



Research Institute for Future Media Computing    Institute of Computer Vision  
未来媒体技术与研究所    计算机视觉研究所



# 图像检索评价标准

授课教师：文嘉俊

邮箱：wenjiajun@szu.edu.cn

2024年春季课程

# ■ 图像检索中常用的评价标准

## ◆ 图像检索考察的是什么

- 检索得到的图像与查询的信息相似
  - 把图像检索问题转变为图像分类问题
  - 使用图像分类的评价标准来评价结果
    - Precision (查准率、准确率)、Recall (查全率)、F值
- 排在前面的检索图像更重要
  - 体现图像检索任务的独特性
  - 融入排序的考察方式
    - Precision Recall曲线 (PR曲线)、MAP (Mean Average Precision)、NDCG (Normalized Discounted Cumulative Gain)

# 常用的评价标准：查

## ◆ 查准率

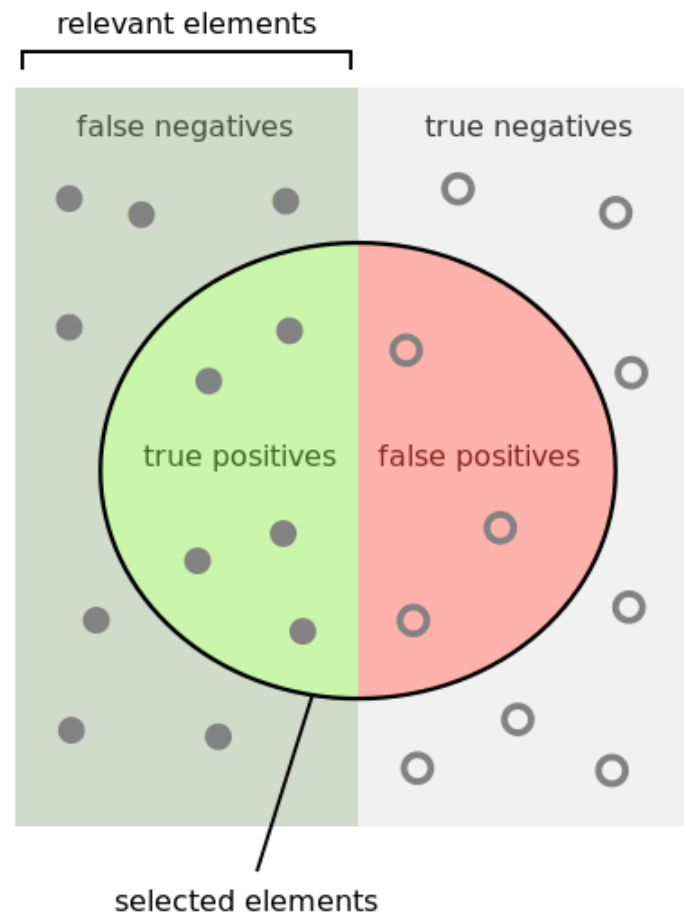
信息检索系统中，检索出的图像与查询相关的图像的比例 (the images retrieved that are relevant to the information retrieval system )

公式:  $\text{Precision} = (N_{\text{relevant}} \cap N_{\text{retrieved}}) / N_{\text{retrieved}}$

## ◆ 查全率

信息检索系统中，与查询相关的图像中，检索出的图像的比例 (fraction of the images that are relevant to the information retrieval system )

公式:  $\text{Recall} = (N_{\text{relevant}} \cap N_{\text{retrieved}}) / N_{\text{relevant}}$



How many selected items are relevant?

Precision =



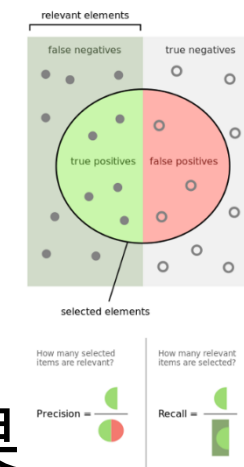
How many relevant items are selected?

Recall =



查准率：检索出来的部分，有多少是与查询图像是相关的，要除以检索出来的部分  
查全率：与查询相关的图像中，有多少被检索出来，它占相关的比例是多少？

# 思考题



- ◆ 数据库有100个图片，有40个与查询词相关，进行检索时，检索得到了30个检索结果，这些结果都属于与查询词相关的，Precision是多少，Recall是多少？

$$\text{Precision} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{retrieved}} = 30/30 = 100\%$$

$$\text{Recall} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{relevant}} = 30/40 = 75\%$$

- ◆ 数据库有100个图片，有40个与查询词相关，进行检索时，检索得到了40个检索结果，其中有30个结果与查询词相关的，Precision是多少，Recall是多少？

$$\text{Precision} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{retrieved}} = 30/40 = 75\%$$

$$\text{Recall} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{relevant}} = 30/40 = 75\%$$

$$\text{Precision} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{retrieved}}$$

$$\text{Recall} = (N_{\text{relevant}} \wedge N_{\text{retrieved}}) / N_{\text{relevant}}$$

# 常用的评价标准：AveP, MAP

## ◆ AveP强调排列在前的图像更重要

$$AveP = \frac{\sum_{k=1}^n (P(k) \times rel(k))}{\text{number of relevant documents}}$$

其中 $rel(k)$ 表示第 $k$ 个文档是否相关，若相关则为1，否则为0， $P(k)$ 表示前 $k$ 个文档的准确率。 $AveP$ 的计算方式可以简单的认为是：

$$AveP = \frac{1}{R} \times \sum_{r=1}^R \frac{r}{\text{position}(r)}$$

其中 $R$ 表示相关文档的总个数， $\text{position}(r)$ 表示，结果列表从前往后看，第 $r$ 个相关文档在列表中的位置。比如，有三个相关文档，位置分别为1、3、6，那么 $AveP = \frac{1}{3} \times (\frac{1}{1} + \frac{2}{3} + \frac{3}{6})$ 。在编程的时候需要注意，位置和第 $i$ 个相关文档，都是从1开始的，不是从0开始的。

## ◆ MAP就是AveP按照query的数量进行平均，得到平均结果

$$MAP = \frac{\sum_{q=1}^Q AveP(q)}{Q}$$

## ■ 思考题

- ◆ 假设两个查询图，查询图1有4张相关图，查询图2有5张相关图。某系统对于主题1的4张相关图分别的排位是：1, 2, 4, 7。对于查询图2，只检索出3张图，相关排位是1, 3, 5。请计算MAP。

## ■ 答案

- ◆ 查询图1：平均准确率为  
 $(1/1 + 2/2 + 3/4 + 4/7)/4 = 0.83$
- ◆ 查询图2：平均准确率为  
 $(1/1 + 2/3 + 3/5 + 0 + 0)/5 = 0.45$   
 $MAP = (0.83 + 0.45)/2 = 0.64$

# 思考题

## ◆ NDCG (Normalized Discounted Cumulative Gain) 的方法如何进行计算?

在MAP计算公式中，文档只有相关不相关两种，而在NDCG中，文档的相关度可以分多个等级进行打分。

$$N(n) = \underbrace{Z_n}_{\text{Normalization}} \underbrace{\sum_{j=1}^n}_{\text{Cumulating}} \underbrace{(2^{r(j)} - 1)}_{\text{Gain}} \underbrace{/\log(1+j)}_{\text{Position discount}}$$