

## 基于无监督学习的风洞压敏漆图像配准算法

刘康<sup>1,2,3</sup>, 孙熊伟<sup>1,2,3\*</sup>, 施海亮<sup>1,2,3\*\*</sup>, 王先华<sup>1,2,3</sup>, 叶函函<sup>2,3</sup>, 程晨<sup>1,2,3</sup>, 朱锋<sup>2,3</sup>, 吴时超<sup>2,3</sup><sup>1</sup>中国科学技术大学, 安徽合肥 230026;<sup>2</sup>中国科学院合肥物质科学研究院安徽光学精密机械研究所, 安徽合肥 230031;<sup>3</sup>中国科学院通用光学定标与表征技术重点实验室, 安徽合肥 230031

**摘要** 为满足压强敏感涂料(简称“压敏漆”)图像准确、快速的配准需求,提出一种基于无监督学习的压敏漆图像配准方法——PIR-Net。首先,针对风洞环境中零件的大程度形变问题,构造多尺度的网络框架,实现压敏漆图像由粗到细的配准;其次,构造类似U-Net结构的卷积网络,实现图像宏观结构特征和局部细节特征的融合;最后,设计基于结构相似性的损失函数并引入梯度场约束,以提高特征稀疏图像的配准精度。实验结果表明,相较于领域内的典型配准方法,所提方法的归一化互相关指标提高了16.4%,均方误差提高了19.1%。

**关键词** 机器视觉; 压敏漆图像; 非刚性变形; 图像配准; 无监督学习

中图分类号 TP391 文献标志码 A

DOI: 10.3788/AOS231885

## 1 引言

基于压强敏感涂料(PSP;简称“压敏漆”)的测压方法被广泛应用于风洞环境中零件的表面压力测量<sup>[1-4]</sup>。涂有压敏漆的零件表面在激发光照下会发出荧光,根据荧光强度图像利用Stern-Volmer公式就能反演出压力<sup>[5]</sup>。然而,风洞环境中零件存在位移和非刚性变形,导致公式中有风和无风图像比值的计算产生误差。对有风和无风图像的精确配准是实验数据处理的基础和核心。

典型的图像配准方法包括基于特征匹配<sup>[6]</sup>、迭代最优化<sup>[7-10]</sup>以及基于变换域的方法<sup>[11]</sup>,而提升针对特定场景的图像配准方法的适用性和准确性仍然是一个技术难点。压敏漆图像仅包含明亮的模型发光区域和暗黑的背景区域两种明显的组分,且零件表面平滑、少有纹理,导致压敏漆图像特征相对稀疏,现有依赖特征点的方法难以适用。针对压敏漆图像非刚性变形问题,Liu等<sup>[12]</sup>提出了基于多项式变换的方法,以模型的测压孔为标志点建立 $n$ 阶多项式进行配准,然而受限于标志点数量,整体多项式无法准确描述不同位置的非刚性变形。Fujimatsu等<sup>[13]</sup>通过手动点选测压孔和边角等特征点建立两幅图像的变换关系,然而单次实验图像数量达到数万幅,手动点选耗时费力且存在定位误差,无法满足时效性要求。Suzuki等<sup>[14]</sup>提出一种基于边界检测的特征点自动化检测方法,该方法虽然增

加了关键点的检测数量,但是集中分布在边缘的特征点难以实现对整体结构的非刚性配准。

深度学习技术的快速发展为图像配准提供了新思路,已被广泛应用于医学影像<sup>[15-17]</sup>、遥感图像<sup>[18]</sup>等领域。深度学习技术主要分为有监督配准和无监督配准两个方向<sup>[19]</sup>。有监督配准方法的损失函数依赖标志点或真实变形场的先验信息<sup>[20-21]</sup>,然而在风洞环境中制作标志点较为困难且会破坏零件的气动特性,模拟变形场也不能描述真实的非线性变形<sup>[22]</sup>。无监督配准方法在网络训练时无需先验信息,以基准和变形图像的相似性为损失函数<sup>[23]</sup>,如de Vos等<sup>[24]</sup>将其应用于医学图像领域,在训练阶段就能够实现模型的无监督学习。基于该思路,Balakrishnan等<sup>[25-26]</sup>开发了一种用于磁共振(MR)脑图谱的图像配准方法——VoxelMorph,Han等<sup>[27]</sup>提出一种双通道网络框架并用于MR-CT图像配准。然而,这些方法难以处理大程度、宽动态范围的位移和非线性变形,随着卷积层的加深,浅层像素的影响会迅速减弱<sup>[28]</sup>,通常需要使用仿射变换后的预处理图像进行训练<sup>[25]</sup>,直接应用会导致配准误差。

压敏漆图像存在大程度、宽动态范围的非刚性形变且图像特征稀疏,单次实验图像数量大,现有方法无法满足其准确、快速的配准需求。因此,本文提出一种基于无监督学习的压敏漆图像快速配准方法PIR-Net(PSP image registration network)。首先,针对风洞中零件大程度的偏移和非刚性形变问题,构造了多尺度

收稿日期: 2023-12-04; 修回日期: 2024-02-02; 录用日期: 2024-02-23; 网络首发日期: 2024-03-13

基金项目: 中国科学院合肥物质科学研究院院长基金(YZJJ202404-TS)

通信作者: \*xiongweisun@163.com; \*\*hlshi@aiofm.ac.cn

配准网络,以实现由粗到细的配准策略;其次,构造类 U-Net 结构的卷积网络来学习从输入图像到输出变形场之间的复杂映射,利用多层结构和跳连接实现压敏漆图像宏观结构特征和局部细节特征的融合;最后,针对压敏漆图像特征稀疏的特性,设计基于结构相似性的损失函数并引入梯度场约束来提高配准的全局最优化。

## 2 原理与方法

压敏漆测压实验中,利用相机来采集零件表面荧光强度的变化,实验过程及数据处理流程如图 1 所示。将风洞启动时相机采集的有风图像(windy image)记为变形图像  $M$ ,风洞停止时相机采集的无风图像(windless image)记为固定图像  $F$ ,实验数据处理的首要任务是对有风和无风图像进行配准。

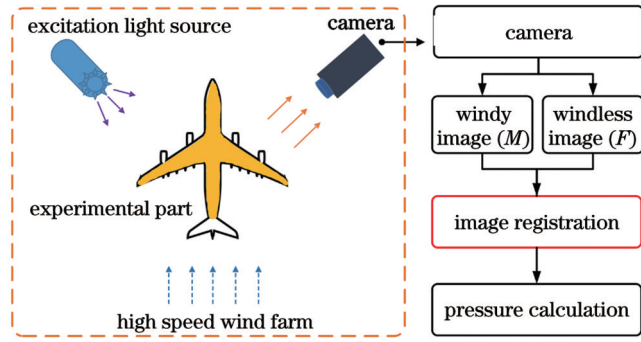


图 1 PSP 实验数据处理流程

Fig. 1 Flow of PSP experiment data processing

实现  $M$  和  $F$  配准的关键是计算出由  $F$  到  $M$  的变形场。传统的非刚性图像配准方法基于最大化相似度量度的原理,对每一幅变形图像进行独立优化,通过持续迭代得到最优的变换参数,配准函数可以表示为

$$\hat{\phi} = \operatorname{argmax}_{\phi} S_{\text{sim}}[F, M(\phi)], \quad (1)$$

式中: $\phi$  为变形场,表示从  $F$  到  $M$  的像素点之间的映射; $\hat{\phi}$  为  $\phi$  的最优值; $M(\phi)$  为使用变形场对变形图像进行变换后的配准图像; $S_{\text{sim}}[F, M(\phi)]$  表示  $F$  和  $M(\phi)$  的相似度。

基于深度学习的配准把传统方法中对每个图像对的独立优化用一个深度学习模型  $g_{\theta}$  替换, $\theta$  为模型的参数。通过卷积网络直接学习有风-无风图像对与变形场之间的关系,配准函数可以表示为

$$\phi = g_{\theta}(F, M). \quad (2)$$

PIR-Net 中  $g_{\theta}$  被定义为一个多尺度的配准网络,在网络训练阶段通过学习大量有风-无风图像对数据不断地评估  $M(\phi)$  和  $F$  之间的损失来更新  $\theta$ ,最终得到最优的网络参数,整个过程不需要额外的监督信息。

在应用阶段,将  $F$  和  $M$  输入到已训练好的深度学习模型  $g_{\theta}$  中,可以直接获得非刚性配准的变换参数  $\phi$ ,最后通过变形场映射得到配准图像。

### 2.1 多尺度配准框架

PIR-Net 构造多级网络结构用于解决风洞环境中试验件大程度、宽动态范围的非刚性形变问题,在满足时效性的基础上实现了由粗到细的配准。同时,设计了多级损失结构来对每一级网络进行优化,实现了特征的层次化提取和学习,网络的整体结构如图 2 所示。配准网络分 3 个阶段学习和提取图像的多尺度结构性特征,每个阶段分别利用不同尺度的卷积网络学习从输入图像到输出变形场之间的复杂映射,得到对应尺度的配准图像,并将其与基准图像的结构相似度作为当前尺度下的损失来优化网络。

在构建的 PIR-Net 三层网络结构中,第一层网络是对变形图像进行粗配准。具体如下:对  $F$  和  $M$  进行  $1/4$  下采样,输入到第一个卷积网络中计算变形场  $W_1$ ;对  $W_1$  进行 4 倍上采样,使其尺寸与  $F$  相同,根据变形场把  $M$  校正为  $M_{d1}$ ,并与固定图像进行第一次对齐;将  $M_{d1}$  和  $F$  的相似度作为第一级卷积网络的损失,反向传播对网络参数进行优化。第二层网络是对  $F$  和  $M_{d1}$  执行进一步的配准。具体的方法是对  $F$  和  $M_{d1}$  进行  $1/2$  下采样,使用第二层卷积网络计算变形场  $W_2$ ,根据  $W_2$  得到第二幅校正图像  $M_{d2}$ 。相似地,将  $M_{d2}$  和  $F$  的相似度作为第二层卷积网络的损失,优化其网络参数。第三层网络是执行  $F$  和  $M_{d2}$  更精细的配准。这一层网络保持原图大小输入到卷积网络中计算变形场  $W_3$ ,最终得到图像  $F$  和  $M$  的配准结果。这三层网络结构共同构成了 PIR-Net 的多级配准策略,旨在处理不同尺度上的变形配准,通过多尺度特征提取和级联回归实现整体配准效果的优化。

### 2.2 卷积网络结构

为了实现 PIR-Net 中图像的多尺度配准和变形场的端到端计算,构造了类似 U-Net<sup>[29]</sup> 结构的卷积网络作为各尺度下的配准子结构,网络结构如图 3 所示。使用类似 U-Net 的卷积网络学习输入图像到输出变形场的映射方法,其更适用于多种场景下压敏漆图像中宽动态范围的非刚性形变的配准问题。同时,其多层结构和跳连接实现了图像的宏观结构特征和局部细节特征的融合,提高了对非刚性形变的适应性。相较于其他卷积模型,类似 U-Net 的卷积网络具有结构简单、训练和推理速度快等优势,更能满足压敏漆图像准确且快速配准的需求。

卷积网络输入为一组设定尺度的  $F$ 、 $M$  图像,输出为变形场。图 3 中矩形表示二维特征,矩形左侧标记为分辨率,上方为特征维数。在编码和解码阶段都使用了二维卷积,卷积核大小为  $3 \times 3$ ,步长为 2,卷积后面连接一个 ReLU 层,提高了网络的非线性学习能力,

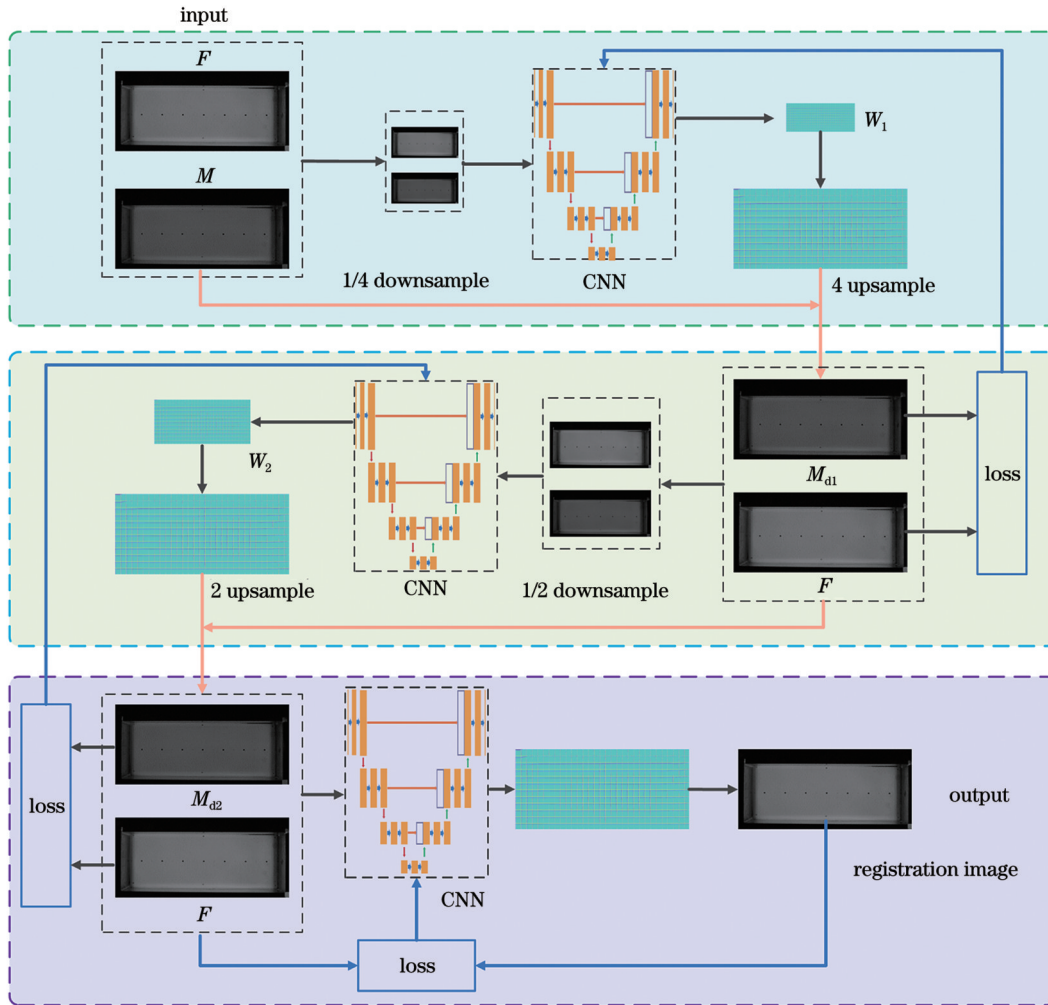


图 2 PIR-Net 结构

Fig. 2 Structure of PIR-Net

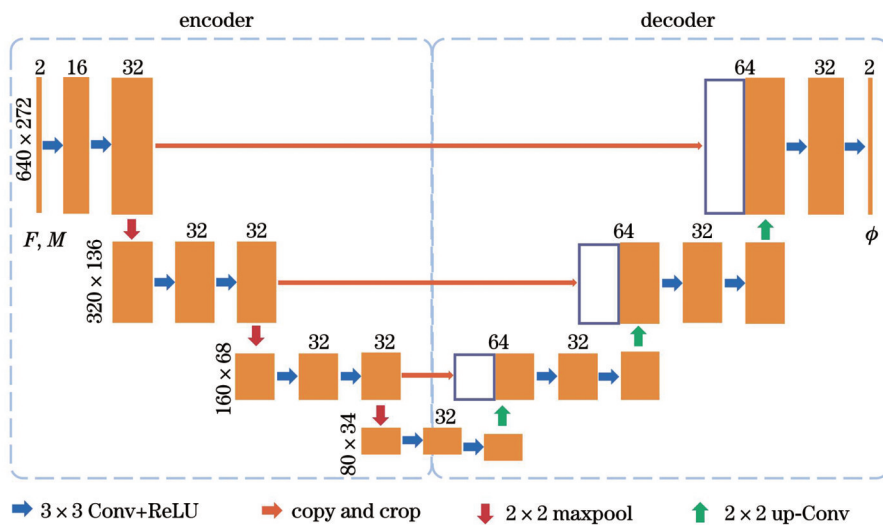


图 3 卷积网络结构

Fig. 3 Structure of convolutional networks

防止梯度消失。在本实验中,输入的  $F$ 、 $M$  均为  $640 \times 272$  大小的单通道图像,但是网络结构不受特定输入图像大小的限制。

卷积网络前半部分的作用是特征提取,后半部分的作用是特征融合。编码阶段堆叠  $F$ 、 $M$  到  $640 \times 272 \times 2$  大小,进行两次卷积(Conv)操作以增加特征的



维度,接着通过一次  $2 \times 2$  的池化(maxpool)进行下采样以增强特征的稳定性。按照相同的操作,经过三次卷积和下采样实现更宏观的结构特征编码,最后的特征尺寸为  $80 \times 34 \times 32$ ,联合的多层结构满足了配准图像对多尺度特征的需求。解码阶段首先对  $80 \times 34 \times 32$  的特征进行一次上采样(up-Conv),使图像尺寸增大 2 倍,恢复特征图的原始分辨率以匹配左侧特征维度;然后,通过跳连接(copy and crop)操作,叠加左右两侧特征图,将编码阶段学习到的特征传播到生成配准层中,实现位置信息与语义信息的融合;最后,将底层特征的维度减小一半,保证与上一层的特征通道数一致,通过编码和解码两个阶段得到  $640 \times 272 \times 2$  的变形场。

### 2.3 变形场映射

变形场是固定图像到变形图像的点对点映射,根据卷积网络输出的变形场,利用坐标映射可以得到校正图像。变形场每一点包含的两个数据  $m, n$  分别表示这点在  $(x, y)$  方向上的位移。对于大小为  $[w, h]$  的二维图像,其变形场尺寸为  $[w, h, 2]$ ,对于校正图像中任意像素点  $p(x, y)$ ,对应变形图像中的位置应为  $p'(x', y')$ ,其中  $x' = x + m_p, y' = y + n_p$ 。由于像素点位置坐标在映射过程中会产生小数,此处引入双线性插值的方法来计算  $p$  点的像素值,计算公式为

$$M_d(p) = \sum_{k \in Z(p')} M(k) (1 - |x_{p'} - x_k|) (1 - |y_{p'} - y_k|), \quad (3)$$

式中:  $M_d(p)$  表示校正图像中  $p$  点的像素值;  $Z(p')$  表示点  $p'$  的 4 个邻域位置;  $M(k)$  表示变形图像中邻域位置的像素值;  $x_{p'}$  和  $y_{p'}$  分别表示  $p'$  点的坐标值;  $x_k$  和  $y_k$  分别表示  $k$  点的坐标值。根据变形场的坐标映射,利用像素插值就可以计算出校正图像每一个像素位置的像素值,从而得到配准图像。

### 2.4 损失函数

本文在基于图像相似度损失的基础上引入空间平滑项,进一步约束图像配准过程中结构性偏差带来的损失。针对压敏漆图像特征稀疏的特性,以梯度场强度为基础构建了配准损失的注意力机制,引导匹配的全局最优化。损失函数可以表示为

$$L(F, M, \phi) = -L_{\text{sim}}[F, M(\phi)] [1 + \lambda_2 L_{\text{grad}}(F)] / 2 + \lambda_1 L_{\text{smooth}}(\phi), \quad (4)$$

式中:  $L_{\text{sim}}[F, M(\phi)]$  为配准图像与固定图像的结构相似度,其目的是最大化两幅图像的相似性;  $\lambda_1$  和  $\lambda_2$  为正则化项;  $L_{\text{smooth}}(\phi)$  表示变形场的平滑性,表征变形场的平滑和局部取向一致性;  $L_{\text{grad}}(F)$  为固定图像的梯度强度信息。

图像相似度  $L_{\text{sim}}$  使用局部归一化互相关进行描

述,计算公式为

$$L_{\text{sim}} = \frac{1}{W \times H \times n^2} \times \sum_{p \in \Omega} \sum_{p_i} \frac{[F(p_i) - \hat{F}(p)] \times [M_d(p_i) - \hat{M}_d(p)]}{\sigma_{F(p)} \times \sigma_{M_d(p)}}, \quad (5)$$

式中:  $W, H$  分别表示图像的宽度和高度;  $p_i$  表示点  $p$  在  $n \times n$  邻域上的像素点,本实验将  $n$  设置为 9;  $\Omega$  表示图像中所有像素点的集合;  $\hat{F}(p)$  和  $\hat{M}_d(p)$  分别为固定图像和配准图像中  $p$  点邻域内像素值的平均值;  $\sigma_{F(p)}$  和  $\sigma_{M_d(p)}$  分别为固定图像和配准图像中  $p$  点邻域内像素值的方差,计算公式为

$$\begin{cases} \sigma_{F(p)} = \sqrt{\frac{1}{n^2} \sum [F(p_i) - \hat{F}(p)]^2} \\ \sigma_{M_d(p)} = \sqrt{\frac{1}{n^2} \sum [M_d(p_i) - \hat{M}_d(p)]^2} \end{cases}. \quad (6)$$

变形场平滑性约束  $L_{\text{smooth}}(\phi)$  描述变形场的光滑性和拓扑结构的一致性,用于表征变形场的约束情况,避免局部相似度过高而产生宏观的结构性配准错误,破坏图像原有的特征分布。变形场平滑性约束的计算公式为

$$L_{\text{smooth}}(\phi) = \|\nabla \phi(p)\|^2. \quad (7)$$

对于固定图像  $F$ , 计算图像的梯度  $\partial F$ , 然后对数值进行归一化。图像梯度是图像变化的幅值,体现了图像的轮廓和边缘特征。

## 3 实验结果与分析

### 3.1 实验数据及实验参数设置

风洞实验中,在风洞启动前采集一幅图像作为固定图像,启动风洞后在连续的时间内使用高速相机采集一系列图像作为变形图像。本文引入两组典型零件的压敏漆实验图像,每组构建了 20000 个图像对,用于配准网络的训练和测试,图像尺寸均为  $640 \times 272$ , 格式为 16 位灰度图。配准过程只使用图像对信息,无外部监督信息辅助。

第一组典型零件为具有较高结构强度的壳体零件,如图 4 所示;第二组为表面特征稀疏且工件刚性低的薄板零件,如图 5 所示。图 4(c) 和图 5(c) 为无风图像和有风图像的叠加结果,直观地展示了有风图像中零件相对于无风图像存在的位移和变形。从叠加图像中零件边缘轮廓和分布在中间的测压孔的偏移可以看出,有风图像相对于无风图像整体存在较大的非刚性变形。

实验中使用的计算机内存为 16 GB,搭载了英特尔 i7 处理器、RTX3070 显卡。使用 Pytorch 框架基于 Python 3.7 进行配准网络的构建、训练和测试。网络使用 ADAM 优化器,学习率为  $1 \times 10^{-4}$ , 一个 batch 为一个图像对。



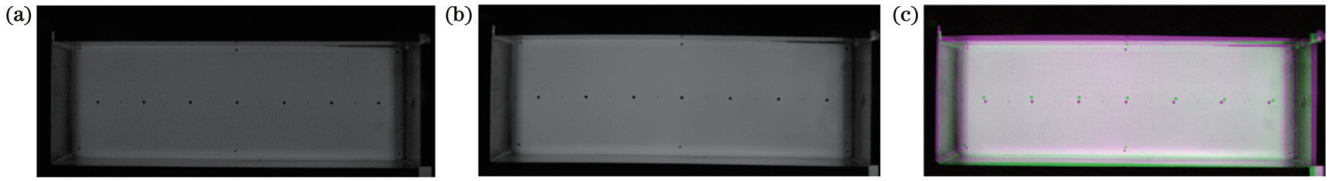


图 4 壳体零件PSP图像。(a)有风图像;(b)无风图像;(c)叠加显示图像

Fig. 4 PSP images of shell part. (a) Windy image; (b) windless image; (c) superimposed display image

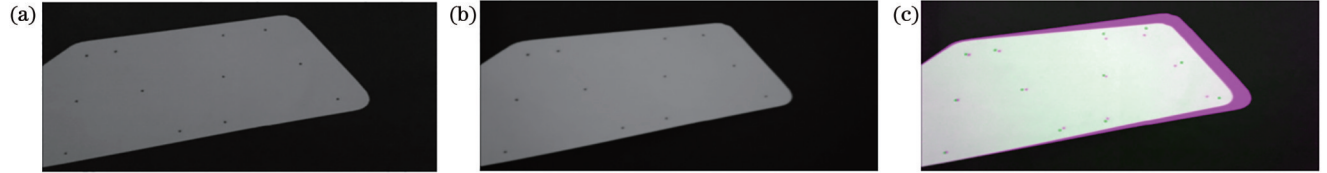


图 5 薄板零件PSP图像。(a)有风图像;(b)无风图像;(c)叠加显示图像

Fig. 5 PSP images of sheet part. (a) Windy image; (b) windless image; (c) superimposed display image

### 3.2 对比方法及评价指标

为了评估本文方法的有效性,将本文方法与压敏漆图像配准典型算法进行对比分析,对比算法包括基于特征的匹配算法、基于灰度和B样条结合的MI-Bspline算法、基于深度学习的配准模型Voxelmorph、CycleMorph、BIRGU-Net<sup>[15]</sup>和LRN<sup>[30]</sup>。

引入配准精度和计算时间对图像配准结果进行评价,配准精度使用3种常用的定量指标进行评价:均方误差(RMSE)、归一化互相关(NCC)指标和目标配准误差(TRE)。RMSE的计算公式如式(8)所示,值越小表明待评估图像的匹配质量越好。

$$R_{\text{MSE}} = \sqrt{\frac{1}{W \times H} \sum_{i=1}^W \sum_{j=1}^H [F(i, j) - M_d(i, j)]^2}, \quad (8)$$

式中: $F(i, j)$ 和 $M_d(i, j)$ 分别表示配准图像和固定图像; $W$ 和 $H$ 分别表示图像的宽度和高度。归一化互相关指标的取值范围为 $[-1, 1]$ ,数值越大表示像素灰度值的线性相关性越强,配准图像与固定图像的匹配程度越高,如式(5)所示。

TRE反映同一点在两幅图像中的差异,数值越小表明待评估的配准图像的配准精度越高。TRE的计算公式为

$$T_{\text{RE}} = \sqrt{(x - x')^2 + (y - y')^2}, \quad (9)$$

式中: $(x', y')$ 表示配准图像上点的坐标; $(x, y)$ 表示固定图像上点的坐标。

### 3.3 图像配准结果

使用本文方法和对比方法对压敏漆图像进行配准,通过视觉主观评估配准结果的结构一致性,壳体零件和薄板零件的配准结果分别如图6和图7所示。为了更清楚地展示配准结果,在图像上选择5个典型区域,取样位置如图6(a)红色矩形区域所示,放大展示匹配误差,如area 1~area 5列所示。从典型区域的放

大结果可以看出,风场中的零件存在较大的非刚性形变,有风图像误差图的每一个区域都存在明显的错位,且错位的大小与方向各不相同。

从两组实验的配准结果来看,基于特征匹配的方法对于刚性变形实现了一定的校正,如图6(b)的area 3和area 5所示,但是对于area 2和area 4中存在的非刚性变形无法进行很好的配准,这是因为执行刚性配准策略放大了图7(b)中area 1的误差。

MI-Spline展示了基于迭代的配准结果,该方法是以图像互信息为目标函数的最优化算法,主要关注图像像素的灰度信息,对坐标关系的关注较小。图6(c)area 2与area 4的非刚性变形得到一定的配准,但是area 3与area 5中仍存在较大的变形;图7(c)中area 2和area 5区域没有得到较好的改善。可见,该方法对于压敏漆图像的适用性较差,往往无法优化到最佳配准参数,导致难以完成良好的配准甚至可能放大误差。

基于深度学习的配准模型Voxelmorph、CycleMorph、BIRGU-Net和LRN的配准结果相似,对于小变形可以完成对齐任务,但是对于变形较大的区域无法完成配准。图6中area 1、area 2和area 4的配准效果较好,基本完成了对齐任务,但是在area 3和area 5区域存在较大的偏差。图7中area 3得到了一定的配准,但是area 2和area 5变形情况几乎没有得到改善。

在PIR-Net配准结果中,两组实验数据的5个区域都基本完成了对齐,说明所提出的方法在处理压敏漆图像中复杂场景和大程度形变的配准时具有更强的鲁棒性和更高的准确性。

### 3.4 配准精度分析

为了进一步量化配准的准确性,分别使用RMSE、NCC、TRE评价指标对本文方法和前文所述对比方法在两组风洞数据上的配准结果进行比较。RMSE和NCC评价结果如表1和表2所示,TRE评价结果如表3和表4所示。

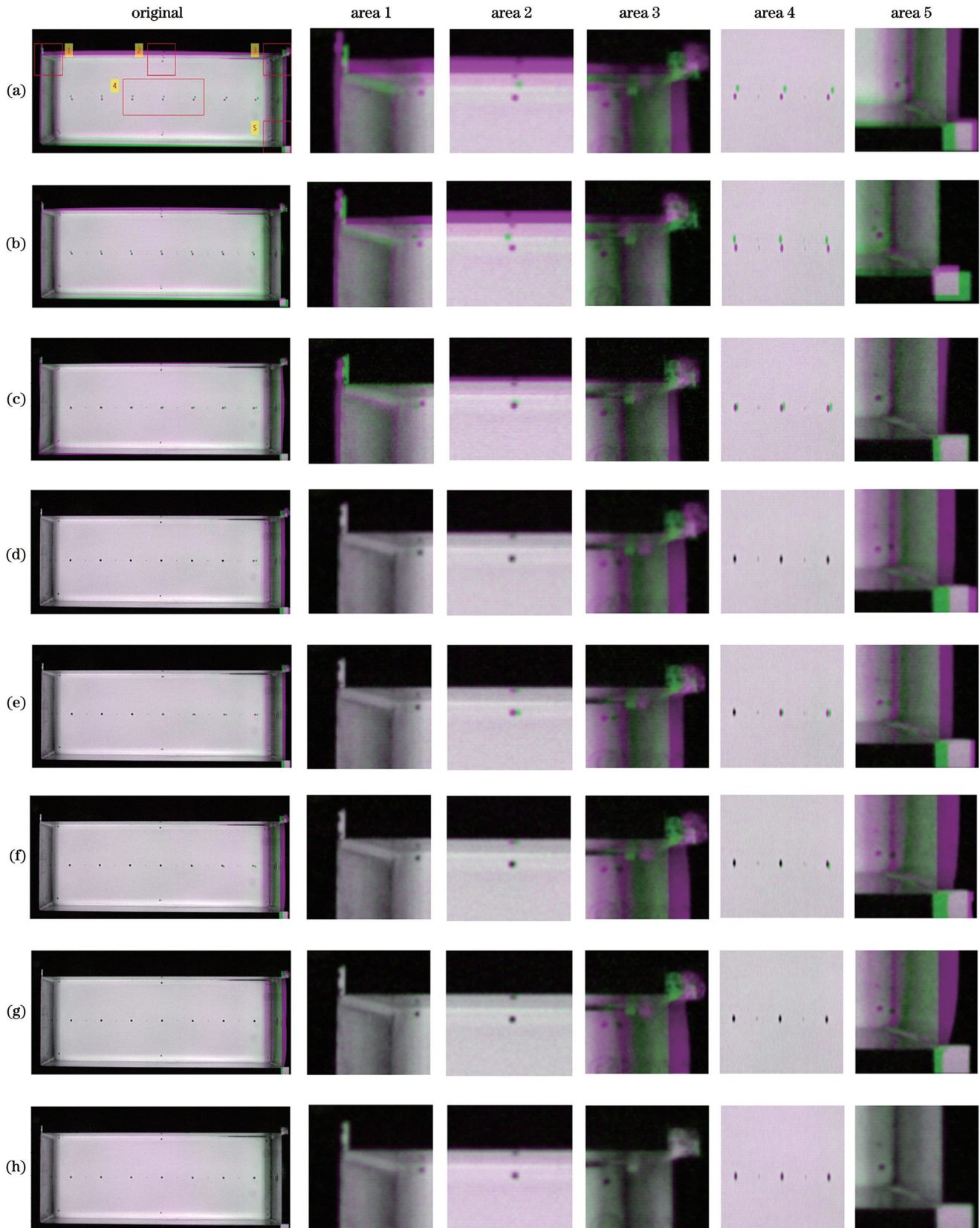


图 6 壳体零件配准误差图。(a)有风图像误差图;(b) Feature Matching 配准误差图;(c) MI-Spline 配准误差图;(d) Voxelmorph 配准误差图;(e) CycleMorph 配准误差图;(f) BIRGU-Net 配准误差图;(g) LRN 配准误差图;(h) PIR-Net 配准误差图

Fig. 6 Registration error images of shell part. (a) Error images of windy image; (b) error images of Feature Matching registration; (c) error images of MI-Spline registration; (d) error images of Voxelmorph registration; (e) error images of CycleMorph registration; (f) error images of BIRGU-Net registration; (g) error images of LRN registration; (h) error images of PIR-Net registration



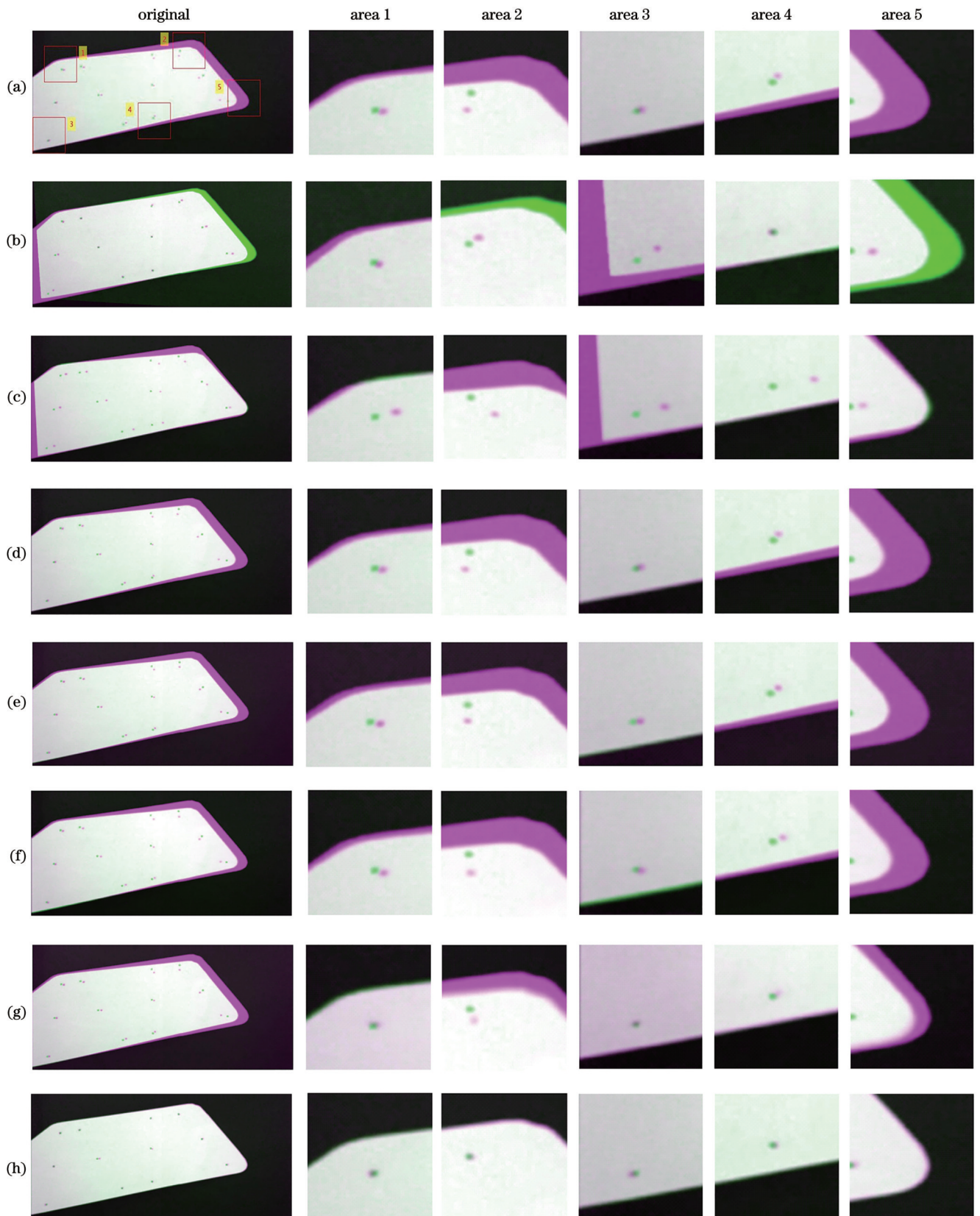


图 7 薄板零件配准误差图。(a)有风图像误差图;(b) Feature Matching 配准误差图;(c) MI-Spline 配准误差图;(d) Voxelmorph 配准误差图;(e) CycleMorph 配准误差图;(f) BIRGU-Net 配准误差图;(g) LRN 配准误差图;(h) PIR-Net 配准误差图

Fig. 7 Registration error images of sheet part. (a) Error images of windy image; (b) error images of Feature Matching registration; (c) error images of MI-Spline registration; (d) error images of Voxelmorph registration; (e) error images of CycleMorph registration; (f) error images of BIRGU-Net registration; (g) error images of LRN registration; (h) error images of PIR-Net registration



表 1 壳体零件配准精度对比  
Table 1 Comparison of registration accuracy of shell part

Method	Original	Feature Matching	MI-Spline	Voxelmorph	CycleMorph	BIRGU-Net	LRN	PIR-Net
RMSE	0.1338	0.1544	0.1043	0.0721	0.0646	0.0715	0.0624	<b>0.0505</b>
NCC	0.1722	0.1451	0.1196	0.3348	0.3483	0.2443	0.3512	<b>0.4087</b>

表 2 薄板零件配准精度对比  
Table 2 Comparison of registration accuracy of sheet part

Method	Original	Feature Matching	MI-Spline	Voxelmorph	CycleMorph	BIRGU-Net	LRN	PIR-Net
RMSE	0.1784	0.1574	0.1423	0.1696	0.1635	0.1681	0.1182	<b>0.0479</b>
NCC	0.0243	0.0237	0.0398	0.0269	0.0261	0.0255	0.0392	<b>0.0596</b>

表 3 壳体零件 TRE 对比  
Table 3 TRE comparison of shell part

Method	TRE /pixel					
	Point 1	Point 2	Point 3	Point 4	Point 5	Point 6
Original	7.07	5.09	4.12	9.21	8.06	8.54
Feature Matching	4.47	4.24	2.82	5.83	5.65	5.09
MI-Spline	3.16	2.82	4.24	3.60	6.08	3.16
Voxelmorph	1.41	1.21	1.63	8.06	6.32	7.07
CycleMorph	1.20	1.07	1.41	5.09	4.56	5.04
BIRGU-Net	1.48	2.11	1.66	9.05	7.07	8.06
LRN	1.21	2.23	2.13	8.24	3.04	7.28
PIR-Net	<b>1.13</b>	<b>0.71</b>	<b>1.10</b>	<b>1.33</b>	<b>1.41</b>	<b>1.27</b>

表 4 薄板零件 TRE 对比  
Table 4 TRE comparison of sheet part

Method	TRE /pixel					
	Point 1	Point 2	Point 3	Point 4	Point 5	Point 6
Original	6.08	4.12	3.60	11.18	10.29	19.10
Feature Matching	3.16	8.06	15.26	8.61	11.04	15.03
MI-Spline	13.34	15.52	19.13	18.86	14.31	9.05
Voxelmorph	7.14	5.38	4.12	13.15	12.21	18.02
CycleMorph	6.08	5.04	4.56	10.04	8.60	16.12
BIRGU-Net	5.39	4.12	4.47	10.19	9.21	14.03
LRN	4.31	3.61	2.82	8.54	4.24	6.71
PIR-Net	<b>2.24</b>	<b>1.70</b>	<b>1.41</b>	<b>2.23</b>	<b>2.11</b>	<b>2.62</b>

从表 1 和表 2 可以看出,所提出的 PIR-Net 在 RMSE 和 NCC 两个评价指标上显著优于对比方法。相对于传统方法,PIR-Net 的 RMSE 指标提升了 51.6%,NCC 指标提升了 181.7%,主要原因是压敏漆图像存在非刚性变形且特征稀疏,基于特征匹配和迭代最优化的方法都不能很好地实现整体配准。相较于其他最新的深度学习配准方法,PIR-Net 引入了多尺度网络结构和损失注意力机制,因此在大形变区域有着更好的适应性,实现了 NCC 指标 16.4% 的提升和

RMSE 指标 19.1% 的提升。

压敏漆图像配准的目的是提高两幅图像比值的准确性,因此不仅要求图像结构相似度高,还注重对应点的精准对齐,具体可以通过 TRE 指标进行评价。本文从图像中选取 6 个典型位置来评估模型针对关键点的匹配效果,点的位置分布如图 8 和图 9 所示。计算图像中关键点的 TRE 以评估匹配的稳定性,结果如表 3、表 4 所示。

从表 3 和表 4 可以看出,PIR-Net 在 6 个关键点上

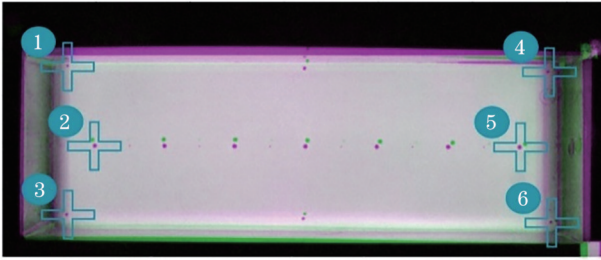


图 8 壳体零件关键点分布

Fig. 8 Distribution of key points of shell part

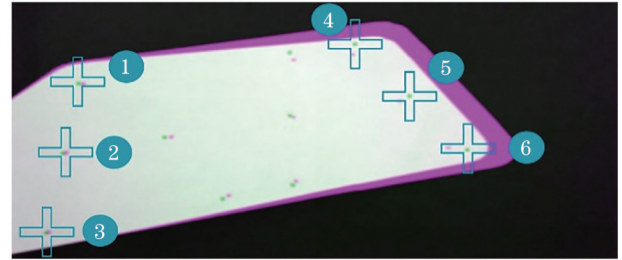


图 9 薄板零件关键点分布

Fig. 9 Distribution of key points of sheet part

的平均配准误差都是最小的,相对于其他深度学习方  
法减少了 62.8%,相对于传统方法减少了 69.8%,实验  
结果证明了本文方法在解决压敏漆图像配准问题上优  
异性。

为进一步说明算法在配准误差控制方面的优势,

将实验中的最大误差位置和平均误差进行对比展示,  
如图 10 和图 11 所示。得益于平滑项约束和注意力机  
制的综合作用,所提方法的最大误差和平均误差数值  
更接近,误差的分布相对平滑,且在误差限约束上更具  
优势。

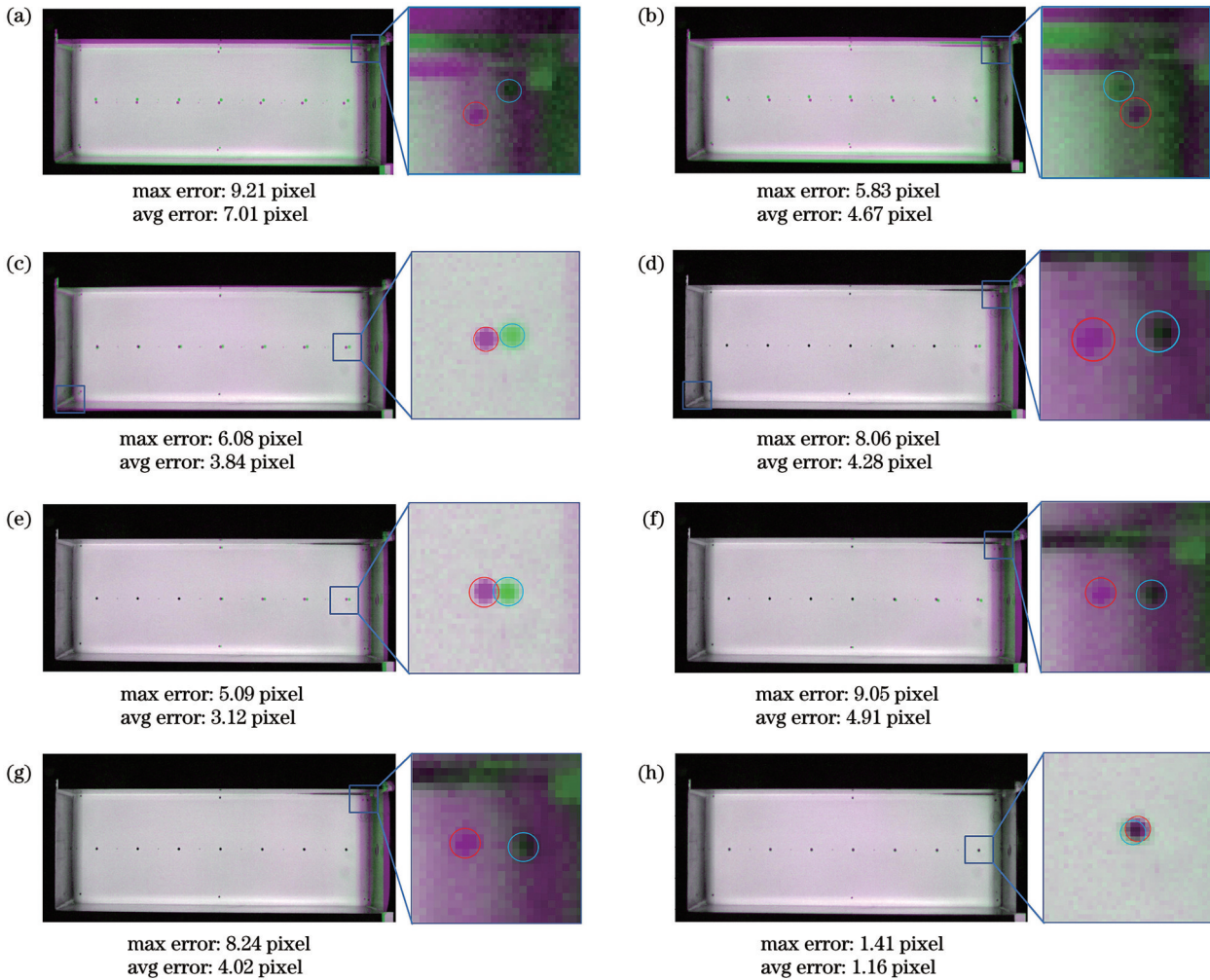


图 10 壳体零件配准结果的最大误差位置和平均误差值。(a)有风图像误差图;(b) Feature Matching 配准误差图;(c) MI-Spline 配准误差图;(d) Voxelmorph 配准误差图;(e) CycleMorph 配准误差图;(f) BIRGU-Net 配准误差图;(g) LRN 配准误差图;(h) PIR-Net 配准误差图

Fig. 10 Maximum error positions and average error values of shell part registration results. (a) Error images of windy image; (b) error images of Feature Matching registration; (c) error images of MI-Spline registration; (d) error images of Voxelmorph registration; (e) error images of CycleMorph registration; (f) error images of BIRGU-Net registration; (g) error images of LRN registration; (h) error images of PIR-Net registration

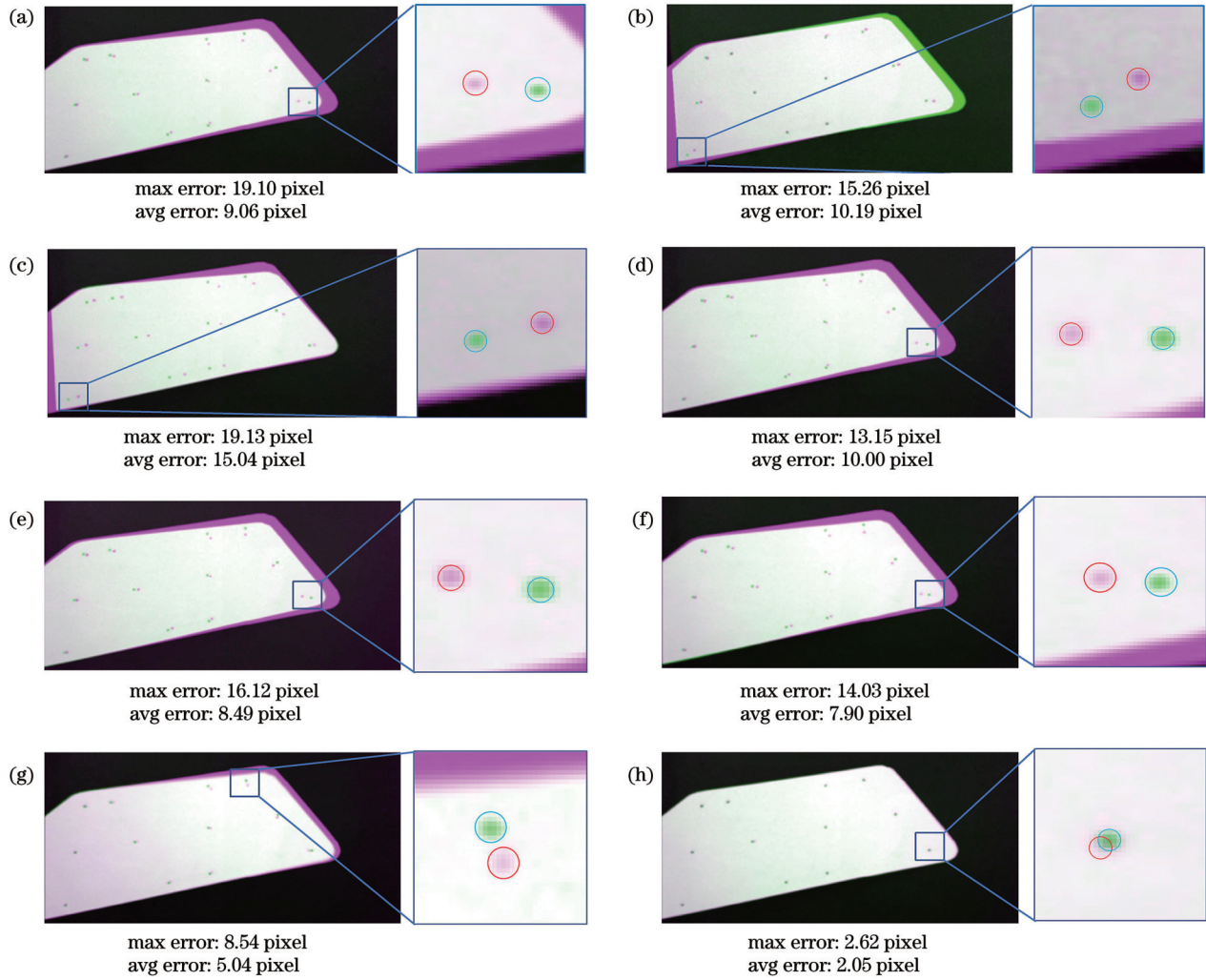


图 11 壳体零件配准结果的最大误差位置和平均误差值。(a)有风图像误差图;(b) Feature Matching 配准误差图;(c) MI-Spline 配准误差图;(d) Voxelmorph 配准误差图;(e) CycleMorph 配准误差图;(f) BIRGU-Net 配准误差图;(g) LRN 配准误差图;(h) PIR-Net 配准误差图

Fig. 11 Maximum error positions and average error values of sheet part registration results. (a) Error images of windy image; (b) error images of Feature Matching registration; (c) error images of MI-Spline registration; (d) error images of Voxelmorph registration; (e) error images of CycleMorph registration; (f) error images of BIRGU-Net registration; (g) error images of LRN registration; (h) error images of PIR-Net registration

### 3.5 配准时效分析

本文方法与对比方法的平均配准时间如表 5 所示。使用 GPU 条件下,本文方法与其他基于深度学习的方法同时具有最优性能表现,皆可在 1 s 内完成两个配准的计算;基于特征匹配的方法需要数秒才能完成配准;MI-Spline 的迭代策略则需要约 10 min 或更多时

间,运行时间差异较大;基于深度学习的配准方法所需要的时间远小于传统经典方法。对比其他深度学习的方法,PIR-Net 每组配准时长增加不到 0.1 s,并不影响处理的时效性,但是微小的配准时间差异换来了显著的配准精度和配准稳定性的提升,这在实际应用中具有重要的意义。

表 5 不同方法的配准时间比较

Table 5 Comparison of registration time for different methods

Method	Feature Matching	MI-Spline	Voxelmorph	CycleMorph	BIRGU-Net	LRN	PIR-Net
Registration time for shell part /s	1.76	754	0.41	0.39	0.42	0.45	0.48
Registration time for sheet part /s	1.75	761	0.43	0.42	0.43	0.47	0.51

### 3.6 消融实验

多尺度配准框架中串联卷积模型的数量是配准网

络抑制大程度、宽动态范围形变的关键参数。为了深入评估基础卷积模型的数量对整体配准精度的影响,



优化其在实际应用中的性能,利用 PIR-Net 设计了级联不同数量卷积模型的消融实验。

实验通过逐步增加卷积模型的数量构建了一系列不同配置的多级联网络,利用 RMSE 和 NCC 两种配准精度评价指标和配准时间指标对这些网络结构的配准性能进行评估,结果如表 6 所示。对于配准精度指标,当模型的数量从 1 个增加到 3 个时,配准精度达到最大值,随着模型数量继续增加,配准精度开始

降低。究其原因,当网络层数过多时,第一个卷积模型输入图像的尺寸变得很小,以致于难以提取图像的深层语义信息,从而导致高层网络匹配性能滑坡。对于配准时间指标,随着级联网络数量的增加,网络变得更加复杂,配准时间增大。结合压敏漆实验图像数据特性,综合考虑配准时间和配准精度两个指标,得到 PIR-Net 采用三层卷积模型级联作为配准框架的最佳策略。

表 6 级联卷积模型的数量对配准精度和时间的影响

Table 6 Impact of the number of cascaded convolutional models on registration accuracy and time

Number of cascading CNN models	Shell part			Sheet part		
	RMSE	NCC	Registration time /s	RMSE	NCC	Registration time /s
1 scale (1)	0.0741	0.3370	0.41	0.1786	0.0237	0.43
2 scale (1/2, 1)	0.0658	0.3476	0.45	0.1168	0.0386	0.48
3 scale (1/4, 1/2, 1)	0.0505	0.4087	0.48	0.0479	0.0596	0.52
4 scale (1/8, 1/4, 1/2, 1)	0.0496	0.4020	0.51	0.0483	0.0588	0.54

## 4 结 论

提出一种基于无监督学习的风洞压敏漆图像配准方法。首先,针对风洞环境中零件大程度、宽动态范围的偏移和非刚性形变问题,通过实验分析和验证构造了适用于该场景的三级联配准网络框架,实现了由粗到细的配准策略。其次,构造了类似 U-Net 结构的卷积网络来融合图像的宏观结构特征和局部细节特征,提高了对非刚性形变的适应性。最后,设计了基于图像相似度的损失函数范式并引入梯度场约束,提高了特征稀疏图像的配准精度。在实测压敏漆图像数据上将本文方法与 MI-Spline、Voxelmorph、CycleMorph、BIRGU-Net 和 LRN 等配准方法进行比较,结果验证了本文方法在配准时间、准确性、稳定性上的优异性,并在量化指标 RMSE、NCC、TRE 上实现了远超现有方法的配准表现。本研究将为解决压敏漆图像配准问题提供了更准确和可靠的方案。

### 参 考 文 献

- [1] Egami Y, Hasegawa Y, Matsuda Y, et al. Ruthenium-based fast-responding pressure-sensitive paint for measuring small pressure fluctuation in low-speed flow field[J]. *Measurement Science and Technology*, 2021, 32(2): 024003.
- [2] Running C L, Sakaue H, Juliano T J. Hypersonic boundary-layer separation detection with pressure-sensitive paint for a cone at high angle of attack[J]. *Experiments in Fluids*, 2019, 60(1): 23.
- [3] Disotell K J, Nikoueeyan P, Naughton J W, et al. Global surface pressure measurements of static and dynamic stall on a wind turbine airfoil at low Reynolds number[J]. *Experiments in Fluids*, 2016, 57(5): 82.
- [4] Sugioka Y, Koike S, Nakakita K, et al. Experimental analysis of transonic buffet on a 3D swept wing using fast-response pressure-sensitive paint[J]. *Experiments in Fluids*, 2018, 59(6): 108.
- [5] Friedl F, Krah N, Jähne B. Optical sensing of oxygen using a modified Stern-Volmer equation for high laser irradiance[J]. *Sensors and Actuators B*, 2015, 206: 336-342.
- [6] Xiang Y M, Wang F, You H J. OS-SIFT: a robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2018, 56(6): 3078-3090.
- [7] Studholme C, Drapaca C, Iordanova B, et al. Deformation-based mapping of volume change from serial brain MRI in the presence of local tissue contrast change[J]. *IEEE Transactions on Medical Imaging*, 2006, 25(5): 626-639.
- [8] 储光涵, 范大昭, 董杨, 等. 结合图论的异源影像点云配准方法[J]. *光学学报*, 2023, 43(12): 1228006.
- [9] 贾仁庆, 殷高方, 赵南京, 等. 浮游藻类细胞显微明场图像与荧光同步测量图像配准方法研究[J]. *中国激光*, 2022, 49(24): 2407202.
- [10] 沈春梅, 刘凡, 朱佳乐. 基于结构光和 CT 的背部点云配准算法[J]. *激光与光电子学进展*, 2023, 60(22): 2210007.
- [11] Shen C M, Liu F, Zhu J L. Back point cloud registration algorithm based on structured light and CT[J]. *Laser & Optoelectronics Progress*, 2023, 60(22): 2210007.
- [12] Dong Y Y, Jiao W L, Long T F, et al. An extension of phase correlation-based image registration to estimate similarity transform using multiple polar Fourier transform[J]. *Remote Sensing*, 2018, 10(11): 1719.
- [13] Liu T S, Sullivan J P, Asai K, et al. Pressure and temperature sensitive paints[M]. Cham: Springer Nature, 2021.
- [14] Fujimatsu N, Tamura Y, Fujii K. Improvement of noise filtering and image registration methods for the pressure sensitive paint experiments[J]. *Journal of Visualization*, 2005, 8(3): 225-233.
- [15] Suzuki K, Inoue T, Nagata T, et al. Markerless image alignment method for pressure-sensitive paint image[J]. *Sensors*, 2022, 22(2): 453.
- [16] Cao C H, Cao L, Li G, et al. BIRGU Net: deformable brain

- magnetic resonance image registration using gyral-net map and 3D Res-Unet[J]. *Medical & Biological Engineering & Computing*, 2023, 61(2): 579-592.
- [16] 李文举, 孔德卿, 曹国刚, 等. 基于训练-推理解耦架构的 2D-3D 医学图像配准[J]. *激光与光电子学进展*, 2022, 59(16): 1610015. Li W J, Kong D Q, Cao G G, et al. 2D-3D medical image registration based on training-inference decoupling architecture [J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1610015.
- [17] 林立昊, 易见兵, 曹锋, 等. 多尺度并行全卷积神经网络的肺计算机断层扫描图像非刚性配准算法[J]. *激光与光电子学进展*, 2022, 59(16): 1617004. Lin L H, Yi J B, Cao F, et al. Non-rigid registration algorithm of lung computed tomography image based on multi-scale parallel fully convolutional neural network[J]. *Laser & Optoelectronics Progress*, 2022, 59(16): 1617004.
- [18] 刘磊, 李元祥, 倪润生, 等. 基于卷积与图神经网络的合成孔径雷达与可见光图像配准[J]. *光学学报*, 2022, 42(24): 2410002. Liu L, Li Y X, Ni R S, et al. Synthetic aperture radar and optical images registration based on convolutional and graph neural networks[J]. *Acta Optica Sinica*, 2022, 42(24): 2410002.
- [19] Haskins G, Kruger U, Yan P K. Deep learning in medical image registration: a survey[J]. *Machine Vision and Applications*, 2020, 31(1): 8.
- [20] Yang X, Kwitt R, Styner M, et al. Quicksilver: fast predictive image registration: a deep learning approach[J]. *NeuroImage*, 2017, 158: 378-396.
- [21] Rohé M M, Datar M, Heimann T, et al. SVF-net: learning deformable image registration using shape matching[M] // Descoteaux M, Maier-Hein L, Franz A, et al. *Medical image computing and computer assisted intervention-MICCAI 2017*. Lecture notes in computer science. Cham: Springer Nature, 2017, 10433: 266-274.
- [22] Li H M, Fan Y. Non-rigid image registration using fully convolutional networks with deep self-supervision[EB/OL]. (2017-09-04)[2023-11-12]. <https://arxiv.org/abs/1709.00799>.
- [23] Kim B, Kim D H, Park S H, et al. CycleMorph: cycle consistent unsupervised deformable image registration[J]. *Medical Image Analysis*, 2021, 71: 102036.
- [24] de Vos B D, Berendsen F F, Viergever M A, et al. End-to-end unsupervised deformable image registration with a convolutional neural network[M] // Cardoso M J, Arbel T, Carneiro G, et al. *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Lecture notes in computer science. Cham: Springer Nature, 2017, 10553: 204-212.
- [25] Balakrishnan G, Zhao A, Sabuncu M R, et al. An unsupervised learning model for deformable medical image registration[C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18–23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 9252-9260.
- [26] Balakrishnan G, Zhao A, Sabuncu M R, et al. VoxelMorph: a learning framework for deformable medical image registration[J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1788-1800.
- [27] Han R, Jones C K, Lee J, et al. Deformable MR-CT image registration using an unsupervised, dual-channel network for neurosurgical guidance[J]. *Medical Image Analysis*, 2022, 75: 102292.
- [28] Li S H, Sui X C, Luo X D, et al. Medical image segmentation using squeeze-and-expansion transformers[EB/OL]. (2021-05-20)[2023-11-12]. <https://arxiv.org/abs/2105.09511>.
- [29] Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation[M] // Navab N, Hornegger J, Wells W M, et al. *Medical image computing and computer-assisted intervention-MICCAI 2015*. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [30] Ho T T, Kim W J, Lee C H, et al. An unsupervised image registration method employing chest computed tomography images and deep neural networks[J]. *Computers in Biology and Medicine*, 2023, 154: 106612.

## Unsupervised Learning Based Image Registration of Wind Tunnel Pressure Sensitive Paint Image

Liu Kang<sup>1,2,3</sup>, Sun Xiongwei<sup>1,2,3\*</sup>, Shi Hailiang<sup>1,2,3\*\*</sup>, Wang Xianhua<sup>1,2,3</sup>, Ye Hanhan<sup>2,3</sup>,  
Cheng Chen<sup>1,2,3</sup>, Zhu Feng<sup>2,3</sup>, Wu Shichao<sup>2,3</sup>

<sup>1</sup>University of Science and Technology of China, Hefei 230026, Anhui, China;

<sup>2</sup>Anhui Institute of Optics and Fine Mechanics, Hefei Institutes of Physical Science, Chinese Academy of Sciences, Hefei 230031, Anhui, China;

<sup>3</sup>Key Laboratory of General Optical Calibration and Characterization Technology, Chinese Academy of Sciences, Hefei 230031, Anhui, China

### Abstract

**Objective** Pressure sensitive paint (PSP) technology is a non-contact optical pressure measurement method utilized extensively for surface pressure measurement of parts in wind tunnel environments. The surface of a part coated with PSP fluoresces under excited light conditions, and the pressure results can be inverted using the Stern-Volmer formula. This formula requires the ratio of the windy image to the windless image, but the displacement and non-rigid deformation of the part in the wind tunnel environment will result in computational errors in the division of non-corresponding points. Consequently, the accurate registration of windy and windless images is fundamental to processing PSP experimental data. Typical PSP images comprise only two distinct components: a bright light-emitting region and a black background region,

leading to sparse image features and a relatively limited number of feature points, which makes it difficult to apply typical registration methods directly. Moreover, as the number of images in a single experiment exceeds tens of thousands, conventional non-rigid registration methods are often slow and insufficient for fast registration requirements. Consequently, there is an urgent demand to develop a new method that can register images accurately and swiftly without relying on marker points.

**Methods** To achieve the demand for accurate and fast registration of PSP images, we propose a registration method based on unsupervised learning. The method does not require *a priori* information and directly learns an end-to-end from image pairs to deformation fields. The registration network structure incorporates multiple scales of structures, through a multi-cascade approach, facilitating a coarse-to-fine registration of PSP images. Furthermore, we have designed a new loss function based on the structural similarity of images, which maximizes the similarity between the registration image and the input fixed image. In our study, two sets of PSP experimental images of typical parts, each comprising 20000 image pairs, are introduced for the training and testing of the registration network. The image sizes are  $640 \times 272$ , and the format is a 16 bit greyscale image. The registration process utilizes only image pair information, without the assistance of external supervisory information. To assess the efficacy of this method, we compare and analyze it with conventional algorithms currently used for PSP image alignment. These conventional algorithms include feature-based matching algorithms, MI-Bspline algorithms that combine greyscale and B-spline, and deep learning-based alignment models such as Voxelmorph, CycleMorph, BIRGU-Net, and LRN. The image registration methods are also evaluated based on registration accuracy and time. Registration accuracy is measured by three common quantitative metrics: root mean square error (RMSE), normalized correlation (NCC), and target alignment error (TRE).

**Results and Discussions** Evaluating and analyzing our method with the conventional methods, it is evident from the registration results of shell parts and thin plate parts in Figs. 6 and 7 that the five regions of the two sets of experimental data in the PIR-Net registration results have essentially completed the registration. This suggests that the method possesses a stronger robustness in handling complex scenes and large deformation alignments in PSP images and enhancing accuracy. To further quantify the accuracy of the registration, we utilize the RMSE and NCC indices to evaluate the results (Tables 1 and 2). The tables indicate that the PIR-Net significantly outperforms the comparison methods in both the RMSE and NCC metrics. Compared to the conventional methods, the RMSE index is improved by 51.6% and the NCC index is improved by 181.7%. This improvement is primarily attributed to the non-rigid deformation and feature sparsity in the PSP image. Neither the feature matching nor the iterative optimization-based methods can effectively address these issues, leading to sub-optimal overall registration. Compared to other deep learning-based registration methods, PIR-Net demonstrates superior adaptability in large deformation regions due to its multi-scale network structure and attention mechanism. This results in a 16.4% improvement in the NCC and a 19.1% improvement in the RMSE. To further illustrate the advantages of the algorithm in registration error control, we compare and demonstrate the maximum error position and the average error in the experiments. Due to the combination of the smoothing term constraint and the attention mechanism, it exhibits a more consistent distribution of error, with a relatively smooth error limit constraint (Figs. 10 and 11). The average time for each method's registration is experimentally demonstrated in Table 5. Our method outperforms other conventional methods. Compared to other deep learning methods, the registration time of PIR-Net is slightly longer. This is primarily due to the use of multi-scale registration. However, using a very small difference in registration time for a higher accuracy of registration is a good compromise between time performance and accuracy of registration, which is more practical.

**Conclusions** We introduce an unsupervised learning-based registration method for the PSP image registration. This method directly learns the end-to-end mapping from image pairs to deformation fields. It designs a multiscale network structure and a coarse-to-fine registration strategy to address the issue of large offsets and non-rigid deformations in wind tunnel environments. Additionally, it incorporates a novel loss function paradigm based on image similarity, which enhances image registration in feature-sparse scenarios. Comparing with typical alignment methods such as MI-Spline and Voxelmorph on two sets of PSP images, the experimental results prove that our method achieves a far better registration performance than the existing methods in visual evaluation and RMSE, NCC, and TRE metrics, under the premise of ensuring the performance of the algorithm. This method provides a reliable solution to the PSP image registration problem.

**Key words** machine vision; pressure sensitive paint image; non-rigid deformation; image registration; unsupervised learning