

Um Método de Segmentação de Imagens para Cenas Naturais

Christiano Augusto Caldas Teixeira
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais, Brasil
teixeira@dcc.ufmg.br

Resumo—Este trabalho apresenta um método de segmentação para imagens de cenas naturais. Um detalhado estudo estatístico destas cenas naturais se mostrou de grande importância no seu entendimento, tornando possível uma geração de conhecimento eficaz à partir de descritores simples. Esse estudo permitiu o desenvolvimento de um processo de segmentação baseado em classificação de regiões promissor.

Keywords—Processamento de Imagens; Visão Computacional; Segmentação de Imagens;

I. INTRODUÇÃO

A segmentação de imagem é descrita como o processo responsável por dividir uma dada imagem em regiões homogêneas, que compartilhem determinada similaridade. Esse processo é de grande relevância em visão computacional e processamento e análise de imagens, possuindo grandes desafios ainda em abertos e largamente estudados. O motivo dessa grande importância pode ser justificado pelo fato da segmentação ser o primeiro passo na análise e processamento de imagens [1].

O produto da segmentação serão regiões ou objetos considerados relevantes para contexto no qual o método foi aplicado, descartando qualquer tipo de informação sem importância na imagem. O grande desafio é conseguir determinar quais atributos em uma imagem podem ser considerados relevantes e, quais podem ser descartados. Infelizmente essa não é uma tarefa simples, já que atributos relevantes para uma classe de imagens podem não ser relevantes para outras. Além disso, ainda existe o problema da subjetividade humana, que faz com que cada indivíduo possui uma percepção única do mundo. Isso significa que pessoas diferentes poderão considerar conjuntos diferentes de atributos relevantes em uma mesma imagem.

Ao longo do tempo, inúmeros métodos para segmentação foram propostos e estudados, cada um se objetivando em solucionar os problemas discutidos [2]. Devido à grande complexidade desses problemas, principalmente relacionados à percepção humana, esses métodos propostos são restritos à determinadas classes de imagens. Isso significa que a grande maioria dos métodos são capazes de realizar a segmentação com eficiência para determinados tipos de imagens, mas não conseguem manter a mesma eficiência quando se utiliza imagens de outra classe.

Devido a essa falta de métodos eficientes para diferentes classes ou tipo variados de imagens, e a existência de uma grande quantidade de métodos eficientes, mas restritos a determinados contextos, trabalhos envolvendo métodos adaptativos e capazes de inferir decisão vêm sendo propostos. Esses métodos têm como principal objetivo encontrar, dentre um conjunto de métodos de segmentação, aquele mais eficiente para determinada imagem [3] [4].

A. Objetivos

O desenvolvido de um método de segmentação capaz de identificar elementos de cenas naturais ou de estruturas construídas pelo homem em uma imagem seria de grande utilidade para uma vasta gama de métodos e sistemas que trabalham com imagens. Isso se deve ao fato desses elementos estarem presentes em uma grande quantidade de imagens encontradas. Por exemplo, fotos dos mais diversos tipos tendem a conter uma grande quantidade de elementos naturais ou de estruturas construídas pelo homem em seu conteúdo e, na grande maioria das vezes representando o fundo da cena. Dessa forma, um método capaz de segmentar esses conteúdos pode ser de grande ajuda para processamentos posteriores.

Motivado por essas observações, esse trabalho tem como principal objetivo, desenvolver um método de segmentação de imagens para cenas naturais e cenas de estruturas construídas pelo homem. Espera-se que esse método seja capaz de reconhecer esses elementos com precisão, apresentando um baixo número de falsos positivos e falsos negativos.

II. METODOLOGIA

A segmentação de imagens não é simples e atualmente existem várias técnicas distintas para diferentes tipos de imagens e diferentes objetivos. Devido a isso, é necessário realizar um estudo detalhado sobre a classe de imagens com o qual se pretende trabalhar, para que seja possível determinar a abordagem para a segmentação que permitirá alcançar os objetivos propostos. O processo de segmentação proposto pode ser dividido em várias etapas, como mostra a figura 1.

O início do processo se dá com a análise e estudo das classes das imagens de interesse. Esse estudo é realizado através da extração de descritores dessas, de modo

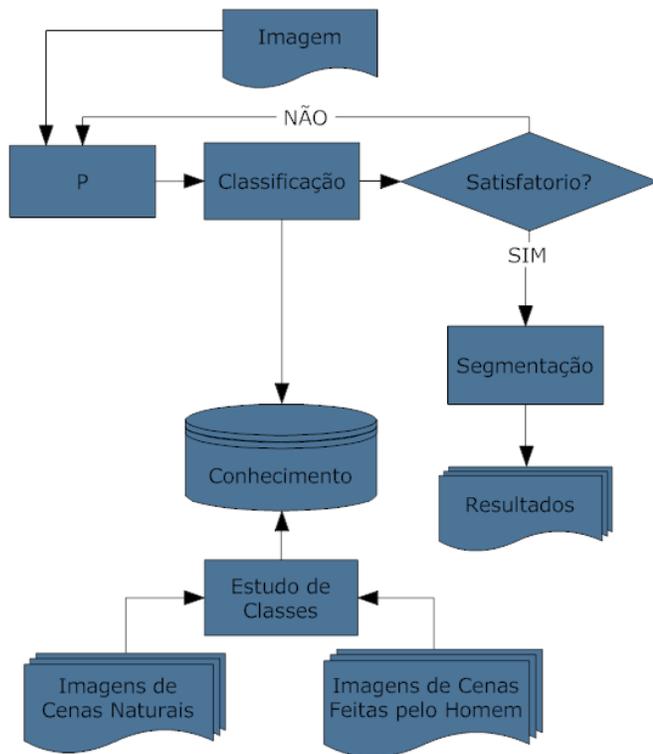


Figura 1. Processo de segmentação proposto para imagens de cenas naturais e de estruturas construídas pelo homem.

a encontrar padrões capazes de diferenciar as imagens das diferentes classes. Essa etapa do processo será responsável por gerar o conhecimento necessário para a classificação e, posteriormente a segmentação.

Com o estudo realizado, a segmentação se inicia com a inserção de uma imagem de entrada no processo. Essa imagem é subdividida em regiões por um processo P, onde cada uma dessas regiões é classificada de acordo com o conhecimento das classes previamente estudadas. Caso a classificação não seja satisfatória¹, a região em questão é submetida novamente ao processo P, sendo dividida em regiões menores a serem classificadas separadamente. Quando a classificação é então satisfatória, a segmentação é realizada, extraindo da imagem, as regiões de cada classe. Os produtos do processo serão várias imagens, cada um contendo elementos de cada uma das classes estudadas. As seções a seguir irão detalhar cada etapa do processo.

A. O Estudo de Imagens

O estudo de imagens é de fundamental importância para o modelo de segmentação proposto, uma vez que é responsável por encontrar as principais diferenças entre cada uma das classes de imagens.

¹Uma classificação é dita satisfatória quando há um certo grau de certeza ao determinar a qual classe a região pertence.

O estudo se inicia com a definição das classes de interesse, que serão o foco do estudo. Estas são então analisadas separadamente de forma a determinar os descritores que apresentam maior diferença nas imagens. Por exemplo, ao se trabalhar com determinadas classes, descritores baseado em cor podem não ser relevantes se todas as imagens possuírem os mesmos padrões de cores. Então uma classificação baseado nesses atributos não será eficiente, devendo ser usado descritores mais qualificados para o contexto.

É possível observar que esse estudo é realizado sobre um conjunto de imagens, onde suas classes são conhecidas à priori. Dessa forma, todo o conhecimento é extraído a partir das amostras disponíveis para o método. É importante perceber então que as imagens de amostra de cada classe do estudo, terão direto impacto na classificação, o que torna necessário uma amostragem bem eficiente, buscando imagens capazes de capturar todos os principais elementos de cada classe.

1) **Estudo de caso: Cenas Naturais:** O principal foco deste trabalho é o desenvolvimento de um método para segmentação de cenas naturais e artificiais. Para isso, é importante analisar com cuidado cada uma dessas imagens para assim encontrar os melhores descritores para distinção das classes.

Todo o estudo de caso aqui descrito foi realizado sobre o trabalho de Geisler[5], que apresenta uma detalhada análise estatística de cenas naturais. Este trabalho tem como principal objetivo determinar padrões estatísticos presentes em cenas naturais, na tentativa de se obter um melhor entendimento do funcionamento do sistema visual de animais. Apesar do aspecto biológico do trabalho, os estudos estatísticos encontrados foram a principal referência para o método de segmentação desenvolvido.

Mesmo este trabalho apresentando uma análise de vários padrões estatísticos de cenas naturais, o método de segmentação proposto apenas utilizou dois: a intensidade luminosa e o contraste. O trabalho [6] do mesmo autor, apresenta o estudo estatístico desses dois padrões detalhadamente, mostrando a importância de cada descritor na imagem e, como esses foram obtidos e calculados.

2) **Amostragem e Definição das Classes:** Para o estudo de classe, um total de 70 fotos foram tiradas, sendo 60 fotos de cenas naturais e 10 fotos de cenas contendo exclusivamente elementos construídas pelo homem. Todas as fotos foram tiradas com a câmera fotográfica HP Photosmart 435 em resolução 2048x1536 pixels.

As imagens foram divididas em duas grandes classes, imagens de cenas naturais e imagens de cenas construídas pelo homem. Com o objetivo de aumentar a precisão do estudo, as imagens de cenas naturais foram divididas em subclasses, uma vez que essas apresentam elementos com grandes diferenças de contraste e intensidade luminosa. As classes estudadas foram:

- Cenas Naturais: 70 Fotos

- Terreno: 25 fotos
- Vegetação: 20 fotos
- Céu: 15 fotos

- Cenas Construídas pelo Homem: 10 fotos

A imagem 2 apresenta uma típica imagem de cada uma das classes escolhidas.

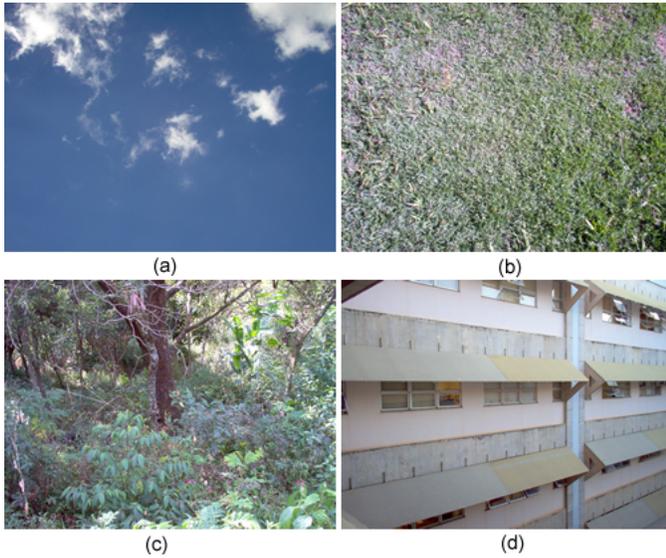


Figura 2. Imagens típicas de cada uma das quatro classes escolhidas. (a) Imagem típica da classe céu. (b) Imagem típica da classe terreno. (c) Imagem típica da classe vegetação. (d) Imagem típica da classe estruturas construídas pelo homem.

3) **O Descritor Intensidade Luminosa:** A luminosidade é um estímulo fundamental capturado pelos sistemas visuais e seus padrões são bem uniformes em cenas naturais [6].

A intensidade luminosa é medida pelo valor de cada pixel na imagem, representando a quantidade de luz capturada pela câmera em cada ponto da cena. Dessa forma, a luminosidade de uma imagem é dada pela fórmula:

$$I = \frac{1}{N} \sum_{i=1}^N I_i,$$

onde N é o número total de pontos da imagem e I_i a intensidade do ponto i desta.

Apesar da intensidade luminosa ser bem distinta nos diferentes padrões de imagens naturais, o processo de captura das câmeras tendem a aplicar um pré-processamento em cima das fotos de cada cena, de modo a normalizar todos os pixels dentro de um padrão. Isso pode fazer com que várias imagens de classes diferentes possuam luminosidades muito próximas. Dessa forma, outra medida baseada na luminosidade foi utilizada: a faixa de intensidade luminosa.

A faixa de intensidade luminosa é definida como a diferença entre as maiores e as menores intensidades na imagem. Esse descritor irá determinar como é a variação de luminosidade ao longo de cada região desta, o que fornece a

precisão necessária para o estudo. A faixa de luminosidade usada foi a de 95%, referente às diferenças do percentil 97,5 e o percentil 2,5.

4) **O Descritor Contraste:** A medida de contraste utilizada foi o contraste RMS. A escolha dessa medida se deve ao fato desta ser padrão para o cálculo do descritor, além dos seus resultados serem os mais próximos da detecção de contraste do sistema visual humano [6].

O contraste RMS é definido pelo desvio-padrão das intensidades de pixel e, sua fórmula é descrita da seguinte forma:

$$C_{rms} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (I_i - \bar{I})^2}$$

onde, I_i é o valor das intensidades normalizado para 1, N o número total de pixels na imagem e \bar{I} a intensidade média da imagem, também normalizada.

B. O processo P

O processo P é responsável por dividir a imagem em regiões menores para que essas possam ser classificadas. Essa etapa do processo é fundamental para o desempenho do sistema, já que classificar uma região é menos trabalhoso que classificar cada ponto da imagem. Além disso, analisando imagens de cenas naturais e construídas pelo homem, percebe-se que os elementos de cada uma das classes se encontram concentrados em regiões distintas da imagem e não espalhadas ao longo desta. Isso torna a divisão ainda mais adequada para o processo proposto.

Para esse trabalho, optou-se por um método de divisão baseado em QuadTree. A QuadTree é uma árvore em que cada nó possui sempre quatro filhos. Cada imagem será então dividida em regiões de mesmo tamanho para que a etapa de classificação possa ser realizada. Cada uma dessas regiões poderá ser dividida em quatro novas regiões menores se necessário e assim, sucessivamente.

C. A Classificação

A etapa de classificação, como o próprio nome já diz, irá classificar uma dada região da imagem em função da distribuição de seus aspectos para cada uma das classes estudadas. Para que a essa etapa possa ser realizada, a quantização dos aspectos dessa região devem ser realizados da mesma forma que realizados para as imagens de estudo.

Com base nas distribuições obtidas em cada região, é possível determinar quão próximos esses dados estão de cada uma das classes estudadas. Para determinar a distância entre a região e cada classe utilizou-se a distância de Mahalanobis. Essa distância é baseada na correlação entre as variáveis, sendo útil para calcular a similaridade entre amostras. A distância de Mahalanobis é dada pela fórmula:

$$d(\vec{x}, \vec{y}) = \sqrt{(\vec{x} - \vec{y})^T S^{-1} (\vec{x} - \vec{y})}$$

Onde \vec{x} e \vec{y} são os vetores de variáveis obtidos para cada amostra e S a matriz de covariância.

Com a distância entre a região da imagem e as classes estudadas calculada, determina-se a classe com menor distância como sendo a classe da região. Essa classificação deve então ser analisada de modo a determinar a sua precisão, ou seja, se essa foi satisfatória. Esse análise é feita a partir dos valores das duas menores distâncias obtidas, onde a diferença destes valores deve ser maior que determinado limiar definido. Essa restrição remove qualquer ambigüidade na classificação, uma vez que se existirem mais de uma classe próxima da região entende-se que essa pode fazer parte de qualquer uma das duas, tornando a classificação insatisfatória.

Ao termino da classificação, verifica-se o resultado obtido. Em caso de uma classificação satisfatória, a etapa de segmentação é executada, extraindo a região da imagem. Caso contrário a região é dividida pelo processo P, onde cada sub-região é classificada separadamente.

D. A Segmentação

Com o processo de classificação realizado, a segmentação se torna muito simples. Essa etapa do processo será responsável por extrair da imagem original, cada uma das regiões da imagem que pertençam à determinada classe, gerando como produto uma ou mais imagens, cada uma contendo elementos de uma mesma classe.

III. EXPERIMENTAÇÃO

Com toda a metodologia definida, o método de segmentação pôde ser desenvolvido. Como descrito anteriormente, a primeira etapa do processo será a etapa de estudo de imagens, responsável por gerar o conhecimento necessário para a segmentação. É necessário então que essa etapa seja executada para que a segmentação possa ser realizada.

É importante observar que os descritores das imagens são extraídos sobre regiões de tamanho específico desta e, este tamanho pode influenciar no valor destas medidas. Por exemplo, em regiões muito pequenas, espera-se que a medida de contraste seja mais baixa que a mesma medida em uma região maior, uma vez que quanto menor o tamanho, menor será o número de elementos e conseqüentemente, menor variação de luminosidade. Devido a esse problema, o estudo de imagens deverá extrair informações da imagem de acordo com diversos tamanhos, de modo que a classificação possa ser feita a partir do tamanho mais adequado.

Inicialmente, foram escolhidos quatro tamanhos diferentes para regiões, sendo esses quadrados com lados 8, 16, 32 e 64. O estudo foi então realizado trabalhando com as imagens da base de dados divididas em regiões para cada tamanho escolhido. Os gráficos das figuras 3 e 4 mostram os valores médios de cada descritor para todas as regiões das imagens de cada classe. Observa-se que tanto para a faixa de luminosidade, como para o contraste, os valores apresentam

valores muito próximos, o que pode dificultar a classificação em alguns casos. Dessa forma, optou-se por trabalhar com regiões de tamanhos 32x32 e 64x64.

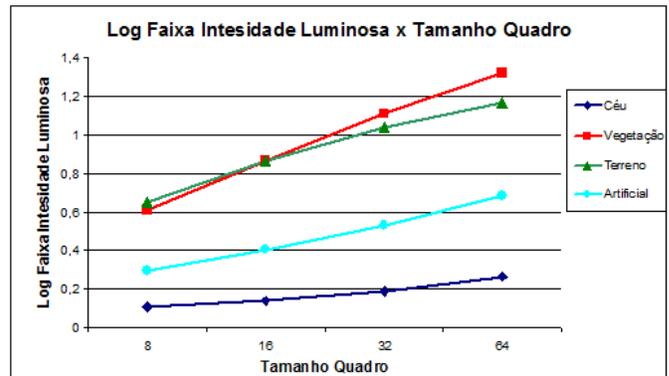


Figura 3. Gráfico exibindo a variação do Log da Faixa de Luminosidade para quadros de tamanho 8, 16, 32 e 64.

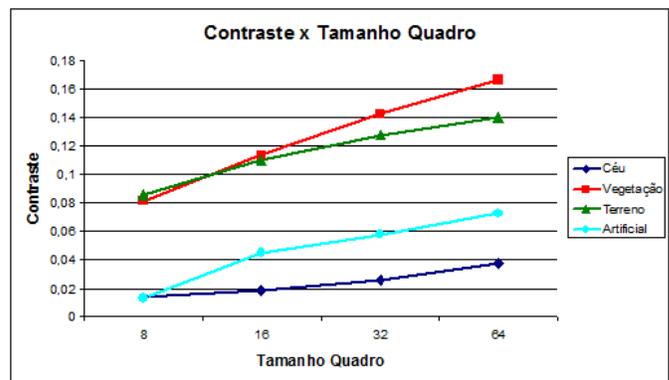


Figura 4. Gráfico exibindo a variação do Contraste RMS para quadros de tamanho 8, 16, 32 e 64.

Definidos os tamanhos do estudo, o processo P irá dividir a imagem em regiões de tamanho 64 (o maior tamanho escolhido) para a classificação. Sempre que constatada uma classificação insatisfatória, a região será dividida em quatro novas regiões, cada uma sendo classificada individualmente. Caso a região tenha tamanho igual a 32, essa não será mais dividida devido aos problemas discutidos anteriormente. Dessa forma, essa região não poderá ser classificada e conseqüentemente, não será segmentada.

Por fim, resta determinar o limiar de satisfabilidade da classificação, responsável por definir quando será necessário dividir a imagem em novas regiões ou quando essa poderá ser segmentada. Esse limiar será o valor mínimo aceitável da diferença entre as duas menores distancias, e seu valor é estimado para cada imagem entrada. Devido à dificuldade em se determinar de forma automática esse valor, esse foram definidos manualmente na tentativa de melhorar os resultados obtidos em cada caso.

A. Resultados

Com a metodologia definida e o conhecimento necessário para a classificação já gerado, o processo de segmentação foi executado como descrito na seção II. Os testes contaram com um total de 15 imagens, todas contendo elementos de duas ou mais classes estudadas. A experimentação do processo foi realizada apenas com imagens entrada contendo elementos das classes estudadas, de modo a medir a capacidade de classificação com mais precisão. Espera-se então que todos os elementos da cena sejam classificados e um pequeno número seja descartado devido a uma classificação insatisfatória.

A figura 5 mostra um dos resultados obtidos após a execução do método proposto usando a imagem 5(a) como entrada. Essa imagem contém elementos de três das quatro classes estudadas: céu, vegetação e estruturas criadas pelo homem. Analisando as imagens 5(b), 5(c), 5(d) e 5(e) com os resultados obtidos, percebe-se que grande parte dos elementos foram classificados corretamente, mas também um grande número de falsos positivos foram gerados, principalmente relacionados com vegetação. Em contra partida, a imagem 5(c) apresenta um número muito pequenos de elementos, uma vez que não existe nenhum elemento de terreno na imagem de entrada.

A figura 6 exibe outro resultado, onde a imagem entrada 6(a) contém elementos das mesmas três classes do resultado anterior. A partir das imagens 6(b), 6(c), 6(d) e 6(e) percebe-se que existe um grande número de falsos positivos relacionados principalmente com vegetação e estruturas construídas pelo homem. As demais classes apresentaram uma classificação satisfatória, onde o céu foi identificado quase por completo e um número muito pequeno de elementos de terreno foram segmentados erroneamente.

Um último resultado apresentado na figura 7, possui como entrada a imagem 7(a) que contém apenas elementos de céu e estruturas construídas pelo homem, gerando as imagens 7(b), 7(c), 7(d) e 7(e) como resultado. Esse resultado mostra que o método classificou com grande precisão elementos de céu e a grande maioria de elementos construídos pelo homem, mas ainda apresentou a mesma falha dos outros dois resultados anteriores: elementos de estruturas construídas pelo homem classificadas como vegetação.

IV. CONCLUSÃO

A segmentação de imagens é um dos grandes desafios nas áreas de visão computacional e processamento de imagens devido à sua grande importância para essas. O desenvolvimento de um processo para segmentação eficaz é de grande dificuldade e, devido a isso, grande parte dos métodos desenvolvidos são restritos à certos contextos. No desenvolvimento de um processo de segmentação de imagens de cenas naturais, a análise precisa das classes imagens de se mostrou de grande valia para o processo, permitindo determinar os aspectos relevantes dessas classes para o

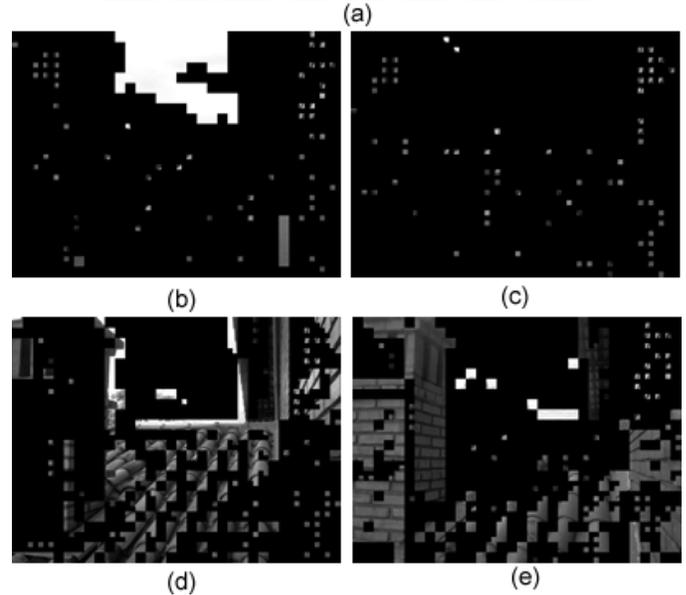


Figura 5. Resultado do processo de segmentação de imagens proposto para uma imagem entrada contendo elementos da classe céu, vegetação e construídos pelo homem. (a) Imagem entrada. (b) Classe Céu segmentada da imagem. (c) Classe Terreno segmentada da imagem. (d) Classe vegetação segmentada da imagem. (e) Classe Estrutura construída pelo Homem segmentada da imagem.

estudo de imagens. Isso tornou possível um estudo eficiente e conseqüentemente, uma classificação precisa na maioria dos casos. Além disso, percebe-se que o uso de descritores simples usados adequadamente podem ser responsáveis por um método de segmentação de imagens com resultados satisfatórios.

A partir dos resultados apresentados, percebe-se que, apesar da capacidade do método segmentar grandes regiões da imagem com precisão, ainda existem algumas falhas, principalmente relacionada à interposição de classes. Algumas classes se encontram muito próximas, o que faz com que o algoritmo apresente erros ao classificar elementos presentes nestas. Esse fato torna necessária a avaliação de outros métodos na tentativa de aumentar a precisão dos resultados e diminuir essa ambigüidade encontrada. Trabalhos futuros podem considerar novos aspectos estatísticos



(a)



(b)



(c)



(d)



(e)

Figura 6. Resultado do processo de segmentação de imagens proposto para uma imagem entrada contendo elementos da classe céu, vegetação e construídos pelo homem. (a) Imagem entrada. (b) Classe Céu segmentada da imagem. (c) Classe Terreno segmentada da imagem. (d) Classe vegetação segmentada da imagem. (e) Classe Estrutura construída pelo Homem segmentada da imagem.

de cenas naturais e artificiais além dos aqui descritos. Além disso, métodos de classificação baseados em inferência Bayesiana podem ser utilizados ao invés da distancia entre as amostras. O uso de inferência pode reduzir significativamente o número de falsos positivos e negativos, além de simplificar a estimação do limiar de satisfabilidade. Finalmente, o uso de tamanhos de janelas maiores do que os aqui descritos também podem beneficiar o método, tanto do ponto de vista de desempenho do sistema, como do ponto de vista da eficiência do resultado uma vez que a classificação de regiões maiores pode reduzir a quantidade de falsos negativos na classificação de uma classe.

REFERÊNCIAS

[1] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 1st ed. ADDISON-WESLEY, 1992.



(a)



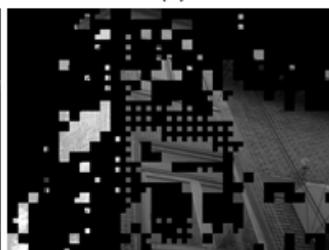
(b)



(c)



(d)



(e)

Figura 7. Resultado do processo de segmentação de imagens proposto para uma imagem entrada contendo elementos da classe céu e construídos pelo homem. (a) Imagem entrada. (b) Classe Céu segmentada da imagem. (c) Classe Terreno segmentada da imagem. (d) Classe vegetação segmentada da imagem. (e) Classe Estrutura construída pelo Homem segmentada da imagem.

[2] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294, 1993.

[3] J. Liu and Y. Yang, "Multiresolution color image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 689–700, 1994.

[4] H. Zhang, S. Cholleti, S. A. Goldman, and J. E. Fritts, "Meta-evaluation of image segmentation using machine learning," in *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Washington, DC, USA: IEEE Computer Society, 2006, pp. 1138–1145.

[5] W. S. Geisler, "Visual perception and the statistical properties of natural scenes," *Annual Review of Psychology*, vol. 59, no. 1, pp. 167–192, 2008. [Online]. Available: <http://dx.doi.org/10.1146/annurev.psych.58.110405.085632>

- [6] R. A. Frazor and W. S. Geisler, "Local luminance and contrast in natural images," *Vision Research*, vol. 46, 2006.