

# Relatório projeto 2-2: Contagem de Pessoas

Elisa Malzoni  
Visão Computacional  
2018.2

## Objetivos

Esse projeto teve como objetivo a contagem de pessoas em uma multidão. Utilizando o modelo de codebook descrito no artigo *Real-time foreground-background segmentation using codebook model* - K. Kim et al..

## Metodologia

Para a separação de *foreground-background* primeiro foi preciso “treinar” o modelo com imagens de fundo, para que fosse possível a detecção do fundo nas imagens com pessoas.

O algoritmo utilizado cria para cada pixel da imagem um ou mais *codewords*, amostras do mesmo pixel são agrupadas em conjuntos de acordo com sua cor e brilho. A determinação como fundo é feita se um pixel satisfaz as seguintes condições:

- se a distorção de cor for menor que um limiar determinado;
- se o brilho está entre os brilhos daquele *codeword*.

## Construção do *codebook*

Seja  $\mathcal{X}$  a sequência de imagens para o treinamento de um único pixel que composto de  $N$  vetores RGB:  $\mathcal{X} = \{x_1, x_2, x_3, \dots, x_N\}$ . E  $C = \{c_1, c_2, c_3, \dots, c_L\}$  o *codebook* de um pixel com  $L$  *codewords*. Cada pixel pode ter tamanhos diferentes pois depende da variação da amostra.

Cada  $c_i$ ,  $i = 1 \dots L$ , é composto de um vetor RGB  $v_i = (R'_i, G'_i, B'_i)$  e uma tupla  $aux_i = (I_{min}, I_{max}, f_i, \lambda_i, p_i, q_i)$ , onde:

$I_{min}$ : brilho mínimo de todos os pixels de daquele *codeword*;

$I_{max}$ : brilho máximo de todos os pixels de daquele *codeword*;

$f_i$ : frequência de ocorrência do *codeword*;

$\lambda_i$ : maior intervalo durante o treinamento que o *codeword* não ocorreu;

$p_i$ : primeira ocorrência do *codeword*;

$q_i$ : última ocorrência do *codeword*.

O algoritmo feito percorria uma sequência de imagens de treinamento que para cada pixel verifica no seu *codebook* se o *codeword* já é existente, conforme sua distorção de cor e o brilho também, é preciso o atualizar. Caso não esteja no *codebook* é necessário adicioná-lo. E depois atualizar a frequência daquele *codeword*.

## Distorção de cor e brilho

A distorção de cor (*colordist*) é dada a partir do cálculo da norma ao quadrado do RGB de um pixel de entrada e da norma ao quadrado do RGB do *codeword*, da seguinte maneira:

$$\begin{aligned} \|x_t\|^2 &= R^2 + G^2 + B^2 \\ \|v_i\|^2 &= R'^2 + G'^2 + B'^2 \\ \{x_t, v_i\}^2 &= (R'R + G'G + B'B)^2 \\ p^2 &= \frac{\{x_t, v_i\}^2}{\|v_i\|^2} \\ \text{colordist}(x_t, v_i) &= \sqrt{\|x_t\|^2 - p^2} \end{aligned}$$

Já para a do brilho é preciso o vetor RGB do pixel, o máximo e mínimo daquele pixel que vem do *codeword*. Assim:

$$\begin{aligned} I_{low} &= \alpha I_{max} \\ I_{hi} &= \min \left\{ \beta I_{max}, \frac{I_{min}}{\alpha} \right\} \\ x_t &= \sqrt{R^2 + G^2 + B^2} \end{aligned}$$

$$\text{brightness}(x_t, I_{min}, I_{max}) = \text{true se } I_{min} \leq x_t \leq I_{max}, \text{ false caso contrário}$$

## Filtro temporal

O refinamento do *codebook* pode ser feito através de um filtro temporal que mantém no *codebooks* os *codewords* onde  $\lambda$  de um seja menor que a metade do número de imagens de treinamento. Separando assim *codewords* que podem ser de objetos do primeiro-plano.

## Detecção de primeiro-plano

Para detecção do primeiro-plano uma imagem com pessoas foi utilizada para verificar se as pessoas iriam ficar destacadas. Para isso foi preciso verificar novamente o

*codebook* do pixel para ver se o *codeword* já existe, dependente da distorção de cor e brilho, e atualizá-lo. Caso já exista é fundo, caso contrário é primeiro-plano. Pixels do primeiro plano foram pintados de branco e os do fundo de preto.

## Contagem das pessoas

Para contagem depois que a imagem já estava dividida entre fundo e primeiro plano, foi feita uma erosão e depois uma dilatação para retirada de ruídos, ficando assim somente com os pixels das pessoas brancos. Após isso, foi achado os contornos da imagem, filtrou-se contornos com área maior que 100, determinando assim as pessoas.

## Implementação

Para implementação foram utilizadas 50 imagens de treino da View\_001, escolhidas a dedo de maneira a tentar cobrir a maior gama de iluminações diferentes. O conjunto de *codebooks* foi salvo em um arquivo por meio da biblioteca *pickle* no arquivo *codebook.p*, que depois foi lido para detecção. Porém como o arquivo ficou muito grande não foi possível adicioná-lo ao git, assim [aqui](#) se encontra o arquivo.

Também o filtro temporal não funcionou com a implementação feita, já que uma vez feita não detectava o primeiro-plano. Mesmo assim o que seria a implementação está no arquivo fonte.

A detecção num primeiro momento foi feita com uma imagem do dataset de background (S0\_BG) que possuía duas pessoas que detectou e contou as duas. Depois utilizou-se uma imagem S1\_L1 da multidão, porém não detectou o primeiro plano de maneira correta.