

K均值聚类算法判断单目图像深度层次

张惠丽¹, 石 炜^{2*}, 耿宇聪²

¹包头职业技术学院电气工程系, 内蒙古 包头

²内蒙古科技大学机械工程学院, 内蒙古 包头

Email: yanstone8697@sohu.com

收稿日期: 2021年3月22日; 录用日期: 2021年4月16日; 发布日期: 2021年4月23日

摘 要

针对单目图像深度层次的判断问题提出一种基于K均值聚类算法判断单目图像深度层次的方法。通过图像增强和图像滤波的方式达到优化图像提高对比度, 保留细节特征以及滤除图像噪声的目的。根据K均值聚类算法, 通过将灰度值相似的像素聚类, 进行图像的分割, 从而实现深度层次的划分。本次实验选取不同数目的聚类中心进行深度层次的划分并将结果进行比较。实验结果表明, 利用K均值聚类的方法判断深度层次在一些几何场景中取得了比较好的成果。

关键词

K均值聚类, 单目图像, 图像分割, 深度层次

Determining the Depth Level of Monocular Image Based on K-Means Clustering Algorithm

Huili Zhang¹, Wei Shi^{2*}, Yucong Geng²

¹Department of Electrical Engineering, Baotou Vocational & Technical College, Baotou Inner Mongolia

²School of Mechanical Engineering, Inner Mongolia University of Science and Technology, Baotou Inner Mongolia

Email: yanstone8697@sohu.com

Received: Mar. 22nd, 2021; accepted: Apr. 16th, 2021; published: Apr. 23rd, 2021

Abstract

Aiming at the problem of judging the depth levels of monocular images, a method based on the

*通讯作者。

K-means clustering algorithm to judge the depth levels of monocular images is proposed. Through image enhancement and image filtering, the purpose of optimizing the image to improve the contrast, retaining the detailed features and filtering the image noise is achieved. According to the K-means clustering algorithm, the image is segmented by clustering pixels with similar gray values to achieve depth level division. In this experiment, different numbers of clustering centers are selected for in-depth division and the results are compared. Experimental results show that the depth level judgment method proposed in this paper has achieved relatively good results in some geometric scenes.

Keywords

K-Means Clustering Algorithm, Monocular Image, Image Segmentation, Depth Level

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

由于单目视觉系统的特点, 使得其图像的深度层次分析一直是图像处理的研究热点之一。单目机器视觉在众多领域中具有广泛的应用, 包括工业、农业、交通、畜牧业、海洋、勘测等各个方面。在众多的应用领域, 单目采集图像的分析处理引起各个学者的广泛研究, 包括目标跟踪与识别、人脸识别、场景理解、图像分类、视频检索等[1]-[6]。由于单目视觉影像内包含了比较丰富的像素信息, 对目标物的精细、精准分类及深度层次的分析提供了数据支撑和保障。在有关深度判断的计算机视觉研究中, 根据使用的摄像机数量分成单目, 双目以及多目, 双目立体视觉是比较常用的深度信息提取方法[7]。双目视觉利用两个不同摄像机成像时相同点的位置差异来估计距离, 这种方法对距离的估计精度较高, 但是计算量较大。因此, 单目图像深度估计的方式就被提了上来。但是, 单目图像深度估计是一个病态问题, 原因在于单张 RGB 图像对应的真实深度情况有无数千[8], 而且图像中并没有稳定的线索可以判断各种可能性。机器视觉和人眼视觉不同, 人类单眼也可通过视角大小、遮蔽关系、阴影关系、空气透明度以及眼肌肉的紧张程度等线索判断深度关系。对于采集到的图像中目标物是曲面, 或多目标的空间纵深位置关系, 是单目视觉的研究难点。因此如何充分合理利用单目图像来判断目标物的深度层次是亟待解决的问题。

王琪等采用 K 均值聚类算法对源图像进行粗分割, 然后采用形态学理论对粗分割结果图进行细化处理[9]。Wei 等提出一种基于粒子群的 K 均值聚类算法来对图像进行分割, 避免在寻优过程中陷入局部最优解[10]。朱永红提出了一种基于 K 均值聚类对陶瓷窑炉烧成带火焰图像进行彩色分割的方法, 同时给出了相应的 K 均值聚类算法[11]。以上学者主要针对彩色图像进行分割和修复, 对于图像的深度层次没有进一步的研究。而很多情况下采集图像主要是灰度图, 研究重点主要是精准分析目标物的各种参数及特征。因此本文根据图像不同深度的灰度值存在差异, 提出一种利用 K 均值聚类算法将图像依据灰度值的不同进行聚类分割, 由此获得二维场景深度信息, 进而进行单目图像深度层次的判断。

2. 图像预处理

2.1. 图像增强

由于受到场景条件以及拍摄的条件等因素的影响, 常常会导致拍摄的效果不佳, 此外, K 均值聚类

算法对于区分较为明显的区域分类效果更好。因此为了改善图像的效果,需要采取图像增强技术等措施。通过对比,本文采用直方图均衡化的技术来改善图像效果。这种方法主要通过改变灰度直方图来达到图像增强的目的,经过直方图均衡化处理使得像素尽可能多的占有灰度级并且均匀分布,通过将灰度值的动态范围增大,从而让图像的对比度增大。在本文的研究中,直方图均衡化可以增强不同区域灰度值的差异性,为后续依据灰度值的不同进行区域分割做好准备。图像增强效果如图 1 所示。图 1(a)为原始图像,图 1(b)为原始图像直方图,图 1(c)为增强后的图像,图 1(d)为增强后的直方图。

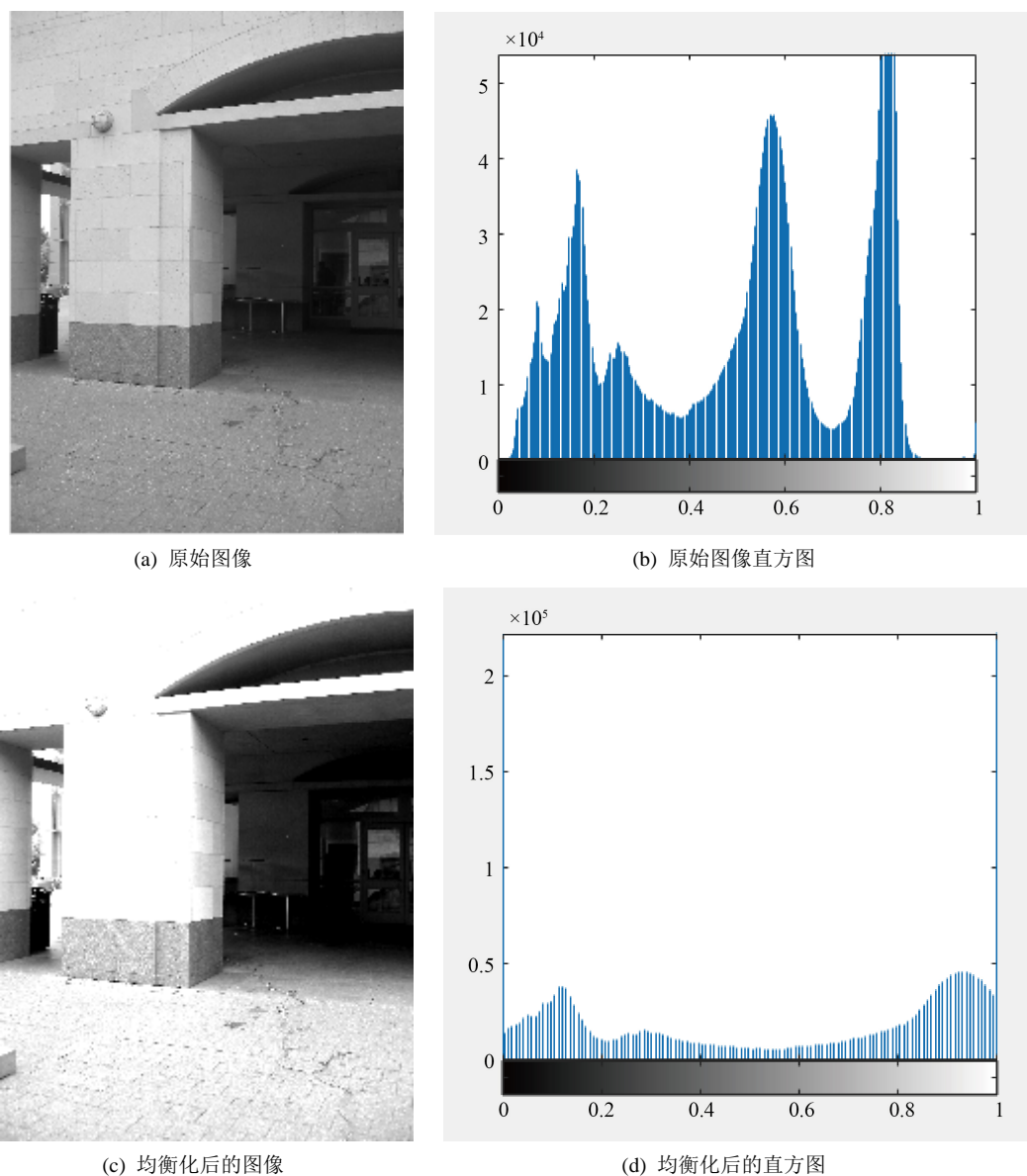


Figure 1. Image enhancement
图 1. 图像增强效果

2.2. 图像去噪

在获取图像后通常不能直接对图像直接开始进一步操作,主要因为其可能存在一些质量问题。图像

去噪操作可以将噪点去除,进而将图像的质量进行改进。 K 均值聚类算法由于其自身的原因,对于噪声比较敏感,孤立点对该算法的影响非常大,因此必须对图像去噪处理。由于噪声的来源有很多种,目前并没有去噪方法可以通用于所有的噪声,因此不同情况应当采取不同的滤波方式。中值滤波对于一些随机噪声具有非常理想的降噪能力,是一种非线性平滑技术,可以认为是一种统计排序滤波器。

中值滤波器以原图中的某一点 (m, n) 为中心的邻域内所有像素的统计排序中值作为 (m, n) 点的相应。因此一般的噪点经过中值滤波器处理后将会被去除,且因滤波器导致图像变得模糊的情况比较轻微。因此本文在研究中采用中值滤波的方式进一步优化图像。采用中值滤波器对图像处理的结果如图 2 所示。图 2(a)为原图,图 2(b)为采用 3×3 中值滤波模板的效果图,图 2(c)是采用 5×5 中值滤波的效果图。

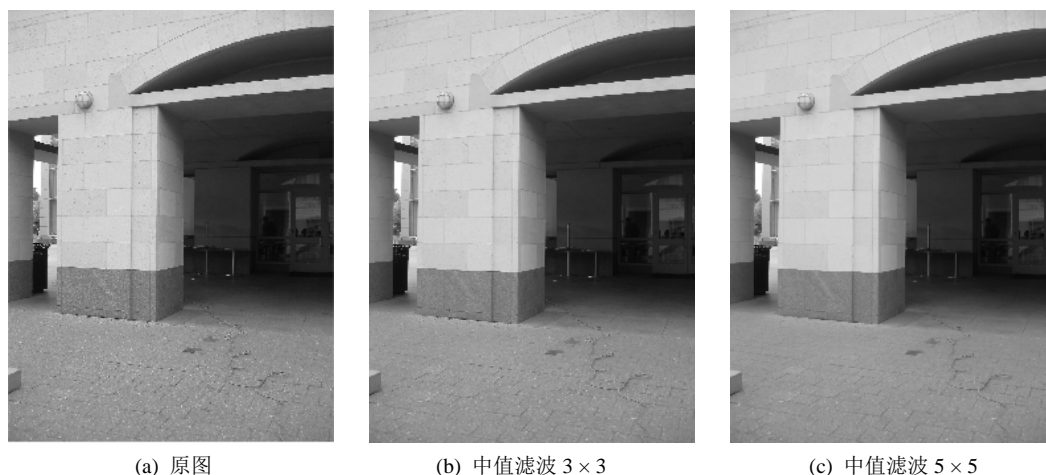


Figure 2. Mean filtering effect
图 2. 均值滤波效果

由图 2 可以看出中值滤波 5×5 模板的滤波效果较 3×3 的滤波效果更好,对于图像的细节保留也比较完整。但并不是模板越大越好,模板过大会不可避免的导致图像模糊程度加深,因此对于窗口的选择应当以实际情况作为标准。

3. K 均值聚类算法判断深度层次

3.1. K 均值聚类算法

均值聚类算法是一种迭代求解的聚类分析算法,是一种无监督学习方法,可以发掘关联规则。 K 均值聚类使用最大期望算法求解的高斯混合模型在正态分布的协方差为单位矩阵,且隐变量的后验分布为一组狄拉克 δ 函数时所得到的特例。该算法有简单快速效率高的特点,但只对于区分明显的区域分类效果好,对噪声敏感,因此本文前部分对于图像的预处理也是为解决该算法的缺陷所设计。

3.2. 算法流程

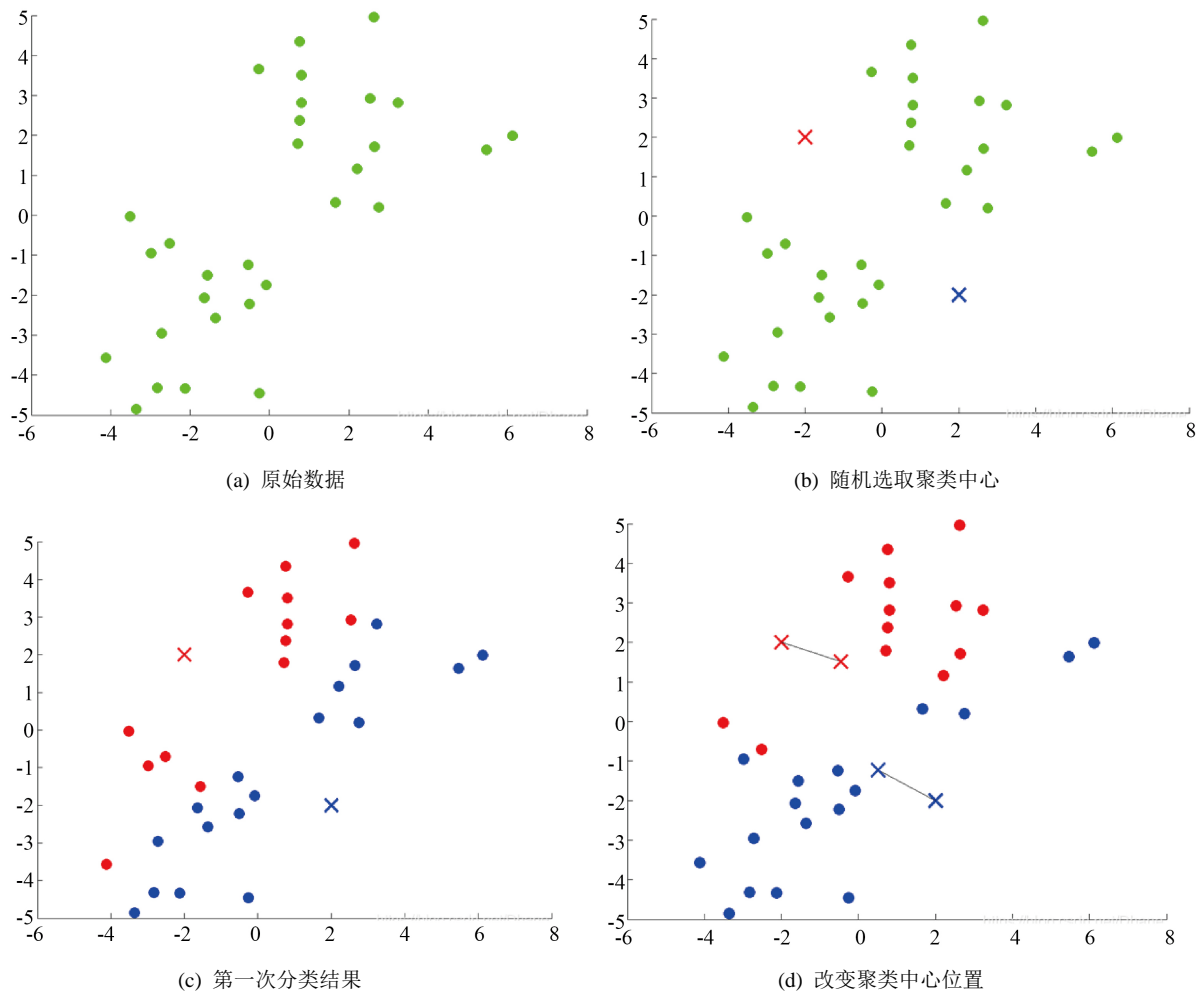
实验所使用的 K 均值聚类算法将像素点分成 K 类,先随机选取 K 个点作为中心点,计算每个点到 K 和中心点的距离,并将该点分给最近的中心点,形成 K 个簇,当所有的点都被分配完成后,每个聚类的中心的位置会被再次计算并改变,之后不断重复上述的过程,直到达到某个终止的要求为止。终止条件可以有多种,比如聚类中心不再发生变化、没有(或最小数目)对象被重新分配、或者达到设定的迭代次数等。用公式来表示,将簇划分为 (C_1, C_2, \dots, C_k) ,最小化平方误差 E 表示为:

$$E = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|_2^2 \quad (1)$$

其中 μ_i 是簇 C_i 的均值向量，表达式为：

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x \quad (2)$$

假设拿到了图 3(a)所示的数据，目的是要将里面的绿点其分成两类。本文采用的算法首先随机选取两个点作为分类的聚类中心，如图 3(b)所示，红×和绿×分别代表两个聚类中心。算出每个点到分类中心的距离，根据距离红×或蓝×的远近将点分别进行标记，这就完成了第一次分类，如图 3(c)所示，但这显然不是我们想要的结果，因此需要重新计算并改变中心的位置。如图 3(d)所示，两个中心点的位置发生了变动，相应的每一个点到中心点的位置也发生了变动，因此这里要重新将所有点依据距离聚类中心近的方式将其重新归类，这样就离分类的目标更加接近了一步。之后再次重新计算聚类中心位置，并根据点到聚类中心的距离重新将点进行归类，并重复以上步骤，不断重新计算，直到聚类中心位置不在发生改变或者完成预定的迭代次数为止，这样就完成了将上图的点分成两类的目标。图 3(e)所示即为最终分类的结果图。



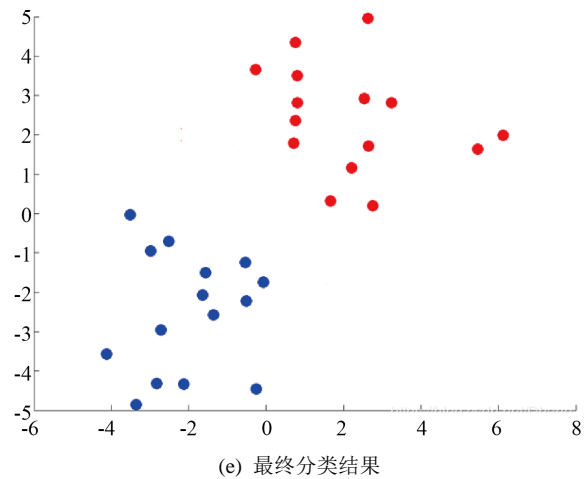


Figure 3. Flow chart of clustering algorithm

图 3. 聚类算法过程图

从图 3(e)可以看出, 经过聚类分类后, 不同深度层次的特征点分为两部分, 并且有好的分界区间。本次实验依据图像中的物体在不同距离的灰度值差异, 用 K 均值聚类算法将图像划分成若干部分, 实现判断图像深度层次的目的。

4. 实验结果

为与真实场景进行对照, 实验采用 Make3D 数据集中的样本进行对比, 该数据集是由 Ashutosh Saxena [2]等人在做 Make3D 任务时所采集, 该样本每张图像深度值均由激光雷达采集, 最大的深度为 81 米, 超过 81 米则映射为 81 米。样本中包括人造环境和自然环境等。本实验以 Matlab 作为实验平台, 分别给实验对象赋予不同 K 值进行实验, 进而得到图像深度层次。表 1 是本次实验的实验结果。

Table 1. Experimental result

表 1. 实验结果

原图	真实深度图	2 聚类中心划分结果	3 聚类中心划分结果	4 聚类中心划分结果

表 1 所示的实验结果从左往右分别是经预处理的灰度图、Make3D 给定的真实深度信息图、 $K = 2$ 得到的深度层次图、 $K = 3$ 得到的深度层次图以及 $K = 4$ 得到的深度层次图。随着聚类中心的生长, 图像在

深度层次上被区分开来, 表 1 中的 4 聚类中心划分结果图红色代表离单目摄像机距离近, 而蓝色代表距离远。

5. 结论

根据单目图像中不同深度的灰度值不同的角度出发, 采用 K 均值聚类算法将灰度值相近的像素进行聚集分类, 并与 Make3D 数据集给定的深度信息进行对比, 实现了深度层次的划分。从实验结果可以经过 K 聚类均值算法深度层次分析后, 图中的像素点的深度可以区分开来, 为进一步计算不同像素点的深度提供了依据。该算法对于较为单一的几何场景进行深度层次的划分时取得了非常理想的效果。

基金项目

2018 年内蒙古自治区自然科学基金项目(2018LH050248); 2018 年内蒙古自治区高等学校科学技术研究项目(NJZY18149)。

参考文献

- [1] Saxena, A., Chung, S.H. and Ng, A.Y. (2007) 3D Depth Reconstruction from a Single Still Image. *International Journal of Computer Vision*, **76**, 53-69. <https://doi.org/10.1007/s11263-007-0071-y>
- [2] Saxena, A., Sun, M. and Ng, A.Y. (2009) Make 3D: Learning 3D Scene Structure from a Single Still Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **31**, 824-840. <https://doi.org/10.1109/TPAMI.2008.132>
- [3] Saxena, A., Schulte, J. and Ng, A.Y. (2007) Depth Estimation Using Monocular and Stereo Cues. *Proceedings of the 20th international joint conference on Artificial intelligence*, 2197-2103.
- [4] Saxena, A., Sun, M. and Ng, A.Y. (2007) Learning 3-D Scene Structure from a Single Still Image. 2007 *IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, 14-21 October 2007, 1-8. <https://doi.org/10.1109/ICCV.2007.4408828>
- [5] Saxena, A., Chun, S.H. and Ng, A.Y. (2005) Learning Depth from Single Monocular Images. *Proceedings of the 18th International Conference on Neural Information Processing Systems*, 1161-1168.
- [6] Saxena, A., Sun, M. and Ng, A.Y. (2007) 3-D Reconstruction from Sparse Views Using Monocular Vision. 2007 *IEEE 11th International Conference on Computer Vision*, Rio de Janeiro, 14-21 October 2007, 1-8. <https://doi.org/10.1109/ICCV.2007.4409219>
- [7] Horn, B.K.P., 著. 机器视觉[M]. 王亮, 蒋欣兰, 译. 北京: 中国青年出版社, 2014.
- [8] Eigen, D., Puhrsch, C. and Fergus, R. (2014) Depth Map Prediction from a Single Image Using a Multi-scale Deep Network. *Proceedings of the 27th International Conference on Neural Information Processing Systems*, Montreal: MIT Press, 2366-2374.
- [9] 裴晨, 徐国彬, 于艺铭, 吴灵, 黄俊轶, 王琪. 基于改进 K 均值聚类的图像修复方法[J]. 包装工程, 2020, 41(23): 255-262
- [10] Wei, X.Q. and Zhang, Y.E. (2020) Image Segmentation Algorithm Based on Dynamic Particle Swarm Optimization and K-Means Clustering. *International Journal of Computers and Applications*, **42**, 649-654. <https://doi.org/10.1080/1206212X.2018.1521090>
- [11] 朱永红, 熊朦, 赵一峰, 王伟. 基于 K 均值聚类陶瓷窑炉火焰图像分割方法[J]. 陶瓷学报, 2016, 37(1): 86-90.