Thermal Image Segmentation in Studies of Wildlife Animals

Mauro dos Santos de Arruda, Bruno Brandoli Machado, Wesley Nunes Gonçalves Departamento de Computação – UFMS CEP: 79907-414, Ponta Porã, MS, Brasil Emails: mauro.arruda@aluno.ufms.br, {bruno.brandoli, wesley.nunes}@ufms.br

José Fernando Rodrigues Júnior Universidade de São Paulo – ICMC/USP CEP: 13566-590, São Carlos, SP, Brasil Email: junio@icmc.usp.br

Resumo—Thermal imaging analysis is an important tool in the study of wildlife animals. The segmentation of thermal images has not extensively explored by the Ecology and Biology communities. In this paper we propose a new approach for segmenting thermal images using the SLIC superpixel algorithm and connected component labeling. Experiments were performed on images taken over the behaviour activity of four mammal species. The results show that our approach has a great potential for partioning animals and background.

I. INTRODUÇÃO

A segmentação de imagens é uma etapa muito importante em sistemas de visão computacional. O objetivo é dividir uma imagem em regiões de interesse e regiões que permanecem como fundo. Entretanto, a segmentação de imagens é fortemente dependente do domínio de aplicação [1]. Durante as últimas décadas, o desenvolvimento de sistemas automáticos de visão computacional cresceu rapidamente usando câmeras visuais. Tais câmeras são consideradas dispositivos tradicionais que capturam a luz visível produzindo imagens coloridas ou em níveis de cinza. Em contrapartida, as câmeras térmicas são sensores passivos, ou de espectro não visível, que absorvem a radiação infravermelha emitida por todos os objetos da cena. Dessa maneira, imagens termográficas tem se tornado muito importante em análise de imagens por superarem problemas de iluminação.

A detecção e o monitoramento de animais silvestres usando imagens térmicas têm atraído a atenção de vários pesquisadores das áreas de Biologia e Ecologia. Existem três motivos para isso: a temperatura corporal dos animais evidencia-os durante qualquer hora do dia, e as imagens térmicas são um método seguro e não-invasivo [2], podendo por exemplo, determinar se um animal está doente sem a necessidade de sedar os animais. Isso faz com que se reduza significativamente o nível de estresse dos animais.

Existem diversos artigos na literatura que abordam a análise de imagens térmicas para monitorar animais. EnCoordenador de Projetos Laury Cullen IPÊ – Instituto de Pesquisas Ecológicas CEP: 13405-045, Piracicaba, SP, Brasil Email: laurycullen@gmail.com

João Henrique Pinheiro Dias, Cristina Cunha Garcia Companhia Energética de São Paulo CEP: 04447-011, São Paulo, SP, Brasil Emails: {joao.dias,cristina.garcia}@cesp.com.br

tretanto, grande parte deles não fazem o uso de métodos automáticos para segmentação dos animais. Em [3] é feita uma revisão de artigos que empregam imagens térmicas para analisar, visualmente, a termoregulação corporal dos animais. Em contraste, o outro grupo de trabalhos usam técnicas computacionais para segmentar os animais nas imagens. Em [4], Christiansen et. al. propuseram um sistema para detectar e identificar qual a espécie do animal usando características extraídas de operadores morfológicos e transformada discreta do coseno. As imagens aéreas foram classificadas usando k-vizinhos mais próximos. A principal desvantagem desse artigo é que o valor da limiarização foi estabelecido de forma manual para detecção dos animais. Em [5] foi proposto um sistema para rastrear animais usando o modelo de mistura de múltiplas Gaussianas nos canais RGB, porém usando imagens visuais. Davis et. al. [6] propuseram um nova abordagem para segmentação de imagens térmicas de pedestres usando mapas de saliência, seguido de um fechamento de regiões corporais com busca A* e preeenchimento das regiões com watershed.

Neste artigo nós propomos a segmentação de imagens térmicas com a presença de animais silvestres usando a abordagem baseada em superpixels. Existem na literatura diferentes abordagens baseadas em superpixels [7], [8], [9]. Aqui nós empregamos a técnica SLIC - Simple Linear Interative Clustering [9] por ela possuir um bom desempenho em termos de tempo e qualidade de segmentação quando comparada em experimentos reais. Basicamente, ela divide a imagem em regiões regulares usando a proximidade de cor em relação ao posicionamento espacial dos pixels. Dessa maneira, nós variamos o principal parâmetro da técnica que define o número de superpixels na imagem, combinando-os as regiões segundo o limiar do espaço de cor dos pixels na imagem. Para os experimentos nós usamos um conjunto de imagens térmicas de animais silvestres, dividido em quatro espécies: onça pintada, cervo do Pantanal, cateto e cachorro-do-mato. As principais contribuições deste artigo são:

Nós desenvolvemos uma abordagem para detectar

animais silvestres em imagens térmicas usando superpixels.

- Nós comparamos o desempenho de diferentes técnicas baseadas em superpixel.
- Nós conduzimos uma avaliação dos parâmetros das técnicas testadas.

Este artigo está dividido em cinco seções. A Seção II apresenta como as imagens térmicas são capturadas, seguido da descrição do algoritmo SLIC. Na Seção III é descrita a abordagem proposta para o monitoramento de animais usando imagens térmicas. A validação dos experimentos e a comparação dos resultados com outros algoritmos são discutidos na Seção IV. Finalmente, as conclusões e trabalhos futuros são apresentados na Seção V.

II. METODOLOGIA

A. Imagens Térmicas

Imagens térmicas registram a radiação infravermelha emitida pelos objetos com uma temperatura acima do zero absoluto. Conforme a intensidade emitida por um objeto aumenta, maior o contraste dele na imagem térmica. Assim, câmeras térmicas podem ser usadas durante o dia e a noite. As câmeras térmicas foram desenvolvidas inicialmente para o uso da indústria militar em tarefas que exigiam visão noturna [10], [11]. Entretanto, nas últimas décadas as imagens térmicas têm sido usadas em uma vasta gama de aplicações, tais como inspeção da qualidade de alimentos [12], detecção por calor em processos industriais [13], detecção de vazamento de gás ou focos de incêndios [14] e monitoramento em estudos com animais [15].

O uso de imagens térmicas possui algumas vantagens sobre problemas clássicos da visão computacional. As imagens capturadas por câmeras visuais dependem da fonte de energia visível no espectro eletromagnético, tornando-se bastante sensíveis às condições de iluminação e sombras. Por outro lado, imagens térmicas tem seus próprios desafios, incluindo as condições do clima, distância do objeto de interesse em relação à câmera e o efeito *halo* que aparece ao redor de objetos muito quentes ou muito frios. Na Figura 1 são mostradas duas imagens para a mesma cena no monitoramento de animais, uma delas capturada com o sensor térmico e a outra com o sensor CCD visível. As propriedades térmicas do animal em relação ao fundo são bastante diferentes. Nota-se que na segunda linha, dificilmente o animal na sombra poderia ser identificado usando abordagens visuais tradicionais.

B. SLIC - Simple Linear Interative Clustering

O SLIC [9] (Simple Linear Interative Clustering) é uma das técnicas baseadas em gradiente que divide a imagem em regiões com número aproximado de pixels, também chamadas de superpixels. A técnica possui dois parâmetros para determinar o número de pixels de cada região. O primeiro parâmetro determina a quantidade desejada de superpixels k na imagem, enquanto m controla o nível de compactação do superpixel definido pela relação entre a similaridade dos atributos de cor e a proximidade espacial. Basicamente, o algoritmo usa cinco dimensões representadas no vetor $\nu = [L, a, b, x, y]$, tal que os valores [L, a, b] são os atributos que representam a cor no espaço CIELAB e [x, y] são as coordenadas espaciais dos pixels.

O algoritmo SLIC inicia com k centros na imagem $C_k = [L_k, a_k, b_k, x_k, y_k]$, amostrados em uma grade regular S e espaçados em intervalos $S = \sqrt{N/k}$, tal que N é a quantidade de pixels na imagem. Os centros são movidos para a posição com menor valor de gradiente resultante da vizinhança de 3×3 pixels, evitando assim pixels ruidosos ou borda. Em seguida, o algoritmo calcula a distância euclidiana d de cada pixel i em relação aos centros C_k de uma região de busca de tamanho $2S \times 2S^{-1}$.

Para agrupar os pixels no espaço de atributos com cinco dimensões, a medida da distância euclidiana dos vetores Labxy se torna inconsistente. Dessa forma, as distâncias são calculadas separadamente para a proximidade de cor (Equação 1) e da posição espacial (Equação 2).

$$d_c = \sqrt{(L_k - L_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2}$$
(1)

$$d_s = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2}$$
(2)

Em seguida, é necessário normalizar a proximidade de cor em relação à distância espacial, somando-se as distâncias euclidianas para o espaço de cor d_c e da posição espacial d_s dos pixels. A distância total é então calculada pela Equação 3, dada por:

$$d = d_c + m \frac{d_s}{S} \tag{3}$$

tal que m permite ponderar a relação da proximidade de cor e posição espacial dos pixels. Para valores de m maiores, a proximidade espacial é privilegiada e resulta em superpixels mais compactos, enquanto valores para m mais baixos os grupos privilegiam a similaridade da cor dos pixels, criando superpixels não regulares em forma e tamanho.

Cada pixel *i* é então associado ao centro do grupo mais próximo. Em seguida, um novo centro é definido por meio da média dos atributos de cada pixel $[L_i, a_i, b_i, x_i, y_i]$ de todos os pixels pertencentes àquela região. O critério de convergência do algoritmo é determinado pelo erro residual entre a posição do centro da última iteração e a posição do centro atual. Em alguns casos são usados apenas um número específico de iterações ao invés do erro. A próxima seção descreve a abordagem proposta baseada em segmentação usando a definição de superpixels SLIC.

III. UMA NOVA PROPOSTA PARA O MONITORAMENTO DE Animais Silvestres usando Segmentação de Imagens Térmicas

Nesta seção nós apresentamos uma nova proposta para a segmentação de imagens térmicas com o objetivo de monitorar a atividade de animais silvestres. A nossa abordagem é baseada no algoritmo de segmentação SLIC - *Simple Linear Iterative Clustering* que foi proposto em [9]. A ideia aqui consiste

¹[9] definiu a área de busca $2S \times 2S$ para aumentar a velocidade de processamento.



Figura 1. Exemplos de imagens térmicas usadas no monitoramento de animais. Na primeira cena é mostrado três catetos, e na segunda três macacos. Na coluna (a) são mostradas as imagens térmicas e, na coluna (b), são mostradas as mesmas imagens visuais.

em agrupar os pixels da imagem em regiões que possuam similaridade baseada na emissão de calor dos corpos dos animais. Uma vez que o calor do corpo é maior, os atributos de cor são definidos com mais intensidade segundo um mapa de calor da câmera térmica. O superpixel permite usar as características locais dos pixels da imagem, ao contrário do algoritmo k-médias [16] que considera todo espaço de busca no agrupamento.

A Figura 2 mostra o diagrama da nossa abordagem. No passo (1) é a realizada a captura das imagens usando uma câmera térmica. No passo (2), são gerados os superpixels na imagem por meio dos número de grupos desejado. Uma vez que o número de superpixels é definido parametricamente, o algoritmo é inicializado distribuindo os centros em intervalos regulares na imagem. Além do número de superpixels desejados, pode-se variar o nível de compactação dos grupos. Os pixels são então agrupados usando a métrica de distância euclidiana em um região limitada de busca, considerando um espaço de cinco dimensões [labxy] extraído de cada pixel, tal que [lab] é o vetor de atributos de cor no espaço CIELAB e [xy] é a coordenada espacial dos pixels. Cada pixel é então associado ao centro do grupo mais próximo e o novo centro é escolhido por meio da média de todos os pixel pertencentes àquela região. Isto é repetido até que o critério de convergência do algoritmo seja alcançado. Aqui, nós não usamos o erro residual entre a posição do centro da última iteração e a posição do centro atual por, em alguns casos, levarem um número muito alto de repetições. Assim, foi definido um valor fixo de iterações, assim como pode ser praticado no algoritmo kmédias.

Tanto o valor do número de superpixels, quanto o valor de ponderação pode formar grupos mais compactos e homogêneos. Por fim, no passo (3), nós convertemos a imagem para níveis de cinza e, em seguida, nos consideramos as regiões de maior quantidade de pixels usando o algoritmo de componentes conectados, com vizinhança-de-8 pixels, e a saída são regiões rotuladas na imagem.

IV. EXPERIMENTOS

A. Conjunto de Imagens

O conjunto de imagens térmicas é composto por cenas com dimensão de 1149×861 pixels. As imagens foram capturadas com a câmera térmica FLIR E60, com lente FOL18. Os parâmetros da câmera térmica foram ajustados para os valorespadrão de 0,98 de emissividade, 60% de umidade relativa e 25° C de temperatura atmosférica. As imagens termográficas foram capturadas em temperatura ambiente, i.e., elas não sofrem influência do efeito *halo*. As 40 imagens analisadas de animais silvestres são divididas em quatro espécies silvestres de mamíferos: onça pintada, cervo do Pantanal, cateto e cachorro-do-mato.



Figura 2. Diagrama da nossa proposta para segmentação de imagens térmicas.

B. Métricas de Avaliação

Para avaliar a nossa abordagem, este trabalho empregou as métricas descritas no trabalho publicado em [1]. Basicamente, o resultado da segmentação de imagens foi comparado com as imagens de referência, chamadas de ground-truth. A comparação foi feita para cada pixel entre a imagem segmentada e a de referência, resultando em quatro possíveis rótulos: (1) verdadeiro positivo, (2) falso positivo, (3) verdadeiro negativo e (4) falso negativo. Um pixel pode ser verdadeiro positivo (vp) ou verdadeiro negativo (vn) quando o resultado da segmentação corresponde ao ground-truth. Por outro lado, um falso positivo (fp) ocorre quando o resultado da segmentação é diferente da imagem de referência, bem como ocorre para um falso negativo (fn). Uma vez quantificado o número dos rótulos para cada imagem, foram calculadas três métricas de avaliação: a acurácia (AC), a precisão (PR), a taxa das predições positivas (TPP). Tais medidas são formalizadas na Equações 4 e 5.

$$AC = \frac{vp + vn}{vp + fn + fp + vn} \quad , \quad PR = \frac{vp}{vp + fp} \qquad (4)$$

$$TPP = \frac{vp + fp}{vp + fn + fp + vn} \tag{5}$$

C. Resultados

A abordagem proposta aqui foi comparada com a técnica de Mori [8] e *Normalized Cut* (NCut) [7]. Os experimentos foram realizados sobre o conjunto de imagens térmicas de animais silvestres (descrito na Seção IV-A). Conforme as imagens de monitoramento foram capturadas em temperatura ambiente, elas não sofreram a influência do efeito *halo*. As Figuras 3, 4 e 5 mostram os resultados para as técnicas SLIC, de Mori e *Normalized Cut*, respectivamente. Cada técnica foi avaliada quantitativamente por três métricas de segmentação (descritas na Seção IV-B). Para todos os gráficos, no eixo-x é mostrado o número k de superpixels, variando de 100 até 3500, exceto para a técnica NCut devido seu elevado custo computacional. No eixo-y é mostrado o valor médio de acerto segundo a métrica de avaliação.

A técnica *Normalized Cut* segmenta uma imagem por meio da minimização da função custo de um grafo ponderado. A

função custo é calculada pelo número total de arestas de cada partição do grafo. Por diversos anos esta técnica foi considerada o estado-da-arte em segmentação de imagens. Porém, ela demanda um alto custo computacional quando trata imagens com dimensões maiores que 300x300 pixels, pois o tamanho do grafo aumenta consideravelmente, uma vez que sua complexidade é $O(N^3)$, onde N é o número de pixels. Assim, só foi possível gerar os resultados para as quantidades de grupos $k = \{100, 200\}$. Para valores maiores a técnica causa falta de memória. Vale mencionar que a técnica NCut trabalha sobre imagens em níveis de cinza, assim para esse caso, nós convertemos o conjunto de imagens para níveis de cinza.

A técnica de Mori [8] define os superpixels da imagem usando um grafo que calcula a similaridade dos grupos usando histogramas de cores sobre o espaço de cores CIELAB. O agrupamento dos pixels similares é feito pelo algoritmo kmédias usando a distância da escavadeira. Uma limitação desta técnica também é o custo computacional na ordem de $O(N^8)$, ao contrário da técnica SLIC que é O(N), cujo N é o número de pixels da imagem.

O resultado quantitativo da métrica acurácia, mostrado na Figura 3(a), é ligeiramente superior para valores de k pequenos (para $k \le 200$), se comparado com as técnicas de Mori e NCut. A diferença no gráfico é pequena devido ao número de pixels total da imagem ser grande, por exemplo, cada imagem tem por volta de 989 mil pixels. Nesse caso, se a técnica errar um grupo, ou superpixels, o resultados ainda tende a ser próximo pela grande quantidade de pixels. Para a métrica do TPP - taxa das predições positivas - o valor da medida quantitativa se mantém constante para quase todos os valores de k.

V. CONCLUSÃO

Neste artigo nós propomos uma nova abordagem para a detecção de animais silvestres baseada na segmentação de superpixels. Uma vez particionados os objetos de interesse na imagem, nós segmentamos as regiões com maior número de pixels usando componentes conexos com conectividade-de-8 vizinhos. A segmentação de animais teve uma acurácia e precisão satisfatórias para a segmentação dos animais analisados. Nós avaliamos o principal parâmetro que particiona os objetos em regiões. A técnica SLIC apresentou performance ligeiramente superior para os animais analisados baseado



Figura 3. Resultados usando a nossa proposta baseada na técnica SLIC [9].



Figura 4. Resultados usando a técnica de Mori proposta em [8].



Figura 5. Resultados usando a técnica Nomalized Cut [7].

nas métricas de validação. Embora a diferença numéria seja pequena pela grande quantidade de pixels, a técnica SLIC teve melhor desempenho quantitativo, i.e. visual, em regiões próximas as pastas, definindo melhor a silhueta do animal. Vale resssaltar que para valores altos de superpixels, a relação espacial prevalece sobre a proximidade de cor, produzindo superpixels menores que podem não aderir as regiões de fronteira da imagem. Cabe ainda ressaltar que a captura das imagens térmicas podem sofrer variações dependendo das condições climáticas e temperatura da espécie do animal. Isso pode afetar os resultados da abordagem proposta aqui no que diz respeito à qualidade da segmentação da silhueta corporal do animal. Como trabalhos futuros, nós pretendemos capturar e testar a mesma abordagem em dias com precipitação e até outras espécies de animais, tais como aves.

AGRADECIMENTOS

Os autores agradecem o suporte financeiro da CESP através do projeto: SENSORIAMENTO REMOTO NO LEVAN-TAMENTO DE GRANDES MAMÍFEROS SILVESTRES. CÓDIGO ANEEL PD- 0061-0028/2011 e ao IPÊ – Instituto de Pesquisas Ecológicas pela execução.

REFERÊNCIAS

- [1] B. B. Machado, W. N. Gonçalves, H. Pistori, J. de Andrade Silva, K. P. de Souza, B. Toledo, and W. Tessaro, "Segmentação de gestos e camundongos por subtração de fundo, aprendizagem supervisionada e watershed," Curitiba, PR, pp. 303–324, September 2012.
- [2] I. Knizkova, P. Kunc, K. Gürdil, Y. Pınar, Ç. Selvi et al., "Applications of infrared thermography in animal production," Anadolu Journal of Agricultural Sciences, vol. 22, no. 3, pp. 329–336, 2007.

- [3] J. Cilulko, P. Janiszewski, M. Bogdaszewski, and E. Szczygielska, "Infrared thermal imaging in studies of wild animals," *European Journal of Wildlife Research*, vol. 59, no. 1, pp. 17–23, 2013. [Online]. Available: http://dx.doi.org/10.1007/s10344-012-0688-1
- [4] P. Christiansen, K. A. Steen, R. N. Jørgensen, and H. Karstoft, "Automated detection and recognition of wildlife using thermal cameras," *Sensors*, vol. 14, no. 8, pp. 13778–13793, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/8/13778
- [5] D. Tweed and A. Calway, "Tracking multiple animals in wildlife footage," in *16th International Conference on Pattern Recognition*, vol. 2, 2002, pp. 24–27.
- [6] J. Davis and V. Sharma, "Background-subtraction in thermal imagery using contour saliency," *International Journal of Computer Vision*, vol. 71, no. 2, pp. 161–181, 2007. [Online]. Available: http://dx.doi.org/10.1007/s11263-006-4121-7
- [7] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, Aug 2000.
- [8] G. Mori, "Guiding model search using segmentation," in *Tenth IEEE International Conference on Computer Vision*, vol. 2, Octubre 2005, pp. 1417–1423.
- [9] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Susstrunk, "Slic superpixels compared to state-of-the-art superpixel methods," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2274–2282, November 2012. [Online]. Available: http://dx.doi.org/10. 1109/TPAMI.2012.120
- [10] M. Kastek, R. Dulski, P. Trzaskawka, T. Sosnowski, and H. Madura, "Concept of infrared sensor module for sniper detection system," in 35th International Conference on Infrared Millimeter and Terahertz Waves, September 2010, pp. 1–2.
- [11] R. Siegel, "Land mine detection," *IEEE Instrumentation Measurement Magazine*, vol. 5, no. 4, pp. 22–28, December 2002.
- [12] D. S. J. R. Vadivambal, "Applications of thermal imaging in agriculture and food industry - a review," *Food and Bioprocess Technology*, vol. 4, no. 2, pp. 186–199, 2011. [Online]. Available: http://dx.doi.org/10.1007/s11947-010-0333-5
- [13] Y.-M. H. Ng, M. Yu, Y. Huang, and R. Du, "Diagnosis of sheet metal stamping processes based on 3-d thermal energy distribution," *IEEE Transactions on Automation Science and Engineering*, vol. 4, no. 1, pp. 22–30, January 2007.
- [14] J.-H. Hwang, S. Jun, S.-H. Kim, D. Cha, K. Jeon, and J. Lee, "Novel fire detection device for robotic fire fighting," in *International Conference* on Control Automation and Systems (ICCAS), October 2010, pp. 96– 100.
- [15] G. Tattersall and V. Cadena, "Insights into animal temperature adaptations revealed through thermal imaging," *The Imaging Science Journal*, vol. 58, no. 5, pp. 261–268, 2010.
- [16] J. A. Hartigan and M. A. Wong, "Algorithm as 136: A k-means clustering algorithm," *Applied statistics*, vol. 28, no. 1, pp. 100–108, 1979.