

关于 Matting 问题的报告

201250070 郁博文

1. 问题描述

图像 matting 是计算机视觉中一项重要任务，旨在将图像中的前景对象从背景中准确地分离出来。这一任务对于图像编辑、虚拟现实、视频合成等应用至关重要。图像抠图方法分为三种类型，基于采样，基于传播和基于学习[1]。

传统的 matting 方法采用了多种技术和策略。这些方法利用图像中像素的颜色、纹理、边缘等信息来推断每个像素的 matting 参数。一些传统方法基于统计建模、优化算法或者机器学习方法来估计这些参数。其中一些方法需要人工标记的训练样本进行模型训练，而另一些方法则依赖于预先设定的背景条件。

除了传统方法外，蓝幕（Blue Screen）matting 是一种广泛应用于影视特效的技术。它利用拍摄过程中的背景色彩差异，通过色彩键控技术将前景对象从背景中分离出来。这种方法通常在拍摄中使用蓝色或绿色背景，并通过背景差分来实现前景的提取。

尽管传统 matting 方法在某些场景下具有一定的效果，但它们面临一些挑战。传统方法可能在复杂的图像场景中无法准确地分离前景和背景的细节部分。此外，传统方法可能需要大量的人工标记样本或先验知识，并且在处理复杂纹理、半透明物体或边缘模糊的图像时可能表现不佳。

需要注意的是，近年来深度学习方法在图像 matting 任务中取得了重要的进展，但在本报告中将专注于传统的 matting 算法，不包括深度学习方面的内容。

2. 发展趋势

在计算机视觉中，matting 是一个受约束程度高、难度大的问题，需要用户交互才能解决这一复杂问题[1]。

早期，人们通过手工绘制或使用图形编辑软件中的基本工具，如钢笔工具或魔棒工具，进行抠图。这种方法需要耗费大量的时间和精力，并且在处理复杂的边缘或细节时往往不够精确。随着计算机视觉和图像处理技术的发展，人们开始使用基于颜色和阈值的分割方法进行抠图。这些方法基于颜色或灰度值的差异来分割前景和背景基于边缘的方法可能会产生不准确的结果。之后，人们开始将一些机器学习的算法运用到抠图技术中，例如 KNN 算法[3]。它提供了比非局部匹配更快、更好的结果。

随着计算机图形和计算能力的不断提升，人们对实时和交互式 matting 处理的需求也在增加。这意味着未来的 matting 算法将更加注重高效的实时处理和用户交互体验[1]。同时，深度学习和神经网络已经在计算机视觉领域取得了巨大的成功，对 matting 技术也有着积

极的影响。未来的 matting 算法可能会更多地利用深度学习方法，通过大规模数据集的训练来提高 matting 的准确性和鲁棒性。例如，将卷积神经网络应用于抠图[2]、

3. 代表性方法

3.1 KNN

KNN Matting 是一种基于 K 最近邻算法的抠图技术，旨在通过利用图像中的局部颜色信息来实现高质量的图像抠图。该技术利用图像中像素之间的相似性，将未知像素分为前景和背景，从而生成一个透明度（alpha）图，以实现精确的抠图效果。利用非局部原则，通过使用 K 近邻(KNN)来匹配非局部邻域，并提供了一个简单快速的算法，给出了稀疏用户标记的竞争结果。[1]

KNN Matting 的算法原理主要分为以下几个步骤。首先从图像中选择一组具有已知前景和背景的像素样本作为训练数据。这些样本应该足够代表图像中的前景、背景和未知区域。对于每个样本像素，提取其局部颜色特征。一种常用的方法是使用颜色直方图表示局部颜色特征。对于图像中的每个未知像素，计算其与训练数据样本的颜色特征之间的距离，并选择 K 个最近邻的样本。基于 K 个最近邻样本的透明度信息，使用线性方程组求解或其他方法估计未知像素的透明度值。根据估计的透明度值，将图像的前景和背景进行分离，生成一个具有透明度信息的抠图结果。[3]

然而，需要与机器学习中的 KNN 算法一样，K 值的取值会很大影响结果。

3.2 Bayesian

Bayesian Matting 预先假设输入图像已经被分割成三个区域：“背景”、“前景”和“未知”，同时背景和前景区域已经被保守地描绘出来。算法的目标是在给定图像未知区域内每个像素的观察颜色 C 的情况下，求解前景色 F、背景色 B 和不透明度 α 。由于 F，B 和 C 各有三个颜色通道，我们有一个有三个方程和七个未知数的问题。[4]事实上，这也就将算法转化成了一个优化问题，所以可以用贝叶斯推断来实现。

Bayesian Matting 的步骤如下：首先从图像中选择一组具有已知前景和背景的像素样本作为训练数据。这些样本应该足够代表图像中的前景、背景和未知区域。对于每个样本像素，提取其颜色值、梯度信息以及其他可能的特征。这些特征可用于建立贝叶斯模型进行透明度估计。根据训练数据和先验知识，建立颜色和梯度的联合概率分布模型。这个模型可以捕捉前景、背景和未知区域的颜色和梯度分布特征。对于图像中的每个未知像素，根据其颜色和梯度特征，利用贝叶斯推断估计其透明度值。这个过程涉及到先验模型、似然模型和先验知识的组合，以最大化后验概率。根据估计的透明度值，将图像的前景和背景进行分离，生成一个具有透明度信息的抠图结果。

3.3 Closed Form Solution

这个算法的核心思想是通过将自然图像抠图问题转化为一个线性系统求解的问题，并通过最小化能量函数来联合优化图像的颜色和抠图参数。相比于传统的图像抠图方法，该算法无需用户交互，并能够提供更准确的抠图结果。[5]

具体来说，该算法步骤如下：首先对输入图像进行预处理，包括颜色空间转换和参数初始化。将图像表示为 RGB 颜色空间，并初始化前景、背景和未知区域的参数。然后使用图像中已知的前景和背景像素，通过最大似然估计建立起一个 GMM 模型，用于对图像的颜色进行建模。通过将颜色和抠图参数联合优化，建立一个线性系统。该线性系统的目标是最小化能量函数，以获得每个像素点的抠图结果。通过求解线性系统，可以得到每个像素点的抠图参数。这可以通过矩阵运算或迭代方法来实现。最后利用获得的抠图参数，对未知区域进行插值，并生成最终的抠图结果。根据前景和背景的权重，将抠图结果与背景进行融合，得到最终的抠图图像。

3.4 Blue Screen

Blue Screen Matting 用于将前景对象从一个图像中提取并放置到另一个图像或背景中。它是一种基于颜色键（color keying）的抠图方法，其中蓝色背景通常用作抠图的背景。具体的工作原理如下。首先选择一个具有均匀蓝色背景的图像作为合成的背景。在一个独立的环境中拍摄前景对象，通常是一个人或物体，将其置于一个蓝色背景前。确保前景对象与蓝色背景之间存在明显的颜色差异。使用相机或摄像机捕获前景对象和蓝色背景的图像。将前景图像和背景图像进行预处理，包括色彩校正、去噪等步骤，以提高合成的质量。使用图像处理软件，通过颜色键（color key）技术，将蓝色背景从前景图像中抠掉。通常是通过选择蓝色区域并将其设为透明或替换为其他背景来实现。将抠图后的前景对象与所选的背景图像进行合成。可以通过图像编辑软件中的图层叠加或融合技术来实现。对合成图像进行必要的优化和调整，以确保前景对象与背景之间的色彩和光照一致性。[6]

关于 Blue Screen 的应用不必多说，在电影等领域经常见到，至于为什么使用蓝色。选择蓝色作为背景颜色主要是因为蓝色在自然图像中较少出现于前景对象的颜色中。这种颜色差异使得在后期处理中更容易将蓝色背景与前景对象区分开来。

3.5 Learning Based Digital Matting

Learning Based Digital Matting 的主要思想是将数字抠图问题视为机器学习中的半监督学习任务[7]，产生了一种基于局部学习的方法和一种基于全局学习的方法，以更好地适应基于潦草的抠图和基于 trimap 的抠图。基本步骤如下：首先准备带有已知前景和背景的图像对作为训练数据。这些图像对通常需要包含前景对象、背景和对应的抠图结果（即真值标签）。然后从每个图像对中提取特征，这些特征可以包括颜色、纹理、边缘等信息。这些特征可用于描述前景对象和背景之间的差异和关联。使用训练数据和特征，训练一个机器学习模型。常见的模型包括支持向量机（Support Vector Machine, SVM）、随机森林（Random

Forest)、神经网络等。训练的目标是通过输入图像的特征来预测每个像素点的抠图结果。对训练的模型进行优化和调整,以提高抠图的准确性和鲁棒性。这可以包括模型参数的调整、特征选择、数据增强等方法。使用训练好的模型对新的图像进行抠图预测。通过输入图像的特征,模型可以自动预测每个像素点的前景或背景概率,从而生成抠图结果。

致谢

感谢任桐伟老师的授课,使我对计算机视觉领域有了初步的认知。

感谢 chatgpt 模型在论文推荐时给我的参考,同时在理解算法方面给予了我帮助。

参考文献

- [1] J. Boda and D. Pandya, "A Survey on Image Matting Techniques," 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 2018, pp. 0765-0770, doi: 10.1109/ICCSP.2018.8523834.
- [2] S. M. Assari, H. Idrees, and M. Shah, "Re-identification of Humans in Crowds using Personal, Social and Environmental Constraints," vol. 1, pp. 626–643, 2016.
- [3] Q. Chen, D. Li, and C. K. Tang, "KNN matting," IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 9, pp. 2175–2188, 2013.
- [4] TY - CONF TI - A Bayesian approach to digital matting T2 - Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 SP - II EP - II AU - Yung-Yu Chuang AU - B. Curless AU - D. H. Salesin AU - R. Szeliski PY - 2001 DO - 10.1109/CVPR.2001.990970 JO - Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 IS - SN - 1063-6919 VO - 2 VL - 2 JA - Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001 Y1 - 8-14 Dec. 2001 ER -
- [5] A. Levin, D. Lischinski and Y. Weiss, "A Closed-Form Solution to Natural Image Matting," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 2, pp. 228-242, Feb. 2008, doi: 10.1109/TPAMI.2007.1177.
- [6] Smith, Alvy Ray and James F. Blinn. "Blue screen matting." *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques* (1996): n. pag.
- [7] Y. Zheng and C. Kambhamettu, "Learning based digital matting," 2009 IEEE 12th International Conference on Computer Vision, Kyoto, Japan, 2009, pp. 889-896, doi: 10.1109/ICCV.2009.5459326.