

DOI:10.3876/j.issn.1000-1980.2023.01.016

基于残差网络的河流表面时空图像测速法

李华宝¹,张振¹,陈林¹,孟健²,孙英军³,崔文浩¹

(1.河海大学计算机与信息学院,江苏南京 211100; 2.杭州市水文水资源监测中心,浙江杭州 310016;
3.浙江水文新技术开发经营公司,浙江杭州 310016)

摘要:针对在耀光、紊流、降雨等复杂含噪场景下时空图像中有效纹理特征变得模糊,使得现有纹理主方向检测算法精度受限的问题,结合深度学习的思想,提出了一种基于残差网络回归模型的时空图像测速(ResNet50-STIV)法,并借助残差网络回归模型强大的非线性学习能力构建了回归预测函数。通过构建人工合成数据集和包含复杂场景时空图像的天然河流数据集对残差网络回归模型进行试验,结果表明:提出的残差网络回归模型在人工合成数据集下的检测精度可达到 0.1° ,对于天然河流数据集,具有残差结构的ResNet回归模型的检测精度优于VGG16;从模型层数看,基于ResNet50的回归模型能较好地平衡检测精度以及执行效率,在正常场景下的检测精度达到 0.7° ,而在耀光、紊流、降雨场景下能控制在 1.3° 以内,ResNet50-STIV优于现有的时空图像测速法;与流速仪法在多场景下表面流速比测的最大相对误差小于12%。

关键词:水流测量;时空图像测速;纹理主方向;残差网络;深度学习

中图分类号:P332.4 文献标志码:A 文章编号:1000-1980(2023)01-0118-11

Surface space-time image velocimetry of river based on residual network

LI Huabao¹, ZHANG Zhen¹, CHEN Lin¹, MENG Jian², SUN Yingjun³, CUI Wenhao¹

(1. College of Computer and Information Engineering, Hohai University, Nanjing 211100, China;
2. Zhejiang Hydrology New Technology Development and Management Company, Hangzhou 310016, China;
3. Hangzhou Hydrology and Water Resources Monitoring Center, Hangzhou 310016, China)

Abstract: Aiming at the problem that the effective texture features in the space-time image become blurred in complex noisy scenes such as flare, turbulence and rainfall, which limits the accuracy of the existing main direction detection algorithms of texture, a space-time image velocimetry method based on residual network regression model is proposed by combining the idea of deep learning, and the powerful nonlinear learning ability of the model is used to construct a regression prediction function. The residual network regression model is tested by constructing synthetic datasets and natural river datasets containing space-time images of complex scenes. The experimental results show that the detection accuracy of the proposed model under the synthetic datasets can reach 0.1° . For the natural river datasets, the detection accuracy of the ResNet regression model with residual structure is better than that of VGG16; in terms of the number of model layers, ResNet50-STIV can better balance the detection accuracy and execution efficiency. The detection accuracy in normal scenarios can reach 0.7° , while in flare, turbulence, and rainfall scenarios, it can be controlled within 1.3° , and ResNet50-STIV is better than the existing STIV method. Meanwhile, the maximum relative error is less than that of the instrument method in the surface velocity ratio measurement of multiple scenarios by 12%.

Key words: flow measurement; space-time image velocimetry; main orientation of texture; residual network; deep learning

我国中小河流众多,洪水频发,及时且准确地获取流速、流量等水文信息对灾害防治至关重要^[1-2]。然而在洪水暴发期河流流速快、漂浮物多,容易造成仪器损毁并威胁到测量人员的人身安全,使得流速仪法等传统接触式测流方法难以正常施测^[3]。近年来基于图像的测流方法发展迅速^[4-5],因测量效率高及安全性强的特点得到广泛应用^[6-8]。该方法通过架设于河岸一侧的摄像机拍摄河流水面得到视频图像,通过估计图像

基金项目:浙江省水利厅科技计划(RB2037);杭州市农业与社会发展一般项目(20201203B103);中央高校基本科研业务费专项(B200202187);江苏省水利科技项目(2021070);中国博士后科学基金面上项目(2019M651673)

作者简介:李华宝(1996—),男,硕士研究生,主要从事时空图像测速法研究。E-mail:191307020017@hhu.edu.cn

通信作者:张振(1985—),男,副教授,博士,主要从事光电成像与图像法测流技术等研究。E-mail:zz_hhuc@163.com

引用本文:李华宝,张振,陈林,等.基于残差网络的河流表面时空图像测速法[J].河海大学学报(自然科学版),2023,51(1):118-128.

LI Huabao, ZHANG Zhen, CHEN Lin, et al. Surface space-time image velocimetry of river based on residual network[J]. Journal of Hohai University(Natural Sciences), 2023, 51(1): 118-128.

序列中水流示踪物的运动矢量大小并转换到世界坐标系下得到流速数据。基于图像的测流方法按运动矢量估计方法的不同主要分为大尺度粒子图像测速法 (large-scale image velocimetry, LSPIV)^[9] 和时空图像测速法 (space-time image velocimetry, STIV)^[10] 两类。

LSPIV 是一种二维瞬时运动矢量估计方法,通常以自然光为光源,以植物碎片、泡沫、细小波纹等天然漂浮物作为水流示踪物,通过对图像进行正射校正来消除相机在倾斜视角下拍摄造成的图像透视畸变,基于互相关匹配算法估计连续两幅图像中矩形分析区域内水流示踪物的平均运动矢量。LSPIV 已应用于多种场景下河流表面流速监测和流量估算,但也存在一定的局限性:①互相关算法所需的计算时间长,难以满足实时流场测量的需求;②参数的设置过程较为复杂,并且学习成本高;③通过固定式或手持式测量的方式采集到的图像通常为斜视图像,图像中不同区域下的空间分辨率大小不同,由于最终的流场计算需要采用较大的采样窗口导致流场空间分辨率偏低。

与 LSPIV 相比,STIV 用单像素宽的测速线代替矩形分析窗口作为分析区域,因此 STIV 具有更高的空间分辨率和计算效率,更适合小倾角拍摄的岸基式系统。STIV 的关键在于纹理主方向 (main orientation of texture, MOT) 的精确估计。现有的纹理主方向检测算法按图像的处理域不同可分为两大类:一是梯度张量法 (gradient tensor method, GTM)^[10]、二维自相关函数法 (two-dimensional autocorrelation function, QESTA)^[11] 等空域方法,二是基于快速傅里叶变换 (fast Fourier transform, FFT) 的频域方法 (FFT-STIV)^[12]。GTM 将时空图像划分为若干个局部分析窗口,计算每个窗口的方向,然后基于互相关算法得到每个窗口的权重,进而计算得到纹理主方向。但在时空图像中存在较为明显的干扰纹理时,GTm 的计算结果易与真实值产生较大的偏差。QESTA 首先通过标准化滤波器对时空图像进行滤波,然后计算时空图像中图像强度的二维自相关函数分布,最后根据相关性较高区域的方向角与原时空图像中有效纹理主方向存在互余的关系,计算该方向角以间接获得纹理主方向。QESTA 能够很好地消除时空图像中垂直方向的干扰条纹,但检测包含非垂直方向干扰条纹的时空图像时精度较低。基于 FFT 法的 STIV (FFT-STIV) 基于傅里叶变换的自配准性质,将在空间中检测示踪物纹理主方向的问题转换为频域中搜索频谱主方向来解决。FFT-STIV 对于短随机噪声具有较好的鲁棒性,但在流速脉动的紊流条件下检测结果易出现较大误差。为了有效滤除噪声的干扰,提升纹理主方向检测算法的精度,Zhao 等^[13-14] 在时空图像的幅度谱中采用频域扇形滤波器^[15],仅保留扇形区域与河流水面示踪物运动相关的频谱,滤除其他区域无用的背景噪声,以提升时空图像的纹理清晰度。扇形滤波器的参数主要包括方向角、通带夹角和半径,其中滤波器的方向角通过在幅度谱中采用单像素宽的直线进行径向积分求最大值的方式确定,并设置固定大小的通带夹角及半径构建扇形滤波器。滤波后,GTm、QESTA 和 FFT 法 3 种纹理主方向检测方法的精度都有提升,其中 FFT 法表现最好。扇形滤波器能否有效地滤除噪声的关键在于参数的设置,但对于滤波器的参数现有算法尚未做到自适应选取,很大程度上依赖于人工设置并且所需设置的参数较多,人工调参很难兼顾各种复杂场景下滤波器参数的设置。

近年来,深度学习发展迅速,在头部姿态估计、年龄估计、目标检测、人脸特征点定位等^[16-19] 典型的图像回归问题中都有着很好的表现,深度神经网络能够从大量训练数据中有效获取复杂非线性的映射关系。基于深度卷积神经网络 (deep convolutional neural network, DCNN) 的回归模型采用 DCNN 实现特征提取和回归,并且在特征表达和泛化性能方面都取得了巨大的成功,可直接学习输入到输出的映射函数,不需要复杂的建模,简单高效。Jourabloo 等^[20] 提出了一种基于卷积神经网络的级联回归模型,级联 6 个卷积神经网络构建回归模型对特征点进行预测,解决了极限姿态下的人脸特征点定位问题。Chen^[21] 提出一种基于 DCNN 方法,实现了人脸外观年龄的估计。可以看出,基于 DCNN 的回归模型在处理图像回归任务时有着出色表现,但目前尚未应用到 STIV 的纹理主方向检测问题中。因此,针对现有 STIV 算法无法解决复杂含噪场景下时空图像纹理主方向的精确估计问题,本文采用 DCNN 中性能出色的残差网络 (ResNet)^[22] 构建回归模型,从数据集的精确性、代表性出发构建了人工合成数据集和攀枝花水文站数据集进行试验,验证了所提模型的可行性以及有效性,并通过流速比测试验证了所提方法应用于实际测量的有效性。

1 图像测速法

1.1 传统的 STIV

STIV 主要包括合成时空图像、检测时空图像的纹理主方向、计算世界坐标系下的流速矢量 3 个步骤。

如图1所示,时空图像合成的步骤包括:首先采集 m 帧图像序列,然后在图像中沿水流的运动方向设置一组单像素宽、 L 像素长的测速线,最后对每一条测速线以 xOt 为直角坐标系合成 $L \times m$ 的时空图像。从图1(b)可以看出,时空图像具有显著的纹理特征,将纹理的主要走向和纵坐标轴的夹角 δ 定义为纹理主方向。 δ 的值由河流流向及流速决定,常见的角度区间包括 $5^\circ \sim 88^\circ$ 及 $92^\circ \sim 175^\circ$ 。设物平面中示踪物在时间 T 内沿着某条测速线方向运动了距离 D ,在图像坐标系下表现为在 τ 帧内运动了 d 像素,则对应的流速 v_s 可以表示为

$$v_s = \frac{D}{T} = \frac{d\Delta s}{\tau\Delta t} = \tan\delta \frac{\Delta s}{\Delta t} = v_o \Delta s \quad (1)$$

式中: v_o 为光流运动速度; Δs 为测速线上的物像尺度因子; Δt 为相邻帧图像时间间隔。 v_s 、 v_o 之间仅存在一个由 Δs 表示的缩放关系,其值通过流场定标解算得到。

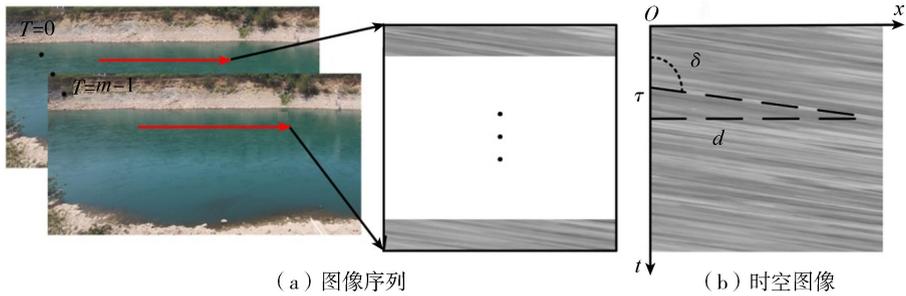


图1 时空图像合成

Fig. 1 Synthesis of space-time image

1.2 基于残差网络的 STIV

STIV 的关键在于纹理主方向检测,然而传统的纹理主方向检测算法无法解决复杂含噪场景下纹理主方向精确估计的问题,并且实现的过程较为复杂,需要对时空图像进行预处理。针对该问题,本文将检测纹理主方向的过程视为图像回归预测问题:

$$\delta = F(M) \quad (2)$$

式中: M 为待检测的时空图像; $F(M)$ 为回归预测函数。采用残差网络构建回归模型,凭借其强大的图像特征提取和回归能力构建预测函数,不需要对时空图像进行预处理,直接从时空图像中回归出纹理主方向以解决复杂含噪场景下纹理主方向估计的问题。基于残差网络回归模型的 STIV 流程如图2所示,将基于残差网络的回归模型代替传统检测算法,将时空图像输入到训练好的模型中即可检测出纹理主方向,进而计算出流速矢量。

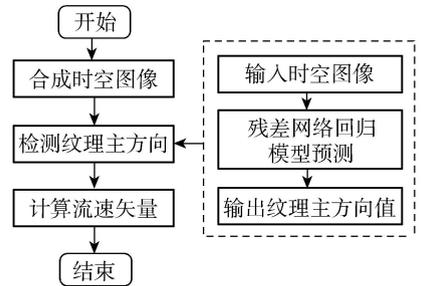


图2 基于残差网络的 STIV 流程

Fig. 2 Flow chart of STIV based on residual network

2 残差网络回归模型

2.1 残差网络原理

卷积神经网络学习和推理过程可以看作是从带有特定标签的数据中学习高维且抽象的特征,并泛化到相同类型但未知的数据。对于视觉任务则需要更深层的网络来提取更高维的特征,理论上网络层数越多就意味着能提取到不同等级的特征越丰富,越能提取到抽象的语义信息。但层数的不断增加,网络会出现梯度弥散或爆炸和网络退化的问题,导致深层的网络训练效果反而比浅层的差。

残差网络思想是假设在一个深层次的网络中,存在一个最优的网络层数,那么该网络中就包含了一部分多余的网络层,定义为冗余层。将这些冗余层设置为恒等层,这样能够完成输入和输出的恒等映射,并且在网络训练过程中能自适应地学习出恒等层。由此,神经网络的层数可以达到几十上百层甚至千层。图3为残差块的具体结构,恒等

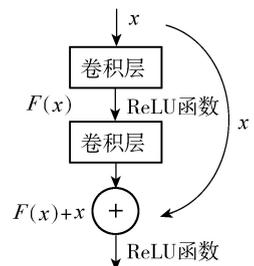


图3 两层恒等残差块示意图

Fig. 3 Schematic map of two-layer identical residual block

残差块中存在一个跳跃连接称作 shortcut 连接, 残差块的输出 $H(x)$ 包括了残差映射 $F(x)$ 以及输入 x 的叠加。假设需要拟合的函数映射是 $H(x)$, 那么可以定义另一个残差映射 $F(x)$, 且 $F(x) = H(x) - x$, 优化残差映射比优化原始函数映射要更加容易。因为假设在恒等映射即 $H(x) = x$ 是最优的极端情况下, 残差网络只需要将残差映射学习为 0 即 $F(x) = 0$, 即不需要使用叠加的非线性网络去拟合恒等映射, 因此残差网络能很好地解决在搭建深层次网络时面临的梯度弥散和网络退化问题。

2.2 模型的搭建

本文在 ResNet50 原始模型上进行改进, 原始模型的全连接层用于分类任务, 而本文需要进行回归任务。图 4(a) 显示 ResNet50 中每个恒等残差块中卷积核的顺序是 1×1 、 3×3 和 1×1 , 第一层 1×1 的卷积核起到减少输入特征通道数的作用, 第三层 1×1 的卷积层则增加输出特征的通道数, 确保通过瓶颈层前后的通道数相同, 并且能够减少中间层 3×3 的卷积核的计算量。本文改进的基于 ResNet50 的回归模型 (以下简称 ResNet50-R) 如图 4(b) 所示, 模型首先对输入时空图像进行卷积计算, 之后通过设计好的 4 个残差块依次叠加 3 次、4 次、6 次和 3 次, 最后经回归层输出纹理主方向角度值。回归层由多个全连接层以及 ReLU 函数组成, 全连接层的作用是将特征提取层学习到的分布式特征映射到样本的标记空间, 即回归出角度值。

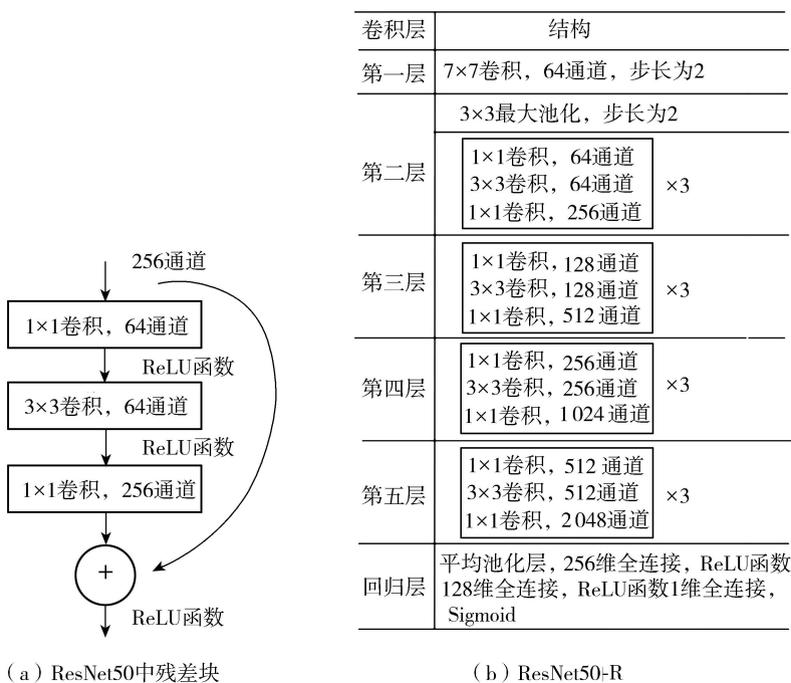


图 4 残差块与模型结构
Fig. 4 Residual block and model structure

3 试验和结果分析

数据集对于模型的训练起着关键的作用, 数据集的质量往往能够直接影响到模型最终的性能, 针对目前 STIV 中缺乏宽量程、高精度纹理主方向的时空图像数据集的问题, 本文基于河流水面背景及示踪物的图像特征, 合成纹理主方向真值已知的纹理图像来模拟无噪声干扰的时空图像, 构建了人工合成数据集。此外, 从数据集的代表性出发, 通过攀枝花水文站搭建的视频测流系统采集了不同光照及水流气象条件下的时空图像, 构建了攀枝花水文站数据集。

3.1 人工合成数据集

河流水面图像中, 水体呈现为一个偏暗且灰度分布均衡的背景。背景灰度强度主要取决于水深、水体成分以及环境的光照强度。为了使得合成的纹理图背景灰度值与真实时空图像接近, 对典型时空图像 (图 5(a)) 进行灰度直方图分析 (图 5(c)), 可以看出灰度值大部分处于 165 附近, 因此首先生成一个灰度值为 165 的背景图, 背景图的像素矩阵记为 $\mathbf{B} = (B_{ij})$ (图 5(b), B_{ij} 为像素矩阵中第 i 行、第 j 列的像素值), 大小为 224×224 像素, 作为叠加各种类型纹理特征的背景。

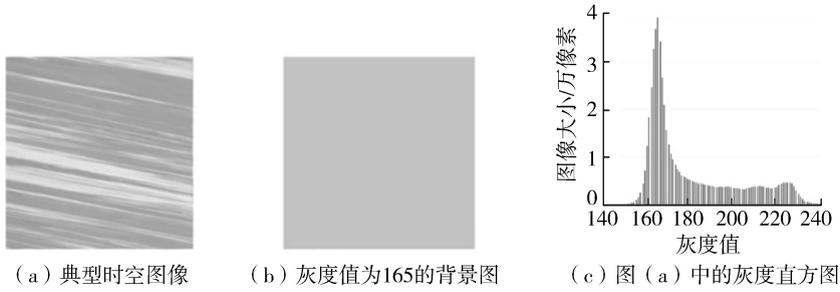


图5 背景图的生成

Fig. 5 Generation of background image

通过叠加二维正弦函数生成指定方向及粗细的纹理:

$$H_{ij} = B_{ij} + I \sin[w(a_i + b_j)] \quad (i \geq 0, j \leq 224) \quad (3)$$

其中 $a_i = \sin\alpha$ $b_j = \cos\alpha$

式中: H_{ij} 为合成纹理图像的像素矩阵第 i 行、第 j 列的像素值; w 、 α 分别为间距及方向参数, 分别决定了纹理的间距及方向; I 为对比度系数, 能调节纹理的对比度。试验中设置 w 、 I 的取值范围均在 $[0, 1]$ 之间, 其中 w 以 0.02 为步长从 0.02 逐渐增大至 0.8 共 40 个值, I 以 0.3 为步长从 0.3 增大至 0.9 共 3 个值。通过不同 w 和 I 的组合可生成 120 种纹理进而得到 120 张不同的纹理图像, 生成的纹理图像示例如图 6 所示。将生成的 120 张图像随机挑选 100 张用于构建训练集, 其余 20 张用于构建测试集。依据常见的时空图像纹理主方向范围 ($5^\circ \sim 88^\circ$ 和 $92^\circ \sim 175^\circ$, 共 168 个整数角度值), 以 1° 为步长 (角度分辨率), 通过改变 α 的取值得到纹理特征相同但纹理主方向不同的图像。每一种类型的纹理图像可生成 168 种不同纹理主方向的图像, 最终人工合成数据集的构成为训练集 16800 张, 测试集 3360 张。

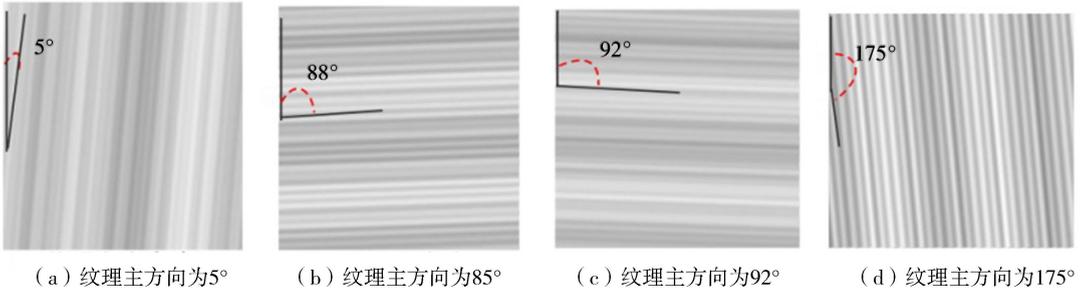


图6 合成纹理图像示例

Fig. 6 Example of composite texture image

3.2 攀枝花水文站数据集

攀枝花水文站测流断面如图 7(a) 所示, 水流折冲形成的表面波和漩滚模式为图像法测流技术提供了良好的天然示踪条件。视频测流系统安装在河流右岸站房一侧的边坡上 (图 7(b)), 位于缆道流速仪断面和水尺断面中间, 对应起点距 2.9 m, 高程 1007.8 m, 俯仰角 19.8° , 以上标定参数用于免像控的河流水面流场定标^[23]。



图7 攀枝花水文站测流断面

Fig. 7 Flow measurement section of Panzhuhua Hydrological Station

通过视频测流系统采集了从 2020 年 7 月到 2021 年 8 月近 1 年的数据。设置测速线的长度为 750 像素,视频的持续时间为 30 s,每秒 25 帧,合成的时空图像的大小为 750×750 像素。为了使攀枝花水文站数据集能够尽可能地覆盖真实河流的各种场景,构建数据集的过程中挑选了正常、紊流、耀光、障碍物和降雨 5 种常见场景下的时空图像,如图 8 所示。挑选了正常场景下 150 张时空图像,其中 100 张用于构建训练集,50 张用于构建测试集;其他 4 种场景下各挑选了 70 张时空图像,其中 50 张用于构建训练集,其余则用于构建测试集。由于正常场景的时空图像在实际测量中出现频次较高,故数据集中该场景下的时空图像占比较大。纹理主方向标记的过程(图 9)如下:首先将时空图像进行二维离散傅里叶变换到频域,在频域中与河流示踪物运动相关的纹理表现为倾斜的有效能量线,然后人为地在幅度谱中设置一条与有效能量线平行的线段,根据线段两个端点在幅度谱上的坐标 (x_1, y_1) 、 (x_2, y_2) 计算出斜率,进而得到频谱主方向 θ :

$$\theta = \arctan[(y_2 - y_1)/(x_2 - x_1)] + 90^\circ \quad (4)$$



图 8 不同场景下的时空图像

Fig. 8 Space-time image in different scenes

最后根据纹理主方向与 θ 存在的正交关系得到纹理主方向。

为了获取宽量程纹理主方向的攀枝花水文站数据集,通过数据增强的方式来扩充数据集,根据纹理主方向范围区间共 168 个整数角度值($5^\circ \sim 88^\circ$ 和 $92^\circ \sim 175^\circ$),采用旋转的方式获取多种纹理主方向的时空图像,旋转时角度的步长为 1° ,每一张时空图像经过旋转可得到 168 种纹理主方向的时空图像,并将旋转后的时空图像的大小裁剪为一致,均为 224×224 像素。最终攀枝花水文站数据集正常、紊流、耀光、障碍物、降雨 5 种场景训练集图像数分别为 16 800 张、8 400 张、8 400 张、8 400 张和 8 400 张,测试集图像数分别为 8 400 张、3 360 张、3 360 张、3 360 张和 3 360 张,训练集和测试集图像总数分别为 50 400 张和 21 840 张。

3.3 试验平台与评价方法

试验平台采用 Intel(R) Xeon(R) Gold 5218 CPU@2.30 GHz,显卡为 Quadro RTX 400;操作系统为 Ubuntu 18.04,编程语言为 Python3.6,深度学习框架为 Pytorch1.5.1。试验中超参数的设置如下:初始的学习率为 0.001,每次迭代计算的样本大小为 64,迭代次数为 200,采用均方误差(mean squared error, MSE)作为损失函数,Adam 算法迭代优化模型。

在深度学习领域,对于回归预测问题,常用的模型评价参数是平均绝对误差(mean absolute error, MAE)^[24],即对预测值和真实标记值直接计算绝对误差,作为算法评价标准,该值越小,说明结果越接近真实值,模型的估计效果越好。试验中采用 VGG16、ResNet18 以及 ResNet34 作为对比的网络模型,为了将它们应用于回归任务,均在原始网络的全连接层上进行修改,增加回归层(图 4(b))。

3.4 人工合成数据集试验

人工合成数据集具有纹理一致性强、纹理主方向真值精确已知的特点,故不存在标记的误差。为了验证 ResNet50-R 应用于纹理主方向估计问题的可行性,采用基于 VGG16、ResNet18、ResNet34 的回归模型(以下简称为 VGG16-R、ResNet18-R、ResNet34-R)和 ResNet50-R 对人工合成数据集进行训练,试验过程中保存模型训练得到的最优参数,然后用最优的模型对测试集进行测试,得到 VGG16-R、ResNet18-R、ResNet34-R 和 ResNet50 的 MAE 分别为 0.12° 、 0.11° 、 0.11° 和 0.08° 。可以看出,各个模型均取得很高的检测精度,整体上 MAE 均在 0.1° 左右,ResNet50-R 精度达到 0.1° 以内,说明残差网络回归模型能够准确地学习到从纹理特征

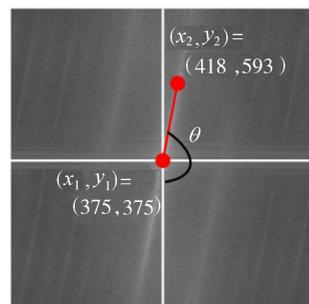


图 9 人工标记纹理主方向
Fig. 9 Main direction of spectrum by manual marking

到纹理主方向之间的非线性映射关系,从而给出高精度的预测结果,证明了 ResNet50-R 应用于纹理主方向估计问题的可行性。

3.5 攀枝花水文站数据集试验

与人工合成数据集相比,攀枝花水文站数据集更具有代表性,因为该数据集的时空图像均由真实河流视频合成得到,并且包含了多种典型场景下的时空图像。试验得到 VGG16-R、ResNet18-R、ResNet34-R、ResNet50-R 的 MAE 分别为 0.92° 、 0.82° 、 0.81° 和 0.80° ,可以看出,ResNet-R 的 MAE 均小于 VGG16-R 的 MAE,ResNet50-R 取得最好的检测精度,说明对于纹理主方向估计难度更大的攀枝花水文站数据集,残差网络依然能充分学习到多种场景下时空图像的纹理特征,从而整体上给出更高的检测精度,体现出残差网络的优越性。

为证明 ResNet50-R 的有效性,与传统方法 FFT 法进行对比,结果如表 1 所示。可以看出,在正常以及障碍物场景下即纹理特征较为明显的条件下,FFT 法的表现优于 ResNet50-R,但在复杂含噪声条件下如耀光、紊流、降雨,时空图像中纹理较为模糊,FFT 法检测结果容易出现粗大误差,导致整体的 MAE 偏大,而 ResNet50-R 通过在大量数据集上训练,学习到不同场景下时空图像中有效的纹理特征,从而能够在纹理清晰度较差的条件下给出更为准确的检测精度。图 10

表 1 不同场景下的 MAE 对比

Table 1 MAE comparison in different scenes

场景	MAE/($^\circ$)				
	VGG16-R	ResNet18-R	ResNet34-R	ResNet50-R	FFT 法
正常	0.66	0.65	0.62	0.61	0.31
障碍物	0.58	0.44	0.43	0.40	0.34
耀光	0.93	0.82	0.84	0.72	9.75
紊流	1.29	1.15	1.19	1.08	15.42
降雨	1.42	1.36	1.19	1.29	19.48

—— 人工标记的纹理主方向 - · - ResNet50-R 检测的纹理主方向

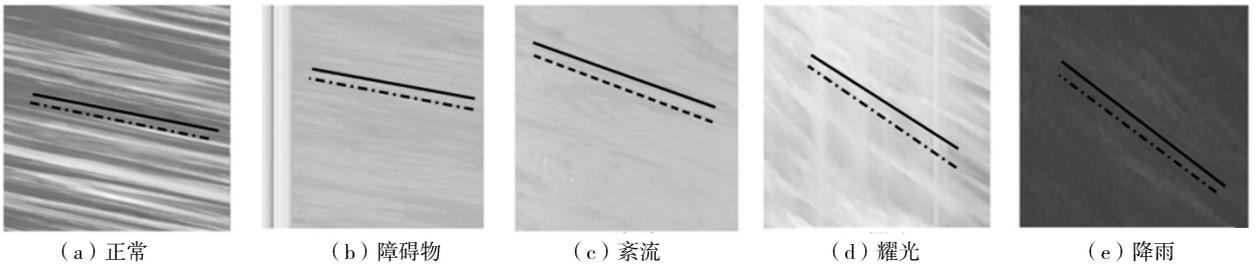


图 10 人工标记与 ResNet50-R 检测结果对比

Fig. 10 Comparison between manually marked value and ResNet50-R predicted value

是不同场景下人工标记与 ResNet50-R 的检测结果。可以看出,即使是在紊流、耀光及降雨的条件下,ResNet50-R 依然能给出较为准确的估计值,没有出现粗大误差,说明构建的残差网络回归模型能对噪声具有较好的鲁棒性,能够同时应用于多个场景,适用性好。

VGG16-R、ResNet18-R、ResNet34-R、ResNet50-R 及 FFT 法检测 100 张时空图像的检测时间分别为 6.59 s、3.40 s、3.82 s、3.87 s 和 4.80 s,可以看出,由于 VGG16-R 参数量大,故执行效率在深度学习模型中最低。除 VGG16-R 外,ResNet18-R、ResNet34-R 和 ResNet50-R 的执行效率均优于 FFT 法,表明残差网络回归模型的检测速度能满足实际测量的需求。与 ResNet18-R、ResNet34-R 相比,ResNet50-R 在攀枝花水文站的数据集上取得最好的检测精度,但模型的执行效率并没有明显下降,综合平衡各模型在检测精度与执行效率的表现,ResNet50-R 整体的性能更好。

3.6 流速比测试验

为了检验 ResNet50-R 应用于复杂含噪场景下的有效性,挑选了晴天和雨天两种气象条件的河流视频进行试验,试验用到的视频数据都来自于攀枝花水文站。试验中将基于 ResNet50-R 的 STIV(以下简称 ResNet50-STIV)与传统方法 FFT-STIV、人工目测值进行对比分析。人工目测值获取流程为:先通过人工标记的方式获取时空图像的纹理主方向并作为参考值,通过纹理主方向可以计算得到光流运动矢量,然后根据时空图像对应测速线上的物像尺度因子将光流运动矢量转换为流速。类似地,ResNet50-STIV 和 FFT-STIV 检测出纹理主方向之后按该流程计算得到对应的流速。此外,为进一步验证 ResNet50-STIV 在实际流速测量中应用的有效性,开展了与 GB 50179—2015《河流量测验规范》中标准流速仪法的流速比测试验。

3.6.1 晴天条件

晴天条件下河流视频的采集时间是 2020 年 7 月 29 日 9 时,视频的时长为 30 s,现场测流断面见

图 7(a), 可以看出, 河流的流向为自左向右, 河流整体的水流示踪条件受到多种因素的干扰, 如近岸紊流、障碍物遮挡、水面耀光等, 其中最为明显的干扰因素是耀光。测流的断面为矩形框区域, 测量了起点距从 45 ~ 160 m 的 9 条测速线, 包含了紊流(测速线 1)、障碍物(测速线 1)、正常(测速线 2、3、4、5)、耀光(测速线 6、7、8、9)和紊流(测速线 8、9)场景, 对应的时空图像如图 11 所示。

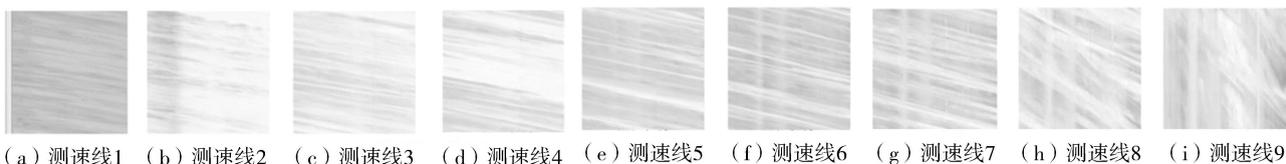


图 11 晴天条件的时空图像

Fig. 11 STIs under sunny conditions

采用人工标记方式、ResNet50-STIV 和 FFT-STIV 3 种方法分别对时空图像进行纹理主方向检测得到纹理主方向并转换为流速, 结果如表 2 所示。从表 2 的纹理主方向检测结果可以看出, 测速线 2 ~ 7 上的误差最小, 测速线 1、8、9 的误差偏大。测速线 1 位于近场, 同时受到了近场紊流和管道的遮挡, 因此两种方法的检测误差略大。测速线 2 ~ 5 受到干扰较少, 纹理清晰且方向的一致性较好, 测速线 6、7 垂直方向上叠加了由于水面耀光引起的干扰, 但整体的纹理特征依然较为明显, 因此对于测速线 2 ~ 7 对应的时空图像 FFT 法与 ResNet50-R 均能给出较为准确的纹理主方向。测速线 8、9 同时受到了耀光与紊流的影响, 导致时空图像的纹理清晰度急剧下降, 此时 FFT 法检测结果的误差分别为 2° 和 24.84°, 出现了较大误差, 而 ResNet50-R 经过含有多种复杂含噪场景的时空图像数据集的训练, 学习到提取有效纹理特征的能力并且对噪声有一定的鲁棒性, 进而给出了较为可靠的纹理主方向检测结果, 纹理主方向的误差在 2.1° 以内。

表 2 晴天条件纹理主方向检测结果和流速测量结果

Table 2 MOT detection results and flow velocity measurement results under sunny conditions

测速线	起点距/m	纹理主方向/(°)			纹理主方向绝对误差/(°)			流速/(m · s ⁻¹)		流速相对误差/%	
		FFT 法	ResNet50-R	人工目测值	FFT 法	ResNet50-R	FFT-STIV	ResNet50-STIV	人工目测值	FFT-STIV	ResNet50-STIV
1	45	81.45	80.45	80.89	0.56	0.44	2.48	2.19	2.33	6.43	-5.84
2	65	81.97	82.41	82.20	0.23	0.21	3.91	4.15	4.03	-2.98	2.98
3	75	80.57	80.94	80.80	0.23	0.14	3.92	4.07	4.01	-2.23	1.65
4	85	80.01	80.18	80.48	0.47	0.30	4.15	4.22	4.36	-4.82	-3.14
5	95	78.95	78.54	78.80	0.15	0.26	4.16	4.01	4.11	1.22	-2.43
6	115	77.39	77.34	77.52	0.13	0.18	4.29	4.27	4.34	-1.55	-1.64
7	130	75.43	74.41	74.45	0.98	0.04	4.20	3.91	3.92	-7.14	-0.26
8	145	68.10	66.32	66.10	2.0	0.22	3.08	2.96	2.79	10.39	6.16
9	160	18.55	41.31	43.39	24.84	2.08	0.44	1.14	1.25	-64.80	-8.64

注: 纹理主方向绝对误差为 FFT 法、ResNet50-R 的纹理主方向检测结果与人工目测值的绝对误差, 流速相对误差为 FFT 法、ResNet50-STIV 测量的流速与人工目测值的相对误差。下同。

由表 2 的流速结果可见, 整体上 ResNet50-STIV 得到流速结果与人工目测值更为接近, 相对误差在 9% 以内, 具有较好的鲁棒性, 而 FFT-STIV 在同时受到多种噪声干扰的场景下, 纹理主方向的检测结果出现较大偏差导致得到的部分流速值不可靠。

3.6.2 雨天条件

雨天条件下河流视频的采集时间是 2020 年 8 月 24 日 9 时, 雨天的拍摄条件较为恶劣, 光线较暗并且受到降雨的影响, 水流示踪物的可见性较差, 测量了起点距从 50 ~ 160 m 的 9 条测速线, 主要包含了紊流(测速线 1、8、9)、降雨(测速线 1 ~ 9)场景, 对应的时空图像如图 12 所示。

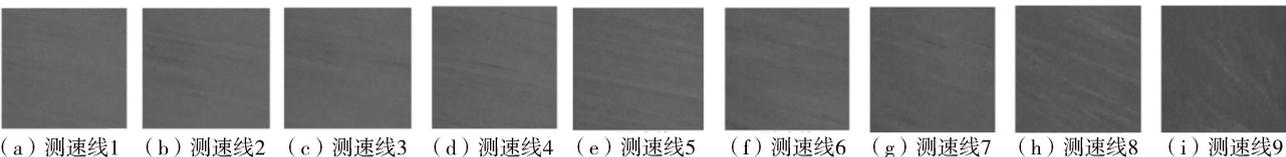


图 12 雨天条件的时空图像

Fig. 12 STIs under rainy conditions

3种方法得到的纹理主方向和流速分别如表3所示,在测速线1、8、9上的时空图像由于同时受到降雨和紊流的影响,有效的纹理特征变得十分模糊并且纹理方向一致性较差,此时对于依赖于纹理清晰度的FFT法,由于时空图像的空域中有效的纹理特征少使得频域中有效信号的能量较弱,信噪比较低,导致纹理主方向检测结果中出现了 54.07° 、 10.78° 以及 13.95° 的偏差。而ResNet50-R由于具备较强的特征提取能力,虽然纹理特征较为微弱,依然能提取到有效的纹理特征,进而学习到从纹理特征到角度空间的复杂非线性映射函数,因此能给出较好的检测结果,误差在 2.5° 以内。测速线2~7虽然也受到降雨的干扰,纹理的清晰度有一定的下降,但由于位于中泓区域,流速稳定,使得示踪物形成的纹理较为明显,因此FFT法和ResNet50-R均能较为准确地检测出纹理主方向。

表3 雨天条件纹理主方向检测结果和流速测量结果

Table 3 MOT detection results and flow velocity measurement results under rainy conditions

测速线	起点距/m	纹理主方向/ $^\circ$			纹理主方向绝对误差/ $^\circ$			流速/ $(\text{m} \cdot \text{s}^{-1})$			流速相对误差/%	
		FFT法	ResNet50-R	人工目测值	FFT法	ResNet50-R	FFT-STIV	ResNet50-STIV	人工目测值	FFT-STIV	ResNet50-STIV	
1	50	23.95	76.91	78.02	57.07	1.11	0.20	2.02	2.21	-90.95	-8.33	
2	60	80.44	79.53	80.01	0.43	0.48	3.28	2.99	3.16	3.80	-5.32	
3	75	80.42	80.46	80.78	0.36	0.32	3.66	3.69	3.81	-3.94	-3.33	
4	85	79.31	78.42	78.31	1.00	0.11	3.69	3.41	3.37	9.50	1.36	
5	95	78.50	79.18	78.75	0.25	0.43	3.92	4.18	4.01	-2.23	4.26	
6	105	77.02	77.20	77.21	0.19	0.01	3.80	3.86	3.87	-1.81	-0.03	
7	120	74.55	74.91	74.68	0.13	0.23	3.71	3.81	3.74	-0.80	1.84	
8	140	76.12	67.76	65.34	10.78	2.42	4.75	2.87	2.56	85.55	12.27	
9	160	23.30	39.24	37.25	13.95	1.99	0.56	1.06	0.99	-43.43	7.47	

由表3可以看出,在示踪条件较好的测速线(2~7)上,FFT-STIV和ResNet50-STIV流速分布都接近于人工目测值,此时ResNet50-STIV的流速相对误差在5.32%以内。对于受到噪声影响较大的测速线(1、8、9),FFT-STIV的流速结果就会出现异常值,但ResNet50-STIV能够给出较准确的流速值,流速相对误差在12.3%以内,验证了ResNet50-STIV应用于雨天条件下的有效性。

3.6.3 流速仪法比测试验

为了验证ResNet50-STIV应用于实际流速测量的有效性,于2020年8月21日8:25—9:25在攀枝花水文站开展了与目前GB 50179—2015《河流流量测验规范》中的标准流速仪法的比测试验。测速期间水位保持在998 m,为高水条件,水流示踪物及表面流速较为稳定;天气为阴天,水面无风,也无明显耀光干扰。试验采用9时拍摄的一段30 s视频测量了起点距55~175 m的9条测速线,包含了正常(测速线2、3、4)、近岸紊流(测速线1、9)和障碍物(测速线5、6、7、8、9)场景,对应的时空图像如图13所示。

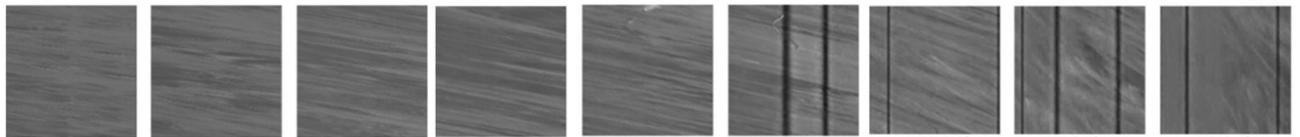


图13 比测试验中的时空图像

Fig. 13 STIs in comparison experiment

从表4纹理主方向检测结果可以看出,正常场景下的误差最小,障碍物和紊流场景下误差较大。测速线1虽受到紊流干扰,但有效纹理特征依然较为明显。测速线5和6尽管分别受到铅鱼流速仪与黑色线缆的局部遮挡,但由于中泓区域流速稳定、示踪物形成的图像纹理清晰且一致性较好,因此在测速线1~6上,FFT法与ResNet50-R都能较为准确地检测出纹理主方向。测速线7、8、9在受到线缆遮挡的同时存在对岸倒影的干扰,表现为有效纹理模糊并且一致性差,存在多个干扰方向的条纹;此外,测速线9处于远场,分辨率较低且受到紊流的影响。此时在测速线7、8、9上FFT法的检测结果出现了较大的偏差,从 3.30° ~ 33.82° ,在测速线9处误差最大,而ResNet50-R由于具备提取到有效纹理特征的能力,给出了较为可靠的纹理主方向估计值,误差在 2.1° 以内。

表 4 流速仪法比测试验纹理主方向检测结果和流速测量结果

Table 4 MOT detection results and flow velocity measurement results of flow meter method comparison test

测速线	起点距/m	纹理主方向/(°)			纹理主方向绝对误差/(°)			流速/(m·s ⁻¹)			流速相对误差/%	
		FFT 法	ResNet50-R	人工目测值	FFT 法	ResNet50-R	FFT-STIV	ResNet50-STIV	流速仪法	FFT-STIV	ResNet50-STIV	
1	55	82.71	82.92	82.60	0.11	0.32	3.51	3.62	2.99	8.23	11.37	
2	65	81.66	82.16	81.54	0.12	0.62	3.62	3.86	3.30	1.14	7.58	
3	90	80.61	80.41	80.69	0.08	0.27	4.49	4.40	4.26	-2.82	-4.93	
4	105	79.32	79.50	79.30	0.18	0.20	4.55	4.63	4.22	-0.59	0.95	
5	120	78.18	78.93	78.08	0.20	0.85	4.75	5.08	4.65	-5.82	-0.43	
6	135	77.55	76.91	77.44	0.14	0.53	4.99	4.74	4.41	4.33	-1.13	
7	155	73.10	69.33	69.80	3.30	0.47	4.17	3.36	3.35	-17.71	-7.76	
8	165	70.52	58.79	56.80	13.72	2.01	3.79	2.22	1.83	90.95	11.75	
9	175	72.02	37.73	38.20	33.82	0.50	3.51	1.10	1.13	259.01	-10.13	

注:流速相对误差为 FFT-STIV、ResNet50-STIV 测量的流速与流速仪法测量的流速的相对误差。

为保证流速仪法测量的稳定性,其转子应完全入水,因此实际测量了水深为 0.2 m 的流速作为比测参考值。从表 4 可以看出,相比流速仪法,ResNet50-STIV 测量的断面流速分布呈现出趋势相同但数值略偏大的规律。这是由于 STIV 测量的是表面水流,与流速仪法所测的水深之间存在一个流速系数,由对数或指数型河道垂线流速分布规律所决定^[25]。这里用各垂线流速系数的均值 0.92 对滤波后的流速进行校正,得到相对误差范围为 0.43% ~ 11.75%。在测速线 1 处的流速误差较大,为 11.37%,但由于纹理主方向检测值和人工目测值的误差为 0.32°,在合理范围内,判断其主要误差源于紊流引起铅鱼偏角过大导致流速仪测量值偏小。

4 结 语

鉴于现有 STIV 无法解决复杂含噪场景下纹理主方向精确估计的问题,本文提出了一种基于残差网络回归模型的 STIV。ResNet50-R 在人工合成数据集下的检测精度可到 0.1°,验证了其应用于纹理主方向估计问题的可行性;在攀枝花水文站数据集下,具有残差结构的 ResNet 回归模型优于 VGG16 回归模型,并且 ResNet50-R 能较好地平衡检测精度以及执行效率,在正常场景下的检测精度可到 0.7°,复杂含噪场景下可到 1.3°,表明通过在包含障碍物、紊流、耀光、降雨复杂含噪场景的数据集上进行训练,ResNet50-R 能够学习到此类场景下时空图像的有效纹理特征,进而抑制环境扰动造成的粗大误差,从而提高方法的场景适用性。在单次表面流速比测试验中,ResNet50-STIV 的相对误差可达到浮标法测流精度,其端到端的实现方式比传统方法更简捷,可作为图像法测流技术的一种新方法。

本文提出的方法虽然在复杂含噪场景下的表现优于 FFT-STIV,但测量精度仍然有提升的空间,未来将从两个方面进行改进:一是对数据进行预处理,通过自适应边缘检测增强图像的纹理特征,使得构建的回归模型更易于学习到复杂变化场景下时空图像的纹理特征,进而提升检测精度;二是优化数据集,包括引入频域滤波技术辅助纹理主方向的标记以进一步提升人工标记的精度,提高数据集纹理主方向的分辨率;三是采集数量更多以及场景更丰富的时空图像进一步扩充数据集,使得数据集更具有代表性。

参考文献:

- [1] 万定生,王坤,朱跃龙,等. 中小河流洪水预报智能调度平台关键技术[J]. 河海大学学报(自然科学版),2021,49(3):204-212. (WAN Dingsheng, WANG Kun, ZHU Yuelong, et al. Key technologies of intelligent flood forecasting and dispatching platform for small and medium-sized basins[J]. Journal of Hohai University (Natural Sciences), 2021, 49(3):204-212. (in Chinese))
- [2] 朱跃龙,赵群,余宇峰,等. 基于时空特征挖掘的流量过程智能模拟方法[J]. 河海大学学报(自然科学版),2021,49(1):7-12. (ZHU Yuelong, ZHAO Qun, YU Yufeng, et al. Intelligent simulation method of runoff process based on spatiotemporal feature mining[J]. Journal of Hohai University (Natural Sciences), 2021, 49(1):7-12. (in Chinese))
- [3] 何秉顺,李青. 山洪灾害防御技术现状与发展趋势探索[J]. 中国水利,2014(18):11-13. (HE Bingshun, LI Qing. Exploration on present situation and developing tendency of mountain flood disaster prevention technology[J]. China Water Resources, 2014(18):11-13. (in Chinese))
- [4] 杨聃,邵广俊,胡伟飞,等. 基于图像的河流表面测速研究综述[J]. 浙江大学学报(工学版),2021,55(9):1752-1763.

- (YANG Dan, SHAO Guangjun, HU Weifei, et al. Review of image-based river surface velocimetry research [J]. Journal of Zhejiang University (Engineering Science), 2021, 55(9):1752-1763. (in Chinese))
- [5] 张振, 徐枫, 王鑫, 等. 河流水面成像测速研究进展[J]. 仪器仪表学报, 2015, 36(7):1441-1450. (ZHANG Zhen, XU Feng, WANG Xin, et al. Research progress on river surface imaging velocimetry [J]. Chinese Journal of Scientific Instrument, 2015, 36(7):1441-1450. (in Chinese))
- [6] JODEAU M, HAUET A, PAQUIER A, et al. Application and evaluation of LS-PIV technique for the monitoring of river surface velocities in high flow conditions [J]. Flow Measurement and Instrumentation, 2008, 19(2):117-127.
- [7] LE COZ J, HAUET A, PIERREFEU G, et al. Performance of image-based velocimetry (LSPIV) applied to flash-flood discharge measurements in Mediterranean rivers [J]. Journal of Hydrology, 2010, 394(1/2):42-52.
- [8] TSUBAKI R, FUJITA I, TSUTSUMI S. Measurement of the flood discharge of a small-sized river using an existing digital video recording system [J]. Journal of Hydro-Environment Research, 2011, 5(4):313-321.
- [9] FUJITA I, MUSTE M, KRUGER A. Large-scale particle image velocimetry for flow analysis in hydraulic engineering applications [J]. Journal of Hydraulic Research, 1998, 36(3):397-414.
- [10] FUJITA I, WATANABE H, TSUBAKI R. Development of a non-intrusive and efficient flow monitoring technique: the space-time image velocimetry (STIV) [J]. International Journal of River Basin Management, 2007, 5(2):105-114.
- [11] FUJITA I, NOTOYA Y, TANI K, et al. Efficient and accurate estimation of water surface velocity in STIV [J]. Environmental Fluid Mechanics, 2019, 19(5):1363-1378.
- [12] ZHEN Z, HUABAO L, YANG Z, et al. Design and evaluation of an FFT-based space-time image velocimetry (STIV) for time-averaged velocity measurement [C]//Proceedings of the 14th IEEE International Conference on Electronic Measurement & Instruments. New York: IEEE, 2019:503-514.
- [13] ZHAO H Y, CHEN H, LIU B Y, et al. An improvement of the space-time image velocimetry combined with a new denoising method for estimating river discharge [J]. Flow Measurement and Instrumentation, 2021, 77:101864.
- [14] FUJITA I, SHIBANO T, TANI K. Application of masked two-dimensional Fourier spectra for improving the accuracy of STIV-based river surface flow velocity measurements [J]. Measurement Science and Technology, 2020, 31(9):094015.
- [15] LIANG Lili, YE Shihuo, SHI Guangming. Arbitrary-orientated fan filter and its application in direction detection of images [J]. Acta Photonica Sinica, 2011, 40(12):1815-1819.
- [16] 桑高丽, 陈虎, 赵启军. 一种基于深度卷积网络的鲁棒头部姿态估计方法 [J]. 四川大学学报(工程科学版), 2016, 48(增刊1):163-169. (SANG Gaoli, CHEN Hu, ZHAO Qijun. Robust head pose estimation based on deep convolution neural networks [J]. Journal of Sichuan University (Engineering Science Edition), 2016, 48(Sup1):163-169. (in Chinese))
- [17] LIU X, LI S X, KAN M N, et al. AgeNet: deeply learned regressor and classifier for robust apparent age estimation [C]//Proceedings of 2015 IEEE International Conference on Computer Vision Workshop. New York: IEEE, 2015:258-266.
- [18] LUO P, TIAN Y L, WANG X G. Switchable deep network for pedestrian detection [C]//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2014:899-906.
- [19] ZHANG J, SHAN S G, KAN M N, et al. Coarse-to-fine auto-encoder networks (CFAN) for real-time face alignment [C]//Proceedings of the 13th European Conference on Computer Vision. Zurich: Springer, 2014:1-16.
- [20] JOURABLOO A, LIU X M. Large-pose face alignment via CNN-based dense 3D model fitting [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016:4188-4196.
- [21] CHEN J C, KUMAR A, RANJAN R, et al. A cascaded convolutional neural network for age estimation of unconstrained faces [C]//Proceedings of the 2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems. New York: IEEE, 2016:1-8.
- [22] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition [C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016:770-778.
- [23] 张振, 吕莉, 石爱业, 等. 基于物像尺度变换的河流水面流场定标方法 [J]. 仪器仪表学报, 2017, 38(9):2273-2281. (ZHANG Zhen, LYU Li, SHI Aiye, et al. River surface flow field calibration method based on object-image scaling [J]. Chinese Journal of Scientific Instrument, 2017, 38(9):2273-2281. (in Chinese))
- [24] KUO P H, HUANG C J. A high precision artificial neural networks model for short-term energy load forecasting [J]. Energies, 2018, 11(1):213.
- [25] 付辉, 杨开林, 王涛, 等. 对数型流速分布公式的参数敏感性及其取值 [J]. 水利学报, 2013, 44(4):489-494. (FU Hui, YANG Kailin, WANG Tao, et al. Analysis of parameter sensitivity and value for logarithmic velocity distribution formula [J]. Journal of Hydraulic Engineering, 2013, 44(4):489-494. (in Chinese))