

AVA数据集 简介与分析

主体内容

AVA数据集的标注有3种：

1. Aesthetic annotations: 每张图片平均200个人打分。参与打分的人包括专业/业余摄影师。
2. Semantic annotations: 总共66个语义tag。约200,000张图至少一个tag, 150,000张图两个tag
3. Photographic style annotations: 14种摄影风格-Complementary Colors (949), Duotones (1,301), High Dynamic Range (396), Image Grain (840), Light on White (1,199), Long Exposure (845), Macro (1,698), Motion Blur (609), Negative Image (959), Rule of Thirds (1,031), Shallow DOF (710), Silhouettes (1,389), Soft Focus (1,479), Vanishing Point (674).

一个概念：challenge = title + short description

其中title可以理解为中心词，short description则是对中心词的简单描述



Figure 1. A sample challenge entitled “Skyscape” from the social network www.dpchallenge.com. Images are ranked according to average score and the top three are awarded ribbons.

关于AVA数据集的分析

1: aesthetic preference

- 1. 分数在2~8之间的图占99.77%，这些图的分数都呈现高斯分布；极高/极低分数的图呈现非高斯分布
- 2. 4~6分的图的方差比极高/极低分数的图的方差小
- 3. 相同均分的图，如果是高方差的，通常使用创新型的摄影技巧；低方差的，通常使用传统摄影技巧。但是低分的图+方差也小=这个图真的很烂。参考👇这张四宫格。

		variance	
		low	high
mean	low	poor, conventional technique and/or subject matter	poor, non-conventional technique and/or subject matter
	high	good, conventional technique and/or subject matter	good, non-conventional technique and/or subject matter

Table 3. Mean-variance matrix. Images can be roughly divided into 4 quadrants according to conventionality and quality.

2: semantic content and aesthetic preference

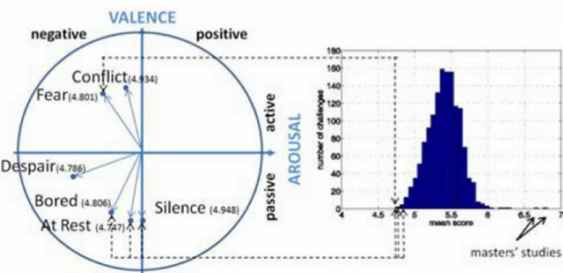


Figure 6. Challenges with a lower-than-normal average vote are often in the left quadrants of the arousal-valence plane. The two outliers on the right are masters' studies challenges.

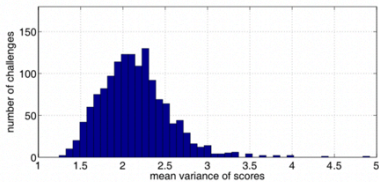



Figure 7. Histogram of the mean variance of score distributions over all challenges. Free studies tend to have low-variance score distributions.

- 1. 📌 (左图)
 - 看右边的那张直方图。横坐标是challenge的编号，纵坐标是这个challenge

的均分。可以发现challenge的均分均匀分布在rating scale的中点附近。

（什么是rating scale的中点？）企图分析challenge-description之间的语义联系：

- 高分的challenge：那些均分特高的challenge之间无语义关联
 - 有两个特殊的challenge：只有在之前的challenge中获奖的人才被允许参加的。这些challenge的均分都特别高。（即：参赛选手的实力普遍都高咯）
 - 看左边的圆圈。有两个坐标轴：arousal & valence（中文意思不知道） The dimension of **valence** ranges from highly positive to highly negative, whereas the dimension of **arousal** ranges from calming or soothing to exciting or agitating. 让专业摄影师去描述低分的challenge。
 - 结论：（？） themes in the left quadrants of the arousal-valence plane (see Figure 6) bias the aesthetic judgments towards smaller scores. themes会使美学判断的分数偏低。
2. （右图）可以发现方差为2左右的challenge的个数是最多的。对于free study challenges（那些不限定主题的图），通常low-variance score distributions；challenge的描述越详细，variance越大，参与challenge会花更多的心思在“如何使照片符合主题”上

一些数据集使用技巧

1. 训练数据越多越好
2. 适当删掉一些训练数据，指ambiguous image，也就是分数在5分左右的图，不好打分的图。只在训练数据中删，不在test集中删。可以适当提高网络的特征提取能力。
3. 先按内容分类，后打分
4. 高分图都使用了相对一致的摄影技巧；对于不同style的照片，需要不同的特征来描述