Image Processing III
Watershed

Aadil Anil Kumar
Otmane Sabir

26/4/2020

Introduction

The third homework assignment required us to implement the similarity metric algorithm "mutual information" while following certain guidelines which could be summarized to the following list:

1. Implement the algorithm.
2. Choose 3 images from the internet and 1separate the green and red channels as two separate gray scale images each. Crop the green channel by cutting 20 pixels from the left and right sides, respectively - resulting in a cropped green channel image with 40 pixels smaller than the red channel image. Then virtually move the red channel images in x-direction over the corresponding green channel image in 41 steps from the left to right, and compute the mutual information of the overlapping regions for every step.
3. Plot the mutual information as a function of the x-position of the red channel image for all three chosen red/green image pairs.

Contents

1 Mutual Information

Mutual information is a quantity that measures a relationship between two random variables that are sampled simultaneously. In particular, it measures how much information is communicated, on average, in one random variable about another. For example, suppose X represents the roll of a fair 6-sided die, and Y represents whether the roll is even (0 if even, 1 if odd). Clearly, the value of Y tells us something about the value of X and vice versa. That is, these variables share mutual information. [1]

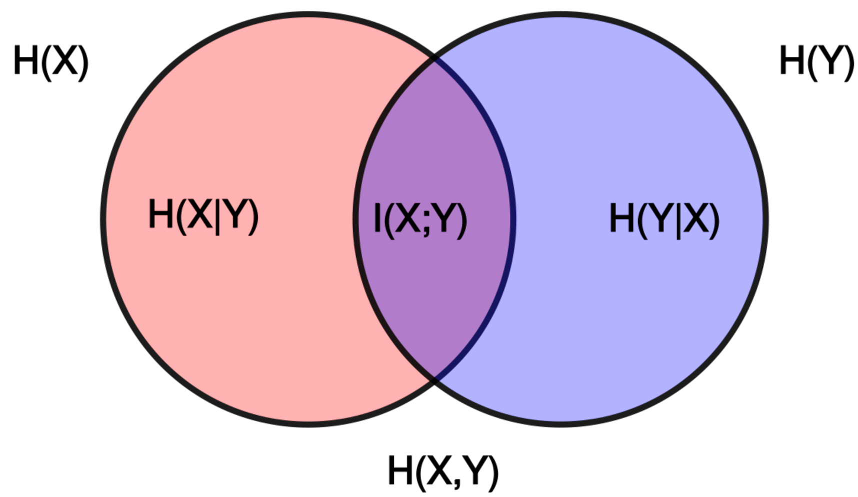

Figure 1: Mutual Information Visualization

1.1 Formal Definition

DEFINITION:[2][3]

The formal definition of the mutual information of two random variables X and Y, whose joint distribution is defined by $P(X, Y)$ is given by

$$I(X;Y) = \sum_{x \in X} \sum_{y \in Y} P(x, y) \log \frac{P(x, y)}{P(x)P(y)}$$

Such that, $P(X)$ and $P(Y)$ are the *marginal distributions* of X and Y obtained through the marginalization process which can be defined as :

$$p_X(x_i) = \sum_j p(x_i, y_i)$$

$$p_Y(y_j) = \sum_i p(x_i, y_i)$$

2 Shannon Entropy & Mutual Information

In our implementation, we rely on calculating the shannon entropy in order to estimate the mutual information. Intuitively, some may ask what is entropy?

The entropy is the expected value of the self-information, the self-information quantifies the level of information or surprise associated with one particular outcome or event of a random variable, whereas the entropy quantifies how "informative" or "surprising" the entire random variable is, averaged on all its possible outcomes.

2.1 Formal Definition : Shannon Entropy

DEFINITION:[4]

The formal definition of the entropy of a discrete random variable X with possible values $\{x_1, ..., x_n\}$ and probability mass function $P(X)$ is

$$H(X) = -\sum P(x_i) \log_b P(x_i)$$

where b is the base of the logarithm used. Common values of b are 2, Euler's number e, and 10. In the case of $P(x_i) = 0$ for some i, the value of the corresponding summand $0 \log_b(0)$ is taken to be 0.

2.2 Formal definition : Mutual Information

DEFINITION:[3]

If we consider pairs of discrete random variables (X,Y), then formally, the mutual information can be defined as :

$$I(X : Y) = H(X) + H(Y) - H(XY)$$

with $H(X)$, $H(Y)$ the Shannon entropy of X an Y, and $H(XY)$ the Shannon entropy of the pair (X, Y).

References

[1] Entropy and Mutual Information:
AS Kornillov
`https://people.cs.umass.edu/ elm/Teaching/Docs/mutInf.pdf`

[2] Mutual Information:
Wikipedia
`https://en.wikipedia.org/wiki/Mutual`$_{i}nformation$

[3] Mutual Information :
Quantiki :
`https://www.quantiki.org/wiki/mutual-information`

[4] Entropy (Information Theory) :
Wikipedia :
`https://en.wikipedia.org/wiki/Entropy`$_{(}information_{t}heory)$