

A Survey of Perceptual Texture Similarity Learning

Chuansong Li¹, Ying Gao², Lin Qi², Feng Gao², Yan Liu³

¹Qingdao Tax Service, State Taxation Administration, Qingdao Shandong

²Department of Computer Science & Technology, Ocean University of China, Qingdao Shandong

³Information Technology Department, Qingdao Vocational and Technical College of Hotel Management, Qingdao Shandong

Email: gaoying@stu.ouc.edu.cn

Received: Dec. 13th, 2019; accepted: Dec. 30th, 2019; published: Jan. 6th, 2020

Abstract

As a basic attribute of the surface of the object, the texture contains rich image information such as color and texon. People use perceptual similarity to measure the similarity between different textures. Texture similarity is widely used in texture recognition and material recognition, which is one of the key technologies of object recognition and scene understanding. Accurate prediction of texture perception similarity can help visual tasks such as texture retrieval and texture labeling to keep consistent with the results of human perception. Reliable similarity data can be obtained through psychophysical experiments, and researchers usually estimate the texture similarity by using the distance measure between the features of the textures. This paper focuses on texture perceptual data acquisition, texture computational feature extraction and texture perceptual similarity estimation, reviews the development of texture perceptual similarity learning and common processing methods, and analyzes the future development of texture perceptual similarity learning combining with the convolutional neural networks.

Keywords

Texture Similarity, Perceptual Similarity, Perceptual Data, Feature Extraction

纹理感知相似性学习综述

李传松¹, 高颖², 亓琳², 高峰², 刘焱³

¹国家税务总局青岛市税务局, 山东 青岛

²中国海洋大学信息科学与工程学院, 山东 青岛

³青岛酒店管理职业技术学院信息工程技术学院, 山东 青岛

Email: gaoying@stu.ouc.edu.cn

收稿日期: 2019年12月13日; 录用日期: 2019年12月30日; 发布日期: 2020年1月6日

摘要

作为物体表面的一种基本属性，纹理图像包含了纹理颜色、纹理基元等丰富的图像信息。在计算机视觉研究领域，人们使用感知相似性来度量不同纹理之间的相似程度，研究人类对纹理图像的视觉感知。纹理相似性度量广泛应用于纹理识别和材质识别，是对象识别和场景理解的关键技术之一。可靠的感知相似性数据可以通过心理物理学实验获得，研究人员通过计算特征之间的距离度量估计纹理感知相似性。本文重点从纹理感知数据获取、纹理计算特征提取和纹理感知相似性估计三个方面回顾了纹理感知相似性学习的发展和常见的处理方法，并结合卷积神经网络对纹理感知相似性学习的对未来发展趋势作了分析。

关键词

纹理相似性，感知相似性，感知数据，特征提取

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

纹理泛指物体面上的花纹或线条，纹理图像(如图 1)广泛存在于世界上，是人类视觉系统最常接收和处理的视觉信息对象。纹理图像包含了纹理颜色、纹理基元等丰富的图像信息，在人们对图像进行分析与理解的时候，提供了十分重要的内容信息，因此，纹理图像分析被广泛应用于物体识别，材料识别与场景识别中。纹理图像包含的丰富的图像信息，很多抽象的信息很难用直观的语言进行描述。随着对纹理图像研究的不断深入，研究人员试图将纹理图像与人类视觉感知相结合，使计算机能够像人类一样去感知纹理图像所包含的信息。探索人类视觉感知纹理图像的研究工作主要集中在纹理感知属性和纹理感知语义这两个方面。纹理感知属性是人们对纹理图像的主观认知，通常是较为抽象的图像信息，例如纹理粗糙度、纹理方向性、纹理光泽度、纹理密度等。而纹理感知语义则是人们对纹理图像的描述，主要包括气泡状的、网格状的、螺旋状的、蜂巢状的等，都是比较具象的描述词。在计算机视觉领域，人们使用感知相似性来衡量不同的纹理图像之间的差异。



Figure 1. Typical texture images

图 1. 常见的纹理图像

关于纹理感知的研究最早可以追溯到 20 世纪六十年代, Julesz [1]使用数字计算机随机显示的图片进行了视觉辨别实验, 研究在自发的视觉过程中, 被试如何观察和理解并排在一起的具有相同视觉属性(包括亮度, 对比度, 颜色等)的两幅图像。近年来, 国内外研究人员在纹理图像视觉感知领域都有了很大的进展。通过一系列的心理物理学实验, 研究人员获得了人们视觉感知纹理图像的直接数据, 通过对数据进行分析整理, 找寻人类理解图像的普遍规律。目前比较常见的研究纹理相似性的心理物理学实验主要包括自由分组实验、感知特征评分和成对比较实验。心理物理学实验对被试、实验环境、实验时间的要求都比较高, 想要获得准确可靠的心理物理学实验结果, 需要耗费大量的时间与人力。如何有效的利用实验获得的数据, 对没有标注的纹理图片进行准确有效的标注是十分值得研究的一个问题。目前, 解决上述问题的主要方法是设计合适的算法对图像进行特征提取, 对所提取的图像特征进行分析和处理, 让计算机代替人们对未标注的图像进行标注。

随着机器学习和深度学习的兴起与发展, 深度特征, 即通过深度网络提取出来的纹理图像特征在很多图像分析任务中都取得了很好的效果, 深度特征在一定程度上提高了纹理感知相似性估计的准确率, 但是距离预期目标仍有较大的差距。本文从纹理感知数据获取、纹理计算特征提取和纹理感知相似性估计三个方面回顾了纹理感知相似性学习的发展和常见的处理方法, 并对纹理感知相似性学习的未来发展趋势作了分析。

2. 纹理感知数据的获取

心理物理学(Psychophysics)是研究人类对物理刺激(Physical Stimulus)与该刺激所引起的感觉在数量化方面的研究的心理学领域。作为心理学的一个分支, 心理物理学学者认为, 人类的一切感觉, 包括视觉、听觉等等, 都可以通过感觉强度与刺激强度之间的关系进行描述。纹理感知主要研究不同的纹理图像对人类视觉所产生的刺激与影响, 即人类视觉系统如何描述和感知纹理图像的。纹理感知在场景理解和数据可视化等领域具有重要的意义。该领域的研究工作最早是由 Julesz [1]等人展开的, 他们使用数字计算机随机显示的图片进行了视觉辨别实验, 研究在自发的视觉过程中, 被试如何观察和理解并排在一起的具有相同视觉属性(包括亮度, 对比度, 颜色等等)的两幅图像, 并分析他们观察图像时所用的空间统计特征。早期研究人员在进行心理物理学实验的时候, 使用的研究图像大多为点、线或者字母等微小图案作为纹理基元, 进行随机或规则排列形成的人造纹理图像。

随着人们对纹理感知研究的不断深入, 包括 Brodatz [2], OuTex [3], CURET [4]在内的很多自然纹理数据库(如图 2)被建立了起来, 研究者开始针对特定的纹理图像进行视觉感知机制的研究。

案例一: Tamura [5]在 Brodatz 纹理数据库的基础上设计了心理物理学实验, 要求被试对纹理图像进行感知评分, 分别从粗糙度, 方向性, 规则性, 粗糙度, 对比度和线状或者团状这六个感知属性的特征, 进行相似性判断和感知量评分。研究者针对被试对不同纹理图像的感知打分情况进行了纹理图像属性的分析, 研究了通过感知量评分进行纹理相似性分析的方法。通过与人类对纹理图像的感知相似性的判断结果进行对比分析, 检验了上述六种感知特征对于纹理感知相似性的描述能力。

案例二: Amadasun [6]同样在该数据库上进行了类似的实验, 定义了包含粗糙度, 对比度、复杂性和纹理强度在内的五种不同的纹理图像的属性, 要求被试对每幅纹理图像的纹理属性进行量化, 针对每一种纹理属性定义了近似的计算形式, 通过纹理相似性度量的结果来研究人类对对纹理图像进行视觉感知的程度。由于人类视觉感知是一个非常复杂的过程, 虽然该研究取得了一定的成果, 但并没有达到预期效果。

案例三: Rao [7]提出了一种全新的实验思路, 即自由分组实验。自由分组实验在 Brodatz 数据库上完成, 实验要求被试再无任何提示的情况下, 将数据库中的 56 幅图像根据自己的理解进行分组, 分组的个

数没有限制，每组的图像数量也同样没有限制。被大多数人分在了同一组的图像被认为具有高度的相似性，同时，没有被分在一组的图像被认为不具有相似性。通过对所有参与实验的被试的分组结果进行汇总分析，可以得到所有图像两两之间的相似性，构建了 56×56 的相似性矩阵。研究人员对得到的纹理感知相似性矩阵进行了降维分析，最终得到了三维的有效感知空间。研究人员认为，三维的感知空间的三个维度，分别对应了纹理图像最为重要的三个感知特性，分别是重复性、方向性和复杂性。尽管自由分组实验可以获得非常可靠的纹理感知相似性数据，但实验耗时较长，且无法在大数量的样本上进行实验。

案例四：Clark [8]提出了一种基于自由分组实验的等距离映射算法，该实验在纹理数据集 Pertex 上进行。在进行自由分组得到相似性矩阵之后，使用等距离映射算法 Isomap，保留相似性矩阵中距离比较相近的点，忽略距离比较远的点，计算每两个点之间的距离，从而实现了从纹理对之间的感知相似性到广义的感知相似性矩阵的转换。

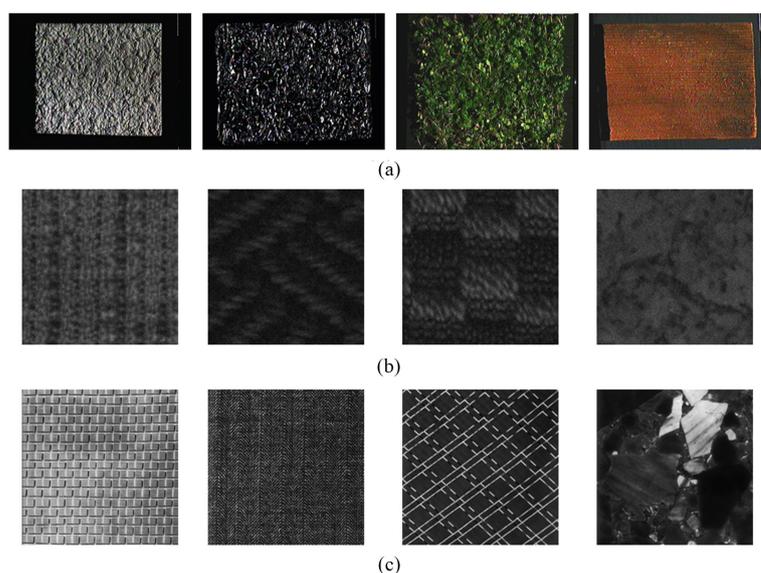


Figure 2. Texture samples of natural texture database. (a) is the texture samples from the CURET database, (b) is the texture samples from the Outex database, and (c) is the texture samples from the Brodat database

图 2. 自然纹理数据库纹理样本。其中图(a) 为 CURET 数据库的纹理样本，图(b)为 Outex 数据库的纹理样本，图(c)为 Brodatz 数据库的纹理样本

案例五：Liu [9]在过程纹理库上进行了纹理感知相似性数据获取的实验，使用了自由分组与合并分组的方法。实验中，20 名被试需根据自己的理解对 450 幅纹理图像进行分组，分组的个数不限，每组的纹理图像数目也没有限制。在第一次分组完成以后，要求被试对分好的纹理组进行合并，并给出每次合并的自信度。进行多次合并分组，直至被试认为剩下的组不能再次进行合并为止。将每次的分组结果与被试的自信度相乘并累加即可得到感知相似性的数据，该实验成功避免了上面提及的相似性数据稀疏的问题。图 3 展示了从过程纹理库中选取的纹理对其对应的感知相似性数据。其中 vSim 表示感知相似性值，该值通过心理物理学实验获得，实验中多名被试根据自己对图像的理解对两幅图像是否相似进行标注，根据所有被试的标注结果可以计算得到图像对之间的感知相似性值。

上述的感知评分实验和自由分组实验可以有效的帮助研究人员获得人类视觉感知属性的研究数据，实现了针对纹理图像中的抽象信息的分析与理解。研究表明，在检索实验中，感知描述符的表现要优于可视描述符。纹理感知建模在特征选择和纹理图像分析任务中都有很大的价值。

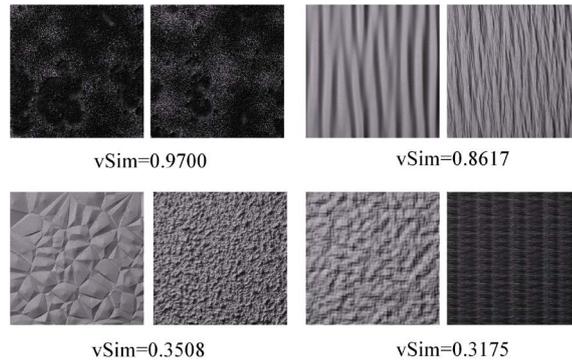


Figure 3. The paired texture of the procedural texture dataset and the corresponding perceptual similarity values, where vSim represents the perceptual similarity value

图 3. 过程纹理库的纹理对及相应的感知相似性数值，其中 vSim 表示感知相似性值

3. 纹理计算特征提取

纹理特征提取是成功进行纹理图像分割、纹理图像分类以及纹理图像检索等任务的最为关键的环节，提取特征的成功与否直接影响后续任务的完成情况。在提取纹理特征的时候，人们试图找寻能够代表纹理特性的度量模式，通过算法分析从中获取可以用于纹理描述的有效特征。特征提取有很多种方法，Mirmehdi [10]将纹理特征归为四个类别，分别是基于滤波器的特征，统计型纹理特征，结构型纹理特征和模型特征。

基于滤波器的特征：基于滤波器的特征提取方法主要是计算图像与滤波器进行卷积后的响应的能量或者方差。通过滤波器或滤波器组等等，对纹理图像进行变换，将纹理图像从空间域或者频率域转换到变换域，提取在纹理图像发生变换前后保持相对平稳的信息作为特征值。在选取滤波器的时候，可以使用罗伯特交叉算子，Prewitt 算子，Canny 算子，Sobel 算子，Shen-Castan 算子提取纹理图像中的线，边缘，点等等，也可使用正交滤波器，基于离散余弦变换的通道滤波器，Gabor 滤波器和小波变换等提取更多其他的图像信息。

统计型纹理特征：统计型纹理特征研究纹理区域中的统计特征，通常用来描述图像灰度值在空间分布的情况。局部统计一般是使用统计特征提取的纹理图像的信息，全局统计可以经过局部统计计算得到，这都可以用来表示纹理的特征。其中，灰度均值和灰度直方图都是最简单的表示图像一阶特征的形式，只需要统计像素值的分布。图像像素之间的关系可以通过自相关函数计算得到，作为二阶统计特征，以灰度共生矩阵(Gray-level Co-occurrence Matrix, GLCM)为代表，是建立在估计图像的像素及其邻域内灰度二阶组合条件概率密度基础上的一种方法。灰度共生矩阵所提取的纹理特征具有较好的鉴别能力，但是对于像素级的纹理识别任务具有一定的局限性，而且计算代价比较昂贵，耗时较长。

结构型纹理特征：结构型纹理特征是基于“纹理基元”进行纹理特征分析的一种方法。该方法认为纹理图像是由许多纹理基元构成，纹理基元的形状、方向、大小和数量的不同决定了纹理不同的表现形式。纹理基元的空间分布规律可以通过基元之间的几何关系和统计属性进行表示。结构型纹理特征中最为经典的一个特征提取方法是局部二值模式(Local Binary Patterns, LBP) [11]，该方法提取的是纹理图像的局部结构特征，反映了每个像素与周围像素之间的关系，在纹理分析中被广泛应用并且取得了很好的实验效果。

模型特征：模型型纹理特征主要是用一些纹理模型去描述纹理，使用算法提取纹理图像特征的时候，研究人员假设纹理图像中的基元分布是由分布式模型产生的，分布式模型可通过具体的参数设置进行控

制。因此，可以通过分析纹理图像来估计产生该纹理的计算模型的参数，将计算得到的参数作为纹理图像的特征，即可进行后续的分析与实验。纹理模型特征的提取方法主要包括随机场方法和分形方法，用来描述纹理的模型主要包括分形模型，自回归模型和马尔科夫随机场模型等等。

几种特征提取的方法都是手工设计提取的特征，以 Gabor 和 LBP 为例，目前已经在物体识别和图像分类等领域取得了很好的实验效果，但是由于不同的任务需求对特征的要求不同，同样的纹理图像特征在不同的任务中会表现出一定的差异性，想要取得较好的实验结果需要对手工特征进行进一步的调整和参数设置。研究人员希望能够找到一种鲁棒性的纹理表示方法，使得提取出来的纹理特征可以同时适用于多种纹理分类任务，通过深度学习方法所提取出来的深度特征就是这样的一种纹理特征。

深度特征主要是指通过多层卷积网络提取的特征，是一种对数据的高级表示形式。传统的机器学习技术在处理原始形式的纹理图像数据时候表现能力有限。深度学习是机器学习研究中的一个新领域，其动机在于建立并模拟人类大脑进行分析学习，实现模仿人脑的机制来分析和理解包括图像、声音和文本在内的各种数据。深度学习一般使用多层的网络结构，其中每一层网络结构都由多个非线性的模块构成。每一个模块负责将该层的输入数据通过变换转换为更高层、更抽象的表示，并将其输出结果作为新的输入数据，传送到网络的下一层再次进行变换。原始图像数据被输入网络之后，经过逐层处理，不断变换，最后被输出为一种抽象的特征。深度学习的常用模型包括自动编码器(AutoEncoder)、稀疏编码(Sparse Coding)、限制波尔兹曼机(Restricted Boltzmann Machine, RBM)、深信度网络(Deep Belief Networks, DBN)和卷积神经网络(Convolutional Neural Networks, ConvNet)等。

4. 纹理感知相似性估计

目前关于纹理图像感知相似性的估计工作大多围绕纹理图像的计算特征展开。Dong [12]评估了 51 种不同的计算特征在预测纹理感知相似性时的效果，通过使用距离度量的方法，计算不同纹理计算特征间的欧氏距离(Euclidean Distance)或者卡方距离(Chi-square distance, χ^2)，将计算得到的特征距离作为感知相似性的估计量。实验表明，计算特征并不能很好的拟合心理物理学实验获得的感知相似性的实验数据。研究人员尝试使用机器学习的方法和技术，试图建立纹理图像计算特征空间和感知空间之间的关系。他们通过心理物理学实验获得了纹理感知的实验数据，利用多维尺度分析及等距离特征映射等子空间变换技术，构建了纹理感知空间。随后，他们使用支持向量机(Support Vector Machines, SVM)等技术，建立起纹理计算特征空间到纹理感知空间的非线性映射[13][14][15]。实验证明，相比于使用距离度量的方法拟合纹理图像的感知相似性，构建纹理感知空间有助于提高纹理图像感知相似性的预测。但是在构建纹理感知空间的时候，感知空间的维度是需要经过大量的实验得到，且不同类型的纹理图像的感知空间可能有所差异，构建的纹理感知空间无法在不同的数据集上得到应用，具有一定的局限性。此外，在子空间变换中，变换的过程是不可控的，很有可能丢失部分真实感知空间中的有效信息。

卷积神经网络(Convolutional Neural Networks, CNNs)是人工神经网络的一种，在语音分析、图像识别和图像理解等领域都取得了不错的实验效果。卷积神经网络的出现受到了动物视觉感知系统的启发，研究人员通过对动物视觉皮层细胞的研究，提出了感受野(Receptive Field)和神经认知机(Neocognitron)，试图将视觉模型系统化，使计算机能够高准确率地进行物体识别。卷积神经网络是一个多层的神经网络，由多个阶段的可训练的结构构成，每一阶段的输入和输出被称为特征图(Feature Map)。如果输入一幅 RGB 彩色图像，那么每个特征图就会以代表每个颜色通道的二维数组的形式进行表示。在输出层，每个特征图表示从输入的所有位置上提取出的一类特定的特征。卷积神经网络可以直接接收图像作为网络的输入，相比于传统算法，省略了单独提取图像特征的过程，用于在识别图像的时候，对于平移、缩放等变换具有高度不变性。

一般的卷积神经通常由卷积层(Convolutional Layer)、非线性层(Non-linearity Layer)、特征池化层(Feature Pooling Layer)和全连接层(Full-connect Layer)组成。卷积层也称为滤波器组层(Filter Bank Layer),包含了多个可以学习的滤波器(卷积核)。每个滤波器都包含一个空间邻域(通常为一个较小的矩形)和对该邻域内的图像像素定义的操作,被用于提取图像局部区域内感兴趣的特征。非线性变换层主要负责对卷积之后的结果进行非线性激活,常见的激活函数有 Sigmoid 函数和 ReLU 函数等等。将上一层输出的特征图进行下采样,再通过激活函数处理得到下采样层的特征图,通过非线性激活函数的变换,模型能够更好地拟合非线性映射,更好地保留有效的图像特征。卷积神经网络的上述特点,使其能很好地应用于图像分析等任务,因此在卷积神经网络的基础上构建了新的网络模型,也同样可以用来学习纹理图像的感知相似性。结构合理的端到端的卷积神经网络可以有效建立纹理图像与纹理感知相似性之间的关系,有效预测纹理图像的感知相似性。

5. 结束语

纹理包含纹理颜色、纹理基元等丰富的视觉信息,是一种重要的视觉线索,被广泛应用于包括纹理检索、纹理标注等多种视觉任务中。研究人员试图将纹理图像与人类视觉感知相结合,使计算机能够像人类一样去感知纹理图像所包含的信息。纹理感知相似性的预测工作得到了越来越多的关注。本文重点从纹理感知数据获取、纹理计算特征提取和纹理感知相似性估计三个方面回顾了纹理感知相似性学习的发展和常见的处理方法,希望能对相关领域的研究人员和工程技术人员有所裨益。在未来研究中,纹理感知相似性预测工作可以在特征提取方法上进行进一步的研究,设计鲁棒的纹理特征,并与深度学习相结合,构建端到端的卷积神经网络,有效建立纹理图像与纹理感知相似性之间的关系,更好地学习纹理感知相似性。

参考文献

- [1] Julesz, B. (1962) Visual Pattern Discrimination. *IRE Transactions on Information Theory*, **8**, 84-92. <https://doi.org/10.1109/TIT.1962.1057698>
- [2] Строениекани (1965) Textures: A Photographic Album for Artists and Designers. *Leonardo*, **1**, 91-92.
- [3] Ojala, T., Maenpää, T., Pietikainen, M., et al. (2002) Outex—New Framework for Empirical Evaluation of Texture Analysis Algorithms. *Object Recognition Supported by User Interaction for Service Robots*, **1**, 701-706.
- [4] Dana, K.J., Nayar, S.K., van Ginneken, B. and Koenderink, J.J. (1997) Reflectance and Texture of Real-World Surfaces. *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 17-19 June 1997, 151-157.
- [5] Tamura, H., Mori, S. and Yamawaki, T. (1978) Textural Features Corresponding to Visual Perception. *IEEE Transactions on Systems, Man, and Cybernetics*, **8**, 460-473.
- [6] Amadasun, M. and King, R. (1989) Textural Features Corresponding to Textural Properties. *IEEE Transactions on Systems, Man & Cybernetics*, **19**, 1264-1274. <https://doi.org/10.1109/21.44046>
- [7] Rao, A.R. and Lohse, G.L. (1996) Towards a Texture Naming System: Identifying Relevant Dimensions of Texture. *Vision Research*, **36**, 1649-1669. [https://doi.org/10.1016/0042-6989\(95\)00202-2](https://doi.org/10.1016/0042-6989(95)00202-2)
- [8] Clarke, A.D.F., Halley, F., Newell, A.J., Griffin, L. and Chantler, M. (2011) Perceptual Similarity: A Texture Challenge. In: Hoey, J., McKenna, S. and Trucco, E., Eds., *Proceedings of the British Machine Vision Conference*, BMVA Press, Guildford, 120.1-120.0. <https://doi.org/10.5244/C.25.120>
- [9] Liu, J., Dong, J., Cai, X., Qi, L. and Chantler, M. (2015) Visual Perception of Procedural Textures: Identifying Perceptual Dimensions and Predicting Generation Models. *PLoS ONE*, **10**, e130335. <https://doi.org/10.1371/journal.pone.0130335>
- [10] Mirmehdi, M., Xie, X. and Suri, J. (2008) Handbook of Texture Analysis. World Scientific, Singapore, 424. <https://doi.org/10.1142/p547>
- [11] Ojala, T., Pietikäinen, M. and Mäenpää, T. (2002) Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **24**, 971-987. <https://doi.org/10.1109/tpami.2002.1017623>

-
- [12] Dong, X. (2014) Perceptual Texture Similarity Estimation.
 - [13] Gao, Y., Wang, L. and Chan, K.L. (2005) Learning Texture Similarity with Perceptual Pairwise Distance. *Proceedings of the 4th International Workshop on Texture Analysis and Synthesis, in Conjunction with the 10th IEEE International Conference on Computer Vision (ICCV)*, 83-88.
 - [14] Long, H.Z., Leow, W.K. and Chua, F.K. (2000) Perceptual Texture Space for Content-Based Image Retrieval. In: *Multimedia Modeling*, World Scientific, Singapore, 167-180. https://doi.org/10.1142/9789812791993_0012
 - [15] Long, H. and Leow, W.K. (2001) Perceptual Texture Space Improves Perceptual Consistency of Computational Features. *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, **2**, 1391-1396.