

Filtragem de Cores em Exames de Citologia Oncótica

A. B. Canto, L.V. Guimaraes e A. Susin

Departamento de Eng Elétrica/UFRGS, Porto Alegre, Brasil

Resumo - Este artigo apresenta um método de filtragem de imagens baseado na medida de distância entre os pixels. A distância é calculada no espaço de cores RGB com técnicas da análise vetorial. O filtro apresentado foi aplicado na automação de exames papanicolau, onde a identificação de regiões de interesse passa pelo processamento de grande volume de informações de cor. A utilização do ângulo dos vetores de cor para mensuração da distância entre os pixels dispensa a utilização de outros espaços de cor, melhorando o desempenho da filtragem e reduzindo o tempo total de processamento.

Palavras-chave: filtragem de cor, proximidade de cor, RGB, Papanicolau..

Abstract - This article presents an image filtering method based on a measure of pixel distance. The distance measures are calculated on RGB color space and vector analysis. The proposed filter automates pap smear cytopathology tests, where the identification of ROIs (Regions Of Interest) requires process a large amount of color information. The use of vector angles to find the pixel proximity does not need color space conversions, improving the filtering performance and reducing the global time.

Key-words: color filtering, color proximity, RGB, Papanicolau.

Introdução

A importância do desenvolvimento de tecnologias que proporcionem uma redução de custos na realização de exames preventivos é diretamente proporcional ao número de indivíduos que potencialmente afetados pela doença diagnosticada. Não obstante os progressos que se vem alcançando no sentido de popularizar os exames de citologia oncótica (Papanicolau), existe ainda uma parcela importante da população que será beneficiada à medida que se desenvolvam tecnologias que aumentem a confiabilidade e reduzam os custos de tais exames. O percentual relativamente baixo de ocorrências de doenças (menos de 5%), impõe ao laboratorista longos e tediosos períodos entre casos de interesse, possibilitando que alguma anormalidade passe despercebida em consequência da monotonia da cena visual [1].

Um dos grandes desafios enfrentados num projeto que se proponha a trabalhar com aquisição de imagens microscópicas é o grande volume de dados gerados para cada lâmina examinada; as imagens de uma única lâmina podem chegar a Gbytes, com impacto direto sobre os custos de armazenamento, de processamento e de transmissão de dados [2].

A observação de que o nome popular Papanicolau derivou-se do método de coloração das lâminas é um forte indicativo da relevância que as informações de cor possuem neste exame. A filtragem de cor é uma das alternativas de pré-processamento com fundamental importância na identificação de regiões de

interesse, normalmente muito menores que a área total da imagem [3].

O objetivo deste trabalho é a apresentação de técnicas de filtragem por cor, utilizadas na etapa de pré-processamento de imagens adquiridas para o exame de citologia oncótica. Os métodos apresentados são desenvolvidos dentro do espaço de cores RGB, com técnicas de análise vetorial.

Metodologia

O método proposto associa o espaço de cores RGB ao espaço vetorial tridimensional, possibilitando a utilização de conceitos da geometria analítica para o tratamento de cores.

Na metodologia proposta, uma cor (RGB) é tratada como um vetor

$$\vec{C} = (R, G, B),$$

onde

$$\vec{C} = R.\vec{r} + G.\vec{g} + B.\vec{b} \quad (1)$$

e

$$\vec{r} = (1,0,0) \quad (2)$$

$$\vec{g} = (0,1,0)$$

$$\vec{b} = (0,0,1)$$

são base ortogonal de representação do espaço tridimensional de cores.

Módulo x Intensidade: o módulo de um vetor

$$|\vec{A}| = (Ra, Ga, Ba) \quad (3)$$

é calculado pela equação (4) :

$$|\vec{A}| = \sqrt{Ra^2 + Ga^2 + Ba^2} \quad (4)$$

Na metodologia utilizada, o módulo de um vetor RGB é associado à intensidade.

Colinearidade x Igualdade de cor: Dois vetores

$$\vec{A} = (Ra, Ga, Ba) \quad (5)$$

$$\vec{B} = (Rb, Gb, Bb)$$

são colineares se atenderem a condição especificada na equação (6) a seguir

$$\frac{Ra}{Rb} = \frac{Ga}{Gb} = \frac{Ba}{Bb} \quad (6)$$

Dois vetores colineares possuem mesma cor.

Produto Escalar x Similaridade de cor: dados dois vetores

$$\vec{A} = (Ra, Ga, Ba) \quad (7)$$

$$\vec{B} = (Rb, Gb, Bb)$$

o seu produto escalar é calculado através da equação (8):

$$\vec{A} \cdot \vec{B} = Ra \cdot Rb + Ga \cdot Gb + Ba \cdot Bb \quad (8)$$

A equação (9), apresenta a relação entre o produto escalar de dois vetores, seus módulos e o ângulo formado entre eles:

$$\vec{A} \cdot \vec{B} = |\vec{A}| \cdot |\vec{B}| \cdot \cos \alpha \quad (9)$$

onde α é o ângulo formado entre os dois vetores.

Proximidade de cor: é definida como o cosseno do ângulo formado dois vetores RGB. A proximidade máxima de dois vetores RGB assume o valor 1, correspondente a dois vetores colineares.

Dados dois vetores

$$\vec{A} = (Ra, Ga, Ba) \quad (10)$$

$$\vec{B} = (Rb, Gb, Bb)$$

a equação (11) permite calcular a proximidade P :

$$P = \frac{Ra \cdot Rb + Ga \cdot Gb + Ba \cdot Bb}{|\vec{A}| \cdot |\vec{B}|} \quad (11)$$

Método de inspeção visual da dispersão de cores: Através deste método inspeciona-se no espaço tridimensional a dispersão de cores de uma determinada imagem. As etapas utilizadas neste processamento são as seguintes:

1. Leitura da imagem a analisar (Figura 1). A Figura 1 é um exemplo de imagem proveniente de amostras de tecido cervical em uma preparação para exame papanicolau, cedidas pelo Instituto de Patologia (parceiro de projeto).
2. Normalizar em módulo as cores da imagem. A Figura 2 é o resultado da normalização dos vetores de cor para o módulo 255.

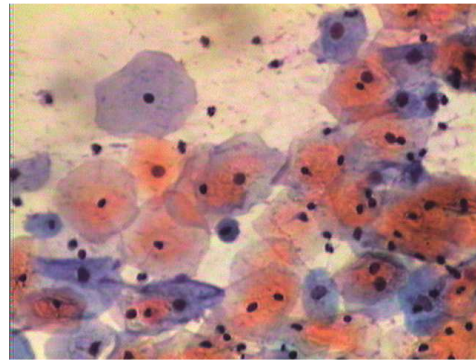


Figura 1 – Imagem analisada

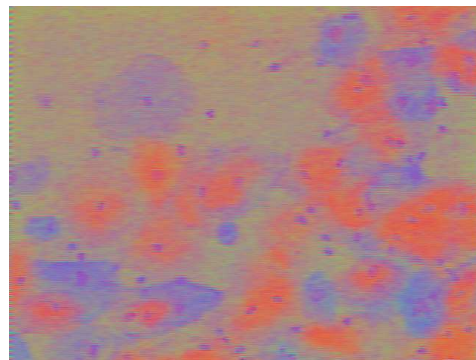


Figura 2 – Efeito da normalização de cores

3. Criação de um histograma, no qual se identifique o número de ocorrências de cada cor normalizada
4. Desenho no espaço tridimensional das cores existentes na imagem analisada. O número de ocorrências de cada cor é representado pela intensidade utilizada no desenho. A Figura 3 mostra as cores da Figura 2 posicionadas no espaço tridimensional RGB.

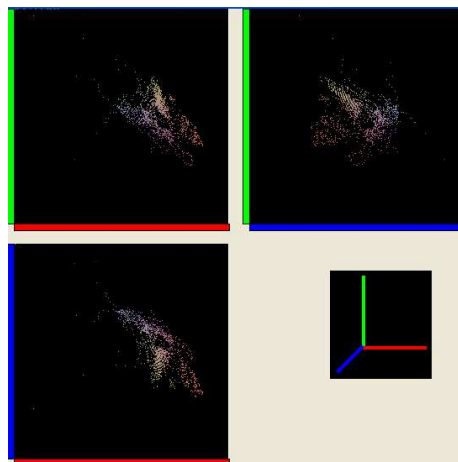


Figura 3 - Distribuição de cores no espaço RGB

A Figura 4 mostra um zoom da distribuição de cores nos eixos RB, onde é possível visualizar a distribuição espacial das cores.



Figura 4 - Zoom da projeção dos eixos RB

Através do método de inspeção visual da dispersão de cores é possível ter uma visão geral sobre a forma como as cores de uma determinada imagem se distribuem no espaço tridimensional, visando a especificação das cores que serão utilizadas no processo de filtragem.

Metodologia de filtragem por proximidade: O objetivo deste método é isolar áreas de interesse, identificadas pela proximidade de vetores RGB. O método consiste nos seguintes passos:

1. Definir um ou mais conjunto nomeado de vetores RGB. A Figura 5 mostra um exemplo, em que foram definidos 3 conjuntos: *Pink*, *Repr* e *Blue*. A figura mostra também as cores selecionadas para compor cada um destes conjuntos

Name	Description	
Pink	Pink Colors	
Blue	Blue Colors	
Repr	Reproved colors	

Figura 5 - Conjuntos nomeados de vetores RGB

2. Analisar cada pixel da imagem sobre a qual se deseja aplicar o método de filtragem. Esta análise deve avaliar a proximidade do vetor RGB associado ao pixel, identificando o conjunto que possui maior proximidade. O resultado desta etapa é um mapa classificando cada pixel segundo as categorias definidas pelos conjuntos de vetores.

Ambiente de desenvolvimento: O sistema de aquisição de imagens utilizado é composto de um microscópio convencional, uma câmera SONY SSC-C104 NTSC conectada a uma placa de aquisição baseada no chipset BT878/879 com barramento PCI. Os métodos computacionais

utilizados para gerar a Figura 2 (bem como as demais figuras e testes de desempenho) foram desenvolvidos no ambiente Borland C++.

Resultados

O algoritmo de filtragem por cor foi aplicado a três grupos de 11 imagens, classificadas, nas categorias *alta*, *média* e *baixa*, conforme a percepção de laboratoristas quanto a área útil ocupada pelas células. Após a filtragem, avaliaram-se os percentuais de pixels aprovados sobre o número total de pixels da imagem. A Tabela 1 apresenta resultados onde os percentuais de pixels aprovados na filtragem podem ser utilizados para quantificar esta percepção.

Tabela 1 - Resultado da filtragem

Classific. Laborat.	Pixel aprovado / Pixel total		
	Média	Máximo	Mínimo
Alta	51,0%	60,3%	42,1%
Baixa	4,8%	9,9%	0,9%
Média	23,7%	30,3%	18,7%

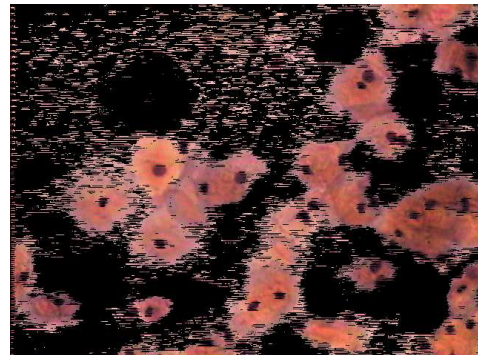


Figura 6 - Pixels classificados na categoria "Pink"

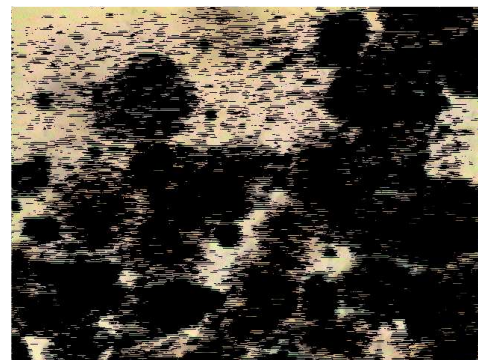


Figura 7 - Pixels classificados na categoria "Repr"

O método desenvolvido mostrou-se eficaz não só no que diz respeito a isolar grupos de pixels classificados pelas cores, como também

com relação à fidelidade das cores após a filtragem.

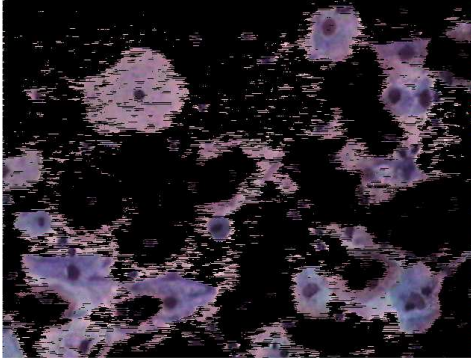


Figura 8 - Pixels classificados na categoria "Blue"

As figuras 6, 7 e 8 mostram o resultado da classificação da imagem original, apresentada na Figura 1. O processamento foi realizado utilizando-se o método de filtragem por proximidade, os conjuntos nomeados de cores mostrados na Figura 5 e o conceito de proximidade apresentado na equação (11).

Discussão

O algoritmo apresentado propõe a realização de filtragem de cores com base no conceito de proximidade que, por sua vez, é definido para a metodologia de análise vetorial de cores no espaço RGB. O mesmo algoritmo pode ser implementado utilizando-se outras definições de proximidade, utilizando-se, por exemplo, a dimensão matiz do espaço HSI.

Esta alternativa, no entanto, apresenta desvantagens de desempenho e resolução cor/saturação com relação à abordagem proposta.

Desempenho – No padrão HSI, a distância entre as cores $C_a = (R_a, G_a, B_a)$ e $C_b = (R_b, G_b, B_b)$ é calculada através da equação (12):

$$distância = |H(R_a, G_a, B_a) - H(R_b, G_b, B_b)| \quad (12)$$

onde [4]

$$H(R, G, B) = \quad (13)$$

$$\cos^{-1} \left[\frac{\frac{2R - G - B}{3(R + G + B)}}{\sqrt{\frac{2}{3} \sqrt{\frac{9(R^2 + G^2 + B^2) - 3(R + G + B)^2}{9(R + G + B)^2}}}} \right]$$

No espaço vetorial RGB, a proximidade é calculada através da equação (14), cujo processamento toma cerca de um terço do tempo tomado pela equação (13).

$$proximidade = r_a r_b + g_a g_b + b_a b_b \quad (14)$$

onde

$$(r_a, g_a, b_a) = (R_a, G_a, B_a) / \|R_a, G_a, B_a\| \quad (15)$$

$$(r_b, g_b, b_b) = (R_b, G_b, B_b) / \|R_b, G_b, B_b\|$$

e

$$\|R, G, B\| = \sqrt{R^2 + G^2 + B^2} \quad (16)$$

A utilização ângulo formado entre os vetores de cor quantifica adequadamente diferenças decorrentes de variações de matiz e saturação, permitindo fugir dos complexos cálculos de arccos [5].

Potencial de diferenciação pela saturação – O cálculo das distâncias com base na dimensão matiz, do sistema HSI, implica uma perda decorrente da impossibilidade de diferenciação de pixels com níveis de saturação distinto. A Figura 9 é uma imagem de hemácias em uma amostra preparada para contagem de reticulócitos. Observa-se que, após a normalização de módulos, as principais diferenças se encontram na saturação. O filtro desenvolvido, utilizando um cálculo de distância de cores com base nos ângulos RGB mostra-se robusto a medida que é capaz de detectar variações de matiz, de saturação ou uma combinação das duas componentes.

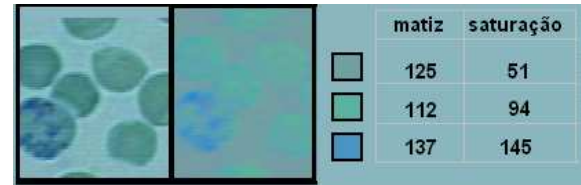


Figura 9 - Variações de matiz e saturação

Conclusões

Na automação de exames de patologia oncológica (Papanicolau) a identificação de cores possui uma especial relevância, tanto por aspectos de diagnóstico, em que há necessidade de identificar regiões com predominância de azul, como no que diz respeito a fatores técnico-econômicos, que impõem restrições quanto ao volume de informações e tempo de processamento. Os filtros desenvolvidos para esta aplicação fizeram uso de uma sistemática baseada no conceito de *proximidade de cores*, definido em função do ângulo entre vetores RGB. Esta técnica viabiliza ganhos de desempenho que poderão ser explorados em quaisquer aplicações críticas, tais como sistemas de tempo real, por exemplo [6].

A metodologia de filtragem por proximidade pode operar ainda com qualquer outro algoritmo de cálculo de distância entre pixels, possibilitando o desenvolvimento de filtros parametrizáveis para outras aplicações em que as condições de contorno extrapolem apenas a homogeneidade da cor [7].

Este processo de filtragem mostra-se sensível a quaisquer deslocamentos de cor decorrentes de variações dos corantes, de iluminação, de ajustes automáticos de ganho nas câmeras ou qualquer outro aspecto que potencialize uma distribuição espacial de cores distinta daquela utilizada originalmente nos conjuntos nomeados de vetores RGB. Esta sensibilidade pode ser compensada através do uso de múltiplos parâmetros de filtragem ou da ponderação da filtragem de cor dentro de uma lógica difusa.

Agradecimentos

Agradecemos ao bolsista Vinicius Cristino Souza pelo suporte nas atividades de programação e ao CNPQ pelo financiamento parcial deste trabalho.

Referências

- [1] Rosenthal, D. L. (1998) "Automation and the endangered future of the Pap test", In: *J Natl Cancer Inst*, May; 90: 738 - 749
- [2] Jornsten, R., Bin Yu (2002) "Compression of DNA microarray images", In: *Biomedical Imaging, 2002. Proceedings. 2002 IEEE International Symposium*, 7-10 July 2002 p.38 - 41
- [3] Begelman, G., Lifshits, M., Rivlin, E., (2006) "Visual Positioning of Previously Defined ROIs on Microscopic Slides", In: , *IEEE Trans.*

Information Technology in Biomedicine, 7-10 July, p. 38 - 41

- [4] Gonzales, R.C., Woods, R.E. (1992), *Digital Image Processing*, Addison-Wesley Publishing Company
- [5] R. D. Dony and S. Wesolkowski, (1999) "Edge Detection on Color Images Using RGB Vector Angle", In: *Proc. IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, Edmonton, Canada
- [6] Jian-Feng Li., Kaun-Quan Wang; Zhang, D., (2002) "A new equation of saturation in RGB-to-HSI conversion for more rapidity of computing", In: *Machine Learning and Cybernetics, 2002. Proceedings. International Conference*, Nov., p.1493 - 1497 vol.3
- [7] Wesolkowski, S., Fieguth P., (2003) "Color image segmentation using connected regions", In: *Electrical and Computer Engineerin., IEEE CCECE 2003. Canadian Conference*, 4-7 May 2003 Volume 2, Page(s):1203 - 1206

Contato

Prof. Alberto Bastos do Canto Filho; Dep Engenharia Elétrica – UFRGS; E-Mail: alberto.canto@ufrgs.br; fone: (51) 3316 3351; Av. Osvaldo Aranha, 103 - Bairro Bom Fim CEP: 90035-190 - Porto Alegre - RS - Brasil