

A Statistical Approach to Material Classification Using Image Patch Exemplars

摘要:

本文主要研究从不同光照，视角下的单一图像的材质分类。材料的分类可以由像素点周围（比如说 3×3 的像素块）的联合密度分布来实现，且这种方法比滤波器的效果好得多。我们改进了基于基元的表示使其更适用于马尔可基夫随机场的邻居分布的建模。我们从训练集中获取这种特征表示然后用测试集进行分类。我们提出了 3 中这样的特征表示并与使用滤波器组的结果进行比较。这一分类效果比 Leung 和 Malik, Cula 和 Dana 等使用的滤波器组分类效果更好。我们也把我们的方法在 UIUC, Microsoft Textile 数据集上实验过。为什么基于邻居的特征能更好地区分有大量结构的纹理，为什么滤波器组的效果没有基于块的方法更好，我们将给出结论。

关键词：材质识别，3D 纹理，纹理基元，图像块，滤波器组

引言:

本文我们的目标是从不同光照，视角下的单一图像进行材料分类。这一课题的难点就在于图像同类之间的差异大，异类之间的差异小这一问题，没有合适，鲁棒的模型能满足这些变化。如果图像的外在条件未知则使这一问题更困难。

早期对纹理分类的研究主要集中于纹理二值图像的模式识别。随后，这种纹理分类加入了合成的 2D 变换。由于光照和拍摄角度的变化，这一分类随后就被 3D 变换的现实纹理的分类取代。现在的研究主要集中于现实生活中所有纹理的分类，而不仅仅是特定的材料分类。另一研究趋势是如何发现纹理的规律信息。

纹理分类的平常问题已经被基于滤波响应的方法很好的解决了，由于这一问题变得更复杂，通过建立更多的滤波相应的表示也能很好的解决。通常是使用大量的，多尺度的滤波器组来提取纹理特征。

本文我们质疑了滤波器组在纹理分类上的绝对统治作用。我们使用可选择的基于像素周围密度的联合分布来表示图像块。

我们首先研究了基于块的图像表示的长处。[A Statistical Approach to Texture Classification from Single Images] 中提出的 VZ 算法使用 MR8 滤波器取得了很好的 3D 纹理分类效果。这一方法取得了一下两种效果：1，使用周围邻居像素（最小从 3×3 大小）的方法取得了很好的分类效果；2，对任何固定大小的邻居，图像块比滤波器组取得了更好的效果。图像块表示的方法对 CURET 图像库的分类取得了比 MR8 滤波器组更好的效果。同样我们也在 UIUC, San Francisco 和 Microsoft Textile 图像库上做了实验。

为什么小的图像块可以更好的区分纹理，是否滤波响应特征比合成的图像块更适合纹理分类呢？我们将讨论这两个问题。最后，我们发现纹理经合成和去噪后的结果更加证明了这一新的特征表示可以从高维的图像块中学习到。

研究背景:

纹理上的研究一般集中在 5 个领域:

1, 合成 2, 分类 3, 分割 4, 压缩 5, 纹理的形状。

前 4 个通常使用小波变换和滤波器组，小波变换更多用在纹理压缩上，滤波器组则在纹理合成，分类，分割上用的比较多。

这几个领域的成功主要源于滤波响应的完全的统计表示的学习。这一完全表现在：1，滤波响应分

布的学习，相比低阶矩的分布；2，联合分布（共生）滤波响应的学习，相比每个滤波器的独立分布；3，相比以前的研究在多尺度多方向上使用更多的滤波器。

这些滤波响应是从训练集中获得然后通过聚类 and 直方图来表示，然后利用这些分布来进行分类，分割和合成。比如说：分类可以通过将新图像的分布和从训练集学习到的模型比较；同样地，合成意义通过建立一个与目标纹理有同样分布的纹理来实现。就其本身而言，滤波器组的使用无所不在且不成问题。

然后，尽管有很多先验的经验和证据表明滤波器组和小波能取得很好的效果。

1. CURET数据库：

CURET图像库包含了61种材料的图像，其中很多都是我们日常能看到的。他有粗糙的，奇特的，人造的等各种各样的纹理。

图像库中每种材料是在205中不同视角和光照下拍摄的，反射，阴影和其他一些因素的变化影响都能表现出来，这使CURET比起其他一些人的收集对于训练一个分类器更有挑战性。

尽管CURET数据库已经成为一个标准并被广泛应用于评价分类的效果，它仍然有一定的局限性。这主要与图像拍摄的方式和纹理的选择有关。对于前者，大多数图像没有显著的尺度变换且只有少数的平面旋转。对于纹理的选择，最显著的缺点是只有极少数的材料同一纹理有多种实例，因此同类间的差距不能被完全研究。因此很难形成一个标准，但CURET始终是用于纹理分类最大最全的数据库之一。

所有的61种材料将在本文第4,5节中的实验用到。每种材料有118张图片，方位视角都小于60度。我们从每种材料中挑选92张出来，然后从中心点截取一个200*200像素大小的块出来作为我们实验所用的图像。对选中的块我们将其转化成灰度然后归一化成均值为0，标准差为1。因此，我们的实验没有使用颜色信息，我们使图像对仿射变换和光照具有不变性。我们挑选后的CURET图像库有61*92=5612张图像。其中每个类随机选择46张作为训练集，其余的46张作为测试集。

VZ 分类器回顾：

分类问题可以这样解决：给定一张从不同光照和视角的纹理图像，把它分到一个提前知道的纹理类别中。[Representing and Recognizing the Visual Appearance of Materials Using Three-Dimensional Textons] 一文中提出了很多模型，滤波响应基元，卡方距离下的最近邻分类等等。随后的算法，比如 [3D Texture Recognition Using Bidirectional Feature Histograms] 中提出的BFH分类器和 [A Statistical Approach to Texture Classification from Single Images] 中提出的VZ分类器，将被本文使用。[Statistical Estimation of Histogram Variation for Texture Classification], [Class-Specific Material Categorisation] 和[On the Significance of Real-World Conditions for Material Classification] 中保留了MR8滤波器组但用SVM或高斯-贝叶斯分类器代替了最近邻分类器。

VZ分类器算法可以分成两个步骤：训练阶段，通过建立训练图片的滤波响应的统计描述来学习纹理模型；分类阶段，新的图像通过与学习到的各个类的模型进行比较来分类。

在训练阶段，训练图片通过与选定的一个滤波器组进行卷积运算来生成滤波响应。对每个纹理类别我们将他们的滤波响应进行聚类。这些类中心叫做基元，所有的基元构成一个词典。对于每一张训练图片我们可以通过标记每个像素与哪个基元最近来学习到一个模型。这个模型是每个基元归一化的频率直方图，每个纹理类别是由它的训练图片的模型表示。

在分类阶段，我们用这些学习到的模型来将一张新的图片分到1-61类别中。过程如下：先生成图像的滤波响应，然后对每个像素用基元词典的基元标记。然后将生成的基元频率直方图归一化，这样就生成了图像的S-vector。最后，我们使用最近邻分类算法来将其分到相应的类别中。我们用卡方距离来计算两个频率直方图的距离。

1. 滤波器组：

MR8滤波器组有38个滤波器但只有8个滤波响应，它包括一个高斯，高斯-拉普拉斯(LOG) 滤波器，一

个6个方向和3种尺度的边缘滤波器和bar 滤波器。高斯和高斯-拉普拉斯(LOG) 滤波器的滤波响应我们直接用，而每个尺度上的edge 和bar 滤波响应我们只用所有方向的最大值。 这样我们只得到8个滤波响应，而且它们都具有旋转不变性。MR4滤波器组只使用一个尺度，MRS4通过选择既在方向又在尺度上的最大值来确保旋转和尺度不变性。

2. 预处理：

在学习与分类之前，我们使用一下预处理手段：

3. 详细步骤：

我们从每个纹理类别中随机选取13张图片用K-means算法来聚类生成基元词典。我们从CURET数据库中的61个类别，每个类设定10个基元，这样我们的基元词典一共就有 $61*10=610$ 个基元。

在此基础上，当把2806张测试图片分到61个类别的时候，VZ分类器在MR8滤波器组上达到了96.93%的正确率。MR8滤波器组在每个类40个纹理基元的时候达到最好的97.43%的正确率。

基于块的分类器：

本节我们介绍使用原图像中的图像块