

# ANÁLISE DISCRIMINANTE PARAMÉTRICA PARA RECONHECIMENTO DE DEFEITOS EM TÁBUAS DE EUCALIPTO UTILIZANDO IMAGENS DIGITAIS<sup>1</sup>

Joseph Kalil Khoury Junior<sup>2</sup>, Francisco de Assis de Carvalho Pinto<sup>3</sup>, Nerilson Terra dos Santos<sup>4</sup>, Ricardo Marius Della Lúcia<sup>5</sup> e Eduardo Eiji Maeda<sup>6</sup>

RESUMO – A indústria de madeira tem dispensado especial atenção às etapas de classificação e seleção de madeira serrada. Sistemas de Visão Artificial têm sido propostos para automação dessas etapas na indústria. A identificação de características apropriadas para discriminar os defeitos da madeira em imagens digitais é um dos maiores desafios no desenvolvimento desta tecnologia. O objetivo deste trabalho foi avaliar, por meio de técnicas de análise multivariada, a capacidade de discriminar defeitos em tábuas de eucalipto, utilizando-se as características de percentis de imagens coloridas. Foram realizadas análises discriminantes linear e quadrática para classificação de defeitos e madeira limpa em imagens digitais de tábuas de eucaliptos. As características de percentis do histograma das bandas do vermelho, verde e azul, retiradas de dois tamanhos de blocos de imagens, foram utilizadas para desenvolvimento e teste das funções discriminantes. Foram usados 492 blocos, contendo os 12 defeitos estudados e madeira limpa, retirados das imagens de 40 tábuas amostradas aleatoriamente. As características foram analisadas com seus valores originais, escores dos componentes principais e escores das variáveis canônicas. Os menores erros globais de classificação foram 19 e 24% para funções discriminantes lineares com os escores das variáveis canônicas para tamanho de bloco de 64 x 64 e 32 x 32 pixels, respectivamente. Tendo em vista a magnitude desses erros, as características de percentis foram consideradas adequadas para discriminar defeitos e madeira limpa em imagens digitais.

Palavras-chave: Processamento de imagem, reconhecimento de padrões e classificação de madeira serrada de eucalyptus.

## PARAMETRIC DISCRIMINANT ANALYSIS FOR RECOGNITION OF DEFECTS IN EUCALYPTUS LUMBER USING DIGITAL IMAGES

*ABSTRACT – The lumber industry has given special attention for lumber grading and selection stages. Machine Vision Systems have been proposed as a technological solution for automation of these stages. The proper feature selection for discriminating defect and clear wood is one of the most challenging in the development of such technology. The objective of this work was to evaluate, using multivariate analysis, the discriminating power of color images percentis. In this work, linear and quadratic discriminant analysis were accomplished for classification of defects and clear wood in digital images of eucalyptus lumber. The percent features of the histogram for the red, green and blue bands, from two sizes of image blocks were used for developing and testing the discriminant functions. 492 blocks were used, containing the 12 studied defects and clear*

---

<sup>1</sup> Recebido em 16.12.2003 e aceito para publicação em 25.11.2004.

<sup>2</sup> Departamento de Engenharia Agrícola da UFV, Viçosa - MG. 36570-000. E-mail: <kalil@ufv.br>.

<sup>3</sup> Departamento de Engenharia Agrícola da UFV, Viçosa - MG. E-mail: <facpinto@ufv.br>.

<sup>4</sup> Departamento de Informática da UFV, Viçosa - MG. E-mail: <nsantos@dpi.ufv.br>.

<sup>5</sup> Departamento de Engenharia Florestal da UFV, Viçosa - MG. E-mail: <rdlucia@ufv.br>.

<sup>6</sup> Graduação em Engenharia Agrícola e Ambiental da UFV, Viçosa - MG. E-mail: <eduardoeiji@zipmail.com.br>.



*wood, derived from images of 40 lumbers randomly sampled. The features were analyzed with their original values, scores of the principal components and scores of the canonical variables. The smallest global misclassification errors were 19% and 24% for linear discriminant function with the canonical variable scores using block sizes of 64x64 and 32x32 pixels, respectively. The percent features were considered appropriate to discriminate defects and clear wood in digital images.*

*Key words: Image processing, pattern recognition, eucalyptus lumber grading.*

## 1. INTRODUÇÃO

Na seleção de tábuas para fins de fabricação de móveis e projeto de interiores, a aparência é um requisito importante. Classificando e selecionando as tábuas, pode-se conseguir um produto com a aparência desejada, mesmo em madeiras oriundas de florestas plantadas sem manejo adequado. Dessa forma, a indústria de madeira serrada tem dispensado especial atenção às etapas de classificação e seleção de tábuas, com isso agregando valor ao seu produto. Essas etapas são, geralmente, realizadas visualmente por classificadores humanos. A etapa de classificação é considerada atividade estressante, em razão da exigência de concentração para identificar e mensurar defeitos nas tábuas e classificá-las por tempo prolongado.

Os sistemas que utilizam as informações, adquiridas por meio da energia eletromagnética refletida pelos objetos, para automatizar processos são conhecidos como Sistema de Visão Artificial (SVA). Um SVA para inspeção de superfícies consiste das seguintes etapas: aquisição, processamento e análise de imagens (PHAM e ALCOCK, 1998).

Na primeira etapa são utilizados diversos tipos de sensores para essa tarefa, sendo os sensores ópticos os mais empregados. A etapa de processamento da imagem tem como objetivo a segmentação (realce) e reconhecimento (identificação e mensuração) dos objetos de interesse (KLINE et al., 1992). Na etapa de análise, a informação é transformada de sua forma pictorial para expressões, que afetam a classificação da tábua.

Um SVA, quando implementado na indústria, não apresenta o mesmo desempenho da etapa de desenvolvimento, reduzindo a exatidão da classificação. Geralmente, isso ocorre devido ao fato de os algoritmos desenvolvidos para esses sistemas não se ajustarem às variações e imprevistos, que acontecem no meio industrial. No entanto, esses sistemas têm se mostrado superiores à baixa exatidão de classificadores humanos (HUBER et al., 1985). A maior dificuldade desse tipo

de tecnologias dá no desenvolvimento de algoritmos na etapa de processamento das imagens, por estarem diretamente relacionados ao erro do sistema (PHAM e ALCOCK, 1998; KLINE et al., 2001).

As câmeras coloridas, com fotossensores sensíveis nas bandas do vermelho (Vm), verde (Vd) e azul (Az), são um dos métodos de aquisição de imagens mais estudados por pesquisadores em classificação de superfícies de madeiras (SZYMANI e MCDONALD, 1981; KLINE et al., 1998; HANEY et al., 1994; LU, 1997; KAUPPIINEN, 1999). Essas pesquisas têm demonstrado que o espaço de cor Vm, Vd e Az vem alcançando melhor desempenho na classificação de madeiras, em comparação com câmeras monocromáticas.

Lebow et al. (1996) estudaram as características das curvas espectrais adquiridas por espectrorradiômetro na faixa do visível de oito tipos de defeitos em tábuas de abeto. Realizaram análises discriminantes linear e quadrática dos dados originais das reflectâncias, com resolução espectral de 10 nm, e dos escores dos componentes principais. A função discriminante quadrática obteve melhores resultados do que a linear, tanto dos dados originais quanto dos escores dos componentes, apresentando erros de classificação global dos defeitos de aproximadamente 3%. Eles recomendaram a investigação do uso de câmeras coloridas para classificação de tábuas em serrarias, graças aos bons resultados encontrados com as características da reflectância.

Koivo et al. (1989) desenvolveram um algoritmo para classificar nove classes de defeitos em carvalho-vermelho. Imagens foram adquiridas por câmera monocromática com 255 níveis de cinza. Inicialmente, esses pesquisadores verificaram que as funções discriminantes lineares, utilizando média e desvio-padrão dos níveis de cinza, separaram somente três classes. Assim, propuseram outras cinco características, obtidas a partir da modelagem da textura de blocos da imagem. Utilizando uma árvore de decisão com os três componentes principais desse vetor de características,

foi possível diferenciar todas as nove classes de interesse (casca, madeira limpa, podridão, furo, nós, inclusão mineral, rachadura, mancha e esmoado), sendo o erro de classificação de 2,8%.

As características de texturas melhoram a exatidão de classificação. Entretanto, características texturais são de aquisição complexa, tornando-se inviáveis na utilização de SVA que objetivam a classificação em tempo real.

Kauppinen (1999) desenvolveu um classificador k-NN para seis tipos de defeitos de tábuas de pinos, utilizando características de percentis do histograma das bandas Vm, Vd e Az, sendo o erro de classificação de 34%. O alto erro foi atribuído à dificuldade em definir algumas classes de defeitos da amostra, como os tipos de nós – o nó preto, por exemplo, por exemplo, é uma subcategoria do nó seco. Concluiu que os percentis são simples de calcular e possuem discriminações satisfatórias, permitindo implementação em tempo real.

O objetivo deste trabalho foi avaliar, por técnicas de análise multivariada, a capacidade de discriminar defeitos em tábuas de eucalipto, utilizando-se as características de percentis de imagens coloridas.

## 2. MATERIALE MÉTODOS

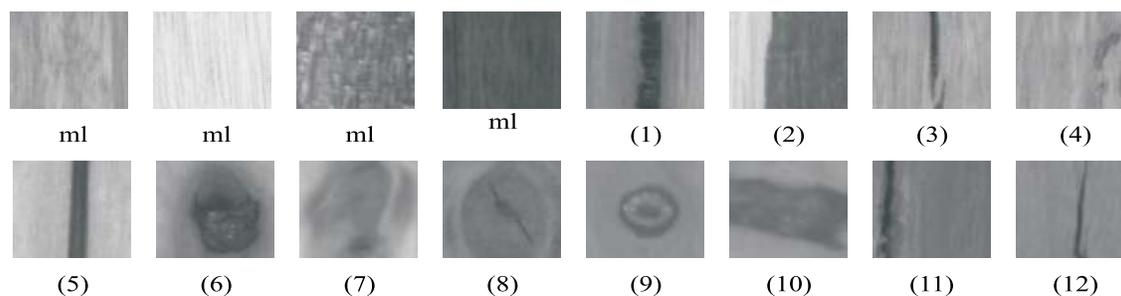
Foram amostradas 40 tábuas de eucalipto secadas ao ar e previamente aplainadas, provenientes do Laboratório de Propriedades Físicas e Mecânicas da Madeira da UFV. As amostras foram escolhidas de forma que estariam presentes, além de madeiras limpas de coloração mais clara ou mais escura, com ou sem alburno, os defeitos visuais mencionados na norma NBR 14806

(ABNT, 2002b): (1) bolsa de goma, (2) esmoado, (3) fendilhado, (4) furos de inseto, (5) medula, (6) nó cariado, (7) nó sadio, (8) nó sadio corrompido, (9) nó solto encaixado, (10) nó solto oco, (11) podridão e (12) racha.

Para aquisição das imagens, foi utilizada a câmera colorida Duncantech, modelo MS3100, conectada a um computador por meio de uma placa de aquisição de vídeo National Instruments, modelo PCI 1424. A altura da câmera foi de 0,9 m da superfície da face larga da tábua e a resolução espacial, de 0,24 x 0,24 mm por pixel. O sistema foi iluminado por dois refletores com difusor de luz com lâmpadas halógenas de 300 watts, instaladas em refletores com difusor de luz. O iluminamento na superfície da face larga da tábua foi de 2.500 lux.

As imagens das tábuas, de 1.392 pixels na horizontal e 1.039 pixels na vertical, foram adquiridas nas suas duas faces largas. Feita a varredura das faces largas das tábuas, obtiveram-se 594 imagens. Destas, foram cortados blocos da imagem que representaram as classes de defeito e madeira limpa, obtendo-se a seguinte amostra: 41 de madeira limpa de alburno, 48 de madeira limpa clara, 24 de madeira limpa clara áspera, 74 de madeira limpa escura, 25 de bolsa de goma, 9 de esmoado, 16 de fendilhado, 34 de furos de inseto, 54 de medula, 46 de nó cariado, 34 de nó sadio, 29 de nó sadio corrompido, 28 de nó solto encaixado, 9 de nó solto oco, 5 de podridão e 16 de racha.

Foram testados dois tamanhos de blocos de imagens, de 64 x 64 e 32 x 32 pixels. Na Figura 1, são mostrados exemplos de blocos de imagens, que representam as classes de defeitos e madeira limpa estudadas.



**Figura 1**– Exemplo de blocos de imagens de 64 x 64 pixels de (mL) madeira limpa, (1) bolsa de goma, (2) esmoado, (3) fendilhado, (4) furos de inseto, (5) medula, (6) nó cariado, (7) nó sadio, (8) nó sadio corrompido, (9) nó solto encaixado, (10) nó solto oco, (11) podridão e (12) racha.

**Figure 1** – Sample of 64 x 64 pixel images blocks of (ml) clear wood, (1) pitch pocket, (2) wane, (3) seasoning check, (4) insect holes, (5) pith, (6) unsound knot, (7) sound knot, (8) rotten sound knot, (9) encased loose knot, (10) loose knot, (11) decay e (12) splits.

### Características das imagens

As características retiradas de cada bloco da imagem foram 12 percentis das bandas Vm, Vd e Az, construindo, assim, um vetor de 36 características. Utilizaram-se os percentis de 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90 e 95 %. A diferença entre os percentis 2 e 95% foi usada para normalizar o vetor.

$$C_{Bd}^y = \frac{P_{Bd}^y}{P_{Bd}^{95\%} - P_{Bd}^{2\%}}$$

em que  $C_{Bd}^y$  = característica normalizada do percentil  $y$  na banda  $Bd$ ;  $P_{Bd}^y$  = nível de cinza da banda  $Bd$  do percentil  $y$ ;  $y = 2, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90$  e 95 %; e  $Bd$  = bandas Vm, Vd e Az.

A partir do vetor de características originais, foi realizado o teste de médias  $T^2$  de Hotelling, a fim de verificar se a separação entre as classes de interesse foi significativa, no nível de significância de 0,01. A separação significativa não implica, necessariamente, boa classificação.

A eficácia de um método de classificação pode ser avaliada independentemente de qualquer teste de separação. No entanto, se a separação não for significativa, a procura de uma regra de classificação seria, provavelmente, infrutífera (REGAZZI, 2000).

Após o teste de médias, foi realizado o descarte de características pelo método passo a passo ("stepwise"). Nesse método, o poder das características em discriminar as classes foi avaliado pelo critério de Wilks Lambda, utilizando-se um nível de significância de 0,15. O procedimento STEPDISC do programa computacional SAS foi empregado para essa tarefa (SAS, 1999).

As características selecionadas no descarte formaram um vetor de variáveis, assumido com distribuição normal multivariada para o desenvolvimento dos classificadores estatísticos.

### Classificadores estatísticos

Os classificadores estatísticos são derivados do teorema de Bayes. Nesse teorema, determina-se a probabilidade *a posteriori*,  $P(w_j|X)$ , de que um indivíduo  $X$  seja classificado na classe  $w_j$ . Essa probabilidade é obtida para cada uma das classes  $w_j$ , por meio do produto da probabilidade *a priori* de ocorrência da classe,  $P(w_j)$ , pela probabilidade condicional dada as características do vetor  $X$  dentro da classe  $w_j$ ,  $P(X|w_j)$ .

O indivíduo é classificado na classe que apresentar maior valor de  $P(w_j|X)$ .

O desenvolvimento dos classificadores, também chamado de fase de treinamento, pode ser realizado de duas maneiras, considerando-se uma co-variância para cada classe ou que todas as classes possuem co-variância comum (JOHNSON e WICHERN, 1999). A probabilidade  $P(w_j|X)$ , quando se considera uma co-variância para cada classe, é obtida pela função discriminante quadrática:

$$D_j(X) = -\frac{1}{2}(X - \bar{X}_j)^T (\hat{\Sigma}_j)^{-1} (X - \bar{X}_j) - \frac{1}{2} \ln |\hat{\Sigma}_j| + \ln [P(w_j)]$$

em que  $D_j(X)$  = valor da função discriminante quadrática da classe  $w_j$  dado  $X$ ;  $X$  = vetor de característica de um indivíduo;  $\bar{X}_j$  = vetor de média da classe  $j$ ;  $\hat{\Sigma}_j$  = estimativa da matriz da co-variância da classe  $j$ ;  $|\hat{\Sigma}_j|$  = determinante da estimativa da matriz da co-variância da classe  $j$ ; e  $P(w_j)$  = probabilidade *a priori* de ocorrência da classe  $j$ .

Quando se considera co-variância comum nas classes, que é obtida pela média das co-variâncias destas, a função discriminante reduz a seguinte combinação linear:

$$D_j(X) = -\frac{1}{2} \bar{X}_j^T (\hat{\Sigma}_c)^{-1} \bar{X}_j + X^T (\hat{\Sigma}_c)^{-1} \bar{X}_j + \ln P(w_j)$$

em que  $D_j(X)$  = valor da função discriminante linear da classe  $w_j$  dado  $X$ ;  $\hat{\Sigma}_c$  = matriz de co-variância comum das classes  $w_j$ 's. Assim, dos valores encontrados de  $D_1(X), D_2(X), \dots, D_k(X)$  das  $k$  classes, o maior valor dos  $D_k(X)$ 's será onde o indivíduo  $X$  terá maior probabilidade de pertencer. Os classificadores foram testados, baseando-se nas funções discriminantes quadrática e linear.

As estimativas dos parâmetros das funções discriminantes ( $\bar{X}_j, \hat{\Sigma}_j$ ) foram adquiridas dos vetores de características das amostras dos blocos de imagem. A probabilidade *a priori*  $P(w_j)$  foi estimada pela frequência de ocorrência das classes na amostra. O procedimento DISCRIM do programa SAS foi utilizado para obtenção das funções discriminantes (SAS, 1999).

Neste trabalho foram testadas funções discriminantes com as características originais e também com os escores dos componentes principais e com os escores de variáveis canônicas. A análise de componentes principais e variáveis canônicas são

técnicas que encontram combinações lineares dessas características que podem carregar grande parte da variabilidade das características originais com um número reduzido de variáveis descorrelacionadas (KHATTREE e NAIK, 2000).

Na análise de componentes principais, obtém-se outro grupo de variáveis que são ordenadas, de tal forma que as primeiras variáveis, ou primeiros componentes principais, carregam as maiores variâncias dos dados. Essas combinações são encontradas a partir da diagonalização da matriz de co-variância ou correlação. Com isso, é possível reduzir a dimensionalidade das variáveis utilizando-se somente os primeiros componentes principais. Neste trabalho, os classificadores foram desenvolvidos e testados com os escores dos componentes principais da matriz de correlação das médias das classes (JOHNSON e WICHERN, 1999). O número de componentes principais, utilizados no desenvolvimento das funções discriminantes, foi definido pela variância acumulada de no mínimo 90% dos escores das médias entre as classes.

Na análise de variáveis canônicas, obtém-se outro grupo de variáveis que são ordenadas, de tal forma que nas primeiras variáveis canônicas carrega as maiores variâncias entre classes, considerando a variância dentro das classes. Essas combinações são encontradas a partir da diagonalização de uma matriz que tem informação da variância entre e dentro das classes, formada pelo produto da matriz de soma de quadrado de tratamento e da inversa da matriz de co-variância comum. Assim, é possível reduzir a dimensionalidade das variáveis, utilizando somente as primeiras canônicas. Neste trabalho, os classificadores foram desenvolvidos e testados com os escores das variáveis canônicas (JOHNSON e WICHERN, 1999). O número de variáveis canônicas, utilizadas no desenvolvimento das funções discriminantes, foi definido pela variância acumulada de, no mínimo, 90% de seus escores.

Os erros de classificação das funções discriminantes ocorreram por validação cruzada, deixando uma amostra de fora (leave-one-out). A partir das  $n$  observações da amostra, foram realizados  $n$  treinamentos e  $n$  testes de validação. Para cada treinamento de tamanho  $(n-1)$  observações, uma observação fica de fora para teste de validação (SAS, 2003), de forma que utilizará toda a amostra para treinamento e teste das funções

discriminantes. Assim, o erro de classificação foi calculado pelo número de observações classificadas, incorretamente, nas observações de teste, em relação ao número de observações da amostra. A opção "crosvalidate" do procedimento DISCRIM executa essa análise (SAS, 1999).

### 3. RESULTADOS E DISCUSSÃO

Neste trabalho foram realizadas análises discriminantes linear e quadrática de defeitos e madeira limpa em imagens de tábuas de eucalipto. As características de percentis foram extraídas do histograma das bandas do vermelho, verde e azul de blocos de imagens com 64 x 64 pixels (resolução espacial de 15,4 x 15,4 mm) e 32 x 32 pixels (resolução espacial de 7,7 x 7,7 mm).

Nos Quadros 1 e 2 são apresentados os valores de  $F_0$  calculado para o teste  $T^2$  de Hotelling dos vetores de médias das classes de defeitos e madeira limpa estudadas, quanto a dois tamanhos de blocos.

As diferenças entre as médias das classes 1, 6 e 9 e das classes 7 e 8 não foram significativas ( $p > 0,01$ ). Segundo a norma de classificação de tábuas de eucalipto (ABNT, 2002a), os defeitos 1, 6, 9 e 10 são igualmente quantificados e pontuados para definição das classes de qualidade das tábuas, assim como os defeitos 7 e 8. Dessa maneira, podem-se agrupar esses defeitos sem proporcionar prejuízo à classificação final da tábua.

Nos Quadros 3 e 4 são apresentados os resultados do teste  $T^2$  de Hotelling dos vetores de médias após o agrupamento das classes 1, 6, 9 e 10, bem como das classes 7 e 8, para dois tamanhos de blocos. Os resultados das diferenças entre as médias dessas novas classes foram significativos ( $p < 0,01$ ). Assim, o desenvolvimento de um classificador para discriminar essas classes pode ser viável.

Os resultados do descarte de características pelo método passo a passo são apresentados no Quadro 5. Quanto às características retiradas de blocos de 64 x 64 pixels, ocorreu redução de 36 para 21 variáveis; para blocos de 32 x 32 pixels, resultou uma redução de 36 para 18 variáveis. Tais resultados indicam a dificuldade de se utilizar uma câmera com apenas uma banda espectral na classificação de madeiras, pois, mesmo após o descarte, características de todas as bandas são necessárias para discriminar as classes.

**Quadro 1** – Resultados do teste T<sup>2</sup> de Hotelling em 12 classes de defeitos e madeira limpa para blocos de imagens de 64 x 64 pixels

**Table 1** – Results of T<sup>2</sup>-Hotelling statistics for 12 wood defect classes and clear wood for 64 x 64 pixel image blocks

Classes	1	2	3	4	5	6	7	8	9	10	11	12	ml
1		5,32	5,05	5,69	3,39	<b>1,18<sup>ns</sup></b>	3,96	2,74	<b>1,18<sup>ns</sup></b>	2,61	3,72	6,76	14,69
2			7,05	5,42	6,31	5,14	4,41	4,55	5,77	3,66	6,00	7,52	6,87
3				4,58	4,25	6,44	5,04	3,95	4,96	5,46	3,67	2,33	12,06
4					5,48	5,72	2,91	2,99	5,17	3,98	4,05	4,67	5,14
5						5,94	3,25	2,90	4,15	3,97	5,04	6,61	19,69
6							4,12	2,83	<b>1,49<sup>ns</sup></b>	1,99	4,15	7,74	21,01
7								<b>0,73<sup>ns</sup></b>	3,92	3,23	4,28	5,33	9,41
8									2,91	3,01	3,94	4,46	9,89
9										2,39	3,93	6,00	16,16
10											3,98	5,45	6,99
11												3,89	5,70
12													11,12
ml													

1 = bolsa de goma; 2 = esmoado; 3 = fendilhado; 4 = furos de inseto; 5 = medula; 6 = nó cariado; 7 = nó sadio; 8 = nó sadio corrompido; 9 = nó solto encaixado; 10 = nó solto oco; 11 = podridão; 12 = racha; e ml = madeira limpa.  
Classes i de uma linha e j de uma coluna não diferem (p>0,01) se o valor de Fo(i,j) é seguido de ns.

**Quadro 2** – Resultados do teste T<sup>2</sup> de Hotelling em 12 classes de defeitos e madeira limpa para blocos de imagens de 32 x 32 pixels

**Table 2** – Results of T<sup>2</sup>-Hotelling statistics for 12 wood defect classes and clear wood for 32 x 32 pixel image blocks

Classes	1	2	3	4	5	6	7	8	9	10	11	12	ml
1		3,51	8,23	6,81	3,63	2,50	5,50	4,11	<b>1,41<sup>ns</sup></b>	3,68	2,88	7,91	12,34
2			7,62	6,61	5,36	3,28	4,62	4,82	4,18	4,26	3,60	6,34	6,97
3				5,96	6,47	10,09	7,57	4,29	7,09	7,47	3,75	3,83	12,40
4					4,58	7,89	4,00	2,12	5,07	5,70	2,43	6,02	7,55
5						5,86	2,76	1,86	2,53	4,73	2,41	6,04	11,77
6							4,96	4,26	1,83	2,36	3,31	9,34	17,47
7								<b>1,65<sup>ns</sup></b>	3,65	4,37	2,54	6,59	6,67
8									2,28	4,17	2,43	3,98	7,73
9										3,13	2,63	6,90	10,33
10											3,60	7,05	6,92
11												3,20	2,28
12													10,23
ml													

1 = bolsa de goma; 2 = esmoado; 3 = fendilhado; 4 = furos de inseto; 5 = medula; 6 = nó cariado; 7 = nó sadio; 8 = nó sadio corrompido; 9 = nó solto encaixado; 10 = nó solto oco; 11 = podridão; 12 = racha; e ml = madeira limpa.  
Classes i de uma linha e j de uma coluna não diferem (p>0,01) se o valor de Fo(i,j) é seguido de ns.

### Classificadores estatísticos

Os resultados dos erros de classificação da validação cruzada, para os blocos de imagens de 64 x 64 pixels, são apresentados nos Quadros 6 e 7, respectivamente, para as funções discriminantes lineares e quadráticas. Os escores das variáveis canônicas tenderam a apresentar menores erros de classificação, comparativamente com as variáveis originais e com os escores dos componentes principais. Observou-

se tendência de menores erros globais para as funções discriminantes quadráticas, quando foram utilizados componentes principais. Entretanto, as funções lineares com variáveis canônicas não apresentaram erro de 100% em nenhuma das classes, indicando maior poder de generalização do classificador linear sobre o quadrático. Tal problema também pode ter ocorrido devido ao número de observações relativamente pequeno de algumas classes (2, 3, 11 e 12), que tenderam a apresentar maiores erros nas funções discriminantes quadráticas.

**Quadro 3** – Resultados do teste T<sup>2</sup> de Hotelling para oito classes de defeitos e madeira limpa para blocos de imagens de 64 x 64 pixels

**Table 3** – Results of T<sup>2</sup>-Hotelling statistics for 8 wood defect classes and clear wood for 64 x 64 pixel image blocks

Classes	1/6/9/10	2	3	4	5	7/8	11	12	ml
1/6/9/10		5.91	7.05	7.83	7.25	6.97	4.31	8.71	40.30
2			6.89	5.41	6.21	4.92	5.98	7.43	6.90
3				4.47	4.22	5.20	3.69	2.32	11.84
4					5.39	3.74	4.06	4.65	5.14
5						4.11	5.08	6.60	19.47
7/8							4.39	5.74	16.35
11								3.91	5.69
12									11.02
ml									

1 = bolsa de goma; 2 = esmoado; 3 = fendilhado; 4 = furos de inseto; 5 = medula; 6 = nó cariado; 7 = nó sadio; 8 = nó sadio corrompido; 9 = nó solto encaixado; 10 = nó solto oco; 11 = podridão; 12 = racha; e ml = madeira limpa.  
Classes i de uma linha e j de uma coluna não diferem (p>0,01) se o valor de Fo(i,j) é seguido de ns.

**Quadro 4** – Resultados do teste T<sup>2</sup> de Hotelling para oito classes de defeitos e madeira limpa para blocos de imagens de 32 x 32 pixels

**Table 4** – Results of T<sup>2</sup>-Hotelling statistics for 8 wood defect classes and clear wood for 32 x 32 pixel image blocks

Classes	1/6/9/10	2	3	4	5	7/8	11	12	ml
1/6/9/10		3,92	10,56	9,45	5,98	7,75	3,15	10,09	30,11
2			7,32	6,47	5,21	4,98	3,54	6,14	6,85
3				5,89	6,29	6,59	3,70	3,85	11,89
4					4,54	3,55	2,43	6,02	7,35
5						2,52	2,43	5,96	11,57
7/8							2,55	5,94	11,38
11								3,18	2,27
12									9,91
ml									

1 = bolsa de goma; 2 = esmoado; 3 = fendilhado; 4 = furos de inseto; 5 = medula; 6 = nó cariado; 7 = nó sadio; 8 = nó sadio corrompido; 9 = nó solto encaixado; 10 = nó solto oco; 11 = podridão; 12 = racha; e ml = madeira limpa.  
Classes i de uma linha e j de uma coluna não diferem (p>0,01) se o valor de Fo(i,j) é seguido de ns.

**Quadro 5** – Características selecionadas pelo método passo a passo de blocos da imagem de 64 x 64 e 32 x 32 pixels

**Table 5** – Selected features by the stepwise method of image blocks of 64x64 and 32x32 pixels

Blocos	Características
64x64	$C_{Vm}^{2}, C_{Vm}^{10}, C_{Vm}^{30}, C_{Vm}^{60}, C_{Vm}^{70}, C_{Vm}^{90}, C_{Vd}^{5}, C_{Vd}^{10}, C_{Vd}^{40}, C_{Vd}^{60}, C_{Vd}^{70}, C_{Vd}^{95}, C_{Az}^{5}, C_{Az}^{10}, C_{Az}^{30}, C_{Az}^{40}, C_{Az}^{50}, C_{Az}^{60}, C_{Az}^{70}, C_{Az}^{90}, C_{Az}^{95}$
32x32	$C_{Vm}^{5}, C_{Vm}^{10}, C_{Vm}^{20}, C_{Vm}^{30}, C_{Vm}^{95}, C_{Vd}^{5}, C_{Vd}^{20}, C_{Vd}^{30}, C_{Vd}^{60}, C_{Vd}^{70}, C_{Vd}^{80}, C_{Az}^{2}, C_{Az}^{5}, C_{Az}^{10}, C_{Az}^{20}, C_{Az}^{40}, C_{Az}^{60}, C_{Az}^{70}$

O menor erro global encontrado foi de 19%, tanto para a função discriminante linear quanto para a quadrática com variáveis canônicas. Porém, com os problemas apresentados em algumas classes com as funções quadráticas, conclui-se que, para tamanhos

de blocos de imagens 64 x 64, as funções lineares foram as mais adequadas.

A contribuição de uma variável canônica ou de um componente principal altera relativamente pouco a variância acumulada (Quadros 6 e 7). No entanto,

o efeito no erro de algumas classes chega a cair pela metade. Assim, para discriminar alguns defeitos, foi necessária a utilização de variáveis que carregassem mais de 90% da variância acumulada.

No Quadro 8, apresenta-se a matriz de confusão da classificação, realizada por função discriminante

linear a partir de oito variáveis canônicas. Os defeitos 3 (fendilhado) e 12 (racha) foram confundidos. Esses defeitos são semelhantes, sendo o fendilhado uma racha superficial (ABNT, 2002b). Tal problema poderia ser minimizado ao acrescentar na regra de decisão a localização do defeito, pois a racha ocorre nas

**Quadro 6** – Erro de classificação, em porcentagem, de cada classe e global para função discriminante linear com blocos de imagens de 64 x 64 pixels

**Table 6** – Overall and class misclassifications, in percentage, for linear discriminant function with 64 x64 pixel image blocks

Número de Variáveis <sup>a</sup>	Classes									
	1/6/9/10	2	3	4	5	7/8	11	12	ml	Global
Originais										
21 (100,0)	21	44	56	50	30	27	60	69	6	23
Componentes Principais										
5 (100,0)	23	67	50	88	50	25	100	94	10	31
4 (99,8)	12	100	44	91	52	57	100	94	11	33
3 (99,5)	3	89	100	85	100	35	100	100	11	35
2(98,7)	2	100	100	94	100	38	100	100	11	36
1(96,5)	1	100	100	100	100	49	100	100	7	37
Variáveis Canônicas										
8 can (100)	19	22	50	41	27	25	20	44	6	19
7 can (98,9)	18	22	50	50	26	25	20	50	6	20
6 can (96,8)	17	22	56	68	26	25	20	44	5	20
5 can (93,7)	16	22	56	71	26	30	40	38	5	21

1-bolsa de goma ; 2-esmoado; 3-fendilhado; 4-furos de inseto; 5-medula; 6-nó cariado; 7-nó sadio; 8-nó sadio corrompido; 9-nó solto encaixado; 10-nó solto oco; 11-podridão; 12-racha; e ml- madeira limpa.

a - entre parêntesis é apresentada a variância acumulada.

**Quadro 7** – Erro de classificação, em porcentagem, de cada classe e global para função discriminante quadrática com blocos de imagens de 64 x 64 pixels

**Table 7** – Overall and class misclassifications, in percentage, for quadratic discriminant function with 64 x64 pixel image blocks

Número de Variáveis <sup>a</sup>	Classes									
	1/6/9/10	2	3	4	5	7/8	11	12	ml	Global
Originais										
21 (100,0)	10	100	100	82	26	24	100	100	4	25
Componentes Principais										
5 (100,0)	4	67	44	62	37	21	100	88	5	20
4 (99,8)	4	67	38	59	67	30	100	94	10	26
3 (99,5)	2	33	100	53	80	27	100	100	19	32
2(98,7)	2	89	100	82	81	21	100	100	12	31
1(96,5)	3	100	100	100	100	22	100	100	6	33
Variáveis Canônicas										
8 can (100)	13	100	56	44	30	32	100	63	3	21
7 can (98,9)	14	78	50	47	30	32	100	63	3	21
6 can (96,8)	12	67	38	59	30	29	100	63	3	20
5 can (93,7)	10	44	25	56	26	30	100	56	3	19

1-bolsa de goma ; 2-esmoado; 3-fendilhado; 4-furos de inseto; 5-medula; 6-nó cariado; 7-nó sadio; 8-nó sadio corrompido; 9-nó solto encaixado; 10-nó solto oco; 11-podridão; 12-racha; e ml- madeira limpa.

a - entre parêntesis é apresentada a variância acumulada.

extremidades das tábuas. A classificação do defeito 4 (furo de inseto) também apresentou, relativamente, erro elevado. Provavelmente, isso ocorreu devido ao fato de a resolução dos pixels da imagem não ter sido suficiente para identificação dos furos. Furos de inseto não são frequentes em toras de madeiras selecionadas para serraria.

Os resultados dos erros de classificação da validação cruzada, para os blocos de imagens de 32x32 pixels, são apresentados nos Quadros 9 e 10, respectivamente nas funções discriminantes lineares e quadráticas. Os classificadores para blocos de 32 x 32 pixels apresentaram tendências semelhantes às dos blocos de 64 x 64 pixels, porém com erros de classificação mais altos. Menores

dimensões de blocos, entretanto, possibilitam uma melhor caracterização da dimensão e formato dos defeitos.

Os erros apresentados pelas funções discriminantes lineares com os escores das variáveis canônicas foram considerados adequados para discriminar defeitos e madeira limpa em imagens digitais. Esses erros indicam um desempenho melhor do que os dos classificadores humanos estudados por Huber et al. (1985).

Os classificadores com base nos percentis apresentaram erros superiores aos de características texturais utilizadas por outros autores (KOIVO e KIM, 1989; PACKIANATHER e DRAKE, 2000). No entanto, os percentis são mais adequados para a implementação em um sistema em tempo real, pela facilidade de extração da imagem.

**Quadro 8** – Matriz de confusão para função discriminante linear com oito variáveis canônicas para blocos de imagens de 64 x 64 pixels

**Table 8** – Confusion matrix for linear discriminant function with 8 canonical variables for image blocks of 64 x 64 pixels

Classes	1/6/9/10	2	3	4	5	7/8	11	12	ml	Total	Erro
1/6/9/10	88	5	3	0	5	4	0	3	0	108	19%
2	0	7	0	0	0	1	0	0	1	9	22%
3	1	0	8	1	1	0	0	5	0	16	50%
4	0	0	0	20	1	1	0	1	11	34	41%
5	4	0	3	0	38	6	0	0	3	54	30%
7/8	7	0	1	0	4	47	0	1	3	63	25%
11	0	0	0	0	0	0	4	0	1	5	20%
12	2	0	0	1	1	1	0	9	2	16	44%
ml	1	0	0	3	1	6	0	0	176	187	6%
									492		19%

**Quadro 9** – Erro de classificação, em porcentagem, de cada classe e global para função discriminante linear com blocos de imagens de 32 x 32 pixels

**Table 9** – Overall and class misclassifications, in percentage, for linear discriminant function with 32 x 32 pixel image blocks

Número de Variáveis <sup>a</sup>	Classes									
	1/6/9/10	2	3	4	5	7/8	11	12	ml	Global
Originais										
21 (100)	20	78	44	56	41	41	80	56	9	27
Componentes Principais										
6 cp (100)	21	100	63	91	54	38	100	44	10	32
5 cp (99,9)	21	100	63	91	54	38	100	44	10	32
4 cp (99,9)	21	100	38	97	63	33	100	81	10	33
3 cp (99,5)	4	100	100	100	94	41	100	94	10	36
2 cp (99,0)	4	100	100	100	94	41	100	100	12	37
1 cp (97,9)	1	100	100	100	100	56	100	100	12	39
Variáveis Canônicas										
8 can (100)	18	67	44	53	39	38	80	50	8	25
7 can (98,3)	18	56	44	53	37	33	80	50	9	24
6 can (96,3)	20	56	38	53	35	38	100	50	8	25
5 can (93,2)	14	56	38	53	41	60	100	50	7	26

1-bolsa de goma ; 2-esmoado; 3-fendilhado; 4-furos de inseto; 5-medula; 6-nó cariado; 7-nó sadio; 8-nó sadio corrompido; 9-nó solto encaixado; 10-nó solto oco; 11-podridão; 12-racha; e ml- madeira limpa.  
a - entre parêntesis é apresentada a variância acumulada.

**Quadro 10** – Erro de classificação, em porcentagem, de cada classe e global para função discriminante quadrática com blocos de imagens de 32 x 32 pixels

**Table 10** – Overall and class misclassifications, in percentage, for quadratic discriminant function with 32 x 32 pixel image blocks

Número de Variáveis <sup>a</sup>	Classes									
	1/6/9/10	2	3	4	5	7/8	11	12	ml	Global
Originais										
21 (100)	18	100	100	56	52	37	100	100	5	29
Componentes Principais										
5 cp (100)	9	89	44	56	65	46	100	69	7	28
4 cp (99,8)	9	78	31	59	63	44	100	81	11	29
3 cp (99,5)	7	56	25	88	63	62	100	88	13	33
2 cp (98,7)	7	67	38	97	65	63	100	100	17	37
1 cp (96,5)	6	89	81	97	67	56	100	100	10	35
Variáveis Canônicas										
8 can (100)	14	100	50	59	44	41	100	56	5	25
7 can (98,3)	13	78	38	47	43	44	100	56	5	24
6 can (96,3)	15	67	31	50	50	40	100	50	5	24
5 can (93,2)	13	44	25	53	52	49	100	44	4	24

1-bolsa de goma; 2-esmoado; 3-fendilhado; 4-furos de inseto; 5-medula; 6-nó cariado; 7-nó sadio; 8-nó sadio corrompido; 9-nó solto encaixado; 10-nó solto oco; 11-podridão; 12-racha; e ml- madeira limpa.

a - entre parêntesis é apresentada a variância acumulada.

#### 4. CONCLUSÕES

As funções discriminantes com os escores das variáveis canônicas tenderam a apresentar menores erros de classificação do que com as variáveis originais e com os escores dos componentes principais.

As funções discriminantes quadráticas tenderam a exibir erros de classificação global menores do que as funções lineares, porém a função linear apresentou melhor distribuição de erros entre as classes de defeitos.

Características retiradas de blocos de 64 x 64 pixels tenderam a mostrar menores percentuais de erros de classificação do que em blocos de 32 x 32 pixels.

Os menores erros globais de classificação foram 19 e 24% para funções discriminantes lineares com variáveis canônicas para tamanho de bloco de 64 x 64 e 32 x 32 pixels, respectivamente.

As características de percentis foram consideradas adequadas para discriminar defeitos e madeira limpa em imagens digitais, podendo ser empregadas em Sistemas de Visão Artificial.

Apesar de a magnitude do erro ter sido considerada satisfatória, outros tipos de classificadores e a inclusão de outros parâmetros poderiam ser considerados, em trabalhos semelhantes, como classificadores não-paramétricos e posição do defeito, respectivamente.

#### 5. AGRADECIMENTOS

Este trabalho teve apoio financeiro da Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG). Os autores receberam bolsa de estudo e pesquisa da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

#### 6. REFERÊNCIAS BIBLIOGRÁFICAS

- ARACRUZ PRODUTOS DE MADEIRA S.A. **Descrição do produto**. Nova Viçosa: 2000. 9p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – **ABNT**. Madeira serrada de eucalipto – Requisitos, NBR 14806. Rio de Janeiro: 2002a. 11 p.
- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS – **ABNT**. Madeira serrada – Terminologia, NBR 12551. Rio de Janeiro, 2002b. 10p.
- HANEY, L. et al. Color matching of wood with a real-time machine vision system. **ASAE Paper Nº. 943579**, ASAE MEETING PRESENTATION, St. Joseph: 1994, 15 p.

- HUBER, H. A.; McMILLIN, C. W.; McKINNEY, J. P. Lumber defect detection abilities of furniture rough mill employees. **Forest Products Journal**, v.35. n.11/12, p.79-82, 1985.
- JOHNSON, R. A.; WICHERN, D. W.; **Applied Multivariate statistical analysis**, 4. ed., Rio de Janeiro: Prentice-Hall, 1999. 816 p.
- KAUPPINEN, H. **Development of a color machine vision method for wood surface inspection**. Oulu: Department of Electrical Engineering and Infotech Oulu, University of Oulu, Dissertation, 1999. 138p. <http://herkules.oulu.fi/isbn9514254244/> 11 nov. 2003.
- KHATTREE, R.; NAIK, D. N. **Multivariate data reduction and discrimination with SAS<sup>R</sup> software**. Cary: SAS Institute, 2000. 558p.
- KLINE, D. E. et al. Performance of color camera machine vision in automated furniture rough mill systems, **Forest Products Journal**, v.48, n.3, p. 38-45, 1998.
- KLINE, D. E. et al. **Lumber scanning system for surface defect detection**. St. Joseph: ASAE, 1992. (ASAE Paper, 923582).
- KLINE, D. E.; ARAMAN, P. A.; SURAK, C., Evaluation of an automated hardwood lumber grading system. In: INTERNATIONAL CONFERENCE ON SCANNING TECHNOLOGY AND PROCESS OPTIMIZATION FOR THE WOOD INDUSTRY, 2001, Seattle. **Proceedings...** Berkeley: Wood Machining Institute, 2001. p.141-151.
- KOIVO, A. J.; KIM, C. W. Automatic classification of surface defects on red oak boards. **Forest Products Journal**, v.39, n.9, p. 22-30, 1989.
- LEBOW, P. K. et al. Classification of wood surface features by spectral reflectance, **Wood and Fiber Science**, v.28, n.1, p. 74-90, 1996.
- LU, Q. **A real-time system for color sorting edge-glued panel parts**. 1997. 119f. Dissertation (M.S. in Electrical Engineering), Faculty of the Virginia Polytechnic Institute and State University, Blacksburg, 1997.
- PACKIANATHER, M. S.; DRAKE, P. R., Neural networks for classifying images of wood veneer part 2. **International Journal of advanced manufacturing Technology**, v.16, n.6, p.424-433, 2000.
- PHAM, D.T.; ALCOCK, R.J. Automated grading and defect detection: a review. **Forest Products Journal**, v.48, n.4, p34-42, 1998.
- REGAZZI, A. J. **Análise multivariada**, notas de aula INF 766, Viçosa, MG: Universidade Federal de Viçosa, 2000. v.2.
- SZYMANI, R.; MCDONALD, K. A. Defect detection in lumber: state of the art. **Forest Products Journal**, v.31, n.11, p. 34-44, 1981.
- SAS INSTITUTE. **SAS OnlineDoc, Version 8**. Cary: 1999.
- SAS/STAT. **The Discrim Procedure Overview**. <http://www.id.unizh.ch/software/unix/statmath/sas/sasdoc/stat/chap25/index.htm>. 30 nov. 2003.