# Recap of last lecture

**Summary**

- Matching and Correspondence
- Invariances beyond scale (rotation, 3D viewpoint, partial occlusion, illumination)
- Using keypoints and 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

VE Tag = Very Examinable

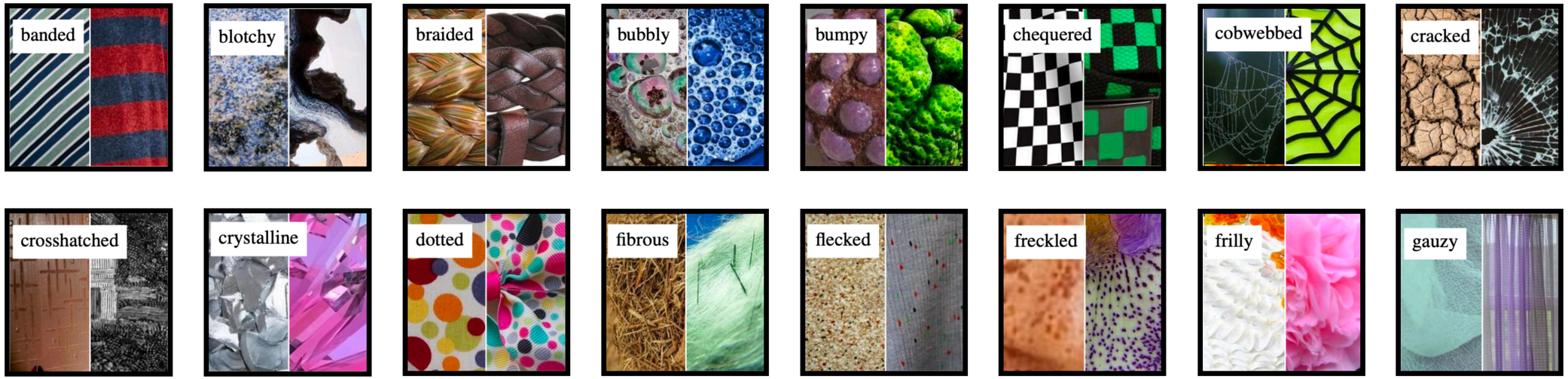
NE Tag = Non Examinable

# Image Textures

## What is a texture?

**Definition:** a texture is a visual pattern on an infinite 2-D plane which, at some scale, has a stationary distribution[1] (**note**: there isn't a universally accepted definition of texture, but this one is useful)

## Example textures




and many others....

Definition: [1]Efros 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

# Historical context: preattentive vision NE

## Texture and vision

Perception of **texture** is studied 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

## Gestalt

In the early 20th century, the *Gestalt school* of psychologists studied how humans perform perceptual grouping, and proposed a set of "laws" that govern how preattentive vision will group elements together

**Example: the law of symmetry**

{ }[ ]{ }

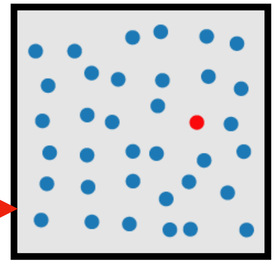
Six elements, but we see three groups

**Example: the law of proximity**

Many dots, but we see four groups

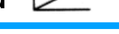
Preattentive processing is the kind of fast, parallel vision that happens (mostly) before attention is involved

Question: Is there a red dot present?



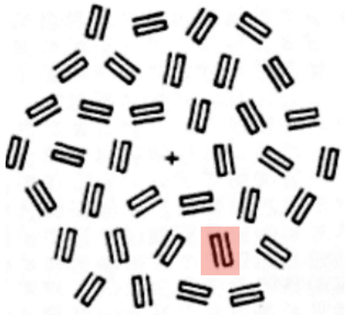
## Julesz Textons

In 1960s -1980s, Bela Julesz studied the statistics of the textures humans could differentiate preattentively. He found the answers could not be explained by Gestalt rules.

Find  among 

Constant time

Find  among 



Time grows linearly with number of elements

Julesz defined textons to be the basic units recognised by preattentive vision that enable texture discrimination (these included oriented bars, crossings, terminators and other low-level structures).

**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

## Describing textures

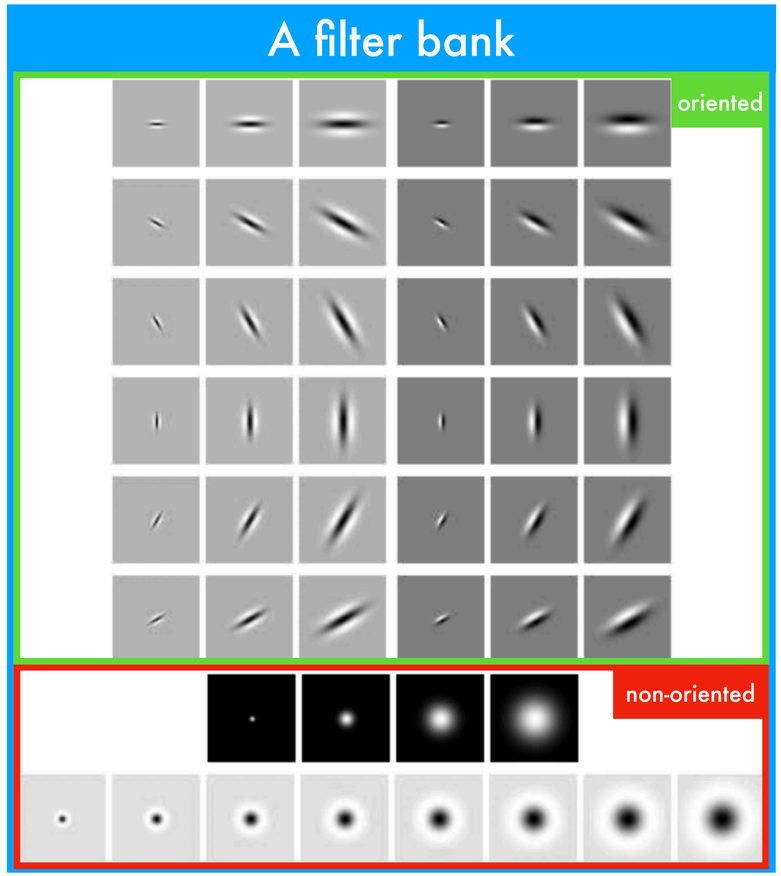
One way to characterise texture is through its response to a filter bank.

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 *concatenated responses* of all of the filters in the filter bank at a pixel.




## Malik Textons

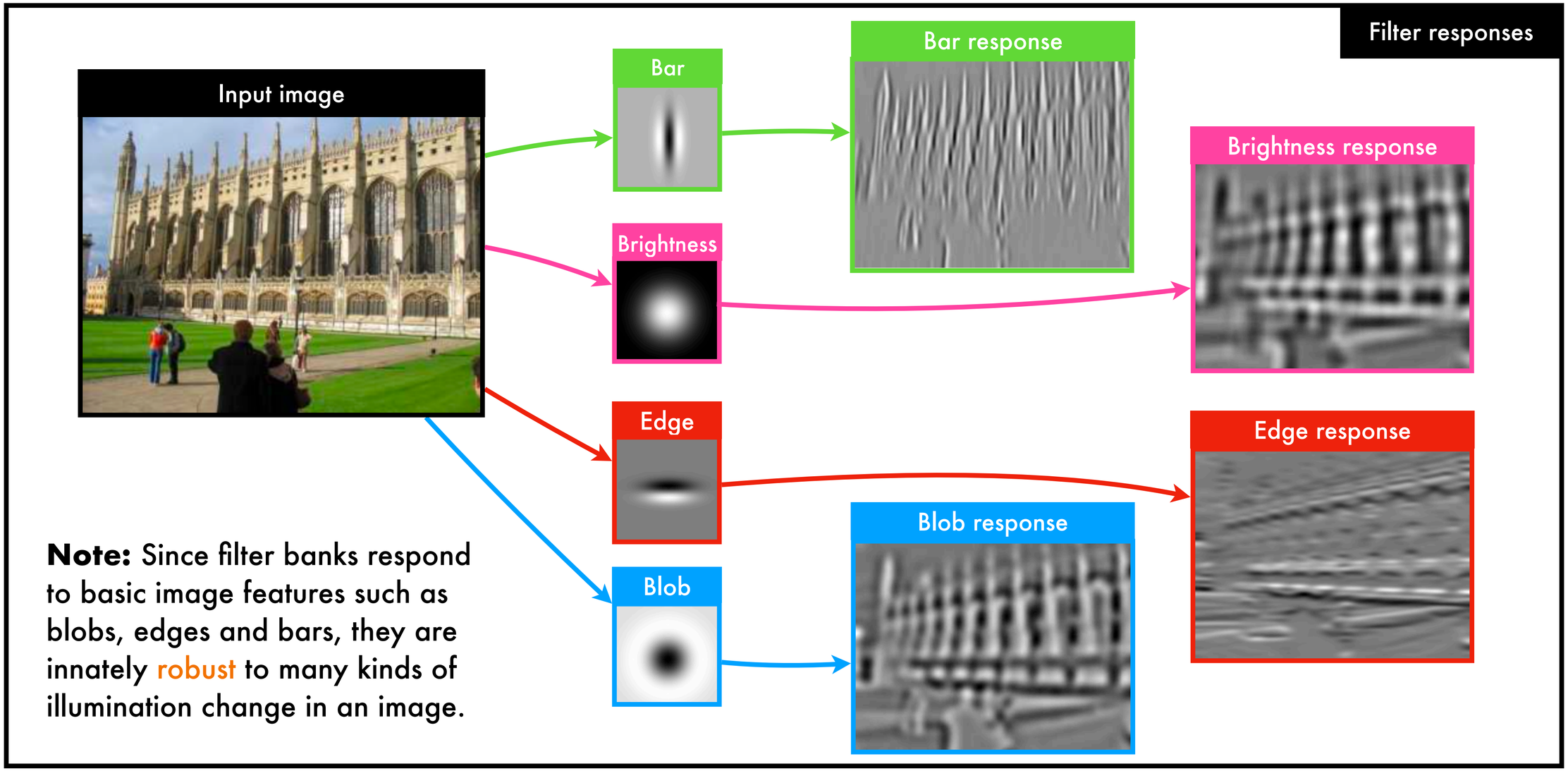
Julesz provided a qualitative definition of *textons*, rather than a mathematical one.

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.

# Link to Deep Learning

**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, you 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.