

Trabalho 2

Prof. Mateus Mendelson
mendelson@unb.br

1. Introdução

Este trabalho tem como objetivo apresentar de forma básica alguns conceitos sobre classificação de padrões e processamento de imagens. Para tal, será implementado um algoritmo que auxilia um drone a identificar áreas de pouso seguras. Suponha que o drone esteja equipado com uma câmera na sua parte inferior voltada para o chão, que gera imagens como as apresentadas na Figura 1. O algoritmo deve decidir se a área sob a qual o drone está é (a) segura para pouso (grama) ou (b) perigosa e não deve ser utilizada para pousar (asfalto). Para cada imagem, deve ser calculada a frequência de ocorrência de cada código ILBP [1] e quatro métricas GLCM [3]. A partir destas frequências e métricas, um classificador é utilizado para decidir se a área é segura ou não para pouso.



(a)



(b)

Figura 1.: Imagens adquiridas pela câmera do drone referentes a duas possíveis situações: (a) grama e (b) asfalto.

2. Descritor de textura

Descritores de textura são algoritmos que, dada uma imagem de entrada, fornecem um conjunto de valores capazes de descrever aquela imagem. Dentre os diversos descritores, há o ILBP (do inglês *Improved Local Binary Pattern*), descrito na Seção IV-A em [1] e a GLCM (do inglês, *Gray Level Co-occurrence Matrix*) descrita em [3].

3. Base de dados

Baixe os arquivos texto fornecidos junto com o trabalho. Note que há duas classes: *asphalt* (asfalto) e *grass* (grama). Cada classe possui um conjunto de 50 arquivos texto, que armazenam os valores de cada pixel de sua respectiva imagem em nível de cinza.

Para cada classe, selecione de forma aleatória 25 imagens, que irão compor o conjunto de treinamento. As 25 imagens restantes comporão o conjunto de teste. Desta forma, deverão ser montados dois conjuntos:

- Treinamento: 25 imagens da classe “asfalto” e 25 imagens da classe “grama”; e
- Teste: 25 imagens restantes da classe “asfalto” e 25 imagens restantes da classe “grama”.

4. Treinamento do sistema

Cada arquivo texto que compõe o conjunto de treinamento deve ser lido e carregado de tal forma que a memória necessária para armazenar a imagem seja alocada dinamicamente.

Após ser carregada, o algoritmo ILBP deve ser aplicado na imagem, gerando o vetor de frequências de cada código ILBP. Deve-se, então, realizar o cálculo das GLCMs para cada uma das direções da vizinhança de 8 de um píxel. A partir de cada uma das 8 GLCMs geradas, é necessário calcular seus respectivos valores de contraste, energia e homogeneidade [4]. O vetor ILBP e as 24 métricas geradas das GLCMs devem ser concatenadas, formando um vetor único de características, que deve ser normalizado.

Após o cálculo do vetor de cada imagem do conjunto de treinamento, deve-se calcular a média dos vetores de cada classe. Dessa forma, cada classe estará caracterizada pela média dos seus vetores de frequência

normalizados.

a. Normalizando um conjunto de dados

Sendo x um vetor de dados, calcula-se o vetor de dados normalizados \hat{x} da seguinte forma:

$$\hat{x}_i = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

, sendo $\min(x)$ o menor valor encontrado no vetor x , $\max(x)$ o maior valor encontrado do vetor x . Ainda, \hat{x}_i e x_i representam o i -ésimo elemento do vetor \hat{x} e x , respectivamente.

5. Conjunto de teste

Cada elemento do conjunto de teste deve ter o seu vetor de frequências normalizado calculado. Este vetor será utilizado para classificar aquela imagem como sendo "grama" (área segura para pouso) ou "asfalto" (área perigosa para pouso).

6. Classificador

Para todos os elementos do conjunto de teste, deve-se calcular a distância euclidiana [2] entre o seu vetor de características normalizado e os dois vetores de médias obtidos no treinamento do sistema. A imagem será classificada como pertencente à classe que estiver mais próxima de seu descritor. A Figura 2 mostra um exemplo de separação de padrões.

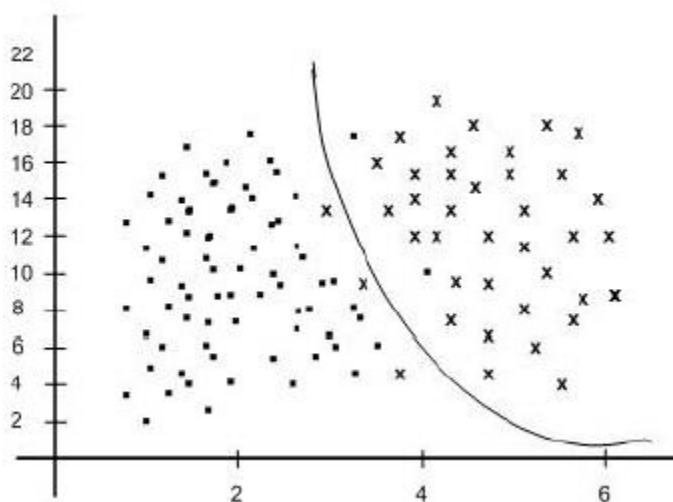


Figura 2.: Representação visual de como diferentes pontos de diferentes classes podem se posicionar no espaço. A linha divisória indica um possível limiar para separar as classes.

7. Métricas

Por fim, o sistema deve imprimir na tela as seguintes métricas:

- Taxa de acerto: a porcentagem de imagens do conjunto de teste que foram classificadas corretamente;
- Taxa de falsa aceitação: a porcentagem de imagens do conjunto de teste que pertencem à classe "asfalto", mas que foram classificadas como sendo "grama"; e
- Taxa de falsa rejeição: a porcentagem de imagens do conjunto de teste que pertencem à classe "grama", mas que foram classificadas como sendo "asfalto".

Referências

- [1] SILVA, M. M. E.; ZAGHETTO, A; ZAGHETTO, C., *Detecção de Fraudes em Leitores de Impressões Digitais sem Contato Utilizando Descritores de Texturas e Redes Neurais Artificiais*, in XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 2016, Santarém, Pará. Anais do XXXIV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais, 2016.
- [2] Wikipedia: https://pt.wikipedia.org/wiki/Dist%C3%A2ncia_euclidiana, em 01 de novembro de 2017.
- [3] Mathworks: <https://www.mathworks.com/help/images/ref/graycomatrix.html>, em 17 de abril de 2018.

- [4] Mathworks: <https://www.mathworks.com/help/images/ref/graycoprops.html>,
em 17 de abril de 2018.