Evelyn Sjafii, ID: 914017341

Alicia Siu, ID: 914005515

ECS174 Computer Vision

June 4, 2019

Problem Set 3: Bag of Visual Words

***Part I. Short Answer Problems***

1. What exactly does the value recorded in a single dimension of a SIFT keypoint descriptor signify?

***Answer:*** The value represents the magnitude of one of the 8 gradient directions. The SIFT keypoint descriptor has dimension of 128 because each SIFT descriptor consists of 8 gradient directions in a 4x4 bin (4x4x8 = 128). The following image is from: <https://www.oreilly.com/library/view/computer-vision-with/9781788299763/f3a93025-1f6c-47db-a76d-6e73d06328a4.xhtml>



2. A deep neural network has multiple layers with non-linear activation functions (e.g., ReLU) in between each layer, which allows it to learn a complex non-linear function. Suppose instead we had a deep neural network without any non-linear activation functions. Concisely describe what effect this would have on the network. (Hint: can it still be considered a deep network?)

***Answer:***  If instead we had a deep neural network without any non-linear activation functions, this neural network would be calculating linear combinations of linear functions. In other words, if we had a linear function, and we applied a layer, the result would still be a linear function. No matter how many layers are applied, the output would still be a linear function and hence, the neural network cannot learn. The non-linear functions, on the other hand, will generate different non-linear models, and continue to allow the network to learn more complex functions.

***Part II: Programming Problem – video search with bag of visual words***

**1.**