Defect Detection in Raw Hide and Wet Blue Leather

Hemerson Pistori & William A. Paraguassu & Priscila S. Martins & Mauro P. Conti
UCDB - Dom Bosco Catholic University, Campo Grande, Brazil
GPEC - Research Group in Engineering and Computing

Mariana A. Pereira & Manuel A. Jacinto
EMBRAPA - Brazilian Agricultural Research Corporation, Campo Grande, Brazil
EMBRAPA Beef Cattle Research Unit

This paper presents an important problem for the Brazilian economy, the classification of bovine raw hide and leather, and argues that this problem can be handled by computer vision and machine learning techniques. Some promising results, using standard techniques, like color based models and cooccurrence matrix based texture analysis, are reported. The paper also presents what seems to be the first major training and testing dataset of bovine raw hide and leather digital images.

1 INTRODUCTION

The importance of the bovine production chain for the Brazilian economy is a well acknowledged fact, as it has the largest cattle herd in the world (Matthey, Fabiosa, and Fuller 2004). In 2005, Brazil displaced the United States as the world's number one beef exporter (Valdes 2006). However, only recently government and private agents started planning proper policies to improve Brazilian role in the bovine leather markets.

The quality of Brazilian leather is far bellow the level of other important players. While in the United States, 85% of the leather are of highest quality, in Brazil, only 8.5% acchieve this classification (da Costa 2002). Nonethless, most of the problems that affects leather quality, like the use of barbed wire, brand marks made from hot iron in improper places, wrong transportation and flaying methods, have simple solutions.

Besides some lack of information about the leather importance for Brazilian economy, the main problem seems to be the absence of clear and fair policies and practices for raw hide (the untanned hide of a cattle) and leather commercialization. In most cases, bovine raw hides are treated as a by-product in the slaughtery, and for the raw hides, cattle ranchers are paid a fixed proportion directly related to the carcass weight. The creation of a remuneration system that rewards higher quality would stimulate more investments on leather care during cattle raising, when 60% of the problems that decrease leather value happens (da Costa 2002; Gomes 2002).

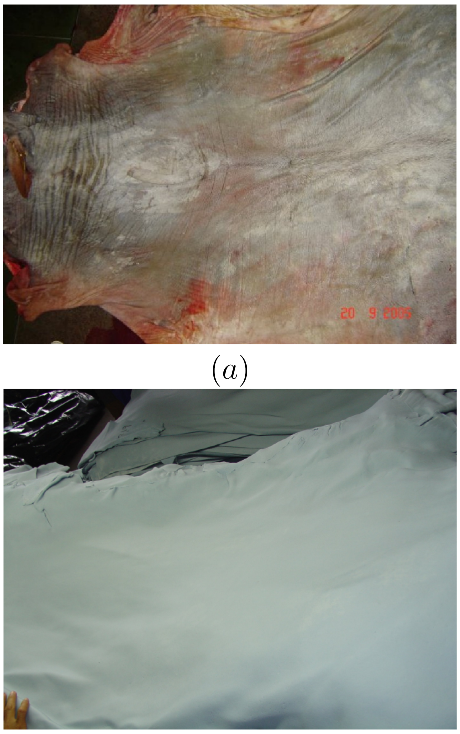
A fair remuneration system depends on standardized, non subjective and reliable classification and grading schemes. Brazilian government, by 2002, made a first attempt to create a national grading system for bovine raw hide [1]. A recent research conducted by the Brazilian Agricultural Research Corporation, EMBRAPA, suggests some improvements to the system and recommends the pursuing of automation to increase reliability. The work described in this paper is part of a scientific research and technological development project, DTCOURO [2], which envision the development of a completely automated system, based on computer vision, for bovine raw hide and leather classification and grading.

One of the DTCOURO project's goal is to propose and provide comparative studies of preprocessing, feature extraction, feature selection, segmentation and classification techniques. Among the feature extraction algorithms that will be experimented and compared are the ones based on Gabor Filters (Grigorescu, Petkov, and Kruizinga 2002), Windowed Fourier Transforms (Azencott, Wang, and

[1] Normative instruction number 12, December 18th, 2002, Brazilian Ministry of Agriculture, Livestock and Food Supply

[2] DTCOURO is an acronym, in Portuguese, for the project named Leather Defect Detection System, which has a Portuguese language website available at http://ommited for blind review

Younes 1997), Wavelets (Sobral 2005), Cooccurrence Matrices (Jobanputra and Clausi 2004), Interaction Maps (Chetverikov 2000), Local Binary Patterns (Menp 2002) and Color Models. Figures 1.(a) and 1.(b) show images from raw hide and wet blue leather [3]. A large image dataset of defects in hides and leather is being prepared and used to train and test supervised machine learning algorithms. Classifiers based on Support Vector Machines, Artificial Neural Networks and Bayesian Inference, among others, will be implemented and experimented.



(a)

(b)

Figure 1: (a) Bovine raw hide and (b) Wet blue leather samples

In this work, we present the first version of the leather defect dataset and the software that makes easier the creation of ground truth images and automatic comparison of different computer vision and machine learning techniques. We also show the results from the first defect inspection experiments, using cooccurrence matrices features and supervised learning. The experiments were based on a 2000 samples dataset, taken from 16 different raw hide and wet blue leather pieces and containing 4 defect types: brand marks made from hot iron, tick marks, cuts and scabies. The results were very promising, with an overall correct classification rate above 94%. The experiments with bovine raw hide images, a much more complex problem than wet blue and finished leather defect inspection, are rarely cited in the specialized literature, which concentrates on tanned leather. Furthermore, some important raw hide defects for the Brazilian leather productive chain are not always of interest in different countries where this type of defects are not so common, and usually not considered.

The next section presents some work directly related to leather inspection based on computer vision. Section 3 details important aspects of raw hide and wet blue leather defects. A brief review of texture analysis based on cooccurrence matrices can be found in Section 4. The following sections explain the experiments, discuss the results and presents the conclusions and future developments.

2 RELATED WORK

Yeh and Perng propose and evaluate semi-automatic methods for wet blue raw hide defects extraction and classification. The system has been tested in a large tannery for six months with reliable and effective results, when comparing with human specialists. Their work also presents an interesting taxonomy for leather defects classification and grading, based on the shape and size of affected area. The defects are classified as (1) thin spots: hair root, pinhole, putrid spot, dermatitis; (2) circular spots: thorn scratch, nail mark, chrome stain, slat stain, cured stain, putrefied; (3) thin line: vein, wring felt mark; (4) strips: score knife, neck wrinkle; (4) holes: dig damage, grub hole, bullet mark; (5) patterns: brand mark; and (6) irregulars: wart, contamination, pipe grain, flay mark, putrefied, scratch, chafe mark, gear mark, parasitic speckled (tick, mange), dung stain (Yeh and Perng 2001). The main contribution of the work is a fully quantified grading system, called *demerit count reference standard for leather raw hides*, but the authors also point out that one of the drawbacks of their proposal is the need for human, specialized intervention, for counting the total number of demerits on a wet blue raw hide.

A leather inspection method, based on Haar's wavelets, is presented by Sobral. The smoothing component has been tuned for each kind of defect, using a manually classified defect sample. The system is reported to perform in real time, at the same level of an experienced operator (Sobral 2005) and to outperform previous methods, based on Gabor filters, like the one described in Kumar and Pang (Kumar and Pang 2002) . Although not clearly stated in Sobral's paper, the system seems to have been experimented only on finished leather, a much simpler problem than raw hide or wet blue raw hide defect extraction.

A dissimilarity measure based on χ^2 criteria has been used to compare gray-level histograms from sliding windows (65x65 pixels) of a wet blue raw hide image to an averaged histogram of non-defective sam-

[3] Wet blue leather is a hide that has been tanned using chrominus sulphate. It is an intermediate stage between untanned and finished leather

ples in (Georgieva, Krastev, and Angelov 2003). The results of the χ^2 test and an experimentally chosen threshold are used to segment defective regions of the raw hide. The approach has not been used to identify the defect type. The segmentation of defective regions from wet blue raw hide images, using histogram and cooccurrence based features, has been investigated in (Krastev, Georgieva, and Angelov 2004). This work also proposes the use of fuzzy logic to model leather defects, but do not give sufficient details or experimental results supporting the proposal.

3 RAW HIDE AND WET BLUE LEATHER DEFECTS

Bovine leather undergoes a long way from cattle raising to final industrial production of leather goods, like furniture, footwear, belts and so on. The problems that affect leather quality begin when the animal is still alive, and include, (1) cuts resulting from barbed wired, in-fighting among male members and thorn scratches and cuts; (2) brand marks made for ownership purposes, using hot iron; (3) holes and spots from infections and infestations, caused by ticks, horn flies, manges and bot flies, among others; (4) abscesses resulting from wrong vaccination techniques and natural growth marks or excess weight related problems, like furrows and wrinkles (Barlee, Lanning, and McLean 1999; Roberts and Etherington 1981).

During transportation, the animal skin may suffer deep injures from nails and wood splints in the truck. Before tanning, three important process, which can also cause leather damage, happen: bleeding, skinning and curing. Insufficient bleeding can cause vain marks, while wrong skinning techniques may result in flaying cuts that, in some cases, may turn unusable otherwise valuable parts of the leather. As the raw hide is subjected to putrefaction, as soon as the animal dies, the raw hide must suffer a curing process to protect it until the tanning process begins, which can take months. Improper curing may lead to rotting and putrefaction. Defects during tanning and post-processing are much less common, as they are controlled by the tanneries, which have in the leather quality their main business.

Even without defects, bovine raw hide has a very complex surface, presenting different textures, colors, shapes and thickness. Besides, in order to be useful, automatic classification system should function in very different environments, like farms, slaughteries and tanneries. Blood or water drops may turn the task even more difficult if the raw hide is to be classified just after skinning or cleaning. Classification of living animals must deal with shadows from hair and the natural bovine anatomy.

Figures 2 and 3 illustrate the diversity of shapes, colors and texture that arise from some important de-

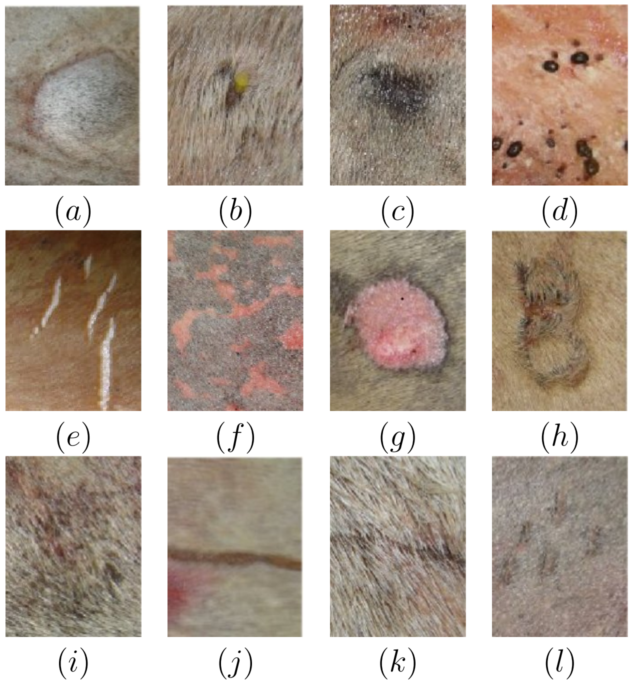


Figure 2: Examples of defects in raw hides: (a) vaccine abscess (b) bot fly open wounds (c) bot fly closed wounds (d) ticks marks (e) wrinkles (f) photosensibility (g) flay mark (h) brand marks made from hot iron (i) horn fly wounds (j) open cuts (k) closed cuts (l) scabies

fects that happens in Brazilian raw hides and leather, decreasing their market value. Images from raw hides, taken after skinning and before tanning, are shown in Figure 2. Figure 3 shows images from leather in the first stage of the tanning process, which are called wet blue leather.

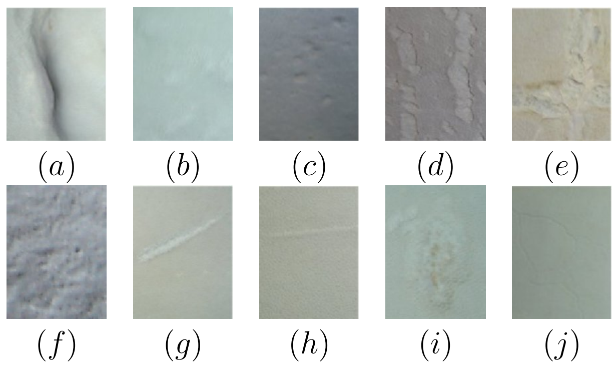
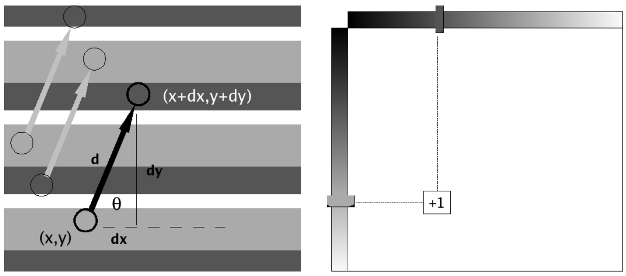


Figure 3: Examples of defects in wet blue leather: (a) vaccine abscess (b) bot fly wounds (c) ticks marks (d) wrinkles (e) brand marks made from hot iron (f) Haematobia irritans wounds (g) open cuts (h) closed cuts (i) scabies (j) veining

4 GRAY SCALE COOCCURRENCE MATRICES

Image segmentation based on features extracted from gray-scale coocurrence matrices, GLCM, is a common and largely used technique in texture analysis (Singh and Singh 2002; Jobanputra and Clausi 2004; Latif-Amet, Ertuzun, and Ercil 2000). As wavelets, windowed Fourier or Gabor filter approaches, coocurrence matrices can represent information related to the frequential distribution content of the original, spatially represented, image. Several GLCMs must be constructed for each sliding window that scans the image during segmentation. Each GLCM has an associated angle and displacement, related to the direction and frequency that will be represented by this GLCM. The number of different angles and displacements, and consequently, the number of GLCMs to be constructed depends on the problem and computer power available.



(a) (b)

Figure 4: Construction of a GLCM: (a) Angle and displacement parameters in spatial domain (original image) (b) Gray-level coocurrence matrix

Figure 4 illustrates the construction of a GLCM for a fixed angle, θ, and displacement, d. The GLCM, G, is an $m \times m$ accumulator, where m is the number of gray levels. For each pixel, (x, y), of the original image (or image window), the accumulator cell (i, j) is incremented, where $i = I(x, y)$, $j = I(x + dx, y + dy)$ and $I(.)$ is the pixel gray level. Interpolation must be used for certain angles and displacements, as $dx = d \times cos(\theta)$ and $dy = d \times sin(\theta)$ may be non- integers. Usually, the values of the GLCMs are not directly used as texture features, but some statistics and indices calculated from them, like entropy, contrast, angular second moment, inverse difference moment, energy and homogeneity.

5 LEATHER DATASET AND EXPERIMENTS

The DTCOURO's dataset includes, currently, images from 258 different pieces of raw hide and wet blue leather showing 17 different defect types. The images have been taken using a five megapixel digital camera during technical visits to slaughteries and tanneries located in the region of Mato Grosso do Sul, Brazil, by September 2005. A total of 66 pictures are from tanned leather in the wet blue stage, the other ones are from the raw hide stage.

For this first experiment, a set of four types of defect has been chosen: tick marks, brand marks made from hot iron, cuts and scabies. These defects have been chosen because they are very common in Brazil. Sixteen images, both from raw hide and wet blue leather, containing these defects (two image for each defect, both in raw hide and wet blue leather), were used. The defects were manually segmented with the help of a specialist and small samples from defective and non-defective areas have been extracted, using the software developed in DTCOURO's project. The samples used in these experiments are of windows of 10x10, 20x20, 30x30 and 40x40 pixels. A total of 2000 samples have been generated in this way, 400 for each defect and 400 for non-defective regions. Figure 5 illustrates some of the 40x40 size samples, from wet blue leather (first line) and raw hides (second line), that were used in the experiments.

(a) (b) (c) (d) (e)

Figure 5: Examples of 40x40 image windows used in the wet blue leather (first line) and raw hide (second line) experiments: (a) ticks (b) brand marks (c) cuts (d) scabies (e) non defective

From each sample, a set of 63 texture and 3 color features were extracted. The color features are the mean values of the histograms for the hue, saturation and brightness in HSB color space. The texture features are the entropy, inverse difference moment, dissimilarity, correlation, contrast, angular second moment and inverse difference entropy, calculated from the grey level coocurrence matrices, with angles 0, 45 and 90 (in degrees) and displacements 1, 5 and 10 (in pixels). A 10-fold cross-validation scheme was used to train and test three supervised learning approaches for this five class classification problem: (1) support vector machines (Keerthi, Shevade, Bhattacharyya, and Murthy 2001), (2) normalized Gaussian radial basis function network (Figueiredo 2000) and (3) k-nearest neighbours (Aha, Kibler, and Albert 1991). The experiment was repeated 5 times, resulting in a total of 50 runs, for each learning approach and each window size. The implementations for the machine-learning algorithms were taken from the Weka free-software, which were also used to produce performance statistics (Witten and Frank 2005).

The sequential minimal optimization algorithm (Keerthi, Shevade, Bhattacharyya, and Murthy 2001), implemented in Weka, has been used to train the support vector machines, with a third-order polynomial kernel. Pairwise classification were used to enable the application of this 2-class classification algorithm to our multi-class problem. Ridge estimators and logistic regression (Cessie and van Houwelingen 1992) were used to learn the normalized Gaussian radial basis functions, clustered by k-means, with k=5 (4 defects and 1 clean leather). The k-nearest neighbour approach was tested with 5 neighbours, weighted by the inverse of their distance. All the parameters were experimentally chosen to enhance percentage correct classification rates for each learning scheme, using Weka's implementation.

6 RESULTS AND DISCUSSION

The correct classification rates, for each classifier, using defective (in four different types) and non-defective images from wet blue leather and raw hides are shown, respectively, in Tables 1 and 2. A two-tailed, t-student test, was used to infer improvement or degradation using support vector machines (SVM) performance as the null hyphoteses. It is clear, from the tables, that all three classifiers achieve excellent correct classifications rates when 20X20 or larger window size were used. Support vector machines achieved good results (above 94%) even with 10X10 windows.

Table 1: Percentual of correct classification in wet blue leather, for the three learning approaches experimented: Support Vector Machines (SVM), Radial Basis Functions Networks (RBF) and Nearest Neighbours (KNN).

Data Set	SVM	RBF	KNN
10X10	97.44	87.40 •	90.42 •
20X20	99.82	98.40 •	99.36
30X30	100.00	99.58	99.66
40X40	100.00	99.74	100.00

◦ - improvement, • - degradation

Table 2: Percentual of correct classification in raw hides

Data Set	SVM	RBF	KNN
10X10	94.32	81.18 •	91.56 •
20X20	99.88	96.50 •	99.20 •
30X30	100.00	99.06	99.90
40X40	100.00	99.92	100.00

◦ - improvement, • - degradation

These results must be considered with some care and as a first, initial, experiment that, nonetheless, should encourage further research. Each defect were taken from only two different pieces of leather and do not represent, mainly in the case of raw hides, all the possible configuration in which the defect could appear, as, for instance, different hair sizes, colors and directions. It is very important to note that processing time were not an issue in these initial experiments and the parameter extraction phase, applied to high resolution images of the full leather or hide piece, can take almost half an hour.

7 CONCLUSION

Some experiments have been conducted in order to verify the applicability of texture analysis and machine learning techniques to the problem of defect detection and classification in bovine hides and leather. Support vector machines trained with the sequential minimal optimization algorithm using attributes extracted from cooccurence matrices presented the highest correct classification rates. The detection of defects in bovine raw hides was not, as far as the authors are concerned, handled in previous work, at least using computer vision approaches. One of the main contributions of this work is to present an important, but not extensively studied, problem.

The dataset will be enlarged, in the near future, with images from defects in live animals and skins with longer hair from the southeast regions of Brazil. Experiments with the complete dataset and other attributes extraction and learning techniques should also be conducted. It must be investigated if the same classification accuracy could be achieved with less computationally expensive approaches, other than cooccurence matrices.

Acknowledgments

This work has been funded by the Dom Bosco Catholic University, UCDB, and the Brazilian Studies and Projects Funding Body, FINEP. The author of this paper and some of his advisees hold scholarships from the Brazilian National Counsel of Technological and Scientific Development, CNPQ

REFERENCES

Aha, D. W., D. Kibler, and M. Albert (1991). Instance-based learning algorithms. *Machine Learning 6*, 37–66.

Azencott, R., J. Wang, and L. Younes (1997, February). Texture classification using windowed Fourier filters. *19*(2), 148–153.

Barlee, R., D. Lanning, and W. McLean (1999). The manufacture of leather. *Journal of Designer Bookbinders 19*, 48–59.

Cessie, S. L. and J. C. van Houwelingen (1992). Ridge estimators in logistic regression. *Applied Statistics 41*(1), 191–201.

Chetverikov, D. (2000). Structural defects: General approach and application to textile inspection. In *ICPR00*, Volume 1, pp. 521–524.

da Costa, A. B. (2002, Dezembro). Estudo da competitividade de cadeias integradas no brasil: Impactos das zonas de livre comrcio. Technical report, Instituto de Economia da Universidade Estadual de Campinas.

Figueiredo, M. A. (2000). On gaussian radial basis function approximations: Interpretation, extensions, and learning strategies. In *Proceedings of the 15th International Conference on Pattern Recognition*, Volume 2, pp. 2618–2622.

Georgieva, L., K. Krastev, and N. Angelov (2003). Identification of surface leather defects. In *CompSysTech '03: Proceedings of the 4th international conference conference on Computer systems and technologies*, New York, NY, USA, pp. 303–307. ACM Press.

Gomes, A. (2002). Aspectos da cadeia produtiva do couro bovino no Brasil e em Mato Grosso do Sul. In *Palestras e proposies: Reunies Tcnicas sobre Couros e Peles, 25 a 27 de setembro e 29 de outubro a 1 de novembro de 2001*, pp. 61–72. Embrapa Gado de Corte.

Grigorescu, S., N. Petkov, and P. Kruizinga (2002). Comparison of texture features based on Gabor filters. *IEEE Trans. on Image Processing 11*(10), 1160–1167.

Jobanputra, R. and D. Clausi (2004). Texture analysis using gaussian weighted grey level co-occurrence probabilities. In *Proceedings of the Canadian Conference on Computer and Robot Vision - CRV*, pp. 51–57.

Keerthi, S., S. Shevade, C. Bhattacharyya, and K. Murthy (2001). Improvements to platt's SMO algorithm for SVM classifier design. *Neural Computation 13*(3), 637–649.

Krastev, K., L. Georgieva, and N. Angelov (2004). Leather features selection for defects' recognition using fuzzy logic. In *CompSysTech '04: Proceedings of the 5th international conference on Computer systems and technologies*, New York, NY, USA, pp. 1–6. ACM Press.

Kumar, A. and G. Pang (2002, March). Defect detection in textured materials using gabor filters. *IEEE Transactions on Industry Applications 38*(2).

Latif-Amet, A., A. Ertuzun, and A. Ercil (2000, May). An efficient method for texture defect detection: Sub-band domain co-occurrence matrices. *Image and Vision Computing 18*(6), 543–553.

Matthey, H., J. F. Fabiosa, and F. H. Fuller (2004, May). Brazil: The future of modern agriculture. *MATRIC*.

Menp, M. P. T. O. T. (2002). Multiresolution gray scale and rotation invariant texture classification with local binary patterns. In *IEEE Transactions on Pattern Analysis and Machine Intelligence 24*, pp. 971–987.

Roberts, M. and D. Etherington (1981). *Bookbinding and the Conservation of Books: A Dictionary of Descriptive Terminoloy*. Library of Congress.

Singh, M. and S. Singh (2002). Spatial texture analysis: a comparative study. In *ICPR02*, pp. I: 676–679.

Sobral, J. L. (2005, September). Optimised filters for texture defect detection. In *Proc. of the IEEE International Conference on Image Processing*, Volume 3, pp. 565–573.

Valdes, C. (2006, April). Brazil emerges as major force in global meat markets. *Amber Waves - The Economics of Food, Farming, Natural Resources and Rural America*.

Witten, I. H. and E. Frank (2005). *Data Mining: Practical Machine Learning Tools and Techniques*. San Francisco: Morgan Kaufmann.

Yeh, C. and D. B. Perng (2001). Establishing a demerit count reference standard for the classification and grading of leather hides. *International Journal of Advanced Manufacturing 18*, 731–738.