

UFSM – PPGE

Disciplina: Reconhecimento de Padrões

Professor: José Antônio T. B. da Costa

Aluna: Gabriela Carla Bauermann

Trabalho 2

Criar um conjunto de amostras de treinamento, extrair características e classificar usando a média de classe mais próxima.

Introdução

A imagem na Figura 1 foi usada para a extração de amostras e posterior classificação.

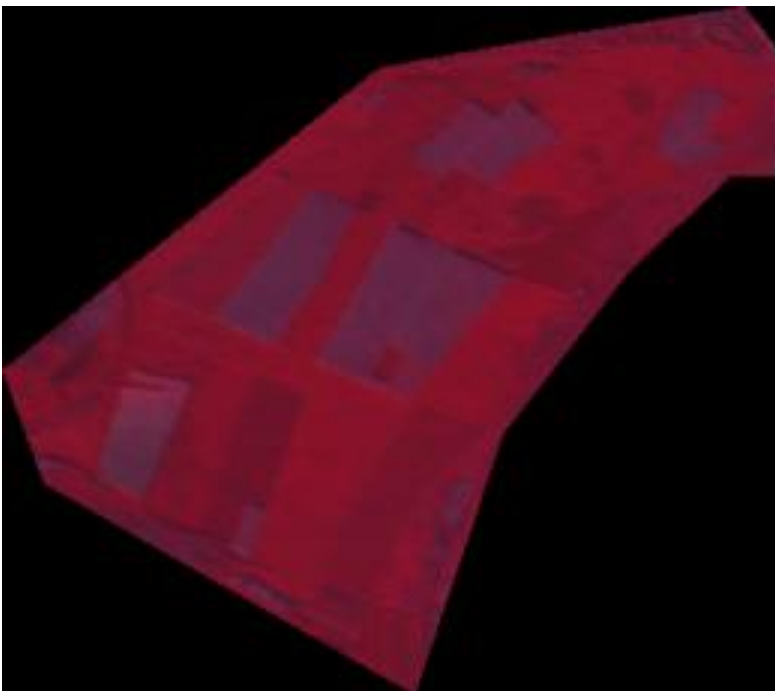


Figura 1: Fragmento de uma imagem CBERS/CCD, na composição 432, considerada adequada para a visualização de florestas. A área é parte de uma fazenda produtora de eucaliptos. A parte não pertencente à área de interesse foi preenchida com o valor zero.

Objetivos

Criar um conjunto de treinamento e utilizá-lo para classificar uma imagem com um classificador por média euclidiana mais próxima.

Avaliar o método proposto para a aplicação em imagens de florestas obtidas do satélite CBERS.

Metodologia

Todo o processamento de dados foi realizado utilizando Matlab 7.0.

Classes

Foram definidas cinco classes (sendo uma correspondente ao rejeito) :

- Solo exposto (Se) – que contém solo e florestas muito jovens;
- Florestas claras (Fc) – as áreas de vermelho intenso, na imagem;
- Florestas escuras (Fe) – as áreas mais escuras;
- Florestas naturais (Fn);
- Rejeito (Nc).

Amostras

Foram coletadas oito amostras retangulares de cada classe. As características estatísticas de cada conjunto de amostras estão agrupadas na Tabela 1.

Tabela 1: Características estatísticas

	n	Média			Desvio-padrão			Amplitude		
		IR	R	G	IR	R	G	IR	R	G
Se	2187	104.34	37.447	68.219	5.7948	2.8807	2.3007	55	20	20
Fc	3049	134.25	18.045	41.182	3.2929	0.65858	0.78076	32	7	9
Fe	2071	114.07	18.835	41.87	3.3646	0.56722	0.94127	29	6	9
Fn	920	112.4	21.127	46.935	6.0498	1.8924	2.3945	129	34	69

* n = número total de pixels.

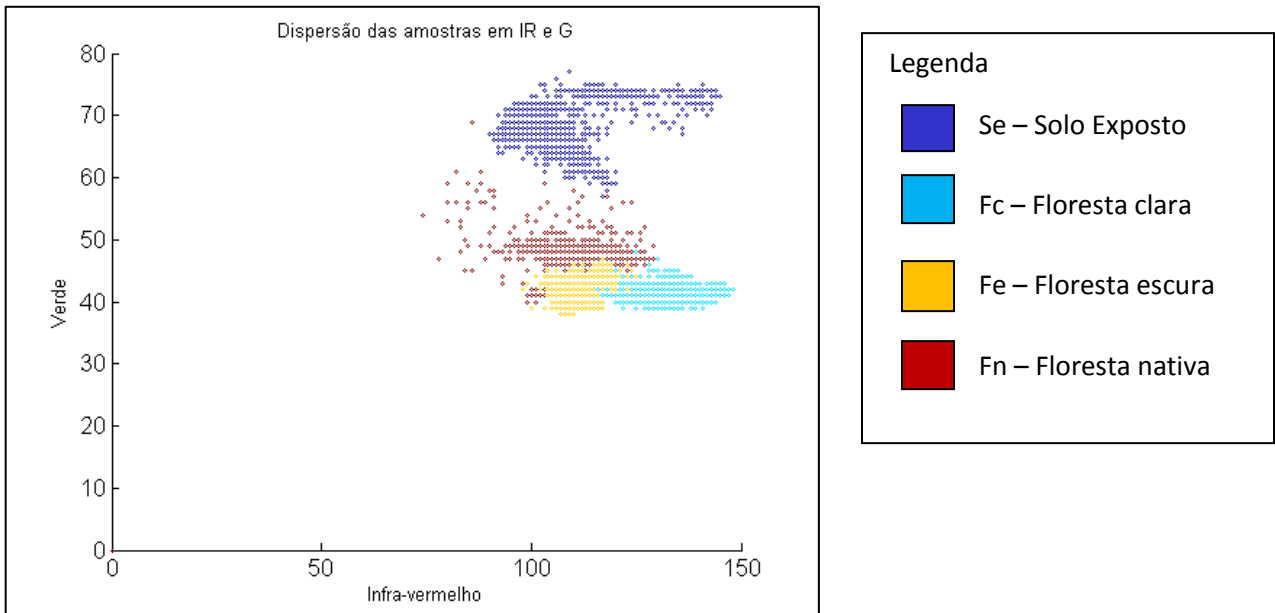


Figura 2: Gráfico de dispersão das amostras em função do valor dos pixels nos canais de Infra-vermelho e verde.

As três combinações possíveis entre os canais de cores foram testadas, e a que melhor separa as amostras (apesar de certa sobreposição entre as florestas nativa, escura e clara) foi a do verde com o infra-vermelho (Figura 2).

Implementação

O classificador foi implementado no Matlab usando o seguinte código:

```
function C = NearestClassMean (A,clas,x,y);

% Classificador por média de classe mais próxima,
% para duas variáveis, sendo estas dois planos de cores de uma imagem.
% A é uma imagem RGB,
% class é a matriz formada pelos vetores com as características das classes
% (construídos por 'stat_tb2.m'), colocados em linhas um abaixo do outro.
% x e y são os dois planos (bandas) escolhidos para a comparação
% C será uma imagem-label
% C = zeros([size(A,1) size(A,2)]);

A = double(A);

% calcula distancias maximas para cada classe
for w = 1:size(clas,1),
    % usando amplitude
    dm(w) = clas(w,(x+7))*2/3
    % usando dev. padrao
    % dm(w) = clas(w,(x+5))*3;
end

% desativa o aviso de divisao por zero
warning off MATLAB:divideByZero

for i = 1: size(A,1),
    for j = 1: size(A,2),
        if (A(i,j,x) == 0) & (A(i,j,y) == 0);
            % continua zero
        else
            % calcula distancia
            d = distancias (A(i,j,:), clas, x, y);
            % testa distancias máximas
            comp = d <= dm;
            d = d./comp;
            % decide label
            if all(d==inf),
                L = 5;
            else
                L = find(d == min(d));
            end
            % transfere label
            C(i,j)=L;
        end
    end
end
end
end
% reativa o aviso de divisao por zero
warning on MATLAB:divideByZero
```

```

function d = distancias (p, clas, x, y)

for w = 1:size(clas,1)
    d(w) = sqrt(double(((p(x)-clas(w,(x+1)))/clas(w,(x+4)))^2 + ((p(y)-clas(w,(y+1)))/clas(w,(y+4)))^2));
end

-----

function x = stat_tb2 (S);

% roteiro de cálculos das características estatísticas do conjunto de
% amostras no segundo trabalho de rec. de padroes

% formato do vetor de características x:
% [n media_IR media_R media_G sd_IR sd_R sd_G amplitude_IR amplitude_R amplitude_G]
%

x = zeros(1,10);

% 1. cálculo do número de pixels
x(1)=size(S,2);

% médias:
x(2) = mean(S(:,1));
x(3) = mean(S(:,2));
x(4) = mean(S(:,3));

% desvio-padrão
x(5) = std(S(:,1));
x(6) = std(S(:,2));
x(7) = std(S(:,3));

% amplitude
x(8) = max(S(:,1))-min(S(:,1));
x(9) = max(S(:,2))-min(S(:,2));
x(10) = max(S(:,3))-min(S(:,3));

```

Resultados

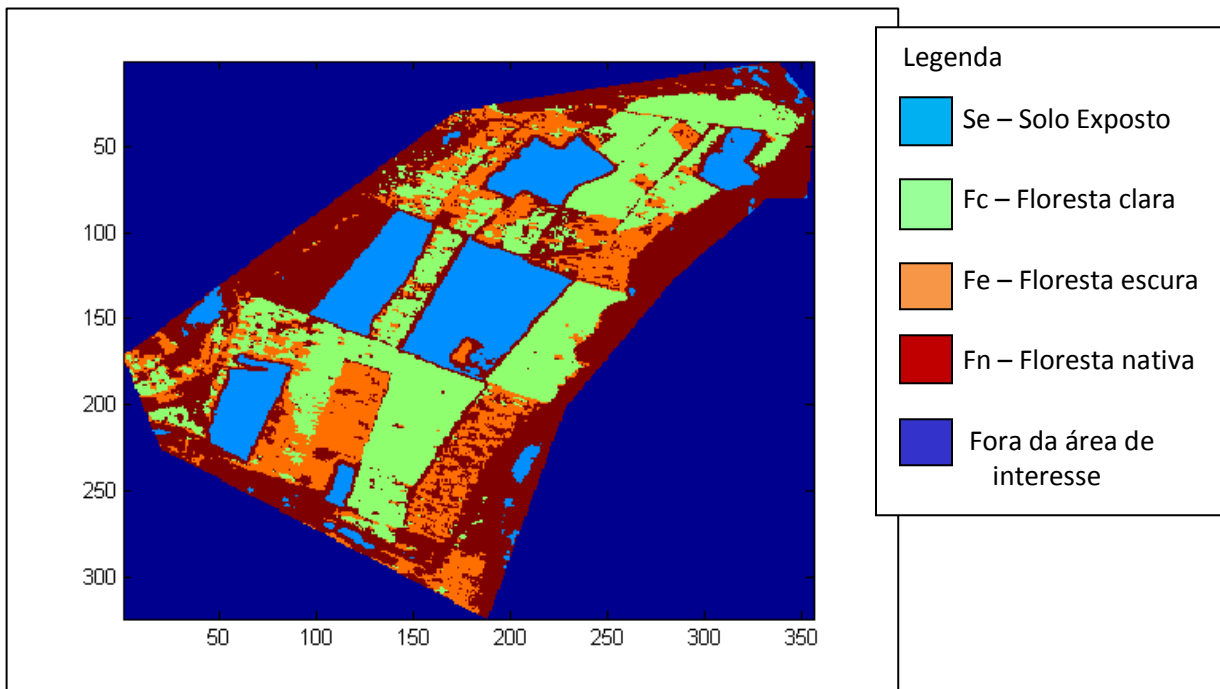


Figura 3: Resultado da classificação utilizando 2/3 da maior amplitude como limite de distância de inclusão.

Na Figura 3, está o resultado da classificação utilizando 2/3 da amplitude maior (a do Infra-Vermelho) como limite mínimo para inclusão de um pixel em cada classe. Neste caso, não houve rejeito, o que sugere que este limite não é adequado. Os limites entre talhões foram classificados como Floresta nativa, quando deveriam ser incluídos na classe de rejeito.

Por outro lado, a classe Solo exposto foi classificada com perfeição, o que já era esperado porque foi a classe que ficou completamente distinta das outras no gráfico de dispersão. A área de floresta nativa foi toda classificada corretamente, mas áreas que não fazem parte desta classe foram incluídas nela.

As áreas de floresta clara e floresta escura foram as de pior resultado, em especial a segunda, que se confundiu com a primeira e com a floresta nativa.

Com o objetivo de obter resultados melhores, foi realizada uma nova classificação utilizando o desvio-padrão do Infra-vermelho como limite (as tentativas incluíram o desvio-padrão multiplicado por 1, 2 e 3). A figura 4, mostra os resultados obtidos com o desvio-padrão multiplicado por dois.

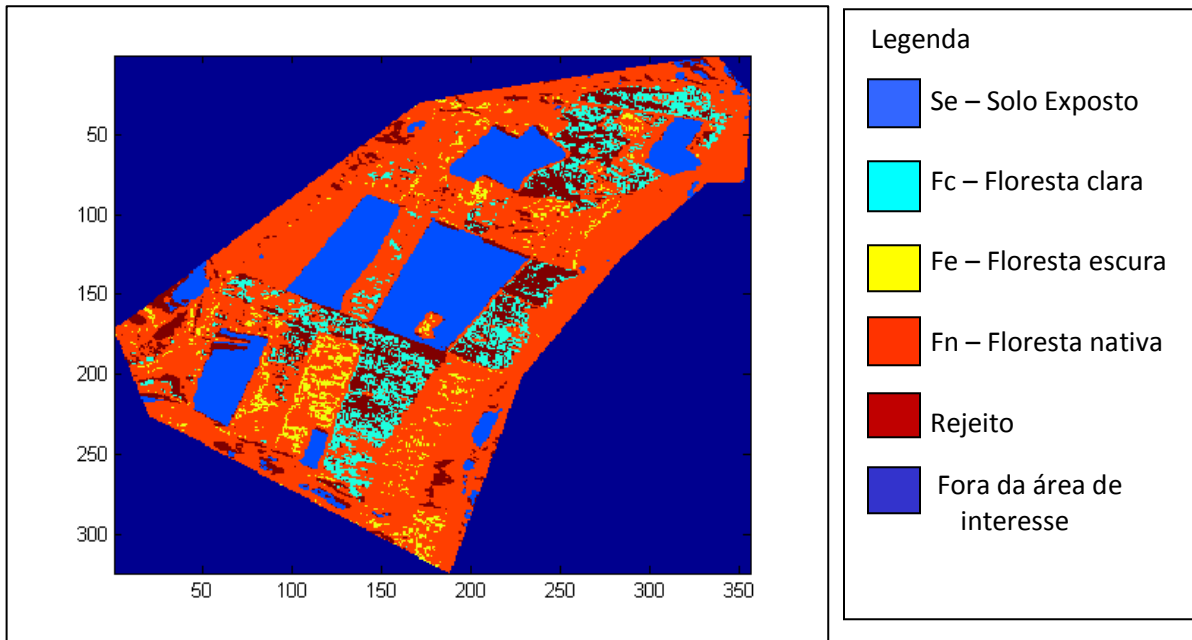


Figura 4: Resultado da classificação utilizando o dobro do desvio-padrão como limite de distância de inclusão.

O segundo resultado foi semelhante ao primeiro ao classificar solo exposto, mas foi muito inferior em todas as outras classes. Apenas algumas divisas foram classificadas como rejeito, e houve prejuízo geral em todas as outras classes.

Conclusão

O método de média euclidiana mais próxima funcionou bem para a única classe que não se sobrepunha às outras no espaço de características estudado (Se). No entanto, considerando a sobreposição entre classes como parte do problema a ser resolvido, este método, sozinho, não pode ser considerado uma solução adequada.