

## 101 | The Web 2018论文精读：如何对商品的图片美感进行建模？

2018-05-25 洪亮劫

AI技术内参

[进入课程 >](#)



讲述：初明明

时长 08:01 大小 3.67M



“万维网大会”（The Web Conference 2018）前身叫作“国际万维网大会”（International World Wide Web Conference），从 1994 年开始举办，已有 20 多年的历史了，在 Google 学术排名上，是“信息系统”排名第一的国际顶级学术会议。

从万维网大会最初举办开始，这个会议就成为了互联网方面独一无二的权威学术会议。会议包含搜索、推荐、广告、数据库、信息提取、互联网安全等诸多领域的优秀论文，每年都吸引着上千名世界各地的学者和工程师来分享他们的最新研究成果。

2018 年的万维网大会于 4 月 23 日~27 日在法国里昂举行。整个会议收录了 171 篇论文，还有 27 个研讨班（Workshop）、19 个讲座（Tutorial）、61 个展板论文（Poster）和 30 个演示（Demo）。

万维网大会的一大特点就是论文成果涵盖了非常广的领域。要在这些论文中找到有价值的学习信息是一件非常耗时、很辛苦的任务。这里给你分享几篇我认为今年这个会议上最有价值的论文，希望能起到抛砖引玉的作用。

今天，我们就来看一篇优秀论文提名，题目是《基于美感的服装推荐》（[Aesthetic-based Clothing Recommendation](#)）。这篇论文一共有六位作者，除了两位分别来自新加坡国立大学和美国的埃默里大学之外，绝大多数作者都来自清华大学。

## 论文的主要贡献

在现代的电商推荐系统中，商品特别是服装服饰的图片，其美观和质量是用户进行购买决策的关键参考因素。不少过去的商品推荐系统已经考虑了图片的属性，特别是尝试同时利用图片信息和文字信息来实现**多模（Multi-Modal）数据理解**的目的，从而能够进行更加智能的推荐。不过，当前的大多数方案都只是考虑基本的图片特性。

从思路上来说，大多数的类似工作都是利用某种深度神经网络提取图片特性，然后和其他特性（例如我们说过的文本信息）加以组合，从而能够扩宽我们对商品信息的提取。这样提取出来的图像特征自然没有显式地对图像的“美感”（Aesthetic）进行建模。

这篇文章的作者们认为，商品图片的“美感”是非常重要的属性，针对美感进行建模会有更显著的商品推荐效果。概括来说，这篇论文的一个贡献就是提供了一种模型，来对图片的美感和一般性的图片语义特性同时进行建模。这是一个在过去的工作中都没有的创新点，我们接下来会详细说明——这个模型的架构。

当作者们提取出了图片的美感信息以后，接下来的一个问题就是如何利用这些特性。这篇论文使用了**张量分解（Tensor Factorization）**的思路。我们在前面介绍推荐系统的时候曾经提到过，张量分解是一种很有效且常用的**利用上下文语义信息**的推荐模型。和一些之前的工作类似，这里作者们采用了三维的张量来表达用户、商品和时间之间的关系。同时，作者们还把图片信息有效地结合到了张量分解中，从而能够利用美感信息来影响推荐结果。

## 论文的核心方法

了解了这篇论文的大体思路以后，我们现在来看看论文的第一个核心部件：**如何利用深度神经网络来提取图片的美感信息？**

首先，这篇论文提出的模型假设对于每一个商品，我们都有一个**综合的美感标签**，并且还有一个**细节标签**来表达这个商品图案的“**图像风格**”（Style）。美感的综合标签是一个 1~10 的打分，而图像风格则是文字的图像特征，比如“高曝光”、“对比色”等。那么，我们需要一个神经网络模型，来同时对美感标签和细节的图像风格进行建模。

具体来说，文章提出的模型分为了两个层次。第一个层次是用来解释细节的图像风格。在本文采用的数据中，一共有 14 种图像风格，作者们就用了 14 个**子网络**（Sub Network）来针对这些风格。每个风格都对应一个独立的**子神经网络**。每一个子神经网络都是标准的“**卷积网络**”（CNN）。他们的目标是尽可能地学习到特性来表示每个细节的图像风格。

当我们有了第一层的 14 个子网络之后，再把这些子网络学习到的特性都整合起来，形成**中间特性层**，然后再经过一个卷积网络，从而学习到一个对商品的整体美感评分进行解释的神经网络。

在文章中，作者们提到这两个层次的神经网络并不是分类进行训练的，而是在**一个整体中进行训练**。意思就是说，我们同时训练底层的针对图像风格的 14 个子网络的参数，以及高层次的针对美感评分的网络的参数。

当我们得到了图片的美感信息之后，下一步，就来看一下**如何利用张量分解来进行商品推荐**。

相比于传统的张量分解，在这篇文章中，作者们提出了一种新颖的，针对商品推荐的张量表达模式，叫作“动态协同过滤”（Dynamic Collaborative Filtering），或简称 **DCF**。

DCF 认为，每一个用户对于某个商品的购买取决于两个方面的因素。第一，用户是否对这个商品有喜好。第二，这个商品是不是符合时间维度上面的“流行度”。作者们认为，只有当这两个条件同时满足的时候，也就是用户喜欢某个当季的商品时才会做出购买的决定。因此，作者们使用了**两个矩阵分解**来分别代表这两个假设。

第一个矩阵分解是针对用户和商品这个矩阵，这里我们会学习到用户对商品的**喜好度**。第二个矩阵分解是针对时间和商品这个矩阵，这里我们会学习到时间和商品的**流行度**。然后，作者把这两个矩阵分解（或者说是把两个矩阵）相乘，这就得到了一个张量，来表达**用户在时间维度上对商品的喜好**。

那么，如何把刚才学习到的图片美感信息给融入到这个新的张量学习框架下呢？作者们是这么做的，针对我们刚才所说的两个矩阵分解进行“扩展”。

刚才我们说，这个张量分解是基于一个假设，那就是用户在时间维度上的购买决定取决于，用户是否对这个商品有喜好，以及这个商品是不是符合时间维度上面的“流行度”。我们用了两个矩阵分解来表达这两个假设。每一个矩阵分解都是把一个大的矩阵分解成两个向量，比如用户和商品的矩阵就被分解为用户特性和商品特性。

基于此，作者们就在这个用户和商品的矩阵后面，再加上一个商品和图片美感信息矩阵，用来混合这两种信息。也就是说，我们刚才的第一个假设，用户对商品的好感，就被扩展成了**两个矩阵的加和**，用户和商品矩阵以及商品和图片信息矩阵，这两个矩阵的加和依然是一个矩阵。同理，时间和商品的流行度，被扩展成了时间和商品矩阵以及商品和图片信息矩阵的加和。也就是说，新的模型是两个矩阵的乘积组成的张量分解，而这里的每个矩阵分别又是两个矩阵的加和。这就是作者们最终提出的模型。

## 方法的实验效果

作者们在亚马逊的衣服数据集上做了实验来验证模型的有效性。这个亚马逊的数据集由将近四万的用户、两万多的商品和超过二十七万的购买信息构成。除了这篇文章提出的模型以外，作者们还比较了一些其他算法，例如完全随机的算法、只推荐最流行的商品、传统的矩阵分解模型以及只有基本图像信息但没有美感信息的算法。文章汇报了排序的精度 NDCG 以及“召回”（Recall）等指标。

从实验效果来看，这篇文章提出的模型要明显好于矩阵分解以及只有基本图像信息的算法，表明针对产品的图像美感进行建模是有价值的。并且，作者们提出的新的张量分解方法也被证明是切实有效的。

## 小结

今天我为你讲了今年万维网大会的一篇优秀论文。文章介绍了如何对商品的图片美感进行建模，以及如何把提取到的信息融入到一个基于张量分解的推荐系统中。

一起来回顾下要点：第一，我们详细介绍了这篇文章要解决的问题以及贡献；第二，我们简要地介绍了文章提出方法的核心内容；第三，我们简单分享了一下模型的实验成果。

最后，给你留一个思考题，有没有在没有标签情况下对图片的美感进行建模的呢？



欢迎你给我留言，和我一起讨论。



# AI 技术内参

你的360度人工智能信息助理

洪亮劼

Etsy 数据科学主管  
前雅虎研究院资深科学家



新版升级：点击「👤 请朋友读」，10位好友免费读，邀请订阅更有**现金**奖励。

© 版权归极客邦科技所有，未经许可不得传播售卖。页面已增加防盗追踪，如有侵权极客邦将依法追究其法律责任。

上一篇 100 | 文本情感分析中如何做意见总结和搜索？

下一篇 102 | The Web 2018论文精读：如何改进经典的推荐算法BPR？

## 精选留言

💬 写留言

由作者筛选后的优质留言将会公开显示，欢迎踊跃留言。