

# 图像分割特征选取和分割方法综述

谢威宇

2020 年 1 月

## 摘要

图像分割是现今计算机视觉领域中最困难、最有意义的问题之一。高效的图像分割算法的提出对未来许多行业都有巨大的促进意义。本文介绍并详细分析了图像分割领域的特征信息选取方法和各种常用算法，包括无监督方法和有监督方法。介绍了方向梯度直方图、数据降维等提取图形特征信息的方法。同时对 K 均值算法、means-shift 算法、随机游走算法、支持向量机等算法做了详细介绍。

**关键字:**、图像分割、计算机视觉、数字图像处理

## Abstract

Image segmentation is one of the most difficult and meaningful problems in computer vision today. The proposal of an efficient image segmentation algorithm will have great promotion for many industries in the future. This paper introduces and analyzes in detail the feature selection methods and various commonly used algorithms in the field of image segmentation, including unsupervised methods and supervised methods. Methods of extracting graphic feature, such as histogram of oriented gradient and data reduction, are introduced. K-means algorithm, means-shift algorithm, random walk algorithm, support vector machine and other algorithms are introduced as well.

**Keywords:** Image Segmentation, Computer Vision, Digital Image Process

## 1 前言

图像分割，又被称为像素级分类，是把图像中的像素按照其语义类别进行分类的一个过程。近年来，图像分割已经渐渐成为计算机视觉、计算机数

字图像处理的潮流方向。在图像和视觉处理任务中，另外两个主要方向是图像分类和图像检测。图像分类指把图片按照其内容分到一类，图像检测则是检测图像中出现的物体并标记。它们工作的对象都是整张图片。而图像分割的任务是把每一个像素都分到它应在的类别中去。这些方向之间也会有结合的工作。图像分割有非常广泛的应用，例如检测路牌（Maldonado-Bascon 等 [7]），结肠息肉分割（Cohen 等 [2]）。图像分割也为自动驾驶领域提供了基石。对于自动驾驶来说，场景分析需要依靠图像的语义分割结果才能进行进一步的工作。

近年来，由于深度学习和神经网络的兴起，图像分割算法的精度得到了极大的提高。通常来说，我们把深度学习涌现之前的图像分割方法称为传统方法。本文将介绍着重介绍和比较这些方法。

## 2 图像分割传统方法

在深度学习以及神经网络被运用在图像分割之前，图像分割最重要的主题就是图像特征信息的提取和分类。在计算机视觉和数字图像处理领域，图像的特征信息就是指那些和解决图像问题有关的关键信息。这和机器学习还有模式识别领域内的特征并无二致。有许多多种多样的特征信息被用在图像分割的领域。

### 2.1 图像特征信息的选取

在图像分割领域，特征信息的选取对算法结果来说十分重要。这里介绍一些常用的全局特征和局部特征信息选取方法。

#### (1) 像素颜色

图片的像素颜色是最常用的图像特征信息。像素颜色有很多空间，例如 RGB 空间、HSV 空间和灰度空间。一个典型图像的颜色使用 RGB 空间来描述，但是如果使用其他颜色空间作为特征信息则可能获得更好的图像分割结果。RGB、YcBcr、HSL、Lab 和 YIQ 都是一些常用的颜色空间。没有任何一种颜色空间在所有算法上都优于其他颜色空间，这已经被证明了 [1]。然而在实践中最常用的颜色空间还非 RGB 和 HSI 莫属。这可能是因为 RGB 在各大编程语言和框架中都有较好的支持，使用 HSI 能够让分割结果少受亮度的影响。

### (2) 方向梯度直方图

方向梯度直方图 (Histogram of oriented gradient, HOG) 把图像当作一个离散函数  $I: \mathbb{N}^2 \rightarrow \{0, 1, \dots, 255\}$ , 这个函数把一个像素位置  $(x, y)$  映射到一个颜色上。它首先被 N. Dalal 等人提出 [4]。所以对于这样表示的图片, 每个像素都有  $x$  和  $y$  两个方向的梯度。这样这个图片就被转化成为了两个同尺寸的梯度图, 然后就可以作为图像分割的特征信息使用。

### (3) 尺度不变特征转换

尺度不变特征转换 (Scale-invariant feature transform (SIFT)) 可以在空间尺度中寻找极值点, 并提取出其位置、尺度、旋转不变数, 此算法由 David Lowe 在 1999 年所发表, 2004 年完善总结 [6]。后续的论文中也有许多基于 SIFT 改进的论文, 例如 SURF 将 SIFT 的许多过程近似, 达到加速的效果; PCA-SIFT 利用主成分分析降低描述子的维度, 减少内存的使用并加快配对速度。此特征信息应用范围包含物体辨识、机器人地图感知与导航、影像缝合、3D 模型建立、手势辨识、影像追踪和动作比对。

### (4) 图片降维

高分辨率的图片有着非常多的像素, 对高分辨率图片的特征提取通常会提取出数百万计的特征信息。这让算法的调试和训练变得十分缓慢, 并且更多的特征信息不一定就囊括了更多的特征。对此问题一个简单的解决方法就是缩放原分辨率的图片, 让它变成低分辨率的图片。另外一种方法是主成分分析 (Principal components analysis, PCA)。主成分分析的思路就是找到一个让所有向量投影之后信息损失最小的超平面。具体关于主成分分析的介绍请看 [10]。主成分分析的一个主要问题在于它有可能损失一些关键信息。

还有其他更多的降维的方法和技巧, 请参照 [12]。

## 2.2 基于阈值的图像分割

基于阈值的分割方法是最简易、最常用的方法之一。读入一副灰度图像, 根据设定的阈值进行过滤, 从而得到对图像的分割。图像若只有目标和背景两大类, 那么只需要选取一个阈值进行分割, 此方法成为单阈值分割; 但是如果图像中有多个目标需要提取, 单一阈值的分割就会出现作物, 在这种情况下就需要选取多个阈值将每个目标分隔开, 这种分割方法相应的成为多阈值分割。

## 2.3 无监督图像分割

图像分割领域的算法主要可分为两种类型：监督算法和无监督算法。虽然无监督算法不能真正提取语义信息，但是它可以提炼数据并将结果作为监督算法的输入。非语义的算法能检测图像中相近的区域和这些区域的边界。

### (1) 聚类算法。

只要输入特征信息，聚类算法就可以直接运用在图像上对像素进行分类。有两种聚类方法：K-means 和 mean-shift 算法。

K-均值聚类是无监督方法中的典型方法。给定特征空间数据和聚类数目，K-means 算法就能自动地分类出数个聚类。K-均值首先随机在特征空间中生成  $k$  个聚类中心，将点归类于和它最近的一个中心点。而后逐步地移动聚类中心使其更符合要求的，直到算法达到全局停止条件。

另外一个算法是由 D. Comaniciu 等人提出的 mean-shift 算法 [3]。基本的 mean-shift 聚类要维护一个与输入数据集规模大小相同的数据点集。初始时，这一集合就是输入集的副本。然后对于每一个点，用一定距离范围内的所有点的均值来迭代地替换它，并且可以使用任意核来计算均值。与之对比，k-均值把这样的迭代更新限制在（通常比输入数据集小得多的） $K$  个点上，而更新这些点时，则利用了输入集中与之相近的所有点的均值。Mean-shift 算法找到的聚类中心具有最高的局部点密度。

### (2) 基于图的图像分割。

基于图的图像分割算法是把图像转换成一个图来处理。通常，像素转换成图中的顶点，像素之间的关系（比如颜色的差距）转换成顶点核顶点之间带权重的边。这些边可以选择 4 领域的（上下左右）或者 8 领域的（加上对角线）。即使这样，这种边仍然太多。一种减少边数目的方法是通过移除一些权值很小的边，从而建立一个最小生成树。去掉边之后，剩下的连通分量就是图像的分割。

### (3) 随机游走。

随机游走算法是基于图算法的一种。通常这种算法分为以下步骤：首先先把种子点分布到图像上，用来表明不同的物体。对每个像素点，通过方向梯度直方图 (HOP) 计算它们随机游走到达每一个种子点的概率。这个像素将被分配到到达概率最高的种子点，完成图像分割。

### (4) 主动轮廓模型。

主动轮廓模型 (Active Contour Models) 是一种能按照给定的粗糙边界

将图像分割成具有光滑边界部分的算法。这通过降低预定能量函数的值来实现。这篇工作最开始提出了这个方法 [5]。

#### (5) 区域生长

区域生长 (Region Growth) 是从一组代表不同生长区域的种子像素开始, 接下来将种子像素邻域里符合条件的像素合并到种子像素所代表的生长区域中, 并将新添加的像素作为新的种子像素继续合并过程, 知道找不到符合条件的新像素为止, 该方法的关键是选择合适的初始种子像素以及合理的生长准则。

#### (6) 区域分裂合并

区域生长是从某个或者某些像素点出发, 最终得到整个区域, 进而实现目标的提取。而分裂合并可以说是区域生长的逆过程, 从整幅图像出发, 不断的分裂得到各个子区域, 然后再把前景区域合并, 得到需要分割的前景目标, 进而实现目标的提取。

#### (7) 分水岭分割。

分水岭模型 (Watershed Segmentation) 接收一副灰度图像作为输入, 并把它转化为高度图。高度较低的地方就是盆地, 盆地和盆地之间较高的地方就是分水岭。此算法从高度最低的像素开始灌水, 如果有两个分隔的盆地被连通, 那么就找到了一个分水岭。像这样灌水直到淹没整个图像, 就求出了所有的分水岭, 完成了图像分割。然而这个算法有两个问题, 一是许多局部最小值会导致太多的分割部分, 二是一个高原会导致很厚的分水岭出现。此算法的详细介绍请参照 [8]。

## 2.4 随机决策森林

随机决策森林最开始由何天琴等人提出 [11], 这种分类器使用了一种叫做集成学习的技巧。集成学习就是指多个分类器同时进行学习并且把它们的输出结合到一起。在随机决策森林的场景下, 分类器就是决策树。决策树是一棵在每个内部节点都有相关特征信息来判断选择哪一个后代、并且叶节点就是最终分类的树结构。相对于其他分类器 (SVM、NN) 来说, 随机决策森林可以处理任意尺寸的特征信息, 并且训练和分类速度较快。决策树及其扩展已经被很充分地研究, 并且有许多训练算法被提出。训练模型一些可能的超参数包括分割好坏的度量、决策树的数目和决策树的深度限制。一般来说, 在分类训练中, 新的节点被加到决策树中, 直到每一个叶子节点只包含一个类别。

对随机决策森林在图像分割中的应用，请参考此篇工作 [9]。

## 2.5 支持向量机

支持向量机 (SVM) 是一个已经被充分研究的二元线性分类器。给定一组训练实例，每个训练实例被标记为属于两个类别中的一个或另一个，SVM 训练算法创建一个将新的实例分配给两个类别之一的模型，使其成为非概率二元线性分类器。SVM 模型是将实例表示为空间中的点，这样映射就使得单独类别的实例被尽可能宽的明显的间隔分开。然后，将新的实例映射到同一空间，并基于它们落在间隔的哪一侧来预测所属类别。除了进行线性分类之外，SVM 还可以使用所谓的核技巧有效地进行非线性分类，将其输入隐式映射到高维特征空间中。

## 2.6 马尔可夫随机场

马尔可夫随机场 (Markov Random Fields) 是一种被广泛运用于计算机视觉领域的概率无向图模型。在图像分割中，将所有输入的特征信息和输出的像素信息分别作为一个节点，然后通过图上的概率运算得出最终的分割结果。这篇工作在医学图像中运用了马尔可夫随机场 [13]。

# 3 总结

本文从多个方面介绍并总结了图像分割领域的特征选取和常用传统算法。近年来基于深度学习和神经网络的方法也非常之多，未来需要对此类方法作一个详细的调查。图像分割作为三大计算机视觉的主要问题，现在还有许多关键问题仍待解决。例如对提高图片语义分割算法的可靠性和性能等。在未来自动驾驶、智慧医疗等热门领域，图像分割必将得到广泛的应用，解决各行各业的实际问题。

## 参考文献

- [1] H.D. Cheng, X.H. Jiang, Y. Sun, and Jingli Wang. Color image segmentation: Advances and prospects. *Pattern Recognition*, 34:2259–2281, 12 2001.

- [2] Assaf Cohen, Ehud Rivlin, Ilan Shimshoni, and Edmond Sabo. Memory based active contour algorithm using pixel-level classified images for colon crypt segmentation. *Computerized Medical Imaging and Graphics*, 43:150 – 164, 2015.
- [3] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5):603–619, May 2002.
- [4] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*, volume 1, pages 886–893 vol. 1, June 2005.
- [5] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, Jan 1988.
- [6] David G. Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2):91–110, Nov 2004.
- [7] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. Lopez-Ferreras. Road-sign detection and recognition based on support vector machines. *IEEE Transactions on Intelligent Transportation Systems*, 8(2):264–278, June 2007.
- [8] Jos B.T.M. Roerdink and Arnold Meijster. The watershed transform: Definitions, algorithms and parallelization strategies. *Fundam. Inf.*, 41(1,2):187–228, April 2000.
- [9] Florian Schroff, Antonio Criminisi, and Andrew Zisserman. Object class segmentation using random forests. 01 2008.
- [10] Jonathon Shlens. A tutorial on principal component analysis, 2014.
- [11] Tin Kam Ho. Random decision forests. In *Proceedings of 3rd International Conference on Document Analysis and Recognition*, volume 1, pages 278–282 vol.1, Aug 1995.

- [12] L.J.P. van der Maaten, E. O. Postma, and H. J. van den Herik. Dimensionality reduction: A comparative review, 2008.
- [13] YY Zhang, Michael Brady, and S.A. Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation maximization algorithm. *IEEE transactions on medical imaging*, 20:45–57, 02 2001.