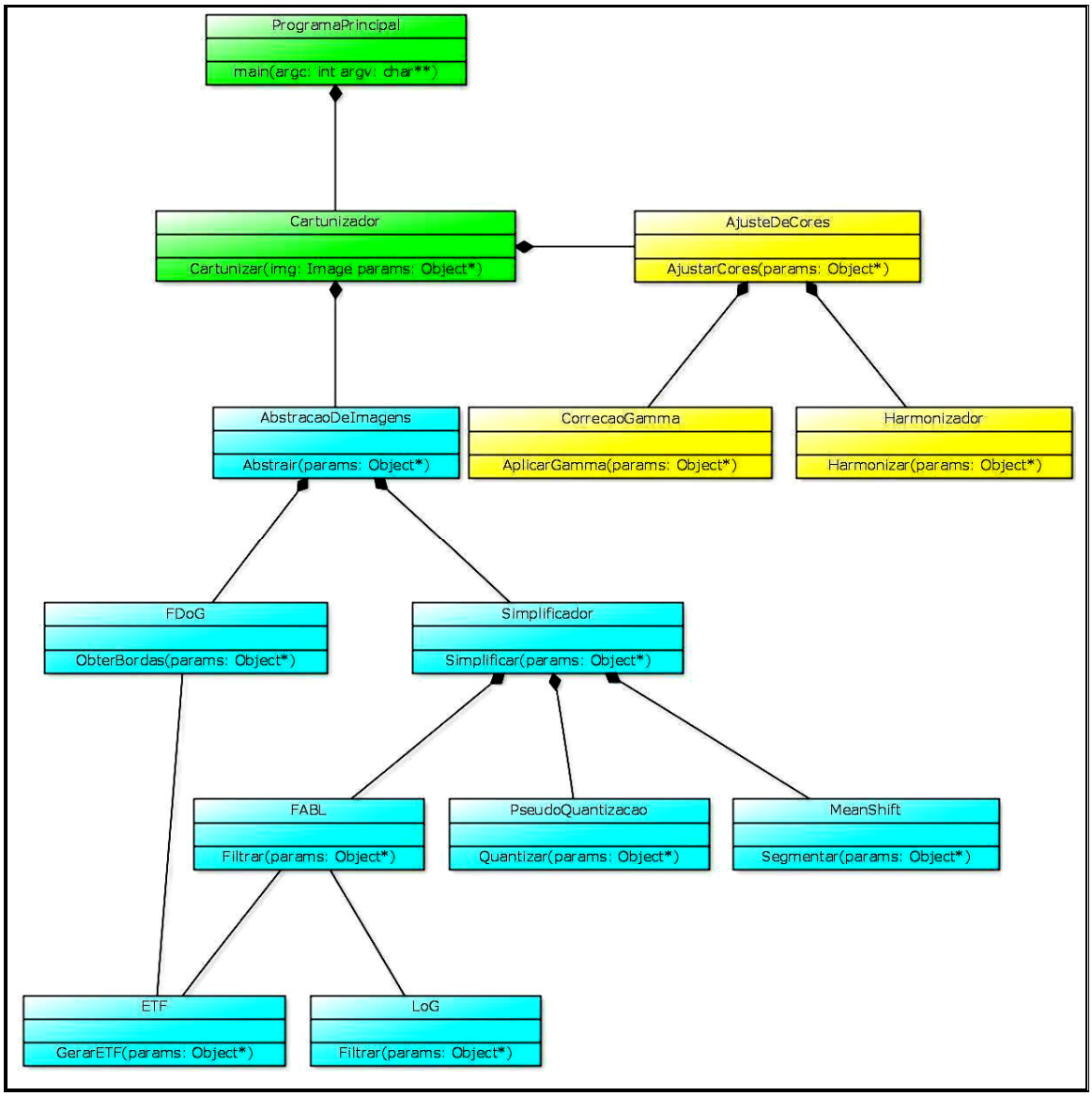


Figura 3.12 – Diagrama de classes da aplicação de *software* desenvolvida. As classes de cor verde são as classes de controle do processo, as classes de cor azul implementam a fase de *abstração da imagem* e as classes de cor amarela implementam a fase de *ajuste de cores*.

[7] Software de segmentação disponível em: http://coewww.rutgers.edu/riul/research/code.html

3.5 Considerações Finais do Capítulo

Neste capítulo, foi descrita a abordagem proposta nesta dissertação, destinada à geração automatizada de imagens abstratas ao estilo de cartuns e fundamentada pelos métodos revisados na bibliografia da área. Esta abordagem é dividida em duas fases: (i) *abstração da imagem* e (ii) *ajuste de cores*.

Na fase de abstração da imagem, a imagem de entrada do sistema é simplificada, i.e., são removidos detalhes não relevantes ao entendimento da imagem, e são obtidos os contornos dos elementos importantes da imagem, os quais são sobrepostos na imagem simplificada para evidenciar estes elementos. Para se obter a simplificação da imagem, foram descritos dois métodos diferentes: (i) *por suavização*; e (ii) *por segmentação*.

Na fase de ajuste de cores, a imagem abstrata obtida é submetida ao processo de harmonização de cores. Neste processo, o conjunto de cores que compõe a imagem é modificado com base em *templates* harmônicos de combinações de cores, visando, assim, melhorar o aspecto visual da imagem. Por fim, são aplicadas transformações *gama* aos canais de brilho e saturação da imagem.

No próximo capítulo, são descritos os resultados experimentais do processo de avaliação da abordagem proposta. Contudo, tendo em vista que nenhuma das aplicações de *software* utilizadas para a comparação com a abordagem proposta disponibiliza opções ou ajustes com propósitos similares à técnica de harmonização de cores (descrita na seção 2.2.3), esta não foi utilizada neste processo de validação. Além disto, a técnica de harmonização de cores utilizada não utiliza informações de alto nível (e.g. reconhecimento de elementos da imagem, detecção de pessoas e objetos), logo, a alteração dos matizes das cores permite gerar cores não naturais aos elementos da imagem, e.g. a mudança da cor da pele de uma pessoa presente na imagem para verde. Desta forma, a utilização desta técnica pode, eventualmente, prejudicar o resultado da abordagem.

Capítulo 4

Avaliação Experimental

Enquanto no fotorrealismo a eficácia de processos automáticos pode ser avaliada mediante a verificação da similaridade entre as imagens produzidas pelos métodos computacionais e imagens fotográficas, não se tem conhecimento de uma forma unificada ou objetiva para avaliar a eficácia dos processos de NPR. Na maioria dos artigos científicos da área, a validação limita-se à comparação visual dos resultados obtidos com outros resultantes de técnicas concorrentes (Hertzmann, 2010), fato constatado nesta pesquisa durante a revisão bibliográfica de abstração de imagens (e.g. Kang, Lee e Chui, 2009; Kyprianidis, Kang e Döllner, 2009).

Segundo Hertzmann (2010), a condução de estudos experimentais para a validação de resultados é um dos problemas mais difíceis enfrentados pela comunidade científica da área de NPR. Apesar da limitação supramencionada, alguns métodos foram propostos, no sentido de facilitar e uniformizar a validação do processo de abstração de imagens, dentre os quais podem-se citar experimentos psicológicos relacionados: (i) à facilitação da memorização das imagens (e.g. Winnemöeller, 2006); (ii) ao reconhecimento do conteúdo das imagens (e.g. Winnemöeller, 2006; Wallraven et al., 2007); (iii) ao direcionamento da atenção visual (e.g. Santella, 2005); e (iv) à preferência subjetiva das imagens abstratas produzidas (e.g. Wallraven et al., 2007).

Neste capítulo, a abordagem proposta para a abstração de imagens digitais ao estilo de cartuns é avaliada por meio da confrontação visual de

seus resultados com outros provenientes de abordagens concorrentes. Além disto, com o intuito de validar a qualidade dos cartuns gerados a partir da abordagem proposta, foram conduzidos experimentos subjetivos envolvendo 51 participantes, cujos resultados foram posteriormente analisados estatisticamente a partir dos testes ANOVA e Tukey-Kramer. O método de avaliação adotado neste trabalho insere-se nas categorias de experimentos (ii) e (iv) mencionadas no parágrafo anterior.

4.1 Comparação Visual dos Resultados

Conforme mencionado no Capítulo 2, as abordagens adotadas no processo de abstração de imagens normalmente utilizam técnicas de suavização e segmentação. Para a abstração de imagens por suavização, foi proposto nesta dissertação o filtro FABL, uma variante do filtro FBL utilizado em abordagens de abstração de imagens (e.g. Kyprianidis e Döllner, 2008; Kang, Lee e Chui, 2009; Zhao et al., 2009). No filtro FBL, a informação da direção das bordas é utilizada para reforçar a estilização das bordas preservadas pela filtragem *Bilateral*. Adicionalmente, no filtro FABL é utilizado o caráter adaptativo da filtragem radial proposta por Zhang e Allebach (2008), a fim de manter a nitidez das bordas da imagem.

Nas Figuras 4.1 (B) e (C), são exibidos os resultados das abstrações produzidas a partir do uso dos filtros FBL e FABL, respectivamente.

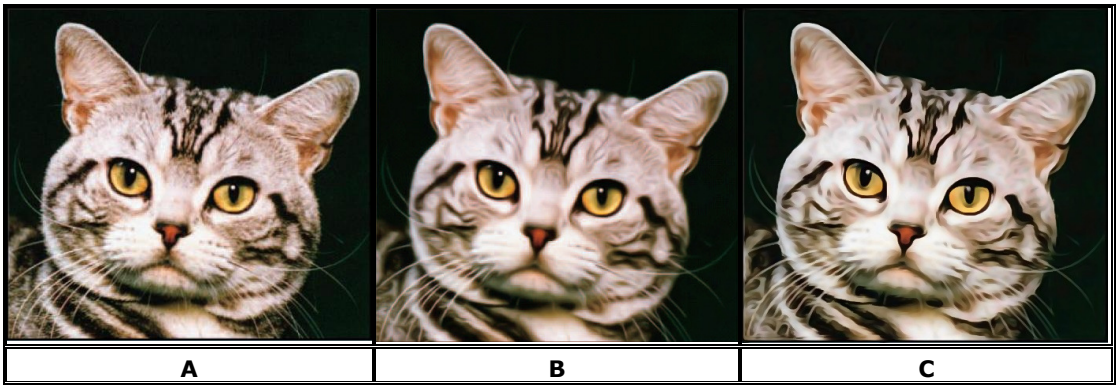


Figura 4.1 – Comparações de abstração por suavização: (A) Imagem original; (B) Filtro FBL; (C) Filtro FABL.

Os resultados apresentados nas Figuras 4.1 (B) e (C) permitem verificar que, além da estilização produzida pelo filtro FBL, o filtro FABL preserva a nitidez dos detalhes representativos da imagem, mantendo aguçados o nariz, os olhos e os pêlos escuros do animal. Esta vantagem também é observada nas Figuras 4.2(B) a (E), quando se compara os resultados obtidos a partir do uso do filtro FABL àqueles provenientes do uso do filtro *Kuwahara* (Kuwahara et al., 1976) e suas variações (atentar para a franja de pêlos do pescoço e região em torno dos olhos do animal, melhor preservados pelo filtro FABL).

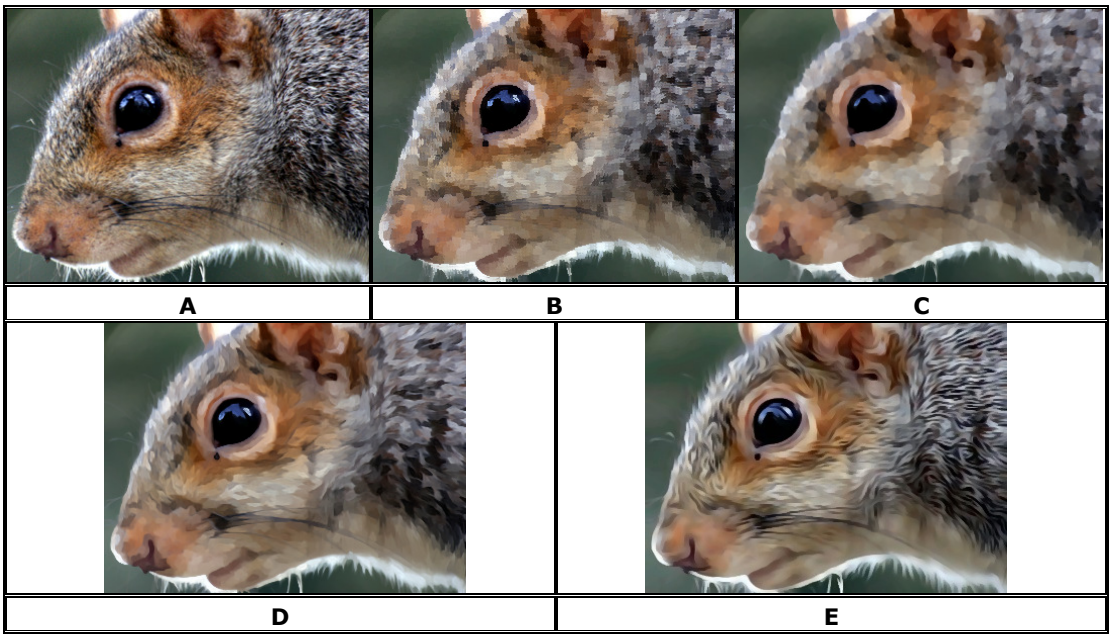


Figura 4.2 – Comparações de abstração por suavização: (A) Imagem original; (B) Filtro *Kuwahara* original; (E) Filtro *Kuwahara* proposto por Papari, Petkov e Campisi (2007); (H) Filtro *Kuwahara* proposto por Kyprianidis, Kang e Döllner (2009); (E) Filtro FABL.

Dentre os trabalhos revisados sobre abstração por segmentação, encontram-se as abordagens de DeCarlo e Santella (2002) e Medhi, Patabandhige e Toyama (2007). Conforme descrito anteriormente, a técnica de segmentação utilizada nesta dissertação foi *Mean Shift*, também utilizada por DeCarlo e Santella (2002), que se destaca pelo

desempenho, precisão e pela disponibilidade de implementações públicas (Estrada e Jepson, 2005; Arbeláez et al., 2009).

Na abordagem de DeCarlo e Santella (2002), a informação contida no plano de fundo (*background*) da imagem é fortemente abstraída, enquanto preserva a informação do primeiro plano (*foreground*) da imagem. Este comportamento é obtido a partir de um modelo de percepção humana e de um registro dos movimentos oculares do usuário, conforme descrito anteriormente.

Por não possuir informações referentes à localização da informação saliente da imagem, a segmentação adotada nesta dissertação não apresenta tal comportamento (vide Figuras 4.3 (B) e (C)), o que sugere a realização de pesquisas futuras sobre a detecção de elementos salientes na imagem.

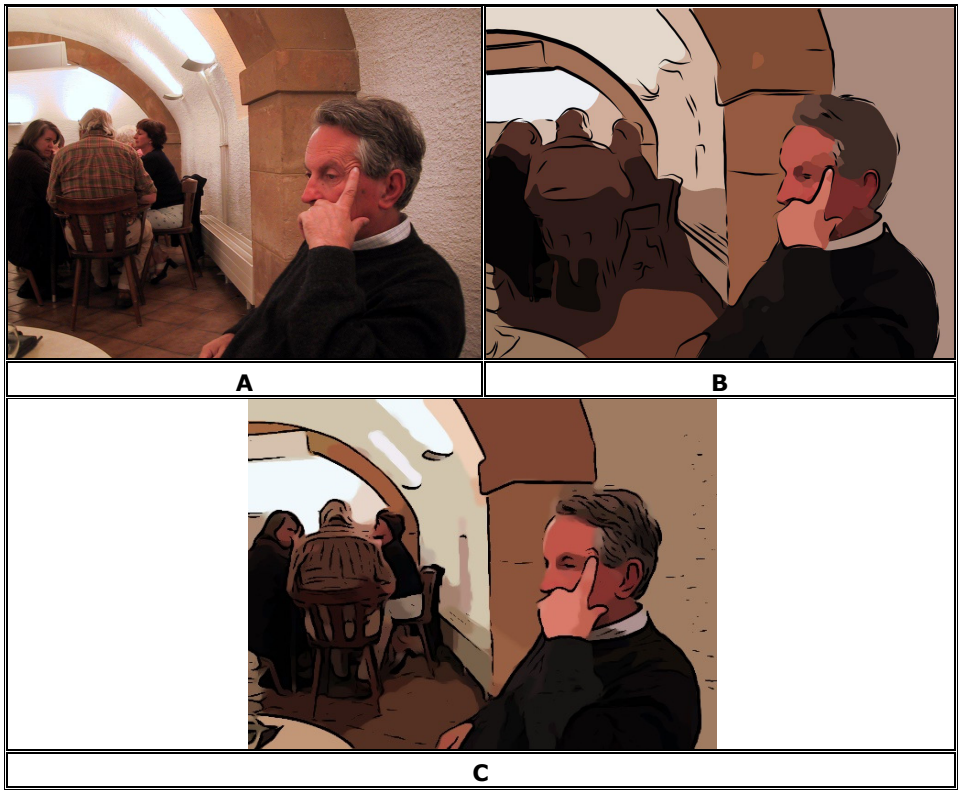


Figura 4.3 – Comparações de abstração por segmentação: (A) Imagem original; (B) Resultado de DeCarlo e Santella (2002); (C) Resultado da abordagem utilizada nesta dissertação.

Com o objetivo de gerar cartuns a partir de imagens digitais, Medhi, Patabandhige e Toyama (2007) utilizaram um processo de segmentação baseada na rotulagem dos componentes conectados, a qual foi auxiliada por uma ferramenta destinada à correção manual das falhas do processo de segmentação. Adicionalmente, a mesma ferramenta auxilia no ajuste da saturação do cartum gerado. Mesmo com o auxílio da ferramenta de segmentação, a técnica de Medhi, Patabandhige e Toyama (2007) apresenta problemas quando se processam imagens com muitos detalhes de textura. Um comparativo entre o método *Mean Shift* - utilizado nesta dissertação - e a abordagem dos autores supramencionados, conforme exemplo ilustrado nas Figuras 4.4 (B) e (C), evidencia as vantagens do método *Mean Shift*.

Figura 4.4 – Comparações de abstração por segmentação: (A) Imagem original; (B) Resultado de Medhi, Patabandhige e Toyama (2007); (C) Resultado da abordagem utilizada nesta dissertação (*Mean Shift*).

A fim de analisar de forma mais rigorosa os resultados obtidos pela abordagem proposta, foi conduzido um processo de avaliação, mediante a estruturação de um experimento subjetivo, o qual é detalhado na próxima seção.

4.2 Experimento Subjetivo

O principal objetivo almejado pelo processo de avaliação discutido nesta seção é demonstrar que a abordagem proposta nesta dissertação é adequada ao propósito ao qual se presta, i.e., a geração de imagens ao estilo de cartuns a partir de imagens fotográficas. Para tanto, foi projetado um experimento com o intuito de avaliar a qualidade visual (i.e., percebida por seres humanos) dos resultados gerados a partir da abordagem proposta.

Neste experimento, foi utilizada a aplicação de *software* implementada para a validação da abordagem proposta nesta dissertação (descrita no Capítulo 3), além de outras aplicações de *software* concorrentes disponíveis. Cabe ressaltar que as aplicações de *software* selecionadas não indicam específica e detalhadamente quais as técnicas computacionais nas quais se fundamentam. Contudo, todas a aplicações selecionadas possuem conjuntos de parâmetros para o controle do nível de abstração equivalentes àqueles existentes na aplicação de *software* implementada para a validação da abordagem proposta. Não obstante ser desejável ter utilizado as abordagens revisadas, não foram encontrados códigos-fonte ou aplicações de *software* publicamente disponíveis associados a tais abordagens. Por tal motivo, o qual não influencia a generalidade do processo de comparação, decidiu-se utilizar, para este fim, aplicações de *software* comerciais.

4.2.1 Softwares Utilizados

Para conduzir o experimento subjetivo, foram utilizadas cinco aplicações de *software* disponíveis na *World Wide Web*, a saber: (i) *MSU Cartoonizer*[8], (ii) *BeFunky*[9], (iii) *KusoCartoon*[10], (iv) *CaricatureSoft*[11] e (v) *LiangZhu*[12].

[8] VirtualDub MSU cartoonizer filter.
Disponível em: http://www.compression.ru/video/cartoonizer/index_en.html.

Na Figura 4.5, são exibidas algumas imagens obtidas a partir das referidas aplicações.

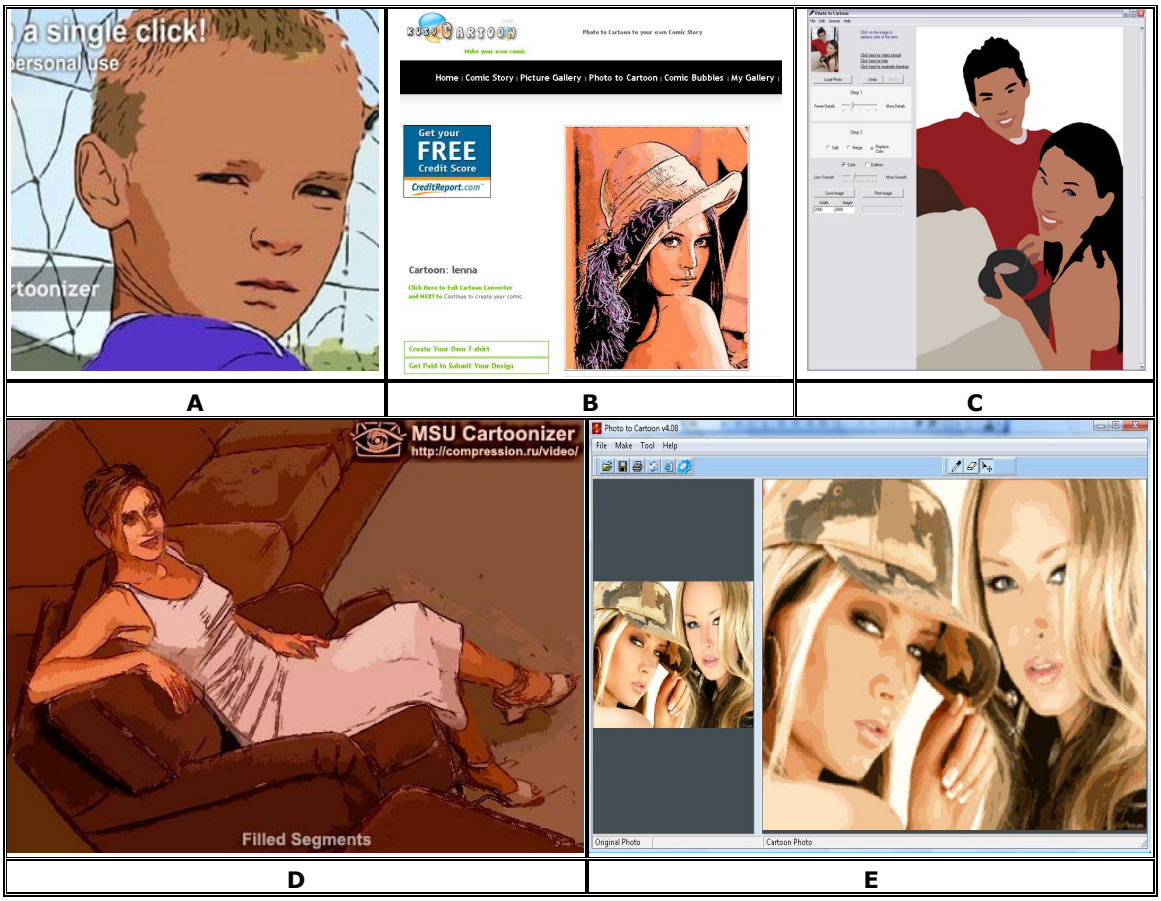


Figura 4.5 – *Softwares* utilizados para comparação: (A) BeFunky; (B) KusoCartoon; (C) CaricatureSoft; (D) MSU Cartoonizer; (E) LiangZhu.

Os cartuns gerados pelas aplicações utilizadas são obtidos por meio de técnicas de transformações das cores e extração dos contornos das imagens. É conveniente reiterar que as aplicações de *software* selecionadas para o processo de comparação de resultados disponibilizam conjuntos de parâmetros que permitem ao usuário a obtenção de

---

[9] BeFunky.com: Funky way to express yourself.
Disponível em: *http://www.befunky.com*.

[10] KusoCartoon PhotoToCartoon Beta Testing.
Disponível em: *http://www.kusocartoon.com*.

[11] CaricatureSoft PhotoToCartoon
Disponível em: *http://www.caricaturesoft.com/products/photo-to-cartoon.html*.

[12] LiangZhu PhotoToCartoon.
Disponível em: *http://www.liangzhuchina.com*.

diferentes níveis de abstração das imagens. Desta forma, para efeito de comparação, foi utilizado na aplicação de *software* associada à abordagem proposta o método da abstração por segmentação, o qual possibilita a obtenção de maiores níveis de abstração.

O experimento subjetivo foi dividido em duas etapas, a saber: (i) *etapa de calibração de parâmetros* e (ii) *etapa de comparação de resultados*, as quais são discutidas mais detalhadamente a seguir.

4.2.2 Calibração de Parâmetros

A etapa de calibração de parâmetros teve como objetivo identificar as melhores configurações de parâmetros para cada uma das aplicações de *software* utilizadas no experimento. Estas configurações foram posteriormente utilizadas para a geração de resultados que serviram de base para a segunda etapa do experimento subjetivo.

Os dados obtidos nesta etapa foram coletados com o auxílio de 20 participantes voluntários. Foram utilizadas 16 imagens de entrada, as quais possibilitaram a geração de um total de 576 imagens de abstrações ao estilo de cartuns.

Com a finalidade de utilizar grupos de imagens semelhantes nas duas etapas conduzidas, o conteúdo destas imagens foi restringido em quatro categorias diferentes, a saber: (i) imagens contendo *animais*; (ii) imagens contendo *uma única pessoa*; (iii) imagens contendo *um grupo de pessoas*; e (iv) imagens de *cenários diversos* (e.g., construções, objetos ou paisagens).

Na etapa de calibração de parâmetros, foram utilizadas 4 imagens em cada categoria (3 imagens fotográficas e uma imagem sintética, gerada por Computação Gráfica). A utilização das imagens sintéticas teve por objetivo apresentar, entre as imagens de cada categoria, uma imagem com poucas informações contidas no *background*, com o intuito de facilitar o ajuste dos parâmetros aos elementos de interesse das categorias. As imagens fotográficas foram obtidas da base de imagens de *Philip*

Greenspun[13], escolhida por ter sido utilizada em outros trabalhos de abstração de imagens (e.g. DeCarlo e Santella, 2002; Kang e Lee, 2008; Kyprianidis e Döllner, 2008).

Para cada aplicação de *software* utilizada no experimento, foram obtidas diferentes configurações de parâmetros, as quais combinam valores de ajuste da transformação das cores, da quantidade dos contornos e do ajuste de saturação. Os parâmetros das aplicações de *software* utilizados no processo de calibração são apresentados no Quadro 4.1.

Quadro 4.1 – Parâmetros de controle dos *softwares*.

Aplicação de Software	Parâmetro	Modificação
BeFunky	Sketch Detail	Quantidade de Contornos
	Color Detail	Quantidade de Cores
CaricatureSoft	Smooth	Suavização da Imagem
	Color Detail	Quantidade de Cores
	Sketch Detail	Quantidade de Contornos
LiangZhu	Scalar	Quantidade de Cores
	Sleek	Suavização da Imagem
	Pencil Pressure	Quantidade de Contornos
MSU Cartoonizer	N. Edges	Quantidade de Contornos
	N. Segments	Quantidade de Cores
	Saturation	Níveis de Saturação
KusoCartoon	Detail	Quantidade de Contornos
	Color	Quantidade de Cores
	Saturation	Níveis de Saturação
Abordagem Proposta	Mean Shift (Sigma Radial)	Quantidade de Cores
	FDoG(Sigma 1)	Quantidade de Contornos
	Saturação(Gama)	Níveis de Saturação

[13] Banco de dados de imagens fotográficas:
Disponível em: *http://philip.greenspun.com*

A quantidade das configurações obtidas foi variada de acordo com a necessidade de cada aplicação, objetivando evitar configurações que gerassem resultados visualmente similares, o que dificultaria o processo de calibração. A quantidade máxima de diferentes configurações de parâmetros para cada *software* foi limitada em 18 e foram escolhidos 3 parâmetros para variação de valores. Na aplicação de *software* BeFunky, foram utilizados apenas 2 parâmetros, por ser a quantidade total de parâmetros disponibilizados nesta aplicação.

Durante o processo de calibração de cada aplicação de *software*, foram apresentadas aos participantes as 16 imagens de calibração selecionadas, para cada uma das quais foram exibidos 6 cartuns, produzidos por diferentes configurações de parâmetros escolhidas aleatoriamente, a fim de evitar vieses. Os participantes deste experimento foram orientados a selecionar os melhores cartuns produzidos - no máximo dois cartuns por imagem, com base nos critérios exibidos no Quadro 4.2.

Foi também recomendada a não seleção de imagens, dentre aquelas apresentadas, caso não atendesses aos critérios de avaliação. Todos os usuários de teste deste experimento participaram da calibração das 6 aplicações de *software*.

Quadro 4.2 – Critérios de avaliação.

Critério	Descrição
1	O cartum deve ser formado principalmente por áreas uniformemente coloridas realçadas por linhas de contornos escuras.
2	No cartum, o assunto principal (objetos, pessoas, materiais etc.) não deve apresentar falhas ou imperfeições graves, como por exemplo, partes faltantes.
3	O conteúdo semântico (como objetos, pessoas, materiais etc.) da imagem original deve ser facilmente reconhecido no cartum.
4	O cartum deve manter um aspecto visual agradável.

A elaboração dos critérios de avaliação foi fundamentada por definições e conceitos de cartum apresentada por Ianeva (2003), Medhi, Patabandhige e Toyama (2007) e Wallraven et al. (2007), assim como por

conceitos de abstração de imagens apresentados no Capítulo 2 desta dissertação. Resumidamente, técnicas de abstração ao estilo de cartuns objetivam produzir imagens formadas por regiões uniformemente coloridas sobrepostas por bordas que enfatizam os contornos. Nestas imagens, o conteúdo importante da imagem deve ser preservado, assim como a facilidade no reconhecimento da informação, por meio de resultados artísticos que sejam visualmente agradáveis ao olho humano.

Na Figura 4.6, é ilustrado um excerto da interface gráfica da aplicação de calibração dos parâmetros utilizada nesta etapa do experimento subjetivo.

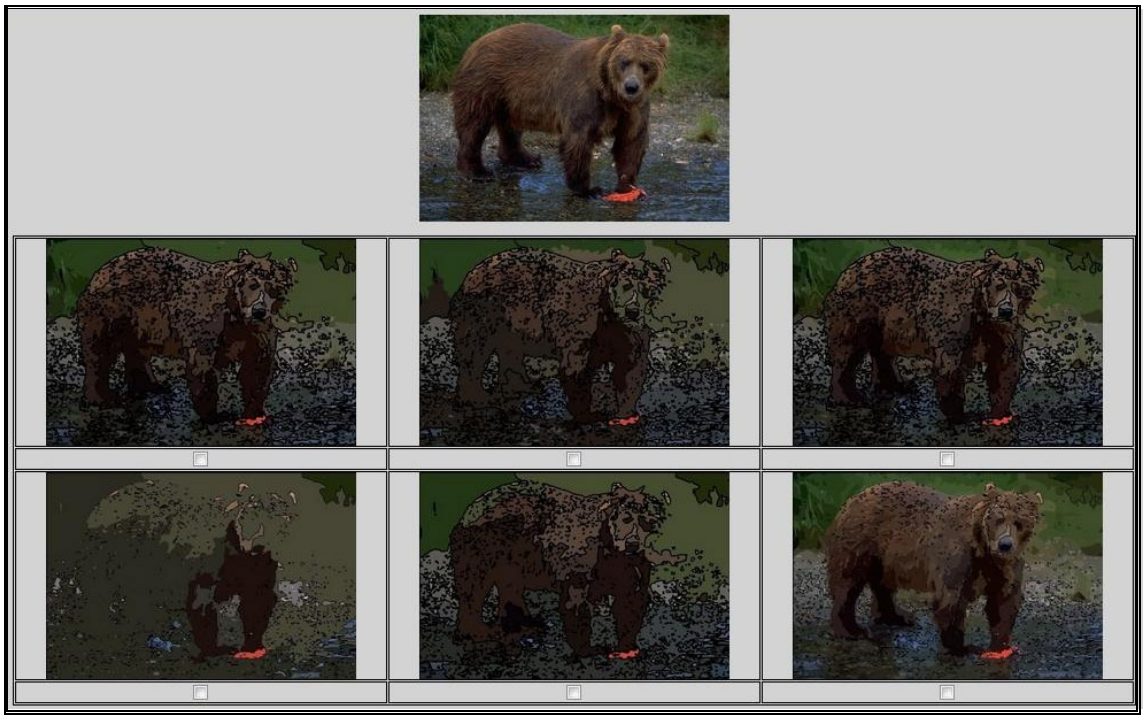


Figura 4.6 – Excerto da interface com o usuário utilizada na etapa de calibração de parâmetros.

Após o processo de aquisição dos dados, foi realizado um processo de triagem dos resultados obtidos em cada uma das aplicações de *software* calibradas. Neste processo, os dados obtidos de cada participante foram sumarizados e agrupados, para obtenção do valor médio da quantidade relativa dos votos recebidos por cada configuração de parâmetros. A

quantidade relativa (Q_{rel}) de votos recebidos por uma configuração de parâmetros c é definida por:

$$Q_{rel} = \frac{Q_{votos}}{Q_{total}} \tag{4.1}$$

em que Q_{votos} representa a quantidade de imagens geradas pela configuração c selecionadas pelo participante, enquanto Q_{total} representa a quantidade total de imagens geradas pela configuração c exibidas ao participante.

Antes de comparar as médias da quantidade relativa dos votos recebidos pelas diferentes configurações de parâmetros, foi formulada e verificada a hipótese de não existir diferença estatisticamente significativa entre estas médias. Para verificação de tal hipótese, foi administrado um procedimento estatístico utilizando o teste F ANOVA fator único (Levine, Berenson e Stephan, 2000).

A condução rigorosa dos testes de hipóteses nesta etapa dificulta a separação entre as melhores configurações de parâmetros e as demais, tornando o teste menos eficaz para identificar as piores configurações. Desta forma, para os testes de hipóteses administrados nesta etapa foi utilizado um nível confiança igual a 95%, comumente recomendado para diversas aplicações estatísticas (Kimothi, 2002; Wheeler e Ganji, 2008). Em contrapartida, na etapa de comparação de resultados foi considerado um nível de confiança maior (99,99%), para obter maior credibilidade ao evidenciar diferenças entre os resultados das diferentes aplicações de *software*.

Na Tabela 4.1, são exibidos os resultados do teste F ANOVA fator único, obtidos com o auxílio da aplicação de *software* Microsoft Office Excel 2003, para os dados associados à calibração de cada uma das aplicações de *software* consideradas. Conforme explicitado nesta tabela, os valores F são maiores do que os F críticos, significando que existem diferenças significativas entre as médias obtidas pelos grupos.

Tabela 4.1 – Teste F ANOVA fator único para a calibração dos *softwares*.

ANOVA: fator único						α =0,05
Software	**Fonte da variação**	**SQ**	**gl**	**MQ**	**F**	**F crítico**
BeFunky	Intergrupos	4,1611	5	0,83222	**34,21981**	**2,293911**
	Intragrupos	2,772461	114	0,02432		
	Total	6,933561	119			
Caricature Soft	Intergrupos	13,04821	11	1,18620	**79,69926**	**1,830818**
	Intragrupos	3,393429	228	0,01488		
	Total	16,44164	239			
LiangZhu	Intergrupos	11,12235	11	1,01112	**26,82383**	**1,830818**
	Intragrupos	8,594444	228	0,03769		
	Total	19,71679	239			
MSU Cartoonizer	Intergrupos	8,024922	11	0,72953	**21,41393**	**1,830818**
	Intragrupos	7,767596	228	0,03406		
	Total	15,79252	239			
KusoCartoon	Intergrupos	6,650344	17	0,39119	**11,35621**	**1,65271**
	Intragrupos	11,78116	342	0,03444		
	Total	18,4315	359			
Abordagem Proposta	Intergrupos	6,379319	17	0,37525	**8,44057**	**1,65271**
	Intragrupos	15,20477	342	0,04445		
	Total	21,58408	359			

Uma vez que o teste F ANOVA fator único evidencia apenas a existência de diferença entre as médias dos grupos, não possibilitando comparações entre pares de grupos, foi utilizado o teste de Tukey (Levine, Berenson e Stephan, 2000) para tal finalidade.

Todas as configurações que evidenciaram diferenças significativas (por meio do resultado do teste de Tukey) à configuração de maior média não foram utilizadas na etapa seguinte, ou seja, na etapa de comparação das aplicações de *software*.

A médias obtidas e os resultados do teste de Tukey para as configurações de parâmetros das aplicações de *software* encontram-se no Apêndice D. Os resultados dos testes de Tukey administrados nesta pesquisa foram obtidos com o auxílio da aplicação de *software* XLSTAT[14].

[14] XLSTAT Statistical Analysis Software. Disponível em: *http://www.xlstat.com/*

Os testes de hipóteses F ANOVA fator único e o teste de Tukey são descritos no Apêndice E desta dissertação.

4.2.3 Comparação de Resultados

Na etapa de comparação dos resultados produzidos a partir das aplicações de *software* utilizadas no experimento subjetivo, foram utilizadas 60 imagens fotográficas obtidas da base de imagens de *Philip Greenspun* (15 imagens para cada categoria). Nesta etapa, gerou-se, a partir de cada aplicação de *software*, apenas um cartum para cada imagem considerada, utilizando a melhor ou uma das melhores configurações de parâmetros obtidas na fase de calibração. Nos casos em que houve mais de uma configuração, a escolha da configuração para a geração do cartum representante foi aleatória.

Nesta etapa, considerou-se um segundo grupo de participantes, composto por 31 voluntários, distintos daqueles que participaram da etapa de calibração. Os critérios apresentados aos participantes desta etapa foram os mesmos critérios apresentados na etapa de calibração dos parâmetros. Similarmente, para cada imagem utilizada foram exibidas seis abstrações ao estilo de cartuns, representantes de cada aplicação de *software* considerada.

Os participantes foram orientados a selecionar no máximo duas das melhores abstrações ao estilo de cartum dentre aquelas que lhes foram apresentadas. Vale a pena reiterar que, nesta etapa, diferentemente da primeira, cada um dos cartuns foi produzido por uma aplicação de *software* diferente. Também é importante ressaltar que, em ambas as fases do experimento subjetivo, a estratégia adotada para a apresentação das abstrações não possibilitava a identificação de quais as aplicações de *software* que haviam sido utilizadas para produzir os cartuns. Além disto, as posições das abstrações em cada uma das telas apresentadas ao participante, durante o teste, variaram aleatoriamente para cada teste realizado.

Na Figura 4.7, é ilustrado um excerto da interface com o usuário utilizada na segunda etapa experimental.

Figura 4.7 – Excerto da interface com o usuário utilizada na etapa de comparação dos resultados das aplicações de *software* utilizadas.

Após a realização do experimento e aquisição dos dados, o valor médio da quantidade relativa dos votos recebidos por cada aplicação de *software* foi calculado similarmente à etapa de calibração de parâmetros.

Tendo em vista que todas as aplicações de *software* foram avaliadas em todas as comparações desta etapa, a quantidade total de imagens produzidas pelas aplicações de *software* corresponde exatamente ao número de imagens fotográficas utilizadas (60 imagens).

Assim, a quantidade relativa (Q_{rel}) de votos obtidos pelas aplicações de *software* em cada participação corresponde a:

$$Q_{rel} = \frac{Q_{votos}}{60} \qquad (4.2)$$

em que Q_{votos} é quantidade de imagens selecionadas pelo participante geradas pela aplicação de *software*. Os valores das médias e desvios padrão da quantidade relativa de votos obtidos por cada aplicação de *software* são exibidos Na Tabela 4.2.

Tabela 4.2 – Média e desvio padrão da quantidade relativa de votos obtidos por cada aplicação de *software*.

Software	Média	Desvio Padrão
BeFunky	0,43871	0,209911229
Caricaturesoft	0,030645	0,036539368
LiangZhu	0,086559	0,091532159
MSUCartoonizer	0,064516	0,090445598
KusoCartoon	0,287097	0,175209381
Abordagem Proposta	0,708602	0,148690019

Similarmente à etapa anterior, foi administrado o teste F ANOVA fator único, a fim de evidenciar a presença de diferenças estatisticamente significativas entre os valores das médias. O resultado deste teste, administrado com o auxílio da aplicação de *software* Microsoft Office Excel 2003, é exibido na Tabela 4.3.

Tabela 4.3 – Teste F ANOVA fator único para a comparação dos *softwares*.

ANOVA: fator único					α = ***0,0001***
Fonte da variação	**SQ**	**gl**	**MQ**	**F**	**F crítico**
Intergrupos	10,98297	5	2,196595	**114,8412**	**5,487052**
Intragrupos	3,442903	180	0,019127		
Total	14,42588	185			

Para este teste, foi considerado um nível de confiança igual a 99,99%, a fim de aumentar o rigor do teste, com o intuito de elevar a credibilidade da evidência de diferenças significativas entre as médias obtidas. Visto que o valor F é maior que o F crítico, evidenciando diferenças entre as médias da quantidade relativa de votos obtidos pelos *softwares*, procedeu-se a comparação entre pares de médias por meio do

teste de Tukey, também utilizando um nível de confiança igual a 99,99%. Uma síntese dos resultados deste teste é apresentada na Tabela 4.4.

Tabela 4.4 – Teste de Tukey para a comparação dos *softwares*.

Teste Tukey				α = ***0,0001***
Contraste	**Erro padrão da diferença**	**Diferença (Dif1)**	**Diferença crítica (Dif2)**	**Significante (Dif1 > Dif2)**
Abordagem Proposta vs Caricaturesoft	0,025	**0,678**	**0,163**	**Sim**
Abordagem Proposta vs MSUCartoonizer	0,025	**0,644**	**0,163**	**Sim**
Abordagem Proposta vs LiangZhu	0,025	**0,622**	**0,163**	**Sim**
Abordagem Proposta vs KusoCartoon	0,025	**0,422**	**0,163**	**Sim**
Abordagem Proposta vs BeFunky	0,025	**0,270**	**0,163**	**Sim**
BeFunky vs Caricaturesoft	0,025	**0,408**	**0,163**	**Sim**
BeFunky vs MSUCartoonizer	0,025	**0,374**	**0,163**	**Sim**
BeFunky vs LiangZhu	0,025	**0,352**	**0,163**	**Sim**
BeFunky vs KusoCartoon	0,025	**0,152**	**0,163**	**Não**
KusoCartoon vs Caricaturesoft	0,025	**0,256**	**0,163**	**Sim**
KusoCartoon vs MSUCartoonizer	0,025	**0,223**	**0,163**	**Sim**
KusoCartoon vs LiangZhu	0,025	**0,201**	**0,163**	**Sim**
LiangZhu vs Caricaturesoft	0,025	**0,056**	**0,163**	**Não**
LiangZhu vs MSUCartoonizer	0,025	**0,022**	**0,163**	**Não**
MSUCartoonizer vs Caricaturesoft	0,025	**0,034**	**0,163**	**Não**

Por meio dos resultados obtidos a partir do teste de Tukey, conforme exibidos na Tabela 4.4, algumas conclusões podem ser formuladas, a saber:

1. Foi evidenciado que a média da quantidade de imagens geradas a partir da abordagem proposta e que foram selecionadas pelos participantes do experimento é estatisticamente diferente das médias das demais aplicações de *software* utilizadas, visto que a

diferença entre estas médias é maior do que a diferença crítica do teste de Tukey;

2. Não foram evidenciadas diferenças significativas (para o nível de confiança utilizado) entre as médias da quantidade de imagens geradas a partir das aplicações de *software* BeFunky e KusoCartoon que foram selecionadas pelos participantes, embora tenham sido evidenciadas diferenças entre as médias obtidas por estas aplicações de *software* e às demais; e
3. Assim como na situação supramencionada, também não foram evidenciadas diferenças significativas entre as médias obtidas pelas aplicações de *software* LiangZhu, CaricatureSoft e MSUCartoonizer, embora tenham sido evidenciadas diferenças entre as médias obtidas por estas aplicações de *software* e às demais.

Assim sendo, foi possível separar as aplicações de *software* em três grupos, por meio da similaridade das médias da quantidade de imagens selecionadas pelos participantes, conforme explicitado no Quadro 4.3.

Quadro 4.3 – Grupos de similaridade obtidos por meio do teste de Tukey.

Softwares	Grupos		
Abordagem Proposta	A		
BeFunky		B	
KusoCartoon		B	
LiangZhu			C
MSUCartoonizer			C
Caricaturesoft			C

A partir da análise do teste de Tukey, apresentado na Tabela 4.4, assim como da análise das médias dos desempenhos das aplicações de *software* apresentadas na Tabela 4.2, pode ser verificado que a aplicação de *software* que implementa a abordagem proposta apresentou média superior e estatisticamente diferente das médias das demais aplicações de

software. Tal fato evidencia que a abordagem proposta é eficaz na produção de cartuns que atendem aos critérios previamente formulados.

4.3 Considerações Finais do Capítulo

Neste capítulo, resultados de outras abordagens para a abstração de imagens foram comparados com os resultados gerados a partir da abordagem proposta. Além disto, foi conduzido um experimento subjetivo comparativo para avaliar a qualidade dos cartuns gerados pela abordagem proposta. Este experimento foi realizado com a participação de 51 usuários de teste, 20 dos quais alocados para a fase de calibração dos parâmetros e 31 para a fase comparação dos resultados produzidos pelas aplicações de *software* consideradas no experimento.

A partir da análise dos resultados obtidos no experimento subjetivo, pôde-se concluir que existem evidências estatísticas de que, para os conjuntos de parâmetros obtidos na fase de calibração aplicados às imagens escolhidas, a qualidade dos resultados gerados pela abordagem proposta é superior à qualidade dos resultados produzidos pelas demais aplicações de *software* comparadas.

Para avaliar a eficácia do processo de abstração, outros experimentos futuros poderão ser conduzidos, com foco, por exemplo, na facilidade de localização e memorização das informações da imagem. Porém, este trabalho teve como principal objetivo obter uma abordagem para geração de imagens cartuns que apresentam bom aspecto visual, limitando-se, pois, na verificação comparativa da qualidade (percebida pelos participantes do experimento subjetivo conduzido) dos cartuns gerados por esta abordagem e pelas demais consideradas no experimento.

Capítulo 5

Considerações Finais

Neste capítulo, é apresentada uma síntese da pesquisa reportada nesta dissertação. Além disto, são apresentadas as contribuições e realizadas propostas para trabalhos futuros, as quais agregam valor à abordagem proposta para a geração de abstrações de imagens.

5.1 Síntese da Pesquisa

A pesquisa desenvolvida nesta dissertação teve como meta central a concepção de uma abordagem para a transformação de imagens fotorrealísticas em imagens abstratas ao estilo de cartuns. A motivação do tema desta pesquisa fundamentou-se no crescente interesse demonstrado na área de PDI, pela indústria de tecnologias da informação (e.g. empresas tais como a Microsoft e a HP), assim como pela indústria cinematográfica e televisiva, conforme comentado nos dois capítulos iniciais deste documento.

Foram conduzidos estudos relacionados ao processo de abstração de imagens e a técnicas de suavização, segmentação e ajustes de cores. Tais estudos foram apresentados no Capítulo 2 desta dissertação e serviram como fundamentação para a concepção da arquitetura da abordagem proposta. Além disto, os referidos estudos fundamentaram a implementação da aplicação de *software* utilizada no processo de validação desta abordagem.

Adicionalmente, foi também documentada uma extensão original do filtro de suavização FBL, utilizado por trabalhos encontrados no estado da arte de abstração de imagens, devidamente citados na revisão bibliográfica. Os detalhes desta extensão e dos módulos que compõem a abordagem proposta foram apresentados no Capítulo 3.

Por fim, os resultados obtidos pelo *software* desenvolvido para a validação da abordagem foram comparados com os resultados de outras aplicações de *software* destinadas à geração de imagens abstratas ao estilo de cartuns. Foi possível concluir a partir dos experimentos subjetivos conduzidos e do processamento estatístico dos dados coletados que as imagens ao estilo de cartuns produzidas pela abordagem proposta apresentaram forte apelo visual a observadores humanos quando comparadas aos resultados equivalentes produzidos pelas demais aplicações de *software* consideradas. Desta forma, pode-se concluir que o objetivo central da pesquisa foi alcançado.

5.2 Principais Contribuições

As principais contribuições deste trabalho foram:

1. A compilação de um corpo bibliográfico relacionado às principais técnicas utilizadas para a abstração de imagens digitais;
2. A proposição de uma abordagem para a abstração de imagens, fundamentando-se no corpo bibliográfico compilado;
3. A proposição de uma extensão original do filtro de suavização FBL, com a finalidade de manter o aguçamento da imagem após sucessivas iterações;
4. A proposição de um experimento subjetivo para a validação da abordagem, utilizando critérios de avaliação fundamentados por definições e conceitos no corpo bibliográfico revisado;
5. A constatação experimental de benefícios obtidos com a utilização das técnicas de suavização e segmentação para o processo de abstração de imagens proposto; e

6. A agregação de módulos para o ajuste das cores da imagem ao processo de abstração.

Vale acrescentar que a abordagem proposta nesta dissertação produz imagens artísticas ou estilizadas, que podem ser utilizadas em diversos cenários, tais como em publicidade e propaganda, revelação fotográfica digital e aplicações cinematográficas e televisivas, dentre outras, assim como para fins de entretenimento.

A utilização do filtro de suavização desenvolvido e proposto nesta dissertação evidenciou a melhoria da nitidez nas imagens abstratas resultantes, em relação ao filtro FBL. Entretanto, o filtro proposto neste trabalho apresenta desvantagem, quando comparado ao filtro FBL, ao ser utilizado em aplicações que objetivam a abstração de imagens em tempo real. Uma vez que o filtro AFBL incorpora ao processo a filtragem do LoG e cálculos adicionais durante o cômputo dos pesos, o tempo de execução do processo de suavização se torna maior do que o tempo gasto pelo filtro FBL.

5.3 Trabalhos Futuros

Durante a pesquisa bibliográfica, constatou-se que a técnica DoG possui resultado aproximado ao LoG. Desta forma, com o intuito de remover o processo de filtragem do LoG no filtro de suavização FABL, propõe-se como trabalho futuro viabilizar a substituição do resultado do filtro LoG pelo resultado do filtro FDoG (utilizado para extração dos contornos dos elementos significantes da imagem).

Adicionalmente, pesquisas adicionais poderão ser conduzidas a fim de estender a aplicação desta abordagem a vídeos digitais, além de viabilizar a substituição do ETF, proposta por Kyprianidis e Döllner (2008), por um método similar de menor custo computacional.

No que concerne ao objetivo de abstração das informações não relevantes da imagem, propõe-se também como trabalho futuro a condução de pesquisas relacionadas à questão da atenção visual. Algumas

abordagens encontradas na revisão bibliográfica (e.g. DeCarlo e Santella, 2002; Zhao et al., 2009) utilizam mapas de saliência para identificar os elementos importantes na imagem. Nestas abordagens, o nível de abstração é controlado para remover mais informação em regiões menos salientes e, assim, direcionar a atenção visual do observador para os elementos importantes da imagem.

Ainda no que concerne ao objetivo de abstração das informações, outra linha de pesquisa que pode ser investigada refere-se à simplificação dos formatos dos elementos utilizando técnicas de vetorização de imagens.

Os módulos de ajuste de cores ora propostos para melhorar o aspecto visual das imagens geradas a partir da abordagem proposta são de caráter subjetivo e realizados pelo usuário do sistema, o que caracterizam a abordagem proposta como semi-automática caso tais módulos sejam (opcionalmente) incorporados ao processo de geração de imagens ao estilo de cartuns. Outra vertente de investigação refere-se a estratégias para a automatização do ajuste de cores por meio de métodos de aprendizagem de máquina, objetivando imitar estilos de imagens produzidas por cartunistas.

Por fim, também se propõe como trabalho futuro a investigação de outros métodos para a validação do processo de abstração de imagens. A validação deste processo normalmente envolve experimentos psicológicos, os quais, segundo Winnemöeller (2006), são difíceis de serem planejados, consomem muito tempo de execução e necessitam de uma análise cuidadosa dos dados. Novas formas de validação para a área de NPR poderão ser concebidas, buscando-se inspiração em outras áreas da Computação, e.g., Interação Homem-Máquina.

Referências

BARASH, D. *Bilateral filtering and anisotropic diffusion: towards a unified viewpoint*. Relatório Técnico da Hewlett-Packard, HPL-2000-18(R.1), 2000.

BOYKOV, Y., JOLLY, M.-P. Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: IEEE INTERNETIONAL CONFERENCE ON COMPUTER VISION, 3., 2001, Vancouver. *Anais...* Vancouver: British Columbia, 2001. p. 105–112.

BOYKOV, Y.; KOLMOGOROV, V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Washington, v. 26, n. 9, p. 1124-1137, 2004.

BRENNAN, S.E. Caricature generator: the dynamic exaggeration of faces by computer. In: *Leonardo 18*, New York, n. 3, p. 170–178, 1985.

BRITTO NETO, L. S.; CARVALHO, B. M. Message in a bottle: stylized rendering of sand movies. In: BRAZILIAN SYMPOSIUM ON COMPUTER GRAPHICS AND IMAGE PROCESSING, 2007, Belo Horizonte. Washington: IEEE CS, 2007. p. 11–18.

CHEN, H. et al. Example-based composite sketching of human portraits. In: INTERNATIONAL SYMPOSIUM ON NON-PHOTOREALISTIC ANIMATION AND RENDERING, 3., 2004, Annecy. *Anais…* Annecy: ACM, 2004. p. 95–153.

COHEN-OR, D. et al. Color harmonization. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 2006, Boston. *Anais…* New York: ACM, 2006. p. 624–630.

COMANICIU, D.; MEER P. Mean Shift: a robust approach toward feature

space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Washington, v. 24, n. 5, p. 603–619, 2002.

CURTIS, J. C. et al. Computer-generated watercolor. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 24., 1997, Los Angeles. *Anais…* Los Angeles: ACM, 1997. p. 421–430.

DECARLO, D.; SANTELLA, A. Stylization and abstraction of photographs. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 29., 2002, San Antonio. *Anais…* San Antonio: ACM, 2002. p. 769-776.

DECAUDINE, P. *Cartoon-looking rendering of 3D-scenes*. Relatório Técnico, INRIA *2919*, 1996.

DENG, Y.; MANJUNATH, B.S.; SHIN, H. Color image segmentation. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION, 1999, Fort Colins. *Anais…* Los Alamitos: IEEE CS, 1999. p. 2446.

DINET, É.; ROBERT-INACIO, F. Color median filtering: a spatially adaptive filter. In: IMAGE AND VISION COMPUTING NEW ZEALAND, 2007, Hamilton. *Anais…* Los Alamitos: IEEE CS, 2007. p. 71–76.

DONITSON, P. P. Quantitative evaluation of edge preserving noise-smoothing filter. *IEEE*, Lincoln, v. 3, p. 1590–1591, 1996..

ESTRADA, F.J.; JEPSON, A.D. Quantitative evaluation of a novel image segmentation algorithm. In: IEEE COMPUTER VISION AND PATTERN RECOGNITION, 2005, San Diego. *Anais…* Los Alamitos: IEEE CS, 2005. p. 1132-1139.

FELZENSZWALB, P.; HUTTENLOCHER, D. Image segmentation using local variation. In: IEEE COMPUTER VISION AND PATTERN RECOGNITION, 1998, Santa Barbara. *Anais…* Los Alamitos: IEEE CS, 1998. p. 98–104.

FINKELSTEIN, A.; RANGE, M. Image mosaics. In: RASTER IMAGING AND DIGITAL TYPOGRAPHY, 1375., 1998, Saint-Malo. *Anais…* Saint-Malo: Springer-Verlag, 1998. p. 11–22.

G. RAMPONI. A rational edge-preserving smoother. In: INTERNATIONAL CONFERENCE ON IMAGE PROCESSING, 1995, Washington. *Anais…* Los Alamitos: IEEE CS, 1995. p. 151–154.

GONZALEZ, C. R.; WOODS, E. R. *Digital image processing*. 3.ed. Upper Saddle River: Prentice Hall, 2008.

GOOCH, B. *Artistic vision: automatic digital painting using computer vision algorithms*. 2001. 41f. Dissertação (Mestrado em Ciência da Computação) –The University of Utah – Utah, USA, 2001.

GOOCH, B. *Human facial illustrations*: creation and evaluation using behavioral studies and functional magnetic resonance. 2003. 102f. Tese (Doutorado em Ciência da Computação) – The University of Utah – Utah, USA, 2003.

GOOCH, B.; GOOCH, A. *Non-photorealistic rendering*. Natick: A. K. Peters/CRC, 2001.

HAEBERLI, P. Paint by numbers: abstract image representations. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 4., 1990, Dallas. *Anais…* Dallas: ACM, 1990. p. 207–214.

HALPER, N. et al. Psychology and non-photorealistic rendering: the beginning of a beautiful relationship. In: MENSCH & COMPUTER, 2003, Stuttgart. *Anais…* Stuttgart: Teubner, 2003. p. 277–286.

HAUTAMÄKI, V. *Efficient color quantization by hierarchical clustering algorithms*. 2005. 86f. Dissertação (Mestrado em Ciência da Computação) – University of Joensuu – Joensuu, FIN, 2005.

HERTZMANN, A. Painterly rendering with curved brush strokes of multiple sizes. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1998, Orlando. *Anais…* New York: ACM, 1998. p. 453–460.

HERTZMANN, A. Non-Photorealistic Rendering and the Science of Art. In: INTERNATIONAL SYMPOSIUM ON NON-PHOTOREALISTIC ANIMATION AND

RENDERING, 2010, Annecy. *Anais…* New York: ACM, 2010. 147–157.

HOLMSTRÖM, J. *Growing neural gas*. 2002. 38f. Dissertação (Mestrado em Ciência da Computação) – Uppsala University – Uppsala, 2002.

IANEVA, T. I. *Detecting cartoons*: a case study in automatic video-genre classification. 2003. 166f. Trabalho de Pesquisa de Doutorado – Universitat de València – Valencia, 2003.

INOUÉ, S; SPRING, K. R. *Video microscopy*: the fundamentals. 2.ed. New York: Plenum, 1997.

JAIN, N. Noise cancellation using adaptive trilateral filter. *International Journal of Recent Trends in Engineering*, v.1, n.1, p. 329–333, 2009.

KANG, H.; LEE, S. Shape-simplifying Image Abstraction. *Computer Graphics Forum*, v. 27, n. 7, p. 1773–1780, 2008.

KANG, H.; LEE, S.; CHUI, C. Flow-based image abstraction. *Visualization and Computer Graphics*, v. 15, n. 1, p. 62–76, 2009.

KIMOTHI, S. K. *The uncertainty of measurements*: physical and chemical metrology impact and analysis. Milwaukee: ASQ Quality, 2002.

KOHONEN, T. The self-organizing map. *IEEE*, New York, v.78, n. 9., p. 1464–1480, 1990.

KURLANDER, D.; SKELLY, T.; SALESIN, D. Comic chat. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1996, New Orleans. *Anais…* New York: ACM, 1996. p. 225–236.

KUWAHARA, M. et al. *Digital processing of biomedical images*. New York: Plenum, 1976.

KYPRIANIDIS, J. E.; DÖLLNER, J. Image abstraction by structure adaptive filtering. In: THEORY AND PRACTICE OF COMPUTER GRAPHICS, 2008, Manchester. *Anais…* [S.l]: Eurographics Association, 2008. p. 51–58.

KYPRIANIDIS, J. E.; KANG, H.; DÖLLNER, J. Image and video abstraction by anisotropic kuwahara filtering. *Computer Graphics Forum*, Munich, v.

28, n. 7, p. 1955-1963, 2009.

LANSDOWN, J.; SCHOFIELD, S. Expressive rendering: a review of nonphotorealistic techniques. *IEEE Computer Graphics and Applications*, Los Amitos, v. 15, n. 3, p. 29–37, 1995.

LEVINE, D. M.; BERENSON, M. L.; STEPHAN, D. *Estatística*: teoria e aplicações. Rio de Janeiro: LTC, 2000.

LEWIS, J.-P. Texture synthesis for digital painting. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1984, Minneapolis. *Anais…* New York: ACM, 1984. p. 245–251.

MARKOSIAN, L. et al. Real-time nonphotorealistic rendering. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1997, Los Angeles. *Anais…* New York: ACM, 1997. p. 415–420.

MARR, D.; HILDRETH, E. Theory of edge detection. *Royal Society of London*, London, v. 207, n. 1167, p. 187-217, 1980.

MEDHI, I.; PATABANDHIGE, C. R.; TOYAMA, K. Cartoon generation for text-free user interfaces. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION, 2007, Rio de Janeiro. *Anais…* Los Alamitos: IEEE CS, 2007.

MEIER, B. Painterly rendering for animation. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1996, New Orleans. New York: ACM, 1996. p. 477–484.

MEZEI, L.; PUZIN, M.; CONROY, P. Simulation of patterns of nature by computer graphics. *Information Processing*, v. 74, p. 861-865, 1974.

MIKOLOV, T. Color reduction using k-means clustering. In: CENTRAL EUROPEAN SEMINAR ON COMPUTER GRAPHICS, 2007, Slovakia. *Anais…* [S.l.: s.n.], 2007.

NEOH, H. S.; HAZANCHUK, A. Adaptive edge detection for real-time video processing using FPGAs. In: GLOBAL SIGNAL PROCESSING EXPO AND CONFERENCE, 2004, Santa Clara. *Anais…* [S.l.: s.n.], 2004.

OLMOS, A.; KINGDOM, F. A. A. Automatic non-photorealistic rendering through soft-shading removal: a colour-vision approach. In: INTERNATIONAL CONFERENCE ON VISION, VIDEO AND GRAPHICS, Edinburgh, 2005. *Anais…* [S.l.]: Eurographics Association, 2005. p. 203–208.

PAPARI, G.; PETKOV, N.; CAMPISI, P. Artistic Edge and Corner Enhancing Smoothing. *IEEE Transactions on Image Processing*, v. *16*, n. 10, p. 2449–2462, 2007.

PARKER, J.R. Extracting vectors from raster images. In: COMPUTERS & GRAPHICS, 1., 1988., Great Britain. *Anais…* Great Britain: Pergamon, 1988. p. 75–79.

PERONA, P.; MALIK, J. Scale-Space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Washington, v. 12, n. 7, p. 629–639, 1990.

PHAM, T. Q.; VLIET, L. J. V. Separable bilateral filtering for fast video preprocessing. In: IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA & EXPO, 2005, Amsterdam. *Anais…* Los Alamitos: IEEE CS, 2005.

PITAS, I.; TSAKALIDES, P. Multivariate ordering in color image filtering. *IEEE Transactions on Circuits and Systems for Vídeo Technology*, v. 1, n. 3, p. 247–259, 1991.

PORIKLI, F. Constant time O(1) bilateral filtering. In: COMPUTER VISION AND PATTERN RECOGNITION, 2008, Anchorage. *Anais…* Los Alamitos: IEEE CS, 2008. p. 1–8.

REDMOND N.; DINGLIANA J. Adaptive abstraction of 3D scenes in real-time. In: EUROGRAPHICS 2007 CONFERENCE, 2007, Prague. *Anais…* [S.l.: s.n.], 2007. p. 77–80.

RZESZUTEK, R.; EL-MARAGHI, T.; ANDROUTSOS, D. Interactive rotoscoping through scale-space random walks. In: IEEE INTERNATIONAL CONFERENCE ON MULTIMEDIA AND EXPO, 2009, New York. *Anais…* Los

Alamitos: IEEE CS, 2009. p.1334-1337.

SAINT-MARC, P.; CHEN, J.-S.; MEDIONI, G. Adaptive smoothing: a general tool for early vision. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, v. 13, n. 6, p. 514-529, 1991.

SAITO, T.; TAKAHASHI, T. Comprehensible rendering of 3-D shapes. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 4., 1990, Dallas. *Anais…* New York: ACM, 1990. p. 197–206.

SANTELLA, A. *The art of seeing:* visual perception in design and evaluation of non-photorealistic rendering. 2005. Tese (Doutorado em Ciência da Computação) – Rutgers University – New Jersey, 2005.

SCHWARZ, M. W.; COWAN, W. B.; BEATTY, J. C. An experimental comparison of RGB, YIQ, LAB, HSV, and opponent color models. *ACM Transactions on Graphics*, New York, v. 6, n. 2, p. 123–158, 1987.

SELIGMANN, D. D.; FEINER, S. Automated generation of intent-based 3D illustrations, In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1991, Las Vegas. *Anais…* New York: ACM, 1991. p. 123–132.

SHI, J.; MALIK, J. Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Inteligence*, Washington, v. 22, n. 8, p. 888–905, 2000.

SMITH, A. R. *Paint*. Memorando Técnico do Instituto de Tecnologia de Nova Iorque, Technical Memo 7, 1978a.

SMITH, A. R. Color gamut transform pairs. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1978, Atlanta. *Anais…* New York: ACM, 1978b. p. 12–19.

SMITH, S. M.; BRADY, J. M. SUSAN – a new approach to low level image processing. *International Journal of Computer Vision*, Hingham, v. 23, n. 1, p. 45–78, 1997.

STRASSMAN, S. Hairy brushes. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1986, Dallas. *Anais…* New York: ACM, 1986. p. 225–232.

STROTHOTTE, T.; SCHLECHTWEG, S. *Non-photorealistic computer graphics*: modeling, rendering and animation. San Francisco: Morgan Kaufmann, 2002.

THOMA, J. *Non-photorealistic rendering techniques for real-time character animation*. 2002. Dissertação (Mestrado em Ciência da Computação) – RWTH Aachen University – Aachen, 2002.

TOOTHAKER, L. E. Multiple comparison procedures. *Sage University Paper series on Quantitative Applications in the Social Sciences*, Newbury Park, n. 89, p. 1–95, 1993.

TOKUMARU, M.; MURANAKA, N.; IMANISHI, S. Color design support system considering color harmony. In: IEEE INTERNATIONAL CONFERENCE ON FUZZY SYSTEMS, 2002, Hawaii. *Anais…* [S.l.]: IEEE, 2002. p. 378–383.

TOMASI, C.; MANDUCHI, R. Bilateral filtering for gray and color images. In: INTERNATIONAL CONFERENCE ON COMPUTER VISION, 1998, Bombay. *Anais…* Los Alamitos: IEEE CS, 1998. p. 839–846.

ULICHNEY, R. Dithering with blue noise. *IEEE*, v. 76, n. 1, p. 56-79, 1988.

WALLRAVEN, C. et al. Evaluation of real-world and computer-generated stylized facial expressions. *ACM Transactions on Applied Perception*, New York, v. 4, n. 3, 2007.

WANG, J. et al. Video tooning. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 2004, Los Angeles. *Anais…* New York: ACM, 2004. p. 574-583.

WEISS, B. Fast median and bilateral filtering. *ACM Transactions on Graphics*, New York, v. 25, n. 3, p. 519–526, 2006.

WEN, F. et al. Color sketch generation. In: INTERNATIONAL SYMPOSIUM

ON NON-PHOTOREALISTIC ANIMATION AND RENDERING, 2006, Annecy. *Anais…* New York: ACM, 2006. p. 47–54.

WHEELER, A. J.; GANJI, A. R. *Introduction to engineering experimentation*. 3.ed. Upper Saddle River: Prentice Hall, 2009.

WHITTED, T. Anti-aliased line drawing using brush extrusion. *ACM Transactions on Computer Graphics*, New York, v. 17, n. 3, p. 151–156, 1983.

WINKENBACH, G; SALESIN, D. H. Computer-generated pen-and-ink illustration. In: ANNUAL CONFERENCE ON COMPUTER GRAPHICS AND INTERACTIVE TECHNIQUES, 1994, Los Angeles. *Anais…* New York: ACM, 2004. p. 91–100.

WINNEMÖLLER, H. *Perceptually-motivated non-photorealistic graphics*. 2006. Tese (Doutorado em Ciência da Computação) – Northwestern University – Illinois, 2006.

WINNEMÖLLER, H.; OLSEN, S.C.; GOOCH, B. Real-time video abstraction. *ACM Transactions on Graphics*, New York, v. 25, n. 3, p. 1221–1226, 2006.

ZENG, K. et al. From image parsing to painterly rendering. *ACM Transactioins on Graphics*, New York, v. 29, n. 1, p. 1–11, 2009.

ZHANG, B.; ALLEBACH, J. P. Adaptive bilateral filter for sharpness enhancement and noise removal. *IEEE Transactions on Image Processing*, v. 17, n. 5, p. 664–678, 2008.

ZHAO, H. et al. Real-time saliency-aware video abstraction. *The Visual Computer*, Secaucus, v. 25, n. 11, p. 973–984, 2009.

Apêndice A

Amostra de Resultados

No Quadro A.1, são mostradas imagens de resultados extraídos dos artigos citados na revisão bibliográfica de abstração de imagens.

Quadro A.1 – Exemplos de resultados de imagens abstraídas por abordagens citadas na revisão bibliográfica.

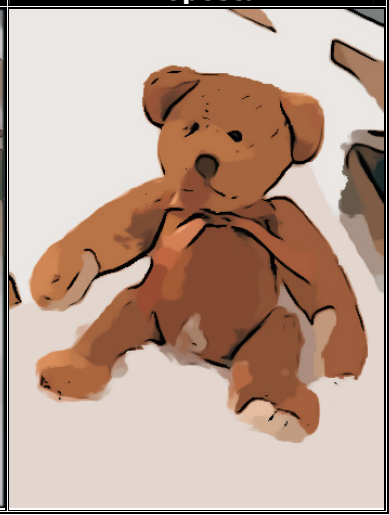
Imagem Original	DeCarlo e Santella (2002)	Abordagem Proposta
Imagem Original	**Olmos e Kingdom (2005)**	**Abordagem Proposta**
		

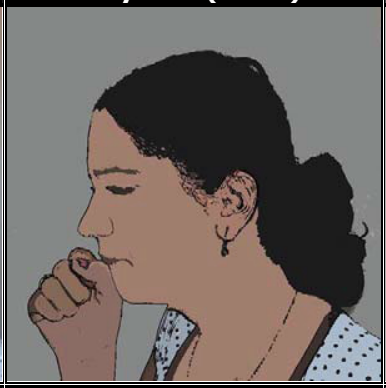
Imagem Original	Wen et al. (2006)	Abordagem Proposta
Imagem Original	**Medhi, Patabandhige e Toyama (2007)**	**Abordagem Proposta**
		
Imagem Original	**Winnemöller, Olsen e Gooch (2006)**	**Abordagem Proposta**

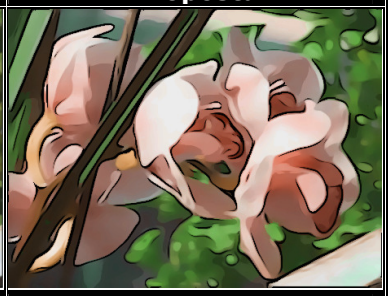
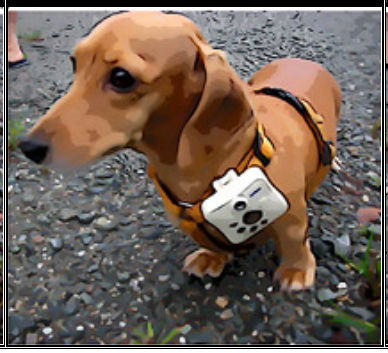
Imagem Original	Redmond e Dingliana (2007)	Abordagem Proposta
Imagem Original	**Papari, Petkov e Campisi (2007)**	**Abordagem Proposta**
		
Imagem Original	**Kyprianidis e Döllner (2008)**	**Abordagem Proposta**

Imagem Original	Kang, Lee e Chui (2009)	Abordagem Proposta
Imagem Original	**Kyprianidis, Kang e Döllner (2009)**	**Abordagem Proposta**
Imagem Original	**Zhao et al. (2009)**	**Abordagem Proposta**
		

Apêndice B

Variações do Filtro Bilateral

Este apêndice contém o detalhamento de três variações do filtro *Bilateral*, originalmente proposto por Tomasi e Manduchi (1998), as quais foram utilizadas como referências nesta dissertação, a saber: (i) filtro *Bilateral Separável*, (ii) filtro *Bilateral Baseado em Fluxo* e (iii) filtro *Bilateral Adaptativo*.

B.1 Filtro Bilateral Separável

A implementação separável do filtro *Bilateral* foi proposta por Pham e Vliet (2005), tendo por finalidade a obtenção, em menor tempo de execução, de um resultado aproximado àquele produzido pelo filtro *Bilateral*. Nesta abordagem, a imagem é filtrada separadamente nas duas dimensões espaciais (x e y). A filtragem *Bilateral Separável* é dividida em duas etapas, a saber: (i) a aplicação do filtro *Bilateral* horizontal, (ii) e a aplicação do filtro *Bilateral* vertical.

Inicialmente, aplica-se o filtro *Bilateral* horizontal para cada linha de pixels da imagem, cujo resultado em um determinado pixel é obtido apenas a partir dos pixels de sua vizinhança na referida linha. Por fim, aplica-se ao resultado desta filtragem o filtro *Bilateral* vertical, que opera de forma análoga à filtragem horizontal, porém aplicado às colunas de pixels da imagem.

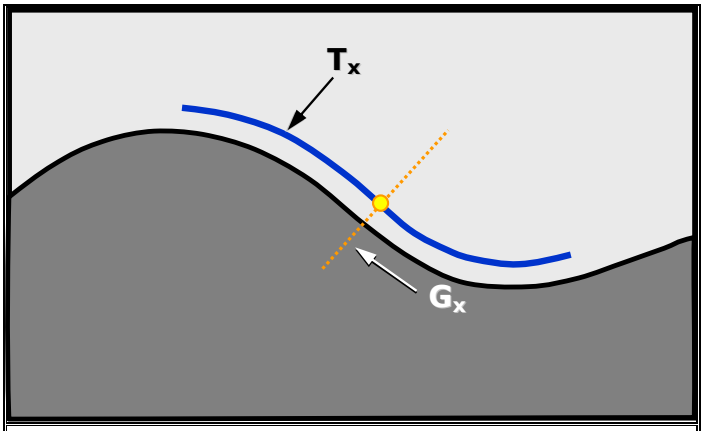
Desta forma, o resultado da aplicação deste filtro em um pixel utilizando uma vizinhança de MxN, será computado por meio de M+N pixels (M pixels na horizontal e, posteriormente, N pixels na vertical), ao invés de M*N pixels (todos os pixels contidos na matriz da vizinhança do pixel).

Segundo resultados experimentais obtidos por Pham e Vliet (2005), o filtro *Bilateral Separável* permite obter resultados próximos àqueles produzidos pelo filtro *Bilateral* convencional, não apresentando perda significativa na qualidade da suavização, enquanto reduz a complexidade computacional.

B.2 Filtro Bilateral Baseado em Fluxo

O filtro *Bilateral Baseado em Fluxo* utilizado nesta dissertação, proposto por Kang, Lee e Chui (2009), utiliza a técnica da filtragem *Bilateral Separável* (Pham e Vliet, 2005). Entretanto, ao invés de ser aplicado nas direções x (horizontal) e y (vertical), o é na direção dos vetores tangente e perpendicular à direção das bordas, a qual é obtida utilizando um campo vetorial chamado de *Fluxo de Tangente das Bordas* (ETF).

Para o cômputo do ETF, inicialmente são obtidos os vetores perpendiculares ao mapa de gradiente computado utilizando os operadores de *Sobel*. Em seguida, estes vetores são suavizados para obtenção do fluxo dos vetores tangentes das bodas. Uma vez computado o ETF, o filtro *Bilateral* é aplicado separadamente na direção da tangente da borda (curva Tx da Figura B.1) e na direção do gradiente da borda (reta Gx da Figura B.1).



Fonte: Adaptado de Kang, Lee e Chui (2009)

Figura B.1 - Direção tangente (Tx) e perpendicular (Gx) ao fluxo de tangente de bordas no ponto X.

Assim, o filtro bilateral divide-se em duas aplicações separadas e consecutivas, orientadas pela direção das bordas, o que permite priorizar a suavização ao longo das tangentes das bordas, a fim de reduzir a complexidade do formato dos elementos (Kang, Lee e Chuí, 2009).

B.3 Filtro Bilateral Adaptativo

O filtro *Bilateral Adaptativo*, proposto por Zhang e Allebach (2008), é uma solução unificada para remover ruído e aguçar a imagem. Para isto, foi proposta uma alteração na fórmula da gaussiana do domínio radial visando o deslocamento do centro (valor médio), assim como uma alteração no valor do sigma (desvio-padrão).

A estratégia utilizada por este filtro resume-se em alterar o comportamento do filtro *Bilateral* nas regiões de bordas da imagem, visando: (i) diminuir o desvio padrão da gaussiana nestas regiões e, conseqüentemente, reduzir a intensidade da suavização; e (ii) priorizar os pixels que se encontram fora da região de transição das bordas por meio do deslocamento do centro da gaussiana na direção contrária à borda.

A função de ponderação radial aplicada ao pixel p para um vizinho s é dada pela Equação (B.1):

$$G(p,s) = e^{\left(-\frac{(I_s - I_p - \zeta(p))^2}{2\sigma^2(p)} \right)} \quad \text{(B.1)}$$

na qual I_p e I_s são os valores correspondentes dos pixels p e s, e $\zeta(p)$ e $\sigma^2(p)$ são funções que retornam os valores do deslocamento da média e do sigma da gaussiana, respectivamente. Estas funções utilizam a resposta do operador *Laplaciano da Gaussiana* (Marr e Hildreth, 1980) no pixel p.

Os coeficientes do operador *Laplaciano da Gaussiana* (LoG) corresponde à segunda derivada da função gaussiana no domínio espacial, obtido pela Equação (B.2):

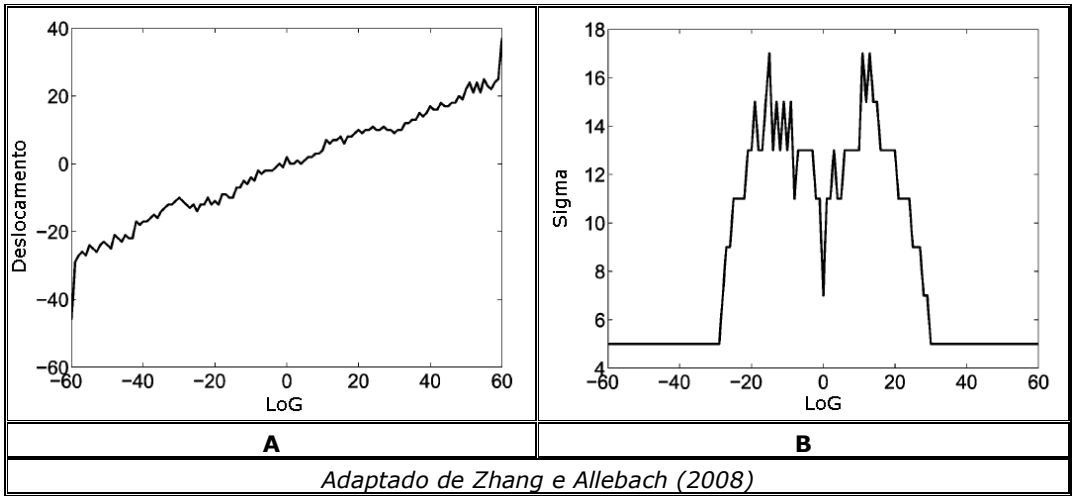
$$LoG(x,y) = \frac{-1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right) e^{\left(-\frac{x2+y2}{2\sigma^2}\right)} - C \quad \text{(B.2)}$$

na qual C é uma constante que iguala a soma dos coeficientes a zero. Para uma máscara de tamanho (2N+1)x(2N+1), C é definida por:

$$C = \frac{1}{(2N+1)^2} \sum_{x=-N}^{N} \sum_{y=-N}^{N} \frac{-1}{\pi\sigma^4}\left(1 - \frac{x^2+y^2}{2\sigma^2}\right) e^{\left(-\frac{x2+y2}{2\sigma^2}\right)} \quad \text{(B.3)}$$

A magnitude da resposta obtida pelo LoG próximo às bordas da imagem é alta, sendo baixa em regiões suaves e no centro da borda. Outra propriedade importante do LoG é o sinal (+/-) da resposta, que indica quando o pixel encontra-se no lado claro da borda (+) ou no lado escuro da borda (-).

Na abordagem proposta por Zhang e Allebach (2008), os valores dos sigmas ($\sigma^2(p)$) e dos deslocamentos ($\zeta(p)$) para as diferentes respostas do LoG foram obtidos por meio de um processo de treinamento. Estes valores são ilustrados na Figura B.2.



Adaptado de Zhang e Allebach (2008)

Figura B.2 - Valores treinados a partir do LoG: (A) Valores do deslocamento da função; (B) Valores dos sigmas da função gaussiana radial.

Em função da resposta do LoG, os valores do sigma e do deslocamento obtidos são inseridos na Equação (B.1).

Apêndice C

Conversões entre Espaços de Cores

Este apêndice contém as equações utilizadas na conversão entre os espaços de cores utilizados nesta dissertação, a saber: RGB, HSV e CIE L*a*b*. Para conversão entre RGB e CIE L*a*b*, é utilizado o espaço de cor CIE XYZ.

C.1 Conversões entre RGB e HSV

Considerando valores normalizados dos canais de cores RGB, i.e., valores restritos ao intervalo contínuo [0..1], os valores dos canais H, S e V são obtidos pelas equações a seguir:

$$V = \max(R,G,B) \tag{C.1}$$

$$S = \begin{cases} \dfrac{V - \min(R,G,B)}{V}, & \text{se } V \neq 0 \\ 0, & \text{se } V = 0 \end{cases} \tag{C.2}$$

$$H = \begin{cases} \textit{indefinido}, & \text{se } V = \min(R,G,B) \\ 60 \times \dfrac{G-B}{V - \min(R,G,B)}, & \text{se } V = R \\ 60 \times \dfrac{B-R}{V - \min(R,G,B)} + 120, & \text{se } V = G \\ 60 \times \dfrac{R-G}{V - \min(R,G,B)} + 240, & \text{se } V = B \end{cases} \tag{C.3}$$

Na hipótese de H assumir valores negativos após a conversão supracitada, este deve ser somado a 360. Na conversão de HSV para RGB, caso $S = 0$, todos os valores de R, G e B tornam-se iguais a V; em caso contrário, são obtidos a partir da Eq. C.4:

$$(R,G,B) = \begin{cases} (V,t,p), & \text{se } H_i = 0 \\ (q,V,p), & \text{se } H_i = 1 \\ (p,V,t), & \text{se } H_i = 2 \\ (p,q,V), & \text{se } H_i = 3 \\ (t,p,V), & \text{se } H_i = 4 \\ (V,p,q), & \text{se } H_i = 5 \end{cases} \tag{C.4}$$

na qual:

$$\begin{aligned} H_i &= \left\lfloor \frac{H}{60} \right\rfloor \bmod 6, \\ f &= \frac{H}{60} - H_i, \\ p &= V(1-S), \\ q &= V(1-fS), \text{ e} \\ t &= V(1-(1-f)S) \end{aligned} \tag{C.5}$$

C.2 Conversões entre RGB e CIE XYZ

As conversões entre estes espaços de cores são obtidas pelas transformações expressas pelas Equações C.6 e C.7.

$$\begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} = \begin{pmatrix} 0{,}4124564 & 0{,}3575761 & 0{,}1804375 \\ 0{,}2126729 & 0{,}7151522 & 0{,}0721750 \\ 0{,}0193339 & 0{,}1191920 & 0{,}9503041 \end{pmatrix} \times \begin{pmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{pmatrix} \tag{C.6}$$

$$\begin{pmatrix} \mathbf{R} \\ \mathbf{G} \\ \mathbf{B} \end{pmatrix} = \begin{pmatrix} 3{,}2404542 & -1{,}5371385 & -0{,}4985314 \\ -0{,}9692660 & 1{,}8760108 & 0{,}0415560 \\ 0{,}0556434 & -0{,}2040259 & 1{,}0572252 \end{pmatrix} \times \begin{pmatrix} \mathbf{X} \\ \mathbf{Y} \\ \mathbf{Z} \end{pmatrix} \tag{C.7}$$

C.3 Conversões entre CIE XYZ e CIE L*a*b*

A conversão do espaço CIE XYZ para CIE L*a*b* é obtida por:

$$\begin{aligned} L^* &= 116 \times f(Y) - 16 \\ a^* &= 500 \times \left[f\left(\frac{X}{0{,}9504}\right) - f(Y) \right] + 128 \\ b^* &= 200 \times \left[f(Y) - f\left(\frac{Z}{1{,}0888}\right) \right] + 128 \end{aligned} \tag{C.8}$$

nas quais:

$$f(t) = \begin{cases} \sqrt[3]{t}, & \text{se } t > 0{,}008856 \\ 7{,}787t + 0{,}137931, & \text{se } t \leq 0{,}008856 \end{cases} \tag{C.9}$$

Por fim, a conversão inversa é obtida por:

$$\begin{aligned} Y &= \begin{cases} f_y^3, & \text{se } f_y > 0{,}206896 \\ \left(f_y - 0{,}13793\right) \times 0{,}128418, & \text{se } f_y \leq 0{,}206896 \end{cases} \\ X &= \begin{cases} 0{,}9504 \times f_x^3, & \text{se } f_x > 0{,}206896 \\ \left(f_x - 0{,}13793\right) \times 0{,}122049, & \text{se } f_x \leq 0{,}206896 \end{cases} \\ Z &= \begin{cases} 1{,}0888 \times f_Z^3, & \text{se } f_Z > 0{,}206896 \\ \left(f_Z - 0{,}13793\right) \times 0{,}139822, & \text{se } f_Z \leq 0{,}206896 \end{cases} \end{aligned} \tag{C.10}$$

nas quais:

$$\begin{aligned} f_y &= \frac{(L^* + 16)}{116}, \\ f_x &= f_y + \frac{a^*}{500}, \\ f_Z &= f_y - \frac{b^*}{200}. \end{aligned} \tag{C.11}$$

Apêndice D

Testes de Tukey - Calibração de Parâmetros

Os testes apresentados neste apêndice representam as comparações feitas entre as diferentes configurações de parâmetros para cada aplicação de *software* calibrada. Estes testes foram conduzidos utilizando um nível de confiança de 95%, tendo sido aplicados após o teste F ANOVA fator único ter evidenciado haver diferenças entre as configurações de parâmetros, por meio dos dados obtidos dos participantes da fase de calibração de parâmetros.

D.1 BeFunky

Na Tabela D.1, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração do *software* BeFunky.

Tabela D.1 – Configurações de parâmetros do BeFunky.

Configuração	Sketch Detail	Color Detail
1	6	20
2	6	35
3	6	50
4	12	20
5	12	35
6	12	50

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.2.

Tabela D.2 – Grupos de similaridades obtidos pelo teste de Tukey para o *software* BeFunky.

Configuração	Média estimada	Grupos			
1	0,588	A			
2	0,275		B		
4	0,175		B	C	
3	0,106			C	D
5	0,078			C	D
6	0,031				D

D.2 CaricatureSoft

Na Tabela D.3, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração do *software* CaricatureSoft.

Tabela D.3 – Configurações de parâmetros do CaricatureSoft.

Configuração	Smooth	Color Detail	Sketch Detail
1	1	1	3
2	1	1	6
3	1	3	3
4	1	3	6
5	1	5	3
6	1	5	6
7	2	1	3
8	2	1	6
9	2	3	3
10	2	3	6
11	2	5	3
12	2	5	6

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.4.

Tabela D.4 – Grupos de similaridades obtidos pelo teste de Tukey para o *software* CaricatureSoft.

Configuração	Média estimada	Grupos			
2	0,732	A			
8	0,544		B		
10	0,363			C	
4	0,305			C	
1	0,100				D
7	0,050				D
12	0,042				D
9	0,035				D
3	0,033				D
6	0,033				D
5	0,011				D
11	0,000				D

D.3 LiangZhu

Na Tabela D.5, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração do *software* LiangZhu.

Tabela D.5 – Configurações de parâmetros do LiangZhu.

Configuração	Scalar	Sleek	Pencil Pressure
1	2	1	80
2	2	1	90
3	2	9	80
4	2	9	90
5	5	1	80
6	5	1	90
7	5	9	80
8	5	9	90
9	8	1	80
10	8	1	90
11	8	9	80
12	8	9	90

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.6.

Tabela D.6 – Grupos de similaridades obtidos pelo teste de Tukey para o *software* LiangZhu.

Configuração	Média estimada	Grupos			
9	0,650	A			
5	0,579	A	B		
10	0,436		B		
11	0,211			C	
6	0,210			C	
12	0,177			C	D
7	0,138			C	D
8	0,080			C	D
1	0,069			C	D
2	0,000				D
3	0,000				D
3	0,000				D

D.4 MSU Cartoonizer

Na Tabela D.7, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração do *software* MSU Cartoonizer.

Tabela D.7 – Configurações de parâmetros do MSU Cartoonizer.

Configuração	N. Edges	N. Segments	Saturation
1	1	1	4
2	1	1	6
3	1	1	8
4	1	9	4
5	1	9	6
6	1	9	8
7	4	1	4
8	4	1	6
9	4	1	8
10	4	9	4
11	4	9	6
12	4	9	8

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.8.

Tabela D.8 – Grupos de similaridades obtidos pelo teste de Tukey para o *software* MSU Cartoonizer.

Configuração	Média estimada	Grupos				
6	0,681	A				
5	0,500	A	B			
3	0,425		B	C		
4	0,307			C	D	
2	0,242			C	D	E
1	0,208				D	E
12	0,207				D	E
11	0,122				D	E
8	0,100					E
9	0,100					E
10	0,094					E
7	0,077					E

D.5 KusoCartoon

Na Tabela D.9, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração do *software* KusoCartoon.

Tabela D.9 – Configurações de parâmetros do KusoCartoon.

Configuração	Detail	Color	Saturation
1	2	1	2
2	2	1	5
3	2	1	8
4	2	3	2
5	2	3	5
6	2	3	8
7	2	9	2
8	2	9	5
9	2	9	8
10	9	1	2
11	9	1	5
12	9	1	8
13	9	3	2
14	9	3	5
15	9	3	8
16	9	9	2
17	9	9	5
18	9	9	8

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.10.

Tabela D.10 – Grupos de similaridades obtidos pelo teste de Tukey para o *software* MSU KusoCartoon.

Configuração	Média estimada	Grupos							
9	0,525	A							
6	0,350	A	B						
8	0,300		B	C					
18	0,288		B	C	D				
15	0,260		B	C	D	E			
17	0,242		B	C	D	E	F		
7	0,230		B	C	D	E	F	G	
5	0,186		B	C	D	E	F	G	H
12	0,150		B	C	D	E	F	G	H
3	0,143			C	D	E	F	G	H
16	0,090				D	E	F	G	H
2	0,081					E	F	G	H
4	0,050						F	G	H
13	0,042						F	G	H
14	0,040						F	G	H
11	0,030							G	H
10	0,025								H
1	0,020								H

D.6 Abordagem Proposta

Na Tabela D.11, são apresentadas as diferentes configurações de parâmetros utilizadas para a calibração da abordagem proposta.

Tabela D.11 – Configurações de parâmetros da abordagem proposta.

Configuração	Mean Shift(Sigma Radial)	FDoG(Sigma1)	Saturação(Gama)
1	3	1	0,8
2	3	1	1
3	3	1	1,2
4	3	2	0,8
5	3	2	1
6	3	2	1,2
7	6	1	0,8
8	6	1	1
9	6	1	1,2
10	6	2	0,8
11	6	2	1
12	6	2	1,2
13	9	1	0,8
14	9	1	1
15	9	1	1,2
16	9	2	0,8
17	9	2	1
18	9	2	1,2

Os resultados obtidos pelo teste de Tukey para estes parâmetros são sumarizados na Tabela D.12.

Tabela D.12 – Grupos de similaridades obtidos pelo teste de Tukey para a abordagem proposta.

Configuração	Média estimada	Grupos				
4	0,571	A				
2	0,460	A	B			
10	0,338		B	C		
3	0,325		B	C		
6	0,319		B	C		
1	0,300		B	C	D	
11	0,225			C	D	E
12	0,220			C	D	E
7	0,219			C	D	E
13	0,170			C	D	E
9	0,158			C	D	E
16	0,150			C	D	E
5	0,138			C	D	E
17	0,117			C	D	E
8	0,110			C	D	E
18	0,108			C	D	E
15	0,083				D	E
14	0,060					E

Apêndice E

Testes de Hipóteses

Este apêndice contém uma descrição do funcionamento dos testes de hipóteses F ANOVA fator único e Tukey-Kramer, utilizado nesta dissertação.

E.1 F ANOVA fator único

O teste F ANOVA fator único verifica a hipótese nula (H_0: $\mu_1=\mu_2=\mu_3=...=\mu_c$) de não existir nenhuma diferença nas médias aritméticas ($\mu_1, \mu_2, \mu_3, ..., \mu_c$) obtidas em c grupos que representam populações cujas medidas foram retiradas aleatória e independentemente, seguem uma distribuição normal e têm variâncias equivalentes. Em caso de rejeição da hipótese nula, assume-se a hipótese alternativa como verdadeira, i.e., H_A: nem todas as médias aritméticas são iguais.

A probabilidade de que o teste de hipóteses rejeite a hipótese nula quando esta for verdadeira é definida pelo nível de confiança (1-α) do teste. Quanto maior o nível de confiança, i.e., quanto menor o valor de α, mais rigorosa será a condução do teste. Em contrapartida, quanto menor o nível de confiança, maior será a *eficácia* do teste, uma vez que se rejeitará mais freqüentemente a hipótese nula.

Para o teste da hipótese nula, calcula-se a estatística F, a partir da razão entre as variâncias ou termos quadráticos de médias *entre grupos* (MQE) e *dentro do grupo* (MQD), i.e.

$$F = \frac{MQE}{MQD} \qquad \text{(E.1)}$$

As variâncias MQE e MQD são valores gerados a partir da divisão da soma dos quadrados das diferenças por seus respectivos graus de liberdade.

A *variação entre grupos*, representada pela *soma dos quadrados entre grupos* (*SQE*), é definida como a soma dos quadrados das diferenças entre a média aritmética da amostra de cada grupo e a média geral, i.e.

$$\textit{Variação entre grupos (SQE)} = \sum_{j=1}^{c} n_j (\overline{X}_j - \overline{\overline{X}})^2 \qquad \text{(E.2)}$$

em que $\overline{\overline{X}}$ é a média geral dos valores obtidos em todas as classes, $\overline{X}_j$ corresponde à média aritmética do grupo j, n_j é o número de observações no grupo j e c é quantidade de grupos (Levine, Berenson e Stephan, 2000). Visto que c é quantidade de grupos comparados, a *SQE* é associada a c - *1* graus de liberdade, logo

$$MQE = \frac{SQE}{c-1} \qquad \text{(E.3)}$$

A *variação dentro do grupo*, geralmente denominada *soma dos quadrados dentro do grupo* (*SQD*), mede a diferença entre cada observação e a média aritmética de seu próprio grupo, acumulando os quadrados dessas diferenças sobre todos os grupos, sendo calculada através da equação.

$$\textit{Variação dentro do grupo (SQD)} = \sum_{j=1}^{c} \sum_{i=1}^{n_j} (X_{ij} - \overline{X}_j)^2 \qquad \text{(E.4)}$$

em que X_{ij} corresponde à *i-ésima* observação no grupo j e $\overline{X}_j$ é a média aritmética do grupo j (Levine, Berenson e Stephan, 2000). No cálculo da variação dentro do grupo, cada um dos c níveis contribui em

$n_j - 1$ graus de liberdade, resultando em $n - c$ graus de liberdade associados à *SQD*:

$$\sum_{j=1}^{c}(n_j - 1) = (n_1 - 1) + (n_2 - 1) + \ldots + (n_c - 1) = n - c \qquad \text{(E.5)}$$

em que n é o somatório de observações de todos os grupos, i.e. $n = n_1 + n_2 + \ldots + n_c$, assim

$$MQD = \frac{SQD}{n - c} \qquad \text{(E.6)}$$

Caso o valor da estatística F (Equação E.1) exceda um valor crítico (F crítico) para o nível de confiança escolhido, a hipótese nula é rejeitada. Os resultados de um teste *ANOVA* são geralmente sintetizados em uma *tabela resumida ANOVA*, cujas entradas incluem as *fontes de variação* (intergrupos, intragrupos), os *graus de liberdade* (gl), as *somas dos quadrados* (SQ), os *quadrados médios* (MQ) e a *estatística* F calculada.

E.2 Teste de Tukey (Tukey-Kramer)

O teste de Tukey permite verificar as diferenças entre as médias obtidas em vários grupos. Todavia, neste procedimento são realizadas comparações múltiplas, permitindo evidenciar quais os pares de grupos que são diferentes entre si.

No teste de Tukey, a hipótese nula é aceita para um par de médias (i e j) quando o módulo da diferença entre estas médias é menor que um valor crítico ($t_{i,j}$), definido por

$$t_{i,j} = \frac{q.\sqrt{MQD \cdot \left(\frac{1}{n_i} + \frac{1}{n_j}\right)}}{\sqrt{2}} \qquad \text{(E.7)}$$

em que MQD é a variância definida pela Equação E.3, n_i e n_j são a quantidade de elementos do grupo i e j respectivamente, e q (estatística

Q) o valor crítico associado ao nível de confiança escolhido para o teste de hipótese (Toothaker, 1993).