

PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ

Processamento e Indexação de Imagens de Eletroforese 2D

Gerson Linck Bichinho

Trabalho apresentado como
requisito parcial para o concurso à
classe de professor titular do
Centro de Ciências Exatas e
Tecnologia da Pontifícia
Universidade Católica do Paraná.

Curitiba, novembro de 1999.

“Pretendendo-se saber qual a natureza das coisas, é necessário procurar conhecê-las; a questão de saber o que as coisas são é, pois, inseparável da questão de saber como se conhecem as coisas como sendo o que são, e particularmente da questão de saber qual a garantia de que as coisas são o que se diz que elas são ao se pretender conhecê-las.”

B. R.

Resumo

Neste trabalho apresentamos um estudo sobre processamento de imagens de eletroforese. A detecção de mudanças nos padrões gerados pelos picos protéicos é de grande importância para a detecção dos mais diversos tipos de anomalias. Apresentamos inicialmente algumas considerações sobre as técnicas de eletroforese unidimensional e alguns resultados que podem ser obtidos através da digitalização dos perfis das bandas.

Atualmente existem diversos bancos de dados de eletroforese bidimensionais, que armazenam e analisam um grande número de imagens. Este tipo de registro está se tornando cada vez mais difundido, fornecendo aos pesquisadores uma poderosa ferramenta na identificação e análise de proteínas. Por outro lado, a grande quantidade de imagens e as pequenas diferenças existente entre elas faz com que a utilização de técnicas de processamento de imagens sejam imprescindíveis para que tenhamos uma análise precisa e ao mesmo tempo rápida das imagens.

A identificação de um determinado padrão protéico é efetuada atualmente através de observação visual. A indexação de imagens proposta neste trabalho, contribuirá no sentido de facilitar a identificação de picos protéicos e conseqüente recuperação de informações, tornando mais eficiente o trabalho de análise do pesquisador.

Abstract

In this work we presented a study on electrophoresis image processing. The detection of changes in the patterns generated by the proteins spot is very important for the detection of several types of anomalies. We discussed some important considerations of the techniques of unidimensional electrophoresis, suggesting results that can be obtained after band digitalisation.

There are several databases of 2D electrophoresis that store and analyze a large number of images. These databases are becoming more and more spread, supplying researchers a powerful tool in the identification and analysis of proteins. So, the great amount of images and the small differences among them make the use of image processing techniques indispensable to achieve precise and fast image analysis.

The identification of a certain protein pattern is made using visual observation. The image indexation described in this work will facilitate the identification of protein spots and consequent information recovery, which permits a more efficient image analysis.

ÍNDICE

RESUMO	3
ABSTRACT	4
ÍNDICE	5
LISTA DE FIGURAS.....	6
CAPÍTULO I.....	7
Introdução.....	7
CAPÍTULO II.....	9
Processamento de Imagens em Eletroforese unidimensional	9
CAPÍTULO III.....	17
Eletroforese bidimensional.....	17
Indexação de imagens.....	19
Bases de Dados existentes	22
Metodologia	24
Aquisição de imagens.....	24
Segmentação.....	24
Sincronização das imagens.....	26
Proposta de Indexação	26
CAPÍTULO IV.....	30
Discussão e Conclusão	30
REFERÊNCIAS BIBLIOGRÁFICAS.....	32
ANEXO I	36
ANEXO II.....	39
ANEXO III.....	40
ANEXO IV	41
ANEXO V	43

Lista de Figuras

Figura 1. Diagrama esquemático, mostrando os principais componentes envolvidos em uma eletroforese uni-dimensional.....	10
Figura 2. Exemplo de imagem digitalizada de eletroforese uni-dimensional com 6 bandas. Neste exemplo, podemos observar uma pequena distorção ocorrida durante a migração das bandas.....	10
Figura 3. Outro exemplo de imagem de eletroforese.....	13
Figura 4. Perfis sobrepostos, referentes as bandas mostradas na Figura 3.....	13
Figura 5. Imagem de uma das bandas juntamente com o perfil da região selecionada.....	14
Figura 6. Curva de perfil originais (sem filtragem) sobrepostas as curvas de perfil após filtragem.....	15
Figura 7. Detalhes de duas das bandas mostradas na Figura 2. Os perfis mostram o ruído do sinal digitalizado.....	15
Figura 8. Exemplo de eletroforese bi-dimensional, mostrando o padrão protéico gerado.....	18
Figura 9. Histograma da imagem mostrada na Figura 8.....	19
Figura 10. Exemplo mostrando a junção de dois ou mais picos protéicos.....	25
Figura 11. Imagem processada, mostrando a localização dos picos protéicos ainda sem a redução completa das regiões a um único ponto.....	26
Figura 12. Geração dos vetores de comportamento do picos protéicos na imagem de referencia.....	27
Figura 13. Em uma outra imagem, delimitamos arbitrariamente a região de interesse e é gerado um outro vetor (sub-padrão).....	28

Capítulo I

Introdução

A percepção de imagens é algo extremamente complexo. As células nervosas da retina, dividem uma imagem em componentes isolados; estes pedaços fragmentados de informação são então remontados no cérebro em uma “imagem” mais simples, coerente com o que está sendo visto. A partir desta imagem, é criada a informação sobre o cenário. Desta maneira, os seres humanos tem um processo muito natural de organizar imagens. Esta organização ocorre baseada na semântica do conteúdo das imagens [RUDEK,98].

Uma das primeiras aplicações das técnicas de processamento de imagens foi utilizada com o objetivo de melhorar a qualidade das imagens de jornais digitalizados que eram transmitidos por cabos submarinos entre Nova York e Londres. Estas imagens foram inicialmente transmitidas em 5 níveis de cinza, evoluindo para 15 níveis no início da década de 20 [GONZALES et al., 92].

Com a evolução das técnicas computacionais, começaram a ser desenvolvidos algoritmos que tem como finalidade principal, melhorar a qualidade gráfica das imagens, facilitando a interpretação.

O processamento de imagens como um todo pode ser dividido em 5 etapas principais: aquisição, pré-processamento, segmentação, reconhecimento e interpretação.

A primeira etapa trata da aquisição, ou seja, um sensor transforma as informações imagem em um sinal elétrico que é posteriormente discretizado, dando origem a uma imagem digital.

Em seguida temos a fase pré-processamento que tem como objetivo retirar imperfeições introduzidas durante o processo de aquisição. Normalmente nesta fase tenta-se melhorar o contraste e eliminar ruídos.

O processo de segmentação trata da separação ou divisão da imagem em suas partes ou objetos constituintes. Dependendo das características e da complexidade

da imagem esta tarefa é geralmente bastante difícil. As técnicas de segmentação permitem a determinação de uma região ou das bordas do objeto sob interesse [GONZALES et al., 92], [JAHNE,97].

O reconhecimento é o processo que classifica um objeto, baseado em informações fornecidas por descritores.

E finalmente temos a interpretação, que envolve a atribuição de um significado a um conjunto de objetos reconhecidos.

A complexidade, a grande quantidade e a diversidade das imagens transforma o problema de processamento e indexação em um grande desafio.

Este estudo foi iniciado a partir da observação das técnicas utilizadas pelos pesquisadores para a análise e interpretação das imagens de eletroforese. Apesar de existirem ferramentas que auxiliem este procedimento, a identificação de alterações dos padrões protéicos ainda está baseado em uma análise visual.

Existem diversas bases de dados com imagens de eletroforese 2D. Estas bases de dados contém um número bastante grande de imagens, com diversos picos precisamente identificados através de técnicas bioquímicas, mas os mecanismos de recuperação destas informações ainda não é eficiente.

Sabemos que as alterações dos padrões estão associadas a situações de normalidade ou anormalidade (doenças), e que certos tipos de tratamento podem produzir alterações destes padrões. Assim sendo, a correta identificação e interpretação dos picos pode ser de vital importância para o entendimento de diversos mecanismos biológicos.

Uma proposta para este processo de indexação é apresentada neste trabalho, a qual possibilitará a localização de imagens com o mesmo padrão, permitindo a identificação baseada em imagens de referência.

Capítulo II

Processamento de Imagens em Eletroforese unidimensional

A eletroforese como utilizada em laboratório, envolve a separação de macromoléculas. Em todas as formas de eletroforese, um campo elétrico será estabelecido entre as bordas da solução tampão, e o potencial elétrico aplicado ocasionará a migração direcionada dos componentes em direção ao ânodo ou cátodo. A velocidade e direção na qual esta migração ocorre depende de diversos fatores, tais como: carga das partículas, força iônica, temperatura, potencial elétrico, tempo, difusão, tamanho e forma das moléculas, concentração do tampão, tipo do suporte. Entretanto, apesar de experimentalmente simples, os resultados obtidos dependem de um grande número de fatores que efetivamente limitam a precisão deste método, dificultando a interpretação dos resultados. Algumas considerações sobre os principais fatores que influenciam na migração das partículas podem ser encontradas no anexo I.

A técnica de eletroforese é largamente utilizada na determinação do peso molecular de proteínas e de pequenos fragmentos de DNA, e também no diagnóstico médico através de técnicas imunoquímicas. Estas aplicações geralmente estão relacionadas à determinação precisa da mobilidade relativa das moléculas.

Alguns dos experimentos envolvendo técnicas de eletroforese, podem demorar mais de 70 horas, e exigir bastante trabalho e material, sendo importante que tenhamos métodos para obtenção de imagens de alta qualidade com a substituição da imagem em papel por uma imagem digital, possibilitando desta forma a utilização de técnicas de processamento de imagens visando melhorar a qualidade das mesmas e facilitar a interpretação.

As bandas formadas após migração são usualmente lidas em um densitômetro, que nos fornece a leitura da densidade ótica correspondente. A quantidade relativa do migrante depende somente da largura da banda e da densidade ótica da mesma. Nas figuras 1 e 2 mostradas a seguir temos um

diagrama esquemático do processo de eletroforese e a imagem obtida respectivamente.

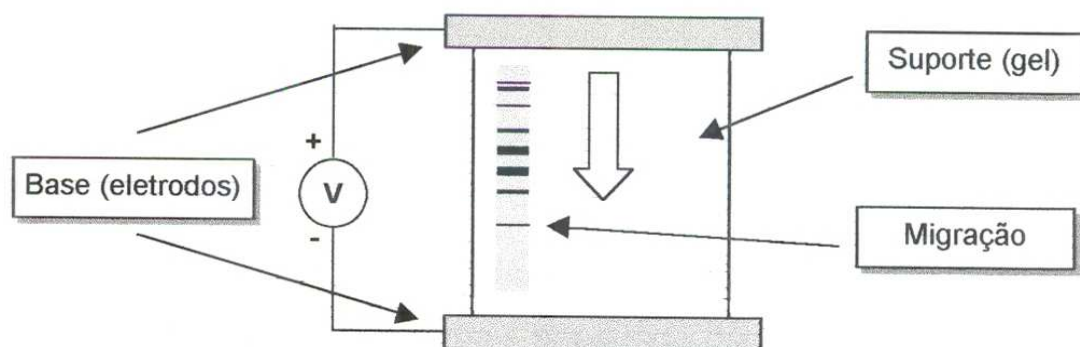


Figura 1. Diagrama esquemático, mostrando os principais componentes envolvidos em uma eletroforese uni-dimensional.

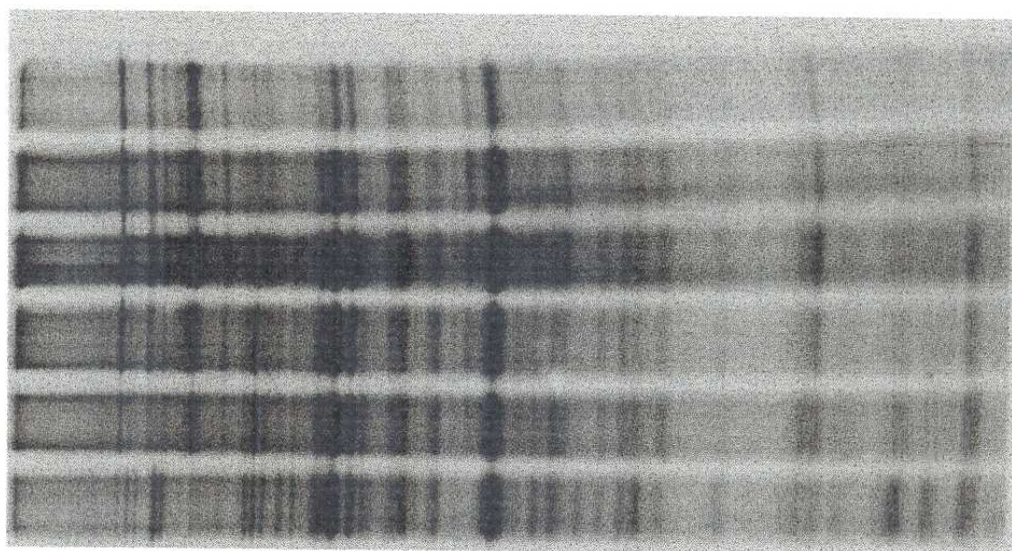


Figura 2. Exemplo de imagem digitalizada de eletroforese uni-dimensional com 6 bandas. Neste exemplo, podemos observar uma pequena distorção ocorrida durante a migração das bandas.

Exemplo clínico: hemoglobinopatias

A eletroforese da hemoglobina é um procedimento de laboratório simples mas muito útil na detecção e identificação de hemoglobinas anormais. Teoricamente podem existir 2200 hemoglobinas anormais como resultado de substituições de aminoácidos simples, mas apenas um terço destes teria carga anormal. Como a eletroforese tem sido usada para a determinação de hemoglobinas anormais, a maioria das 200 variantes que são clinicamente assintomáticas estão associadas com um padrão eletroforético anormal [LEE et al., 98]. A figura mostrada no anexo IV

apresenta a molécula de hemoglobina, e na figura seguinte deste mesmo anexo temos um gráfico esquemático da migração das mesmas, mostrando as anomalias dos aminoácidos e as mudanças no padrão migratório.

Para podermos utilizar as técnicas de processamento de imagens, inicialmente devemos obter uma imagem digital. As principais características dos equipamentos que efetuam esta aquisição são comentadas a seguir:

Diversos tipos de equipamentos podem ser utilizados para aquisição de imagens, a escolha do tipo mais adequado depende do tipo de aplicação, do custo envolvido e da maneira como é gerada a imagem. Como exemplo, podemos citar:

- scanners de mesa podem ser utilizados tanto no modo de reflexão como de transmissão de luz através do suporte, com um custo relativamente baixo. Atualmente os scanners de mesa possuem resolução espacial e em profundidade suficiente para que tenhamos uma boa qualidade da imagem gerada.
- densitômetro é um aparelho que permite a leitura de imagens por transmissão da luz através do gel. Os densitômetros são aparelhos que tem um custo mais elevado, apresentando praticamente a mesma leitura em termos de densidade ótica que as cameras CCD e que os scanners de mesa.
- cameras de video CCD (charged coupled device) tem tido sua utilização aumentada nos últimos anos, principalmente devido à diminuição dos custos deste tipo de equipamento, e também devido ao fato do mesmo gerar uma imagem digital de maneira direta, ou seja não necessitamos de etapas intermediárias no processo de aquisição, fato que poderia diminuir a qualidade final das mesmas. Cameras CCD são também utilizadas na captação de imagens fluorescentes, onde a iluminação artificial interfere na qualidade.

Em muitas situações o gel precisa ser corado e neste caso a imagem formada não é essencialmente em níveis de cinza. Este fato, apesar de levar a procedimentos diferentes em termos de processamento, não altera os resultados finais em nenhum dos métodos utilizados na obtenção das curvas, pois tanto o densitômetro permite um ajuste do comprimento de onda luminosa emitida, quanto o scanner gera imagens coloridas.

Observamos que em muitas situações, podem ocorrer distorções durante a migração. Algumas das razões que geram estas distorções podem ser encontradas no anexo I, e no anexo VI. Apesar do exemplo mostrado apresentar somente uma pequena distorção, podemos encontrar situações onde as distorções são bastante acentuadas, dificultando a análise comparativa entre as bandas: Muitas vezes é difícil de refazer todo o procedimento da eletroforese.

Em [GLASBEY et al., 94], é proposto um algoritmo que elimina os efeitos de alguns tipos de distorção. Outro algoritmo tratando deste problema mas específico a seqüências de DNA foi apresentado por [FUJII e KASHIWAGI, 1992].

Rotineiramente, a comparação e análise visual das bandas é efetuada tomando-se como base as curvas de perfil, por esta razão, acreditamos ser muito mais simples processarmos a curva obtida, e não a imagem, de maneira a obtermos o alinhamento das mesmas. Podemos expandi-las interpolando pontos adequadamente para que todas tenham o mesmo comprimento final e os picos apresentados possam ser comparados, figura 6 e figura 7.

Mostramos a seguir algumas imagens e traçados correspondentes de algumas imagens de eletroforese. As imagens mostradas foram digitalizadas em 300 dpi e o traçado das curvas a partir das imagens originais, sem pré-processamento, diminuindo desta forma possíveis distorções que poderiam ocorrer.

A eliminação da componente do fundo também foi efetuada somente nas curvas de perfil, principalmente devido a simplicidade do processamento. Foi detectado o ponto de mínimo e este valor foi subtraído dos demais.

O traçado das bandas é efetuado de maneira bastante simples, sendo que a seleção da região de interesse é efetuada de maneira manual efetuando-se a marcação dos pontos inicial e final e em seguida realizando-se a leitura dos valores dos pixels na imagem de eletroforese.

Os pontos que definem o perfil não foram utilizados individualmente, mas foram definidos através da média de três pontos vizinhos, perpendiculares à reta definida pelo usuário. Este procedimento visa melhorar a qualidade final do traçado diminuindo o ruído introduzido durante o processo de digitalização das imagens.

A marcação dos pontos para o cálculo da área sob a curva foi efetuada de maneira manual, e tentamos obter a melhor correspondência possível entre os pontos marcados em ambas as curvas, apesar de não termos uma perfeita equivalência, os resultados são bastante satisfatórios.

Medidas efetuadas experimentalmente mostram que mesmo sem ajustes mais precisos, somente eliminando o fundo (background), foram obtidas diferenças de cerca de 1% entre as medidas efetuadas pelo densitômetro e as obtidas calculando-se a área sob as curvas.

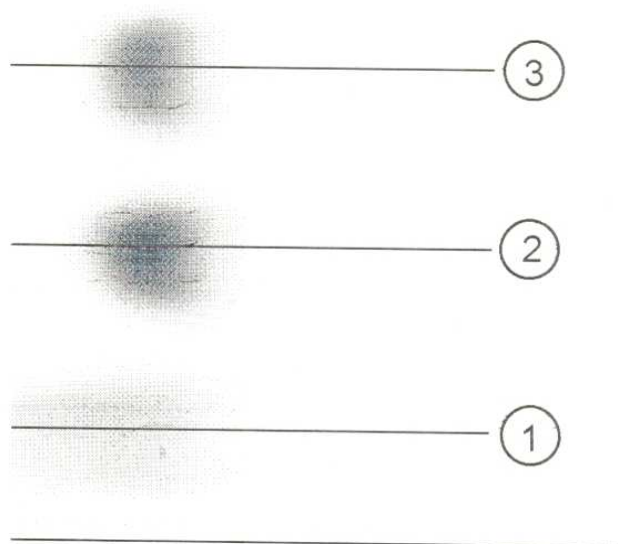


Figura 3. Outro exemplo de imagem de eletroforese.

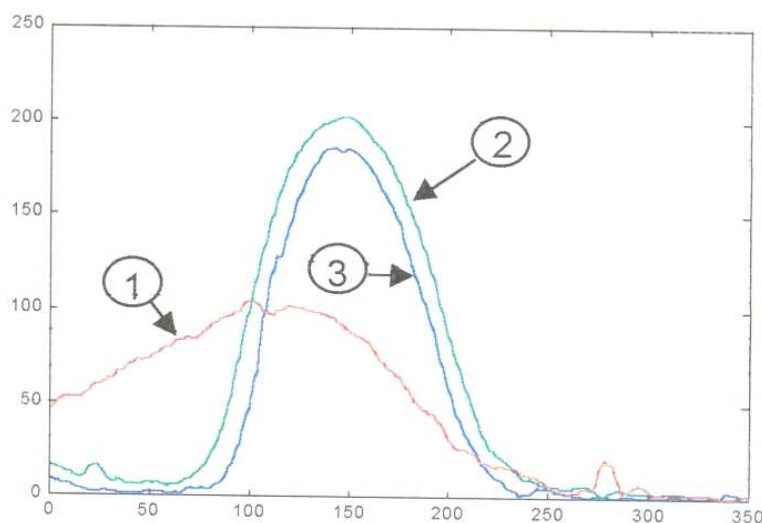


Figura 4. Perfis sobrepostos, referentes as bandas mostradas na Figura 3.

Através da utilização das técnicas de processamento de sinais podemos efetuar facilmente ampliações em partes da imagem, como mostrado na Figura 5

Outra possibilidade interessante é o fato de podermos isolar uma das bandas e efetuar a visualização da curva de perfil ampliada para a região marcada sobre a imagem e vice-versa, conforme mostrado na figura abaixo.

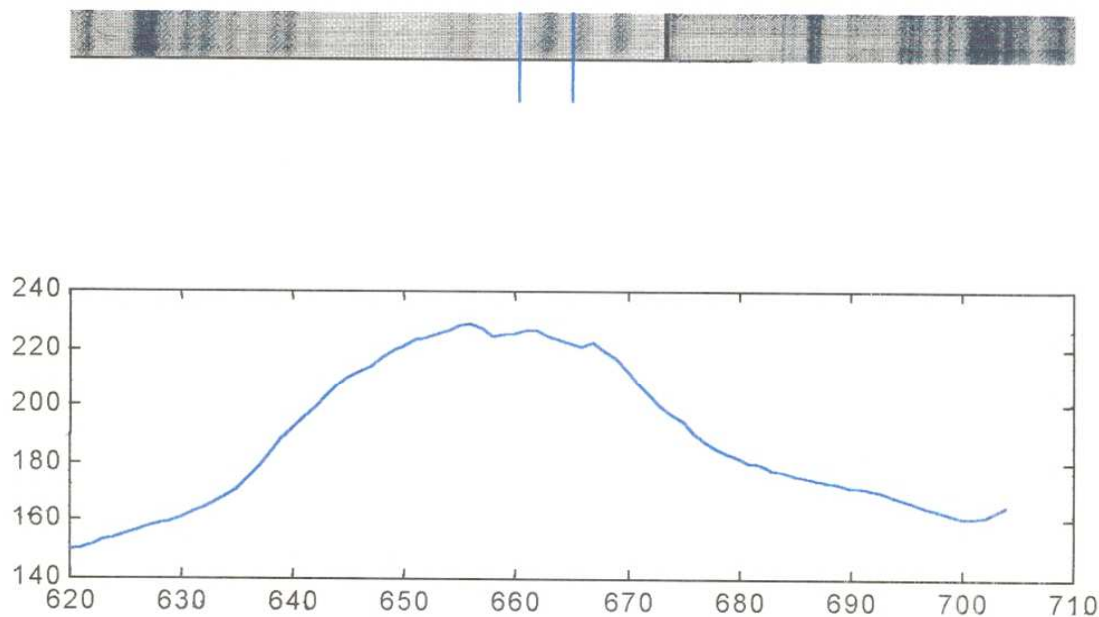


Figura 5. Imagem de uma das bandas juntamente com o perfil da região selecionada.

Na maioria das curvas obtidas através do densitômetros, os valores são normalizados entre 0 e 1.

Na Figura 6 mostrada à seguir, temos a sobreposição das curvas originais com a curva após ter sido submetida a um processo de filtragem. O filtro utilizado calcula a media entre 5 pontos substituindo o ponto central por este novo valor.

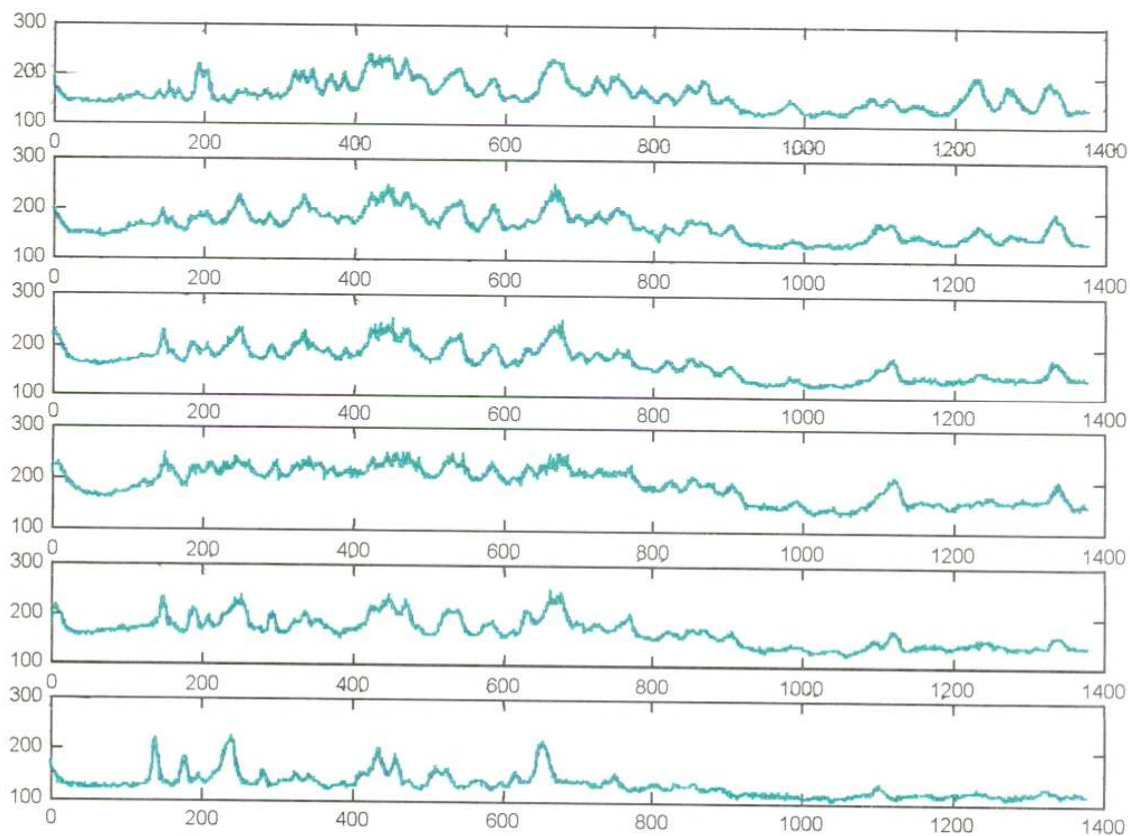


Figura 6. Curva de perfil originais (sem filtragem) sobrepostas as curvas de perfil após filtragem.

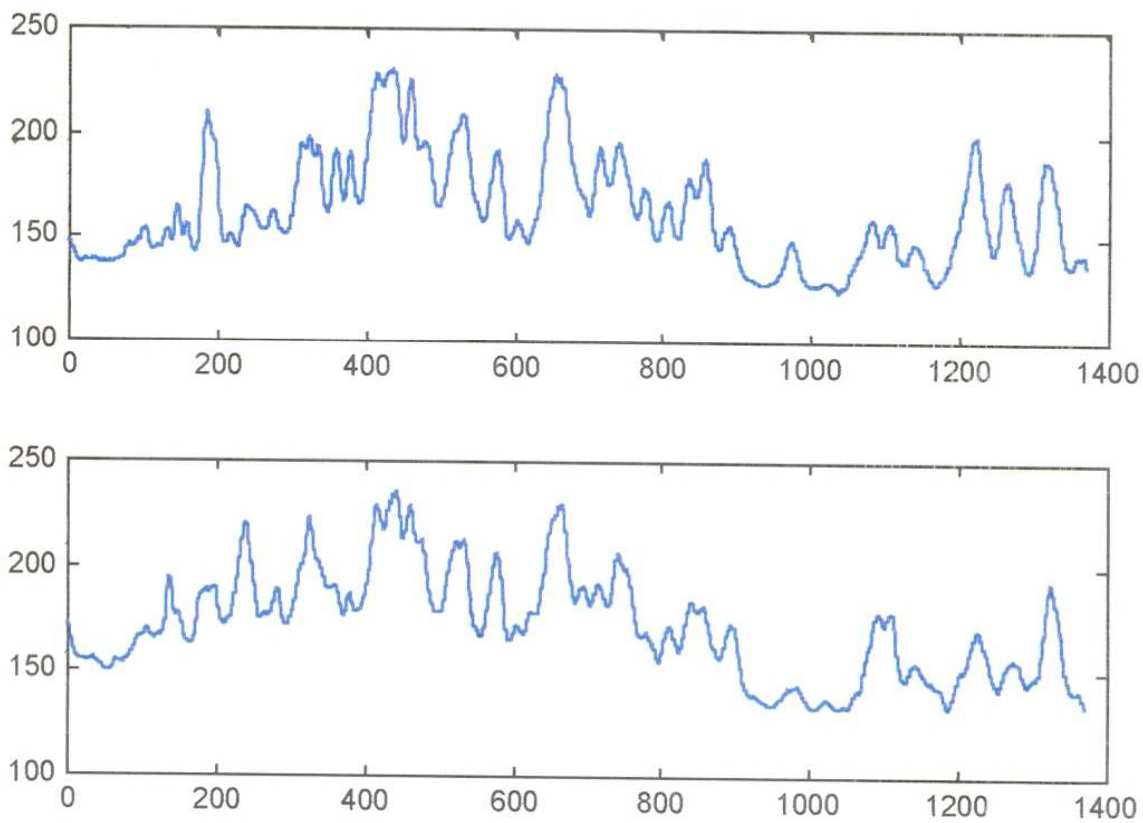


Figura 7. Detalhes de duas das bandas mostradas na Figura 2. Os perfis mostram o ruído do sinal digitalizado.

Nestas considerações iniciais, apresentamos algumas características dos sinais obtidos à partir de imagens de bandas de eletroforese. A grande diferença observada entre as imagens, em termos de tamanho, número de bandas, cor predominante, além da presença de ruído são aspectos importantes, influenciando na qualidade final da imagem e consequentemente no perfil das bandas.

As técnicas de processamento de imagens são importantes ferramentas de auxílio na interpretação e análise dos resultados.

Além disso, em casos onde houveram problemas na migração das proteínas, provocando deformações, em imagens com pouco ou excessivo contraste, com muito ruído, as técnicas de processamento podem minimizar os problemas, possibilitando a análise. Este fato é particularmente importante, em situações onde todo o procedimento eletroforético teria de ser refeito. Existem situações onde é muito difícil refazermos todos os procedimentos envolvidos, seja devido ao tempo extremamente longo necessário, ou ainda, devido a não disponibilidade de tecidos, muitas vezes submetidos a tratamentos que ativam mecanismos bioquímicos com tempo de vida bastante curtos.

Capítulo III

Eletroforese bidimensional

A idéia de criar bases de dados para eletroforese 2D surgiu com as primeiras descrições de separação protéica efetuadas em bactérias de *E. coli* e tecidos de ratos, em 1975. Em 1981, surgiram as primeiras propostas de indexação de proteínas humanas.

Estudos mostraram que eletroforeses efetuadas em pequenos gels não possuíam resolução suficiente para se detectar a totalidade de proteínas encontradas nos tecidos. As estratégias para melhorar a resolução são aumentar o tamanho do gel ou efetuar o pré-fracionamento através de cromatografia. Ambos os métodos mostraram que o número de proteínas pode ser bastante grande, por exemplo, cerca de 10.000 em tecidos de ratos, [BURGGRAF et al., 95].

Um outro pré-requisito importante na construção das bases de dados é a identificação das proteínas à partir das imagens 2D. A utilização de técnicas bioquímicas e mais recentemente de espectrometria de massa permite que os sequenciamentos sejam corretamente efetuados. A eletroforese bi-dimensional, conforme exemplo mostrado na

Figura 8, permite a detecção de milhares de polipeptídeos, separando-os pelo peso molecular e ponto isoelétrico. Esta técnica fornece uma visão da expressão genética celular e é extremamente útil na detecção de proteínas com significancia biológica importante.

Os diversos padrões de concentração das proteínas nos permite documentar e associação existente entre estes padrões e situações de normalidade e anormalidade (doenças). Atualmente vários estudos visando a identificação e quantificação dos padrões foram desenvolvidos. Além disso, bases de dados foram implementadas de maneira a documentar e associar os padrões à situações de normalidade ou anormalidade.

A grande quantidade de pontos a serem analisados torna imprescindível a utilização de ferramentas de auxílio. Por exemplo, no trabalho apresentado por [ALAIYA et al., 99], a utilização de ferramentas que auxiliem na análise permitiu que conclusões importantes quanto a heterogeneidade de tumores ovarianos. Diversos outros estudos encontrados na literatura, também utilizam como ferramenta de análise, o processamento, segmentação, identificação de proteínas, podemos citar como exemplo: [GOLDFARB,99], [KAWCZAK et al., 98], [GIDDINGS et al., 98], [JENSEN et al., 96], [SKOVGAARD et al., 95].

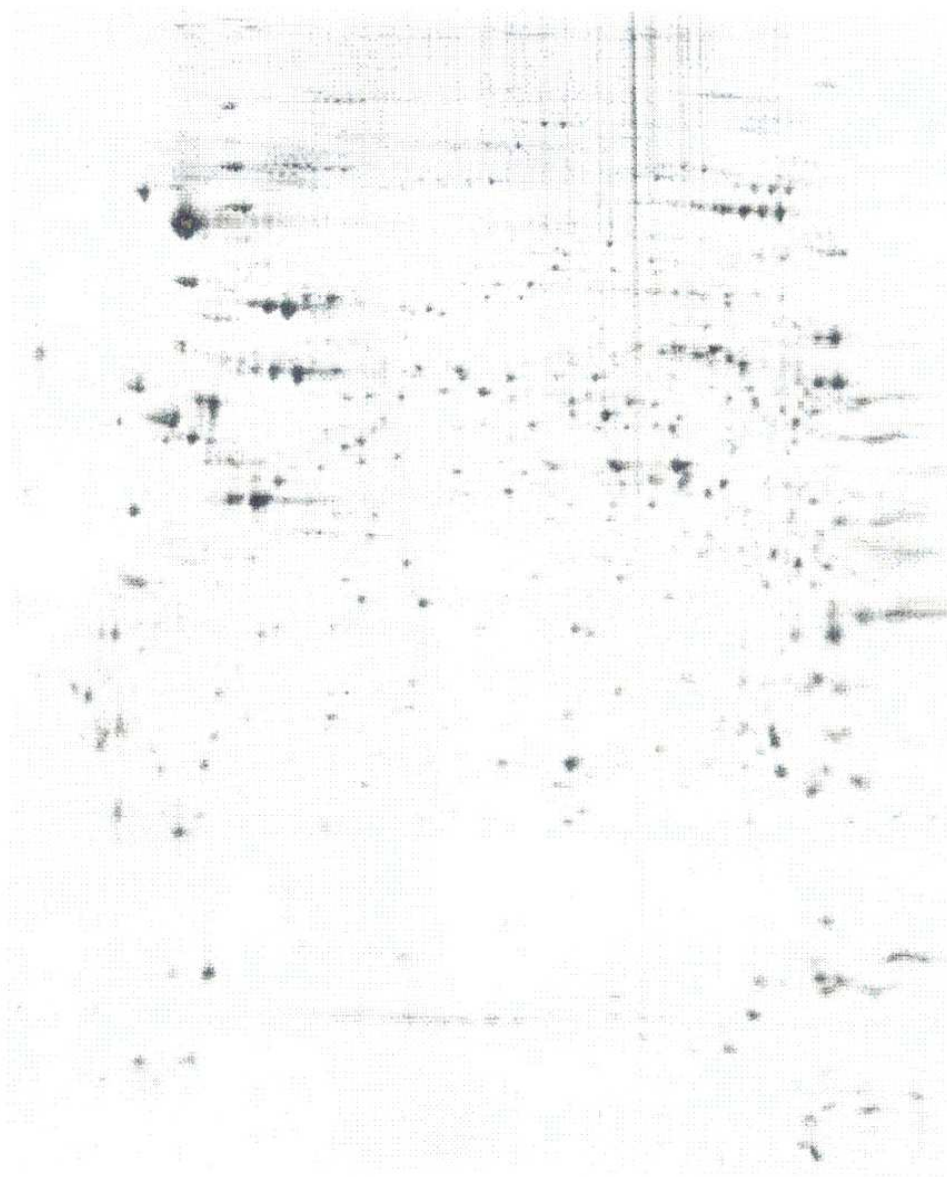


Figura 8. Exemplo de eletroforese bi-dimensional, mostrando o padrão protéico gerado.

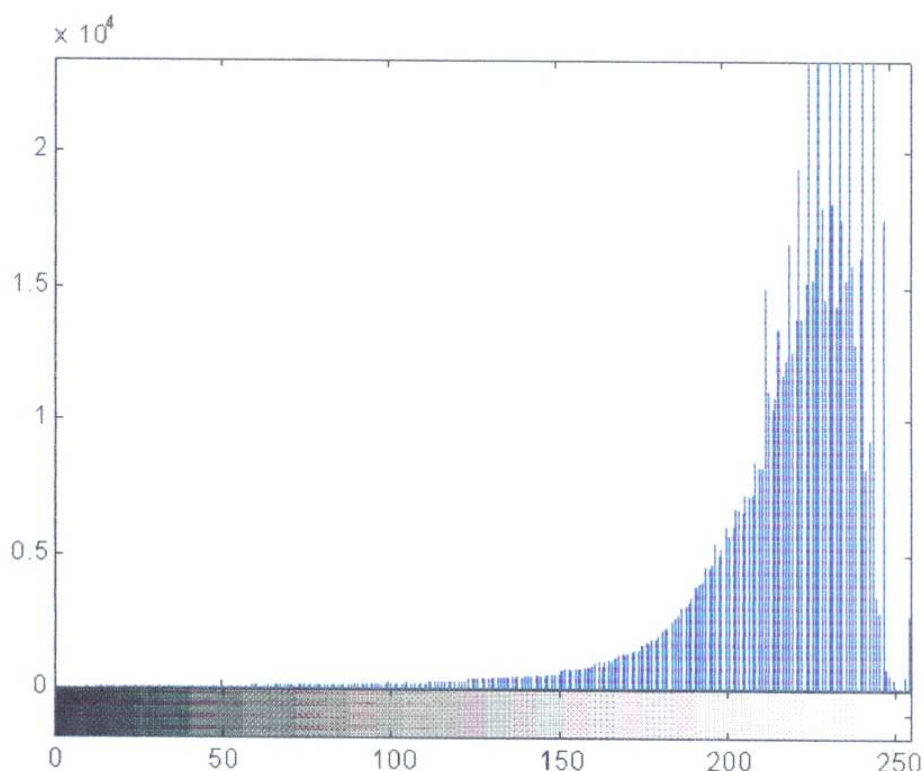


Figura 9. Histograma da imagem mostrada na figura 8.

Indexação de imagens

A grande quantidade de informações do tipo imagem atualmente disponíveis vem fazer com que cada vez mais, pesquisas de novos métodos de busca e organização das informações sejam necessárias para a solução destes novos desafios. Uma ferramenta apropriada de indexação nos permite otimizar o processo de busca. Nos bancos de dados convencionais, o acesso às informações ocorre por meio de consultas a partir de palavras chave, frases, etc. Em sistemas que armazenam outros tipos de informação, tais como, imagens, sons, gráficos, vídeos, as soluções são menos evidentes. A recuperação de informações armazenadas em imagens requer métodos mais sofisticados de recuperação, normalmente tomando-se como base atributos visuais [DARZIGER,90].

O método mais comum utilizado para encontrar imagens contendo uma determinada informação é através da inserção de campos associados contendo descritores textuais. A busca será efetuada através destes descritores. Muitas

imagens são decididamente de natureza diversa, ou seja, podem conter uma grande variedade de tipos de informação, de cores, de formas, e de texturas. A definição destes descritores é feita através da observação direta da imagem e pode ser influenciada pelo que o observador considera importante ao observar a imagem [CAWKELL, 93].

Em sistemas que armazenam imagens, procuramos extrair aquelas que possuam características relevantes. De acordo com [RAVELA et al., 95], a relevância da imagem está associada ao seu contexto, em termos de atributos (cores, texturas, objetos), ou seja, está associada com a forma como o usuário interpreta esta imagem. Decodificar a informação semântica envolve problemas como segmentação, reconhecimento e extração automáticas de atributos, este ainda é um problema bastante complexo.

Normalmente para contornarmos o problema da definição automática dos atributos, adotamos informações adicionais fornecidas pelos usuários ou especialistas. Imagens são associadas a textos ou palavras, como mencionado anteriormente, e estes textos são utilizados para recuperar as imagens. Esta solução muitas vezes não é satisfatória, tendo em vista a quantidade de detalhes que podem ocorrer em algumas imagens. A interpretação dada pelo usuário ao que é ou não relevante pode influenciar significativamente na escolha das palavras que definem a imagem.

De acordo com o afirmado por [LEUNG et al., 95], é necessário criarmos um modelo de dados para a representação sistemática do conteúdo de uma imagem. A partir deste modelo, uma descrição da imagem pode ser construída automaticamente e podemos capturar informações semânticas que serão utilizadas na construção de um índice.

[YIHONG et al., 96], propõe que as imagens sejam indexadas automaticamente usando um método de indexação por histogramas. O histograma de cores carrega informação sobre a distribuição de cores, mas não sobre a localização das mesmas. Existem muitas regiões do histograma que não representam nenhum pixel. Por esta razão, [YIHONG et al., 94] considera apenas os 20 locais de topo como pixels representativos do histograma.

Muitas técnicas de indexação de imagens requerem uma grande quantidade de processamento, espaço adicional para armazenamento e não podem ser aplicados à imagens compactadas. De maneira a minimizar os aspectos mencionados anteriormente, [SMITH et al., 96] propõe que seja efetuada a seleção e a transformação do espaço de cores da imagem, em um novo espaço, onde as cores são quantizadas. [IDRIS et al., 95] apresenta técnicas de indexação baseadas em vetores de quantização.

Diversos outros trabalhos foram desenvolvidos, com abordagens que tratam da indexação baseada em fractais e wavelets, [MENON et al., 96], [GRAPS,95], [MANDAL et al., 96] e [WANG et al., 96].

A identificação das proteínas, e a formação de bases de dados permitindo consultas para a rápida verificação de padrões é de extrema importância, tendo em vista a imensa quantidade e diversidade de tipos tecidos que estão sendo atualmente analisados.

As questões respondidas pelas bandas de eletroforese quanto a possíveis anomalias, está relacionada com as diferenças existentes nas composições de duas ou mais amostras e, na quantidade e propriedades dos diferentes constituintes presentes em um conjunto de amostras [HORGAN et al., 95]. A resposta a estas questões de normalidade ou anormalidade está ligada as posições, tamanho e densidade dos picos e das bandas na imagem. Esta informação é essencialmente visual, os cientistas utilizam sua experiência na análise dos bandas. Muitas vezes este tipo de análise produz resultados satisfatórios, mas em alguns casos necessitamos de resultados mais objetivos, além disso, pode ser cansativo efetuar a análise visual de muitos gels.

Diversos trabalhos foram desenvolvidos com o objetivo de melhorar a qualidade e facilitar a interpretação dos mesmos. [PANEK et al., 97] utilizando classificação adaptativa com utilização de redes neurais, verificou que as redes neurais tem uma performance melhor que análise por agrupamento ou, análise dos componentes principais.

[MASSEROLI et al., 93], descreve um ambiente para análise de imagens de bandas de eletroforese que permite a avaliação e quantificação em termos absolutos

ou relativos e os resultados deste trabalho mostram a compatibilidade entre os valores obtidos através de densitometria.

[REDMAN et al., 91], apresentou um programa (GelReader 1.0) desenvolvido para melhorar a precisão e possibilitar a análise de bandas eletroforéticas unidimensionais, possibilitando a localização de bandas e gerando um relatório com pesos moleculares desconhecidos baseados em um conjunto de padrões. O software forneceu ainda a possibilidade de manipulação da imagem, equalização do histograma, ampliação e comparação entre diferentes bandas.

Diversos trabalhos foram também desenvolvidos com o objetivo de obter a melhor imagem possível no momento da aquisição. Por exemplo, [REINER et al., 91] desenvolveram um programa para aquisição e correção da linearidade da saída da camera; [MALEK et al., 97] descreve um novo método que utiliza uma camera de vídeo CCD na aquisição e quantificação de imagens eletroforéticas.

Um software que permite a aquisição através de camera de vídeo, possibilitando ainda a correção do contraste e de imperfeições geradas por uma iluminação não uniforme está descrito em [D'ARGY et al., 90]. O programa permite o traçado de histograma de densidades em uma área de interesse previamente definida. [YEE, 91] apresenta um trabalho cujo principal objetivo é o de melhorar o contraste e eliminação do fundo (background) da imagem. Permite também efetuar filtragens com o objetivo de melhorar a qualidade de imagens não muito nítidas pela presença de ruído.

Bases de Dados existentes

Existem atualmente diversos sistemas oferecendo serviços que facilitam a identificação, quantificação de picos protéicos em imagens de eletroforese. Alguns destes sistemas possuem somente a função de banco de dados de imagens, permitindo apenas a marcação manual e visualização dos picos protéicos, com apenas algumas ferramentas de busca de informação sobre os picos já identificados. Cada uma destas bases de dados é especializada em determinados tipos de material, tais como: plasma, linfoma, bactérias, músculo, etc. A relação das

principais bases pode ser encontrada em [<http://www-lecb.ncifcrf.gov/EP/table2databases.html>].

O sistema chamado de “Flicker”, proposto pelo NCI (National Cancer Institute) possibilita a comparação de imagens de diferentes fontes através da internet. Este sistema permite a comparação visual entre duas imagens permitindo a identificação de picos referentes a determinada proteína. Foi desenvolvida uma *applet* em JAVA que mostra alternadamente às duas imagens previamente alinhadas. O alinhamento entre as imagens foi efetuado tomando-se como base aspectos morfológicos comuns, ou seja, foram marcados pontos que representam estruturas comuns as duas imagens e que servem de referencia para a transformação geométrica e conseqüente alinhamento espacial das mesmas. Este programa permite também que sejam efetuadas melhorias de contraste e brilho, e que uma região previamente selecionada seja ampliada.

O programa Melanie (The Proteome Browser), permite a extração, análise, identificação, busca e verificação de dados em imagens de eletroforese 2D e em bases de dados associadas [APPEL et al., 97].

O sistemas Gellab-II : 2D Eletrophoretic Gel Analysis System, desenvolvido pelo Laboratory of Experimental and Computational Biology, National Cancer Institute, implementa, além das funções de bases de dados, ferramentas de identificação, extração e quantificação de picos protéicos.

Diversas outras empresas também oferecem serviços de identificação e quantificação de proteínas a partir de imagens de eletroforese 2D. Podemos citar, entre outras: LSB Company, BioImage, Amersham Pharmacia Biotech, Phoretix Ltd, etc.

[BETTENS et al., 97], apresenta um trabalho de segmentação e modelização dos picos de eletroforese.

A seguir descreveremos as etapas necessárias para a obtenção de imagens segmentadas e que apresentem somente os pontos correspondentes aos picos. Esta fase é importante e servirá de base para o método de indexação proposto.

Metodologia

Descreveremos a seguir as principais etapas envolvidas no processamento de imagens de eletroforese 2D.

Aquisição de imagens

No processo de aquisição de imagens, devemos levar em consideração a resolução espacial (horizontal e vertical), definida em número de pixels por unidade de área. Esta densidade espacial deve possuir precisão adequada ao tipo de análise efetuada, ou seja, é evidente que cada banda ou pico eletroforético deve conter um número suficiente de pontos que permita uma boa visualização e processamento.

A resolução em profundidade, fator que define o número de valores de intensidade de um ponto ou *pixel*, é usualmente de 256 (8 bits) ou 65536 valores por pixel (16 bits). A densidade ótica observada na maioria dos gels é menor do que 3 unidades de densidade ótica.

Um outro fator importante é a linearidade do equipamento, é necessário que tenhamos a possibilidade de calibrar o equipamento pois a relação entre a densidade ótica de um objeto e a quantidade de luz que passa através dele não é linear. Desta forma, para que tenhamos uma maior fidelidade, é necessário definir um protocolo de calibração. A maioria dos scanners de mesa é pré-calibrado para a faixa de densidades óticas para a qual foi projetado.

Segmentação

Em termos de processamento de imagens muitos trabalhos foram elaborados com o objetivo de segmentar as imagens. A imagem de eletroforese apresenta um histograma mono-modal, conforme mostrado anteriormente na figura 9, razão pela qual as técnicas clássicas, baseadas em um limiar não fornecem bons resultados.

Uma outra característica observada em imagens de eletroforese é apresentada na Figura 10, onde temos a junção de diversos picos, técnicas de segmentação que consigam separá-los de maneira correta não são simples de serem implementadas.

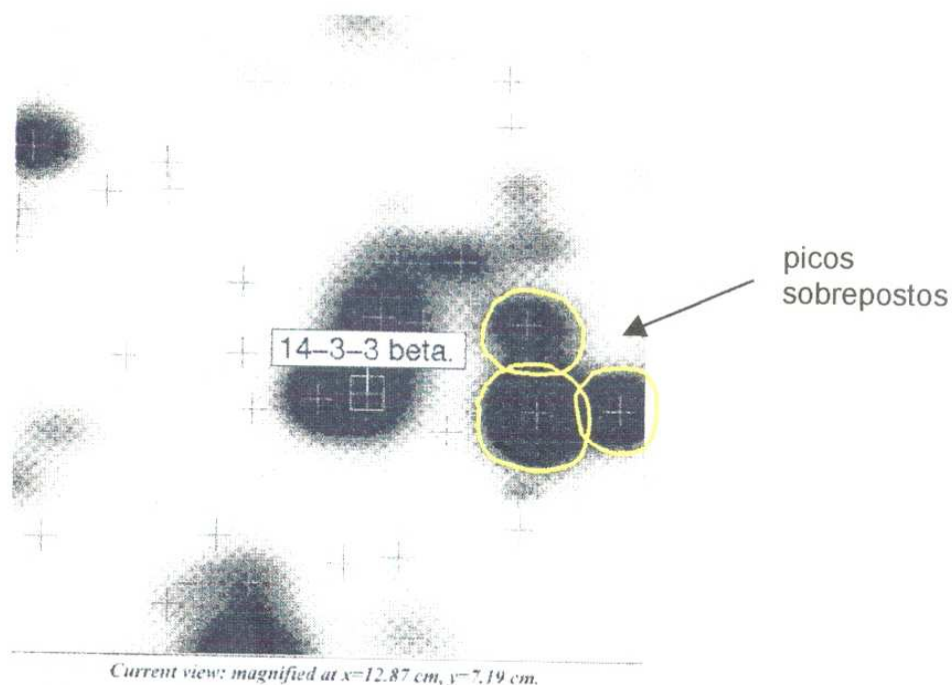


Figura 10. Exemplo mostrando a junção de dois ou mais picos protéicos.

Muitos outros trabalhos, não específicos ao tipo de aplicação que estamos abordando, foram desenvolvidos e apresentam soluções para problemas semelhantes a este. A técnica de *watershed*, dentre as que pesquisamos na literatura, é a que apresenta melhores resultados dadas as características da imagem. Podemos citar, entre outros os trabalhos de [BETTENS et al., 97], [VINCENT et al., 91], [ROERDINK et al., 99], [ROERDINK et al., 97], [YECHEUNG et al., 93] e [BRAGA NETO et al., 95] que descrevem diversas aplicações da técnica de *watershed*. Apesar do processo de segmentação não ser o objetivo principal deste trabalho, sua eficiência e precisão está diretamente relacionada ao sucesso das fases seguintes, relativa a indexação.

Uma outra importante etapa após a segmentação, é a definição dos pontos centrais dos picos pois os mesmos serão utilizados posteriormente na definição dos vetores que expressarão os padrões utilizados na indexação. A redução da região que define os picos a um pixel, pode ser efetuada através da determinação dos centros de gravidade das regiões. No anexo V apresentamos o cálculo dos centros de gravidade em imagens binárias.

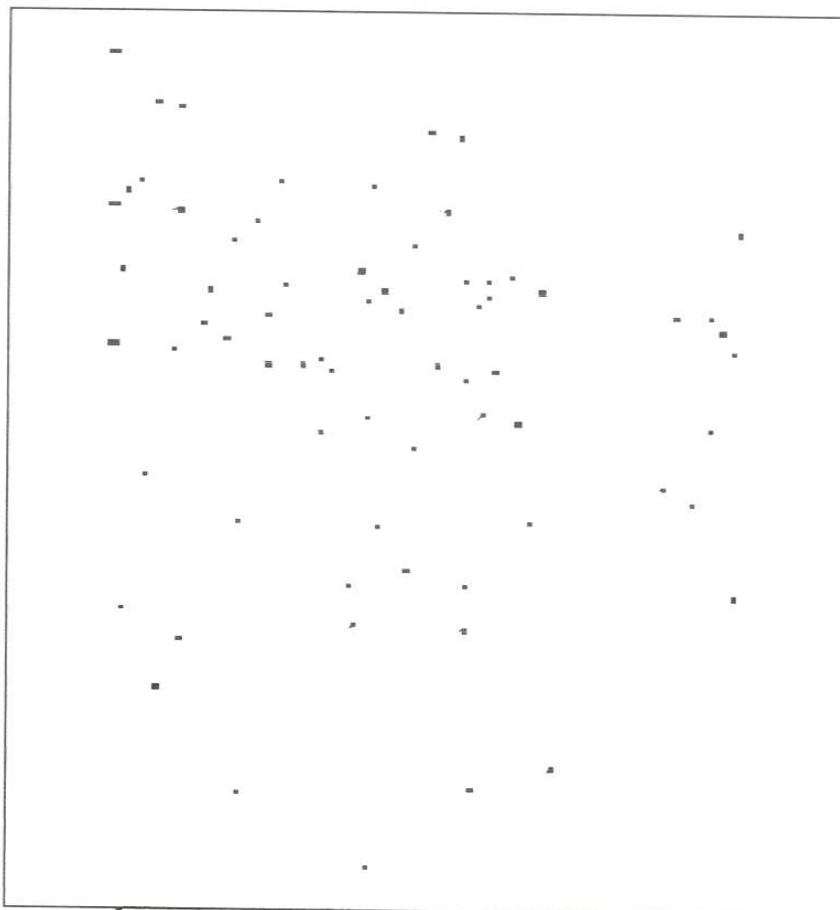


Figura 11. Imagem processada, mostrando a localização dos picos protéicos ainda sem a redução completa das regiões a um único ponto.

Sincronização das imagens

Após a aquisição, as imagens não serão necessariamente iguais. As diferenças de tamanho existente entre elas fazem com que tenhamos que efetuar o redimensionamento da imagem. Na maioria dos sistemas que efetuam o processamento de imagens de eletroforese, um tamanho de referência é definido e são efetuadas marcações nas imagens. Estes pontos de referência serão utilizadas no processo de redimensionamento geométrico. A utilização desta técnica de redimensionamento é importante quando precisamos efetuar a comparação visual entre as imagens.

Proposta de Indexação

Se observarmos as técnicas de indexação de imagens citadas anteriormente, as mesmas utilizam como principal elemento discriminador as informações de cor,

textura e forma. Em se tratando de imagens de eletroforese, estas informações são praticamente iguais em todas as imagens.

Efetuiremos então o mapeamento das posições de cada uma dos picos da imagem de referência, a partir de pontos definidos e identificados, obtidos após segmentação e cálculo dos centros de gravidade.

Cada posição será definida, juntamente com informações relativas as concentrações, forma do pico, diâmetro da base e altura. Apesar de existirem estudos que tratam da modelização dos picos os resultados obtidos ainda não são satisfatórios, pois os mesmos não modelam com precisão todas as características dos picos protéicos encontrados nas imagens de eletroforeses [BETTENS et al., 97].

A seguir, efetuiremos a projeção horizontal e vertical de cada uma dos picos (histograma), conforme mostrado na figura 12. O vetor gerado por estas projeções representa de maneira simplificada o padrão protéico da imagem de referência. Se os picos estiverem em uma mesma linha de projeção, eles podem ficar escondidos. Neste caso, deve ficar registrado nos vetores a existência das coincidências. Isto será feito aumentando-se a amplitude do vetor das projeções em questão. Por exemplo, amplitude 1 significa que não houveram coincidências, amplitude 2 houveram duas coincidências, e assim sucessivamente. O mesmo procedimento será efetuado para a projeção vertical. Estes vetores se tornarão os descritores da imagem de referência.

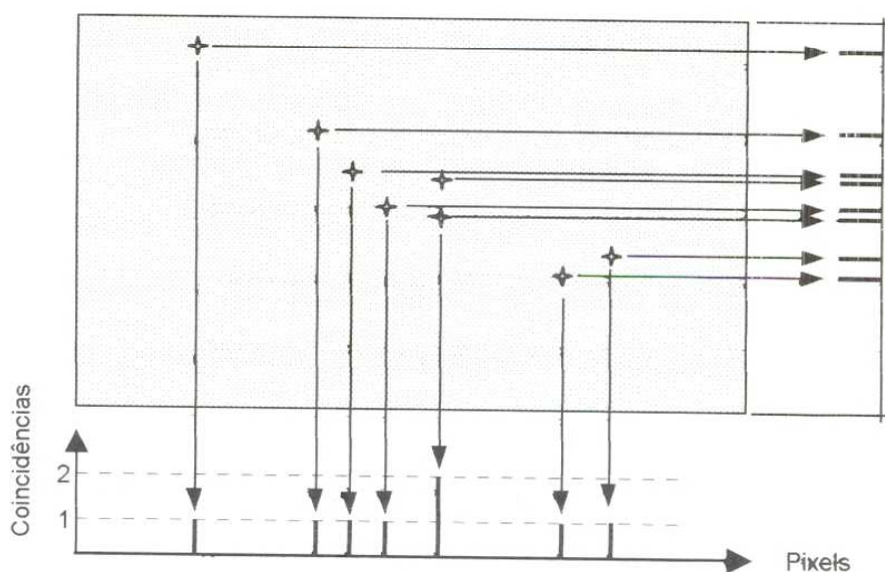


Figura 12. Geração dos vetores de comportamento dos picos protéicos na imagem de referência.

No momento da identificação deste padrão em outra imagem, será efetuada a marcação manual do pico de interesse, a partir daí delimitaremos arbitrariamente a região em torno do ponto, conforme mostrado na Figura 13. A largura desta região deverá ser objeto de estudo futuro, este tamanho está relacionado a quantidade de pontos que determinam os padrões e influenciará de maneira direta no fator discriminante dos mesmos.

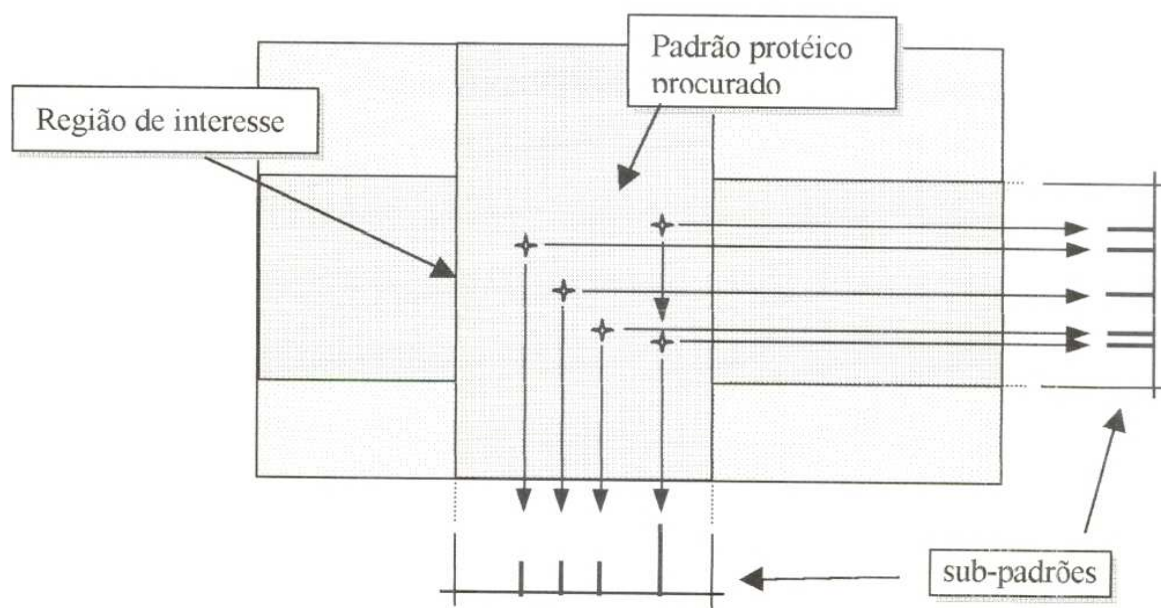


Figura 13. Em uma outra imagem, delimitamos arbitrariamente a região de interesse e é gerado um outro vetor (sub-padrão).

Uma vez definido estes vetores, de acordo com técnica descrita, devemos efetuar a procura deste sub-padrão no padrão gerado pela imagem de referência. Para identificar os padrões gerados, pretendemos utilizar algoritmos de busca descritos na literatura (boyer-moore, knuth-morris-pratt, etc) [WIRTH, 1989]. A otimização dos algoritmos de busca tendo em vista as características dos vetores gerados deverá ser pesquisada, principalmente nos casos de imagens ampliadas e/ou deformadas.

Diversos tipos de deformações podem ocorrer devemos avaliar as diferentes situações para evitar detecções incorretas ou falsas detecções.

Difícilmente, em situações reais teremos um sincronismo perfeito entre as imagens, pois as mesmas são geradas através de procedimentos e equipamentos

diferentes. Apesar disso os padrões gerados são bem definidos, podendo ocorrer pequenas diferenças na distribuição dos picos ao longo do vetor. Estudo dos algoritmos de busca devem ser efetuados para que os mesmos possam ser adaptados as deformações descritas. A adaptação dos algoritmos às variações ocorridas nos padrões, através de alongamentos ou encurtamento devem ser estudadas e avaliadas. Esta adaptação poderá ser efetuada com utilização da marcação manual de pontos de referência ou com métodos automáticos.

Devemos prever todos os tipos de situação e definir as tolerâncias adequadas em caso de não existir uma coincidência perfeita entre os padrões.

Capítulo IV

Discussão e Conclusão

Eletroforese é um método onde moléculas em solução, tais como proteínas e ácidos nucleicos, migram em resposta a um campo elétrico. A velocidade de migração ou mobilidade das partículas, através do campo elétrico, depende da intensidade do campo, da carga elétrica das partículas, do tamanho e forma das moléculas, da viscosidade e temperatura do meio nas quais as moléculas estão se movendo. Como uma ferramenta analítica a eletroforese é simples, rápida e altamente sensível. Ela é principalmente usada no estudo das propriedades de partículas carregadas e também como uma ferramenta de separação.

Procuramos neste trabalho apresentar as principais características da técnica de eletroforese. Acreditamos que pelo fato de estarmos trabalhando em áreas que são multi-disciplinares este conhecimento é fundamental como base para discussão dos problemas existentes. Foram discutidos os principais aspectos relativos a aquisição, processamento e indexação de imagens de eletroforese.

Acreditamos que a determinação de descritores, sob forma de vetores, facilitem e agilizem a procura de características armazenadas em grandes bancos de imagens de eletroforese.

O aperfeiçoamento e adequação dos algoritmos de segmentação deve ser efetuada, pois a qualidade da segmentação influencia diretamente na definição dos pontos representando os picos eletroforéticos.

Os algoritmos de busca, adaptados às características do vetor de comportamento gerado devem ser implementados e avaliados.

Finalmente, acreditamos que as considerações e proposições apresentadas neste trabalho possam servir como base e incentivo à continuação do desenvolvimento de ferramentas que auxiliem na busca de explicações para diversas situações em que estão envolvidas os padrões protéicos de diferentes tecidos, evidenciados através da técnica de eletroforese.

Referências Bibliográficas

- ALAIYA, A. A., FRANZEN, B., MOBERGER, B. et al. Two-dimensional gel analysis of protein expression in ovarian tumors shows a low degree of intratumoral heterogeneity. *Electrophoresis* vol.26, n.6, p.1039-1046, 1999.
- APPEL, R. D., PALAGI, P. M., WALTHER, D. et al. Melanie II - a third-generation software package for analysis of two-dimensional electrophoresis images: I. Features and user interface. *Electrophoresis* vol.18, n.15, p.2724-2734, 1997.
- BETTENS, E., SCHEUNDERS, P., VAN DYCK, D. et al. Computer analysis of two-dimensional electrophoresis gels: a new segmentation and modeling algorithm. *Electrophoresis* vol.18, n.5, p.792-798, 1997.
- BRAGA NETO, ULISSES DE MENDONÇA e SIQUEIRA NETO, WINSTON A. DE. A technique for the automatic detection and classification of mammographic calcifications using the watershed method. *Anais do VIII SIBGRAPI* , p.273-280, 1995.
- BURGGRAF, D. e LOTTSPEICH, F. The further construction of the two-dimensional database of common human proteins. *Electrophoresis* vol.16, n.5, p.831-840, 1995.
- CAWKELL A.E. Picture-queries and picture databases. *Journal of Information Science* vol.19, n.6, p.409-423, 1993.
- D'ARGY, R., SPERBER, G. O., LARSSON, B. S. et al. Computer-assisted quantification and image processing of whole-body autoradiograms. *Journal Pharmacological Methods* vol.24, n.3, p.165-181, 1990.
- DARZIGER, P. N. Picture databases a protical approach to picture retrieval. *Database* vol.13, n.4, p.13-17, 1990.
- GIDDINGS, M. C., SEVERIN, J., WESTPHALL, M. et al. A software system for data analysis in automated DNA sequencing. *Genome Research* vol.8, n.6, p.644-665, 1998.
- GLASBEY, CHRIS A. e WRIGHT, FRANK. An algorithm for unwarping multitrack electrophoretic gels. *Electrophoresis* vol.15, p.143-148, 1994.

- GOLDFARB, M. Two-dimensional electrophoresis and computer imaging: quantitative of human milk casein. *Electrophoresis* vol.20, n.4-5, p.870-874, 1999.
- GONZALES, RAFAEL C. e WOODS, RICHARD E. Digital Image Processing, 1992.
- GRAPS, AMARA. An Introduction to Wavelets. *IEEE Signal & Image Processing* , p.50-61, 1995.
- HORGAN, GRAHAM W. e GLASBEY, CHRIS A. Uses of digital image analysis in electrophoresis. *Electrophoresis* vol.16, p.298-305, 1995.
- IDRIS, F. e PANCHANATHAN, S. Image Indexing Using Vector Quantization. *Proceedings of the SPIE* vol.2420, p.373-380, 1995.
- JAHNE, BERND. Digital Image Processing: Concepts, Algorithms and Scientific Applications. vol.4, n.1, 1997.
- JENSEN, K., KESMIR, C., e SONDERGAARD, I. From image processing to classification: IV. Classification of electrophoretic patterns by neural networks and statistical methods anable quality assessment of wheat varieties for breadmaking. *Electrophoresis* vol.17, n.4, p.694-698, 1996.
- KAWCZAK, J. A., MATHISEN, P. M., DRAZBA, J. A. et al. Digitized image analysis reveals diffuse abnormalities in normal-appearing white matter during acute experimental autoimmune encephalomyelitis. *Journal Neuroscience Research* vol.54, n.3, p.364-372, 1998.
- LEE, G. RICHARD, BITHELL, THOMAS C., FOERSTER, JOHN et al. Wintrobe Hematologia Clínica. vol.1ª ed brasileira, n.1, p.1120-1152, 1998.
- LEUNG, C. H. C. e ZHENG, Z. J. Data Modeling for Efficient Content Indexing. *Proceedings of the International Workshop on Multi-Media Database Management Systems* , p.143-150, 1995.
- MALEK, A. M., IZUMO, S., e ALPER, S. L. Quantitative densitometric analysis using a commercially available handheld CCD digital camera. *Biotechniques* vol.22, n.6, p.1150-1153, 1997.
- MANDAL, M. K., ABOULNASR, T., e PANCHANATHAN, S. Image Indexing Using Moments and Wavelets. *IEEE Transactions on Consumer Electronics* vol.42, n.3, p.557-565, 1996.

- MASSEROLI, M., MESSORI, A., BENDOTTI, C. et al. Automatic quantitative evaluation of autoradiographic band films by computerized image analysis. *Life Sci* vol.53, n.20, p.331-336, 1993.
- MENON, RAGHU P., ACHARYA, RAJ S., e ZHANG, AIDONG. Content Basead Image Query from Image Database Using Spatio-Temporal Transforms and Fractal Analysis Methods. *IEEE International Conference in Image Processing* vol.3, p.863, 1996.
- NIKLAUS, WIRTH, Algoritmos e estruturas de Dados, Prentice/Hall do Brasil, Rio de Janeiro, 1989.
- PANEK, J. e VOHRADSKY, J. Object-oriented developmental environment for image-analysis applications: implementation for 2D gel electrophoretogram analysis. *Comput Appl Biosci* vol.13, n.3, p.257-262, 1997.
- RAVELA, S. e MANMATHA, R. Retrieval from Image databases Using Scale-Space Matching. *CmpSci Technical Report* , p.95-104, 1995.
- REDMAN, T. e JACOBS, T. Electrophoretic gel image analysis software for the molecular biology laboratory. *Biotechniques* vol.10, n.6, p.790-794, 1991.
- REINER, P. e BRENOWITZ, M. Quantitative densitometry of autoradiograms: digital images representative of digital density. *Comput Appl Biosci* vol.7, n.3, p.337-340, 1991.
- ROERDINK, JOS B. T. M. e MEIJSTER, ARNOLD. Segmentation by watersheds: definition and parallel implementation. *Advances in Computer Vision* , p.21-30, 1997.
- ROERDINK, JOS B. T. M. e MEIJSTER, ARNOLD. The Watershed Transform: Definitions, Algorithms and Parallelization Strategies. Institute for Mathematics and Computing Science. vol.IVI 00-9-06, p.1-34, 1999.
- RUDEK, MARCELO. Uma proposta para indexação automática de imagens e reconhecimento de cheques bancários baseadas no vetor de comportamento. *Tese de dissertação de mestrado, CEFET/PR* , 1998.
- SKOVGAARD, I. M., JENSEN, K., e SONDERGAARD, I. from image processing to classification: III. Matching patterns by shifting and stretching. *Electrophoresis* vol.16, n.8, p.1385-1389, 1995.
- SMITH, JOHN R. e CHANG, SHI-FU. Automated Image Retrieval Using Color and Texture. *Columbia University Technical Report* , 1996.

- VINCENT, L. e SOILLE, P. Watershed in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Anal. Mach. Intel.* vol.13, n.6, p.583-598, 1991.
- WANG, X., CHAN, E., MANDAL, M. K. et al. Wavelet-Based Image Using Nonlinear Interpolative Vector Quantization. *IEEE Transactions on Image Processing* vol.5, n.n, p.518-522, 1996.
- YECHENG, W., LEMKIN, P. F., e UPTON, K. A fast spot segmentation algorithm for two-dimensional gel electrophoresis analysis. *Electrophoresis* vol.14, p.1351-1356, 1993.
- YEE, T. Image Processing Software for enhanced visualization of Faint and Noisy Autoradiographic Imagens. *Biotechniques* vol.10, n.6, p.786-789, 1991.
- YIHONG, G., CHUAN, C. H., e XIAOYI, G. Image Index and Retrieval Based on Color Histograms. *Multimedia Tools and Applications* vol.2, n.2, p.133-156, 1996.
- YIHONG, G., HONG, JIANG Z., CHUAN, C. H. et al. Image Database System with Content Capturing and Fast Image Indexing Capabilities. *IEEE International Conference in Image Processing* , p.121-130, 1994.

ANEXO I

Considerações gerais sobre eletroforese

Algumas considerações importantes devem ser observadas, e apesar de serem bastante técnicos, estão relacionados a obtenção de uma imagem de qualidade, por este motivo efetuaremos, a seguir, alguns comentários sobre os principais fatores que influenciam na qualidade final.

Resistência, tensão, corrente e potência

Na eletroforese, um dos parâmetros elétricos é sempre mantido constante. Se a resistência aumenta durante a corrida, os resultados podem diferir, dependendo de qual das grandezas elétricas foi mantida constante:

- Corrente constante (velocidade é diretamente proporcional à corrente), a velocidade é mantida, mas calor é gerado;
- tensão constante: a velocidade diminui, mas não existe geração adicional de calor durante a corrida;
- Potência constante, a velocidade diminui mas a quantidade de calor gerada é constante.

Independentemente do modo que foi selecionado, o principal fator a ser levado em consideração é que devemos minimizar a difusão e a perda da atividade da amostra causada pelo aquecimento e pelo tempo total da corrida.

Relação entre a Carga Elétrica das Proteínas e o pH do meio

As proteínas são componentes anfóteros, isto é, contêm tanto componentes básicos quanto ácidos. Sua carga elétrica é determinada pelo pH do meio em que estão.

A carga da proteína que está relacionada ao pH depende da ionização dos grupos amino ou carboxílicos. Cada proteína tem sua carga elétrica característica e

que dependem do número e do tipo de amino-ácidos contendo grupos amino ou carboxílicos.

Os ácidos nucleicos, diferentemente das proteínas não são anfóteras. Eles permanecem negativos em qualquer pH a que sejam submetidos.

Existe um pH em que não existe carga nas proteínas. Este pH é o ponto isoelétrico ou PI. Cada proteína possui seu PI. Desta maneira, se colocarmos as proteínas em uma solução com pH acima do ponto isoelétrico, a proteína terá uma carga elétrica resultante negativa e migrará em direção ao ânodo. Abaixo do ponto isoelétrico, a proteína é positiva e migrará em direção ao cátodo.

Temperatura

Em todos os estágios do processo eletroforético a regulação da temperatura é importante. A polimerização do gel produz uma reação exotérmica. Se esta reação não for regulada, pode causar irregularidades no tamanho do poro do gel. De maneira a mantermos o tamanho da porosidade do gel uniforme e reprodutível, a reação de ser mantida sob temperatura constante, por esta razão os gels devem ser resfriados por líquido com temperatura constante.

Durante a eletroforese, manter a temperatura constante é essencial para evitar a desnaturação das proteínas. Um outro problema pode ocorrer com alguns tipos de gel, quando o problema de desnaturação não é relevante, e que provoca uma maior mobilidade das amostras no centro do gel e uma menor mobilidade nas laterais, isto provoca distorção na imagem tornando a comparação entre as bandas mais difícil.

Outro fator de problemas relativo a temperatura, esta ligado a alguns tipos de soluções tampão que são desenvolvidas para trabalhar entre 0°C e 25°C, e é preciso encontrar meios de manter a solução dentro desta faixa de temperatura, através de sistemas de resfriamento.

Suporte

Geralmente, a amostra é colocada em um suporte que pode ser papel, acetato de celulose, agarose, gel poliacrilamida. Os suportes mais utilizados em eletroforese

são os gels de agarose e poliacrilamida. Estes dois suportes diferem bastante em suas estruturas físicas e químicas, a porosidade dos gels pode retardar, ou em alguns casos obstruir completamente o movimento das macromoléculas ou fazendo com que somente moléculas pequenas migrem livremente. Preparando um gel com tamanho de poro adequado, o operador pode obter alguma vantagem relativa aos diferentes tamanhos de moléculas. Por exemplo, a hemoglobina possui um peso de 66000 Daltons e a B-lactoglobina de 38000 Daltons. É perfeitamente possível a preparação de um gel que seja mais restritivo a hemoglobina do que a B-lactoglobina.

Também é importante que o suporte escolhido seja eletricamente neutro, esta condição é essencial para evitarmos interações entre o suporte e as amostras que poderiam diminuir a resolução.

Eletroforese é usualmente efetuada em um gel em tubos ou placas. Quando efetuada em tubos, o gel é colocado em um tubo de vidro com cerca de 12 cm de comprimento e 3 à 5 mm de diâmetro interno. A placa de gel é formada por várias camadas, com duas placas planas de vidro separadas por um espaçador. Ambos são montados verticalmente.

ANEXO II

Bases de dados de eletroforese 2D

Swiss-Flash newsletter

<http://www.expasy.ch/www/swiss-flash.html>

CSPI/Scanalytics

<http://www.cspi.com/scanalytics/index.htm>

Kendrick Laboratories

<http://msn.fullfeed.com/kendrick>

Large Scale Biology Corp

<http://www.lsbc.com>

Phoretix Ltd

<http://www.phoretix.com>

Proteome Inc Home Page

<http://www.proteome.com>

ESA Electrophoresis

<http://www.esainc.com/esaapplications/esaelectrophoresisapp.html>

Biolmage

<http://www.bioimage.com>

Amersham Pharmacia Biotech

<http://www.hpb.com/2D/>

ANEXO III

Electrophoretic Societies

The Electrophoresis Society

<http://www-lecb.ncifcrf.gov/ESAB>

Swiss Electrophoresis Society

<http://dcwww.epfl.ch/icp/chel/chel.html>

Société Française d'Electrophorèse

<http://europe.u-strasbg.fr/sfe/sfe.html>

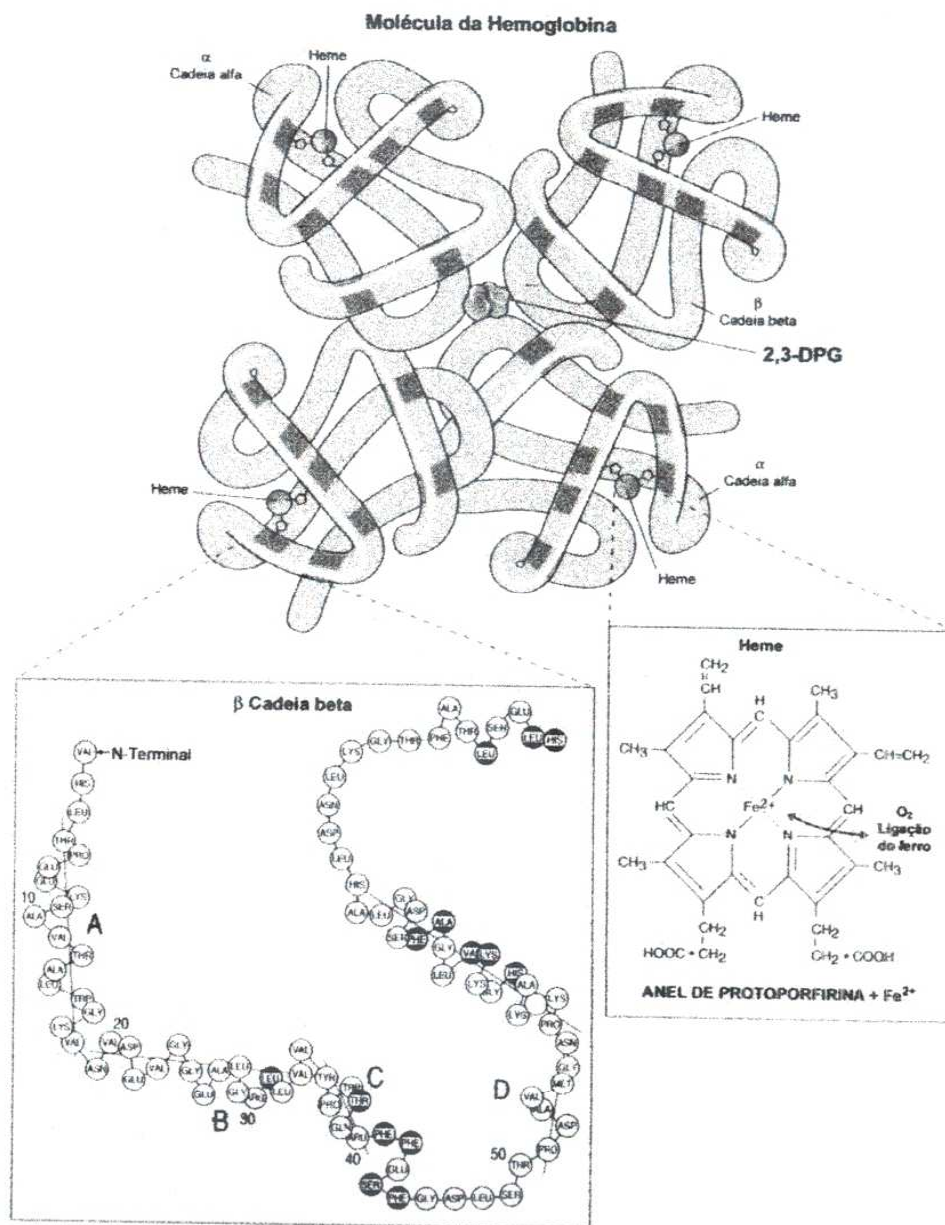
The British Electrophoresis Society

<http://sunspot.bioc.cam.uk/BES.html>

The Japanese Electrophoresis Society of ICES

<http://www.tmig.or.jp/JES>

ANEXO IV



– A molécula da hemoglobina (Adaptado de STEVENS, M.L., 1997, Saunders).

Mobilidades Eletroforéticas Comparativas de Hemoglobinas Mutantes e Cadeias de Globina

Hb & Substituição	Hemoglobina Acetato de celulose T.E.B., pH 8,5				Hemoglobina Ágar citrato pH 6,0				Cadeia de globina mutante Tampões de uréia, mercaptoetanol											
	+	Hb A	Hb S	Hb C	-	+	Hb C	Hb S	Hb A	Hb F	T.E.B., pH 8,9				-	T.E.B., Citrato, pH 6,0				
											+	γ^F β^A	β^S	α^A		+	γ^F β^A	β^S	α^A	-
I																				
$\alpha^{16}\text{Lys} \rightarrow \text{Glu}$																				
N - Baltimore																				
$\beta^{95}\text{Lys} \rightarrow \text{Glu}$																				
J - Baltimore																				
$\beta^{16}\text{Gly} \rightarrow \text{Asp}$																				
Camden																				
$\beta^{131}\text{Gln} \rightarrow \text{Glu}$																				
Hope																				
$\beta^{136}\text{Gly} \rightarrow \text{Asp}$																				
Mobile																				
$\beta^{73}\text{Asp} \rightarrow \text{Val}$																				
Korie Bu																				
$\beta^{73}\text{Asp} \rightarrow \text{Asn}$																				
Alabama																				
$\beta^{39}\text{Gln} \rightarrow \text{Lys}$																				
D - Ibadan																				
$\beta^{87}\text{Thr} \rightarrow \text{Lys}$																				
Montgomery																				
$\alpha^{48}\text{Leu} \rightarrow \text{Arg}$																				
Titusville																				
$\alpha^{94}\text{Asp} \rightarrow \text{Asn}$																				
G - Georgia																				
$\alpha^{95}\text{Pro} \rightarrow \text{Leu}$																				
G - Galveston																				
$\beta^{43}\text{Glu} \rightarrow \text{Ala}$																				
D - Los Angeles																				
$\beta^{121}\text{Glu} \rightarrow \text{Gln}$																				
S																				
$\beta^6\text{Glu} \rightarrow \text{Val}$																				
G - Philadelphia																				
$\alpha^{68}\text{Asn} \rightarrow \text{Lys}$																				
Shimonoseki																				
$\alpha^{54}\text{Gln} \rightarrow \text{Arg}$																				
Gun Hill																				
$\beta^{91-95}\text{deleted}$																				
O - Arab																				
$\beta^{121}\text{Glu} \rightarrow \text{Lys}$																				
E																				
$\beta^{26}\text{Glu} \rightarrow \text{Lys}$																				
C																				
$\beta^6\text{Glu} \rightarrow \text{Lys}$																				
A ₂																				
$\alpha^{16}\text{Gly} \rightarrow \text{Arg}$																				

Mobilidade eletroforética de 22 variantes de hemoglobina pelos quatro métodos descritos no texto. A origem de cada padrão está à margem direita dos painéis e a migração das hemoglobinas está à esquerda. A mobilidade do marcador Hbs A, S e C estão mostradas como linhas verticais. (De Schneider RG *et al.*: Abnormal hemoglobins in a quarter million people. Blood 48:629, 1976, com permissão de Grune & Stratton.)

ANEXO V

Momentos

Para uma imagem digital, o momento de ordem pq é definido por:

$$M_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (1)$$

onde $f(x,y)$ representa o nível de cinza do *pixel*.

Para uma imagem binária, podem ser definidos o valor da área e do centro de gravidade de um objeto a partir do cálculo dos momentos M_{00} , M_{01} e M_{10} . Os *pixels* da equação são aqueles relativos à região ocupada pelo objeto. Nesse caso, os valores de todos os *pixels* são 1 (preto), em contraste com os *pixels* de valor 0 (fundo branco).

Área

Área a uma medida de tamanho e corresponde ao número de *pixels* que constituem um objeto. É equivalente ao cálculo do momento M_{00} .

$$Área = M_{00} = \sum_x \sum_y F(x, y) \quad (2)$$

Centro de Gravidade

O centro de gravidade pode ser obtido através dos momentos M_{00} , M_{01} e M_{10}

$$G(x_g, y_g) \Leftrightarrow \begin{cases} x_g = \frac{M_{10}}{M_{00}} \\ y_g = \frac{M_{01}}{M_{00}} \end{cases} \quad (3)$$

onde:

$$M_{10} = \sum_x \sum_y x f(x, y) \quad M_{01} = \sum_x \sum_y y f(x, y)$$

Troubleshooting

6.0 Troubleshooting 2-D results

Table 26 lists problems that may be encountered in 2-D electrophoresis results, describes the possible causes, and suggests ways to prevent each problem in future experiments.

For troubleshooting problems encountered during the various steps of the 2-D process, refer to the following tables:

- Table 11, page 19. Troubleshooting IPG strip rehydration in Reswelling Tray

- Table 14, page 23. Troubleshooting first-dimension IEF: Multiphor II and Immobiline DryStrip Kit
- Table 17, page 27. Troubleshooting first-dimension IEF: IPGphor
- Table 23, page 33. Troubleshooting vertical second-dimension SDS-PAGE
- Table 25, page 35. Troubleshooting second-dimension SDS-PAGE: Multiphor II flatbed system

TABLE 26. TROUBLESHOOTING 2-D RESULTS

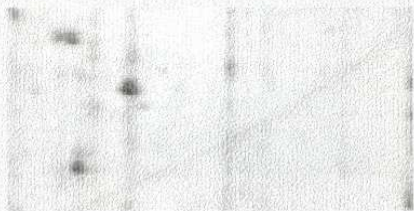
Symptom	Possible cause	Remedy
No distinct spots are visible 	Sample is insufficient.	Increase the amount of sample applied.
	Insufficient sample entered the IPG strip due to poor sample solubilization.	Increase the concentration of the solubilizing components in the sample solution. (See section 2.5, 'Composition of sample solution.')
	Sample contains impurities that prevent focusing.	Increase the focusing time or modify the sample preparation method. (See 'Part I. Sample Preparation.')
	The pH gradient is wrongly oriented.	The pointed end of the Immobiline DryStrip is the acidic end and should point toward the anode (+).
	(Flatbed gel format) IPG strip is placed wrong side down on second-dimension gel.	Ensure that the IPG strip is placed gel-side down (plastic backing upward) on the SDS second-dimension gel.
	Detection method was not sensitive enough.	Use another detection method (e.g., silver staining instead of Coomassie blue staining).
Individual proteins appear as multiple spots or are missing, unclear, or in the wrong position	Failure of detection reagents.	Check expiration dates on staining solutions. Prepare fresh staining solutions.
	Protein carbamylation.	Do not heat any solutions containing urea above 30 °C, as isocyanate, a urea degradation product, will carbamylate proteins, changing their pI.
	Protein oxidation.	DTT in the rehydration and equilibration solutions keeps the disulphide bonds reduced. For additional protection include an iodoacetamide treatment during equilibration prior to the second-dimension separation. Iodoacetamide alkylates the thiol groups to prevent the reduced proteins from reoxidizing.
Spots are vertically doubled, or "twinning" 	(Vertical gel format) IPG strip is not placed properly.	Ensure that the plastic backing of the IPG strip is against the glass plate on the second-dimension gel.

TABLE 26. TROUBLESHOOTING 2-D RESULTS (continued)

Symptom	Possible cause	Remedy
Distortion of 2-D pattern		
	(Vertical gel format) The top surface of the second-dimension gel is not flat.	Immediately after pouring the gel, overlay the surface with water-saturated butanol.
	(Vertical gel format) Uneven polymerization of gel due to incomplete polymerization, too rapid polymerization, or leakage during gel casting.	Degas the gel solution. Polymerization can be accelerated by increasing by 50% the amount of ammonium persulphate and TEMED used. Polymerization can be slowed by decreasing by 33% the amount of ammonium persulphate and TEMED used. Ensure that there is no leakage during gel casting.
	(Flatbed gel format) Moisture on the surface of the second-dimension gel.	Allow ExcelGel to dry for about 5 minutes after removing plastic cover and before applying buffer strips and IPG strip.
	(Flatbed gel format) IPG strip not removed during electrophoresis.	Remove the IPG strip and application pieces from the second-dimension gel when the bromophenol blue dye front has moved away from the IPG strip by 4–6 mm.
	(Flatbed gel format) Air bubbles under the second-dimension gel cause uneven migration due to poor heat transfer.	Ensure that no bubbles are trapped under the second-dimension gel during placement on the cooling plate.
	(Flatbed gel format) Water drops or pieces of buffer strip on the surface of the second dimension gel.	Take care that nothing is dropped or splashed onto the surface of the second-dimension gel.
Horizontal streaking or incompletely focused spots		
	Sample not completely solubilized prior to application.	Be sure that the sample is completely and stably solubilized. Note: Repeated precipitation-resolubilization cycles produce or increase horizontal streaking. See section 2.5, 'Composition of the sample solution,' for general guidelines for sample solubilization.
	Sample is poorly soluble in rehydration solution.	Increase the concentration of the solubilizing components in the rehydration solution. (See section 3.4, 'IPG strip rehydration solution.') Increase concentration of IPG Buffer.

TABLE 26. TROUBLESHOOTING 2-D RESULTS (continued)

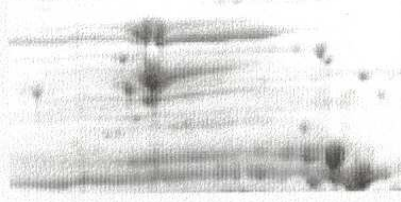
Symptom	Possible cause	Remedy
Horizontal streaking or incompletely focused spots (continued)		
	Interfering substances. Non-protein impurities in the sample can interfere with IEF, causing horizontal streaking in the final 2-D result, particularly toward the acidic side of the gel.	Modify sample preparation to limit these contaminants. (See section 2.4, 'Removal of contaminants that affect 2-D results.')
	Ionic impurities in sample.	Reduce salt concentration to below 10 mM by dilution or desalt the sample by dialysis. Precipitation with TCA and acetone and subsequent resuspension is another effective desalting technique that removes lipids, nucleotides, and other small molecules. Note: Specific and non-specific losses of proteins can occur with dialysis, gel chromatography, and precipitation/resuspension of samples. If the sample preparation cannot be modified, the effect of ionic impurities can be reduced by modifying the IEF protocol. Limit the voltage to 100–150 V for 2 hours, then resume a normal voltage step program. This pre-step allows the ions in the sample to move to the ends of the IPG strip.
	Ionic detergent in sample.	If the ionic detergent SDS is used in sample preparation, the final concentration must not exceed 0.25% after dilution into the rehydration solution. Additionally, the concentration of the non-ionic detergent present must be at least 8 times higher than the concentration of any ionic detergent to ensure complete removal of SDS from the proteins.
	High sample load.	Load less sample. Micropreparative separations require clean sample. Modify sample preparation to limit contaminants. (See section 2.4, 'Removal of contaminants that affect 2-D results.')
	Underfocusing. Focusing time was not long enough to achieve steady state focusing.	Program a low initial voltage and increase voltage gradually. Extend focusing time.
	Overfocusing. Extended focusing times (over 100,000 Vh) may result in electroendosmotic water and protein movement, which can produce horizontal smearing.	Reduce focusing time.

TABLE 26. TROUBLESHOOTING 2-D RESULTS (continued)

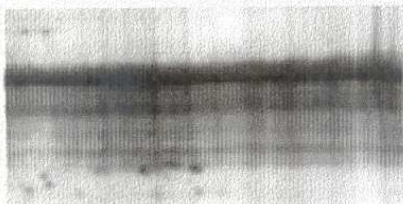
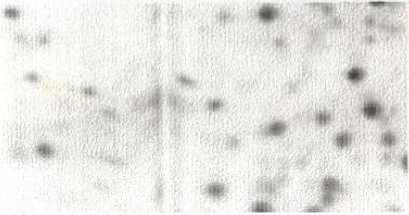
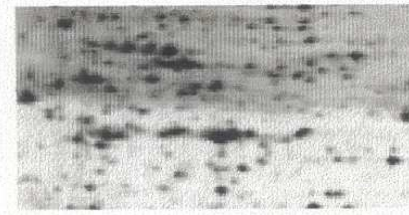
Symptom	Possible cause	Remedy
Horizontal stripes across gel 	Impurities in agarose overlay or equilibration solution.	Prepare fresh agarose overlay and equilibration solution.
Prominent vertical streak at the point of sample application (when loading IPG strips using sample cups) 	(Flatbed gel format) Sample aggregation or precipitation.	Dilute the sample and apply as a larger volume. Program a low initial voltage and increase voltage gradually.
Vertical streaking 	Insufficient equilibration.	Prolong equilibration time.
	(Flatbed gel format) Electroendosmosis.	Add 30% glycerol and 6 M urea to the SDS equilibration buffer. Place application pieces at the end of the strips during second-dimension electrophoresis to absorb excess water.
	Second-dimension buffer solutions prepared incorrectly.	Prepare fresh solutions.
	Insufficient SDS in SDS electrophoresis buffer.	Use 0.1% w/v SDS.
Vertical gap in 2-D pattern  	Impurities in sample.	Modify sample preparation. (See section 2.4, 'Removal of contaminants that affect 2-D results.')
	Impurities in rehydration solution components.	Use only high-quality reagents. De-ionize urea solutions

TABLE 26. TROUBLESHOOTING 2-D RESULTS (continued)

Symptom	Possible cause	Remedy
	Bubble between IPG strip and top surface of second-dimension gel.	Ensure that no bubbles are trapped between the IPG strip and the top surface of the second-dimension gel.
	(Flatbed gel format) Urea crystals on the surface of the IPG strip.	Allow residual equilibration solution to drain from the IPG strip before placing the strip on the second-dimension gel.
	(Flatbed gel format) Bubbles under the IPG strip.	Ensure that the IPG strip is placed firmly on the gel with no air bubbles trapped underneath. Stroke the plastic backing of the IPG strip gently with a pair of forceps to remove trapped bubbles.
Vertical regions of poor focusing		The IPG strip was not fully rehydrated.
		Ensure that the IPG strips are rehydrated with a sufficient volume of rehydration solution.
		Remove any large bubbles trapped under the IPG strip after rehydration solution is applied.
		Check that the rehydration solution is evenly spread along the entire length of the IPG strip.
Poor representation of higher molecular weight proteins	Proteolysis of sample.	Prepare sample in a manner that limits proteolysis and/or use protease inhibitors. (See section 2.2, 'Protection against proteolysis.')
	Insufficient equilibration.	Prolong equilibration time.
	Poor transfer of protein from IPG strip to second-dimension gel.	Employ a low current sample entry phase in the second-dimension electrophoresis run.
	Poor entry of sample protein during rehydration.	Use recommended volume of rehydration solution. (See Tables 10 and 15.)
Point streaking		(Silver staining). Dirty plates used to cast gel or particulate material on the surface of the gel. DTT and other thiol reducing agents exacerbate this effect.
		Properly wash glass plates. Scavenge any excess or residual thiol reducing agent with iodoacetamide before loading the IPG strips onto the second-dimension gel.
Background smear toward bottom of gel	(Silver or Coomassie blue staining) Staining of carrier ampholytes.	Use IPG Buffer as carrier ampholyte mixture. Reduce concentration if necessary.
Background smear toward top of gel	(Silver staining) Nucleic acids in sample.	Add DNase and RNase to hydrolyze nucleic acids. Note: The proteins DNase and RNase may appear on the 2-D map.
High background in top region of gel		Protein contaminant in SDS electrophoresis buffer or dirty electrophoresis unit.
		Make fresh SDS electrophoresis buffer. Clean electrophoresis unit.