

MC920 - Trabalho 5: Aprendizado não Supervisionado

Naomi Takemoto
RA 184849
n184849@dac.unicamp.br

I. INTRODUÇÃO

Neste trabalho foi explorada a aplicação de técnicas de aprendizado não supervisionado para comprimir imagens, alterando sua quantização. Mais especificamente foi utilizado o *k-means* que será descrito com mais detalhes nas seções subseqüentes.

O *baseline* para as análises realizadas aqui foram a documentação da biblioteca *scikit-image* e *scikit-learn* [1].

II. TEORIA

Técnicas de agrupamento buscam particionar grupos em um conjunto de dados segundo algum critério de similaridade, podendo constituir assim uma categoria de classificadores conhecidos como não supervisionados. Isso se deve ao fato de esses processos não requerem dados previamente rotulados.

O *k-means* está entre as técnicas de agrupamento particional, ou seja cada elemento do conjunto pertence exatamente a um grupo (para abordagem não *fuzzy*). Este algoritmo se baseia na ideia de centroides, seu número e localização inicial devem ser previamente definidos. A ideia é que cada centroide, cuja posição será modificada iterativamente, definirá um grupo. Idealmente um grupo deverá possuir grande coesão interna em contraste com uma pouca coesão entre elementos de agrupamentos distintos.

O *k-means* pode ser utilizado para reduzir a quantização de imagens. Ao invés de representá-la, por exemplo usando matrizes RGB, trocam-se as cores por índices de um dicionário (*codebook*), cujos elementos são determinados a partir da aplicação do *k-means*. A depender do número de cores escolhidos para a paleta da imagem final, pode haver uma diminuição de 3 para 1 bytes para representar intensidades da imagem.

III. EXPERIMENTOS

Num primeiro experimento, aplicamos o algoritmo *k-means* para diferentes valores de *k*, e observamos os resultados visualmente, como ilustrado na figura 1. Em alguns pares de imagem, como para *k* = 64 e para *k* = 128, praticamente não há diferenças perceptíveis.

Para uma análise mais quantitativa, comparamos também as medidas de entropia de *Shannon* para as imagens convertidas para escala de cinza. Esta última transformação tinha com objetivo tornar o cálculo mais rápido. Como mostra a tabela I.

k	entropia
4	1.95
8	2.83
16	3.70
32	4.52
64	5.48
128	6.30
original	9.25

Tabela I: Entropia de Shannon para diferentes valores de *k*.

Em teoria da informação a entropia é definida como o negativo da soma dos logaritmos em base 2 das probabilidades de cada intensidade de uma imagem. Quanto maior o número de intensidades possíveis, espera-se que a entropia seja maior, o que significa que prever o valor de um pixel se torna mais difícil. Diz-se então que o sistema carrega mais informação.

Da tabela I ainda é possível observar que a entropia não cresce linearmente com o aumento do número de cores. Nesse caso, para um aumento de 64 para 128, a entropia cresceu pouco, indicando que, se for desejado um alto nível de compressão, é possível escolher 64 cores, ao invés de 128 sem grandes prejuízos.

Diferentes técnicas podem ser utilizadas para modificar a quantização de uma imagem. No segundo experimento realizado, comparamos a definição do *codebook* com o *k-means* com a definição aleatória. Este processo ocorre da seguinte maneira:

- 1) Sortear, dentre as intensidades presentes na imagem original, um vetor de tamanho igual ao número de cores que se deseja na imagem de saída.
- 2) Com o auxílio da função *pairwise_distances_argmin*, presente na biblioteca *scikit-learn*, computar as menores distâncias entre cada ponto do vetor acima com relação às intensidades da imagem original. Definindo, para cada posição da imagem um índice para o *codebook*.

Como pode ser observada na figura 2, a degradação, para um mesmo número de cores, é mais visível para a escolha aleatória. Esse efeito pode ser explicado pelo fato de que a técnica de aprendizado não supervisionado tende a sumarizar nos centroides as informações mais importantes com relação às cores. Isso é feito de tal modo que as cores da paleta do *k-means*, definam grupos cujas intensidades são mais representativas no sentido de que dois elementos da paleta são distantes entre si (maior poder de discriminação) e as intensidades que são mapeadas para cada um desses elementos

são suficientemente próximas para minimizar a percepção da perda na quantização.



Figura 1: Da esquerda para direita, de cima para baixo: imagem para $k = 4$, $k = 8$, $k = 64$ e $k = 128$ respectivamente.

IV. CONCLUSÃO

A relação não linear entre a entropia e a quantidade de cores utilizadas mostra a possibilidade aplicar uma grande taxa de compressão utilizando algoritmos não supervisionados para escolha da nova paleta de cores.

O *k-means* pode ser utilizado para diminuir a quantização de uma imagem, preservando razoavelmente sua aparência para o sistema visual humano. Conforme mostrado nos resultados, a perda de quantização é menos visível do que a escolha aleatória de cores para o *codebook*.

REFERÊNCIAS

- [1] Documentação scikit-learn. Disponível em: https://scikit-learn.org/stable/auto_examples/cluster/plot_color_quantization.html. Acesso em 24 de junho de 2019.

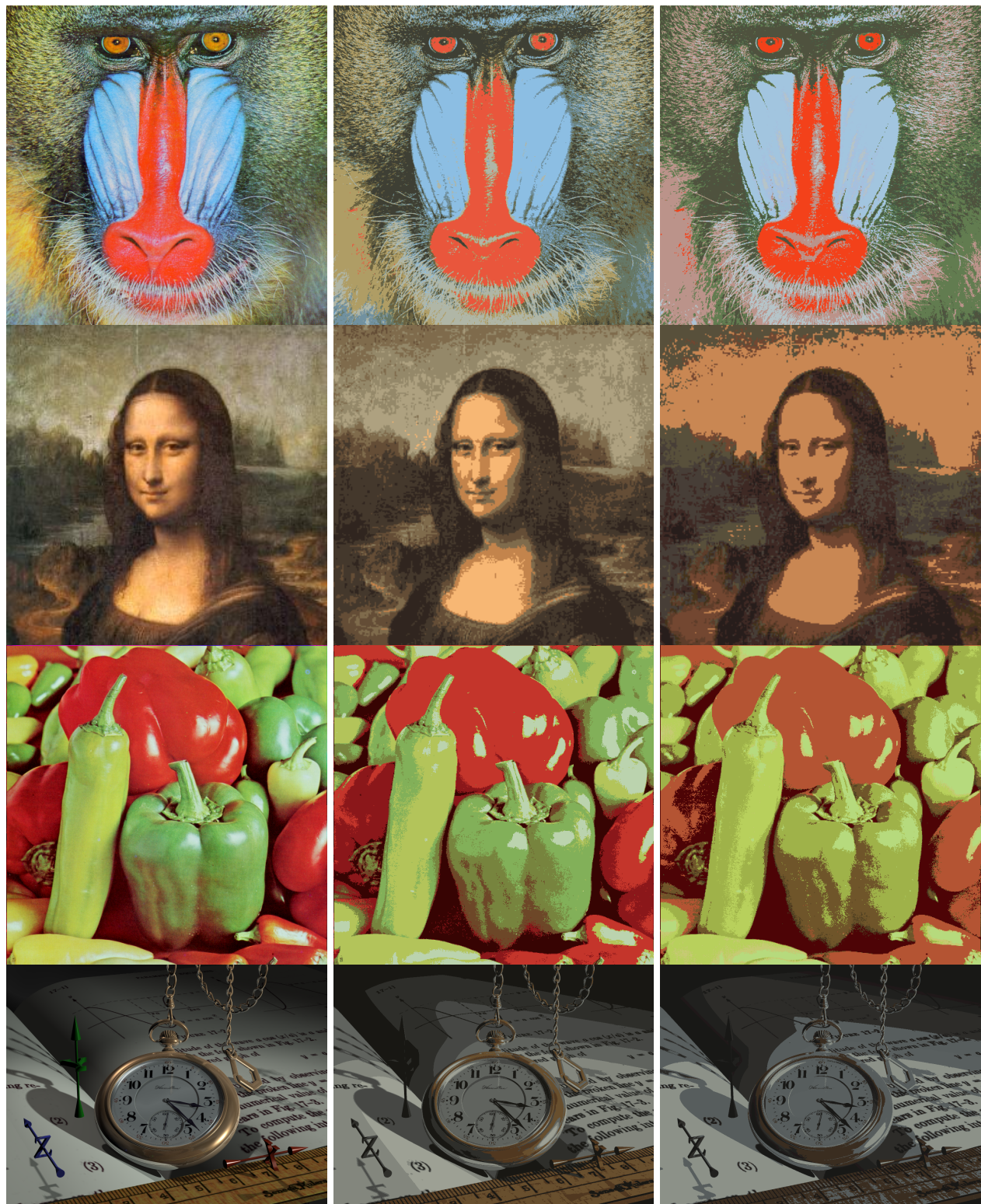


Figura 2: Na primeira coluna a imagem original. Na segunda, uma imagem de 8 cores, escolhidas com base no *k-means*. A terceira coluna, a imagem formada pela escolha randômica de 8 cores, com base nas cores presentes na imagem original.