



(12) 发明专利申请

(10) 申请公布号 CN 117896504 A

(43) 申请公布日 2024. 04. 16

(21) 申请号 202410091719.4

G06V 20/40 (2022.01)

(22) 申请日 2024.01.22

(71) 申请人 京甲云信息技术(南京)有限公司

地址 210000 江苏省南京市建邺区奥体大街69号1栋504-1室

(72) 发明人 周冕 刘敏杰 周柄兆 席梓航

卯水龙 刘志豪

(74) 专利代理机构 安徽华普专利代理事务所

(普通合伙) 34151

专利代理师 高孝强

(51) Int. Cl.

H04N 7/18 (2006.01)

H04N 7/10 (2006.01)

H04B 3/54 (2006.01)

G06N 3/0464 (2023.01)

权利要求书2页 说明书9页 附图6页

(54) 发明名称

基于电力线载波与深度学习融合的电力监控系统及方法

(57) 摘要

本发明提出了一种基于电力线载波与深度学习融合的电力监控系统,包括视频数据采集模块,用于捕获视频数据;电力线载波收发模块,用于接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;视频解码播放模块,用于接收并解析电力线传输的视频数据,并按照用户配置的输出协议发送到外部视频设备;视频高清播放模块,经过视频渐对齐融合网络模块处理后的视频目标帧,被送入图像超分辨率重建网络模块中进行超分辨率重建操作。本发明充分利用电力线作为传输介质,实现了视频数据的短距离传输,降低布线成本和人力施工成本,通过视频超分辨率重建模型应用于电力视频监控中,有效地提升了监控视频的分辨率和视觉效果。



1. 基于电力线载波与深度学习融合的电力监控系统,其特征在于,包括视频数据采集模块,用于捕获视频图像数据以及视频音频数据;

电力线载波收发模块,用于接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;

视频解码播放模块,用于接收并解析电力线传输的视频图像数据以及视频音频数据,经处理后的视频数据按照用户配置的输出协议发送到外部视频设备执行播放控制;

视频高清播放模块,经过视频渐对齐融合网络模块处理后的视频目标帧,被送入图像超分辨率重建网络模块中进行超分辨率重建操作,实现视频高清播放。

2. 根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,电力线载波收发模块由电力线控制芯片和模拟前端构成,芯片物理层采用OFDM调制方式,通过将待发送的信息码元通过位换并减小速率,增加码元周期,以减弱多径干扰的影响。

3. 根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,视频渐对齐融合网络模块包括特征提取模块、特征对齐模块和特征融合模块,特征提取模块利用卷积神经网络对输入视频帧进行有效特征提取,特征对齐、特征融合模块是指目标帧之前与之后的支持帧都是从视频序列最两端,逐帧向中间目标帧对齐和融合。

4. 根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,视频渐对齐融合网络模块的具体步骤为:

低分辨率视频的目标帧表示为 $I_t$ ,则视频帧序列表示为 $\{I_{t+k}\}$ , $r \in [-1, +r]$ ,1为目标帧前侧视频帧数量, $r$ 为目标帧后侧视频帧数量。首先通过卷积对 $I_{t+k}$ 进行特征提取得到 $F_{t+k}$ , $F_{t+k} \in \mathbb{R}^{W*H*C}$ ,特征序列 $\{F_{t+k}\}$ 渐对齐融合过程如下;

令 $F_{t-1}^L = F_{t-1}$ ,使用对齐模块将 $F_{t-1}^L$ 对齐到前侧方向下一帧特征 $F_{t-1+1}$ ,然后使用融合模块将对齐后的 $F_{t-1}^L$ 与 $F_{t-1+1}$ 进行融合,把融合后特征继续对齐到前侧方向下一帧并融合,重复上述操作,直到 $F_{t-1}$ ,得到 $F_{t-1}^L$ ,即目标帧前侧所有特征对齐融合结果,计算如下式所示:

$$F_{t-1}^L = \text{TSA2}(\text{TDAN}(F_{t-1-1}^L, F_{t-i}), F_{t-i}), \quad i=1-1, \dots, 1$$

令 $F_{t+r}^R = F_{t+r}$ ,使用对齐模块将 $F_{t+r}^R$ 对齐到后侧方向下一帧特征 $F_{t+r-1}$ ,然后使用融合模块将对齐后的 $F_{t+r}^R$ 与 $F_{t+r-1}$ 进行融合,把融合后特征继续对齐到后侧方向下一帧并融合,重复上述操作,直到 $F_{t+1}$ ,得到 $F_{t+1}^R$ ,即目标帧后侧所有特征对齐融合结果,计算如下式所示:

$$F_{t+j}^R = \text{TSA2}(\text{TDAN}(F_{t+j+1}^R, F_{t+j}), F_{t+j}), \quad j=r-1, \dots, 1$$

使用融合模块将上述得到的特征 $(F_{t-1}^L, F_t, F_{t+1}^R)$ 进行对齐和融合,得到所有输入序列对齐融合后的特征 $F'_t$ ,计算如下式所示:

$$F'_t = \text{TSA3}(\text{TDAN}(F_{t-1}^L, F_t), F_t, \text{TDAN}(F_{t+1}^R, F_t))$$

式中, $\text{TSA3}(\cdot)$ 表示输入特征图数量为3的融合模块。

5. 根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在

于,对齐模块使用目标帧和每个支持帧的特征,动态预测支持帧的形变偏移量,借助可形变卷积,将这些偏移量添加到支持帧的特征上,从而完成对齐操作。

6.根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,图像超分辨率重建网络模块具体步骤包括:

通过卷积操作将目标帧的通道数提升至中间特征通道数;

借助含有残差模块的特征选择网络进行非线性映射;

再次应用卷积操作以降低通道数;

利用亚像素卷积进行特征图尺寸的扩大,并将其与原图像的上采样结果相加,从而获得最终的超分辨率重建图像。

7.根据权利要求6所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,所述残差模块在激活层之后,首先通过11卷积降低特征图的通道数量,接着利用33卷积进行特征提取和映射,最后使用1\*1卷积进行通道数还原。

8.根据权利要求1所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,还包括特征选择网络模块,该网络由不同深度的分支组成,每个分支由偶数个残差模块构成,随着层次的加深,分支网络的学习能力逐渐增强,根据输入信息的不同,调整不同尺度特征的比例,并将它们进行融合。

9.根据权利要求8所述的基于电力线载波与深度学习融合的电力监控系统,其特征在于,所述特征融合模块具体步骤包括:

对于特征选择模块的各个分支特征 $U_i$ ,通过元素求和的方式获得融合结果 $U$ ;

随后,利用全局平均池化来整合全局信息,生成信道统计信息 $S$ ;

接下来,通过一个全连接层生成一个紧凑特征 $Z$ ,以便进行精确且自适应的特征选择;

用softmax来进行选择,将其与原分支特征 $U_i$ 进行元素级乘法,从而得到 $V_i$ ,最终,通过求和操作,获得融合后的特征 $V$ 。

10.基于电力线载波与深度学习融合的电力监控方法,其特征在于,包括以下步骤:

S1:视频数据采集模块捕获视频图像数据以及视频音频数据;

S2:电力线载波收发模块接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;

S3:视频解码播放模块接收并解析电力线传输的视频图像数据以及视频音频数据,经处理后的视频数据按照用户配置的输出协议发送到外部视频设备执行播放控制;

S4:视频高清播放模块将经过视频渐对齐融合网络模块处理后的视频目标帧送入图像超分辨率重建网络模块中进行超分辨率重建操作,实现视频高清播放。

## 基于电力线载波与深度学习融合的电力监控系统及方法

### 技术领域

[0001] 本发明属于视频监控技术领域,具体涉及一种基于电力线载波与深度学习融合的电力监控系统及方法。

### 背景技术

[0002] 视频监控已广泛应用于生产、管理和检测等各个领域,当前市场上的监控产品主要划分为两类:模拟摄像机和IPC摄像机。模拟摄像机通过同轴电缆传输视频信号,但其缺点在于像素受限,通常为200W像素以下。而另一种是IPC(网络摄像机),它通过有线或无线网络传输视频流。这两种市场产品的共同不足之处在于视频数据传输常借助光纤、宽带网络、双绞线等,从而在生产和安装过程中耗费大量人力和物力资源,同时传统的监控视频的图像分辨率有待提高,以提升对监测对象细节的识别度,在此基础上,提出了一种基于电力线载波与深度学习融合的电力监控系统及方法。

### 发明内容

[0003] 本发明主要提供了一种基于电力线载波与深度学习融合的电力监控系统及方法,用以解决上述背景技术中提出的技术问题。

[0004] 本发明解决上述技术问题采用的技术方案为:

[0005] 基于电力线载波与深度学习融合的电力监控系统,包括

[0006] 视频数据采集模块,用于捕获视频图像数据以及视频音频数据;

[0007] 电力线载波收发模块,用于接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;

[0008] 视频解码播放模块,用于接收并解析电力线传输的视频图像数据以及视频音频数据,经处理后的视频数据按照用户配置的输出协议发送到外部视频设备执行播放控制;

[0009] 视频高清播放模块,经过视频渐对齐融合网络模块处理后的视频目标帧,被送入图像超分辨率重建网络模块中进行超分辨率重建操作,实现视频高清播放。

[0010] 进一步的改进在于:电力线载波收发模块由电力线控制芯片和模拟前端构成,芯片物理层采用OFDM调制方式,通过将待发送的信息码元通过位换并减小速率,增加码元周期,以减弱多径干扰的影响。

[0011] 进一步的改进在于:视频渐对齐融合网络模块包括特征提取模块、特征对齐模块和特征融合模块,特征提取模块利用卷积神经网络对输入视频帧进行有效特征提取,特征对齐、特征融合模块是指目标帧之前与之后的支持帧都是从视频序列最两端,逐帧向中间目标帧对齐和融合。

[0012] 进一步的改进在于:视频渐对齐融合网络模块的具体步骤为:

[0013] 低分辨率视频的目标帧表示为 $I_t$ ,则视频帧序列表示为 $\{I_{t+k}\}$ , $r \in [-1, +r]$ , $l$ 为目标帧前侧视频帧数量, $r$ 为目标帧后侧视频帧数量。首先通过卷积对 $I_{t+k}$ 进行特征提取得到 $F_{t+k}$ , $F_{t+k} \in \mathbb{R}^{W \times H \times C}$ ,特征序列 $\{F_{t+k}\}$ 渐对齐融合过程如下;

[0014] 令  $F_{t-1}^L = F_{t-1}$ , 使用对齐模块将  $F_{t-1}^L$  对齐到前侧方向下一帧特征  $F_{t-1+1}$ , 然后使用融合模块将对齐后的  $F_{t-1}^L$  与  $F_{t-1+1}$  进行融合, 把融合后特征继续对齐到前侧方向下一帧并融合, 重复上述操作, 直到  $F_{t-1}$ , 得到  $F_{t-1}^L$ , 即目标帧前侧所有特征对齐融合结果, 计算如下式所示:

$$[0015] \quad F_{t-1}^L = \text{TSA2}(\text{TDAN}(F_{t-1-1}^L, F_{t-1}), F_{t-1}), \quad i=1-1, \dots, 1$$

[0016] 令  $F_{t+r}^R = F_{t+r}$ , 使用对齐模块将  $F_{t+r}^R$  对齐到后侧方向下一帧特征  $F_{t+r-1}$ , 然后使用融合模块将对齐后的  $F_{t+r}^R$  与  $F_{t+r-1}$  进行融合, 把融合后特征继续对齐到后侧方向下一帧并融合, 重复上述操作, 直到  $F_{t+1}$ , 得到  $F_{t+1}^R$ , 即目标帧后侧所有特征对齐融合结果, 计算如下式所示:

$$[0017] \quad F_{t+j}^R = \text{TSA2}(\text{TDAN}(F_{t+j+1}^R, F_{t+j}), F_{t+j}), \quad j=r-1, \dots, 1$$

[0018] 使用融合模块将上述得到的特征  $(F_{t-1}^L, F_t, F_{t+1}^R)$  进行对齐和融合, 得到所有输入序列对齐融合后的特征  $F'_t$ , 计算如下式所示:

$$[0019] \quad F'_t = \text{TSA3}(\text{TDAN}(F_{t-1}^L, F_t), F_t, \text{TDAN}(F_{t+1}^R, F_t))$$

[0020] 式中,  $\text{TSA3}(\cdot)$  表示输入特征图数量为3的融合模块。

[0021] 进一步的改进在于: 对齐模块使用目标帧和每个支持帧的特征, 动态预测支持帧的形变偏移量, 借助可形变卷积, 将这些偏移量添加到支持帧的特征上, 从而完成对齐操作。

[0022] 进一步的改进在于: 图像超分辨率重建网络模块具体步骤包括:

[0023] 通过卷积操作将目标帧的通道数提升至中间特征通道数;

[0024] 借助含有残差模块的特征选择网络进行非线性映射;

[0025] 再次应用卷积操作以降低通道数;

[0026] 利用亚像素卷积进行特征图尺寸的扩大, 并将其与原图像的上采样结果相加, 从而获得最终的超分辨率重建图像。

[0027] 进一步的改进在于: 所述残差模块在激活层之后, 首先通过11卷积降低特征图的通道数量, 接着利用33卷积进行特征提取和映射, 最后使用1\*1卷积进行通道数还原。

[0028] 进一步的改进在于: 还包括特征选择网络模块, 该网络由不同深度的分支组成, 每个分支由偶数个残差模块构成, 随着层次的加深, 分支网络的学习能力逐渐增强, 根据输入信息的不同, 调整不同尺度特征的比例, 并将它们进行融合。

[0029] 进一步的改进在于: 所述特征融合模块具体步骤包括:

[0030] 对于特征选择模块的各个分支特征  $U_i$ , 通过元素求和的方式获得融合结果  $U$ ;

[0031] 随后, 利用全局平均池化来整合全局信息, 生成信道统计信息  $S$ ;

[0032] 接下来, 通过一个全连接层生成一个紧凑特征  $Z$ , 以便进行精确且自适应的特征选择;

[0033] 用softmax来进行选择, 将其与原分支特征  $U_i$  进行元素级乘法, 从而得到  $V_i$ , 最终, 通过求和操作, 获得融合后的特征  $V$ 。

[0034] 本发明还提供了一种基于电力线载波与深度学习融合的电力监控方法,包括以下步骤:

[0035] S1:视频数据采集模块捕获视频图像数据以及视频音频数据;

[0036] S2:电力线载波收发模块接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;

[0037] S3:视频解码播放模块接收并解析电力线传输的视频图像数据以及视频音频数据,经处理后的视频数据按照用户配置的输出协议发送到外部视频设备执行播放控制;

[0038] S4:视频高清播放模块将经过视频渐对齐融合网络模块处理后的视频目标帧送入图像超分辨率重建网络模块中进行超分辨率重建操作,实现视频高清播放。

[0039] 与现有技术相比,本发明的有益效果是:

[0040] 本发明充分利用电力线作为传输介质,实现了视频数据的短距离传输,该系统在经济方面带来了诸多优势,如降低布线成本和人力施工成本等,在实际效果方面,成功地解决了家庭监控系统布线难题和无线信号弱的问题。通过将设计的视频超分辨率重建模型应用于电力视频监控中,有效地提升了监控视频的分辨率和视觉效果,实验结果表明,这一方案在各个领域均具备可行性,例如停车场、工厂、仓库、家庭以及银行等,总体而言,本方案具备广阔的市场应用前景;

[0041] 本发明采用卷积神经网络对输入视频帧进行特征提取,通过按照时间顺序将输入视频帧序列的支持帧从序列两端对齐到目标帧,结合时间与空间自注意力机制将支持帧特征逐帧融合到目标帧上,接着,引入多尺度特征学习机制,利用多分支特征选择重建网络对经过特征融合的目标帧进行超分辨率重建,在公共数据集和电力视频数据集上对该方法进行了训练,并经过测试验证,测试结果显示,所提出的方法在峰值信噪比和结构相似性指标均值方面分别达到了27.08和0.819,有效地提高了电力监控视频的分辨率和视觉效果。

[0042] 以下将结合附图与具体的实施例对本发明进行详细的解释说明。

## 附图说明

[0043] 图1是本发明的系统流程图;

[0044] 图2是本发明的视频超分辨率重建模型结构示意图;

[0045] 图3是本发明的时间形变对齐模块结构示意图;

[0046] 图4是本发明的时间空间注意力融合模块示意图;

[0047] 图5是本发明的图像超分辨率重建网络结构示意图;

[0048] 图6是本发明的残差模块的改进结构示意图;

[0049] 图7是本发明的特征选择网络结构图;

[0050] 图8是本发明的特征融合模块结构图;

[0051] 图9是本发明的模型性能与参数量之间关系图;

[0052] 图10是本发明的模型重建主观视觉对比效果;

[0053] 图11是本发明的视频帧重建前后对照图。

## 具体实施方式

[0054] 为了便于理解本发明,下面将参照相关附图对本发明进行更加全面的描述,附图

中给出了本发明的若干实施例,但是本发明可以通过不同的形式来实现,并不限于文本所描述的实施例,相反的,提供这些实施例是为了使对本发明公开的内容更加透彻全面。

[0055] 具体参照图1,本实施例提供了一种基于电力线载波与深度学习融合的电力监控系统,包括以下步骤:

[0056] 视频数据采集模块,用于捕获视频图像数据以及视频音频数据;

[0057] 电力线载波收发模块,用于接收视频数据采集模块发送的对应数据,并通过正交频分复用编码将数据调制到电力线上传输;

[0058] 视频解码播放模块,用于接收并解析电力线传输的视频图像数据以及视频音频数据,经处理后的视频数据按照用户配置的输出协议发送到外部视频设备执行播放控制;

[0059] 视频高清播放模块,经过视频渐对齐融合网络模块处理后的视频目标帧,被送入图像超分辨率重建网络模块中进行超分辨率重建操作,实现视频高清播放。

[0060] 在本系统中,视频数据采集模块的摄像头采用了目前市场上常用的IPC摄像头,这领域国内已取得成熟进展,众多主要厂商已拥有自身成熟的解决方案,如TI的达芬奇平台、华为的Hi3516/Hi3518平台等,以及视频解码播放模块采用了华为海思的Hi3518E解决方案,在此不做过多介绍。

[0061] 其中,电力线载波通讯模块作为本设计的核心,主要用于接收IPC摄像头发送的视频网络包,并通过正交频分复用(OFDM)编码将数据调制到电力线上传输,该模块由电力线调制控制芯片和模拟前端构成,优选采用高通QCA6410芯片和AR1500模拟前端,芯片物理层采用OFDM调制方式,通过将待发送的信息码元通过位换并减小速率,从而增加码元周期,以减弱多径干扰的影响。同时,采用循环前缀(CP)作为保护间隔,极大地减少了码间干扰,并确保各信道之间的正交性,从而显著减少信道间干扰。在OFDM中,各载波频谱存在1/2重叠的正交性,提升了OFDM调制方式的频谱利用率。此外,它采用了TurboFEC错误校验,物理层线路速率达到了200Mbps,在电缆传输中,传输距离可达1公里,传输效率为100Mbps。

[0062] 视频高清播放模块基于渐对齐特征融合残差网络的视频超分辨率重建模型结构如图2所示(以5帧视频中第3帧作为目标帧为例),该模型包含视频渐对齐融合网络和图像超分辨率重建网络两个模块。

[0063] 视频渐对齐融合网络由特征提取、特征对齐和特征融合三个分模块组成,其中,特征提取模块利用卷积神经网络对输入视频帧进行有效特征提取。

[0064] 特征对齐和特征融合模块是指目标帧(TR)之前与之后的支持帧(LR)都是从视频序列最两端,逐帧向中间目标帧对齐和融合。

[0065] 视频渐对齐融合网络工作时:

[0066] 低分辨率视频的目标帧表示为 $I_t$ ,则视频帧序列表示为 $\{I_{t+k}\}$ , $r \in [-1, +r]$ ,1为目标帧前侧视频帧数量, $r$ 为目标帧后侧视频帧数量。首先通过卷积对 $I_{t+k}$ 进行特征提取得到 $F_{t+k}$ , $F_{t+k} \in \mathbb{R}^{W*H*C}$ ,特征序列 $\{F_{t+k}\}$ 渐对齐融合过程如下。

[0067] 令 $F_{t-1}^L = F_{t-1}$ ,使用对齐模块将 $F_{t-1}^L$ 对齐到前侧方向下一帧特征 $F_{t-1+1}$ ,然后使用融合模块将对齐后的 $F_{t-1}^L$ 与 $F_{t-1+1}$ 进行融合,把融合后特征继续对齐到前侧方向下一帧并融合,重复上述操作,直到 $F_{t-1}$ ,得到 $F_{t-1}^L$ ,即目标帧前侧所有特征对齐融合结果,计算如式(1)所示。

$$[0068] \quad F_{t-l}^L = \text{TSA2}(\text{TDAN}(F_{t-l-1}^L, F_{t-i}), F_{t-i}), i=l-1, \dots, 1 \quad (1)$$

[0069] 令  $F_{t+r}^R = F_{t+r}$ , 使用对齐模块将  $F_{t+r}^R$  对齐到后侧方向下一帧特征  $F_{t+r-1}$ , 然后使用融合模块将对齐后的  $F_{t+r}^R$  与  $F_{t+r-1}$  进行融合, 把融合后特征继续对齐到后侧方向下一帧并融合, 重复上述操作, 直到  $F_{t+1}$ , 得到  $F_{t+1}^R$ , 即目标帧后侧所有特征对齐融合结果, 计算如式 (2) 所示。

$$[0070] \quad F_{t+j}^R = \text{TSA2}(\text{TDAN}(F_{t+j+1}^R, F_{t+j}), F_{t+j}), j=r-1, \dots, 1 \quad (2)$$

[0071] 使用融合模块将上述得到的特征  $(F_{t-1}^L, F_t, F_{t+1}^R)$  进行对齐和融合, 得到所有输入序列对齐融合后的特征  $F'_t$ 。计算如式 (3) 所示。

$$[0072] \quad F'_t = \text{TSA3}(\text{TDAN}(F_{t-1}^L, F_t), F_t, \text{TDAN}(F_{t+1}^R, F_t)) \quad (3)$$

[0073] 式中,  $\text{TSA3}(\cdot)$  表示输入特征图数量为3的融合模块。

[0074] 该网络旨在对经过特征提取后的特征图进行对齐, 以降低原图像中遮挡、模糊及其他噪声的影响, 采用逐渐对齐的方式, 每次只在相邻的两帧之间进行对齐, 从而增强了对复杂运动图像的对齐鲁棒性。此外, 此方法允许引入更多的支持帧, 利用更多的场景信息, 进而有助于提升模型的重建效果。

[0075] 对齐模块采用一种时间形变对齐网络, 完成运动隐式补偿。在特征级上将支持帧  $F_L$  对齐到目标帧  $F_T$ , 能更精确地估计目标运动和形变, 其过程如图3所示, 采用可形变卷积是在特征提取之后, 以特征级别的方式进行支持帧对齐操作, 该方法涉及使用目标帧和每个支持帧的特征, 动态预测支持帧的形变偏移量, 接着借助可形变卷积, 将这些偏移量添加到支持帧的特征上, 从而完成对齐操作。相对于直接在原图像上进行光流估计对齐, 这一方法不容易受到复杂运动和伪影的影响, 同时具备端到端学习的优势。

[0076] 对齐后图像进行特征融合是视频超分辨率任务中的关键问题, 本实施例提出了一种基于时间空间注意力多帧图像特征融合方法 (TSA), 以2个输入特征图为例 (TSA2) TSA, 其基本结构如图4所示。

[0077] 假设需要融合  $n$  个特征图, 每个特征图表示为  $U (i=1, 2, \dots, n)$ ,  $U \in \mathbb{R}^{W \times H \times C}$ , 则特征融合过程如下。

[0078] (1) 通过对位元素相加方式融合所有输入特征图, 得所有输入特征矩阵和  $U$ ,  $U \in \mathbb{R}^{W \times H \times C}$ , 计算如式 (4) 所示。

$$[0079] \quad U = \sum_{i=1}^n U_i \quad (4)$$

[0080] (2) 对  $U$  进行全局平均池化, 得到通道向量  $S$ ,  $S \in \mathbb{R}^C$ , 计算如式 (5) 所示。

$$[0081] \quad S_c = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W U_c(w, h) \quad (5)$$

[0082] 式中,  $S_c$  和  $U_c$  分别表示  $S$  和  $U$  第  $c$  个特征通道。

[0083] (3) 使用两个全连接层对特征图各通道间相关性进行建模, 第一个全连接层将通道数减小为  $C/r$ ,  $r$  为压缩因子 (本文实验设为4), 第二个全连接层将通道数还原为  $C$ , 获得相关性建模后的特征向量  $Z$ , 计算如式 (6) 所示。

$$[0084] \quad Z = W_2 \cdot (\delta(W_1 \cdot U)) \quad (6)$$

[0085] 式中： $W_1 \in \mathbb{R}^r \times c$ ， $W_2 \in \mathbb{R}^c \times c$ 分别代表两个全连接层的权重矩阵； $\delta$ 代表ReLU激活函数。

[0086] (4) 使用 $1 \times 1$ 卷积分别将 $\{U_i\}$ 尺寸变为 $W \times H$ ，学习每个输入特征矩阵在空间维度上的内部相关性，计算如式(7)所示。

$$[0087] \quad V_i = \text{CNN}_{1 \times 1}(W_3, U_i) \quad (7)$$

[0088] 式中： $\text{CNN}_{1 \times 1}(U_i)$ 表示卷积核为 $1 \times 1$ 卷积层； $W_3 \in \mathbb{R}^{1 \times 1 \times c}$ 为卷积层权重矩阵。

[0089] (5) 将步骤(3)中得到的 $Z$ 分别同 $\{V_i\}$ 相乘，得到 $\{A_i\}$ ， $A_i \in \mathbb{R}^{W \times H \times c}$ ，计算如式(8)所示。

$$[0090] \quad A_i = V_i \cdot Z \quad (8)$$

[0091] (6) 使用sigmoid函数对 $\{A_i\}$ 进行重标定，得到权重矩阵 $\{B_i\}$ ， $B_i \in \mathbb{R}^{W \times H \times c}$ ，这是在特征图每个位置上单独计算，其过程如式(9)所示。

$$[0092] \quad B_{i,w,h,c} = \frac{e^{A_{i,w,h,c}}}{\sum_k^n e^{A_{k,w,h,c}}} \quad (9)$$

[0093] 式中， $(w, h, c)$ 代表像素空间坐标和通道位置。

[0094] (7) 最后将权重向量 $\{B_i\}$ 和各输入特征 $\{U_i\}$ 位相乘得到 $\{\hat{U}_i\}$ ，将 $\{\hat{U}_i\}$ 相加得到最终融合结果 $\hat{U}$ ，计算如(10)所示。

$$[0095] \quad \hat{U} = \sum_{i=1}^n (B_i \odot U_i) \quad (10)$$

[0096] 式中， $\odot$ 表示对位元素相乘。

[0097] 其中， $H$ 和 $W$ 分别表示图的高与宽的分辨率， $S$ 表示上采样倍数(2、4、8等)，在该网络中，所有卷积都使用 $3 \times 3$ 卷积层，中间层特征通道数设为128。

[0098] 该方法的独特之处在于对不同位置的特征进行重新标定，通过学习，为不同位置的特征分配不同的权重，这些“位置”包括图像帧在视频序列中的序列位置(时间位置)、空间位置以及通道位置。

[0099] 经过渐对齐融合网络处理后的视频目标帧，将被送入图像超分辨率重建网络，进行超分辨率重建操作，重建网络结构如图5所示，首先，通过卷积操作将目标帧的通道数提升至中间特征通道数。接着，借助含有残差模块的特征选择网络进行非线性映射。随后，再次应用卷积操作以降低通道数。最后，利用亚像素卷积进行特征图尺寸的扩大，并将其与原图像的上采样结果相加，从而获得最终的超分辨率重建图像。

[0100] 本实施例所采用的残差模块结构，采用深度残差网络进行学习，能够有效地解决深层神经网络中可能出现的梯度消失或梯度爆炸问题，如图6所示，在宽泛激活超分辨率网络(WDSR)残差模块的基础上引入了瓶颈结构进行改进，在激活层之后，首先通过11卷积降低特征图的通道数量，接着利用33卷积进行特征提取和映射，最后使用 $1 \times 1$ 卷积进行通道数还原。这一改进后的残差模块不仅降低了计算负担，同时也增强了模块对通道之间关系的建模能力。

[0101] 根据人类神经元感受野不同的机制，本文在残差模块的基础上设计了一种特征选

择网络模块,如图7所示,该网络由不同深度的分支组成,标记为 $\{fk\}$  ( $[1, 2, \dots, n]$ ),每个分支由 $2(k-1)$ 个残差模块构成,随着层次的加深,分支网络的学习能力逐渐增强,可以根据输入信息的不同,自主学习和动态调整不同尺度特征的比例,并将它们进行融合。这是对残差学习机制的一种扩展。

[0102] 这一网络的特征提取机制类似于GoogleNet的Inception结构,它采用多个具有不同感受野大小的分支网络来提取多尺度特征。然而,与之不同的是,首先,在分支网络的结构设计方面,本模块采用不同数量的残差模块堆叠而成的分支网络;其次,在特征融合方面,本网络采用基于自注意力机制的多尺度特征融合方法。

[0103] 为了融合各分支网络的多尺度特征,本文提出了一种基于自注意力机制的多尺度特征融合方法。其基本思想是赋予模型“注意力”,使其能够忽略与当前任务无关的冗余信息,而专注于某些“感兴趣”的关键信息。这种机制可以有效地从大量稀疏数据中提取最为重要的特征。对于两分支网络的结构示意图,可参考图8。

[0104] 首先,对于特征选择模块的各个分支特征 $U_i$ ,通过元素求和的方式获得融合结果 $U$ ;随后,利用全局平均池化来整合全局信息,生成信道统计信息 $S$ ;接下来,通过一个全连接层生成一个紧凑特征 $Z$ ,以便进行精确且自适应的特征选择;此步骤同时降低了维度以提高效率。在信道维度上,我们使用softmax来进行选择,然后将其与原分支特征 $U_i$ 进行元素级乘法,从而得到 $V_i$ 。最终,通过求和操作,获得融合后的特征 $V$ 。

[0105] 这个融合模块能够自动学习两个分支特征图各自的重要程度,即权重,然后利用这些权重将这些特征图进行加权相加。

[0106] 在视频渐对齐融合网络中,我们采用卷积神经网络对输入视频帧进行特征提取。通过按照时间顺序将输入视频帧序列的支持帧从序列的两端逐帧对齐到目标帧,同时运用时间与空间自注意力机制,将支持帧特征逐帧融合到目标帧上。

[0107] 在图像超分辨率重建网络方面,我们在深度残差网络基础上进行了改进,扩展了残差块结构,增加了中间层的激活宽度。引入多尺度特征学习机制,并构建了一种特征选择网络。我们运用亚像素卷积来扩展特征融合后目标帧的特征图尺寸。最后,将该特征图与原图像上采样结果相加,得到最终的超分辨率重建图像。

[0108] 本实施例中对所构建的模型,分别使用了公共数据集和电力视频数据集进行了训练,在训练过程中,我们使用了Vimeo-90K数据集作为训练数据集,同时使用Vid4和Reds数据集作为基准测试集来评估模型的重建效果。

[0109] 在训练阶段,我们使用双三次插值来获取低分辨率图像,并通过随机翻转和旋转来进行数据增强。我们采用L1距离作为损失函数,进行4倍超分辨率的重建测试,输入数据为 $64 \times 64$ 的低分辨率RGB视频,每批次大小为16,总共进行了60万个批次的训练,所有网络层的学习率初始化为 $4 \times 10^{-4}$ 。

[0110] 以下对本实施例的高清视频进行效果测试

[0111] 1) 模型重建性能

[0112] 在本实施例中,我们将所提出的重建模型与当前代表性的模型,如ESPCN、VESPCN、TOFlow、TDAN和RBPN,在测试数据集Vid4上进行了比较,通过超过4倍的分辨率重建性能来进行客观指标的评估,包括峰值信噪比 (PSNR) 和结构相似性 (SSIM),对比结果如表1所示。

[0113] 表1模型性能客观指标对比表

模型	日历		城市		树叶		行人		平均	
	PSNR	SSMI								
ESPCN	21.76	0.679	25.82	0.670	24.58	0.652	28.05	0.859	25.06	0.715
VESPCN	21.98	0.691	26.17	0.696	24.91	0.673	28.31	0.861	25.34	0.730
[0114] TOFlow	22.47	0.732	26.78	0.740	25.27	0.709	29.05	0.879	25.89	0.765
TDAN	22.98	0.756	26.99	0.757	25.51	0.717	29.50	0.890	26.24	0.780
RBPN	23.99	0.807	27.73	0.803	26.22	0.757	30.70	0.909	27.12	0.818
本文模型	23.91	0.798	27.78	0.812	26.21	0.761	30.77	0.917	27.16	0.821

[0115] 从表格1的数据可以清楚地看出,本实施例中提出的模型在峰值信噪比 (PSNR) 和结构相似性 (SSIM) 方面的均值分别达到27.16和0.821,其重建效果整体上优于所列模型的表现。

[0116] 2) 模型重建效率

[0117] 本实施例分辨率重建模型与对比模型在参数量与客观指标峰值信噪比 (PSNR) 之间的对比关系如图5所示。

[0118] 从图表9可以明显观察到,尽管本文模型的参数量比TDAN和VESPCN多了38%,但其在PSNR指标上比TDAN高出0.92dB,比VESPCN高出1.82dB;与RBPN模型相比,PSNR性能基本相当,但参数量减少了约70%;同时,本文模型在PSNR性能上超越了TOFlow和ESPCN模型,并且参数量更少;总体来看,本文模型在重建性能和参数量之间取得了很好的平衡。

[0119] 3) 模型主观性能

[0120] 通过主观视角的视觉比较,我们可以对比本文模型与VESPCN、TOFlow、TDAN和RBPN模型的重建性能,在图6中,我们选择了测试集中某个视频段的目标帧重建效果进行展示。

[0121] 在图10中,左上角展示的是原始图像,其余部分展示了不同模型的重建效果,与原始图像中的黑色方框相对应。从图11中可以明显观察到,与对比模型相比,本文模型在字母结构方面表现最为完整,能够无歧义地还原该信息,同时能够很好地恢复出细节纹理,因此其重建效果是最优的。

[0122] 电力视频数据集训练

[0123] 在充分训练公共数据集后,本文模型还经过电力视频数据集的适用性训练,以提升其性能。

[0124] 1) 数据集制作与训练

[0125] 本文采用的电力视频数据集主要来自于某220kV高压送电线路和某10kV高压配电线路的监控视频。这些视频片段涵盖了不同季节、不同天气、不同时间段和不同拍摄内容,视频片段的长度随机分布在5帧到60帧之间。该数据集总共包含500个视频片段,其中包括400个用于训练、50个用于验证和50个用于测试。在训练过程中,进行了数据增强操作,训练过程与使用公共数据集的训练过程相一致。

[0126] 2) 模型性能

[0127] 在完成训练后,我们对测试集进行了峰值信噪比 (PSNR) 和结构相似性 (SSIM) 指标的测试。测试结果显示,模型在测试集上的PSNR和SSIM均值分别为27.08和0.819。这些结果与在公共数据集上训练得到的模型性能接近。

[0128] 将训练完成后的模型应用于某高压送电线路监控视频的超分辨率重建任务中,我们从中选择了第4、10和30帧的重建图作为图11所示。这些图展示了模型在实际场景中的重建效果。

[0129] 从图11中可以观察到,在重建前后视频帧的视觉角度下,模型能够成功地重建出更为完整的细节纹理,从而显著提升了监控视频的分辨率和视觉效果。

[0130] 综上,基于电力线载波与深度学习融合的高清电力监控视频系统设计,充分利用电力线作为传输介质,实现了视频数据的短距离传输。该系统在经济方面带来了诸多优势,如降低布线成本和人力施工成本等。在实际效果方面,成功地解决了家庭监控系统布线难题和无线信号弱的问题。通过将设计的视频超分辨率重建模型应用于电力视频监控中,有效地提升了监控视频的分辨率和视觉效果。实验结果表明,这一方案在各个领域均具备可行性,例如停车场、工厂、仓库、家庭以及银行等。总体而言,本方案具备广阔的市场应用前景。

[0131] 尽管已经示出和描述了本发明的实施例,对于本领域的普通技术人员而言,可以理解在不脱离本发明的原理和精神的情况下可以对这些实施例进行多种变化、修改、替换和变型,本发明的范围由所附权利要求及其等同物限定。



图1

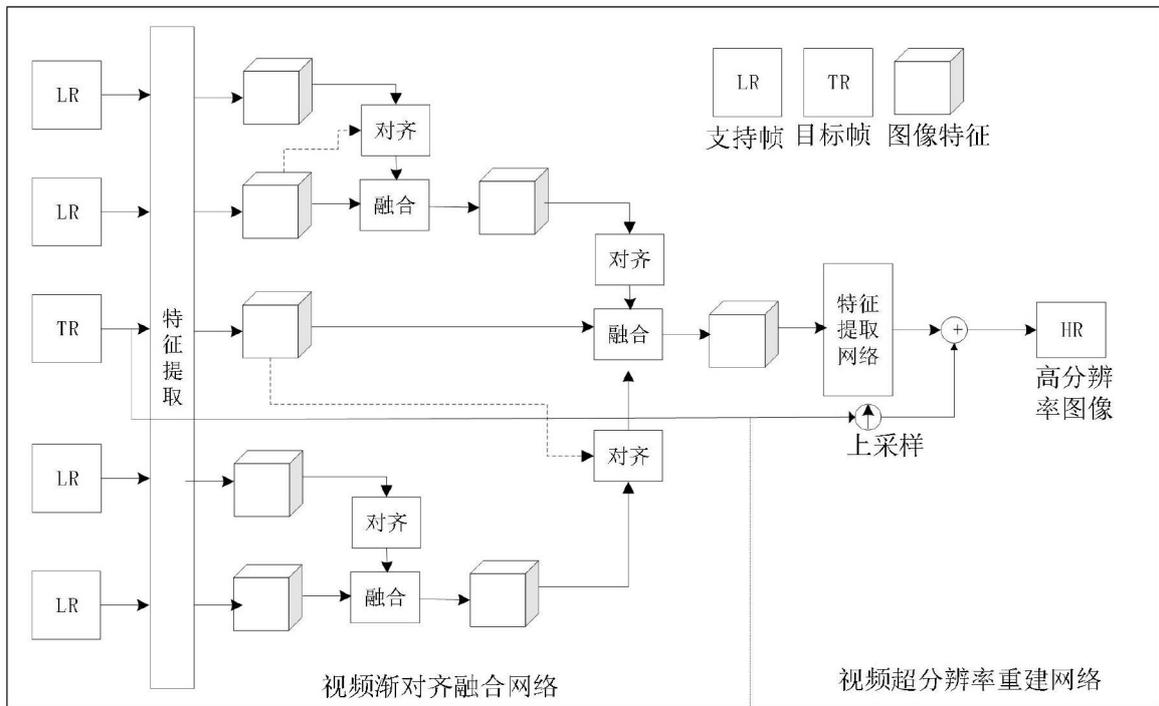


图2



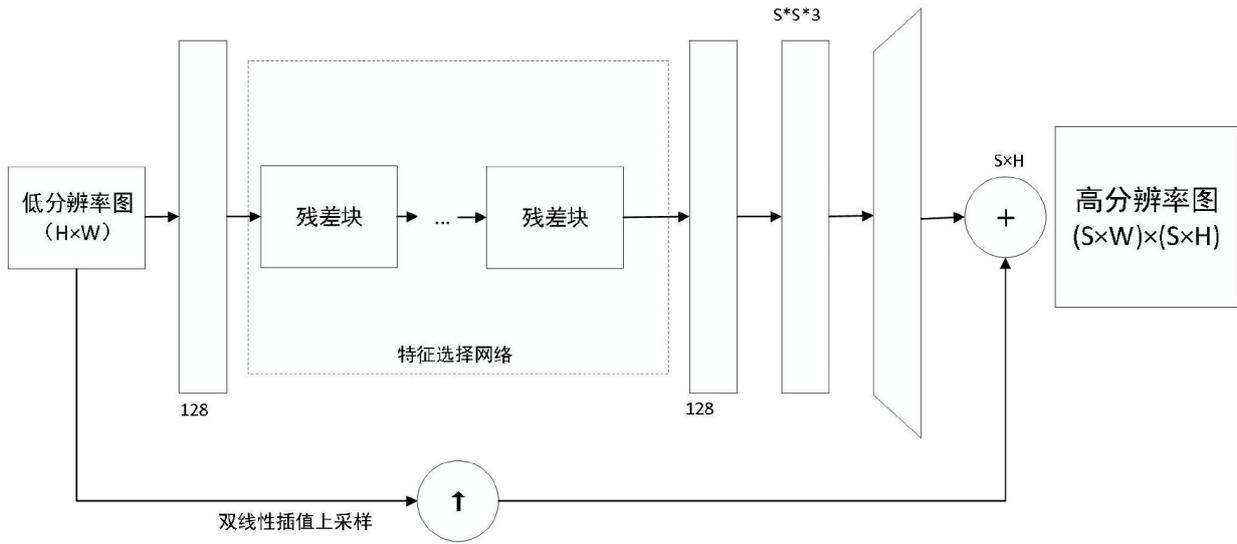


图5

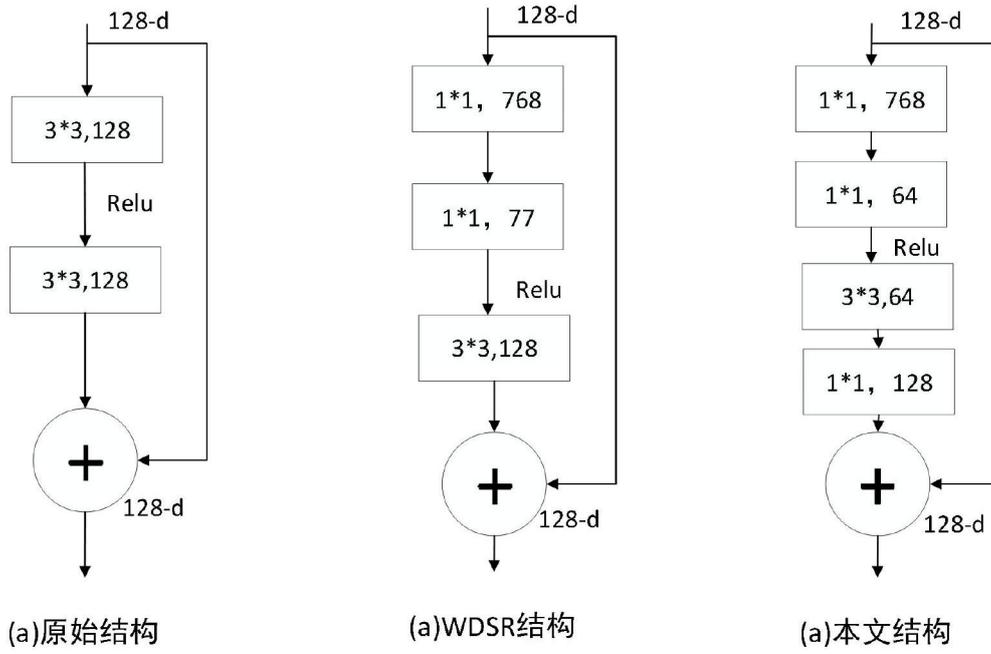


图6

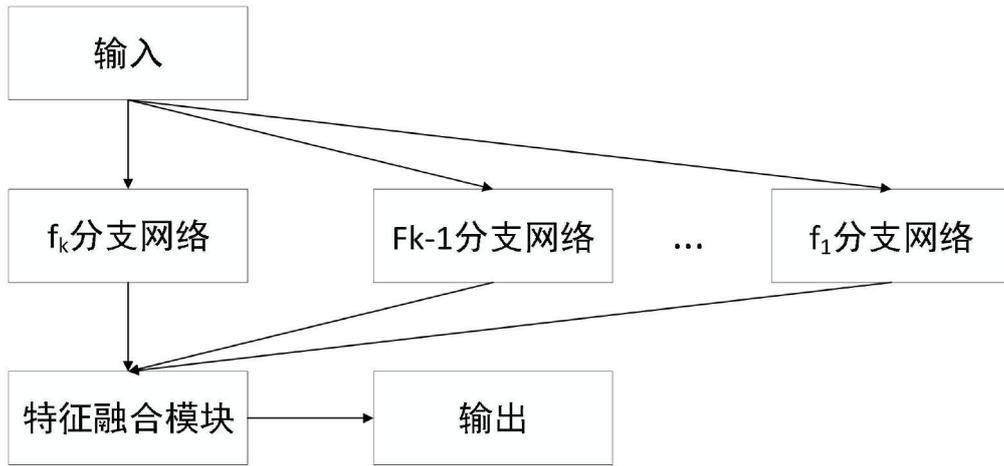


图7

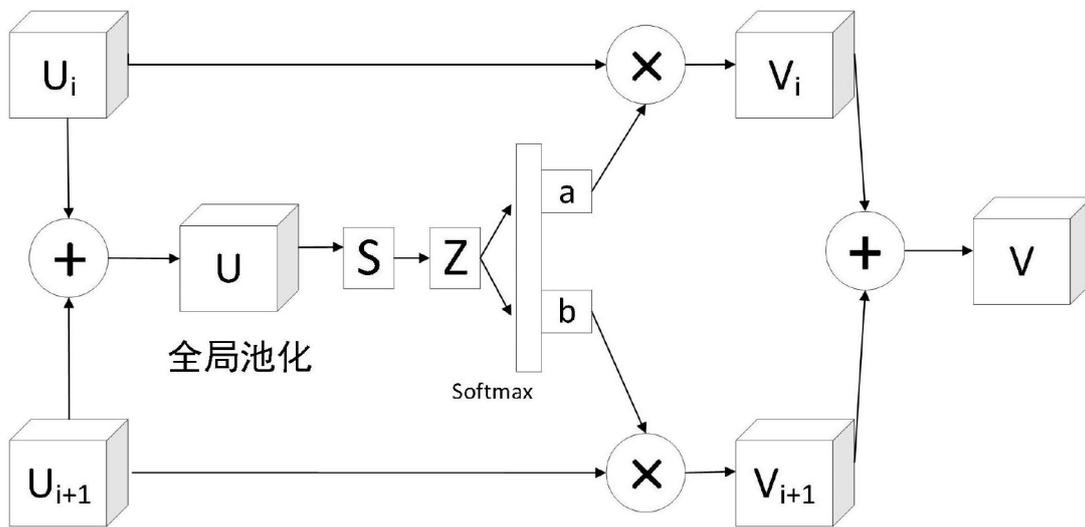


图8

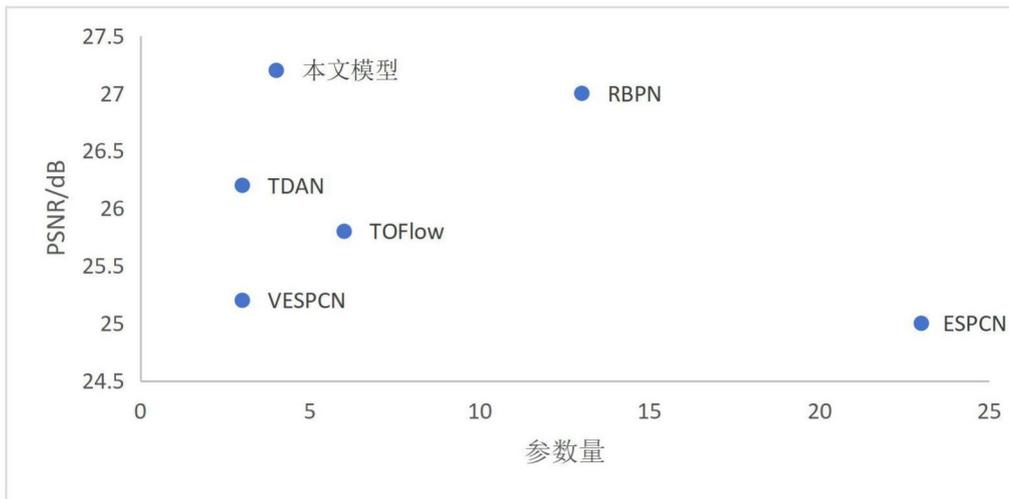


图9

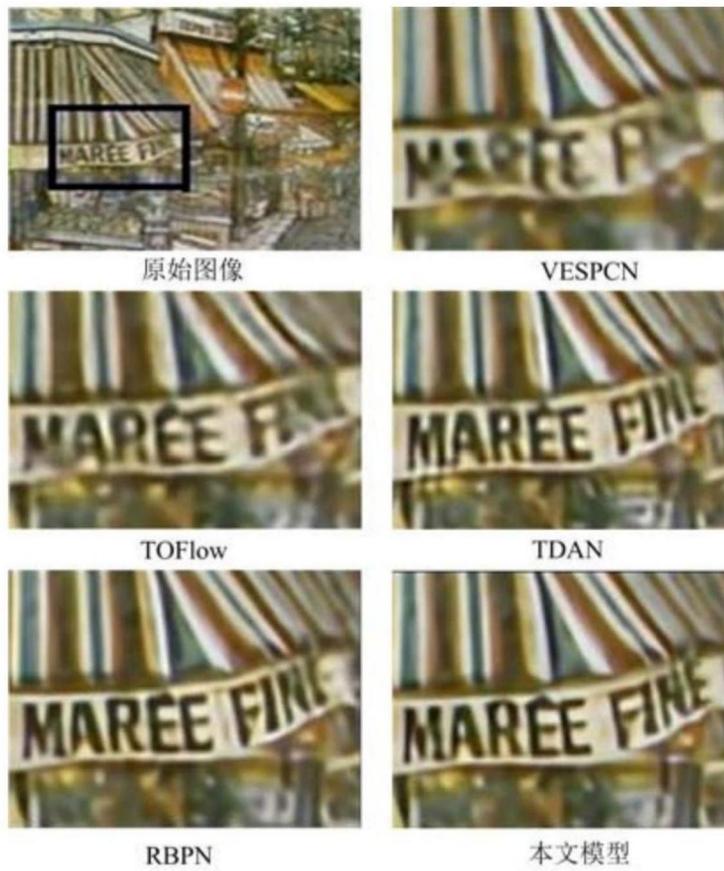


图10

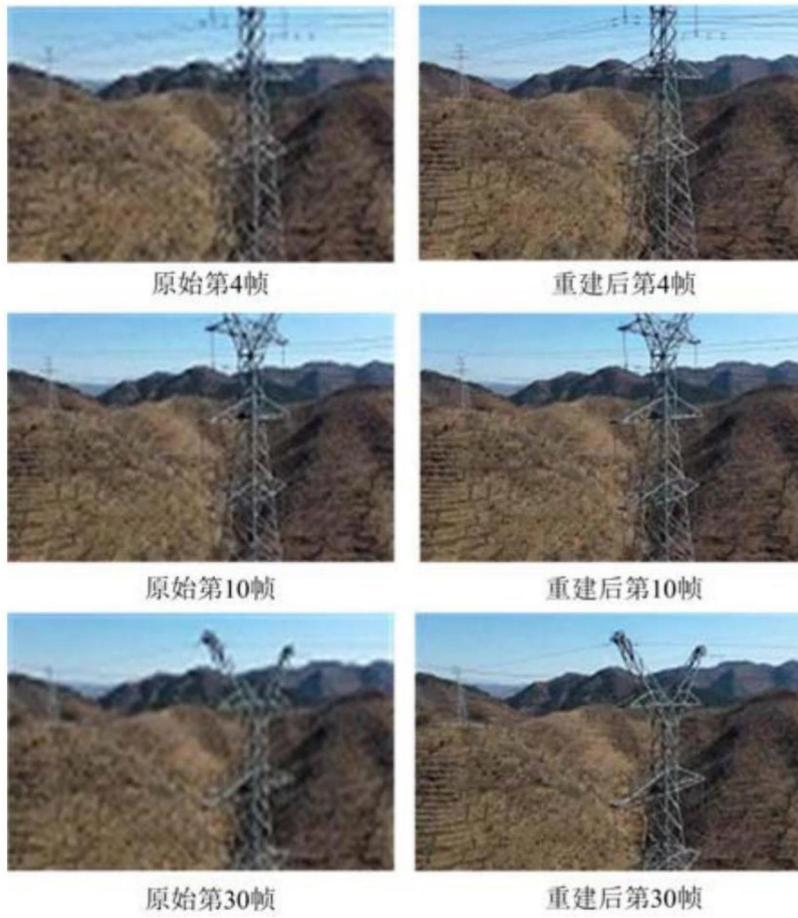


图11