

## HW07: Image representation

1. How does the SIFT descriptor achieve invariance to (a) small deformations, and (b) changes in illumination?

(a) In SIFT we calculate local features based on pixel intensities from a neighbourhood. SIFT tries to find the keypoints with deformation invariant neighbourhoods and these keypoints are used as a representation for an image. When small deformations are applied to images, intensities change position, but not their value with deformations and local neighbourhood structure is still preserved. So the descriptors for the deformed image still matches with the original image and hence SIFT achieve invariance to small deformations.

(b) After localising the keypoints, SIFT computes orientation histograms in a  $16 \times 16$  region around each keypoint. Orientation histograms stores the gradient information in the neighbourhood and these gradients are less sensitive to illumination changes. Each orientation histogram is calculated on  $4 \times 4$  neighbourhoods containing 8 bins. With  $4 \times 4$  equal to 16 histograms with 8 bins each we get a vector of size 128. This normalised vector is invariant to small illumination changes.

2. How does one obtain a dictionary in a bag-of-words representation of images?

In representing images with bag of words, images are represented as bag of local features instead of actual pixels. These features can be any set of hand crafted or learned features such as SIFT are calculated over several local patches. This set of local features for an image is called bag of features. All images are represented by these feature vectors and these feature vectors are used for classifying the images. In the bag of words approach, we are storing the number of occurrences of features and not their actual positions.

3. Give two examples each of categories that can be best described by their shape, and by their texture.

Shape: 1. Pedestrian 2. Cars 3. Chairs and Tables 4. Animals

Texture: 1. Fabric 2. Metal surfaces 3. Wood Surfaces 4. Skin of animals 5. Food surfaces