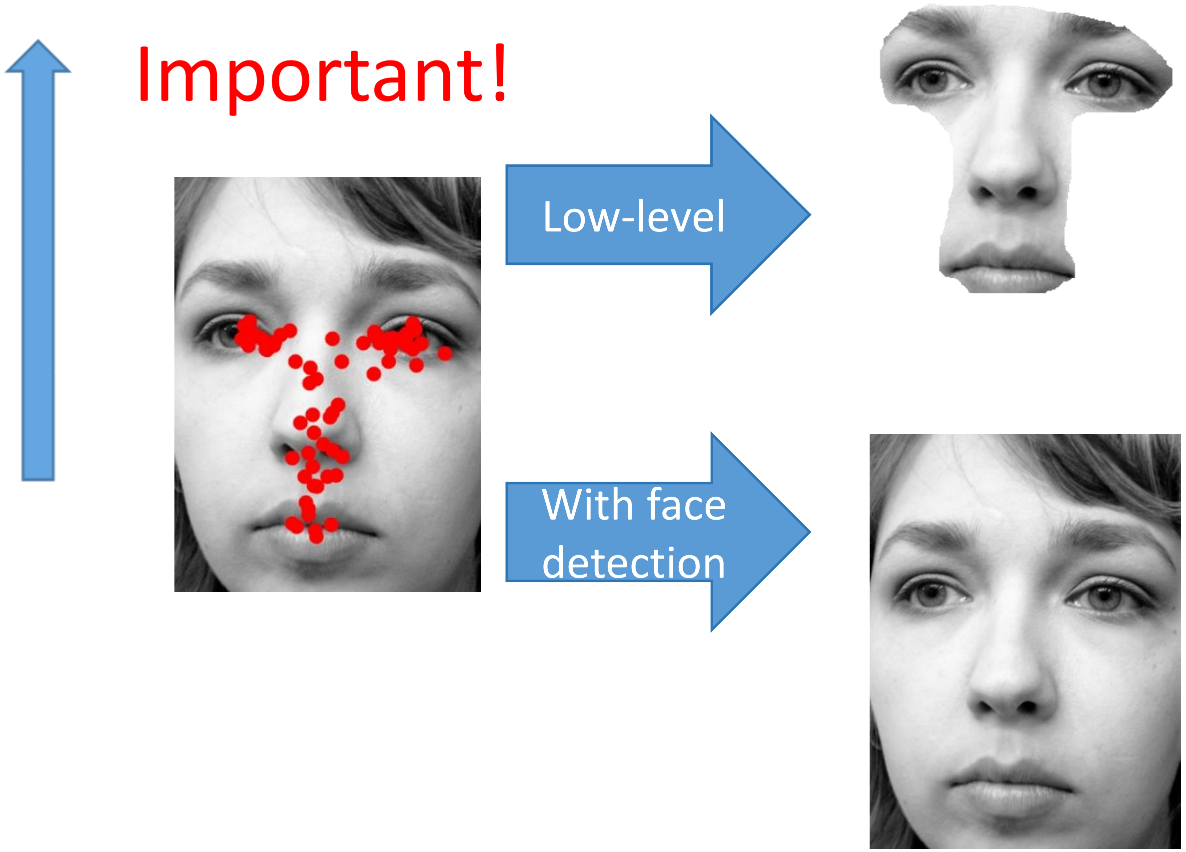
Some Representative Methods for Saliency Detection

Shenghua Gao

ShanghaiTech University

Some Interesting phenomenon in saliency detection

- Low-level(contrast)
  - Color
  - Orientation
  - Size
  - Motion
  - Depth
- High-level
  - People
  - Context



Judd et al, 2009

Outline

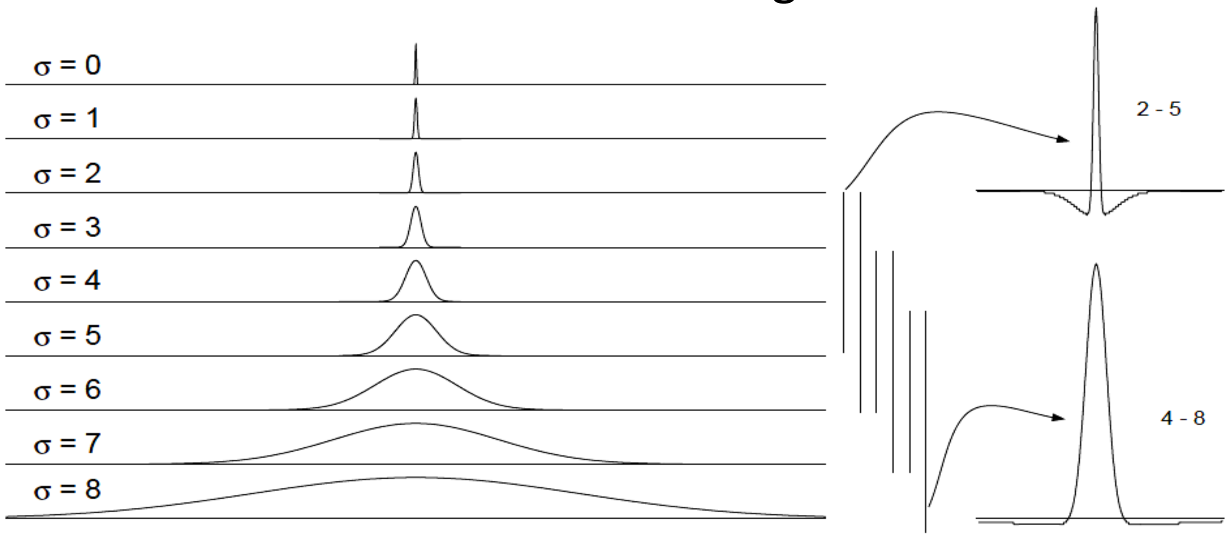
- Button-up approach
  - L. Itti’s approach
  - Frequency-tuned
  - Multi-scale contrast
  - Depth of field
  - Spectral Residual approach
  - Global contrast based
- **Top-down approach**
  - Context-aware
- Information Maximization

L. Itti's approach

- Architecture:

L. Itti’s approach

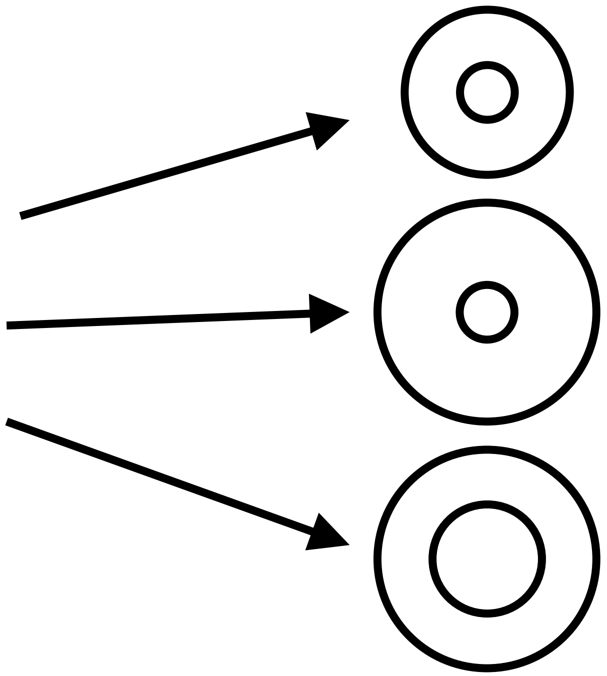
- **Center-surround Difference**
- Achieve center-surround difference through across-scale difference



- Operated denoted by Θ: Interpolation to finer scale and point-to-point subtraction
- One pyramid for each channel: $I(\sigma)$, $R(\sigma)$, $G(\sigma)$, $B(\sigma)$, $Y(\sigma)$ where $\sigma \in [0..8]$ is the scale

L. Itti’s approach

- **Center-surround Difference**
  - **Intensity Feature Maps**
- $I(c, s) = |\ \mathrm{I}(c)\ \Theta\ \mathrm{I}(s)|$
- $c \in \{2, 3, 4\}$
- $s = c + \delta$ where $\delta \in \{3, 4\}$
- So $I(2, 5) = |\ \mathrm{I}(2)\ \Theta\ \mathrm{I}(5)|$
  $I(2, 6) = |\ \mathrm{I}(2)\ \Theta\ \mathrm{I}(6)|$
  $I(3, 6) = |\ \mathrm{I}(3)\ \Theta\ \mathrm{I}(6)|$
  …
- → 6 Feature Maps



L. Itti's approach

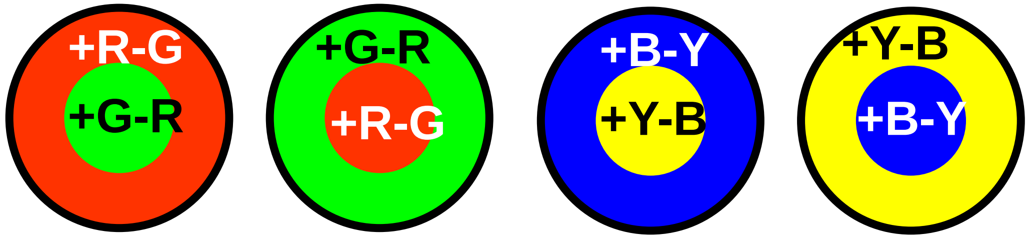
- **Center-surround Difference**
  - **Color Feature Maps**

    Red-Green and Yellow-Blue

- **Center-surround Difference**
  - **Orientation Feature Maps**

    $$O(c,s,\theta) = |O(c,\theta) - O(s,\theta)|$$

- Same c and s as with intensity

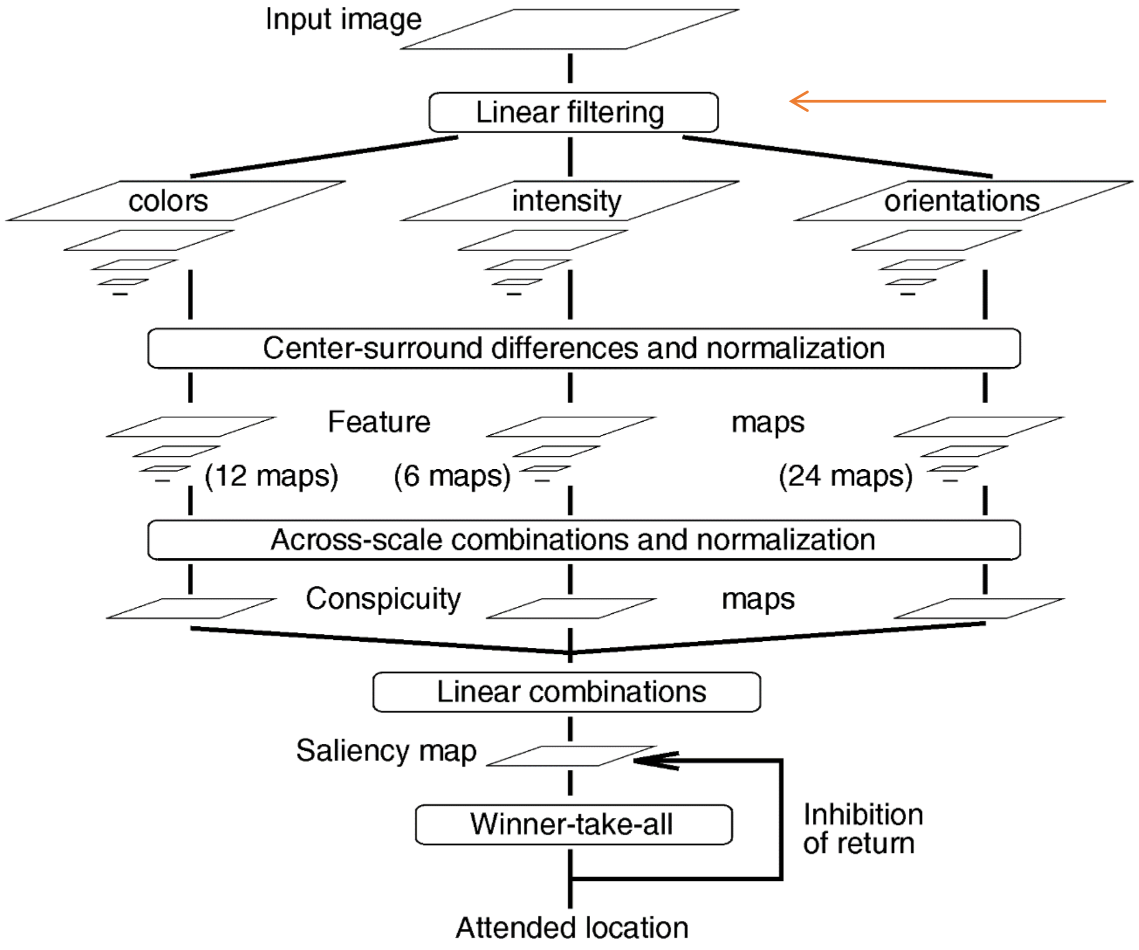


$$RG(c, s) = |\ (R(c) - G(c))\ \Theta\ (G(s) - R(s))\ |$$
$$BY(c, s) = |\ (B(c) - Y(c))\ \Theta\ (Y(s) - B(s))\ |$$

L. Itti's approach

- **Normalization Operator**
- Promotes maps with few strong peaks
- Surpresses maps with many comparable peaks
  1. Normalization of map to range $[0 \ldots M]$
  2. Compute average m of all local maxima
  3. Find the global maximum M
  4. Multiply the map by $(M - m)^2$

A review of Itti's approach



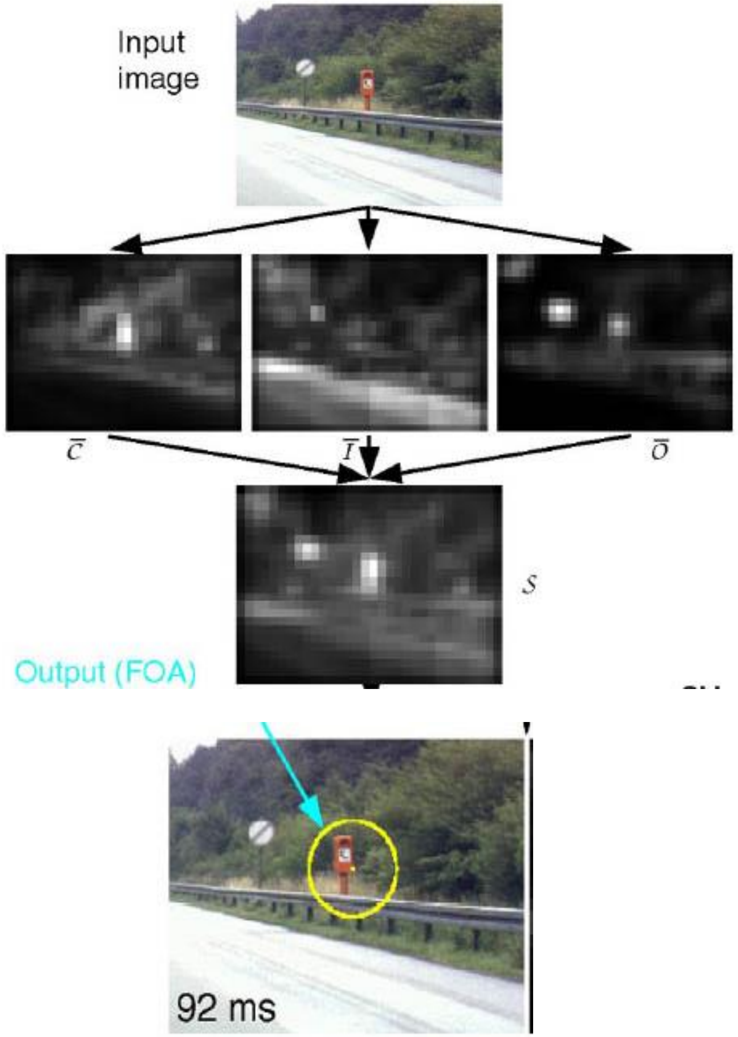
R,G,B,Y

Gaussian Pyramids

Gabor pyramids for θ = {0º, 45º, 90º, 135º}

L. Itti’s approach

Example of Operation:



Outline

- Button-up approach
  - L. Itti’s approach
  - Frequency-tuned
  - Multi-scale contrast
  - Depth of field
  - Spectral Residual approach
  - Global contrast based
- **Top-down approach**
  - Context-aware
- Information Maximization

Detecting single objects

One approach to saliency is to consider saliency as a single object prominent in the image

An Algorithm using this approach is the Spectral Residual Approach

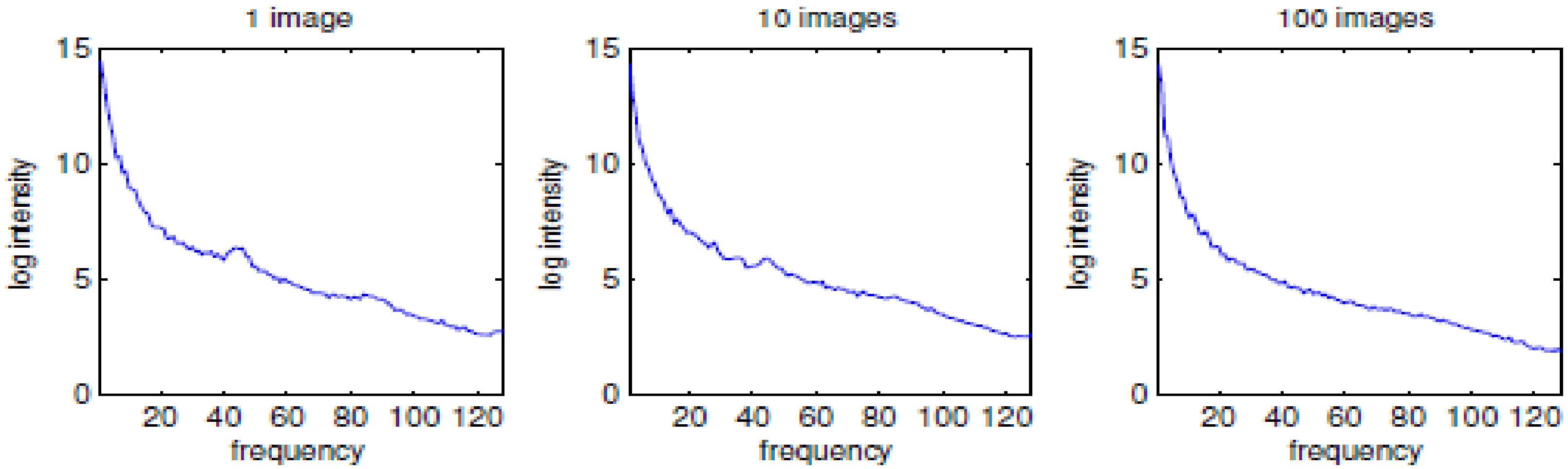
Spectral Residual Approach

What did we say that image Consists of?

That’s right!!! Frequencies

Spectral Residual Approach (1)

Turns out, that if we will take the average frequency domain of many natural images, it will look like this:



Spectral Residual Approach (2)

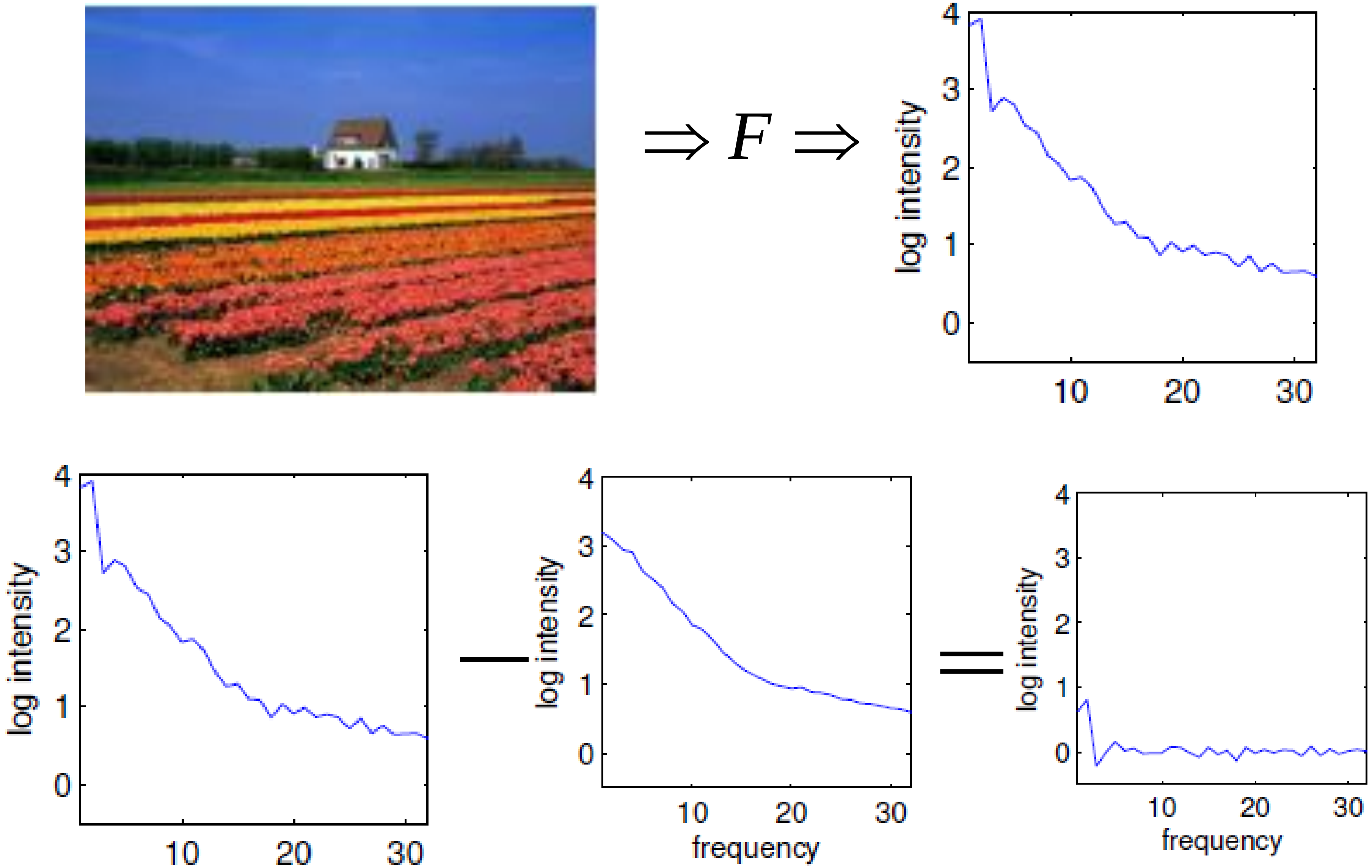
Based on this notion, if we take the average frequency domain and subtract it from a specific Image frequency domain we will get **Spectral Residual**

Spectral Residual Approach

The log spec. ℓ of Image is defined in matlab as:

```
ImageTransform = fft2(Image);
logSpec = log(1+ abs(ImageTransform));
```

Spectral Residual Approach - example



Spectral Residual Approach

h_n will be defined as a blurring matrix sized $n \times n$:

$$h_n(f) = \frac{1}{n^2}\begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 1 \end{pmatrix}$$

Spectral Residual Approach

Generally one takes average over many images to get the average spec but because we have only one image We can convolute it with h_n to get an approximation. Then we can get:

$$spectral\ residual = \ell - (h_n * \ell)$$

Spectral Residual Approach

At this stage, we'll perform inverse fft and go back to
The space domain. In matlab:

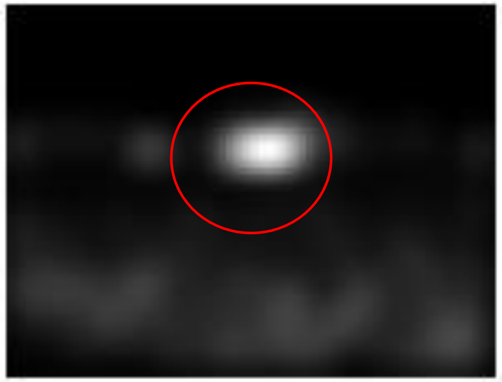
```
SaliencyImage = ifft2(ImageSpecResidual);
```

Spectral Residual Approach

And we will take a threshold to determine the Object map:

$$O = \begin{cases} 1 & if\ spectral\ resadual > threshold \\ 0 & otherwise \end{cases}$$

The saliency map:



Outline

- Button-up approach
    - L. Itti's approach
    - Frequency-tuned
    - Multi-scale contrast
    - Depth of field
    - Spectral Residual approach
    - Global contrast based
- **Top-down approach**
    - Context-aware
- Information Maximization