Image Structure 4 Feature Detection and Matching 4F12: Computer Vision

Instructor: Samuel Albanie

Based on course material authored by Roberto Cipolla



Recap of last lecture

- Matching and Correspondence
- Invariances beyond scale (rotation, 3D viewpoint, partial occlusion, illumination)
- Using keypoints <u>and</u> descriptors
- Raw intensity patches and zero-normalised intensity patches as descriptors
- Histogram of Oriented Gradients and dominant orientations
- The SIFT descriptor, and the nitty gritty details
- Matching features over multiple views



Summary



Image Textures

Definition: a texture is a visual pattern on an infinite 2-D plane which, at some scale, has a stationary distribution¹ (note: there isn't a universally accepted definition of texture, but this one is useful).



Definition: ¹Efros and Leung. "Texture synthesis by non-parametric sampling." ICCV 1999 Images source: Cimpoi et al. Deep filter banks for texture recognition and segmentation, CVPR 2015 What is a texture?



NE Historical context: preattentive vision

Texture and vision Perception of **texture** has received a great deal of attention due to its potential to yield clues about how the visual system is able to group stimuli together in a way that supports their interpretation.





References: M. Wertheimer, "Laws of organization in perceptual forms." (1938); Julesz, Bela. "Textons, the elements of texture perception, and their interactions." Nature 290.5802 (1981): 91-97. Image sources: (red dot) Perception in Visualization, C.G. Healey; https://en.wikipedia.org/wiki/Gestalt_psychology#/media/File:Gestalt_proximity.svg; Zhu et al. "What are textons?." IJCV, 2005







Textons revisited

One way to characterise texture is through its response to a filter bank.

The example on the right consists of 48 filters:

- 8 LoG filters and 4 Gaussian filters at different scales to provide non-oriented responses
- 36 oriented filters at 6 angles, 3 scales, and 2 phases.

The two phases of oriented filters are first and second derivatives of Gaussians on the minor axis and elongated Gaussians on the major axis, and thus detect edges or bars respectively along their major axes.

The **descriptor** is simply the <u>concatenated</u> <u>responses</u> of all of the filters in the filter bank at a pixel.



Describing textures

Malik Textons

Julesz provided a qualitative definition of <u>textons</u>, rather than a mathematical one.

Malik et al. (2001) proposed to redefine textons as the prototypes that result from **clustering** the responses of a filter bank.

Empirically, they found that these tend to correspond to oriented bars and terminators (aligning with elements of the original definition of Julesz).

Reference: J. Malik et al. "Contour and texture analysis for image segmentation." IJCV (2001)





Filter banks visualised



Note: Since filter banks respond to basic image features such as blobs, edges and bars, they are innately robust to many kinds of illumination change in an image.

Blob



Filter responses

Brightness response



Edge response

Blob response







Link to Deep Learning

In our study of image structure, the raw image has been pre-processed through "hand-crafted" feature extractors (for edges, corners, textons).

The feature extractors were not learned directly from data.

Later in the course, we will learn a hierarchy of feature extractors just by looking at examples - from low-level to mid-level invariant representations up to object identities. This is called Deep Learning.

Deep Learning



7

End of Image Structures Lecture 4

