图像处理课程期末大作业

 姜海天
 龙蕾如
 王逸群

 2022 年 6 月 3 日

摘要

本项目实现了多种图像配准算法,图像空间变换包括仿射变换和非线性变换两种。前者的损失函数包括均方误差、负相关系数、负归一化交叉互相关、负互信息、KL 散度五种,并使用带有动量的梯度下降优化算法;后者的损失函数由光流数据约束项和正则化项组成,通过引入辅助变量并使用 Coarse-to-Fine 方法求解。

测试结果表明,对于仿射变换的各种损失函数,均方误差、负相关系数、负互信息的效果都很好,只是需要注意超参数的选择,比如形变程度更大的图像可能需要更大的学习率等等;归一化交叉互相关的计算复杂度太高,而效果并没有在前三者的基础上有明显提升,并不是一个很好的算法;KL 散度只适合一些特定的任务,而且对超参数过于敏感,因此也不是一个很好的算法。

对于非线性变换,在部分数据上的测试结果很好,而在另一部分数据上测试时,由于浮动图像的形变程度较大,加之使用了 Coarse-to-Fine 方法,使得主体结构被严重破坏。因此,在处理实际问题时,应针对特定的图像选择不同的算法。

目录

1	任务	苗述	3
2	算法	苗述	3
	2.1	仿射变换	3
		2.1.1 图像变换	3
		2.1.2 损失函数	4
		2.1.3 优化算法	5
	2.2	非线性变换	6
		2.2.1 图像变换	6
		2.2.2 损失函数	6
		2.2.3 模型求解	7
3	开发	不境	7
4	用户	操作界面	7
5	测试	吉果	8
	5.1	默认参数	8
	5.2	测试图像	8
	5.3	仿射变换	9
		5.3.1 均方误差损失函数	9
		5.3.2 负相关系数损失函数	13
		5.3.3 负归一化交叉互相关损失函数	14
		5.3.4 负互信息损失函数	15
		5.3.5 KL 散度损失函数	16
	5.4	非线性变换	16
6	总结		16
7	组内	分 工	19

1 任务描述

本项目的任务是图像配准。具体而言,输入一张浮动图像和一张参考图像,对浮动图像进行空间变换后输出结果图像。图像配准可以应用于比较或融合针对同一对象在不同条件下获取的图像,例如图像来自不同的采集设备、取自不同的时间、不同的拍摄视角等。

为了实现图像配准,首先需要确定对浮动图像的空间变换方式,即确定假设空间;接着需要确定结果图像和参考图像之间的相似性度量方式,即确定损失函数;有了假设空间和损失函数之后,需要确定优化算法,使损失函数最小化;最后输出结果图像。

2 算法描述

正如第 1节中所述,图像配准算法包括三个部分:图像变换、损失函数和优化算法。 其中,对于图像变换,本项目包含了仿射变换和非线性变换两种。

2.1 仿射变换

2.1.1 图像变换

记变换前浮动图像某点的坐标为 (x,y), 则变换后得到的坐标可记为

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

自由度为6。

为了使得各参数更具有实际意义,可将仿射变换分解为拉伸变换、旋转变换、水平 剪切变换、垂直剪切变换、平移变换的复合,即

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\begin{pmatrix} x'' \\ y'' \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x' \\ y' \end{pmatrix}$$

$$\begin{pmatrix} x''' \\ y''' \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ s_h & 1 \end{pmatrix} \begin{pmatrix} x'' \\ y'' \end{pmatrix}$$

$$\begin{pmatrix} x'''' \\ y'''' \end{pmatrix} = \begin{pmatrix} 1 & s_v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x''' \\ y''' \end{pmatrix}$$

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} x'''' \\ y'''' \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

满足

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} 1 & s_v \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ s_h & 1 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} c_1 & 0 \\ 0 & c_2 \end{pmatrix}$$

此时,参数数量增加,自由度变为7。

实际实现时,使用反向图变换进行从浮动图像到参考图像的变换,学习的参数也相 应变为逆变换的参数。另外,使用双线性插值处理变换产生的浮点数,对于超出图像边 界的点,直接取图像边界点的灰度值。

2.1.2 损失函数

对于仿射变换,本项目实现了五种损失函数,包括均方误差、负相关系数、负归一 化交叉互相关、负互信息、KL 散度。

均方误差 即为平均图像灰度值差。记 $I_0(i,j)$ 和 $I_1(i,j)$ 分别表示结果图像和参考图像 在坐标 (i,j) 处的灰度值,N 表示图像像素数量,则均方误差为:

$$MSE = \frac{1}{N} \sum_{i,j} |I_0(i,j) - I_1(i,j)|^2$$

负相关系数 一般地,对于两个向量 $x,y \in \mathbb{R}^N$,可以计算两者的样本方差与协方差,

$$Var(x) = \frac{1}{N-1} \sum_{i} (x_i - \bar{x})^2$$

$$Var(y) = \frac{1}{N-1} \sum_{i} (y_i - \bar{y})^2$$

$$Cov(x, y) = \frac{1}{N-1} \sum_{i} (x_i - \bar{x})(y_i - \bar{y})$$

进而相关系数 (correlation coefficient) 为:

$$Cor(x,y) = \frac{Cov(x,y)}{\sqrt{Var(x)Var(y)}} \in [-1,1]$$

用于估计两个向量之间的线性相关程度。如,若 Cor(x,y) > 0,则当 x 增加时,y 趋于增加;若 Cor(x,y) < 0,则当 x 增加时,y 趋于减小。本项目中,取两个向量分别为结果图像和参考图像的灰度值,进而取损失函数为负相关系数。

负归一化交叉互相关 [1] 为了更好地提取图像信息,归一化交叉互相关 (Normalized Cross Correlation, NCC) 计算局部窗口中结果图像和参考图像的相关系数并取平均值。值得注意的是,记图像大小为 $H \times W$,窗口大小为 $h \times w$,则计算归一化交叉互相关的时间复杂度为 $\Theta(hwHW)$,比相关系数的 $\Theta(HW)$ 大得多。本项目中,可取损失函数为负归一化交叉互相关,窗口大小默认为 25×25 。

负互信息 一般地,对于一个离散型随机变量 X 及其分布 p(x),熵表示其信息量:

$$H(X) = -\sum_{x} p(x) \log p(x)$$

对于两个离散型随机变量 X 和 Y 及其分布,可以计算两者的联合熵和条件熵:

$$\begin{split} H(X,Y) &= -\sum_{x,y} p(x,y) \log p(x,y) \\ H(X|Y) &= \sum_{x} p(x) H(Y|X=x) = -\sum_{x,y} p(x,y) \log p(x|y) \\ H(Y|X) &= \sum_{y} p(y) H(X|Y=y) = -\sum_{x,y} p(x,y) \log p(y|x) \end{split}$$

其中条件熵表示给定其中一个随机变量的值之后另一个随机变量的信息量。进而互信息 (Mutual Information) 为

$$MI(X,Y) = H(X,Y) - H(X|Y) - H(Y|X) = \sum_{x,y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

用于衡量随机变量之间的相关性。相关性越大,互信息越大。特别地,当随机变量相互独立时,互信息取最小值 0。本项目中,取两个随机变量的分布为结果图像和参考图像的灰度值分布,进而取损失函数为负互信息。

KL 散度 用于衡量两个概率分布 p 和 q 之间的差异, 定义为:

$$KL(p||q) = E_{x \sim p}[\log \frac{p(x)}{q(x)}]$$

在两个分布相同时取最小值 0。本项目中,记参考图像的分布为 p,结果图像的分布为 q,可以计算:

$$KL(p||q) = \sum_{x} p(x) \log \frac{p(x)}{q(x)}$$

特别需要注意的是, KL 散度与分布有关,但与位置无关,因此只适合一些特定的任务。 比如,如果浮动图像和参考图像的灰度值分布相似,则不适合选择 KL 散度作为损失函数;如果分布不相似,则可以尝试选择 KL 散度作为损失函数。

2.1.3 优化算法

本项目对仿射变换使用带有动量的梯度下降优化算法。 记待学习参数为x,损失函数为f(x)。使用差分法估计梯度:

$$\nabla f(x) = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

由于函数值沿梯度相反方向下降最快,可以使用梯度下降算法更新参数:

$$x^+ = x - \alpha \nabla f(x)$$

其中 α 是学习率,用于控制参数更新幅度。

为了加快梯度下降算法的收敛速度,以负梯度的滑动平均作为动量代替负梯度作为 参数的更新方向:

$$\Delta x^{+} = \rho \Delta x + (1 - \rho) \nabla f(x)$$
$$x^{+} = x + \alpha \Delta x^{+}$$

其中 ρ 是动量因子, 通常取 0.9, 本项目中默认取 0。

另外,为了避免参数在最优点附近来回震荡,还可以使用不同的学习率衰减方法, 本项目中使用阶梯衰减。

2.2 非线性变换

本项目的非线性变换采用 TV-L1 光流方法。

2.2.1 图像变换

与仿射变换不同,光流法并不是针对全局进行整体的空间变换,而是考虑每个坐标点的位移。即对于每个坐标点 (x,y),有:

$$\begin{pmatrix} x^+ \\ y^+ \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}$$

2.2.2 损失函数

TV-L1 方法的损失函数由光流数据约束项和正则化项组成 [2]。

光流方程包括以下三个假设: 首先,光照能量保持不变,确保无需考虑各坐标点灰度值的变化,仅需考虑其位移;其次,各坐标点的运动幅度较小;最后,相邻坐标点的运动相似。记浮动图像中坐标点 (x,y) 的光照能量即灰度值为 $I_0(x,y)$,该点运动了 (u,v) 的距离后,在结果图像中的灰度值根据第一条假设满足:

$$I_1(x+u, y+v) = I_0(x, y)$$

进而得到光流数据约束项:

$$\int_{\Omega} |I_1(x+u,y+v) - I_0(x,y)| d\Omega$$

为了确保平滑的位移,加入全变分 (Total Variance, TV) 正则化项:

$$\int_{\Omega} |\nabla u| + |\nabla v| \ d\Omega$$

最终得到损失函数:

$$\int_{\Omega} \lambda |I_1(x+u,y+v) - I_0(x,y)| + |\nabla u| + |\nabla v| \ d\Omega$$

或写作向量形式:

$$\int_{\Omega} \lambda |I_1(\boldsymbol{x} + \boldsymbol{u}) - I_0(\boldsymbol{x})| + |\nabla \boldsymbol{u}| \ d\Omega$$

2.2.3 模型求解

在损失函数中,引入辅助变量 v 和常量 $\theta > 0$,进而损失函数变为:

$$\int_{\Omega} \lambda |I_1(\boldsymbol{x} + \boldsymbol{v}) - I_0(\boldsymbol{x})| + |\nabla \boldsymbol{u}| + \frac{1}{2\theta} (\boldsymbol{u} - \boldsymbol{v})^2 d\Omega$$

当 $\theta \to 0$ 时, $v \to u$ 。由此,可以先固定 u 求解 v,再固定 v 求解 u [3]。求解过程中,利用了各坐标点运动幅度较小的假设,对损失函数中的 $I_1(x+u)$ 在 $x+u_0$ 处进行了一阶 Taylor 展开。

由于实际图像可能并不满足运动幅度较小的假设,需要使用 Coarse-to-Fine 方法: 首先,对图像进行金字塔分层,使得底层的分辨率较低,顶层的分辨率较高,因此图像中较大幅度的运动在顶层显得幅度较小;于是接下来从顶层开始,进行基于 Taylor 展开的模型求解,反复插值迭代直到底层。

Coarse-to-Fine 方法的局限性在于:一方面,图像分层方法和插值方法的选择对结果有很大的影响;另一方面,由于在 Coarse 的过程中,原本图像中的精细结构被忽略了,导致该方法对精细结构的处理不准确。

3 开发环境

以 Python 为编程语言,使用 NumPy 和 SciPy 完成主要的科学计算、实现仿射变换算法,使用 skimage 实现非线性变换算法;使用 PvQt5 完成用户操作界面。

4 用户操作界面



图 1: 操作界面示意图

如图 1所示,打开用户操作界面后,用户可以单击"选择初始图像"从本地文件中选择浮动图像,成功后浮动图像会显示在窗口左上角;单击"选择目标图像"从本地文件中选择参考图像,成功后参考图像会显示在窗口右上角;单击"选择算法"下方的下拉框,选择"均方误差"、"相关系数"、"归一化交叉互相关"、"互信息"、"KL 散度"、"非线性光流"中选择一个算法;输入学习率、学习率衰减策略、训练步数;单击"应用"开始进行图像配准,成功后结果图像会显示在窗口左下角,结果图像与参考图像的差会显示在窗口右下角;单击"重置图像"清空输入与图像显示。

5 测试结果

5.1 默认参数

对于仿射变换,正如第 2.1.1节中所述,共有 7 个参数。由于各个参数的数量级不同,如拉伸变换参数可能小于 1,而平移变换参数可能达到上百,需要对不同的参数设置不同的初始化、学习率、以及第 2.1.3节优化算法中的 Δx 。

为了方便后续统一设置所有参数的学习率,定义学习率比率 γ 。每个参数都有各自的学习率比率,当统一设置学习率 α 为特定的值,每个参数各自的真实学习率为 $\gamma\alpha$ 。 默认参数如表 1所示。

参数名称	第 2.1.1节中的符号	初始化	学习率比率 γ	优化算法中的 Δx
水平拉伸变换参数	c_1	1	1	0.01
垂直拉伸变换参数	c_2	1	1	0.01
旋转变换参数	heta	0	1	0.01
水平剪切变换参数	s_h	0	0.1	0.01
垂直剪切变换参数	s_v	0	0.1	0.01
水平平移变换参数	b_1	0	50000	1
垂直平移变换参数	b_2	0	50000	1

表 1: 默认参数表

5.2 测试图像

本项目使用三组图像来测试算法,如图 2所示。其中,前两组是自然图像,参考图像相同:后两者是医学图像。

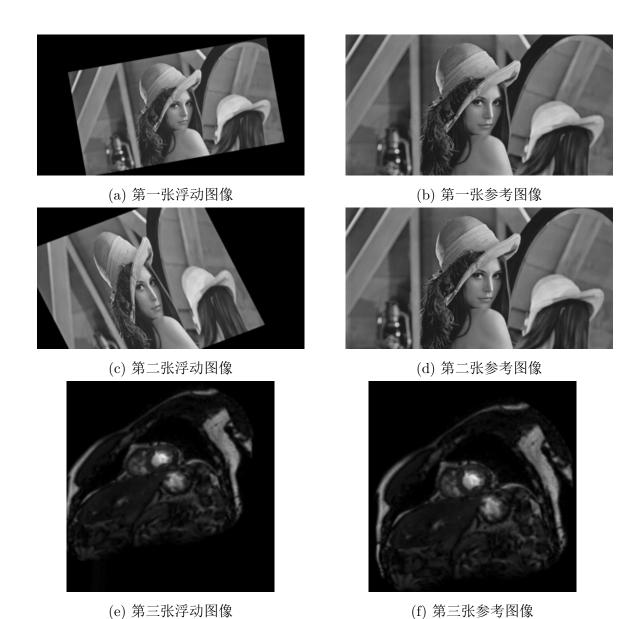


图 2: 测试图像

5.3 仿射变换

5.3.1 均方误差损失函数

对于前两组图像,首先尝试设置初始学习率为 0.2,在第 40 个循环衰减为 0.04,在 第 80 个循环衰减为 0.02,总循环数为 200。测试结果如图 3和 4所示,损失函数变化如图 7的前两条曲线所示。

可以看到,第一张结果图像效果较好,其灰度差图像如图 6右上所示;同样的超参数应用于参考图像相同的第二组图像的结果较差,原因是第二组图像形变程度更大,需要更大的学习率。可见,对于图像配准问题,很难找到适用于各种情况的超参数,这也就是为什么在用户操作界面中加入了用户手动调整超参数的功能。

为了改进第二组图像结果,保持初始学习率为 0.2 不变,改变学习率衰减策略,在

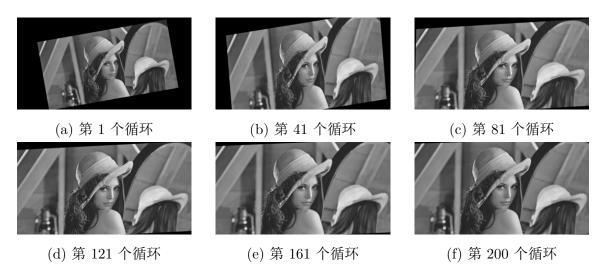


图 3: 均方误差损失函数在第一组图像上的测试结果

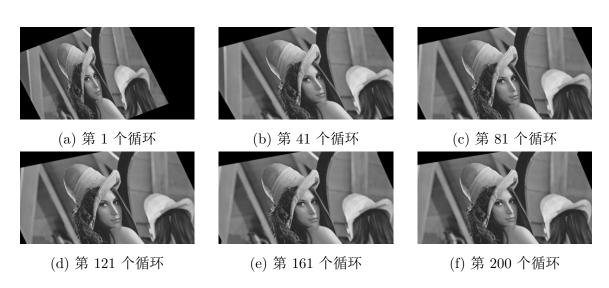


图 4: 均方误差损失函数在第二组图像上的测试结果

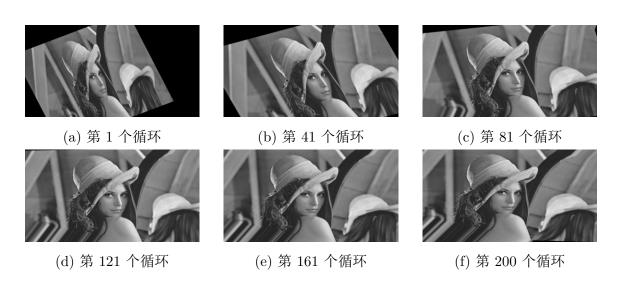


图 5: 均方误差损失函数在第二组图像上的第二次测试结果

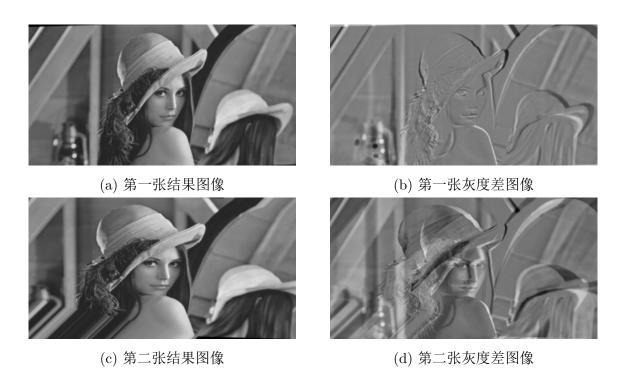


图 6: 均方误差损失函数在前两组图像上的测试结果

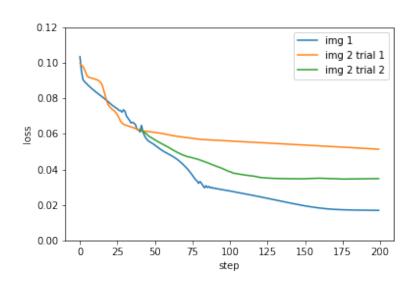


图 7: 均方误差损失函数在前两组图像上的测试结果

第 100 个循环衰减为 0.1,在第 120 个循环衰减为 0.05。测试结果如图 5所示,损失函数变化如图 7最后一条曲线所示,灰度差图像如图 6右下所示。可以看到,第二组图像结果有了很大改进,只是由于浮动图像缺少了左下部分的信息而出现条纹。

接下来,在医学图像上测试均方误差损失函数。设置初始学习率为 0.1,在第 30 个循环衰减为 0.01,总循环数为 40,测试结果如图 8所示。可以看到,效果较好,只是由于浮动图像缺少了左下部分的信息,因此结果图像也缺少了这部分信息。

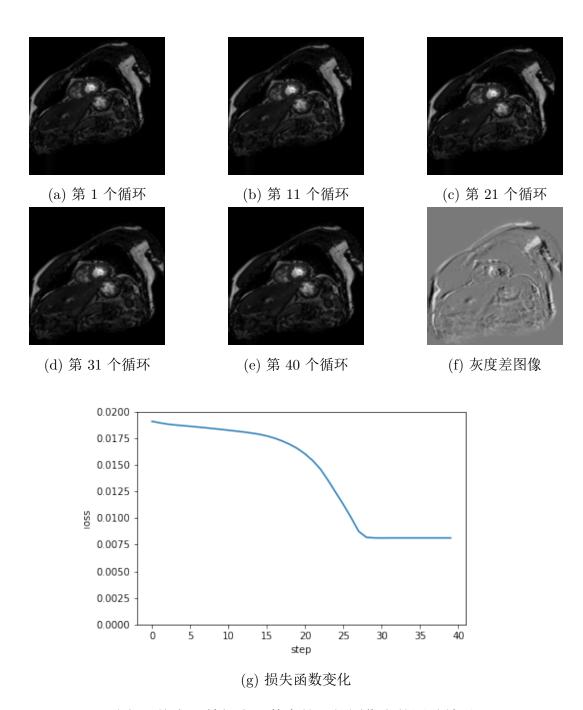


图 8: 均方误差损失函数在第三组图像上的测试结果

5.3.2 负相关系数损失函数

由于前两组测试图像的参考图像相同,本节仅考虑第一组和第三组测试图像。

对于第一组图像,设置初始学习率为 0.01,在第 250 个循环衰减为 0.001,总循环数为 420;对于第三组图像,设置学习率为常数 0.001,总循环数为 100。测试结果如图 9所示,与均方误差损失函数的效果相差不大。

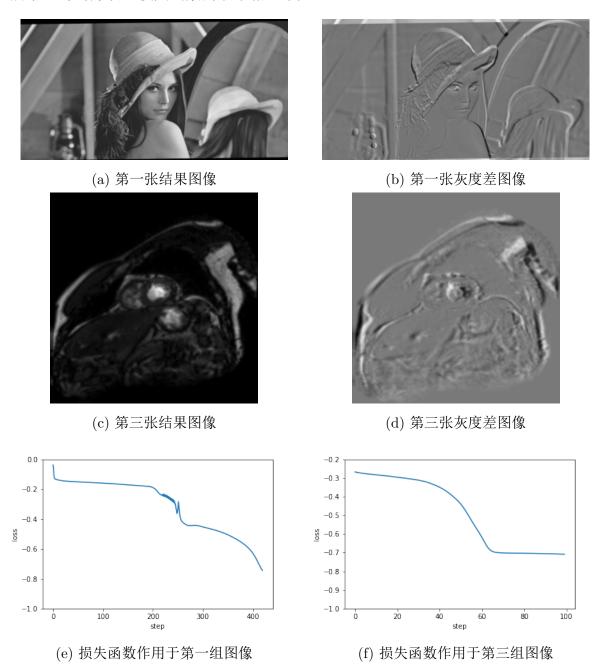
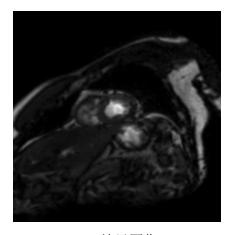


图 9: 负相关系数损失函数的测试结果

5.3.3 负归一化交叉互相关损失函数

正如第 2.1.2节中所述,归一化交叉互相关损失函数的计算复杂度很高,故本节仅考虑第三组测试图像:设置学习率为常数 0.01,总循环数为 100,测试结果如图 10所示,与之前算法的效果相差不大,但是计算复杂度大大增加,并不是一个很好的算法。



(a) 结果图像



(b) 灰度差图像

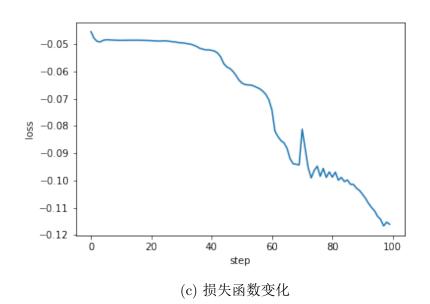


图 10: 负归一化交叉互相关损失函数的测试结果

5.3.4 负互信息损失函数

与第5.3.2节相同,本节仅考虑第一组和第三组测试图像。

对于第一组图像,设置初始学习率为 0.05,在第 10 个循环衰减为 0.025,在第 120 个循环衰减为 0.0125,在第 160 个循环衰减为 0.00625,在第 200 个循环衰减为 0.00125,总循环数为 200;对于第三组图像,设置初始学习率为 0.01,在第 10 个循环衰减为 0.001,总循环数为 80。测试结果如图 11所示,与均方误差损失函数的效果相差不大。

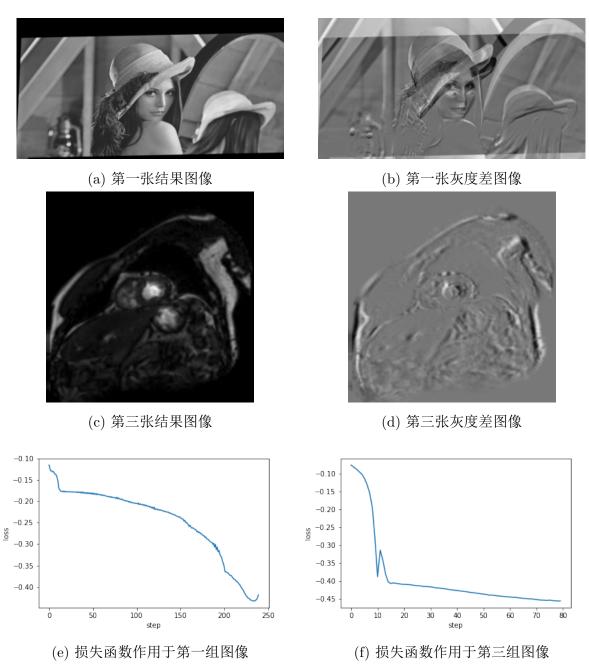


图 11: 负互信息损失函数的测试结果

5.3.5 KL 散度损失函数

正如第 2.1.2节中所述, KL 散度损失函数只适合一些特定的任务。事实上,通过反复测试发现,使用 KL 散度作为损失函数时,算法的收敛情况对超参数过于敏感。以下对第一组图像给出一组超参数和测试结果:

设置初始学习率为 0.01, 在第 10 个循环衰减为 0.001,总循环数为 40,其余超参数如表 2所示。测试结果如图 12所示。可以看到,结果与均方误差等算法差别不大。因而从超参数选择的角度来看,该算法并不是一个很好的算法。

5.4 非线性变换

测试结果如图 13所示。可以看到,第一张结果图像效果很好,只是左下角的物体发生了移位;第二张结果图像效果很不好,主体结构被严重破坏,原因是浮动图像的形变程度较大,一些重要特征被挤压,不适合使用 Coarse-to-Fine 方法进行处理;第三章结果图像效果较好,与仿射变换效果相差不大。

总体而言,对于图像变换方式,仿射变换是全局的,因此无法消除图像之间的局部 几何差异;非线性变换尽管可以改善这一点,比如第一张结果图像,也有可能带来不必 要的结构上的相对移动,比如第二张结果图像。因此,针对特定的图像,应尝试不同的 算法,选择最合适的。

6 总结

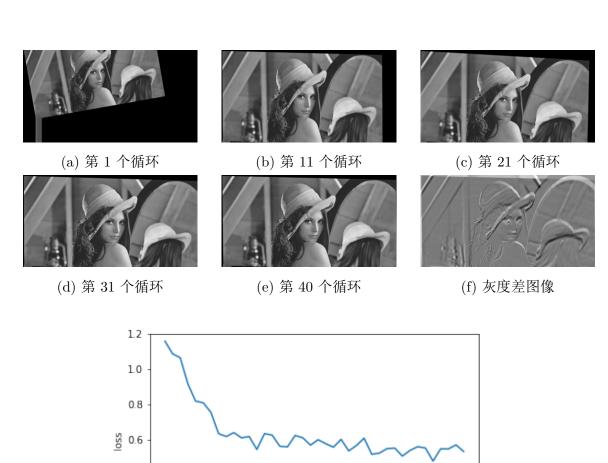
本项目实现了多种图像配准算法,但是难以找到适用于所有问题的算法,也难以找到适用于所有情况的超参数。在处理实际问题时,应针对特定的图像选择不同的算法:若要对图像进行全局变换,而其中的局部几何差异可以忽略不计,则可以选择仿射变换,反之可以尝试非线性变换。同时,超参数选择也非常重要,比如形变程度更大的图像可能需要更大的学习率等等。

对于仿射变换的各种损失函数,测试结果表明,均方误差、负相关系数、负互信息的效果都很好;归一化交叉互相关的计算复杂度太高,而效果并没有在前三者的基础上有明显提升,并不是一个很好的算法; KL 散度对超参数过于敏感,因此也不是一个很好的算法。

最后,考虑到在图 5的结果图像左下角出现的条纹,注意到图像配准并不包含补全缺失信息的功能。若要在结果图像中包含完整的信息,需要在输入的浮动图像中就包含完整的信息,或者结合图像复原算法进行进一步的操作。

参数名称	第 2.1.1节中的符号	初始化	学习率比率 γ	优化算法中的 Δx
水平拉伸变换参数	c_1	1	1	0.01
垂直拉伸变换参数	c_2	1	1	0.01
旋转变换参数	heta	0	1	0.01
水平剪切变换参数	s_h	0	0.01	0.0001
垂直剪切变换参数	s_v	0	0.01	0.0001
水平平移变换参数	b_1	150	1000	10
垂直平移变换参数	b_2	120	1000	10

表 2: 超参数表



(g) 损失函数变化

20

step

25

30

35

40

0.4

0.2

0.0

ò

5

10

15

图 12: KL 散度损失函数的测试结果

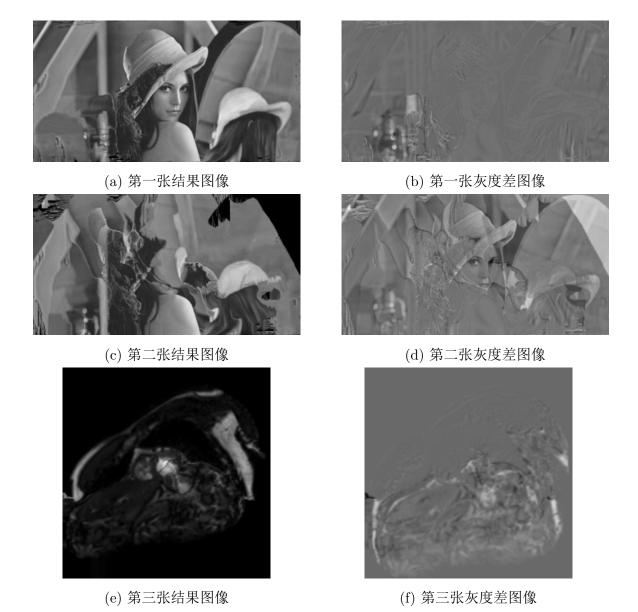


图 13: 非线性变换测试结果

7 组内分工

- 姜海天: 算法实现与实验;
- 龙蕾如: GUI 界面设计、视频制作;
- 王逸群:报告撰写。

参考文献

- [1] Guha Balakrishnan, Amy Zhao, Mert R. Sabuncu, Adrian V. Dalca, and John Guttag. An unsupervised learning model for deformable medical image registration. In 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, jun 2018.
- [2] Leonid I Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: nonlinear phenomena*, 60(1-4):259–268, 1992.
- [3] Christopher Zach, Thomas Pock, and Horst Bischof. A duality based approach for realtime tv-l 1 optical flow. In *Joint pattern recognition symposium*, pages 214–223. Springer, 2007.