

Um Método Automático para o Ajuste de Segmentadores de Imagens Baseados em Informação de Textura e Cor

Fernando Henrique Bezerra Cardoso

Orientador: Herman Martins Gomes

Laboratório de Visão Computacional
Coordenação de Pós-graduação em Informática
Departamento de Sistemas de Computação
Universidade Federal de Campina Grande

13 de agosto de 2009



- 1 Introdução
- 2 Conceitos Fundamentais
- 3 Revisão Bibliográfica
- 4 Método Proposto
- 5 Avaliação Experimental
- 6 Conclusões e Trabalhos Futuros



Detecção de tipos de materiais: pele, grama, céu, asfalto, dentre outros.

Passo intermediário de aplicações em Visão Computacional:

- Detectar Pele:
 - Redução do espaço de busca na localização de faces
 - Informação para detecção de pornografia
- Detectar Grama: Informação para sumarização de vídeos desportivos
- Detectar Céu:
 - Categorização *indoor* × *outdoor*
 - Localização de horizonte



Tarefas para obter um segmentador:

- 1 Construir conjunto de imagens para treinamento
- 2 Determinar o valor dos parâmetros relacionados às imagens
- 3 Ajustar um classificador para os *pixels* das imagens
- 4 Testar o segmentador para mensurar sua qualidade



Em geral, pouca ou nenhuma automatização

Benefícios da automatização:

- Maior agilidade em projetos de Visão Computacional
- Não há desvio de esforços para resolução de tarefas periféricas

Principal contribuição: automatização da obtenção de segmentadores



“Que passos devem ser seguidos para que seja obtido um bom segmentador de imagens para detecção de materiais a partir de um conjunto de imagens em que as regiões do material em questão estão marcadas de maneira confiável?”

Particularidades do enunciado:

- A resposta não é um segmentador, mas um método
- Não há exigência de otimização
- O conjunto marcado manualmente deve ser confiável



Diminuir o tempo exigido nas tarefas que compõem o processo de obtenção de segmentadores de imagens:

- Maior agilidade na obtenção dos segmentadores
- Não é necessária supervisão humana

Geração de detectores para 3 tipos de material: grama, pele humana e céu



Propor um método de automatização da obtenção de segmentadores de imagens baseados em informação de cor e textura para detecção de materiais em fotografias digitais

Objetivos específicos:

- Investigar os aspectos relevantes para detectar materiais
- Implementar um protótipo de programa de computador seguindo o método proposto
- Validar o método proposto, através de estudos de caso
- Obter detectores para cada material dos estudos de caso.



- 1 Introdução
- 2 Conceitos Fundamentais**
- 3 Revisão Bibliográfica
- 4 Método Proposto
- 5 Avaliação Experimental
- 6 Conclusões e Trabalhos Futuros



Conceitos Fundamentais

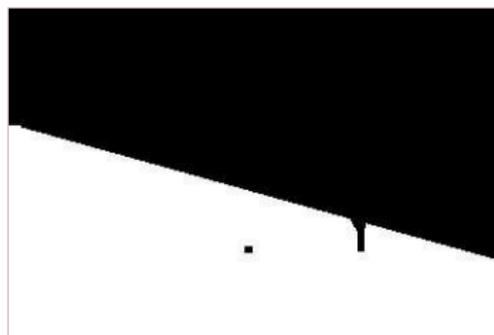
Regiões das Imagens (Classes)

P : abrange todos os *pixels* do material a ser detectado

N : abrange todos os demais *pixels*



(a) Imagem original



(b) P (Branco) / N (Preto)

Figura: Regiões P e N das imagens



Conceitos Fundamentais

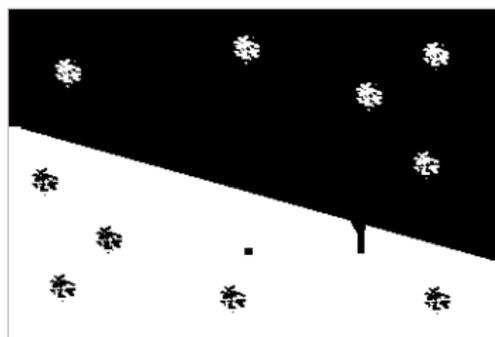
Regiões dos Resultados

T_P : pixels P classificados corretamente

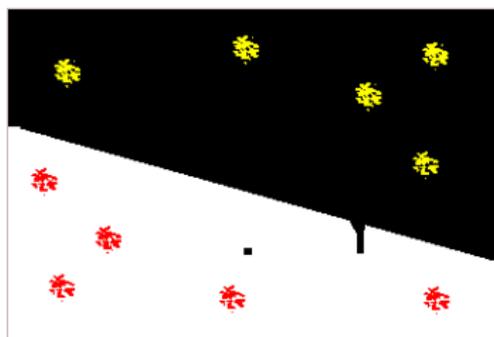
T_N : pixels N classificados corretamente

F_P : pixels N classificados incorretamente como P

F_N : pixels P classificados incorretamente como N



(a) Segmentação hipotética



(b) T_P (Branco) / T_N (Preto) / F_P (Amarelo) / F_N (Vermelho)

Figura: Regiões dos resultados das segmentações



Taxa de Detecção (*Recall*):

$$D_r^B(\zeta) = \frac{\sum_{I \in B} S(T_P(I))}{\sum_{I \in B} S(P(I))}$$

Precisão:

$$P_r^B(\zeta) = \frac{\sum_{I \in B} S(T_P(I))}{\sum_{I \in B} S(T_P(I)) + \sum_{I \in B} S(F_P(I))}$$



Medida-F (*F-measure*):

$$F_m^B(\zeta) = \left(1 + \beta^2\right) \frac{P_r^B(\zeta) \cdot D_r^B(\zeta)}{\beta^2 P_r^B(\zeta) + D_r^B(\zeta)}$$

β : peso relativo entre a taxa de detecção e a precisão

Neste estudo $\beta = 1$, salvo quando explicitamente especificado em contrário



Roteiro

- 1 Introdução
- 2 Conceitos Fundamentais
- 3 Revisão Bibliográfica**
- 4 Método Proposto
- 5 Avaliação Experimental
- 6 Conclusões e Trabalhos Futuros



Dificuldade em encontrar trabalhos com o mesmo objetivo desta dissertação

Projeto PERMM: objetivos *semelhantes*

Content Based Image Retrieval Using Semantic Visual Categories [Town e Sinclair 2000]

- Recuperação de imagens a partir de primitivas semânticas de alto nível
- Características de alto nível são determinantes da relevância na recuperação das imagens
- Construção de interfaces amigáveis de busca



Técnica dividida em três estágios:

- Segmentação por similaridade de cor utilizando diagramas de Voronoi [Sinclair 1999]
- Treinamento de redes neurais MLP, uma para cada tipo de material
- Consultas à base de imagens com a linguagem Oquel [Town e Sinclair 2001]

Diferenças entre o projeto PERMM e esta dissertação:

- Não há mecanismo automático de obtenção de novos detectores de materiais
- Leva em consideração as formas das regiões [Town e Sinclair 2000]



Espaço de Cor:

- Sistema de coordenadas tridimensionais onde cada cor é representada por um único ponto [Gonzalez e Woods 1987]
- Espaço padrão é o **RGB** [Jones e Rehg 2002]: alta correlação entre os 3 canais e mistura entre luminância e crominância
- Espaços com luminância separada: **YCbCr** [Phung et al. 2005], YES, YIQ, YUV, **HSV**, **HSL**, HSI
- Modelos da percepção humana das cores: **CIE XYZ**, **CIE L*a*b*** [Phung et al. 2005] e **CIE L*u*v*** [Vezhnevets et al. 2003]



Profundidade de Cor:

- Quantidade de *bits* com que cada canal de cor é representado
- Quando a profundidade é reduzida, há uma tendência de homogeneização nas cores
- Jones e Rehg obtiveram melhores resultados com 5 *bits* por canal [Jones e Rehg 2002]
- Jayaram *et al.* (2004) determinaram de 5 a 7 *bits*, a depender do espaço de cor



Posicionamento dos *Pixels*:

- Alguns materiais têm tendência a estar em uma determinada região da imagem
- Zafarifar *et al.* (2007) propõem um modelo probabilístico a partir da posição vertical dos *pixels* de céu



Revisão Bibliográfica

Aspectos Relevantes à Detecção de Materiais

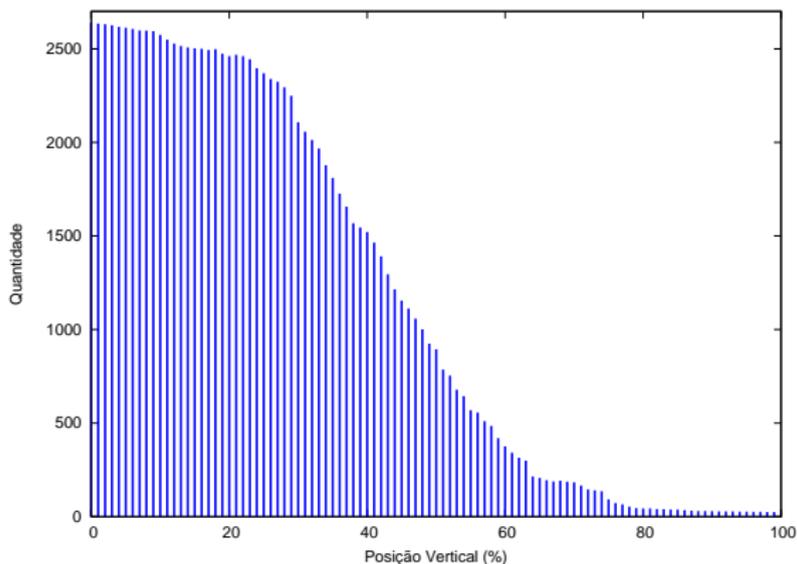


Figura: Histograma da coordenada vertical dos *pixels* de céu



Textura:

- Padrão homogêneo das pequenas variações de cor que caracterizam materiais
- Apenas informação de cor pode confundir detectores
- Se for introduzida a análise de textura, a confusão pode diminuir



(a) Sem textura (b) Textura de grama.

Figura: Comparação visual entre diferentes texturas.

Ineditismo na proposição do método automatizado para obter segmentadores

Vários aspectos que devem ser considerados para detectar materiais

Referências consultadas servem de fonte para estudo comparativo



Roteiro

- 1 Introdução
- 2 Conceitos Fundamentais
- 3 Revisão Bibliográfica
- 4 Método Proposto**
- 5 Avaliação Experimental
- 6 Conclusões e Trabalhos Futuros



Método Proposto

Visão Geral

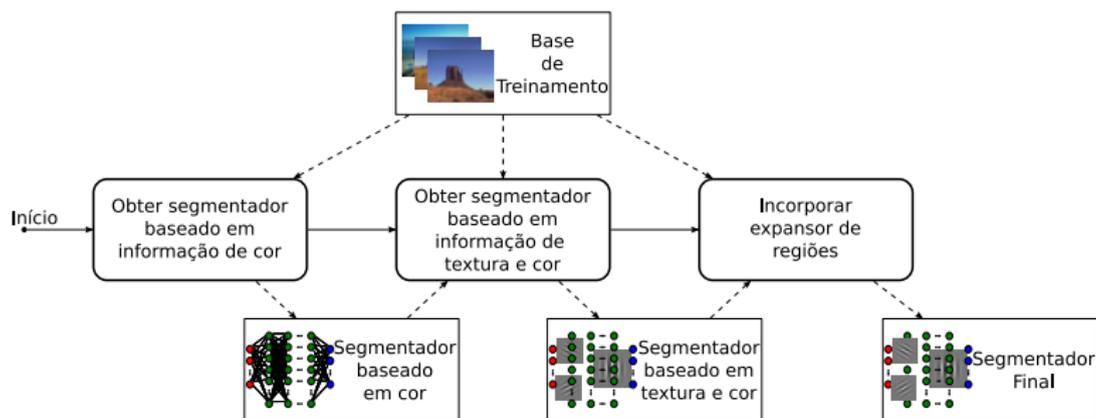


Figura: Visão geral do método proposto

Método Proposto

Fase 1: Segmentador neural baseado em informação de cor

Objetivo: ajustar um segmentador de imagens baseado em informação de cor

Posicionamento dos *pixels* também é levado em consideração

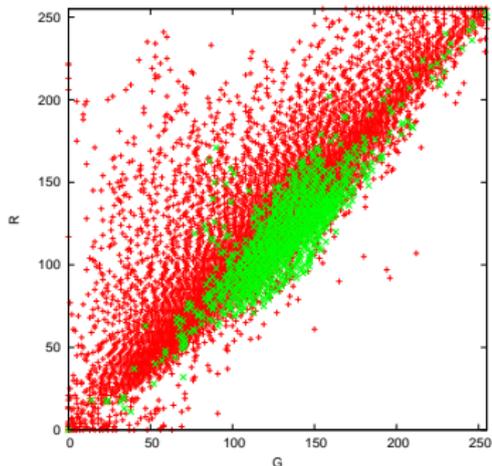
Redes Neurais MLP:

- Padrões não-linearmente separáveis
- Pouca memória para execução (apesar do lento treinamento)
- Não exigem nada a respeito da distribuição dos dados

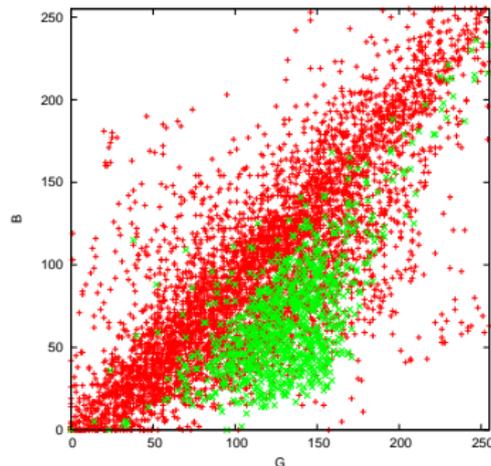


Método Proposto

Fase 1: Segmentador neural baseado em informação de cor



(a) Verde X Vermelho



(b) Verde X Azul

Figura: Distribuições dos pixels de grama e não-grama



Método Proposto

Arquitetura das redes neurais utilizadas

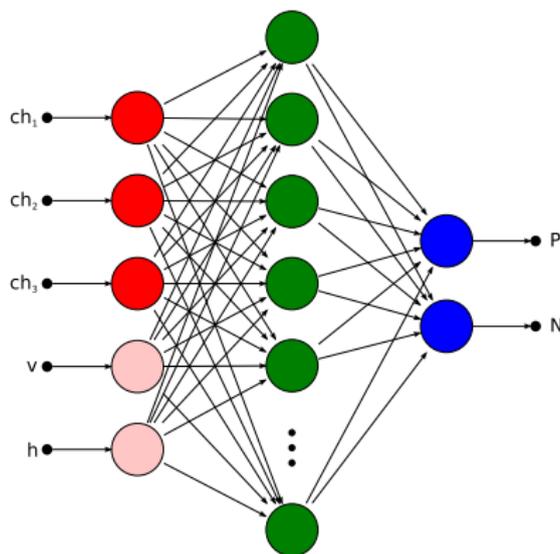


Figura: Representação da arquitetura das redes MLP no âmbito desta dissertação.



Método Proposto

Bases de imagens de treinamento

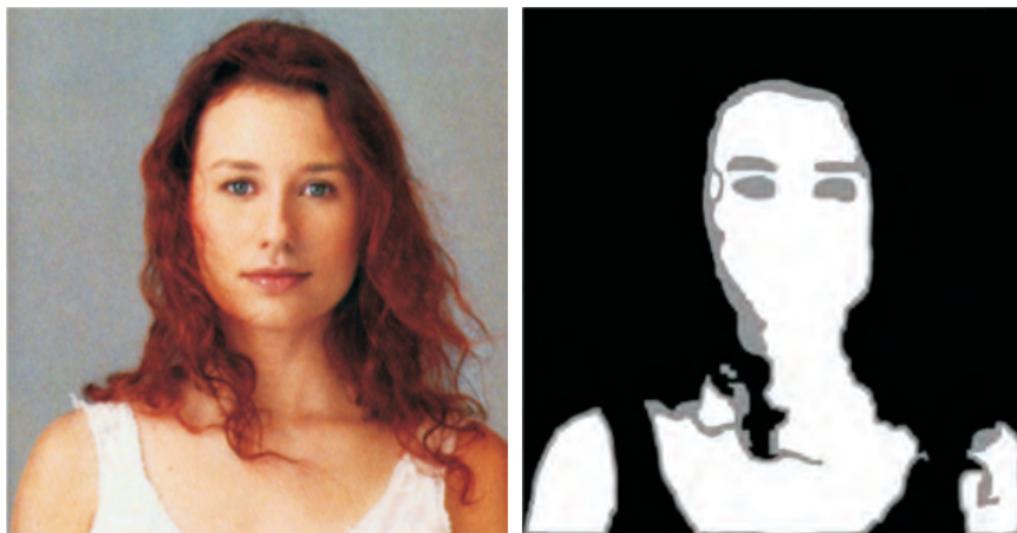


Figura: Exemplo de imagem no conjunto de treinamento.



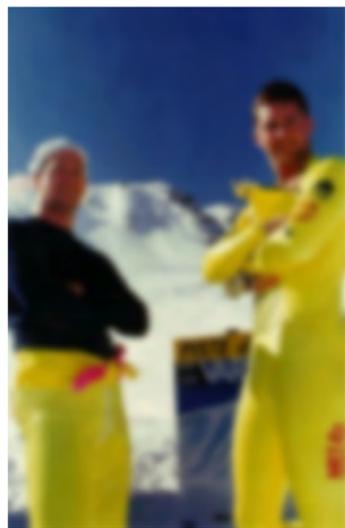
Método Proposto

Pré-processamento das imagens

Filtro Gaussiano: considerar vizinhança dos *pixels*



(a) Imagem Original



(b) Imagem Borrada

Figura: Borragem Gaussiana (9×9)



Método Proposto

Pré-normalização, redução e normalização final

Pré-normalização:

$$\text{norm}(v) = \frac{V - C_{min}}{C_{max} - C_{min}}$$

Redução:

$$\text{reduced}(\text{norm}, \text{levels}) = \frac{\lfloor \frac{\text{norm} \cdot 255 \cdot \text{levels}}{256} \rfloor}{\text{levels} - 1}$$

Normalização Final:

$$i(v_{\text{reduced}}) = 2v_{\text{reduced}} - 1$$



Método Proposto

Análise da posição dos *pixels*

$$i_h(x) = \frac{x}{w} \quad i_v(y) = \frac{y}{h}$$

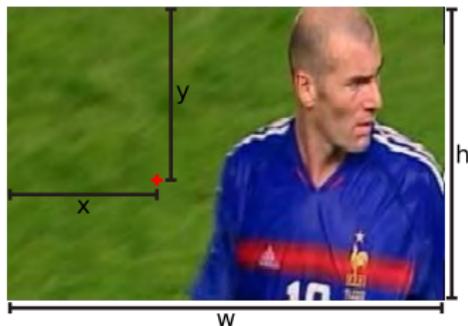


Figura: Distâncias consideradas quando se utiliza o posicionamento dos pixels



Método Proposto

Neurônios de saída

Valores esperados na saída:

- Classe P: $o_1 = +1, o_2 = -1$
- Classe N: $o_1 = -1, o_2 = +1$

Considera-se resultado da rede:

- P: se $o_1 > o_2 + t$
- N: nos demais casos



Inicialmente, todos os *pixels* são amostrados

Ordenação levando em consideração os canais de cor e o posicionamento

São retirados os k -ésimos *pixels*, tais que $k \equiv h \pmod{[h + 1]}$

Filtragem das ambiguidades:

- Ocorrências minoritárias e quantidades iguais das ocorrências majoritárias
- Ocorrências iguais em ambas as classes

Extração de quantidades iguais de padrões para cada classe



Método Proposto

Obtenção do segmentador neural

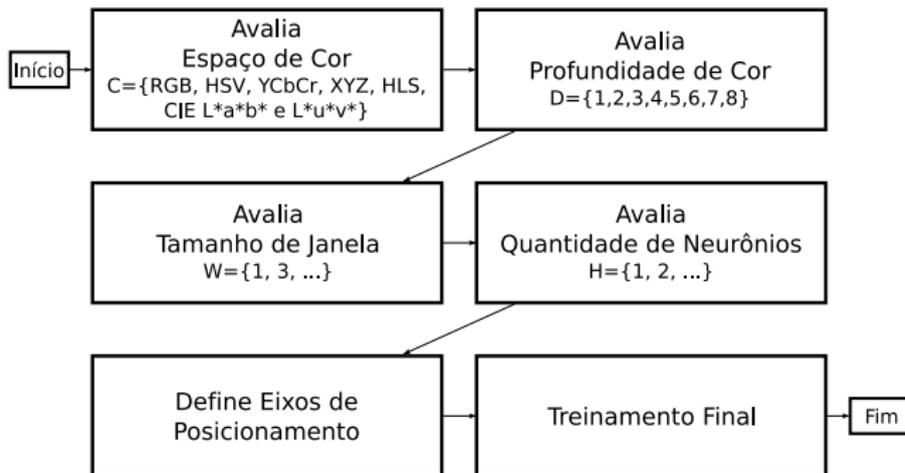


Figura: Principais passos da obtenção do segmentador neural baseado em informação de cor



Método Proposto

Fase 2: Segmentador baseado em informação de textura e cor

Objetivo: incorporar informação de textura ao segmentador da primeira fase.

A meta desta fase é minimizar a taxa de falsos positivos

Técnica proposta por Jiang *et al.* (2007): aplicação de filtros de Gabor com diferentes escalas e orientações

É realizada uma binarização do resultado do filtro de Gabor

Também é feita uma operação lógica “e” entre o resultado da binarização e o resultado da fase 1



Método Proposto

Fase 2: Segmentador baseado em informação de textura e cor



Figura: Segmentação baseada em informação de textura e cor.



Método Proposto

Fase 2: Segmentador baseado em informação de textura e cor

Figura - Etapa	Medida-F	F_{Pr}
Segmentação por Cor	0,9576	0,1917
Aplicação do Filtro de Textura	0,2554	0,0869
Conjunção (Resultado Final)	0,2466	0,0138

Tabela: Segmentação baseada em informação de textura e cor.



Método Proposto

Fase 2: Segmentador baseado em informação de textura e cor

Parâmetros a ajustar:

- M: Orientações dos filtros de Gabor
- N: Escalas dos filtros de Gabor
- l: Limiar inferior para binarização
- u: Limiar superior para binarização

Solução para ajuste de parâmetros: Algoritmo Genético



Objetivo: incorporar ao segmentador um expansor de regiões baseado em informação de cor

Adaptação do algoritmo *floodfill*:

- Com respeito à área conectada devem ser respeitados
 - um limiar b de distância Euclidiana de cor com respeito à borda da região já marcada
 - um limiar a de distância Euclidiana de cor com respeito à média de cor da região já marcada
- Utilização de várias sementes para execução do algoritmo



Método Proposto

Fase 3: Expansor de regiões com *floodfill* adaptado

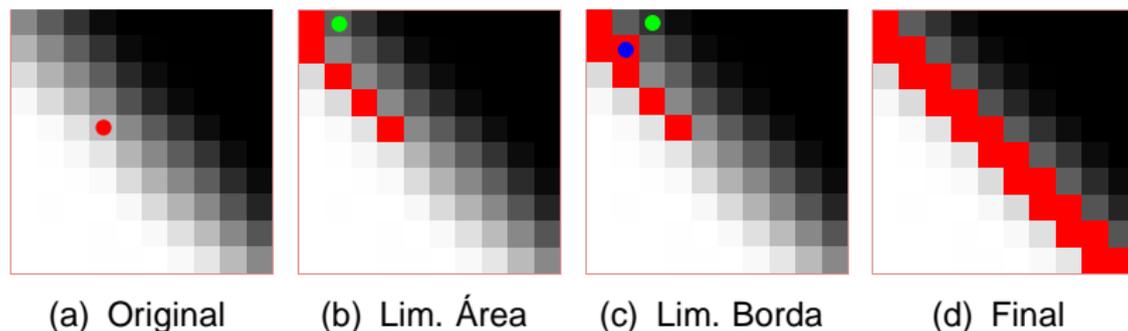


Figura: Algoritmo *floodfill* adaptado

Método Proposto

Fase 3: Expansor de regiões com *floodfill* adaptado



(a) Imagem Original



(b) Resultado Parcial



(c) Resultado Final

Figura: Algoritmo *floodfill* executado para detecção de grama.



Método Proposto

Fase 3: Expansor de regiões com *floodfill* adaptado

Figura - Etapa	Medida-F	F_{Pr}
Segmentação por Cor	0,9576	0,1917
Segmentação com Textura	0,2466	0,0138
Expansão de Regiões	0,9678	0,0380

Tabela: Resultados parciais da segmentação baseada em informação de textura e cor.



Método Proposto

Fase 3: Expansor de regiões com *floodfill* adaptado

Parâmetros a ajustar:

- a: Limiar de aceitação da área
- b: Limiar de aceitação da borda

Solução para ajuste de parâmetros: Algoritmo Genético



Método Proposto

Considerações sobre o método

Considera uma miríade de aspectos que ajudam na separabilidade das classes P e N

Sistematização para o ajuste de todos os parâmetros



Roteiro

- 1 Introdução
- 2 Conceitos Fundamentais
- 3 Revisão Bibliográfica
- 4 Método Proposto
- 5 Avaliação Experimental**
- 6 Conclusões e Trabalhos Futuros



Protótipo desenvolvido na linguagem C++

Bibliotecas utilizadas: OpenCV, libFann e gaLib

MLP: 100 épocas (1000 no treinamento final), $h = 10$, $i = 10$, conectividade total, *bias* em todos os neurônios

Genéticos: probabilidade de *crossover* de 0,9 e de mutação 0,01, com população de 100 indivíduos e limite máximo de 500 gerações



Base de imagens fornecida por Firmino Jr. e Gomes (2007)

Vídeos de partidas da Copa do Mundo de Futebol de 2006

Resolução de 360×240 *pixels*

50 imagens: 25 para treinamento e 25 para teste das métricas



Avaliação Experimental

Estudo de Caso 1: Grama

Fase 1 - Medida-F: 96,76%	
Espaço de Cor	YCbCr
Profundidade	8 <i>bits</i>
Tamanho da Janela	7×7 pixels
Neurônios	11

⇒

Fase 2 - Medida-F: 89,32%	
Escalas	1
Orientações	4
Limiar inferior	0
Limiar superior	7

↓

Fase 3 - Medida-F: 95,61%	
Limiar de área	18,57
Limiar de borda	232,63



Comparativo dos resultados:

Segmentador	Tx. Detecção	Precisão	Medida-F
Automatizado	96,68%	97,55%	97,11%
<i>Naïve Bayes</i>	96,64%	96,93%	96,78%
Firmino Jr. e Gomes	98,12%	92,79%	95,38%
Wang <i>et al.</i> (2004)	44,33%	97,81%	61,01%

Tabela: Resultados obtidos para detecção de grama



Melhor resultado:



(a) Original



(b) Resultado

Figura: Melhor resultado

Resultado médio:



(a) Original



(b) Resultado

Figura: Resultado médio

Pior resultado:



(a) Original



(b) Resultado

Figura: Pior resultado

Base de imagens fornecida por Jones e Regh (2002)

Necessidade de remarcação, devido a má qualidade da segmentação manual

Resolução variada

119 imagens: 60 para treinamento e 59 para teste das métricas



Avaliação Experimental

Estudo de Caso 2: Pele

Fase 1 - Medida-F: 68,81%	
Espaço de Cor	YCbCr
Profundidade	6 <i>bits</i>
Tamanho da Janela	3×3 pixels
Neurônios	13

⇒

Fase 2 - Medida-F: 49,88%	
Escalas	2
Orientações	1
Limiar inferior	3
Limiar superior	19

↓

Fase 3 - Medida-F: 57,84%	
Limiar de área	9,72
Limiar de borda	43,34



Comparativo dos resultados:

Segmentador	Tx. Detecção	Precisão	Medida-F
Automatizado	52,82%	49,31%	51,01%
<i>Naïve Bayes</i>	84,50%	35,17%	49,67%
Rahman <i>et al.</i> (2006)	66,56%	29,40%	40,78%
Herodotou <i>et al.</i> (2000)	83,34%	29,97%	44,09%

Tabela: Resultados obtidos para detecção de pele



Melhor resultado:



(a) Original



(b) Resultado

Figura: Melhor resultado

Resultado médio:



(a) Original



(b) Resultado

Figura: Resultado médio

Pior resultado:



(a) Original



(b) Resultado

Figura: Pior resultado

Base de imagens fornecida por Zafarifar *et al.* (2007)

Diversas imagens com reflexo de céu em água, piscinas, ou corpos azulados

Resolução variada

57 imagens: 29 para treinamento e 28 para teste das métricas



Avaliação Experimental

Estudo de Caso 3: Céu

Fase 1 - Medida-F: 78,42%	
Espaço de Cor	CIE L*a*b
Profundidade	7 bits
Tamanho da Janela	7×7 pixels
Neurônios	7

⇒

Fase 2 - Medida-F: 78,11%	
Escalas	8
Orientações	2
Limiar inferior	0
Limiar superior	19

↓

Fase 3 - Medida-F: 96,18%	
Limiar de área	237,94
Limiar de borda	2,65



Comparativo dos resultados:

Segmentador	Tx. Detecção	Precisão	Medida-F
Automatizado	94,52%	91,21%	92,84%
<i>Naïve Bayes</i>	93,47%	89,60%	91,49%

Tabela: Resultados obtidos para detecção de céu azul



Melhor resultado:



(a) Original



(b) Resultado

Figura: Melhor resultado



Resultado médio:



(a) Original



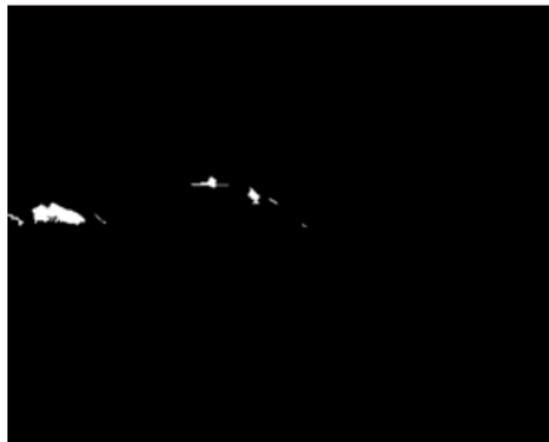
(b) Resultado

Figura: Resultado médio

Pior resultado:



(a) Original



(b) Resultado

Figura: Pior resultado

Efeito do expansor de regiões:



(a) Imagem original



(b) Apenas cor



(c) *Floodfill* adaptado

Figura: Evolução dos resultados da segmentação antes e depois da aplicação do *floodfill* adaptado

A automatização da obtenção de segmentadores é tarefa factível

Para todos os estudos de caso, os segmentadores gerados obtiveram resultados compatíveis (e maiores) que os demais segmentadores testados

Espaço de cor: parâmetro mais importante, com maior variabilidade

Profundidade de cor: parâmetro importante, comportamento crescente

Tamanho da janela: parâmetro com pouca importância, resultados sem variabilidade



Neurônios ocultos: rápida elevação na medida-F, seguida de estabilização

Expansor de regiões é eficaz em situações bem específicas:

- Bastante homogeneidade nas regiões do material
- Fronteiras muito bem definidas



Roteiro

- 1 Introdução
- 2 Conceitos Fundamentais
- 3 Revisão Bibliográfica
- 4 Método Proposto
- 5 Avaliação Experimental
- 6 Conclusões e Trabalhos Futuros**



Conclusões

Objetivo geral atingido: o método proposto é genérico e os seus resultados são compatíveis com os de técnicas específicas

A busca pelos aspectos para separação entre as regiões também é contribuição deste trabalho

Foram gerados e disponibilizados 3 detectores de materiais, que podem ser utilizados livremente pela comunidade acadêmica

Novo enfoque: abordar o problema de detecção de materiais sob a ótica da recuperação de conteúdo



Detector de pele utilizado em projeto de mestrado da UFCG [Leite 2008]:

- “Detecção de Olhos em Imagens com Faces Humanas”
- Fornecimento de critérios para eliminação de falsos candidatos a olhos

Todo o material está disponibilizado na Página Web do LVC



Pontos positivos:

- Não é necessário conhecer técnicas de segmentação de imagens
- MLP: execução rápida e exige pouca memória

Ponto negativo: treinamento demorado e custoso



Realizar análises matemática e estatística mais formais (PCA, por exemplo)

Incorporar técnicas de otimização global para descoberta de parâmetros das MLPs

Utilizar outras técnicas de classificação além de MLP, outros espaços de cores etc. Tornar o método ainda mais genérico



-  Gonzalez, R. C. e Woods, R. E. (1987).
Digital Image Processing.
Addison-Wesley.
-  Jones, M. J. e Rehg, J. M. (2002).
Statistical color models with application to skin detection.
International Journal of Computer Vision, 46(1):81–96.
-  Leite, B. B. (2008).
Detecção de olhos em imagens com faces humanas.
Master's thesis, Universidade Federal de Campina Grande
- UFCG.
-  Phung, S. L., Bouzerdoum, A., e Chai, D. (2005).
Skin segmentation using color pixel classification: Analysis
and comparison.
IEEE Trans. Pattern Anal. Mach. Intell., 27(1):148–154.



-  Sinclair, D. (1999).
Voronoi seeded colour image segmentation.
Technical report, AT&T Laboratories.

-  Town, C. e Sinclair, D. (2001).
Ontological query language for content based image
retrieval.
*In Proceedings of the 2001 IEEE Workshop on
Content-Based Access of Image and Video Libraries -
CBAIVL 2001*, pages 75–80.

-  Town, C. P. e Sinclair, D. (2000).
Content based image retrieval using semantic visual
categories.
Technical report, AT&T Laboratories Cambridge.

-  Vezhnevets, V., Sazonov, V., e Andreeva, A. (2003).



A survey on pixel-based skin color detection techniques.
In Proceedings of the GraphiCon 2003, pages 85–92.



Um Método Automático para o Ajuste de Segmentadores de Imagens Baseados em Informação de Textura e Cor

Fernando Henrique Bezerra Cardoso

Orientador: Herman Martins Gomes

Laboratório de Visão Computacional
Coordenação de Pós-graduação em Informática
Departamento de Sistemas de Computação
Universidade Federal de Campina Grande

13 de agosto de 2009

