

Uma ferramenta para a análise morfológica de embriões de insetos

Luciano A. Digiampietri¹, Beatriz Teodoro¹

¹Escola de Artes, Ciências e Humanidades – Universidade de São Paulo (USP)
Av. Arlindo Bértio, Ermelino Matarazzo – 03828-000 – São Paulo – SP – Brasil

digiampietri@usp.br

Abstract. *This paper presents a tool for the morphological analysis of insect embryos in order to identify the start and end positions of each embryo's stripe. This tool has been tested with real data and helped to validate the influence of some genes on the embryonic development of insects.*

Resumo. *Este artigo apresenta uma ferramenta para a análise morfológica de embriões de inseto que visa a identificar as posições iniciais e finais de cada uma das faixas do embrião. Esta ferramenta foi testada com dados reais e ajudou a validar a influência de alguns genes sobre o desenvolvimento embrionário de insetos.*

1. Introdução e Motivação

Diversas aplicações da ciência atual envolvem o processamento de imagens digitais, como por exemplo, o processamento de imagens médicas ou pesquisas em biologia envolvendo imagens de animais (Andrioli, 2002; Andrioli, 2004; Wang 2004). Nos últimos anos, com o barateamento e a popularização de máquinas fotográficas acopladas a microscópios, este tipo de aplicação se tornou ainda mais relevante.

O processamento de imagens digitais é uma subárea da ciência da computação que apresenta diversas vertentes e diversos desafios, dependendo do tipo de imagem que será utilizado, dos ruídos existentes nas imagens e dos objetivos do sistema que se pretende produzir [Galbiati 1990].

Este artigo apresenta uma ferramenta para o processamento de imagens de embriões de insetos, a fim de extrair características morfológicas dessas imagens. Com essas características, é possível identificar padrões e comparar as imagens auxiliando na identificação de padrões nestas imagens. Em especial, a ferramenta desenvolvida foi utilizada para comparar o posicionamento das faixas de embriões de *Drosophila melanogaster* a fim de verificar se existe uma correlação entre variações genéticas e o posicionamento das faixas. Este fato foi comprovado, conforme pode ser visto em [Andrioli et al. 2012].

O restante deste artigo está organizado da seguinte maneira. A Seção 2 sumariza os trabalhos correlatos. A Seção 3 apresenta a metodologia utilizada. A Seção 4 descreve a ferramenta desenvolvida. Por fim, a Seção 5 contém as conclusões.

2. Trabalhos Correlatos

Atualmente, o processamento de imagens está sendo usado em diversos domínios e, em particular, em projetos de e-Science. Devido a complexidade e diferenças de natureza dos

problemas tratados, cada solução costuma ser bastante específica compartilhando apenas o ferramental básico de processamento de imagens (como conjuntos de filtros tipicamente aplicados às imagens).

Em [Cope et al. 2012], os autores realizaram uma revisão dos principais trabalhos sobre classificação de plantas usando análise morfológica. Os autores identificaram diversos trabalhos interessantes para a realização desta tarefa, porém ainda há diversos desafios para a criação de um sistema preciso e robusto para esta tarefa.

Outra aplicação na qual o processamento de imagens vem sendo bastante usada é na contagem de células, por exemplo, para identificar o número de células tumorais ou a quantidade de vírus ou bactérias em uma dada amostra [Moorman e Dong 2012]. Esta é uma atividade que há algumas décadas vinha sendo realizada manualmente, sendo bastante trabalhoso e susceptível a erros.

Os autores de [Lu et al. 2011] usam o processamento de imagens para a classificação automática de algumas espécies de insetos. Para isto, 17 pontos de referência são extraídos e comparados.

A análise das faixas transversais presentes nos embriões de diversas espécies de insetos tem sido realizada há diversos anos, porém o processo costuma ser feito manualmente [Andrioli et al. 2002]. Este fato motivou o desenvolvimento desta ferramenta que foi testada e validada utilizando-se dados reais e cujos resultados foram utilizados para demonstrar a influência de alguns genes na formação embrionária de moscas da espécie *Drosophila melanogaster* [Andrioli et al. 2012].

3. Metodologia

A metodologia deste artigo foi composta por quatro atividades: estudo da literatura correlata, desenvolvimento da ferramenta, calibração dos parâmetros utilizando-se dados reais e validação dos resultados da ferramenta também utilizando dados reais e com auxílio de um especialista no domínio.

4. Ferramenta Desenvolvida e Discussão dos Resultados

A ferramenta desenvolvida recebe como entrada a imagem de um embrião e retorna as posições inicial e final de cada uma de suas faixas transversais. Para isto, um processamento dividido em nove passos é realizado:

(i) conversão da imagem: a imagem de entrada é convertida para uma matriz de *bytes* onde o valor de cada byte corresponde a média dos tons vermelho e verde do ponto correspondente da imagem original (a utilização dos tons de vermelho e verde foi feita após testes com as diferentes cores e codificações);

(ii) remoção de fundo: o fundo é removido a partir de um filtro que utiliza a mediana dos valores dos bytes e um limiar para identificar se cada byte é ou não fundo;

(iii) aumento do contraste: o contraste é aumentando de forma a maximizar a amplitude dos valores dos bytes da matriz para todo o domínio de um byte (de 0 de 255);

(iv) filtro passa-baixa: um filtro passa baixa é executado sobre toda a matriz de forma a minimizar o ruído da imagem;

(v) identificação das extremidades dos embriões: a imagem bidimensional de embriões de insetos tipicamente tem um formato semelhante a uma elipse, neste passo as extremidades do embrião são identificadas pois esta informação é necessária para a busca das

posições das faixas transversais;

(vi) alinhamento da imagem: neste passo a imagem é rotacionada de forma que as duas extremidades do embrião fiquem alinhadas;

(vii) cálculo do histograma: o histograma de cores da região central da imagem de uma extremidade até a outra é calculado;

(viii) identificação dos vales e picos do histograma: um algoritmo iterativo é utilizado para computar o número de inversões (vales e picos) que ocorrem no histograma. Já que o número de faixas é um parâmetro conhecido do programa, o número de vales mais picos deverá ser igual a duas vezes o número de faixas mais um (ver Figura 1), porém, devido à ruídos remanescentes na imagem é possível que este número seja maior, se isso ocorrer, um filtro passa baixa é iterativamente executado sobre o histograma até que o número de inversões seja igual ao número esperado;

(x) cálculo das posições das faixas: com base nas posições dos vales e picos identificadas no passo anterior, o posicionamento de cada faixa é calculado. Cada faixa fica entre dois vales e a diferença dos valores do histograma é utilizada para aferir o ponto exato de início e fim da faixa correspondente.

A Figura 1 ilustra os conceitos utilizados no algoritmo em uma imagem real na qual há sete faixas transversais.

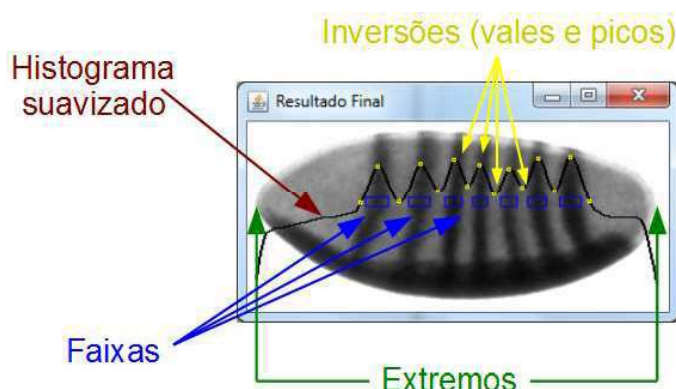


Figura 1. Indicação dos conceitos utilizados na ferramenta

A Figura 2 apresenta o resultado gráfico de algumas das etapas de processamento executadas pela ferramenta. Na primeira linha, a imagem a esquerda corresponde a imagem original; a imagem central contém a imagem convertida utilizando-se as bandas vermelho e verde da imagem original; e a imagem da direita corresponde ao resultado da remoção do fundo. Na segunda linha desta imagem, ao lado esquerdo é apresentada a imagem com o contraste ampliado; ao centro há a imagem suavizada após a utilização do filtro passa-baixa; e, por fim, a direita, é apresentada a imagem final, com o histograma, vales, picos e as faixas identificados.

5. Conclusões

Este artigo apresentou uma ferramenta para a identificação das posições inicial e final das faixas transversais de embriões de insetos.

A ferramenta foi testada com dados reais de embriões de *Drosophila melanogaster* e validada por um especialista do domínio. Os resultados obtidos pela ferramenta

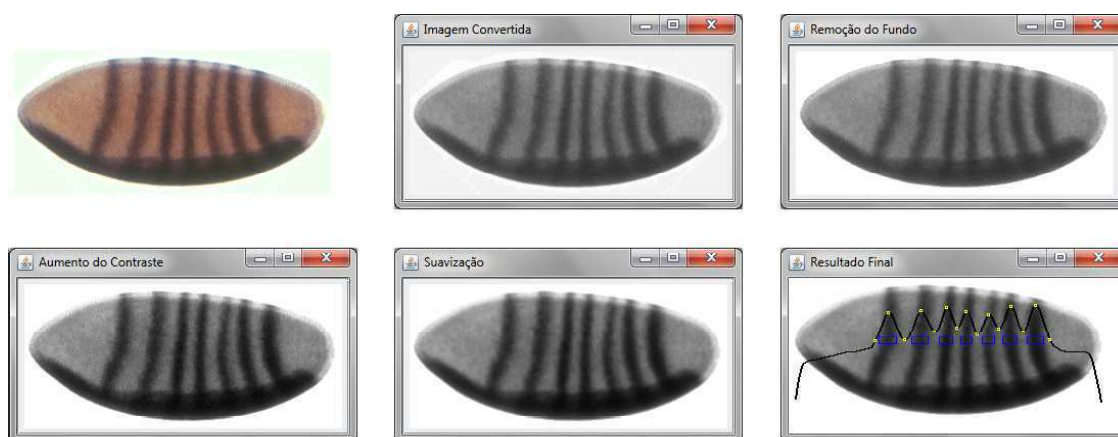


Figura 2. Imagens intermediárias e final resultantes da execução da ferramenta

confirmaram a hipótese de que certos genes têm influência sobre o posicionamento das faixas dos embriões desta espécie [Andrioli et al. 2012].

Como trabalhos futuros pretende-se estender a ferramenta para lidar com o problema de classificação de espécies com base em embriões.

Agradecimentos

O trabalho apresentado neste artigo foi parcialmente financiado pela FAPESP, pelo CNPq, pela CAPES, pelo Programa de Educação Tutorial do MEC e pela Pró-Reitoria de Graduação da Universidade de São Paulo.

Referências

- Andrioli, L. P., Digiampietri, L. A., de Barros, L. P., e Machado-Lima, A. (2012). Hucklebein is part of a combinatorial repression code in the anterior blastoderm. *Developmental Biology*, 361(1):177 – 185.
- Andrioli, L. P. M., Vasisht, V., Theodosopoulou, E., Oberstein, A., e Small, S. (2002). Anterior repression of a drosophila stripe enhancer requires three position-specific mechanisms. *Development (Cambridge)*, 129(21):4931–4940.
- Cope, J. S., Corney, D., Clark, J. Y., Remagnino, P., e Wilkin, P. (2012). Review: Plant species identification using digital morphometrics: A review. *Expert Syst. Appl.*, 39(8):7562–7573.
- Galbiati, L. J. (1990). *Machine Vision and Digital Image Processing Fundamentals*. Prentice Hall.
- Lu, W., Lingwang, G., Zuorui, S., Lili, H., e Qian, X. (2011). Research on landmark extraction technology in identification of fruit flies (diptera: Tephritidae). In *Natural Computation (ICNC), 2011 Seventh International Conference on*, volume 3, pages 1681–1685.
- Moorman, M. e Dong, A. (2012). Automated viral plaque counting using image segmentation and morphological analysis. In *Proceedings of the 2012 IEEE International Symposium on Multimedia, ISM '12*, pages 157–160, Washington, DC, USA. IEEE Computer Society.