

# 实验 2-Detect Common

## 一、实验目的

- ✓ 了解相似物体检测的方法
- ✓ 计算图像局部自相似性描述子，实现图像的相同区域检测

## 二、实验介绍

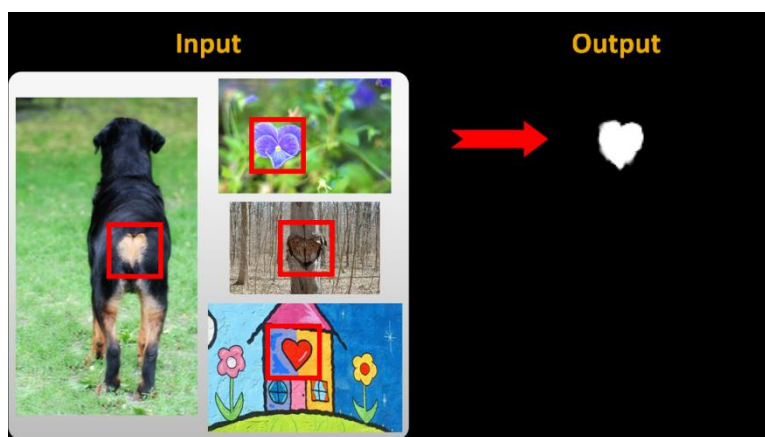


Figure 1 Detect the common of images

随着 ILSVRC 数据库的建立与开放，大批优秀的计算机视觉领域的科研人员投入到 ImageNet 竞赛中，由此引发了近期物体检测等相关领域的研究狂潮。因此，认识与了解物体检测的相关研究工作，是很有必要的。

相同目标检测：给定几张图片，确定这些图片所共有的相同目标。

目标所在的位置与目标的大小都是不确定的，这些都给相同目标检测带来了困难。

Shai Bagon 等人提出一种基于局部自相似性描述子匹配的算法，用于解决相同目标检测的问题。

## 三、实验环境

C/C+, Opencv, C#, Matlab, Python 均可

## 四、实验步骤

### 1、计算像素点的局部自相似描述子

- 对于每个像素点  $I_x$ ，计算该点与以其为中心的图像区域内的其他点的相似性。其中两点  $I_x, I_y$  的相似性度量方式为：记以这两点为中心的图像块分别为  $x, y$ ，则其相似性为  $sim = \exp(-\alpha \|x - y\|^2)$
- 进行对数极坐标变换。设图像区域内某点的直角坐标为  $(x, y)$ ，对数极坐标为  $(\rho, \theta)$ ，则二者的转换关系为  $\rho = \log(\sqrt{(x - x_c)^2 + (y - y_c)^2})$ ， $\theta = \arctan((y - y_c)/(x - x_c))$ 。其中  $(x_c, y_c)$  为区域中心的直角坐标。
- 划分成多个小区域，并用每个小区域的最大值作为描述子的元素
- 这样就可以得到给定点的局部自相似性描述子

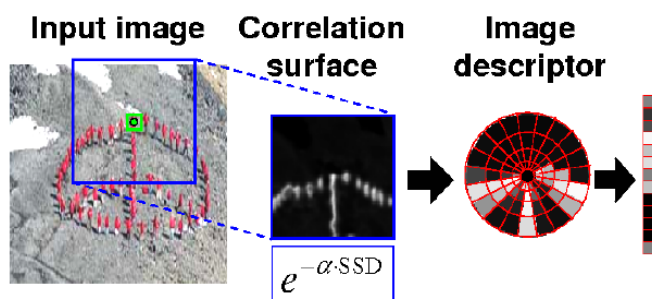


Figure 2 The Local Self Similarity Descriptor

### 2、基于自相似性描述子的目标匹配

- 用矩形框遍历给定的第一张图像，假设有  $n$  个矩形框
- 对于每个矩形框图像区域，基于自相似描述子在其他输入的每张图像中找到最匹配的区域 ( $w$  的建议取值范围为  $[80, 120]$ ,  $h$  的建议取值范围为  $[120, 160]$ ，可根据实验结果自行调试)

$$\begin{aligned} Score(S|\tilde{I}_1, \dots, \tilde{I}_K) &= \sum_{k=1}^K match(S, \tilde{I}_k) \\ &= \sum_{k=1}^K \sum_{i=1}^{w \cdot h} sim(d_i^S, d_i^k) \end{aligned}$$

- 这样，我们可以得到  $n$  对相匹配的矩形框图像区域

### 3、确定图像相同区域

- 从  $n$  个区域对中确定最有意义的目标区域对

首先，在每个给定的图像  $I_k$  中，找到与  $\mathcal{I}$  的最高匹配值  $\max Match(\mathcal{I}, I_k)$ ，在这之前需要依据  $\mathcal{I}$  和  $I_k$  的自相似性描述子得到子图像  $\mathcal{I}'$  与图像  $I_k$  各子图像之间的匹配值：

$$match(\mathcal{I}', I_k^m) = \sum_{i=1}^{w \cdot h} sim(d_i^{\mathcal{I}'}, d_i^{k^m})$$

其中,  $sim(d_1, d_2) = -\|d_1 - d_2\|_p$ , 进而可以计算得到图像  $I_k$  中与  $\mathcal{P}$  的最高匹配值:

$$\max Match(\mathcal{P}, I_k) = \max_m (match(\mathcal{P}, I_k^m))$$

依据  $\mathcal{P}$  和图像  $I_1, \dots, I_K$  之间的匹配值, 可以计算出  $\max Match(\mathcal{P}, I_k)$  的统计显著性, 定义如下:

$$Significance(\mathcal{P} | I_1, \dots, I_K) = \frac{1}{K} \sum_{k=1}^K \frac{[\max Match(\mathcal{P}, I_k) - avgMatch(\mathcal{P})]}{stdMatch(\mathcal{P})}$$

其中,  $avgMatch(\mathcal{P})$  表示图像  $I_1, \dots, I_K$  与  $\mathcal{P}$  的所有匹配结果的平均值,  $stdMatch(\mathcal{P})$  表示这些匹配值的标准差。

在实际算法求解时, 由于不知道目标区域的大小, 因此从给定图像中任意选择一张图像  $I_k$ , 指定一系列不同的  $w$  和  $h$  的值, 计算图像  $I_k$  的子图像  $\mathcal{P}$  与剩下的  $k-1$  张图像的统计显著性, 统计显著性最大的  $\mathcal{P}$  就是最佳的包含公共区域的子图像。

## 五、参考文献

- [1] Detecting and Sketching the Common, by Bagon et.al, CVPR
- [2] Matching Local Self-Similarities across Images and Videos, by Shechtman et.al, CVPR