

Projeto de Graduação

O objetivo do projeto para alunos de graduação é fazer o aluno lidar com o processo de reconhecimento de padrões como um todo, principalmente no ciclo de avaliação e retreinamento de classificadores. Para isto iremos utilizar uma aplicação de processamento de imagens chamada **landmark identification**, cujo objetivo é identificar monumentos históricos em imagens.

1 Descrição do projeto

O problema de **landmark detection** consiste em, dada uma imagem de um monumento histórico, como o Cristo Redentor ou as Pirâmides do Egito, identificar qual é o monumento fotografado. Esta é uma área de pesquisa atual que possui aplicações comerciais bem impressionantes, como o Google Goggles.



→ Torre Eiffel

A identificação de monumentos pode ser vista como um problema de classificação em que a entrada é uma imagem e a saída é o nome do monumento retratado. Um dos desafios nesta área é a representação dos dados, pois a informação de cor contida nos pixels carrega pouco significado sobre o conteúdo da imagem. É necessário, portanto, extrair da imagem informações que possam nos ajudar na tarefa de classificação.

O projeto consistirá na *avaliação e comparação de classificadores* de imagens. Os classificadores estudados seguirão um modelo padrão, porém os métodos envolvidos em cada etapa da classificação e

do treinamento podem variar. O estudo deve “montar” classificadores utilizando diferentes métodos para cada etapa e compará-los utilizando técnicas de avaliação vistas em sala de aula. É permitido implementar classificadores que utilizem outras técnicas, porém é necessário compará-los também com os que utilizem o modelo padrão proposto.

2 Classificadores de imagens

2.1 Modelo Bag of Words

O modelo utilizado para representar as imagens será o *Bag of Words*. Neste tipo de modelo é definido um *vocabulário* ou *dicionário* de palavras de tal modo que todo objeto a ser classificado é representado como um mapa contendo a frequência de cada palavra do *dicionário*. O vetor de características é, portanto, um histograma que marca a frequência que cada palavra ocorre em um objeto.

2.1.1 Definição do vocabulário

Para a utilização do modelo *Bag of Words* em imagens, é necessário definir o vocabulário utilizado e um método de extrair as “palavras” de uma imagem. O método utilizado consiste em duas etapas:

1. Detecção de pontos especiais na imagem, chamados de **KeyPoints**;
2. Determinação de um vetor de descrição para cada **KeyPoint**.

Os **KeyPoints** são pontos da imagem que possuem destaque e tem menor chance de variar quando ocorrem transformações como rotação e escala. Um exemplo comumente usado são cantos dos objetos na imagem. O método de detecção de **KeyPoints** é um dos pontos que devem variar nos classificadores.

Para permitir a comparação entre dois **KeyPoints** é extraída uma descrição que procura dar distâncias grandes a pontos diferentes e distâncias pequenas entre pontos em que um é uma variação pequena do outro. O tipo de descritor utilizado é outro ponto que deverá ser variado nos classificadores estudados.

Após a extração e descrição dos **KeyPoints** é necessário definir quais serão os *bins* do histograma de cada objeto, ou seja, quais serão as palavras do vocabulário. Para isto é feita um agrupamento de descritores parecidos e são selecionados alguns “representantes” de cada grupo. Cada objeto será representado por um histograma desses descritores. Se um descritor encontrado não está presente no histograma, então é calculada a similaridade dele com todos os representantes e é contada uma ocorrência para o representante mais próximo.

2.1.2 Classificadores

O *OpenCV* possui diversos classificadores implementados sob uma interface comum. Devem ser testados vários classificadores no projeto. O classificador deverá receber um vetor de características extraído usando uma combinação de extrator de *KeyPoints* e descritores e devolver como classe o nome do monumento retratado na foto.

2.1.3 Implementação

Tantos os extratores de *KeyPoints* como os descritores e o modelo *Bag of Words* estão implementados como funções no *OpenCV*, a biblioteca de processamento de imagens e visão computacional mais usada. Os slides da apresentação do projeto contém links para as documentações das classes relevantes.

3 Monumentos históricos

O projeto deverá identificar os seguintes monumentos:

1. Cristo Redentor;
2. Pirâmides do Egito;
3. Taj Mahal;
4. Muralha da China
5. Parthenon;
6. Torre Eiffel.

O conjunto de treinamento está disponível no Paca. O conjunto de testes será disponibilizado em breve.

4 Entrega

A entrega consistirá nos programas implementando diversas combinações de classificadores, *KeyPoints* e descritores e em um *script* automático de avaliação que utilize as combinações acima e gere um relatório dos testes feitos. Além disto, será necessário escrever um relatório descrevendo as combinações utilizadas e os resultados obtidos

Os programas deverão usar o OpenCV e podem ser feitos tanto em C++ como em Python.

A entrega dos programas e do relatório final está prevista para o dia **03/12**.

5 Dúvidas

Eventuais dúvidas podem ser postadas no PACA. Se for necessário, o monitor está a disposição para marcar um horário para conversar sobre o projeto.