

# Análise Discriminante de Fisher Aplicada à Detecção de Defeitos em Couro Bovino

Willian Paraguassu Amorim  
Universidade Católica Dom Bosco  
Grupo de Pesquisa em Engenharia e Computação  
Campo Grande - MS, Brasil  
paraguassuec@gmail.com

Hemerson Pistori  
pistori@ucdb.br

## Resumo

*Este artigo apresenta um estudo e aplicação da análise discriminante de Fisher na seleção de atributos baseados em textura e cor, para a detecção de defeitos em couro bovino. Um conjunto de 261 atributos, baseados nos métodos matriz de co-ocorrência, mapas de interação e atributos discretizados dos espaço de cores RGB e HSB, foram extraídos para a utilização da redução de dimensão por Fisher. Experimentos com dados reais indicaram que um vetor com 16 atributos criado a partir da redução dos 261 atributos extraídos, obtiveram a mesma taxa de classificação, diminuindo o tempo de treinamento e o tempo de classificação. Esse trabalho faz parte do projeto DTCOURO, cujo objetivo principal é criar um sistema automático de extração de parâmetros, a partir de imagens digitais, capaz de tornar mais eficiente e preciso o processo de classificação de peles e couros bovinos.*

## 1. Introdução

O Brasil possui o maior rebanho bovino do planeta, e apesar disto não consegue assumir um papel de destaque na indústria do couro bovino. A maior parte do couro produzido no país é de baixa qualidade, sendo que a maioria desse couro é comercializada nos estágios iniciais de processamento, com baixo valor agregado. A baixa qualidade do couro bovino brasileiro tem limitado o desenvolvimento do setor coureiro, afetando também a capacidade do país agregar valor ao produto internamente, e de usufruir das oportunidades oferecidas pela comercialização do couro de melhor cotação no mercado [8].

Inúmeros fatores contribuem no problema da qualidade do couro bovino, entre eles, a inexistência de sistemas de remuneração diferencial pela qualidade da matéria prima

produzida, baixa qualidade do couro, ocorrência de defeitos na fase produtiva, como: exposição aos ectoparasitas e manejo inadequado (ferrão, cerca de arame farpado, marcação a ferro quente em local impróprio, entre outros) [6] [9] [3].

Atualmente a classificação do couro bovino é realizada através de um especialista, que tem a tarefa de inspecionar cada peça do couro e levantar os defeitos encontrados, informando sua classificação<sup>1</sup>. Visando auxiliar o especialista na tarefa da classificação do couro bovino, foi criado uma parceria com a EMBRAPA/Gado de Corte e o GPEC (Grupo de Pesquisa em Engenharia de Computação) para a criação do projeto DTCOURO. O projeto DTCOURO tem como objetivo a construção de um sistema em visão computacional que será capaz de realizar a detecção automática das diversas inconformidades apresentadas no couro, e através destas informações, realizar sua classificação.

Nos artigos [12], [1] e [13], foram realizados alguns experimentos iniciais de classificação de defeitos do couro bovino. O objetivo desses experimentos foi implementar e analisar as técnicas de extração de atributos, como: matriz de co-ocorrência, mapas de interação e atributos de cores, e verificar o desempenho do classificador na discriminação entre as classes de defeitos.

Com base nesses experimentos, verificamos que foram obtidos bons resultados na classificação, mas os tempos de treinamento e classificação não foram tão satisfatórios. Para realizar a detecção automática de defeitos do couro bovino o sistema precisará realizar um varredura na imagem capturando suas amostras e dessa forma efetuar a classificação. Como os tempos de treinamento e classificação foram altos, é preciso encontrar alguma técnica que efetue a redução de atributos, acelerando de modo eficaz esse processo.

Este artigo apresenta alguns experimentos, indicando que o uso da análise discriminante de Fisher (FLDA), pode reduzir significamente o vetor de atri-

---

<sup>1</sup> Em alguns curtumes e frigoríficos são utilizadas as classificações de A (melhor qualidade) até D (pior qualidade).

butos de 261 para 16 atributos melhorando o tempo de treinamento e classificação sem perder sua precisão de discriminabilidade.

Os experimentos foram realizados com base em um grupo de 4 imagens do couro-cru. Para cada imagem foram feitas marcações sobre os defeitos, e a captura de amostras para cada marcação. Os defeitos analisados foram: carapato, esfolia, estria, marca-ferro, risco, regiões sem defeito e o local em que se encontra o couro bovino, chamado aqui de “fundo”. A partir das amostras capturadas, foi realizada a extração de atributos para a geração das bases de aprendizagem, que foram utilizadas na redução de dimensão. Os métodos de extração de atributos utilizados foram: matriz de co-ocorrência, mapas de interação e atributos relacionados a discretização e média para cada componente dos formato de cores RGB<sup>2</sup> e HSB<sup>3</sup>, a partir dos histogramas gerados para cada região de interesse. Nesse artigo iremos descrever somente o método de extração de atributos utilizando mapas de interação, pelo fato de ser menos conhecido e possuir poucas referências.

Esse artigo esta organizado da seguinte forma, na seção 2, é apresentado o método de extração de atributos utilizando mapas de interação que foi usado junto com matriz de co-ocorrência e atributos de cores na criação da base de conhecimento. A seção 3 apresenta a definição e teoria da técnica de Fisher utilizada para redução de atributos. Na seção 4 e 5 são apresentados os experimentos e resultados obtidos e na seção 6 são apresentadas as conclusões.

## 2. Mapas de Interação

Mapas de interação são muito utilizados na análise de pares de pixel de uma imagem [10]. Essa técnica consiste em calcular a diferença absoluta dos valores da intensidade dos pixels, localizado em uma determinada distância e ângulo da imagem. A utilização desse método pode capturar algumas das diferentes variações de textura encontradas no couro bovino. Como a matriz de co-ocorrência, o mapa de interação também trabalha com a intensidade em tons de cinza dos pixels de uma imagem. A Figura 1 ilustra como é aplicado essa técnica na análise de uma imagem.

A técnica consiste em calcular a média das diferenças absolutas dos pixels da imagem, com seus vizinhos  $img(i, j)$ , dado um deslocamento inicial  $d_i$ , deslocamento final  $d_f$ , ângulo inicial  $a_i$  e ângulo final  $a_f$ . Como resultado é criada uma nova matriz chamada mapa polar de interação  $map[n_d/d_{inc}][n_a/a_{inc}]$ , em que  $n_d$  representando a quantidade de distância entre os pares de pixels,  $n_a$  quantidade

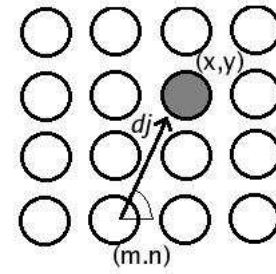


Figura 1. Aplicação do mapa de interação na criação do mapa polar.

de ângulos,  $d_{inc}$  distância de incremento entre os píxel e  $a_{inc}$  ângulo que será incrementado a partir do ângulo inicial e final[5] [4].

## 3. Análise Discriminante de Fisher e Seleção de Atributos

A análise discriminante, também conhecida como discriminante linear de Fisher (FLDA), é uma técnica que está se tornando muito comum em aplicações de visão computacional. Essa técnica utiliza informações das categorias associadas a cada padrão para extrair linearmente as características mais discriminantes. Através da análise discriminante de Fisher podemos também realizar a discriminação entre classes, através de processos supervisionados (quando se conhece o padrão) ou através de processos não supervisionados, em que é utilizado quando não se tem um padrão conhecido. A FLDA é um método supervisionado na concepção estatística e deve ser empregada quando algumas condições forem atendidas, como: (1) as classes sob investigação são mutuamente exclusivas, (2) cada classe é obtida de uma população normal multivariada, (3) duas medidas não podem ser perfeitamente correlacionadas, entre outras [7].

Nos trabalhos [2], [15] e [16] são apresentadas de forma detalhada a utilização da análise discriminante de Fisher. Essa técnica consiste na computação de uma combinação linear de  $m$  variáveis quantitativas, que mais eficientemente separam grupos de amostras em um espaço  $m$ -dimensional. Para isso a razão da variância entre-grupos e inter-grupos é maximizada. A separação intra-classe é realizada através da seguinte fórmula estabelecida por Fisher.

1. Espalhamento intra-classes;

$$S_w = \sum_{j=1}^c \sum_{i=1}^{T_j} (x_i^j - u_j) \cdot (x_i^j - u_j)^t, \quad (1)$$

2 O modelo RGB é um modelo de cor com três cores primárias: vermelho, verde e azul, com um intervalo de valores de 0 até 255.

3 O modelo de cores HSB se baseia na percepção humana das cores e não nos valores de RGB. Suas cores são representadas por H(matiz), S(saturação) e B(brilho).

em que  $x_i^j$  é o  $i$ -ésimo exemplo da classe  $j$ ,  $u_j$  é a média da classe  $j$ ,  $T_j$  é o número de amostras da classe  $j$  e  $c$  é a quantidade de classes;

2. Espalhamento inter-classes;

$$S_b = \sum_{j=1}^c (u_j - u) \cdot (u_j - u)^t, \quad (2)$$

em que  $u$  é a média de todas as classes.

A partir do cálculo de espalhamento intra-classe e inter-classe de um conjunto de amostras, é possível seguir o critério de Fisher, que é maximizar a medida inter-classes e minimizar a medida intra-classes. Uma forma de fazer isso é maximizar a taxa  $S_f = \det(S_b)/\det(S_w)$ . Essa fórmula mostra que se  $S_w$  é uma matriz não singular, então essa taxa é maximizada quando os vetores colunas da matriz  $S_f$  são os autovetores de  $S_w^{-1} \cdot S_b$ .

A redução de atributos é um problema de otimização que busca criar novas características através de uma combinação do conjunto original de características mantendo a eficiência no processo de classificação [11]. A dimensionabilidade do espaço de atributos pode resultar em problemas como alto custo de processamento e a geração do fenômeno conhecido como maldição da dimensionabilidade. Maldição da dimensionabilidade, pode ser caracterizado como uma degradação nos resultados de classificação, com o aumento da dimensionabilidade dos dados [14].

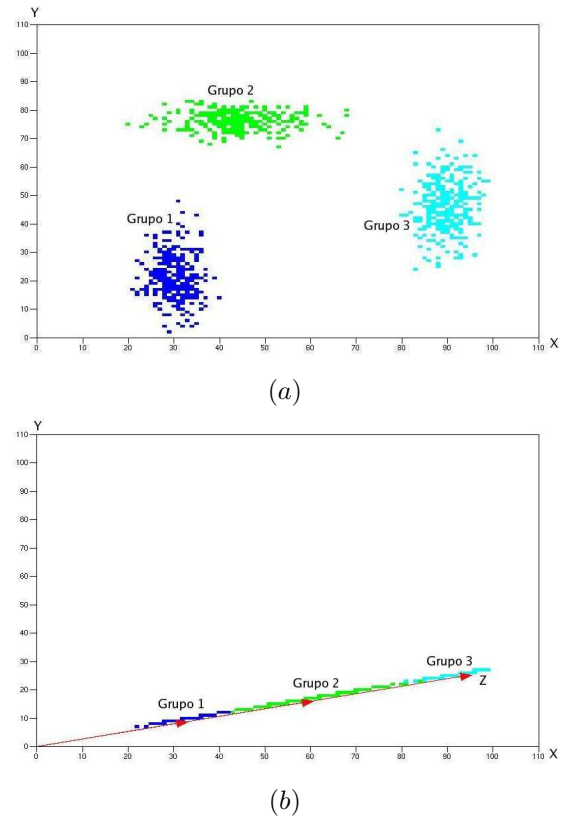
Em análise discriminante de Fisher a redução de atributos é realizado a partir de um conjunto de amostras para  $n$  classes, tendo  $p$  variáveis, com o objetivo de reduzir para  $m$  atributos. Para a redução de atributos por Fisher segue-se a seguinte linha de cálculo;

1. Calcular o espalhamento  $S_w$  e  $S_b$  para  $n$  classes;
2. Maximizar a medida inter-classes e minimizar a medida intra-classes  $S_f$  a partir de  $S_w^{-1} \cdot S_b$ .

A partir de  $S_f$  é possível a redução de atributos com base em seus auto-valores e auto-vetores, em que os atributos selecionados serão os auto-vetores referente aos seus  $m$  maiores auto-valores. A Figura 2(a) ilustra 3 conjuntos de indivíduos representando cada classe de classificação e utilizando dois atributos  $(x, y)$  para sua representação. A Figura 2(b) mostra um exemplo da projeção dos dados sobre um atributo reduzido  $(z)$ , demonstrando também que a discriminabilidade dos dados é mantida.

## 4. Experimentos

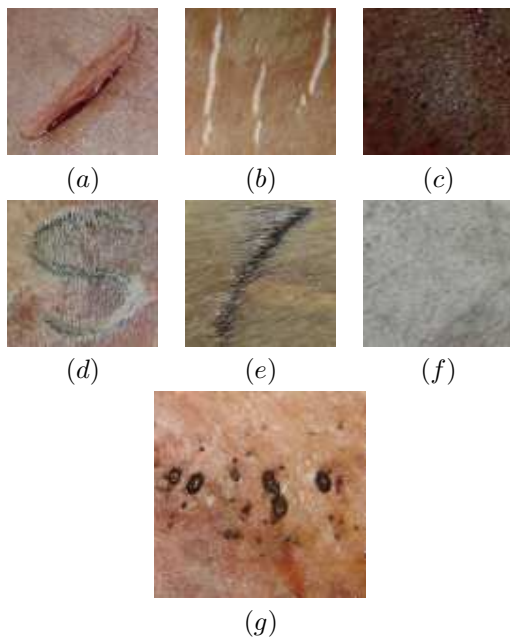
Para a realização desses experimentos foram utilizadas 4 imagens, sendo em cada, contendo os defeitos: carrapato, esfolia, estria, marca-ferro, risco e regiões (sem defeito e fundo). Em seguida, para cada imagem, realizamos



**Figura 2. Exemplo utilizando análise discriminante de Fisher para redução de variáveis (a) conjunto de amostras utilizando 2 atributos  $(x, y)$  e (b) redução de atributos para um único atributo  $(z)$**

as marcações referentes a cada classe. O objetivo dessa regra de marcação foi a necessidade de se obter uma maior variação de regiões sem defeito, fundo, e regiões com defeitos, próximas uma da outra. Para cada marcação foram capturadas todas as amostras possíveis numa escala de 20X20 píxel e com um intervalo de 1 píxel para cada amostra. A quantidade de amostras geradas para cada defeito foram: carrapato (366), esfolia (38), estria (2355), marca-ferro (391), risco (49), sem defeito (2664) e fundo (948). Na Figura 3 são ilustradas algumas amostras dos tipos de defeitos que foram capturados.

A partir de cada amostra capturada, realizamos a extração de atributos, utilizando as técnicas de matriz de co-ocorrência, mapas de interação, atributos discretizados dos modelos de cores HSB e RGB e sua médias. Para matriz de co-ocorrência utilizamos como propriedade de sua técnica variações no ângulo de  $(0^\circ$  a  $360^\circ)$  com intervalo nos ângulos de  $45^\circ$  e distância de 1 pí-



**Figura 3. Exemplo de amostras capturadas dos defeitos do couro cru. (a) esfola, (b) estria, (c) fundo, (d) marca ferro, (e) risco, (f) sem defeito e (g) carrapato**

xel. Em mapas de interação utilizamos variações de ( $0^\circ$  a  $360^\circ$ ) com intervalo inicial  $0^\circ$ , final  $360^\circ$  e incremento de  $45^\circ$  com distâncias de 1 e 2 pixels. As duas técnicas resultaram em duas matrizes, uma contendo o resultado da matriz de co-ocorrência entre os pixels e outra contendo o mapa polar do mapas de interação, e a partir disso foram capturados atributos, como: entropia, momento da diferença inversa, dissimilaridade, correlação, contraste, segundo momento angular e a diferença inversa. Foram extraídos também os valores médios e os histogramas discretizados em 32 intervalos de cada modelo dos formato de cores HSB e RGB. Foram capturados no total 261 atributos, sendo 63 atributos de textura e 198 atributos de cores. A Tabela 1, mostra com mais detalhes os atributos extraídos.

Com base nos 261 atributos extraídos das amostras e junto com a técnica implementada de Fisher, para redução de atributos, foram geradas 261 bases de aprendizagem, sendo em cada contendo  $a_i$  atributos tal que  $1 \leq a_i \leq 261$ . Utilizamos em nossos experimentos os algoritmos de classificação C4.5 e máquinas de vetores de suporte (SMO). O motivo que nos levaram a escolher o C4.5 e o SMO, foi que em experimentos anteriores seus resultados de classificação se comportaram de modo mais eficiente.

Para a realização dos experimentos, escolhemos

Método de Extração	Quantidade de Atributos
Média (H), (S), e (B)	3
Média (R), (G) e (B)	3
HSB (Discretizado cada modelo em 32 intervalos)	96
RGB (Discretizado cada modelo em 32 intervalos)	96
Mapas de Interação	7
Matriz de Co-ocorrência	56
Total	<b>261</b>

**Tabela 1. Atributos extraídos**

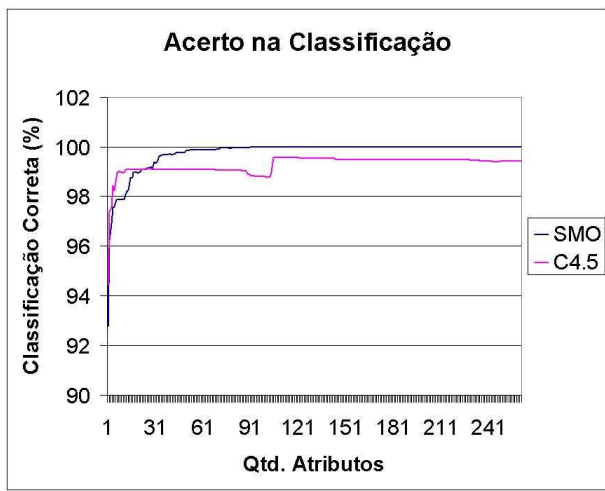
como modo de testes dos resultados a técnica estatística de validação cruzada com configuração de 5 dobras e 6 validações. Para a visualização dos resultados, foram selecionados as seguintes características: taxa(%) de classificação correta e tempos(s) de treinamento e classificação. Dessa forma será possível analisar se o desempenho do classificador irá ser mantido a medida que os atributos são reduzidos utilizando análise discriminante de Fisher.

Para a realização dos experimentos foram utilizados três módulos que pertencem ao projeto DTCOURO, que ajudaram a criar a base de conhecimento, que foram: (1) módulo de marcação e geração de amostras, que foi responsável pela criação do banco de imagens para marcação e geração das amostras específicas para cada tipo de defeito, (2) módulo de experimentos, que realizou a extração dos atributos a partir das amostras geradas e (3) módulo de classificação automática, que foi responsável pela classificação automática dos defeitos coletados. Nos experimentos foi utilizado um computador com processador Pentium 4 - 2.3 GHz com 1024 MB de memória RAM.

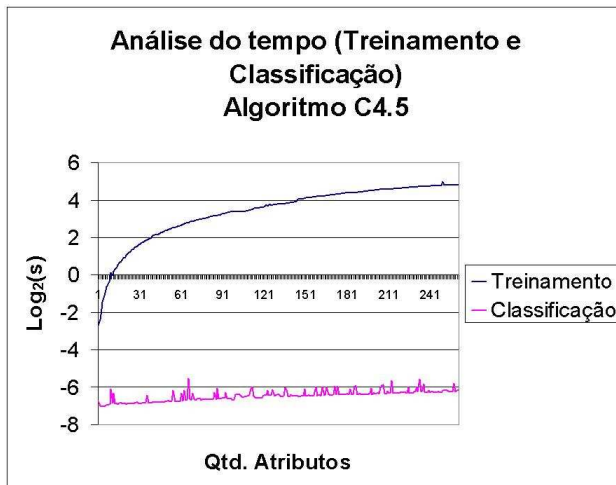
## 5. Resultados

Para avaliar os resultados, mostraremos graficamente o comportamento do classificador a medida que o número de atributos é reduzido. Nos experimentos foram utilizadas 6.811 amostras, para a classificação de 7 classes (carrapato, esfola, estria, marca-ferro, risco, sem defeito e fundo), utilizando os algoritmos de aprendizagem (C4.5) e (SMO). O tempo total para a realização dos 261 experimentos foram de 36 horas para C4.5 e 47 horas para SMO. Os Gráficos 4(a), 4(b) e 4(c), ilustram os resultados com base na porcentagem de acerto e os tempos de treinamento e classificação dos algoritmos C4.5 e SMO.

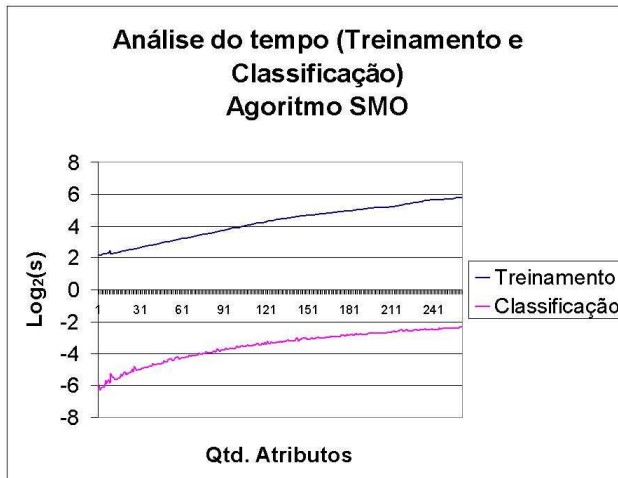
Com base nos resultados, podemos observar que a taxa de acerto do classificador teve um bom desempenho. Mas também acreditamos que as grandes variações



(a)



(b)



(c)

**Figura 4. Comportamento do classificador, (a) classificação correta(%) de C4.5 e SMO, (b) tempo de treinamento e classificação de C4.5 e (c) tempo de treinamento e classificação de SMO.**

nos defeitos marcados e pelas poucas imagens utilizadas poderam influenciar nos resultados. Analisando o comportamento do gráfico 4(a), podemos verificar que mesmo com a redução de atributos, mas com pequenas variações da taxa de acerto, a análise discriminante de Fisher apresentou um bom desempenho, mantendo o nível dos classificadores quase constante. Pelo gráfico também pode ser verificado que a porcentagem de acerto melhora a medida que a quantidade de atributos é acrescentada, isto até uma certa quantidade, em que há uma estabilidade. Para o algoritmo C4.5 a estabilidade é encontrada a partir de 106 atributos com 99,56% de acerto, e para o algoritmo SMO a estabilidade é visualizada a partir dos experimentos com 71 atributos com 100% de acerto. É verificado também que em experimentos com uma baixa quantidade de atributos a técnica de FLDA se comportou de modo eficiente, em que com 16 atributos o classificador apresentou como resultado de classificação 99,12% de acerto.

Os gráficos 4(b) e 4(c) foram apresentados com o objetivo de mostrar que a redução de atributos é uma etapa muito importante para classificação de imagens. A medida que a quantidade de atributos é acrescentada o tempo de treinamento e classificação aumentam, mostrando assim que a utilização de uma técnica para redução de atributos, como análise discriminante de Fisher, pode dar eficiência, velocidade e confiança para um classificador automático.

## 6. Conclusão

A detecção de defeitos e a classificação do couro bovino, com base em inspeção visual, é uma importante etapa da cadeia produtiva do boi. A automatização desse processo, através de sistemas computacionais, pode torná-lo mais preciso, menos subjetivo, menos sujeito a falhas e mais uniforme.

Com base nos experimentos realizados e em seus resultados, verificamos que a técnica de Fisher demonstrou um bom desempenho, mostrando que a partir da combinação estatística de um conjunto de atributos, podem ser criados um sub-conjunto com a mesma eficiência de classificação. Foi também verificado que os módulos implementados auxiliares facilitaram os experimentos para testes de classificação, reforçando as expectativas de que será possível a implantação de um classificador automático de couros bovinos. A baixa quantidade de imagens e a grande variação na semelhança dos defeitos, favoreceram o alto nível de acerto, porém para um experimento ideal, seria necessário um conjunto maior de imagens, em diferentes ambientes.

Durante os experimentos, alguns problemas foram encontrados, como: baixa quantidade de imagens de tipos de defeitos e problemas com um número muito alto de atributos.

tos utilizando a técnica de Fisher, provocando assim, problemas de matrizes singulares. A partir desses resultados alcançados, foi possível abrir novas perspectivas para trabalhos futuros, como: resolver o problema de matriz singular da redução de atributos por Fisher, para uma quantidade  $m$  de atributos e uma quantidade  $n$  de amostras, sendo  $n \leq m$ , e atributos com pouca variância; utilizar técnicas de segmentação para a remoção do fundo das imagens do couro bovino para realização de novos experimentos; utilizar a técnica de Fisher em outras aplicações, como: reconhecimento de texturas, gestos, faces, entre outras.

## Agradecimentos

Este trabalho recebeu apoio financeiro da Universidade Católica Dom Bosco, UCDB, da Agência Financiadora de Estudos e Projetos, FINEP e da Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, FUNDECT. O primeiro autor desse trabalho é bolsista DTI-CNPq e o segundo autor é bolsista de Produtividade em Desenvolvimento Tecnológico Extensão Inovadora do CNPq.

## Referências

- [1] W. P. Amorim, R. H. R. Viana, R. C. B. Rodrigues, and H. Pistori. Desenvolvimento de um software de processamento e geração de imagens para classificação de couro bovino. *Workshop de Iniciação Científica - SIBGRAPI*, 2006.
- [2] P. N. Belhumeur, J. P. Hespanha, and D. J. Kriegman. Eigenfaces vs. fisherfaces: Recognition using class specific linear projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1997.
- [3] G. A. Cardoso, E. and V. S. Lário. Análise da cadeia produtiva de peles e couros no brasil. *Embrapa - Comunicado Técnico*, pages (68):1–4, 2001.
- [4] D. Chetverikov. Structural filtering with texture feature based interaction maps: Fast algorithms and applications. In *ICPR96*, pages II: 795–799, 1996.
- [5] D. Chetverikov. Texture analysis using feature-based pairwise interaction maps. *Pattern Recognition*, 32(3):487–502, 1999.
- [6] A. Gomes. Como melhorar a qualidade do couro. *Gado de corte informa*, page 10, 1997.
- [7] A. Gomes. Aspectos da cadeia produtiva do couro no brasil e em mato grosso do sul. *Embrapa Gado de Corte*, pages 61–72, 2002.
- [8] A. Gomes. Avaliação técnica e operacional do sistema de classificação do couro bovino. *Embrapa Gado de Corte, Relatório Final do Projeto (CNPQ)*, 2005.
- [9] IEL. Instituto euvaldo lodi, cna confederação nacional da agricultura, e sebrae/nacional serviço brasileiro de apoio às micro e pequenas empresas. estudo sobre a eficiência econômica e competitividade da cadeia da pecuária de corte no brasil. page 398, 2000.
- [10] R. Jobanputra and D. Clausi. Texture analysis using gaussian weighted grey level co-occurrence probabilities. In *CRV04*, pages 51–57, 2004.
- [11] J. A. Olivera, L. V. Dutra, and C. D. Renn. Aplicações de métodos de extração e seleção de atributos para classificação de regiões. *Instituto Nacional de Pesquisas Espaciais - INPE*, 2005.
- [12] H. Pistori, W. P. Amorim, M. C. Pereira, P. S. Martins, M. A. Pereira, and M. A. C. Jacinto. Defect detection in raw hide and wet blue leather. *CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications*, 2006,.
- [13] L. Ramirez, W. P. Amorim, and P. S. Martins. Comparação de matrizes de co-ocorrência e mapas de interação na detecção de defeitos em couro bovino. *Workshop de Iniciação Científica - SIBGRAPI*, 2006.
- [14] J. C. Santos, J. R. F. Oliveira, and L. V. Dutra. Uso de algoritmos genéticos na seleção de atributos para classificação de regiões. *Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos - SP - Brasil*, 2004.
- [15] H. Yu and J. Yang. A direct lda algorithm for high-dimensional data - with application to face recognition. *Preprint submitted to Pattern Recognition Letters*, 2000.
- [16] W. Zhao, R. Chellappa, and P. J. Phillips. Subspace linear discriminant analysis to face recognition. *Partially supported by the Office of Naval Research*, 1999.