

# 面向电子助视系统的文档图像拼接算法研究

胥鹏鹏, 刘国强, 郑雅羽

(浙江工业大学 信息工程学院, 杭州 310023)

**摘要:** 电子助视系统是一种运用实时图像处理技术来提高低视力患者阅读能力的仪器; 由于传统电子助视系统使用单摄像头采集图像, 视角较小, 要多次移动助视设备或目标书籍才能完成阅读, 因此用户体验不佳; 针对此问题, 设计了一种基于多路摄像头采集的文档图像拼接算法; 在图像配准步骤, 提出了一种针对文档图像的区域特征描述方法, 提取单个文字区域作为特征区域, 其质心作为特征点, 通过描述特征区域的纹理信息完成特征点的描述, 提高了特征匹配效率; 在图像融合步骤, 提出了一种利用文档图像字符间隙的最优拼接缝搜索算法, 消除了拼接鬼影; 构建多尺度的图像金字塔, 使用多分辨率融合算法, 抑制了融合图像的曝光差异; 从实验结果来看, 算法获得了很好地配准和融合效果, 提升了配准速度, 消除了配准鬼影和曝光差异, 能够获得较广的阅读视角, 可以较好地用于电子助视系统。

**关键词:** 电子助视系统; 文档图像拼接; 特征描述; 最优拼接缝; 多分辨率融合

## Research on Document Image Mosaic Algorithm for Electronic Video Magnify System

Xu Pengpeng, Liu Guoqiang, Zheng Yayu

(College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China)

**Abstract:** Electronic video magnify system is an instrument, using real-time image processing technology to improve the ability of reading of low vision people. However, traditional electronic video magnify system usually uses a single camera to collect document image. Therefore, the acquisition angle is small, can not get a complete chapter image. In view of this problem, this paper designs a multi-channel camera document image stitching algorithm. In the image registration step, an area feature description method for document image is proposed to improve the feature matching efficiency. In the image fusion step, a search algorithm of the optimal splicing seam of document image is proposed to eliminate the splicing ghost. At the same time, the multi-resolution fusion algorithm is applied for optimizing the image exposure differences. According to the experimental results, this algorithm receives a good registration and fusion effect and can be better used in electronic video magnify system, in which this algorithm improves the registration accuracy and speed as well as eliminates the splicing ghost and exposure differences, obtaining a wider reading angle and high image resolution.

**Keywords:** electronic video magnify system; document image stitching; feature description; optimal splicing seam; multi-resolution fusion

## 0 引言

目前全球低视力患者约为 3.14 亿人<sup>[1]</sup>, 每年因视力损害引起的不良事件花费高达 250 亿美元<sup>[2]</sup>, 视力障碍给患者和卫生保健体系带来了巨大的负担, 严重影响低视力人群的身心健康和日常生活质量, 轻者不仅使工作、学习受阻, 重者往往生活不能自理。对于那些手术及药物治疗依然无效的低视力患者, 其中 50% 以上是可以通过使用相关助视器来提高视力的, 而且这是提高患者视力的最后一次机会。

随着计算机技术的发展和运用, 电子助视系统的研究取得了很大进展, 提高了低视力患者的阅读能力。电子助视系统在纸质文档经过摄像头采集后进行文字识别 OCR<sup>[3]</sup>, 接着通过文字放大, 语音播报等方法帮助低视力患者接受信息。通常,

报纸、杂志等版面篇幅过大, 仅仅采用单摄像头进行文档图像采集, 一次只能采集部分篇章, 且采集到的图像视角较小, 分辨率较低。低视力患者在使用时只有移动助视设备或者目标书籍才能阅读完整篇章。但是, 由于低视力患者视觉存在缺陷, 在移动助视设备或目标书籍时并不能很好地定位在需要阅读的位置, 使得助视效果较差。因此, 本文提出使用多个摄像头进行文档图像的采集, 将不同位置的图像进行拼接, 从而获得较广的阅读视角以及较高的图像分辨率, 提高低视力患者的阅读能力。

图像拼接的方法很多<sup>[4]</sup>, 由图像配准和图像融合两个过程来完成。图像配准是确定两幅图像间的相对关系, 目的是为了获取在同一坐标系下的变换关系的单应矩阵。图像配准是图像融合的基础, 通常图像配准算法的计算量非常大, 因此图像拼接技术的发展很大程度上取决于图像配准技术的创新。基于特征的图像拼接方法利用了图像中数量较少, 特征较稳定的一些点、线或边缘等进行匹配, 相对压缩了所需处理的信息量, 使得匹配搜索的计算量小、速度较快, 是目前研究最多、应用最广的一类配准方法<sup>[5]</sup>。图像融合是将两张不同视角的图像融合在一起, 并将两张图像融合部分的差异降低至最小, 从而获取最终的平滑清晰的图像。

收稿日期: 2017-07-14; 修回日期: 2017-08-21。

基金项目: 国家自然科学基金资助项目(61401398)。

作者简介: 胥鹏鹏(1991-), 男, 河南周口人, 硕士研究生, 主要从事视频和图像处理、嵌入式系统应用方向的研究。

通讯作者: 郑雅羽(1978-), 男, 浙江温州人, 博士, 副研究员, 硕士生导师, 主要从事视频和图像处理、深度学习及嵌入式系统方向的研究。

陈志昂<sup>[6]</sup>等人通过对比不同的图像配准方法, 得出了较为通用的配准方法, 即基于 SIFT<sup>[7]</sup> 或 SURF<sup>[8]</sup> 的特征提取的配准方法。苗立刚<sup>[9]</sup>将通用的基于 SIFT 的图像配准方法运用于手持相机的文档图像拼接, 但是由于 SIFT 算法的复杂度较高, 所以拼接所用的时间较长, 因此对于电子助视器这种实时的图像处理系统是不适用的。Tomohiro Nakai<sup>[10]</sup>等人提出了一种基于摄像头的低质量英文文档图片的配准算法 (LLAH), 该算法首先提取文档图片中的特征点, 然后依据局部相邻关键点之间相互组合计算出仿射不变的哈希值对特征点进行描述, 这种方法虽然具有较好的仿射不变性, 但是没有较好的解决不同尺度文档图像拼接问题, 配准精度较低, 且涉及到多次的局部区域面积的计算, 算法复杂度也相对较高, 因此该算法也不能较好的应用于电子助视系统。

Tomohiro Nakai 等人在图像融合时均采用加权融合的方法对重合区域进行处理, 往往会因为图像配准误差产生的部分重叠而使得图像变的模糊, 又称为鬼影<sup>[11]</sup>。James Davis<sup>[12]</sup>等人利用 Dijkstra 算法寻找一条最佳的缝合线, 通过在缝合线两边只取一幅图像的内容实现两幅图像的拼接, 达到消除鬼影的目的。但是由于 Dijkstra 算法复杂度非常高, 所以这种拼接缝的搜索方法不能适用于电子助视系统。

可以看出, 已有的文献均未能很速准确地完成文档图像的拼接。针对前人工作中存在的问题, 本文提出了一种强鲁棒性的文档图像拼接算法, 流程如图 1 所示。本文提出将单个中文字符或者英文单词作为一个特征区域, 区域的质心作为特征点, 提出了基于 ART<sup>[13]</sup>变化的文字区域特征描述方法, 通过比较特征描述子完成特征点匹配, 使用 RANSAC<sup>[14]</sup>特征匹配算法计算出两张图之间的映射关系, 然后使用仿射变化完成图像的配准。同时提出了一种利用文档图像字符间隙来快速搜索的最优拼接缝, 使用了多分辨率的融合方法对曝光差异进行了优化。最终获得色调平衡, 过渡平滑的拼接图像。

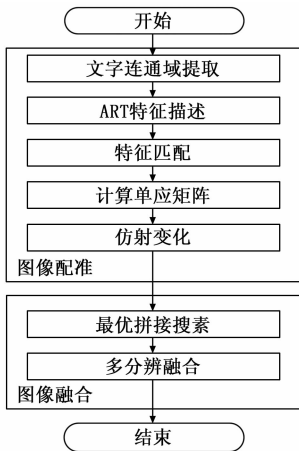


图 1 文档图像拼接流程图

### 1 ART 区域特征描述算法

本文提出的文字区域特征描述方法, 是利用文字的不同纹理信息, 计算文字区域的不同维度 ART 变换系数, 以此构建特征描述向量。这种描述方法具有仿射和尺度不变性、对噪声具有较强的鲁棒性、能够较好地描述复杂纹理等优点。图像特征点的个数即为图像中汉字或单词的个数, 相对于通用的图像

配准方法提取到的大量的特征点, 本文方法的特征点个数较少, 因此在特征提取和描述以及匹配阶段都减少了大量的计算。

#### 1.1 文字连通域提取

本算法采用连通区域分析进行单个文字图像的分割, 使用 Two-Pass 法<sup>[15]</sup>进行连通域的提取。由于中文汉字和英文单词分别由不同偏旁部首和字母组成, 而不同的偏旁部首和字母是不连通的。如果直接进行连通域提取则不能将整个汉字或者单词区域分割出来, 如图 2 (a) 所示。因此, 本算法在连通域提取之前进行一次形态学腐蚀操作, 使得同一汉字的不同偏旁以及单词的不同字母之间互不连通, 不同汉字或者单词之间不连通, 如图 2 (b) 所示。但是仍有些较小的汉字部首以及标点符号被单独提取出来, 这些会导致特征匹配时产生错误匹配, 因此本算法根据连通域的面积将较小的连通域进行剔除, 筛选后的效果如图 2 (c) 所示, 可以看出图中单个汉字和单词已经被提取出来。



图 2 单个文字或单词的连通域提取

#### 1.2 ART 特征描述

对于单字区域的特征描述, 本算法采用基于角径向变换的描述方法。角径向变换 (Angular Radial Transform, ART) 是一种二维区域形状描述矩算子, 源于 MPEG-7 标准。利用 ART 进行特征提取, 具有尺度和旋转的不变性、对噪声具有较强的鲁棒性, 能够较好地描述复杂纹理等优点。

##### 1.2.1 ART 变换

ART 变换具有一组定义在单位圆上正交基函数, 基函数有两部分组成, 它们分别是径向函数以及角向基函数, 定义如下:

$$V_{mn}(\rho, \theta) = R_n(\rho)A_m(\theta) \tag{1}$$

其中:  $R_n(\rho)$  是阶数为  $n$  的径向函数, 定义:

$$R_n(\rho) = \begin{cases} 1, & n = 0 \\ 2\cos(\pi n \rho), & \text{others} \end{cases} \tag{2}$$

$A_m(\theta)$  是阶数为  $m$  的角向基函数, 定义:

$$A_m(\theta) = \frac{1}{2\pi} \exp(jm\theta) \tag{3}$$

通过欧拉变换公式, 可以将 ART 基函数转换在直角坐标系下, 图 3 (a) 和 (b) 分别对应直角坐标系下 ART 基函数实部和虚部的图像, 当  $n$  和  $m$  取的越大, 所对应的基图像的频率越高<sup>[16]</sup>。

令  $f(\rho, \theta)$  是图像在极坐标情况下的灰度函数, 将极坐标下  $(\rho, \theta)$  函数投影到 ART 基函数  $V_{nm}(\rho, \theta)$  上, 即得到极坐标下阶数分别为  $n$  和  $m$  的 ART 系数:

$$F_{nm} = \int_0^{2\pi} \int_0^1 V_{nm}(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (4)$$

本文即是通过  $F_{nm}$  来反映不同频率图像区域的周期性纹理。

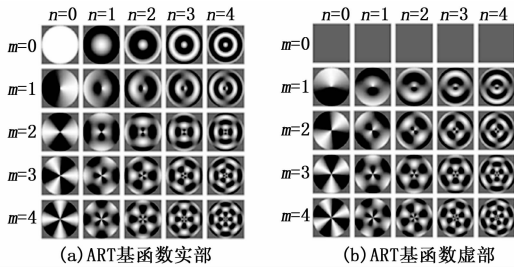


图 3 ART 基函数关于  $n$  和  $m$  设置从 0~4

### 1.2.2 特征描述

在 MPEG-7 标准中直接取  $n=10, m=10$  计算出的作为特征描述子。但是这个描述子仅对复杂的形状具有较好的鲁棒性, 在应用到文字连通域的描述时, 能够较好地描述复杂的汉字, 对于简单的汉字描述效果不好。

因此本文提出了一种针对文字区域纹理描述子。通过计算不同维度的 ART 变换系数, 构建一个多维的特征矩阵  $C_{NM}$ , 定义如下:

$$C_{NM} = \begin{bmatrix} |F_{0,0}| & |F_{0,1}| & \cdots & |F_{0,m}| \\ |F_{1,0}| & |F_{1,1}| & \cdots & |F_{1,m}| \\ \vdots & \vdots & \vdots & \vdots \\ |F_{n,0}| & |F_{n,1}| & \cdots & |F_{n,m}| \end{bmatrix} \quad (5)$$

其中:  $0 \leq n \leq N-1, 0 \leq m \leq M-1$ 。  $F_{nm}$  为不同维度的 ART 变化系数, 由于  $F_{nm}$  为一个复数, 因此本文取复数模作为特征矩阵的元素。

该特征描述子包括了  $N \times M$  个 ART 变换的系数, 反应了  $N \times M$  种不同频率的纹理信息, 在保证提取到高频率的纹理信息的同时, 也获取到了低频率的纹理的信息, 这样就可以保证特征描述子在匹配时有较高的匹配率。

表 1 不同  $M, N$  特征点匹配率 %

M \ N	%			
	3	4	5	6
3	22.4	42.9	46.3	62.2
4	38.3	53.7	68.4	66.5
5	52	70.7	80.6	74.5
6	64.9	74.6	72.1	76.9

本文对不同维度的特征矩阵所对应的文字图像特征匹配的效果进行测试。如表 1 所示, 取不同的  $N$  和  $M$  值, 分别计算不同的  $C_{NM}$  完成特征匹配, 其中, 表中内容是使用 RANSAC 进行误匹配点剔除后剩余的正确的匹配点占总匹配点的比例, 即匹配率。从表中看出, 并不是阶数  $N$  和  $M$  越大, 图像的匹配率越高, 当  $N$  和  $M$  大于 5 之后, 匹配率反而下降, 这是因为对于文字这种纹理较为简单的图像, 不需要进行复杂的细节的描述, 当  $N$  和  $M$  取的过大反而会因为噪声产生误匹配。因

此对于文字图像的区域描述在  $N$  和  $M$  取值为 5, 能够获得最好的匹配效果。

## 2 基于最优拼接缝搜索的多分辨率融合

本文提出一种针对文档图像的最优拼接缝的搜索算法, 消除图像融合时产生的鬼影, 再利用搜索到的最优拼接缝进行图像的多分辨率融合, 消除不同图像融合的曝光差异。

### 2.1 文档图像的最优拼接缝搜索算法

使用拼接缝可以将图像配准之后的重叠区域分割为两部分, 在图像融合时分别从两张原图像找出对应部分进行填充, 如果分割较好, 则图像融合之后可以消除鬼影。对于文档图像, 本文使用不包含文档信息的字符间隔区域作为拼接缝, 保证图像融合时缝合线两侧都是完整的字符。算法分为初选区域划定和最优路径搜索两部分。

#### 2.1.1 初选区域划定

拼接缝是在图像的重叠区域进行搜索, 根据图像配准后两幅图像间的对应关系, 可以确定图像的重叠区域。重叠区域包含多列文字图像, 若在整个重叠区域进行路径的搜索, 则得到大量的可行路径, 因此本算法优先划定一个重合度最高的初选区域, 然后在初选区域中完成路径的快速搜索。计算方法如下所示:

$$R = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |A(i, j) - B(i, j)|} \quad (6)$$

其中:  $M, N$  为初选区域宽和高, 可以根据连通域提取步骤获取的单字面积大小来确定。  $A(i, j), B(i, j)$  为两张图像重叠区域的像素值。将初选区域从左至右滑动, 计算  $R$  值。若值越小, 说明图像间重合程度越好, 也表明了配准效果越好, 因此选定  $R$  值最小的区域作为初选区域。

如图 4 所示为初选区域划定效果图, 其中图 (a) 为重叠区域连通域信息, 本文选取初选区域宽度  $M$  为单字连通域宽度的三倍,  $N$  为初选区域的高度。图 4 (a) 中矩形标注为搜索到的初选区域, 从图中可以看出初选区域一般在空白较多的地方, 因此可以更快速地搜索到最优拼接缝。

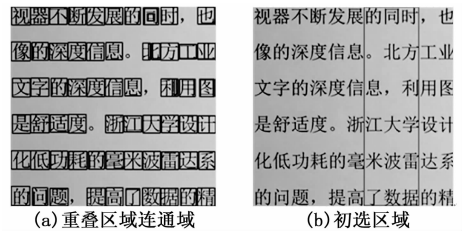


图 4 初选区域划定

#### 2.1.2 最优路径搜索

根据文档图像的特点, 初选区域可以被简单地分割成行、列和单个字。先进行水平投影, 将不同行进行分割。然后对各行进行垂直投影, 得到字符的间隙, 将不同行不同字符的间隙分别连接就可以得到所有的可行路径。

在获取一些可行路径后, 通过制定规则选取最优拼接缝。取字符间隙的中间作为可能通过的通道。在众多的通道中, 字符间隙越大说明该处比较适合选取为拼接缝。而对于字符间隙相同的通道, 要考虑整体上的路径长度, 总体路径长度越短越好。最终得出一个最优的拼接缝。图 5 所示为不同测试图像最优拼接缝搜索的结果。

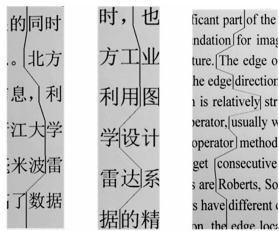


图 5 拼接缝搜索结果

### 2.2 多分辨率图像融合

本文采用多分辨率融合的方法来解决曝光差异问题。多分辨率融合是通过构建拉普拉斯图像金字塔,将原图像分别分解到不同空间频带上,在不同的空间频带上分别进行融合。这样就可以采用不同的融合算子以达到突出特定频带上特征与细节的目的<sup>[17]</sup>,能够将来自不同图像的特征与细节融合在一起。

本算法采用沿最优拼接缝的融合算子进行多分辨融合,使得在消除鬼影的时候也能抑制曝光的差异,融合的公式如下:

$$L_p(i, j) = G_m(i, j)L_{p1}(i, j) + (1 - G_m(i, j))L_{p2}(i, j) \quad (7)$$

其中:  $L_p(i, j)$  为融合后图像  $p$  的  $r$  阶拉普拉斯图像金字塔的像素值。图像  $m$  为以拼接缝为分割线的黑白填充图,如图 6 (a) 所示。 $G_m(i, j)$  为拼接缝图像  $m$  的  $r$  阶高斯金字塔图像的像素值。 $L_{p1}(i, j)$  和  $L_{p2}(i, j)$  分别为扩展了的原始图像  $p_1$  和  $p_2$  的  $r$  阶拉普拉斯金字塔图像对应的像素值,  $p_1$ 、 $p_2$  如图 6 (b)、(c) 所示。

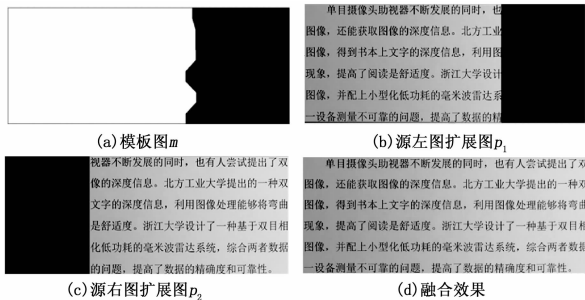


图 6 多分辨率融合输入图像

分别完成  $r$  阶金字塔的融合后,从最高一阶向下扩展,并与其下一阶的图像相加,直到最底一阶,即可得到融合完成的图像如图 6 (d) 所示。

### 3 实验结果及分析

本文实验是在 PC 机上进行,PC 环境为酷睿 i5 2.4 GHz 的 CPU、内存 6 GB。软件环境为 Visual Studio 2013,测试图像的分辨率为  $720 \times 480$ 。

在图像配准步骤实验中,使用本文提出的 ART 区域特征描述算法对图像进行特征提取,然后进行特征向量的匹配。其中匹配效果如图 7 所示为使用中文文档图像、英文文档图像、倾斜图像、不同尺度图像分别进行拼接测试,其中匹配点对用直线连接。从匹配效果中可以看出,本文提出的针对文档图像的 ART 区域特征描述算法可以较好的完成单字区域的特征匹配。

相对于本文中提到的匹配方法,图 8 为本文中提到的通用的特征匹配算法和针对于文档图像的特征匹配算法的匹配效

果。其中图 (a) 为苗立刚等人使用的基于 SIFT 特征匹配的算法。图 (b) 为陈志昂等人使用的 SURF 特征匹配算法。图 (c) 为 Tomohiro Nakai 等人提出的针对文档图像的 LLAH 特征匹配算法。从图中可以明显得出,这些算法产生了大量的误匹配。

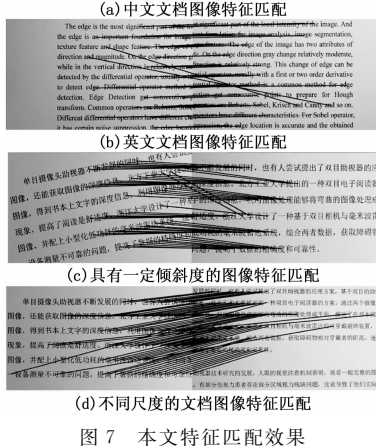


图 7 本文特征匹配效果

果。其中图 (a) 为苗立刚等人使用的基于 SIFT 特征匹配的算法。图 (b) 为陈志昂等人使用的 SURF 特征匹配算法。图 (c) 为 Tomohiro Nakai 等人提出的针对文档图像的 LLAH 特征匹配算法。从图中可以明显得出,这些算法产生了大量的误匹配。

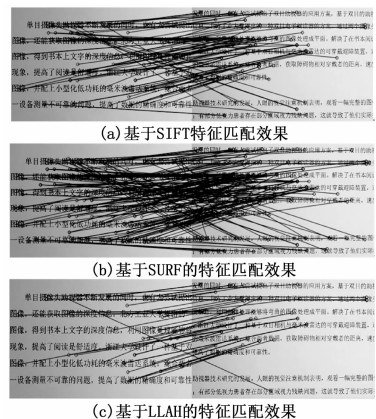


图 8 其他算法特征匹配效果

为了定量地对比特征匹配的匹配率,本文通过比较使用 RANSAC 算法剔除误匹配点前后的匹配点对数,来计算特征匹配率。通过特征匹配的时间来比较特征匹配的效率,结果如表 2 所示。

表 2 不同算法匹配率和耗时对比

	本文	SIFT	SURF	LLAH
左右图特征点数	156/223	2 130/2 627	1 106/1 327	221/324
总匹配点对数	72	365	232	113
RANSAC 后匹配对数	58	183	112	67
匹配率/%	80.6	50.1	48.2	59.2
匹配时间/S	0.53	5.68	3.21	4.68

从表中可以看出, SIFT 和 SURF 通用的图像拼接算法,能够提取到了大量的特征点,但是在匹配的时候产生了大量的误匹配和较长的匹配时间。虽然 LLAH 算法的特征匹配率高于通用的图片拼接算法,但是算法的效率较低。因此本文算法

在匹配率和算法效率上都优于文中提到的其他匹配算法。

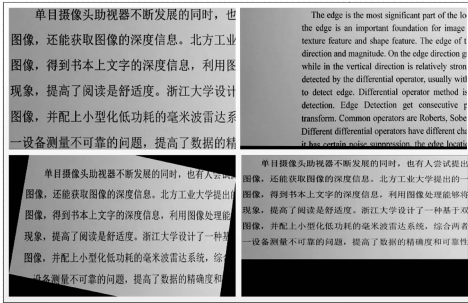


图 9 图像配准结果

为了测试图像的匹配效果, 本文以图 7 中右侧图像为基准图像, 根据匹配到的特征对求解左图配准到右图所要的单位矩阵, 然后对左侧图像进行仿射变化后即配准结果如图 9 所示。

为了定量地分析图像配准的精度, 采用经过变换后的所有匹配点对之间的平均绝对误差 MAD 和均方根误差 RMSE 作为精度评定指标<sup>[18]</sup>, 公式表示如下:

$$MAD = \frac{1}{n} \sum_{i=1}^n \|X_{1i} - \mathbf{H}X_{2i}\| \quad (8)$$

$$RMSE = \frac{1}{n} \sqrt{\sum_{i=1}^n \|X_{1i} - \mathbf{H}X_{2i}\|^2} \quad (9)$$

其中:  $X_{1i}$ 、 $X_{2i}$  为原左右图坐标,  $\mathbf{H}$  为图像配准得到的单位矩阵。

对图 8 中的算法分别进行测试, 测试结果如表 3 所示, 从表中可以看出, 本文算法与 SIFT、SURF 算法的匹配误差均在 1pixel 范围之内, 低于 LLAH 算法, 保证了较高的配准精度。且本文算法的复杂度低于 SIFT、SURF 算法, 因此本文提出的 ART 区域特征描述算法, 更适用于文档图像的配准。

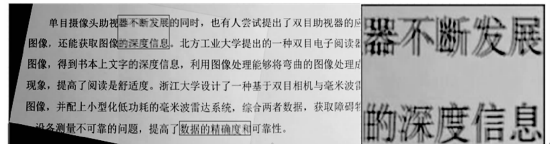
表 3 配准误差对比 pixel

	本文	SIFT	SURF	LLAH
MAD	0.672	0.590	0.689	1.125
RMSE	0.461	0.361	0.467	0.620

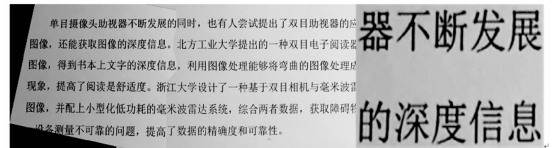
在图像融合步骤实验中, 使用最佳拼接缝对鬼影进行消除, 与 Tomohiro Nakai 使用的加权融合效果进行对比, 如图 10 所示, 图 10 (a) 图为 Tomohiro Nakai 文中融合效果, 产生了鬼影, 右图为对部分鬼影放大图。图 10 (b) 左图为使用最优拼接缝进行融合后的效果图, 能够消除加权融合产生的鬼影, 右图为放大效果图。由于在拼接缝两侧分别取原图像, 所以在拼接缝处出现了明显的过渡现象, 但是在使用了多分辨率算法对拼接缝两侧进行融合之后, 过渡现象得到了明显的抑制, 如图 10 (c) 所示。实验表明, 基于最优拼接缝搜索的多分辨率融合算法更适用于文档图像的融合。

### 4 结论

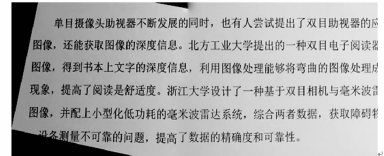
本文算法针对文档图像的特点进行设计, 与前人的算法相比, 在图像配准步骤, 提出了一种针对文本图像的 ART 区域特征描述算法, 大大减少特征点的个数, 在保证较高配准精度的同时, 配准速度得到很大的提升。在图像融合步骤, 提出了一种针对文档图像最优拼接缝的搜索算法, 用来消除图像融合



(a) Tomohiro Nakai 文中使用的融合算法



(b) 本文的最优拼接缝融合



(c) 最优拼接缝两侧使用多分辨率融合

图 10 融合效果对比

产生的鬼影; 使用多分辨率融合算法, 明显抑制由曝光误差产生的图像色彩不平衡问题。但是在使用多分辨率融合的时候的计算量比较大, 融合的时间比较长, 仍需要进一步的优化。

### 参考文献:

[1] Wang B Z, Pesudovs K, Keane M C, et al. Evaluating the effectiveness of multidisciplinary low-vision rehabilitation [J]. *Optometry & Vision Science*, 2012, 89 (9): 1399-1408.

[2] Rein D B. Vision Problems Are a Leading Source of Modifiable Health Expenditures [J]. *Investigative ophthalmology & visual science*, 2013, 54 (14): ORSF18-ORSF22.

[3] Tiwari L, Kumar B, Patnaik T. Robust Camera Captured Image Mosaicking for Document Digitization and OCR Processing [A]. 2014 International Conference on Information Technology (ICIT) [C]. IEEE, 2014: 100-105.

[4] Ghosh D, Kaabouch N. A survey on image mosaicing techniques [J]. *Journal of Visual Communication and Image Representation*, 2016, 34: 1-11.

[5] 张莹, 闫璠, 高赢, 等. 基于 ORB 算法和 OECF 模型的快速图像拼接研究 [J]. *Computer Engineering and Applications*, 2017, 53 (1): 183-189.

[6] 陈志昂, 徐晓刚, 徐冠雷. 图像拼接技术研究 [J]. *计算机科学*, 2015 (S2): 160-161.

[7] Lowe D. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*, 2004, 60 (2): 91-110.

[8] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. *Computer Vision and Image Understanding*, 2008, 110 (3): 346-359.

[9] 苗立刚. 基于手持相机的文档图像拼接算法 [J]. *中国图象图形学报*, 2009, 14 (8): 1656-1662.

[10] Nakai T, Kise K, Iwamura M. Camera-based document image mosaicing using LLAH [A]. DRR. 2009.

[11] Shum H Y, Szeliski R. Construction of panoramic image mosaics with global and local alignment Panoramic vision [M]. Springer New York, 2001: 227-268.