Computer Vision Homework 2, Part1
Spectral Residual Approach Implementation

Jiawei Hou
Shanghaitech University
houjw@shanghaitech.edu.cn

Abstract

I read the spectral residual approach [1], which detects the salient regions based on Fourier transform, and use Matlab to implement the algorithm. In this report, I will describe the algorithmic process and discuss the parameters setting. Finally, I will show that my program runs well results.

1. Introduction

From the information theory point of view: information can be divided into redundant and change part. People's vision is more sensitive to changing parts. Then the image can be decomposed into two parts:

$$H(Image) = H(Innovation) + H(PriorKnowledge) \tag{1}$$

The author found that the average of the log spectrum of a large number of images is proportional to the frequency, which can be expressed as follows:

$$E\{Af\} \propto 1/f \tag{2}$$

In the paper, the author shows that the log spectrum-frequency curve of a large number of images, on log-log scale, almost a straight line. The log spectrum in the text is the natural logarithm of the amplitude spectrum after image Fourier transform. Therefore, the author has proposed that the log amplitude spectrum of an image minus the average log amplitude spectrum is the significant part of the image, which can be simply written as:

2. Program analysis

In this section we analyze the program step by step, and discuss the related parameters setting.

First of all, we need to calculate the Fourier transform of the image f, then get the log amplitude spectrum $L(f)$ of image and the average log amplitude spectrum $L(f) * H$, and its phase spectrum $P(f)$. And the spectral residual of an image can be worked out:

$$R(f) = L(f) - L(f) * H \tag{3}$$

where $L(f)$ is the log amplitude spectrum of image f, $A(f)$ is the average log amplitude spectrum, and $A(f)$ is obtained by averaging $L(f)$ with a $k * k$ smoothing filter H ($k = 3$ in the program). $R(f)$ is the Spectral Residual of the image f. To compute the saliency of the image, we use $exp(R(f) + iP(f))$ to find the natural exp. And do the inverse Fourier transform, and then use a Gaussian blur filter to filter. Finally, we get its significant area:

$$S(x) = [iFFT(exp(R(f) + iP(f)))^2] * G(x); \tag{4}$$

where $G(x)$ is a Gaussian filter.

3. Result analysis

Here we use the giraffe images as inputs to test the code performance.

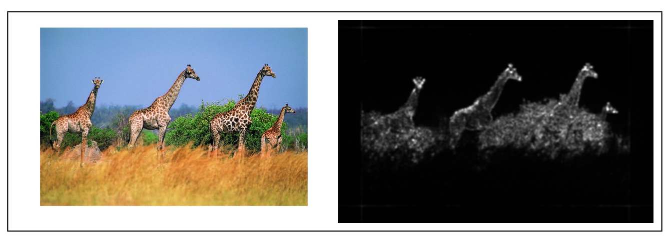

Figure 1. Left: input; Right: the result

Figure 1 parameters: Size of smoothing filter: 3*3;

References

[1] X. Hou and L. Zhang. Saliency detection: A spectral residual approach. pages 1–8, 2007.