

doi:10.11835/j.issn.1000-582X.2019.05.011

基于引导滤波器和加权二维主成分分析的视频融合

徐 丹¹, 巩沛琪², 郭松涛^{1,3}, 王 英¹ 姚 菁⁴

(1.西南大学 电子信息工程学院, 重庆 40715; 2.河南漯河市第一高级中学, 河南 漯河 462000;
3.重庆大学 计算机学院, 重庆 400044; 4.香港大学 理学院, 香港 999077)

摘 要:可见光视频可以提供纹理信息,而红外视频可以提供隐藏的热信息。通过 2 种视频的融合可以提高移动用户的视频观看体验。然而,由于移动设备资源有限,复杂的视频处理任务被迁移到资源(计算、存储和电池资源)相对丰富的微云上执行。基于均值哈希的帧间冗余检测算法被提出,将去冗之后的视频帧传输到微云等待处理,基于引导滤波器和加权二维主成分分析(W2DPCA, weighted two-dimensional principal component analysis)的视频融合算法被提出。该算法首先使用引导滤波器将待融合视频帧分成基层和细层,然后,利用改进的自适应 W2DPCA 融合可见光帧的基层和红外帧的基层。最后,通过组合融合基层和细层来获取融合帧。实验结果表明,帧间冗余检测方法最大限度地减少了微云中冗余数据的传输量,降低了移动设备的能耗。与现有方法相比,提出的视频融合算法得到的融合帧具有与原始帧更多的互信息和更高的结构相似度,同时融合结果也具有较高的整体标准差和峰值信噪比,所以具有更好的整体融合效果。

关键词:冗余检测;视频融合;可视视频;红外视频;微云

中图分类号:TP393

文献标志码:A

文章编号:1000-582X(2019)05-095-13

Video fusion based on guided filter and weighted two-dimensional PCA

XU Dan¹, GONG Peiqi², GUO Songtao^{1,3}, WANG Ying¹, YAO Jing⁴

(1. College of Electronic and Information Engineering, Southwest University, Chongqing 400715, P. R. China; 2. Luohe First Senior High School, Luohe 462000, Henan P.R.China;
3. College of Computer Science, Chongqing University, Chongqing 400044, P. R. China;
4. Faculty of Science, University of Hong Kong, Hong Kong 99077, P. R. China)

Abstract: Visible videos can mainly provide texture information, whereas infrared videos provide hidden hot information. The fusion of both kinds of videos can generate higher quality video viewing experience. However, due to the limited mobile device resources, the complicated video processing tasks are off loaded to the more powerful cloudlet with more sufficient resources (computing, storage, and battery resources) to be performed. In this paper the inter-frame redundancy detection algorithm based on the mean hash is proposed, and the video frames after redundancy removal are transmitted to the cloudlet for processing. The video fusion algorithm based on the guided filter and the weighted two-dimensional principal component analysis (W2DPCA) is proposed. The video frames to be fused are first divided into base layer and detail layer using guided filter. Then, the base layers of visible frame and infrared frame are

收稿日期:2019-01-20

基金项目:国家自然科学基金资助项目(61772432, 61772433)。

Supported by National Natural Science Foundation of China (61772432, 61772433).

作者简介:徐丹(1994—),女,硕士,主要从事数据融合的研究,(E-mail)1320204093@qq.com。

fused by the improved adaptive W2DPCA. Finally, the fused frame is acquired by combination of the fusional base layer and the reserved detail layer. Experiment results show that the redundancy detection method minimizes the amount of redundant data transmission in cloudlets and reduces the energy consumption of mobile devices. Compared with the existing methods, the fused frame obtained by the video fusion algorithm in this paper has more mutual information and higher structural similarity to the original frame, and the fusion result also has higher overall standard deviation and peak signal-to-noise ratio thus having better overall integration effect.

Keywords: redundancy detection; video Fusion; visible video; infrared video; cloudlet

随着视频处理和传输技术的发展,监控系统的应用越来越广泛^[1]。人们期望使用便携式设备来监控周围环境并及时发现火灾等意外情况。由于移动设备的资源有限所以难以进行复杂的计算。为了解决这个问题,将计算密集型任务从资源受限的智能移动设备转移到微云或服务器,以提高移动设备的计算能力。微云位于移动设备和云之间以应对数据路由,因此可以为移动设备提供更近,更强大的计算和存储能力^[2]。另一方面,由于单个传感器不足以提供关于目标场景的完整视频信息,可以通过融合同一场景的多个视频来了解一个场景的全部信息^[3]。因此,同一场景多个视频的融合在监控视频中显得十分重要。由于恶劣环境的影响,可见光视频中的目标物体可能无法被清楚地观察到,相比之下,红外视频可以不受周围环境影响显示隐藏的热物体^[4]。同时可见光视频又可以补充在相应的红外视频中经常丢失的丰富纹理信息。从图 1 中,可以找到可见光视频和红外视频各自的优势。在图 1(a)中,可以清楚地观察到草地的真实纹理,但是阴影中的人很容易被忽略。在图 1(b)中,存在许多与草地真实纹理不同的红外斑点,在红外帧中,可以清楚地检测隐藏在阴影中的人。根据它们的互补优势,可将可见光和红外视频进行融合,以提高整体观看质量^[5-7]。虽然融合后的视频一定程度上丢掉了可见光视频的真实性,降低了红外视频中目标物体的显著性,但其兼具了可见光视频与红外视频的特点,增强了视频的可读性,并可应用于特定场景。

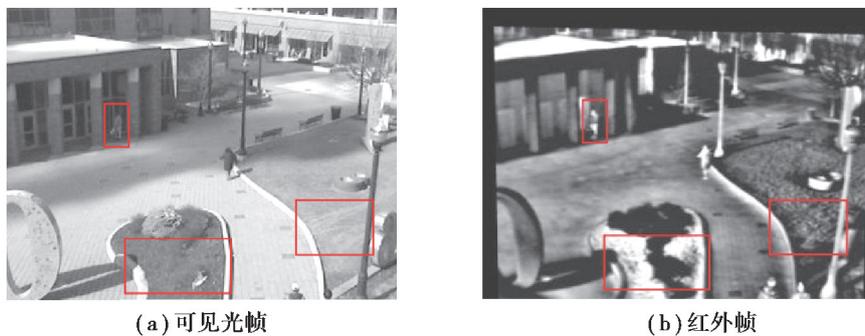


图 1 可见光和红外视频的帧

Fig. 1 Visible and infrared video frames

红外与可见光视频融合在民用、军事和航空等领域具有许多应用。在民用方面,红外与可见光视频融合可以在复杂的背景中清晰检测热物体,实现跟踪定位、火灾预警、隐藏目标的查找等。例如在火灾预警中,普通可见光摄像头并不能透过障碍物观测到其背后的火情,但是红外摄像头可通过观测物体的温度变化,在视频融合后对火灾进行预警。在军事领域,由于周围环境和烟雾弹等的影响,无法对敌方的部署和人员的行踪进行观察和追踪。红外和可见光视频的融合有助于克服这些问题,保证部队的安全并针对敌情做出相应的作战计划。在航空领域,由于天气的影响,例如大雨、大雾会对可见光的传播造成影响,使得拍摄的照片和视频清晰度大大降低,但是红外线的传播并不会收到很严重的影响,所以将红外和可见光视频进行融合可以使可见光视频中被遮挡物体显现出来。

然而,由于计算资源和存储资源的要求,在资源受限的移动设备上进行视频融合仍然是一个挑战^[8]。现有的视频融合方法大多数都是基于帧的融合。这些相关的研究工作,包括基于引导图像滤波器的融合^[9-11],

基于加权最小二乘滤波器的融合^[12],基于双边滤波器的融合^[13],Karhunen-Loève(KL)变换融合^[14],主成分分析融合(PCA, principal components analysis)^[15-16],Brovey 变换融合^[17],HIS 变换融合^[18],拉普拉斯金字塔融合^[19],小波变换融合^[20-22]等。但是,上述算法大多忽略了可见光帧和红外帧之间的差异,它们着重于尽可能多的保留源视频的信息,使红外帧中的纹理信息融合到最终结果中。实际上,红外帧中存在许多温度诱发的斑点,这些斑点与真实的纹理不同,并且它们不应融合到最终结果中。而且现有算法过多依赖于这些纹理斑点,这导致真实可见纹理被这些红外斑点破坏。此外,他们不考虑视频处理时的资源受限问题,而只关注视频的融合。

针对上述问题,笔者将视频融合任务迁移到微云端来执行,首先使用引导滤波器将所有可见光视频和红外视频的帧分解成基层和细层。考虑到红外帧细层包含了无用的纹理信息,所以将其丢弃以便提供更高质量的融合结果。然后将可见光帧和红外帧在微云上进行融合,以克服移动设备的资源受限问题从而得到高质量的融合视频帧。研究的主要贡献有以下几个:

首先,在视频传输时考虑到网络带宽的限制,提出了一种快速的帧间冗余检测算法,以减少传输到云端的帧的数量以及带宽和设备能量的消耗。

其次,为了获得更高质量的帧,提出了一种改进的自适应加权二维主成分分析(W2DPCA)的融合算法,以融合可见光帧和红外帧。此外,丢弃红外帧的细层以防止真实的纹理被无用的红外纹理信息破坏。

最后,实验结果表明,当错误率最低时,提出的冗余检测算法在视频传输过程中实现了 14.9% 的压缩比。而且,W2DPCA 融合算法与其他现有算法相比,可以实现更高质量的融合结果以及更低的全局标准偏差和更高的峰值信噪比。

1 系统模型和问题描述

基于微云的系统包括摄像头采集设备,移动设备(即智能手机,平板电脑,笔记本电脑等)和云端 3 个部分^[23-25],如图 2 所示。在系统模型中,摄像头将捕获的视频通过 Wi-Fi 传输到移动设备。移动设备可以通过移动网络(电信网络)或接入点 2 种方式将视频融合任务卸载到云端。对于第一种方式,移动设备通过 3G 或 4G 与基站(BS)连接。第二种方式,移动设备是通过 Wi-Fi 连接到接入点。基站和接入点都进一步连接到互联网,从而访问云端的可用资源。因此,通过将计算密集型任务外包给云端,可以提高移动设备的处理能力^[26]。

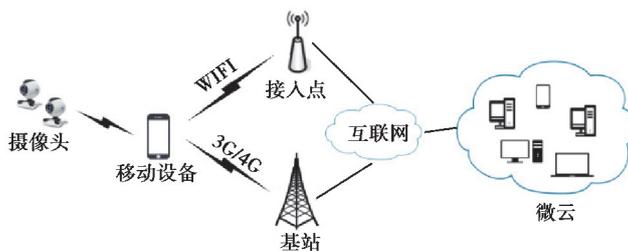


图 2 系统模型

Fig. 2 System model

在该系统中,移动设备不仅可以利用有限的本地资源,还可以访问丰富的云资源。因此,视频融合任务既可以在移动设备上执行也可以在云上执行。具体而言,一旦本地资源不足或不可用时,复杂的计算任务将被卸载到微云,而简单的计算任务将在本地设备中完成。由于微云上强大的计算能力,卸载的方法加速了视频处理并减少了延迟。

不同摄像头采集的可见光视频和红外视频通过 Wi-Fi 传输到移动设备,然后在资源受限的条件下解决可见光视频和红外视频的快速融合问题。考虑到视频传输过程中网络带宽的限制,笔者提出了一种计算复杂度较低的冗余检测算法,以减少帧的传输量,从而减少带宽占用和相应的传输延迟。然后提出了一种基于加权二维 PCA(W2DPCA)的视频融合算法,以获得更高质量的融合视频。

2 提出的算法

详细描述帧间冗余检测算法和视频融合算法。在检测到帧间冗余之后,非冗余视频帧被选择性地发送到云端,以便获得更高质量的视频。

2.1 帧间冗余检测算法

为了满足网络带宽和设备能量的限制,提出了在移动设备上的基于均值哈希的帧间冗余检测算法^[27],以减少传输的冗余帧的数量。冗余检测算法主要分为 3 个步骤。

首先,不考虑原始帧的像素比,将原始帧都缩小成 8×8 的灰度图像。目的是消除高频细节以及尺寸和比例带来的差异。所有像素的平均灰度值计算为

$$\bar{I} = \frac{\sum_{i=1}^8 \sum_{j=1}^8 I(i, j)}{64} \quad (1)$$

然后,将每个像素的灰度值与平均灰度值进行比较。如果该像素的灰度值大于或等于平均值,则表示为 1;否则,记为 0。这样,生成 64 位二进制值,即均值哈希。

最后,通过计算均值哈希值中不同位的个数来获得 2 帧之间的汉明距离。如果距离小于阈值 k ,则意味着 2 个帧近似重复;否则,说明 2 个帧是不同的。选择一个重复帧并将其发送到云端以进行融合。

通过提出的帧间冗余检测算法对图 3 所示的 2 个帧进行检测,先将原始视频帧的大小缩小到 8×8 并转换为灰度图像。通过将每个像素的灰度值与平均灰度值进行比较来获得均值哈希。然后可以获得距离 d 。

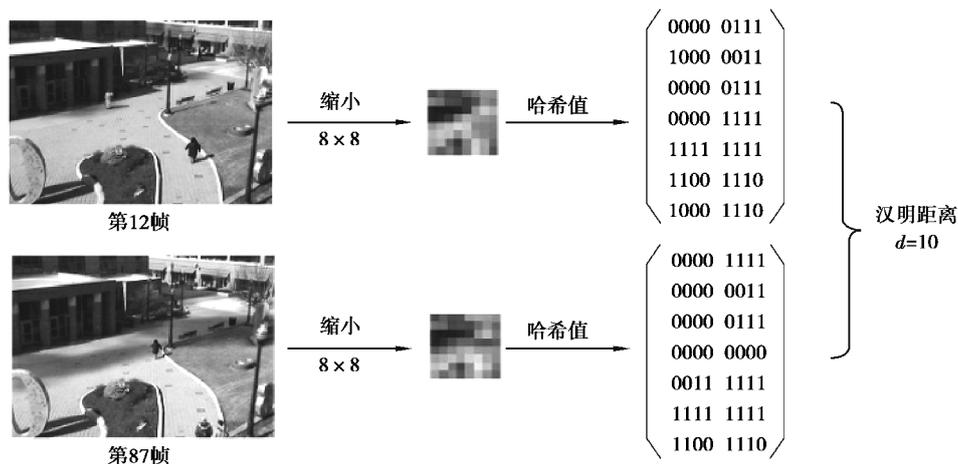


图 3 冗余检测算法的流程图

Fig. 3 The flow chart of redundancy detection algorithm

2.2 基于引导滤波器和 W2DPCA 的融合算法

为了使监控视频不受周围环境的影响,将可见光视频和红外视频进行融合以获得更高质量的结果。因此,提出计算复杂度较低的基于引导滤波器和 W2DPCA 的融合算法。算法中,可见光视频和红外视频进行逐帧融合。该融合算法主要分为 3 个步骤。首先使用引导滤波器对帧进行分层,得到每帧的基层和细层。然后,使用不同的规则分别融合基层和细层,以提取有效信息。最后,组合融合后的基层和细层以获得最终融合视频帧。

2.2.1 使用引导滤波器分解视频帧

可见光视频通常包含丰富的纹理,而这些纹理信息很容易在红外视频中丢失。为了确保红外纹理信息在融合期间不影响可见光纹理,首先使用引导滤波器来获得包含整体强度变化的基层和包含纹理信息的细层。对基层和细层采用不同的方法进行融合,使得融合视频帧不仅保留了可见光帧中有价值的纹理信息,而且还强调了红外帧中包含的热物体。

使用引导过滤器分解可见光帧和红外帧以获得它们的基层和细层。可见光帧和红外帧的基层分别表示为

$$I_B^v = G_{r,\epsilon}(I^v, I^v), \quad (2)$$

$$I_B^i = G_{r,\epsilon}(I^i, I^i), \quad (3)$$

其中: G 是引导滤波函数; I^v 和 I^i 分别表示源可见光帧和红外帧; r 是滤波器的半径; ϵ 是正则化参数。而 I_B^v 和 I_B^i 分别代表可见光帧和红外帧的基层。

然后,可以通过计算源始帧和其基层之间的差来获得细层。因此,可见光帧和红外帧的细层可分别表示为

$$I_D^v = I^v - I_B^v, \quad (4)$$

$$I_D^i = I^i - I_B^i. \quad (5)$$

通过这种方法,基层保留源始视频帧中具有较大方差的区域,细层包含原始帧的纹理信息,并且红外对象总是出现在基础层中。

2.2.2 基于W2DPCA的基层的融合

融合的主要目的是使可见光帧中因为受到环境影响而看不到的物体可以被更明显地看到。由于融合结果很容易受到帧质量的影响,对于高质量的可见光帧,只需要添加红外热对象。否则融合帧需要包含更多红外信息。基于上述考虑,提出了一种自适应W2DPCA算法。下面描述该算法的主要过程。

首先,在可见光帧和红外帧的基层上执行W2DPCA^[28]。矩阵 \mathbf{X} 是可见光帧和红外帧的集合。并且 \mathbf{X} 的协方差矩阵 \mathbf{C} 为

$$\mathbf{C} = \frac{1}{2} \sum_{i=1}^2 (\mathbf{X}^i - \mu)^T (\mathbf{X}^i - \mu), \quad (6)$$

其中: \mathbf{X}^1 和 \mathbf{X}^2 代表可见光帧和红外帧, μ 是2帧的平均值。

计算矩阵 \mathbf{C} 的特征向量,并根据特征值 λ 的大小以降序排序,使得最大特征值对应的特征向量位于矩阵的最左侧。前 d 个特征向量作为特征提取矩阵 \mathbf{V} 的,矩阵 \mathbf{V} 的大小是 $n \times d$ 。

$$\mathbf{V} = (\mathbf{V}_1, \mathbf{V}_2, \dots, \mathbf{V}_d), \quad (7)$$

然后,获得2个特征图像 Y_B^v 和 Y_B^i ,分别由式(8)、(9)表示。 $Y_{B1}^v, Y_{B2}^v, Y_{B3}^v, \dots, Y_{Bd}^v$ 和 $Y_{B1}^i, Y_{B2}^i, Y_{B3}^i, \dots, Y_{Bd}^i$ 分别表示可见光帧和红外帧的前 d 个主成分。

$$Y_B^v = I_B^v \times \mathbf{V} = [Y_{B1}^v, Y_{B2}^v, Y_{B3}^v, \dots, Y_{Bd}^v], \quad (8)$$

$$Y_B^i = I_B^i \times \mathbf{V} = [Y_{B1}^i, Y_{B2}^i, Y_{B3}^i, \dots, Y_{Bd}^i]. \quad (9)$$

随着像素大小的改变,相应帧的权重也会改变。由于特征图像的尺寸较小,因此从特征图像计算可见光帧和红外帧的权重可以减少计算量。

对于可见光帧,需要检测其边缘区域之间的对比度变化。在具有挑战性的环境中,可见光帧的质量可能较差,但它包含的纹理信息是有用。因此,在融合时基于帧质量可以获得自适应的可见光帧的权重,用式(10)计算基于区域方差的可见光帧权重 W^v 。

$$W^v = \sigma^2 = \frac{1}{m \times d} \sum_{i=1}^m \sum_{j=1}^d (Y_B^v(i, j) - \overline{Y_B^v})^2. \quad (10)$$

对于红外帧,较高的像素值表示较热的区域,而较低的像素值表示较冷的区域。在低温下,与高温相比,所有像素的灰度值都保持较低值。但是,热物体始终具有比其他像素更高的灰度值。在融合结果中,有必要强调红外热物体,最简单的方法是通过式(11)对红外帧执行零均值操作以获得红外权重 W^i 。因此,较高像素将被分配较大的权重。

$$W^i = \frac{1}{m \times d} \sum_{i=1}^m \sum_{j=1}^d (Y_B^i(i, j) - \overline{Y_B^i}). \quad (11)$$

然后,可以通过对可见光帧和红外帧的特征图像进行加权平均来获得融合特征图像 Y_B 。该算法重点保留可见光帧的信息,因此融合特征图计算如下

$$Y_B = \frac{W^v Y_B^v + W^i Y_B^i + Y_B^v}{W^v + W^i + 1} \quad (12)$$

最后,融合基层可以使用式(13)近似重构。

$$I_B = Y_B \times V^T, \quad (13)$$

2.2.3 细层的融合

由于红外帧的细节层含有无用的信息,可能会破坏真实的纹理信息,因此可以丢弃。融合细层由式(14)得到

$$I_D = I_D^v. \quad (14)$$

2.2.4 最终的融合视频帧

一旦获得了融合基层和融合细层,根据式(15)可以通过组合融合基层和细层获得最终融合结果,

$$I_F = I_B + I_D. \quad (15)$$

所得到的融合帧既包含了可见光帧丰富的纹理信息,也包含了红外帧中的热目标。

该融合算法通过图 4 进行描述。

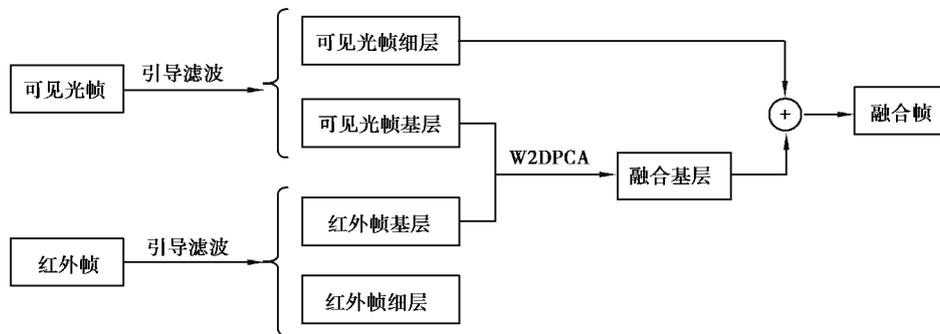


图 4 提出算法的流程图

Fig. 4 The flow chart of presented algorithm

3 性能评估

选择 3 组可见光视频和红外视频,用于对提出的 W2DPCA 融合算法的性能评估。第一组视频是 OSU 色热数据库,每帧的大小为 320×240 。视频帧的融合结果如图 5 所示。第二组红外和可见光视频在屋顶上拍摄的,图像大小为 360×270 。融合结果如图 6 所示。第三组可见光视频和红外视频是夜晚在街上拍摄的,每帧的大小为 632×496 。在图 7 中,可以看到第三组视频帧的融合结果。为了评估提出的冗余检测算法的性能,对视频执行帧间冗余检测,并对实验结果进行客观评估。融合算法与其他 4 种算法进行了比较,即主成分分析融合(PCA),离散小波变换融合(DWT, discrete wavelet transform)^[21],基于熵和灰度累积分布的融合算法(ECD, a fusion algorithm based on entropy and the cumulative distribution of gray levels)^[7]。

3.1 主观评价

从第一组视频中选择一对可见光帧和红外帧,分别如图 5(a)和图 5(b)所示。在图 5(a)中,隐藏在屋檐下的人不容易被看到,但在红外帧中具有较高的灰度值。此外,可以从图 5(b)中看到草地上的一些斑点,这些斑点是温度造成的,不应保留在融合结果中。PCA 和 DWT 的融合结果分别如图 5(c)和图 5(d)中所示,可以发现这些斑点被视为有效信息保留到融合帧中,因此草地上真实的可见光纹理被破坏。图 5(e)和 5(f)分别显示了 ECD 和提出的 W2DPCA 算法的融合结果。可以发现,不仅草地的可见光纹理被成功保存,而且红外热物体也被融合到结果中。与 ECD 相比,W2DPCA 算法保留了草地上更多的纹理信息,而 ECD 的结果更平滑导致了一些纹理信息的丢失。

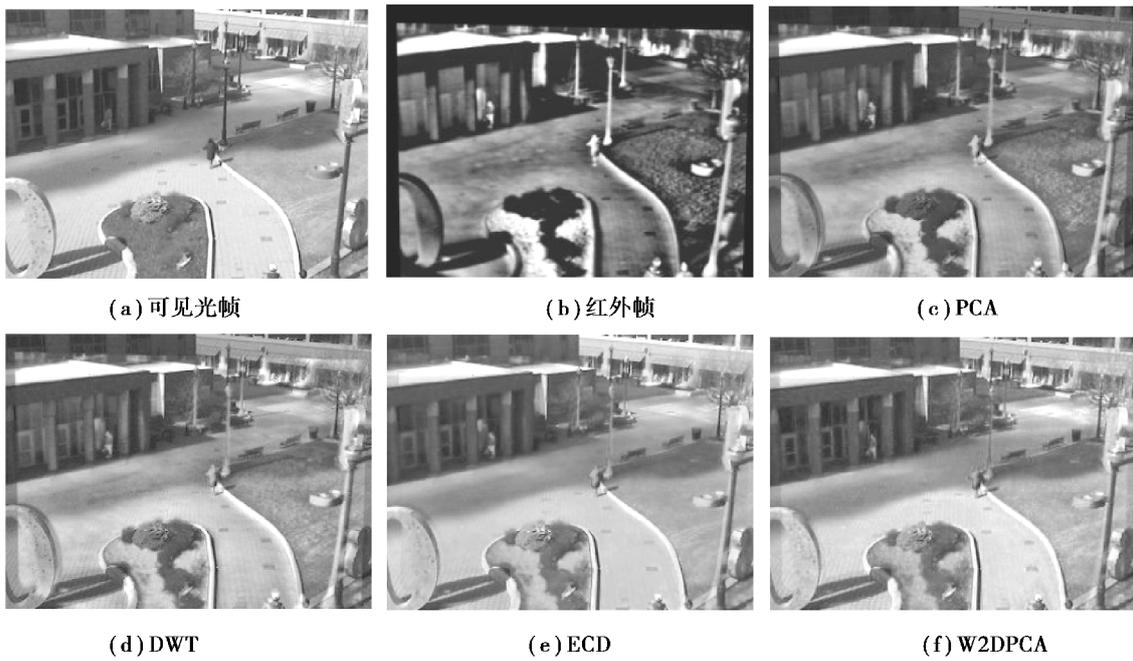


图 5 第一组源始帧和不同算法的融合结果

Fig. 5 The first set of source frames and fusion results of different algorithms

如图 6 展示了第二对视频帧和使用不同融合算法得到的结果。如图 6(a)所示,可见光帧包含屋顶和树木的纹理信息,但是被树木遮挡的人看不清楚。然而,如图 6(b)所示,由树木遮挡的热物体在红外帧中显示出了较高的灰度值。图 6(c)~图 6(f)分别显示了 PCA,DWT,ECD 和 W2DPCA 算法的融合结果。从这些图中可以看出,红外热物体的信息得到了很好的保存。由于添加了太多的红外信息,图 6(c)~图 6(e)所示的 PCA,DWT 和 ECD 的融合结果受到红外噪声的影响,这使得树木的纹理变得模糊。在图 6(f)中的融合结果中保留了可见光帧的更多信息,保留了清晰的纹理和被遮挡的人。

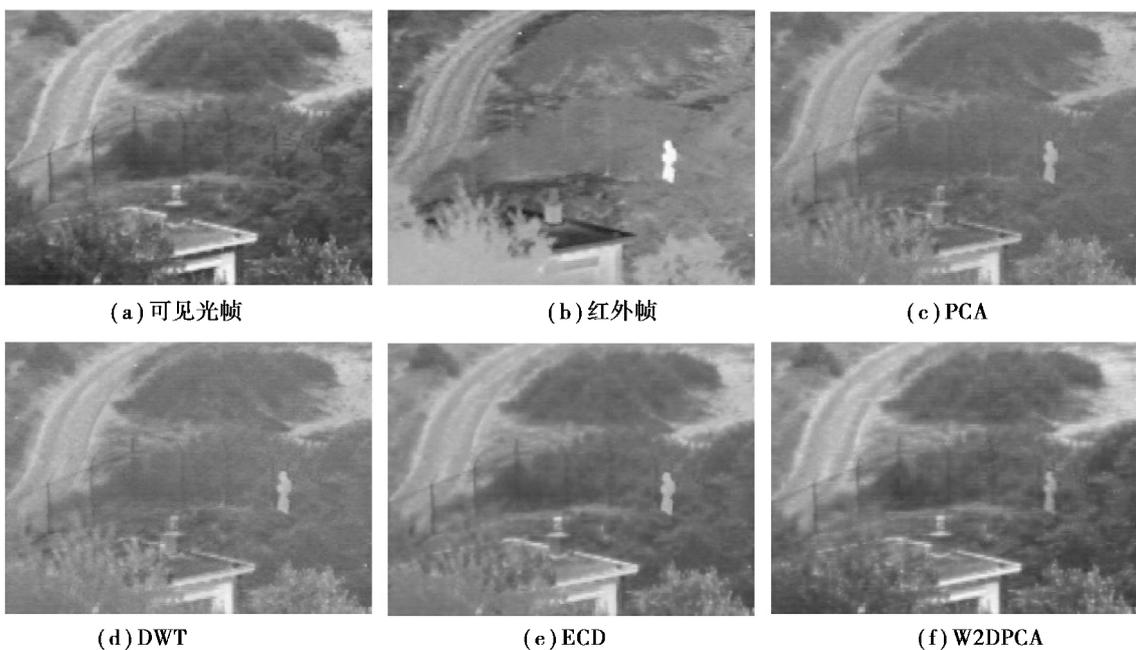


图 6 第二组源始帧和不同算法的融合结果

Fig. 6 The second set of source frames and fusion results of different algorithms

图 7 显示了第三组视频中的 2 对源始视频帧和他们的融合结果。由于缺乏照明条件,图 7(a)所示的可见光帧的质量也相对较差,而图 7(b)所示的红外帧的质量相对较好。PCA 和 DWT 的结果分别如图 7(c)~图 7(d)所示,融合结果更接近红外帧。ECD 和本文提出的 W2DPCA 的融合结果分别如图 7(e)~图 7(f)所示,不仅保留了红外帧中的热物体信息,而且更好地保留了可见光帧的纹理。与 ECD 相比,本文所提出算法的融合结果中汽车和路灯的轮廓更清晰。

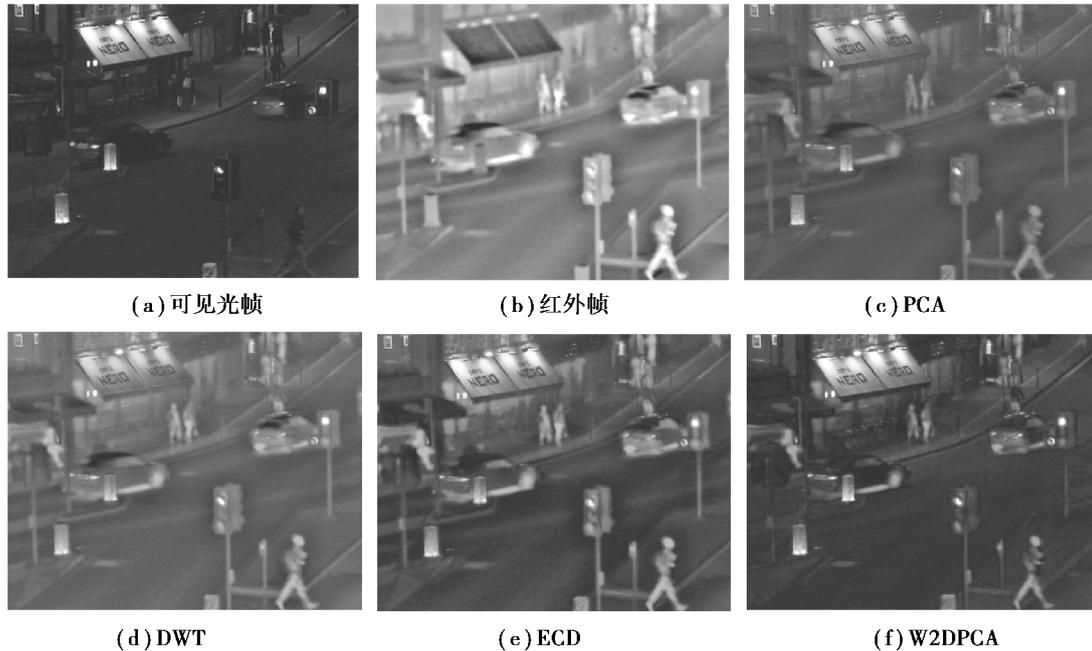


图 7 第三组源始帧和不同算法的融合结果

Fig. 7 The third set of source frames and fusion results of different algorithms

3.2 客观评价

为了客观地评估提出算法的性能,使用压缩比(CR)、错误率 α 和移动设备的能耗 E 作为帧间冗余检测算法的评价指标,给出了视频压缩比、冗余检测错误率和移动设备能耗的定义。

1) 视频的压缩比(CR)定义为帧的传输量与总帧数的比率,即

$$CR = \frac{R}{N}, \quad (16)$$

其中: R 是传输的视频帧的数量; N 表示视频帧的总数量。较低的 CR 值表示传输的视频帧数量较少,这减少了所需的存储和带宽资源以及传输期间的能量消耗。

2) 冗余检测的错误率 α 计算如下

$$\alpha = \frac{e}{N}, \quad (17)$$

其中 e 表示所有重复检测结果错误的帧和未检测到的重复帧的总数量。错误率 α 越低意味着冗余检测算法的性能越好。

3) 移动设备的传输能耗 E

$$E = P_t^c \times \frac{R \times D}{B_t}, \quad (18)$$

其中: P_t^c 是移动设备使用 3G/4G 时的传输功率; B_t 代表数据的平均传输速度; R 表示传输帧的个数; D 表示每一帧的大小。

实验中,为了表明该算法的通用性,采用了 3 组不同的视频数据进行去冗前后传输能耗的对比。其中,

传输功率 $P_i=150\text{ mW}$,传输速度 $B_i=8\text{ Mbps}$ 。由式(18)可以看出传输越少的帧数,移动设备能耗就越小。因此,在视频传输前先进行冗余检测,再将去冗后的视频帧传输到微云端,从而降低移动设备的能耗。

此外,利用以下4个客观质量指标来评估这些融合算法的性能,即全局标准偏差(SD),互信息(MI)^[29],基于结构相似性的度量($SSIM$)^[30]和峰值信噪比($PSNR$)。

1)融合图像 F 的全局标准偏差(SD)定义为

$$SD_F = \sqrt{\frac{\sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \bar{F})^2}{M \times N}}, \quad (19)$$

\bar{F} 是图像 F 的像素均值。一般来说,图像的标准差越大表示图像中包含的信息越丰富。

2)互信息(MI)定义如下

$$MI_{AF} = \sum_{i,j} P_{AF}(i,j) \log \frac{P_{AF}(i,j)}{P_A(i) P_F(j)}, \quad (20)$$

其中: A 和 F 分别表示源图像和融合结果; P_{AF} 是图像 A 和 F 的联合归一化直方图; P_A 和 P_F 分别是 A 和 F 的归一化直方图。 MI 可以测量融合图像与输入图像在灰度分布上的相似度,即源图像在融合图像中保留了信息的多少。

3)结构相似度指标($SSIM$)定义如下

$$SSIM_{AF} = \frac{(2\mu_A\mu_F + C_1)(2\sigma_{AF}^2 + C_2)}{(\mu_A^2 + \mu_F^2 + C_1)(\sigma_A^2 + \sigma_F^2 + C_2)}, \quad (21)$$

其中: μ 和 σ 分别表示源图像的均值和标准差, C_1 和 C_2 是用于避免不稳定性的2个参数,所有参数均采用默认值。 $SSIM$ 值越大表示相似性越大,说明融合图像的质量越好。

4)峰值信噪比($PSNR$)与均方误差(MSE)紧密相关,是用于衡量有效信息与噪音之间的比率,公式如下

$$PSNR = 10 \log \frac{255^2}{MSE}. \quad (22)$$

$PSNR$ 是一个正向指标,它的值越大,意味着融合算法的效果越好。反之, $PSNR$ 值越小,融合效果就越差。

图8给出了阈值 k 对压缩比和冗余检测错误率的影响。可以看出,当阈值 k 较大时,可以检测到更多重复的帧并且压缩比变得更大,但同时存在更高的错误率。当 $k < 2$ 时,有许多冗余帧未被检测到,导致相对较高的错误率。当 $k > 2$ 时,错误率随着 k 的增加而增加。可以看出,在 $k=2$ 时冗余检测的错误率最低。因此,为了保证错误率最低,设定阈值 $k=2$,即当2帧之间的距离小于或等于2时,认为2帧近似重复。

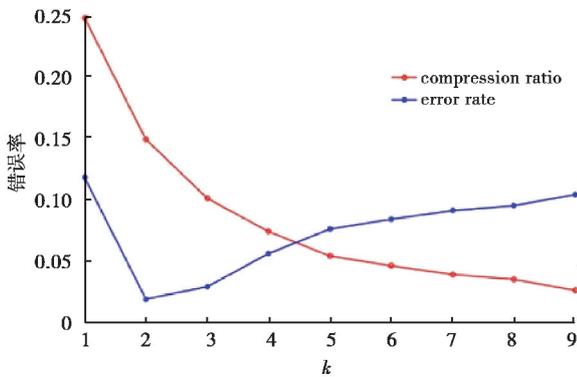


图8 阈值 k 对压缩比的影响

Fig. 8 The influence of the threshold k on the compression ratio

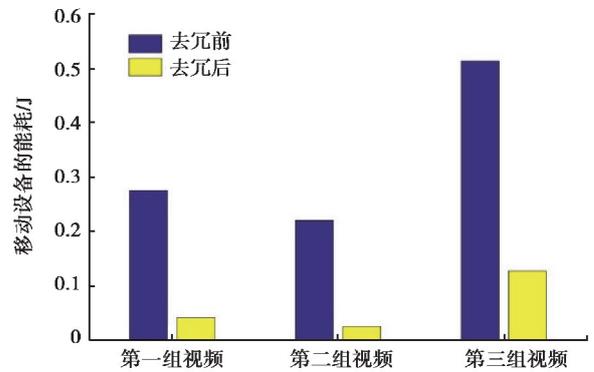


图9 不同视频组去冗前后移动设备的传输能耗对比

Fig. 9 Comparison of transmission energy consumption of mobile devices before and after redundancy removal in different video groups

根据图 9,可以得到以下 2 个结论。1) 3 组视频进行去冗操作之后,可以分别节约了 85.09%、88.64%和 75.24%的传输能耗,这表示冗余检测和视频去冗可以极大地减少移动设备的传输能耗。2) 相比其他视频,第二组视频节约了更多能耗。这是由于该组视频场景变化较少,冗余程度较高,去冗效果更加明显,即传输视频帧数量更少,从而达到了节能的目的。表 1 显示了对于 3 组不同视频帧在使用不同融合算法时得到的实验结果对应的指标的值,其中粗体数据表示最佳值。显然,对于 3 对视频帧,提出的融合算法比其他算法得到的融合帧具有更大的全局标准差。而且提出的算法更倾向于保留可见光帧的细节。由于红外帧包含很多不真实的纹理信息,因此只提取红外帧中最重要的热物体,所以融合结果包含来自红外帧的较少信息。

表 1 对不同的视频帧采用不同的融合算法得到的不同的融合结果的评估

Table 1 Evaluation of fusion results obtained by different fusion algorithms for different video frames

原始视频	评价指标	PCA	DWT	ECD	W2DPCA
第一组视频	SD	50.723 7	54.095 6	55.834 1	60.278 6
	MI_{VF}/MI_{IF}	0.905 1/1.065 6	1.044 5/0.787 5	1.834 5/0.604 5	1.728 0/0.617 3
	$SSIM_{VF}/SSIM_{IF}$	0.518 9/0.714 9	0.653 8/0.565 2	0.832 1/0.348 6	0.832 8/0.325 6
	PSNR	29.960 6	39.458 2	50.273 9	54.822 4
第二组视频	SD	23.186 9	25.951 6	30.347 3	36.405 4
	MI_{VF}/MI_{IF}	0.683 0/0.383 9	0.728 3/0.366 9	1.598 3/0.376 9	1.784 1/0.380 6
	$SSIM_{VF}/SSIM_{IF}$	0.783 8/0.733 4	0.767 8/0.702 9	0.875 7/0.536 0	0.898 4/0.482 0
	PSNR	46.675 8	45.007 0	52.446 3	66.882 7
第三组视频	SD	20.680 6	22.673 8	29.499 8	27.934 1
	MI_{VF}/MI_{IF}	0.400 6/0.928 9	0.322 5/1.089 7	0.508 1/0.991 5	0.938 2/0.454 3
	$SSIM_{VF}/SSIM_{IF}$	0.670 4/0.669 7	0.550 9/0.769 4	0.731 1/0.563 3	0.868 5/0.409 3
	PSNR	45.468 2	38.435 7	45.641 1	58.695 4

此外,如表 1 所示,提出算法的融合结果具有更高的全局标准偏差 SD,更好的 MI_{VF} 和 $SSIM_{VF}$ 值,这意味着原始可见光帧的更多有效信息被保留,并且可以根据视频质量自适应地保留更有效的信息。由于只提取红外帧中的热物体,避免了融合结果中的大量噪声,因此该算法得到的融合结果也具有更高的 PSNR。显然,提出的算法在大多数情况下产生最佳融合结果。

3.3 参数的影响

从 3 组视频中选择 500 对帧,以根据上述 4 个指标来评估不同参数对提出的融合算法性能的影响。图 10 描绘了滤波器窗口半径 r 对 4 个指标的影响。在这个评估中,设置参数 $\epsilon = 10^{-3}$ 。从图 10(a)可以看出当滤波器窗口半径 r 增加时, MI_{IF} 首先减小,即在 $r < 20$ 时,随着 r 的增加, MI_{IF} 的减小,这意味着融合结果中保留的红外帧信息也在减少,融合帧的质量也更差。然后在 $r > 20$ 时略有增加,但是变化不明显。此外,随着 r 的增加, MI_{VF} 、 $SSIM_{IF}$ 和 $SSIM_{VF}$ 的值基本不变。图 10(b)显示随着滤波器窗口半径 r 的增加,全局标准偏差(SD)逐渐减小,PSNR 不变。当 $r=3$ 时,所有指标都得到最大值。

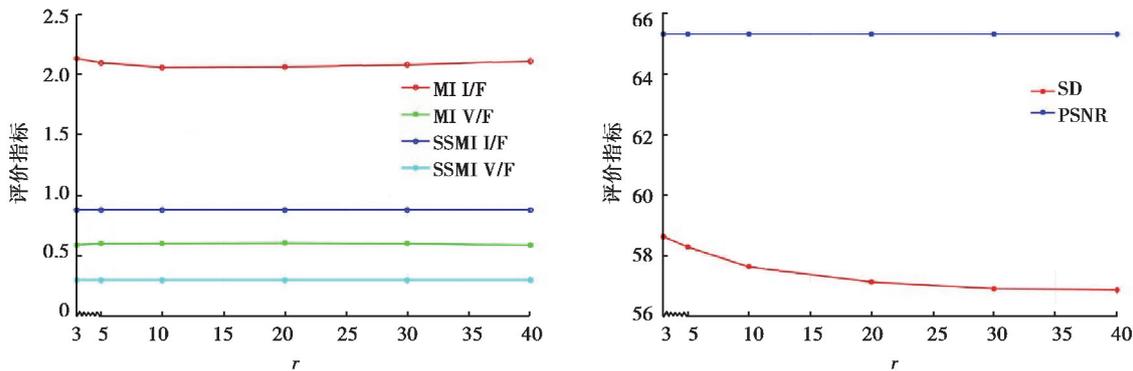


图 10 窗口半径 r 对算法的影响
 Fig. 10 The influence of window radius r on the algorithm

图 11 描绘了参数 ϵ 对 4 个指标的影响。在该评估中,设置窗口半径 $r=3$ 。从图 11(a)中可以发现 MI_{IF} 随着 ϵ 的增加先是保持不变,当 $\epsilon > 10^{-3}$ 时, MI_{IF} 逐渐减小。随着 ϵ 的增加, MI_{VF} 、 $SSIM_{IF}$ 和 $SSIM_{VF}$ 的值均保持不变。图 11(b)显示随着 ϵ 的增加, SD 保持不变, $PSNR$ 逐渐减小,最终趋于稳定。 $PSNR$ 的减少意味着融合帧的质量受到噪声的破坏,整体融合帧的质量在降低。在 $\epsilon = 10^{-3}$ 之前,所有指标都达到最大值。总的来说, r 和 ϵ 的变化对算法的性能影响不大,这意味着该算法相对稳定。选择参数 $r=3, \epsilon = 10^{-3}$ 。

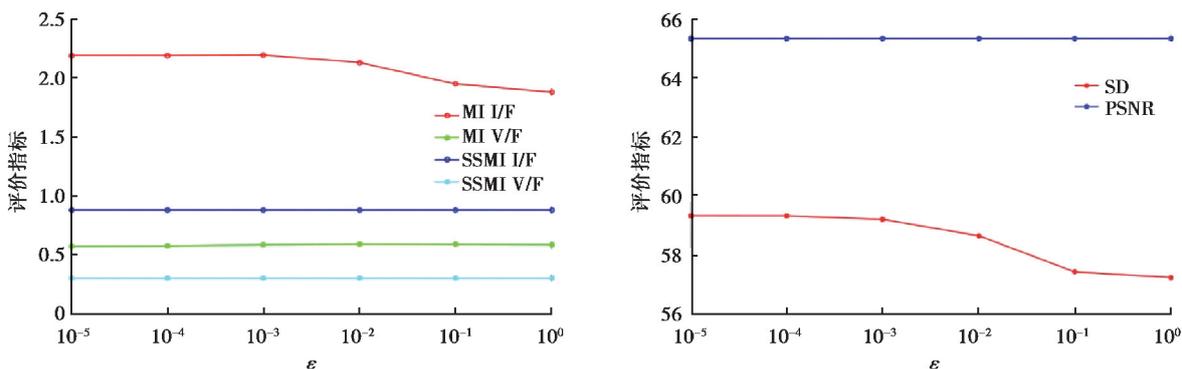


图 11 参数 ϵ 对算法的影响
 Fig. 11 The influence of parameter ϵ on the algorithm

4 总 结

由于移动设备的计算能力,存储和能量供应的限制,视频融合任务被卸载到强大的云端执行。为了减少传输过程中的带宽占用,研究提出了一种帧间冗余检测算法,通过计算帧之间的汉明距离,可以确定这些帧是否近似重复。然后,将非重复帧传输到资源丰富的微云进行融合。这样,大大减少了传输帧的数量和网络带宽的占用。此外,为了获得高质量的融合视频,提出基于引导滤波器和 W2DPCA 的视频融合算法。使用引导过滤器将捕获的可见光和红外视频帧分为基层和细层。采用本文提出的加权自适应 W2DPCA 算法进一步融合视频帧的基层。由于红外帧的细层仅包含一些无用的纹理信息,因此将其丢弃并仅保留可见光帧的细层作为融合基层。然后通过组合融合基层和细层来获的融合帧。实验结果表明,当错误率最低时,提出的冗余检测算法在视频传输过程中实现了 14.9% 的压缩比。与现有的其他算法相比,视频融合算法得到的融合结果更好的保留了可见光帧的真实纹理信息,具有更好的视觉效果。可见光帧与红外帧的融合结果具有与原始帧更多的互信息和结构相似度,同时也具有较好的整体标准差和峰值信噪比。

参考文献:

- [1] Liu Y, Chen X, Wang Z F, et al. Deep learning for pixel-level image fusion: Recent advances and future prospects[J]. Information Fusion, 2018, 42: 158-173.
- [2] Liu Y C, Lee M J, Zheng Y Y. Adaptive multi-resource allocation for cloudlet-based mobile cloud computing system [J]. IEEE Transactions on Mobile Computing, 2016, 15(10): 2398-2410.
- [3] Li S T, Kang X D, Fang L Y, et al. Pixel-level image fusion: a survey of the state of the art[J]. Information Fusion, 2017, 33: 100-112.
- [4] Zhang Q, Liu Y, Blum R S, et al. Sparse representation based multi-sensor image fusion for multi-focus and multi-modality images: a review[J]. Information Fusion, 2018, 40: 57-75.
- [5] Liu Z, Blasch E, John V. Statistical comparison of image fusion algorithms: recommendations[J]. Information Fusion, 2017, 36: 251-260.
- [6] Denman S, Lamb T, Fookes C, et al. Multi-spectral fusion for surveillance systems[J]. Computers & Electrical Engineering, 2010, 36(4): 643-663.
- [7] Hu H M, Wu J W, Li B, et al. An adaptive fusion algorithm for visible and infrared videos based on entropy and the cumulative distribution of gray levels[J]. IEEE Transactions on Multimedia, 2017, 19(12): 2706-2719.
- [8] Zhang H L, Zhang Q, Du X J. Toward vehicle-assisted cloud computing for smartphones[J]. IEEE Transactions on Vehicular Technology, 2015, 64(12): 5610-5618.
- [9] Kong H, Audibert J Y, Ponce J. General road detection from a single image[C]//IEEE Transactions on Image Processing a Publication of the IEEE Signal Processing Society, US: IEEE, 2010: 2211.
- [10] Na Y, Zhao L, Yang Y X, et al. Guided filter-based images fusion algorithm for CT and MRI medical images[J]. IET Image Processing, 2018, 12(1): 138-148.
- [11] Yang Y, Wan W G, Huang S Y, et al. Remote sensing image fusion based on adaptive IHS and multiscale guided filter[J]. IEEE Access, 2016, 4: 4573-4582.
- [12] Jiang Y, Wang M H. Image fusion using multiscale edge-preserving decomposition based on weighted least squares filter[J]. IET Image Processing, 2014, 8(3): 183-190.
- [13] Hu J W, Li S T. The multiscale directional bilateral filter and its application to multisensor image fusion[J]. Information Fusion, 2012, 13(3): 196-206.
- [14] Bavirisetti D P, Dhuli R. Fusion of infrared and visible sensor images based on anisotropic diffusion and karhunen-loevtransform[J]. IEEE Sensors Journal, 2016, 16(1): 203-209.
- [15] Shahdoosti H R, Ghassemian H. Combining the spectral PCA and spatial PCA fusion methods by an optimal filter[J]. Information Fusion, 2016, 27: 150-160.
- [16] Patil U, Mudengudi U. Image fusion using hierarchical pca [C] // International Conference on Image Information Processing, Shimla, India: IEEE, 2011: 1-6.
- [17] Mandhare R A, Upadhyay P, Gupta S. Pixel-level image fusion using broveytransforme and wavelet transform[J]. International Journal of Advanced Research in Electrical Electronics and Instrumentation Engineering, 2013, 2(6): 2690-2695.
- [18] He C T, Liu Q X, Li H L, et al. Multimodal medical image fusion based on IHS and PCA[J]. Procedia Engineering, 2010, 7: 280-285.
- [19] Du J, Li W S, Xiao B, et al. Union Laplacian pyramid with multiple features for medical image fusion [J]. Neurocomputing, 2016, 194: 326-339.
- [20] Wan T, Canagarajah N, Achim A. Segmentation-driven image fusion based on alpha-stable modeling of wavelet coefficients[J]. IEEE Transactions on Multimedia, 2009, 11(4): 624-633.
- [21] Desale R P, Verma S V. Study and analysis of pca, dct and dwt based image fusion techniques[C]// International Conference on Signal Processing Image Processing and Pattern Recognition, 2013, Coimbatore, India: IEEE, 2013: 66-69.

- [22] Gambhir D, Manchanda M. A novel fusion rule for medical image fusion in complex wavelet transform domain[J]. International Journal of Image and Graphics, 2016, 16(4): 16-22.
- [23] Ma Y, Chen J, Chen C, et al. Infrared and visible image fusion using total variation model[J]. Neurocomputing, 2016, 202: 12-19.
- [24] Shahdoosti H R, Ghassemian H. Spatial pca as a new method for image fusion[C]// The 16th CSI International Symposium on Artificial Intelligence and Signal Processing (AISP 2012). Shiraz, Fars, Iran: IEEE, 2012: 90-94.
- [25] Felemban M, Basalamah S, Ghafoor A. A distributed cloud architecture for mobile multimedia services[J]. IEEE Network, 2013, 27(5): 20-27.
- [26] Guo S, Xiao B, Yang Y, et al. Energy-efficient dynamic offloading and resource scheduling in mobile cloud computing[C]// The 35th Annual IEEE International Conference on Computer Communications, 2016. San Francisco, CA, USA: IEEE, 2016: 1-9.
- [27] ChaerulHaviana S F, Taufik M. Comparison of various similarity measures for average image hash in mobile phone application[J]. IOP Conference Series: Materials Science and Engineering, 2017, 190: 12-13.
- [28] Lin Y, Liang Y. A human face recognition method by improved modular 2dpca[C]// International Symposium on It in Medicine and Education, 2011. Guangzhou, China: IEEE, 2012: 7-11.
- [29] Qu G, Zhang D, Yan P. Information measure for performance of image fusion[J]. Electronics Letters, 2002, 38(7): 313-315.
- [30] Wang Z, Bovik A C, Sheikh H R. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.

(编辑 侯湘)