

指导教师： 杨 涛 提交时间： 2016/3/21

CVPR2015 Paper

Translation

No: 01

姓名： 吴 鹏

学号： 2013302581

班号： 10011305



稳健性图像配准的深度稀疏表示法

Yeqing Li , Chen Chen , Fei Yang , and Junzhou Huang

德州大学阿灵顿分校

Facebook 有限公司

摘要

在图像配准中对于相似性度量的定义是十分必要的。本文中，我们将为基于两个及以上图像的配准提出一种新的相似性度量定义。这种方法基于最佳配准图像可以通过深入分析灰度领域与频域变化、分离错误的稀疏张量得到的原理。这种相似性度量方法的一项关键优势体现在严重强度失真的稳健性上，光照强度失真这一现象广泛存在于不同获取方式与光照条件下的医学图像、遥感图像、自然图像中。两种有效的算法被设计用来在一个统一的框架中解决批图像配准问题与对图像配准问题。我们使用了广泛的富有挑战性的数据库来验证自己的理论。实验结果表明，我们提出的算法在稳定性、准确性和高效性上已经超过了九个在合成图像方面的传统或新兴的算法以及一大批现实世界中存在的应用。

1. 引言

图像配准是一项图像处理与计算机视觉领域的基本任务。它的目的是将两个或更多图像整合到同一类中以便于这些图像的进一步处理或比较。精确性与稳定性是衡量一种图像配准好坏的最主要的标准。现已证明仅仅 0.3 像素的平均几何失真就会对像素到像素的图像融合处理产生显著地影响。稳健性被定义为在不同条件下的不同试验中接近准确结果的能力。基于配准过程中所使用的特征，现有的方法可以被分成基于特征配准法与基于像素配准法两类。基于特征配准法依赖于从图像中提取特征，然而，可靠的特征提取的方法仍然是一个悬而未决且被热烈讨论的研究话题。本文中，我们寄希望于直接使用像素值来进行图像配准。此外，我们希望尽可能的成功达到对来源于一系列应用的图像的亚像素级配准。

图像配准的一个关键要素是用能量函数测量相似度。最优相似度应该能够正确指向某个条件

约束下的图像空间。然而，输入图像的不确定性使得找出一个可靠的相似度测量方法变得极为困难。在许多现实世界的应用中，待配准的图像是在不同的时间、地点，不同的获取方式，众多的光照条件取得的。结果是，得到的图像的光照强度有着显著的不同。例如，缓慢变化的亮度偏移往往存在于脑磁共振图像中；对于同一片地区的物体在遥感图像中甚至会完全相反；众多的传感器会有不同的波长灵敏度。糟糕的是，许多现有的基于像素配准法对于这些亮度变化并不稳健，例如广泛使用的平方差（SSD）。

最近，稀疏诱导相似度测量方法已多次成功克服类似的配准难题。在 RASL（稀疏低秩分解的稳健性配准）中，图像被矢量化成一个数据矩阵。这种转化是为了寻找一种配对图像的低秩与稀疏表示方法。两种在线配对方法，ORIA（在线稳健性图像配对）和 t-GRASTA（格拉斯曼稳健性自适应子空间跟踪算法变形）是为了提高 RASL 的可扩展性。所有的这些方法均假设图像中只存在较少的巨大误差且这些误差是可分离的（如：阴影，部分遮挡）。然而，正如我们以后将展示的，许多现实世界图像包含严重的空间变化引起的亮度扭曲。这些变化是很多的所以用这些方法很难分离这些误差。结果，以上方法可能无法找到正确的配对图像从而在这些具有挑战性任务中无法表现出较好的稳健性。

残余复杂性（RC）是最好的记录严重的灰度失真图像的方法之一，它是采用离散余弦变换（DCT）来减少两图像的残余。对于一部分图像，RC 算法需要对它们进行一对对的配对并且得到的结果可能并不是最优的。此外，DCT 变换与逆 DCT 变化需要在每一次迭代中进行，这一行为减缓了图像配对的速度。最后，尽管 RC 算法对于亮度失真是稳健的，它处理被遮挡部分的能力仍是未知的。

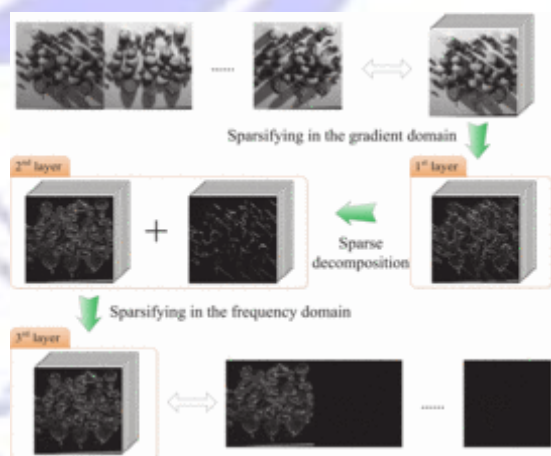
不同于以往的做法每一幅图像矢量化为一个向量，我们把输入图像转换成三维张量以保持图像的空间结构。通过这种做法，最佳配对图像张量可以被分成稀疏频率张量和张量的稀疏误差（见图 1 的更多细节）。重度亮度失真与部分遮挡将在第一第二层被减少并被分离出来，而任何误差会在第三层减少频率张量。我们基于自然图像的类似的深度稀疏表示方法假设了一种新的相似度量方法。对比需要一批输入图像的低秩相似度量方法，新的相似度量方法即使在只有两张输入图像时仍能工作。一种高效的基于增广拉格朗日乘法（ALM）提出了批处理的方法，同时一种回溯式梯度下降法被用来解决一对图像的配对问题。这两种算法在每次迭代中都具有非常低的计算复杂度。我们利用一大批自然图像数据库（包括医疗图像，遥感图像）将新算法对比了 9 种传统的、最新的算法。实验结果表明新算法在解决不同类型的亮度变化时表现出了更好的稳健性并总能比其他算法取得更好的亚像素级准确度。

2. 图像配准的深度稀疏表示

本文中，我们用黑体字表示多维数据。例如， \mathbf{x} 表示多维数据。 \mathbf{X} 表示一个矩阵。 \mathbf{X} 表示 3D 或者三阶张量。 $X_{(i,j,t)}$ 表示第 i 行 j 列和第 t 层。 $\mathbf{X}_{:,:,t}$ 表示整个第 t 层，因此这是一个矩阵。 ℓ_1 表示所有条目的绝对值的总和，它适用于向量，矩阵和张量。

2.1. 批处理模式

为了便于理解我们引入深度稀疏架构的逆架构。假设我们有一批待配准的灰度图像 $I_1, I_2, \dots, I_N \in \mathbb{R}^{w \times h}$ ，其中 N 表示图像总数



图一. 最佳配准图像的深度稀疏表示。首先我

们将图像张量转换成梯度张量（第一层），然后在第二层中分离出稀疏误差张量。重复模式的梯度张量在频域中分析。最后，我们在第三层由傅立叶系数得到一个非常稀疏的频率张量。

首先，我们思考在最简单的情况下，所有的输入图像是相同的并且从一组公式开始转换 $\tau = \{\tau_1, \tau_2, \dots, \tau_N\}$ 。

我们利用公式 $\mathcal{D}_{(:, :, t)} = \mathbf{I}_t, t = 1, 2, \dots, N,$ (1) 将输入图像转换成一个 3D 张量 $\mathcal{D} \in \mathbb{R}^{w \times h \times N}$ 。

删除变换扰动后，部分图像显示出了相同的模式。这样的周期信号在频域中是非常稀疏的。理想情况下，从第二层到最后一片的傅立叶系数都应该是零。我们可以最小化傅立叶系数以得到以下的最佳转换：

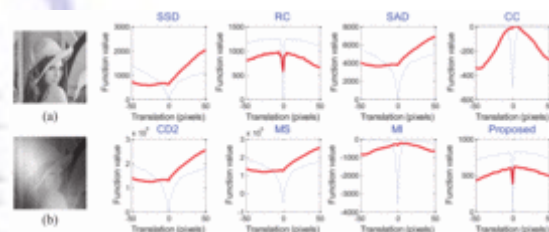
$$\min_{\mathbf{A}, \tau} \|\mathcal{F}_N \mathbf{A}\|_1, \text{ s.t. } \mathcal{D} \circ \tau = \mathbf{A}, \quad (2)$$

其中在第三个方向上， \mathcal{F}_N 表示傅立叶变换

由于图像的失真与部分遮挡，上述模型很难应用在实际情况中。类似以前的情况，相比于失真引起的误差我们假定信噪可以忽略不计。假定 \mathcal{E} 为误差张量，我们可以在它足够稀疏的情况下将它从图像张量中分离出来类似的，我们用 ℓ_1 范数诱导稀疏：

$$\min_{\mathbf{A}, \mathcal{E}, \tau} \|\mathcal{F}_N \mathbf{A}\|_1 + \lambda \|\mathcal{E}\|_1, \text{ s.t. } \mathcal{D} \circ \tau = \mathbf{A} + \mathcal{E}, \quad (3)$$

当 $\lambda > 0$ 时为正则化参数。



图二. 一个用不同的相似度量方法对水平平移的无用配准实例。(a) 图为一张 lena 的 (128×128) 图像。(b) 图为一张严重失真的无用 lena 图像。蓝色曲线为图 (a) 与图(a)之间的配准；红色曲线为图 (b) 与图(a)之间的配准。

上述方法要求的误差 \mathcal{E} 是稀疏的。然而在现实世界中的许多应用中，图像被空间变化的亮度失

真所损坏的。现有的方法如 FasL 和 t-grasta 对于这些非稀疏的错误的分离是无能为力的。新方法的最后阶段的灵感来源于图像梯度（边缘）的位置应该几乎保持不变，即使在严重的亮度失真的情况下。因此，我们在梯度领域对图像进行配准：

$$\min_{\mathcal{A}, \mathcal{E}, \tau} \|\mathcal{F}_N \mathcal{A}\|_1 + \lambda \|\mathcal{E}\|_1, \text{ s.t. } \nabla \mathcal{D} \circ \tau = \mathcal{A} + \mathcal{E}, \quad (4)$$

当 $\nabla \mathcal{D} = \sqrt{(\nabla_x \mathcal{D})^2 + (\nabla_y \mathcal{D})^2}$ 时表示沿二维方向的梯度张量。这是基于假设自然图像的亮度失真是平稳变化的条件下的。

基于这个原理，输入图像可以被稀疏表示在图一所示的一个三层架构中。我们可以称之为图像的深度学习表示法。相比于现在流行的低秩表示法，我们的模型有两个主要优势。第一低秩表示法将每一张图像看做一个一维信号，而我们的模型利用自然图像的空间先验信息（分段光滑）。第二，当输入图像的数量不足以形成一个低秩矩阵时，新方法仍是有效的。接下来我们将演示如何应用新方法在仅有两张输入图像时进行配准。

2.2. 双图模式

在配准两张图像时，新算法可以被进一步简化使得配准过程所需的时间大大缩减。在两点离散傅里叶变换（DFT）中，第一项表示和，第二项表示差。当两张图像已经被配准后所有的步长都远远小于它们的和。我们可以抛弃步长和以求更好的稀疏表示。用 \mathbf{I}_1 表示参考图像， \mathbf{I}_2 表示待配准的原始图像。问题可以被简化成

$$\min_{\mathbf{A}_1, \mathbf{A}_2, \mathbf{E}, \tau} \|\mathbf{A}_1 - \mathbf{A}_2\|_1 + \lambda \|\mathbf{E}\|_1, \quad (5)$$

$$\text{s.t. } \nabla \mathbf{I}_1 = \mathbf{A}_1, \nabla \mathbf{I}_2 \circ \tau = \mathbf{A}_2 + \mathbf{E}.$$

两个常数 ℓ_1 意味着相同的属性，即， \mathbf{E} 残余图像的稀疏。因此，我们可以进一步简化上述能量函数：

$$\min_{\tau} \|\nabla \mathbf{I}_1 - \nabla \mathbf{I}_2 \circ \tau\|_1. \quad (6)$$

一个非常有趣的现象是公式（6）相当于最大限度的减小残留图像的总变化。电视领域已经成功地在许多图像重建和非刚性配准问题上应用了这一发现。

我们在一张无用的配准图像上对新方法与

SSD, RC, SAD（绝对值之和法），CC（相关系数法），CD2, MS, MI（交互信息法）等算法做了比较。Lena 图像与它的平移图像做了配准。图二中的蓝色曲线显示了不同方法所得出的结果，所有这些方法都可以在无变换两图配准中找到最佳配准结果。经过增强亮度失真与缩放后，图二中的图像(b)出现了与原始 lena 图像的不同。红色曲线显示的结果表明只有 RC 算法与新方法能够解决这种亮度失真问题。

3. 算法

3.1. 批处理

由于 τ 变换的非线性，问题（4）很难得到直接解决。我们为每一张图像做了本地的一阶泰勒近似变换：

$$\nabla \mathbf{I}_t \circ (\tau_t + \Delta \tau_t) \approx \nabla \mathbf{I}_t \circ \tau_t + \mathcal{J}_t \otimes \Delta \tau_t \quad (7)$$

其中 $t = 1, 2, \dots, N$ ， $\mathcal{J}_t = \frac{\partial}{\partial \zeta} (\nabla \mathbf{I}_t \circ \zeta) |_{\zeta=\tau_t} \in \mathbb{R}^{w \times h \times p}$ ， τ_t 取决于 p 。

最后一项的张量向量的乘积定义为：

定义 1

张量向量的乘积。张量 $\mathbf{A} \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 与向量 $\mathbf{b} \in \mathbb{R}^{n_3}$ 的乘积是一个矩阵 $\mathbf{C} \in \mathbb{R}^{n_1 \times n_2}$ 。它由公式 $\mathbf{C} = \mathbf{A} \otimes \mathbf{b}$ 得到，其中 $C_{(i,j)} = \sum_{t=1}^{n_3} A_{(i,j,t)} b_{(t)}$ ， $i = 1, 2, \dots, n_1$ ， $j = 1, 2, \dots, n_2$ 。所以，公式（4）可以被写成：

$$\min_{\mathcal{A}, \mathcal{E}, \Delta \tau} \|\mathcal{F}_N \mathcal{A}\|_1 + \lambda \|\mathcal{E}\|_1, \quad (8)$$

$$\text{s.t. } \nabla \mathcal{D} \circ \tau + \mathcal{J} \otimes \Delta \tau = \mathcal{A} + \mathcal{E},$$

这个约束问题可以通过增广拉格朗日乘子法（ALM 算法）解决。增广拉格朗日问题是通过 $(\mathcal{A}^{k+1}, \mathcal{E}^{k+1}, \Delta \tau^{k+1}) = \arg \min_{\mathcal{A}, \mathcal{E}, \Delta \tau} \mathcal{L}(\mathcal{A}, \mathcal{E}, \Delta \tau, \mathcal{Y})$,

$$\mathcal{Y}^{k+1} = \mathcal{Y}^k + \mu^k h(\mathcal{A}^k, \mathcal{E}^k, \Delta \tau^k), \quad (9)$$

迭代得到 $\mathcal{A}, \mathcal{E}, \Delta \tau$ 和 \mathcal{Y} 。

k 是迭代计数器且

$$\mathcal{L}(\mathcal{A}, \mathcal{E}, \Delta \tau, \mathcal{Y}) = \langle \mathcal{Y}, h(\mathcal{A}, \mathcal{E}, \Delta \tau) \rangle + \|\mathcal{F}_N \mathcal{A}\|_1 + \lambda \|\mathcal{E}\|_1 + \frac{\mu}{2} \|h(\mathcal{A}, \mathcal{E}, \Delta \tau)\|_F^2, \quad (10)$$

其中两个张量内积是所有矩阵点的乘积。且

$$h(\mathcal{A}, \mathcal{E}, \Delta \tau) = \nabla \mathcal{D} \circ \tau + \mathcal{J} \otimes \Delta \tau - \mathcal{A} - \mathcal{E}. \quad (11)$$

解决问题（9）的一个通用方法是在同一个时间，最小化一个未知数的影响。而每个子问题都具有

封闭形式的解决方案:

$$\begin{aligned} \mathcal{A}^{k+1} &= \mathcal{T}_{1/\mu^k}(\nabla \mathcal{D} \circ \tau + \mathcal{J} \otimes \Delta \tau + \frac{1}{\mu^k} \mathcal{Y}^k - \mathcal{E}^k) \\ \mathcal{E}^{k+1} &= \mathcal{T}_{\lambda/\mu^k}(\nabla \mathcal{D} \circ \tau + \mathcal{J} \otimes \Delta \tau + \frac{1}{\mu^k} \mathcal{Y}^k - \mathcal{A}^{k+1}) \\ \Delta \tau_t^{k+1} &= \mathcal{J}_t^T \otimes (\mathcal{A}_{(:, :, t)}^{k+1} + \mathcal{E}_{(:, :, t)}^{k+1} - \nabla \mathcal{D}_{(:, :, t)} \circ \tau \\ &\quad - \frac{1}{\mu^k} \mathcal{Y}_{(:, :, t)}^k), \text{ for } t = 1, 2, \dots, N \end{aligned} \quad (12)$$

$\mathcal{T}_\alpha(\cdot)$ 表示阈值 α 的软阈值操作。在 (12) 中的第三个方程中, 我们将张量-矩阵的乘积与张量的转置定义如下:

定义 2

张量-矩阵的乘积。张量 $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 与矩阵 $B \in \mathbb{R}^{n_2 \times n_3}$ 的乘积是向量 $c \in \mathbb{R}^{n_1}$ 。

它由 $c = A \otimes B$ 得到, 其中

$$c_{(i)} = \sum_{j=1}^{n_2} \sum_{t=1}^{n_3} A_{(i,j,t)} B_{(j,t)}$$

定义 3

张量的转置: 张量 $A \in \mathbb{R}^{n_1 \times n_2 \times n_3}$ 的转置是张量 $A^T \in \mathbb{R}^{n_3 \times n_1 \times n_2}$ 。

双图模式的配准算法在算法 1 中得到了总结。令 $M = w \times h$ 为每张图像的像素总数。我们设 $\lambda = 1/\sqrt{M}$, $\mu_k = 1.25^k \mu_0$, $\mu_0 = 1.25/\|\nabla \mathcal{D}\|_2$ 。对内循环, 应用傅里叶变换 (FFT) 公式的复杂度为 $\mathcal{O}(N \log N)$ 。

其他步数的复杂度为 $\mathcal{O}(MN)$ 。因此, 新方法的总的计算复杂度为 $\mathcal{O}(N \log N + MN)$, 对比与在 RASL 中使用 SVD 分解算法时的复杂度 $\mathcal{O}(N^2 M)$ (如果 $M \gg N$)。

算法一 使用 DSR-批处理模式的图像配准

input: Images I_1, I_2, \dots, I_N , initial transformations $\tau_1, \tau_2, \dots, \tau_N$, regularization parameter λ .

repeat

- 1) Compute $\mathcal{J}_t = \frac{\partial}{\partial \zeta} (\nabla I_t \circ \zeta)|_{\zeta=\tau_t}$, $t = 1, 2, \dots, N$;
- 2) Warp and normalize the gradient images:
 $\nabla \mathcal{D} \circ \tau = [\frac{\nabla I_1 \circ \tau_1}{\|\nabla I_1 \circ \tau_1\|_F}; \dots; \frac{\nabla I_N \circ \tau_N}{\|\nabla I_N \circ \tau_N\|_F}]$;
- 3) Use (12) to iteratively solve the minimization problem of ALM:
 $\mathcal{A}^*, \mathcal{E}^*, \Delta \tau^* = \arg \min \mathcal{L}(\mathcal{A}, \mathcal{E}, \Delta \tau, \mathcal{Y})$;
- 4) Update transformations: $\tau = \tau + \Delta \tau^*$;

until Stop criteria

3.2. 双图模式

类似于批处理模式, 可以有

$$\nabla I_2 \circ (\tau + \Delta \tau) \approx \nabla I_2 \circ \tau + \mathcal{J} \otimes \Delta \tau \quad (13)$$

其中 $\mathcal{J} \in \mathbb{R}^{w \times h \times p}$ 表示雅可比行列式。这样, 双图模式是最小化关于 $\Delta \tau$ 的能量函数:

$$E(\Delta \tau) = \|\nabla I_1 - \nabla I_2 \circ \tau - \mathcal{J} \otimes \Delta \tau\|_1 \quad (14)$$

公式 (14) 中的常数 ℓ_1 是非平滑的。我们可以有一个绝对值的近似公式: $|x| = \sqrt{x^2 + \epsilon}$, ϵ 为一个很小的常数 (如 10^{-10})。令 $\mathbf{r} = \nabla I_1 - \nabla I_2 \circ \tau - \mathcal{J} \otimes \Delta \tau$, 我们可以通过链规则得到的能量函数的梯度:

$$\nabla E(\Delta \tau) = \mathcal{J}^T \otimes \frac{\mathbf{r}}{\sqrt{\mathbf{r} \circ \mathbf{r} + \epsilon}} \quad (15)$$

其中 \circ 表示 Hadamard 乘积。注意式 (15) 中的除数是矩阵点。

回溯梯度下降法是用来最简化能量方程 (14), 这已在算法 2 中总结过。令初始步长为 $\eta = 0.8$, $\mu^0 = 1$ 。每次迭代计算的复杂度为 $\mathcal{O}(M)$, 远远小于使用快速余弦变换 (FCT) 的 RC 算法时的复杂度 $\mathcal{O}(M \log M)$ 。类似于批处理模式, 我们使用标准化的图像以排除一般的解决方案。我们在解决批处理模式与双图模式时均使用一个由粗糙到详细的配准框架。

算法 2 使用 DSR-双图模式的图像配准

input: $I_1, I_2, \eta < 1, \tau, \mu^0$.

repeat

- 1) Warp and normalize I_2 with τ ;
- 2) $\mu = \mu^0$;
- 3) Compute $\Delta \tau = -\mu \nabla E(\mathbf{0})$;
- 4) If $E(\Delta \tau) > E(\mathbf{0})$,
 set $\mu = \eta \mu$ and go back to 3);
- 5) Update transformation: $\tau = \tau + \Delta \tau$;

until Stop criteria

4. 实验结果

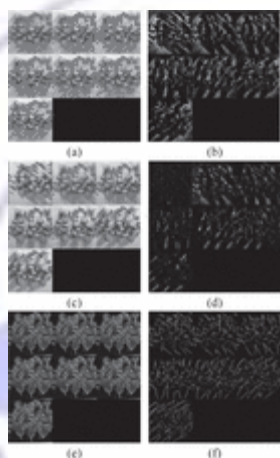
本节中, 我们使用了一系列的应用来验证新算法。并将批处理模式与 RASL、t-GRASTA 算法对比, 双图模式与 RC、SSD 算法对比。新算法的一个最重要的优点是处理随空间变化的亮度失真的自然图像时的稳健性与准确性。正如在图二中显示的, SAD、CC、CD2、MS、MI 算法在处理这类图像问题时表现的相当无力。在以下实验中我们并未使用这些算法。所有试验都是在一台配置为 Intel i7-3770 CPU、12GB RAM 的台式机

上完成的。

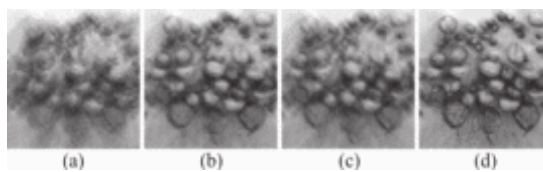
4.1.批图像配准

为了评估新算法在批图像处理模式下的性能，我们采用了一个流行的自然采集图像数据库。我们选中了四个具有最大亮度变化的数据库：“NUTS”，“MOVI”，“FRUITS”，“TOY”。这些数据库在图像配准中是非常具有挑战性的，因为其中包含了 20 种不同的照明环境并且由变化的阴影产生了部分遮挡。四个数据库在使用中还加入了在两个方向上的从均匀分布的 10 像素范围内取得的随机变换图像。

在对“NUTS”数据库进行图像配准后，每种算法的两种核心要素显示在图三中。RASL 与 t-GRASTA 算法没能区别出阴影与较大的误差，而我们能够成功找到最佳配准图像的深度稀疏表示方法。扰动图像和配准结果的平均值如图 4 所示。相比于现有的两种方法，新方法得出的图像有着更清晰的边缘。四个数据库上的定量比较结果被列在表一中，包含超过 20 种运行结果。新算法的整体平均误差都低于 FasL 和 t-grasta 算法。更重要的是，只有新算法能够一直达到亚像素级的准确度。对于 20 张 128*128 像素的图像配准过程来说，新算法与 RASL 算法的大致配准时间均为 7 秒，而 t-GRASTA 算法花费的配准时间大约是 27 秒。尽管我们并未对更大的图像数据库进行试验，但由于更大的数据库具有更高的 SVD 复杂度，RASL 算法将会花费更长的时间。



图三.基于 NUTS 数据库的批图像配准。图 (a) 基于 RASL 算法的低秩元素。图 (b) RASL 算法的稀疏误差。图 (c) t-GRASTA 算法的子空间表示。图 (d) t-GRASTA 算法的稀疏误差。图 (e) 基于新算法的 A 图的可视化结果。图 (f) 新算法的稀疏误差 \mathcal{E}

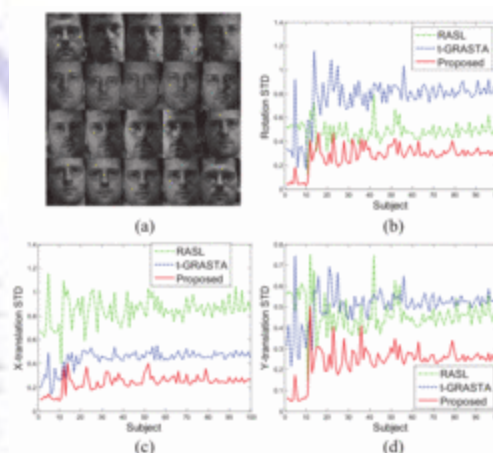


图四.基于 NUTS 数据库的图像配准结果。图 (a) 扰动图像的平均图像。图 (b) 基于 RASL 算法的平均图像。图 (c) 基于 t-GRASTA 算法的平均图像。图 (d) 基于新算法的平均图像。

	RASL	t-GRASTA	Proposed
NUTS	0.670/2.443	1.153/3.842	0.061/0.488
MOVI	0.029/ 0.097	0.568/ 2.965	0.007/0.024
FRUITS	0.050/0.107	1.094/4.495	0.031/0.076
TOY	0.105/ 0.373	0.405/2.395	0.038/0.076

表一.基于数据库的 RASL、t-GRASTA、新算法的平均/最大配准误差。第一张图像被用来评估误差。

我们利用了多饼人脸数据库来评估这三种算法。这个数据库的每一种图像都包含了 20 种不同亮度条件下捕获的图像。我们对第一部分中的前一百个物体加入了 10 度以内的 10 像素级随机人工图像变换。由于最佳配准方式不是唯一的，(如所有图像都有 1 像素以内的误差)。我们比较了配准后的转换图像 STD 值。理想情况下，当所有的扰动都被完全移除时的 STD 应为零。图 5 表明平均图像配准所花的步数超过了 20 步。新算法在配准每一个图像时都比 RASL 与 t-GRASTA 算法表现的要好。

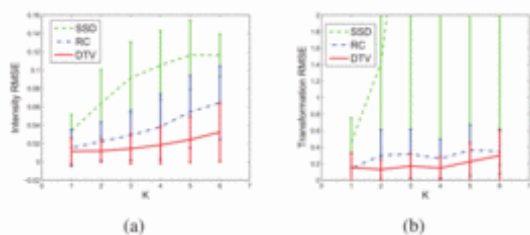


图五. 图 (a) 多饼图像数据库输入的实例。图 (b) 配准后的 STD 转变值。图 (c) 经配准后的 X-变换图像的 STD 值。图 (d) 经配准后的 Y-变换图像的 STD 值。

4.2.双图图像配准

4.2.1 仿真

为了做到定量比较，我们使用了经过随机亮度失真的 lena 图像（图二）与随机仿射变换的 lena 图像（与以前的设置相同的变换）对 SSD、RC、新算法进行了评估。高斯亮度 K 的变化范围为 1~6。无亮度失真的参考图像被用作基准。均方根误差（RMSE）被用作图像亮度变换与图像转换的误差评估的标准。该实验被重复了 50 次，结果记录在图六中。由实验结果可以看出在不同的亮度失真条件下，新算法的表现总是优于 SSD、RC 算法。

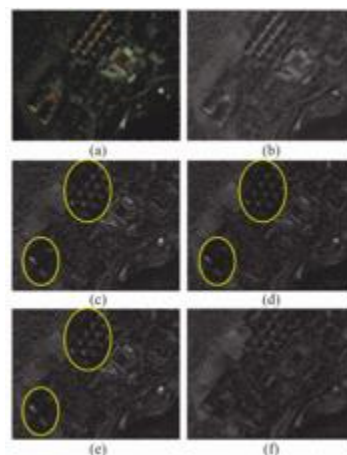


图六. 随机变换扰动和随机亮度失真条件下的配准性能比较。图 (a) lena 图像的亮度均方差 图 (b) lena 图像的仿射变化均方差

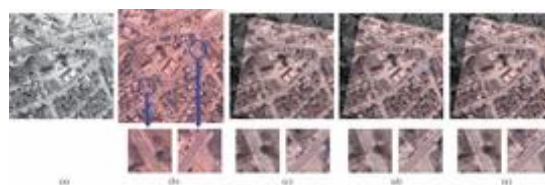
4.2.2 多传感器遥感图像配准

多传感器图像配准是遥感图像融合、变化检测的一项关键预处理方法。在不同的亮度条件下，陆地上的同一个物体的图像可能被多个传感器在不同的时间下捕获。因此，不同亮度下获得的同一个物体图像可能存在显著的差异。在这里，我们对一个多光谱成像卫星 IKONOS 捕获的多光谱图像进行了配准，该图像已在捕获时经过预处理。该多光谱图像具有四个不同波段：蓝、绿、红、近红外，分辨率达到 4 米（见图七中 (a) 图）。平移图像达到了 1 米的分辨率（见图七中 (b) 图）。不同的图像分辨率使得配准过程变得更加困难。从图七中 (c) 图的不同图像中，我们可以观察到在西北方向上存在着偏差。

我们将新算法与 SSD、RC 算法做了对比，对比结果记录在图七中 (d) 图到 (f) 图中。假设图像仅仅经过了纯变换。尽管我们没有实际的地面图像，我们可以从处理后的图像中明显的看到新算法减小了角误差。与此对应的是 RASL、RC 算法并不能在这一点上做到更好。



图七. 多光谱图像与全色图像配准。图 (a) 参考图像。图 (b) 原始图像。图 (c) 配准前差异图像。图 (d) SSD 算法处理的图像。图 (e) RC 算法处理的图像。图 (f) 新算法处理的图像。可视误差被黄色圆圈圈出。在彩色 PDF 文件中可以看到更清晰的 2 倍大的图像。



图八. 航拍图像和数字图像配准。从左到右所示图像分别是：参考图像，原始图像，通过 MATLAB 叠加处理后图像，RC 处理后图像，新方法处理后图像。第二行为放大后的街道 A 和 B。在彩色 PDF 文件中可以看到更清晰的 2 倍大的图像。

我们对航拍图像与数字图像进行了配准。参考图像为 MassGIS 图像。原始图像是没有经过任何特定变换的地表数字图像。输入图像与处理结果显示在图八中。其中 MATLAB 使用了手动选取控制点以便配准而 RC 算法与新算法采用了主动选取控制点。一开始，所有的算法都取得了较好的图像配准结果。更进一步的观察发现新算法相对于其他算法有着更好的准确性。在原始图像中，可以明显观察到街道 A、B 上的两行车道。由于角误差的存在，经过配准处理后 MATLAB 得到的图像中的 B 街道与 RC 得到的 A 街道很是模糊。而即使图像中存在车辆误差，新算法仍然表现出了良好的稳健性。

5. 结语与讨论

本文中，我们为提高图像配准处理的稳健性与准确性提出了一种新的相似度测量方法。新方

法的灵感来源于最佳配准图像的深度稀疏表示。新算法具有以下三个主要优点：（1）相比于现有算法，新算法能够处理严重亮度失真与部分遮挡的图像问题；（2）新算法适用于包含各种类型的图像转换的批图像模式或双图模式的配准；（3）新算法的低计算复杂度使它适用于巨大数据库的图像配准处理。我们已经利用了一大批富有挑战性的图像数据库对新算法进行了测试。令人振奋的实验结果显示新算法在批处理模式与双图模式下的稳健性与准确性超过了现有的先进算法。我们同样展示了新算法可用于减少许多现实应用的图像配准误差。

由于最优化过程中的局部线性化，正如其他对比算法一样，新算法无法有效处理较为显著地图像转换问题。然而，对于许多现实世界应用来说这一问题并没有造成太大影响。例如，遥感图像可能因为地理坐标的不同而产生较大的偏差。对于有显著变换的图像而言，我们可以使用基于FFT的算法进行初步的粗略配准再使用新算法进行更进一步的图像配准。因此，我们并未更进一步的测试出新算法能够同时处理的图像的最大数量。到目前为止，新算法只适用于离线配准。如何将新算法从离线配准模式扩展到在线配准模式将是未来研究的一个有趣的方向。

参考目录

[1] <http://www.robots.ox.ac.uk/~vgg/research/affine/>. 5

[2] <http://www.mathworks.com/help/images/register-an-aerial-photograph-to-a-digital-orthophoto.html>. 7

[3] P. Blanc, L. Wald, T. Ranchin, et al. Importance and effect of co-registration quality in an example of pixel to pixel fusion process. In 2nd International Conference “Fusion of Earth Data: merging point measurements, raster maps and remotely sensed images”, pages 67–74, 1998. 1

[4] E. J. Candes, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM*, 58(3):11, 2011. 4

[5] C. Chen, Y. Li, W. Liu, and J. Huang. Image fusion with local spectral consistency and dynamic gradient sparsity. In

IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 2760–2765, 2014. 6

[6] B. Cohen and I. Dinstein. New maximum likelihood motion estimation schemes for noisy ultrasound images. *Pattern Recognition*, 35(2):455–463, 2002. 1, 3, 5

[7] R. Gross, I. Matthews, J. Cohn, T. Kanade, and S. Baker. Multi-pie. *Image and Vision Computing*, 28(5):807–813, 2010. 6

[8] V. Hamy, N. Dikaïos, S. Punwani, A. Melbourne, A. Latifoltojar, J. Makanyanga, M. Chouhan, E. Helbren, A. Menys, S. Taylor, et al. Respiratory motion correction in dynamic mri using robust data decomposition registration—application to dce-mri. *Medical image analysis*, 18(2):301–313, 2014. 2

[9] J. He, D. Zhang, L. Balzano, and T. Tao. Iterative grassmannian optimization for robust image alignment. *Image and Vision Computing*, 32(10):800–813, 2014. 1, 2, 3, 5

[10] J. Huang, X. Huang, and D. Metaxas. Simultaneous image transformation and sparse representation recovery. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–8, 2008. 1

[11] J. Huang, S. Zhang, H. Li, and D. Metaxas. Composite splitting algorithms for convex optimization. *Computer Vision and Image Understanding*, 115(12):1610–1622, 2011. 3

[12] J. Huang, S. Zhang, and D. Metaxas. Efficient MR image reconstruction for compressed MR imaging. *Medical Image Analysis*, 15(5):670–679, 2011. 3

[13] J. Kim and J. A. Fessler. Intensity-based image registration using robust correlation coefficients. *IEEE Transactions on Medical Imaging*, 23(11):1430–1444, 2004. 3, 5

[14] Y. Li, C. Chen, J. Zhou, and J. Huang. Robust image registration in the gradient domain. In *Proceedings of the International Symposium on Biomedical Imaging (ISBI)*. 2015. 3

[15] J. Ma, W. Qiu, J. Zhao, Y. Ma, A. Yuille, and Z. Tu. Robust L2E estimation of transformation for non-rigid registration. *IEEE Transactions on Signal Processing*, 63(5):1115–1129, 2015. 1

[16] J. Ma, J. Zhao, J. Tian, A. L. Yuille, and Z. Tu. Robust point matching via vector field consensus. *IEEE*

Transactions on Image Processing, 23(4):1706–1721, 2014.

1

[17] A. Myronenko and X. Song. Intensity-based image registration by minimizing residual complexity. *IEEE Transactions on Medical Imaging*, 29(11):1882–1891, 2010.

1, 2, 3, 5, 6

[18] A. Myronenko, X. Song, and D. J. Sahn. Maximum likelihood motion estimation in 3d echocardiography through non-rigid registration in spherical coordinates. In *Functional Imaging and Modeling of the Heart*, pages 427–436. 2009. 3, 5

[19] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(11):2233–2246, 2012. 1, 2, 3, 4, 5

[20] A. Sotiras, C. Davatzikos, and N. Paragios. Deformable medical image registration: A survey. *IEEE Transactions on Medical Imaging*, 32(7):1153–1190, 2013. 1

[21] Space-Imaging. IKONOS scene po-37836. Geoeye IKONOS Scene Data, 2000. 6

[22] C. Studholme, C. Drapaca, B. Iordanova, and V. Cardenas. Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change. *IEEE Transactions on Medical Imaging*, 25(5):626–639, 2006. 1

[23] R. Szeliski. Image alignment and stitching: A tutorial.

Foundations and TrendsR in Computer Graphics and Vision, 2(1):1–104, 2006. 1, 3, 5, 6

[24] C. Thomas, T. Ranchin, L. Wald, and J. Chanussot. Synthesis of multispectral images to high spatial resolution: A critical review of fusion methods based on remote sensing physics. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1301–1312, 2008. 1

[25] G. Tzimiropoulos, V. Argyriou, S. Zafeiriou, and T. Stathaki. Robust FFT-based scale-invariant image registration with image gradients. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(10):1899–1906, 2010. 1, 8

[26] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International Journal of Computer Vision*, 24(2):137–154, 1997. 1, 3, 5

[27] Y. Wu, B. Shen, and H. Ling. Online robust image alignment via iterative convex optimization. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1808–1814, 2012. 1, 2

[28] Y. Zheng, E. Daniel, A. A. Hunter III, R. Xiao, J. Gao, H. Li, M. G. Maguire, D. H. Brainard, and J. C. Gee. Landmark matching based retinal image alignment by enforcing sparsity in correspondence matrix. *Medical image analysis*, 18(6):903–913, 2014. 1

[29] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003.

