

Síntese de Imagens de Exames de Eletroforese de Hemoglobina

Maria C. A. de Andrade, Jugurta Montalvão, Ana C. S. Aragão

Resumo— A eletroforese de hemoglobina é uma das principais ferramentas para o diagnóstico de hemoglobinopatias e talassemias. No entanto, um desafio enfrentado por pesquisadores e estudantes é a falta de disponibilidade de imagens desses exames em bancos de dados abertos. Com isso, esse trabalho apresenta um simulador desenvolvido em GNU Octave para síntese de imagens de exames de eletroforese de hemoglobina.

Palavras-Chave— Síntese de imagens, eletroforese de hemoglobina, modelos fenomenológicos.

Abstract— Hemoglobin electrophoresis is one of the main tools for diagnosing hemoglobinopathies and thalassemias. However, a challenge faced by researchers and students is the lack of availability of images from these exams in open databases. Therefore, this work presents a simulator developed in GNU Octave for synthesizing images of hemoglobin electrophoresis exams.

Keywords— Image synthesis, hemoglobin electrophoresis, phenomenological models.

I. INTRODUÇÃO

A eletroforese consiste na migração de partículas quando submetidas a um campo elétrico, fazendo com que as proteínas migrem para o polo do eletrodo de carga elétrica oposta aos aminoácidos que compõem as moléculas, através da atração eletrostática [1]. Em meio alcalino, por exemplo, os aminoácidos que compõem as moléculas de hemoglobina se tornam negativamente carregados. Sendo assim, ao se conhecer as propriedades físico-químicas de cada tipo de hemoglobina (Hb), é possível fazer a detecção da mesma ao analisar o padrão resultante no suporte onde as amostras são aplicadas.

Em muitos dos trabalhos acadêmicos encontrados na literatura, foi necessária a execução de todo o processo da eletroforese, desde a coleta das amostras até a obtenção final das imagens das tiras do suporte, a exemplo de [2]. Isso requer a colaboração de especialista e aprovação dos respectivos comitês que atuam para garantir a ética e segurança dos processos relacionados ao exame e ao paciente.

Uma alternativa à obtenção de dados próprios consiste no levantamento de imagens em bancos de dados abertos, porém, a quantidade de imagens disponíveis e de livre acesso acaba não sendo suficiente para representar estatisticamente o fenômeno alvo.

Maria C. A. de Andrade, Departamento de Engenharia Elétrica, Universidade Federal de Sergipe, São Cristóvão-SE, e-mail: aragao.carol@academico.ufs.br; Jugurta Montalvão, Departamento de Engenharia Elétrica, Universidade Federal de Sergipe, São Cristóvão-SE, e-mail: jmontalvao@academico.ufs.br; Ana C. S. Aragão, Departamento de Biologia, Universidade Federal de Sergipe, São Domingos-SE, e-mail: analuiza0506@academico.ufs.br.

Para fenômenos cujas leis físico-químicas são bem conhecidas e modeladas, imagens artificiais podem ser sintetizadas de acordo com essas leis, resultando em uma aplicação da técnica mais refinada e, potencialmente, mais poderosa do aumento de dados. Esse trabalho propõe a elaboração de um simulador capaz de gerar imagens realistas de exames de eletroforese de hemoglobina compatíveis com as obtidas na técnica de eletroforese alcalina de hemoglobina em gel de agarose. Com isso, espera-se que esse simulador possa ser utilizado em futuros trabalhos nos quais sejam necessárias imagens desse exame.

II. DESENVOLVIMENTO

Para a elaboração do simulador, foi necessário, a partir do estudo do fenômeno, elaborar um modelo fenomenológico que explica os principais pontos de interesse, sendo as diferentes mobilidades das hemoglobinas e os padrões formados a partir da aplicação da corrente elétrica no suporte de fracionamento. Posteriormente, utilizou-se o modelo para elaborar um algoritmo e implementá-lo na linguagem GNU Octave (<https://octave.org/>).

A. O Modelo

O foco principal do modelo consistiu na capacidade de replicar os principais aspectos da imagem de um exame de eletroforese alcalina de hemoglobina em gel de agarose. Nesse caso, a ideia consistiu em simular as partículas das hemoglobinas como um conjunto de pontos em um plano cartesiano. Sendo assim, ao variar a quantidade de pontos associados a cada constante de velocidade presente no vetor \mathbf{v} , é possível ajustar à quantidade de pontos que migram até determinada região na imagem, construindo assim as zonas correspondentes a cada tipo de hemoglobina. Assim, definiu-se o vetor \mathbf{v} como:

$$\mathbf{v} = [v_1, v_2, v_3 \dots v_N], \quad (1)$$

onde v_N corresponde à velocidade associada a cada tipo de hemoglobina e N à quantidade de velocidades pré-definidas. Em \mathbf{v} são definidas 4 velocidades principais, a partir da comparação entre a imagem gerada e uma imagem de referência real, correspondentes às hemoglobinas A, S, F e C/A2.

Além disso, para a definição da quantidade de partículas que seriam associadas a cada velocidade, construiu-se um vetor \mathbf{c} da seguinte forma:

$$\mathbf{c} = [\mathbf{a}_1(1 \times q_1), \mathbf{a}_2(1 \times q_2), \dots, \mathbf{a}_n(1 \times q_n)], \quad (2)$$

sendo qn a quantidade de pontos e $\mathbf{a}_n(1 \times qn)$ um conjunto de números aleatórios onde cada elemento é uma instância da variável aleatória definida como:

$$A_n = v_n + \sigma Z, \quad (3)$$

onde A_n é um número aleatório; v_n é a velocidade do índice n do vetor \mathbf{v} , que corresponderá à média da distribuição; σ é o desvio padrão da distribuição; Z é uma variável aleatória normal padrão, com média 0 e desvio padrão 1.

Para simular o processo de migração das partículas, as coordenadas de cada ponto devem ser atualizadas considerando as suas respectivas velocidades associadas, bem como os demais fatores que influenciam na movimentação das partículas. Um desses fatores corresponde à difusão, que provoca o espalhamento das zonas em ambos os eixos do plano. A forma selecionada para descrever esse fenômeno consistiu na simulação computacional “Passeio Aleatório” [3], com um viés numa dada direção, simulando assim o efeito do campo elétrico sobre as partículas em dispersão.

B. Geração das Imagens

A conversão do conjunto de pontos em imagem foi realizada a partir da utilização de conceitos relacionados a elaboração de histogramas bidimensionais, os quais são uma representação gráfica da distribuição de dados em um espaço bidimensional, para contar e organizar matricialmente a quantidade de pontos existentes em cada partição do plano. As dimensões dessas partições são definidas a partir da resolução desejada da imagem.

Além disso, para trazer à imagem gerada um aspecto mais próximo do real, foi realizada a aplicação de um filtro digital gaussiano com $\sigma = 1,85$ e dimensões 21×21 . Esse tipo de filtro consegue atenuar as transições rápidas de cores na imagem, de forma que a suavização do conjunto de pontos se assemelha a um espalhamento das amostras.

Por fim, o último passo consistiu na construção de um mapa de cores mais fiel ao que é observado em imagens reais de eletroforese, coradas com Negro de Amido ou outro corante de tom azulado. Realizou-se a modelagem do subespaço desses tons reais das cores no espaço tridimensional do sistema de cores RGB, a partir de uma aproximação desse subespaço por segmentos de reta. Para isso, selecionaram-se 5 pontos de forma que fosse possível utilizá-los para criar intervalos de mapeamento de cores, capazes de imitar um comportamento similar, considerando a capacidade da visão humana de reconhecer a diferença de cores.

III. RESULTADOS

O simulador proposto é capaz de sintetizar imagens com resolução 200×200 contendo a representação de um número variado de amostras, onde cada uma delas possui proporções diversas das hemoglobinas. Em conjunto, é gerado um arquivo .txt contendo valores correspondentes às proporções de cada tipo de hemoglobina presente nas amostras. A Figura 1 mostra um exemplo de imagem sintetizada pelo simulador.

Posteriormente, foi realizada uma enquete anônima com 11 profissionais de áreas correlatas ao exame, sem que fosse revelado se tratar de 4 imagens sintéticas. Nela, os participantes

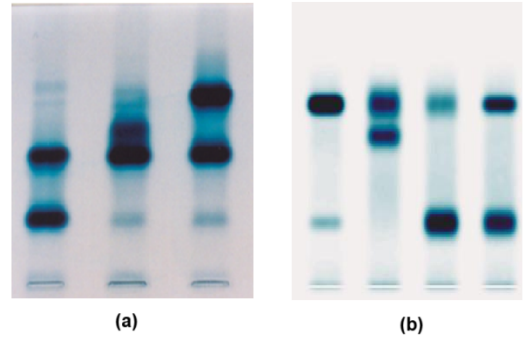


Fig. 1. Comparação entre uma imagem real (a) e uma imagem gerada pelo simulador (b).

deveriam escolher uma opção na escala de 1 a 5, onde 1 correspondia à baixa qualidade das imagens do exame e 5 à alta qualidade. O resultado é apresentado no gráfico da Figura 2.

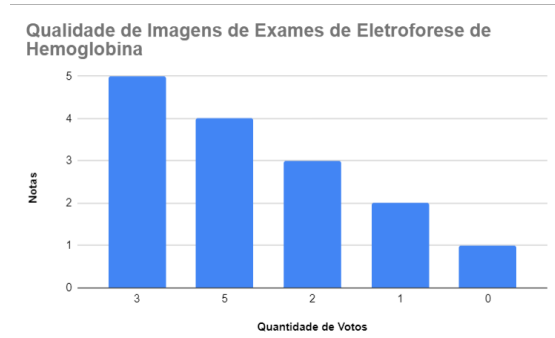


Fig. 2. Resultados da enquete anônima.

É possível observar que cerca de 91% dos indivíduos que responderam à enquete avaliaram que as imagens possuem qualidade de média a alta (notas de 3 a 5). Outro fator observado foi que nenhum dos consultados pareceu ter identificado a imagem como sendo sintética.

IV. CONCLUSÕES

Esse trabalho apresentou o desenvolvimento de um simulador em GNU Octave capaz de realizar a síntese de imagens realistas de exames de eletroforese de hemoglobina. Pretende-se, a partir disso, que a sua utilização possibilite a execução de demais trabalhos relacionados, por exemplo, a proposta de métodos para auxiliar na detecção de hemoglobinas variantes.

REFERÊNCIAS

- [1] Neila Caroline Henrique da Silva, Júlio César Gomes da Silva, Maria Gabriella Nunes de Melo, and Igor Felipe Andrade Costa de Souza, "Principais técnicas para o diagnóstico da anemia falciforme: uma revisão de literatura," *Caderno de Graduação - Ciências Biológicas e da Saúde - UNIT - PERNAMBUCO*, vol. 3, no. 2, pp. 33-33, 2017.
- [2] Muhammad Salman Khan et al. *Deep Learning Assisted Automated Assessment of Thalassemia from Haemoglobin Electrophoresis Images. Diagnostics*, v. 12, n. 10, p. 2405, 2022.
- [3] Vlahos, Loukas, et al. "Normal and anomalous diffusion: A tutorial." arXiv preprint arXiv:0805.0419 (2008).