



VII Workshop de Visão Computacional

Anais do Evento

www.wvc2011.ufpr.br

22 a 25 de maio de 2011 - UFPR - Curitiba

ISSN - 2175-6120

WVC 2011

Anais do VII Workshop de Visão Computacional

Universidade Federal do Paraná

ISSN 2175-6120

Editores:

Luiz Antônio Pereira Neves

Lucas Ferrari de Oliveira

Hugo Vieira Neto

Curitiba, 22 a 25 de Maio de 2011

Sistema de Visão Computacional para Classificação de Grãos de Café por Cor e Forma

Pedro Ivo de Castro Oyama¹; Lúcio André de Castro Jorge²; Carlos Cesar Gomes³

- 1- Bolsista, Embrapa Instrumentação, São Carlos – SP, pedro.oyama@gmail.com
 2- Pesquisador, Embrapa Instrumentação, São Carlos – SP, lucio@cnpdia.embrapa.br
 3- Analista de qualidade, Cooperativa de Café Cooxupé, Guaxupé – MG, cesar@cooxupe.com.br

Abstract

The present work aims to develop a vision system for coffee beans classification of small samples to be used in the lab of coffee quality analysis of the COOXUPÉ, cooperative of coffee producers from Guaxupé, Brasil, in order to replace the current visual inspection made by human specialist, which is slow and is often unable to fulfill the demands of the coffee industry. Moreover, for being made by humans, these methods are more susceptible to failure. The system is a machine vision that uses image processing algorithms to extract color and shape beans attributes from the image samples. An artificial neural network Multilayer Perceptron was used to recognize patterns of color and shape from images of beans. It was possible to accurately classify coffee beans but an improvement in the illumination system is necessary.

1. Introdução

Com o aumento da produção e da demanda pelo produto de qualidade, cada vez mais o número de análises rápidas e precisas dos atributos físicos e fisiológicos dos grãos tem se tornado o gargalo na tomada de decisões rápidas em tempo compatível com o processo produtivo [1].

A análise dos atributos físicos de um lote de grãos é baseada principalmente em caracteres morfológicos das frações componentes da amostra. Dentre os atributos físicos do lote destacam-se as dimensões, a forma, a presença de impurezas tais como restos vegetais, pedras, partículas de solo, frações de grãos menores que sua metade e etc. [5]. A análise física em grãos é um procedimento moroso e dependente da interpretação do analista, pois hoje é feita manualmente.

Os atributos fisiológicos de um lote de sementes devem indicar a sua capacidade de germinar e o seu vigor. No caso do grão, servem para identificar principalmente defeitos baseados na cor da amostra.

Vários testes para avaliação dos atributos físicos e fisiológicos de um lote de sementes são relatados e descritos em [2], [3] e [4] e podem ser aplicados aos grãos.

Com o avanço na área de visão computacional, com a criação de sensores mais rápidos e processamentos mais eficientes, podem ser encontrados trabalhos de classificação de sementes de soja com base na forma [6] e a análise de cor para a separação de grãos em máquinas utilizadas pela indústria alimentícia [7], dentre outros.

Em geral, as aplicações aparecem em máquinas e sistemas para aplicação no processo de produção de produtoras de sementes. No laboratório, a análise de amostras ainda continua sendo feita de forma subjetiva e manual.

Nesse contexto, está sendo desenvolvida uma máquina para a classificação automática de grãos de café através de imagens digitais e neste trabalho serão apresentados os primeiros resultados da classificação feita através da aplicação de algoritmos de processamento de imagem numa amostra de grãos de café com base nos atributos de cor e forma.

2. Metodologia Utilizada

A metodologia utilizada neste trabalho é dividida em: aquisição de imagens da amostra; aplicação de algoritmos de processamento de imagem para detecção dos objetos e extração de suas características; classificação por padrões de cor; definição de descritores de forma; e, finalmente, classificação por padrões de forma.

2.1 Amostragem e Classificação

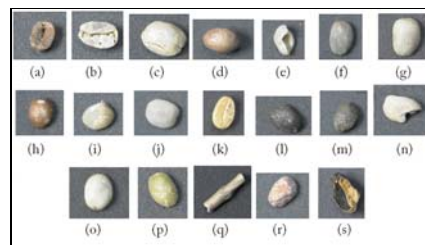


Figura 1. Classes de grãos e impurezas. (a) ardido, (b) brocado, (c) cabeça, (d) côco, (e) concha, (f) fava escura, (g) marago, (h) melado, (i) película, (j) perfeito, (k) pergaminho, (l) preto, (m) preto-verde, (n) quebrado, (o) secador, (p) verde, (q) paus, (r) pedras e (s) cascas.

Foram utilizadas amostras de grãos de diferentes peneiras fornecidas pela Cooperativa Regional de Cafeicultores de Guaxupé Ltda. (Cooxupé), juntamente com os grãos com defeitos e impurezas já separados e classificados por especialistas da cooperativa.

As classes de defeitos e de grãos sadios são apresentadas na Figura 1. As impurezas de uma amostra são paus, pedras e cascas.

2.2 Aquisição de imagens

Para a aquisição de imagens das amostras de grãos de café com iluminação padrão, foi desenvolvida uma máquina de classificação (Figura 2), composta de uma bandeja de vidro – onde são depositados os grãos sob condições controladas de luz. A bandeja é iluminada por duas lâmpadas fluorescentes de 20W por cima e duas por baixo, duas câmeras de vídeo posicionadas cerca de 60cm acima e abaixo da amostra. A câmera de vídeo utilizada é do tipo IP e é conectada ao computador que executa o software através de uma interface *Gigabit Ethernet*. As imagens são capturadas diretamente pelo software e são salvas no formato JPG com resolução de 1280x1024 pixels.

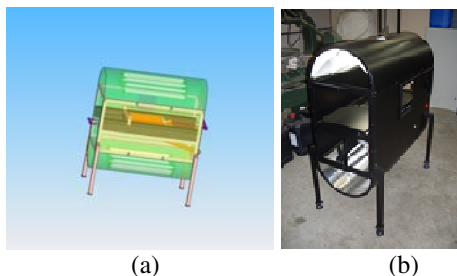


Figura 2: (a) Representação gráfica da máquina de classificar café e (b) Foto da primeira versão construída

Na Figura 3 pode ser observada a tela de captura da imagem, onde são possíveis ajustes de brilho, contraste e iluminação da amostra.

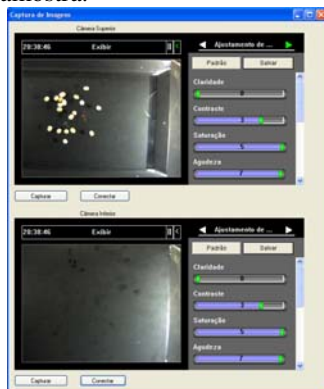


Figura 3: Interface de controle de aquisição

2.3 Processamento das Imagens

O software Qualicafé desenvolvido foi dividido em módulos: pré-processamento, classificação por cor, por forma e por tipo de grão, conforme descrito no diagrama da Figura 4.

Todos os processos de classificação se baseiam numa rede neural do tipo *Multilayer Perceptron* (MLP). Foi utilizada a API Weka, com algoritmos de aprendizado de máquina implementados em Java e distribuído gratuitamente sob a licença GNU [8].

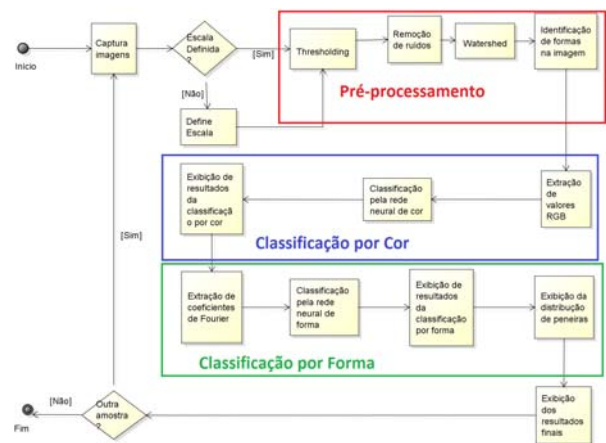


Figura 4: Diagrama de blocos do Qualicafé

As duas MLPs presentes neste trabalho utilizam o algoritmo de *backpropagation*, com taxa de aprendizado de 0,3, momento de 0,2, limiar de validação de 20 e 1000 épocas como condição de parada. Os números de neurônios das camadas de entrada e de saída são iguais, respectivamente, ao número de atributos e ao número de classes definidos e ambas MLPs apresentam apenas uma camada escondida, com n neurônios cada, sendo $n = (\text{número de atributos} + \text{número de classes}) / 2$.

2.4 Pré-processamento

A primeira etapa de processamento é a identificação de grãos e impurezas presentes na amostra. Esta identificação é feita através do contorno de cada objeto encontrado na imagem. Para a identificação dos contornos, primeiramente é feita a segmentação da imagem pela aplicação de um simples *Thresholding*, algoritmo que classifica cada pixel como fundo ou objeto, através da comparação do valor do pixel em escala de cinza com um valor adotado como limiar - o *threshold*. Com a separação do fundo dos objetos, a imagem é então binarizada. Utiliza-se os processos morfológicos de erosão e dilatação, a fim de se eliminar possíveis ruídos. Visando separar grãos que estejam encostados um no outro,

podendo gerar problemas na contagem de grãos, é executado o algoritmo de *watershed*. Para cada *pixel* ‘objeto’ da imagem é calculada a sua distância ao *pixel* ‘fundo’ mais próximo. Então são definidos os pontos com as distâncias máximas locais e cada um desses pontos é dilatado até que se chegue à borda do objeto ou à borda da região de outro ponto sendo dilatado. Dessa forma, cada uma das regiões provenientes de um ponto dilatado define um objeto diferente. A partir da imagem binarizada é finalmente executado um algoritmo para rotular as diferentes formas e se extrair os *pixels* de seus contornos.

Na Figura 5 pode ser visualizada uma imagem de saída do processo de identificação de grãos. Nesta imagem, os contornos dos grãos são pintados de verde e numerados sequencialmente. A rotulação é armazenada juntamente com os *pixels* localizados nos seu interior.

Para se trabalhar com as dimensões reais dos objetos durante a análise por forma, foi desenvolvida uma ferramenta para a indicação da escala da imagem, através da qual se desenha uma linha sobre a imagem e informa-se a distância real que ela representa.

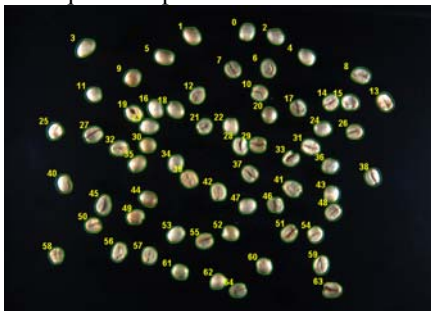


Figura 5. Objetos identificados e rotulados.

2.5 Classificação por cor

A análise de cor dos grãos de café é feita para se distinguir grãos bons de grãos com certos tipos de defeito, como: grãos verdes, pretos, pretos-verdes, ardidos, etc. Essa análise se baseia no reconhecimento de padrões de cor presentes nos grãos, visto que certos defeitos são identificados por uma cor característica do grão, ou de uma parte dele, como por exemplo o tom mais esverdeado que é comum nos grãos-verdes e manchas marrons nos grãos melados. Da mesma forma, as sementes perfeitas também têm uma coloração característica que as definem.

Sendo assim, cada um dos defeitos passíveis de serem identificados pela análise de cor foi associado a uma classe, que juntamente com a classe de grão perfeito formam o conjunto de padrões a serem reconhecidos pela rede neural. As classes são: ardido, brocado, casca, côco, fava escura, melado, pau, pedra, película, perfeito, pergaminho, preto, preto-verde, secador e verde (todas ilustradas na Figura 1). A MLP foi definida com três atributos de entrada: as componentes de cor R, G e B.

Para cada uma das classes, foram capturadas imagens de amostras contendo grãos da classe em questão para execução do treinamento da rede. Foram utilizadas janelas selecionadas manualmente sobre a imagem, cujas áreas apresentavam a coloração característica do defeito do grão (Figura 6), ou um padrão de saudável. Para cada um dos *pixels* que compõem essas áreas foi determinado o valor médio de R, G e B, com base nos *pixels* vizinhos. Esses valores foram então utilizados como entrada para a MLP no seu treinamento. Para cada classe foram coletados cerca de 8.000 *pixels*.

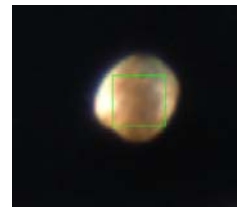


Figura 6. Seleção de amostra em um grão para treinamento da MLP.

A classificação dos grãos de uma amostra é feita se passando como entrada para a MLP treinada os valores médios de R, G e B de cada um dos *pixels* do grão com seus vizinhos. Conforme os *pixels* são classificados eles são pintados na imagem da amostra com uma cor diferente para cada classe, de modo que os padrões de cor que constituem o grão sejam facilmente identificados, como exibido na Figura 7. Quando todos os *pixels* de um grão tiverem sido classificados, a classe atribuída ao maior número de *pixels* é definida como a classe do grão.

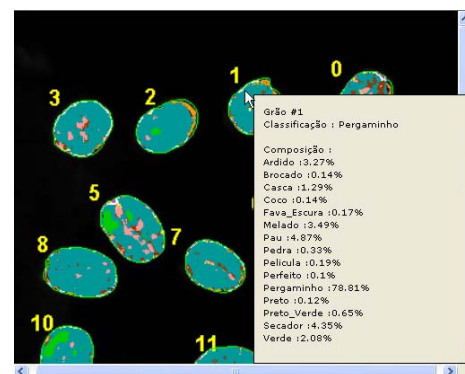


Figura 7. Resultado de uma classificação por cor.

2.6 Descritores de forma utilizados

Para que a MLP possa realizar o reconhecimento de padrões de forma, foram utilizados alguns descritores de forma. Os descritores utilizados foram:

Assinaturas: a distância de um ponto da borda até o seu centro de massa. Os valores foram normalizados pelo maior valor, de forma que cada um variasse entre 0 e 1.

Foram utilizados 36 valores, separados entre si de 10° . Juntamente com as assinaturas também foram utilizados o comprimento e a largura da forma, em milímetros.

Descritores de Fourier: Descritores de Fourier representam a borda de uma região como uma função periódica que pode ser expandida em uma série de Fourier. Uma característica típica dos descritores de Fourier é que geralmente a forma geral da figura é relativamente bem definida a partir de alguns dos termos de menor ordem da expansão. Para este projeto, optou-se por usar os 40 primeiros coeficientes para se definir as formas.

Na Figura 8 podem ser observados exemplos de assinatura e descritores de Fourier para cada tipo de contorno estudado.

A comparação dos modelos consistiu no treinamento de duas MLPs para o reconhecimento de grãos perfeitos, grãos quebrados, paus e pedras, cada uma utilizando-se de um dos modelos. Então as MLPs foram utilizadas na classificação de amostras com grãos conhecidos, para que finalmente pudessem ser calculadas suas taxas de acerto.

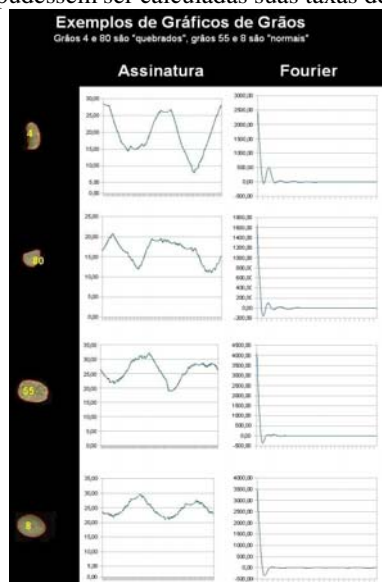


Figura 8: Assinatura e Descritor de Fourier

A MLP treinada com assinaturas, largura e comprimento teve uma taxa de acerto de cerca de 80%, enquanto que para a treinada com descritores de Fourier, a taxa foi de cerca de 60%. Portanto o primeiro modelo foi o adotado.

2.7 Classificação por forma

As impurezas de uma amostra de grãos de café podem ser facilmente reconhecidas através de suas formas, que estão fora do padrão que as sementes apresentam. Como exemplo, pode-se citar os paus, que apresentam um contorno fino e comprido, ou as pedras, que são pequenas e têm um contorno irregular. A análise de forma também é

capaz de indicar grãos com defeitos, como no caso de grãos quebrados, que naturalmente são menores e não possuem o característico formato arredondado de uma semente perfeita. Seguindo essas premissas foi utilizada uma MLP para reconhecer os padrões de forma dos objetos presentes na imagem de uma amostra e indicar de qual classe de grão ou impureza esses padrões são característicos.

Essa MLP foi configurada com 38 atributos, dos quais 36 são assinaturas, uma é a largura e a outra o comprimento (ambos em milímetros) do grão a ser classificado, e 10 classes: cabeça, casca, côco, concha, marago, pau, pedra, perfeito, pergaminho e quebrado (apresentadas da Figura 1). Analogamente ao treinamento da MLP de cor, o da de forma foi feito capturando-se imagens contendo grãos ou impurezas de uma classe específica e então extraindo seus contornos para serem utilizados no cálculo das assinaturas, dos comprimentos e das larguras. Foi coletado um conjunto de cerca de 60 formas para cada uma das classes.

Além de indicar os padrões de forma presentes, também é gerado um histograma com a distribuição dos grãos por peneira, uma classificação utilizada pelas cooperativas de café. Em geral esta classificação é feita por um jogo de peneiras, que separa os grãos pela forma e pelo tamanho. As peneiras têm crivos com diversas medidas e dois formatos diferentes: podem ser oblongos, para separar os cafés mocas, ou circulares, para separar os cafés chatos. As medidas dos crivos das peneiras são dadas em frações de $1/64$ de polegada e o número da peneira corresponde ao numerador da fração. Por exemplo: peneira 19 = $19/64$ de polegada [9].

3. Resultados e Discussão

Com as MLPs devidamente treinadas foram utilizados conjuntos de grãos fornecidos pela Cooxupé em diversos testes com o intuito de se validar os métodos desenvolvidos.

3.1 Identificação dos objetos

Constatou-se que o processo de reconhecimento de formas na imagem apresenta algumas falhas, como ilustrado na Figura 9. A mais perceptível acontece na presença de objetos com coloração muito escura, visto que uma porção deles é rotulada como parte do fundo no processo de segmentação, resultando na extração de formas diferentes das reais. Outro problema ocorre na execução do algoritmo de *watershed*, que algumas vezes divide a forma de um único objeto em várias (o que ocorre com frequência com os paus) e em certas ocasiões traça erroneamente a fronteira entre dois objetos que se encostam. Essas falhas se devem principalmente ao

sistema de iluminação da amostra, que não distribui a luz homogeneamente e ao fato de a superfície sobre a qual a amostra é depositada ter uma tonalidade de cor próxima a de certos objetos de interesse.

Apesar das falhas, de modo geral o algoritmo de reconhecimento de objetos se mostrou muito eficiente. Posicionando-se a amostra na faixa da bandeja onde a iluminação é mais homogênea e são gerados menos reflexos, são detectados praticamente todos os objetos (com exceção das formas muito escuras - defeito que pode ser corrigido alterando-se a cor do fundo usado).

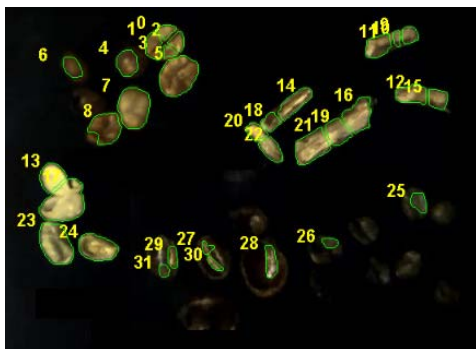


Figura 9. Deficiências na identificação de objetos.

3.2 Classificação por cor

Analisando-se visualmente as imagens capturadas das amostras, verificou-se que as cores apresentavam significativa diferença em comparação àquelas observadas nos objetos reais, tornando mais difícil o reconhecimento de alguns dos padrões que definem algumas das classes das sementes. Somando-se a isso as limitações impostas pelo sistema de iluminação, que dificulta a extração satisfatória de contornos de objetos muito escuros, estão sendo alterados os processos de iluminação construídos no primeiro protótipo.

Para teste apenas, usando condição não-ideal de iluminação, foi classificada uma amostra contendo 15 grãos pergaminho, 15 perfeitos e 15 pedras e calculadas as taxas de acerto para cada uma das classes (Tabela 1).

Tabela 1. Taxa de acerto da classificação por cor.

Classe	Pergaminho	Pedra	Perfeito
Taxa de acerto	100%	80%	27%

Apesar da taxa de acerto para os grãos perfeitos ter sido baixa, do total de sementes perfeitas classificadas, 40% foram classificadas como “secador”, grão visivelmente muito semelhante ao “perfeito”, o que indica que houve certo êxito no reconhecimento do padrão, porém este não foi tão preciso quanto o necessário.

Sendo assim, o sistema de iluminação ainda está sendo modificado para se melhorar o processo de análise.

3.3 Classificação por forma

A extração de contornos de objetos escuros é prejudicada pela iluminação presente. Portanto, as classes de forma que apresentam uma coloração escura (cascas e côcos) não foram usados nos testes para validar o método de classificação por forma, além de serem menos presentes nas amostras. Foram utilizados 15 grãos quebrados, 15 pergaminhos e 15 maragos. As taxas de acerto resultantes estão na Tabela 2.

Tabela 2. Taxa de acerto da classificação por forma.

Classe	Pergaminho	Quebrado	Marago
Taxa de acerto	46,67%	60,00%	40%

3.4 Distribuição de peneiras

Foram conduzidos testes para a obtenção da distribuição por peneira em três amostras de grãos, cada uma proveniente de um processo de separação por peneira com crivos diferentes. As peneiras em questão eram 15, 17 e 19. Nos três casos pôde-se perceber uma maior concentração em torno da furação utilizada na separação, a qual também apresentou em todos os casos o maior número de ocorrências. A figura 10 mostra o histograma obtido da amostra da peneira 19.

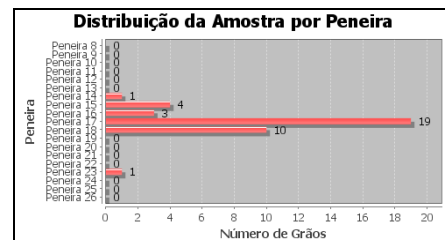


Figura 10. Distribuição por peneira de uma amostra de peneira 17.

3.5 Distinção entre grãos e impurezas

Comparando-se as classificações geradas pelo *software* com as reais classes presentes nas amostras, observou-se que foi capaz de distinguir entre impurezas – paus e pedras – e grãos, sejam eles com defeitos ou não, com uma acurácia satisfatória. Dado que a presença de impurezas na amostra exerce uma depreciação no café muito maior que os grãos defeituosos, essa característica indica um importante resultado. Na figura 11 é mostrado um exemplo de classificação, na qual essa distinção é notada. Em verde estão os objetos corretamente classificados (como paus, pedras ou grãos) em pelo menos umas das duas análises, de cor e de forma, e em vermelho

aqueles cujas classificações não foram corretas em nenhuma delas.

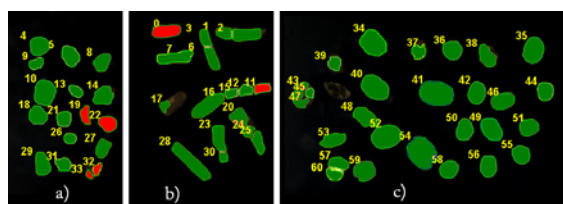


Figura 11. Acertos de classificação de grãos e impurezas a) pedras b) paus c) grãos.

A interface implementada no Qualicafé mostra a classificação do grão, tomado por cima e por baixo, como pode ser observado na Figura 12.

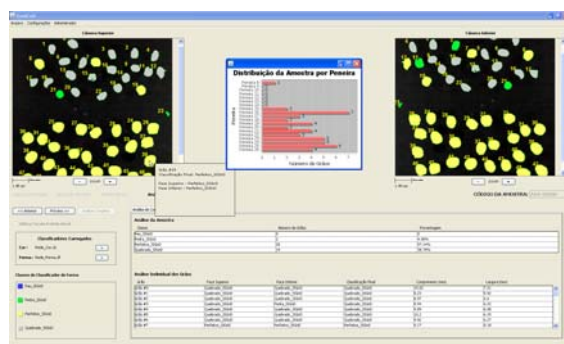


Figura 12: Tela de Resultados por análise

4. Conclusões e trabalhos futuros

Este trabalho buscou desenvolver um software que substituísse o atual processo manual de classificação de grãos de café por um método automático, mais eficiente e confiável. Os resultados obtidos mostraram que as atuais técnicas de processamento de imagem são capazes de extrair da imagem de um grão informações suficientes para definir sua forma e coloração, alguns dos atributos utilizados por especialistas humanos para classificar os grãos. Também se pode concluir que com as classificações geradas pelas redes neurais é possível fazer o reconhecimento de padrões de forma satisfatória.

Mesmo com as dificuldades geradas pela iluminação não-adequada, que acarreta grandes limitações em ambos os processo classificatórios, foi possível se obter bons resultados preliminares – distinguir grãos de impurezas e

identificar certas classes com um acerto significativo – sugerindo que uma melhora no sistema de iluminação ampliará o substancialmente o potencial do método.

5. Agradecimentos

Agradecemos ao apoio financeiro da Cooxupé, Guaxupé, MG, aos alunos Gabriel Marcondes, Daniel Bueno e Alex Rodrigues pelas contribuições na fase inicial do projeto

6. Referências

- [1] BEWLEY, J.D.; BLACK, M. Seeds: physiology of development and germination. 2º ed. New York: Plenum Press, p.199-267. 1994.
- [2] KRZYZANOWSKI, C.F. Relato de testes de vigor disponíveis para as grandes culturas. Informativo ABRATES. V.1, n.2, p.59-61. 1994.
- [3] MARCOS FILHO, J. Utilização de testes de vigor em programas de controle de qualidade de sementes. Informativo ABRATES. v.4, n.2, p.59-61. 1995.
- [4] VIEIRA, D.R.; CARVALHO, N.M. Testes de vigor em sementes. Jaboticabal, FUNEP. 164 p. 1994.
- [5] BRASIL. Ministério da Agricultura e Reforma Agrária. Regras para análise de sementes. Brasília, 1992. 365p.
- [6] KHATCHATOURIAN, O.; PADILHA, F.R.R. Reconhecimento de variedades de soja por meio do processamento de imagens digitais usando redes neurais artificiais Eng. Agríc., Jaboticabal, v.28, n.4, p.759-769, out./dez. 2008.
- [7] MACDOUGALL, D.B. Colour in food: Improving quality. Cambridge: Woodhead publishing, 2002. pp.115-142.
- [8] HALL, M.; EIBE, F.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1. 2009.
- [9] THOMAZIELLO, R.A. Classificação do café. Disponível em: <http://www.megaagro.com.br/cafe/art_class_cafe.asp>. Acesso em: 17 fev. 2011.