

DOI:10.19651/j.cnki.emt.2107453

# 一种基于差异特征驱动的红外与可见光视频 拟态融合方法\*

李向燕 王肖霞 杨风暴

(中北大学 信息与通信工程学院 太原 030051)

**摘要:** 针对固定融合算法无法适应动态场景下视频帧间差异特征变化造成融合效果差的问题,提出了一种基于差异特征驱动的红外与可见光视频拟态融合方法。首先,分别提取了红外与可见光视频序列的3种差异特征;其次,利用改进的融合有效度公式计算不同融合算法对3种差异特征的融合有效度;最后,利用熵权法对融合有效度进行加权合成,进而得到多融合算法的决策评分,确定视频不同序列段上的最优融合算法。实验结果表明,所提出方法相较于在不同序列段选取的最优融合算法RP和MOD,在整段视频序列上融合效果更好,综合客观评价指标上比上述固定算法分别提升了59.925%、2.7608%,为红外与可见光视频融合提供了新思路。

**关键词:** 视频拟态融合;差异特征;熵权法;融合有效度

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 510.4050

## Fusion method of infrared and visible video mimicry based on difference feature driving

Li Xiangyan Wang Xiaoxia Yang Fengbao

(College of Information and Communication Engineering, North University of China, Taiyuan 030051, China)

**Abstract:** Aiming at the problem that the fixed fusion algorithm can not adapt to the difference between video frames in dynamic scene, which results in poor fusion effect and even failure of fusion, a fusion method of infrared and visible video mimicry based on differential feature driver is proposed. Firstly, three different features of infrared and visible video sequences are extracted respectively. Secondly, the fusion validity of different fusion algorithms is calculated by using the improved fusion validity formula, the fusion validity is weighted by entropy weight method, and then the decision score of the fusion algorithm is obtained, and the optimal fusion algorithm on different video sequences is determined. The experimental results show that the proposed method is better than the optimal fusion algorithms RP and MOD in the whole video sequence, compared with the fixed algorithm mentioned above, the comprehensive objective evaluation index is improved by 59.925% and 2.7608%, which provides a new idea for infrared and visible light video fusion.

**Keywords:** video mimicry fusion; difference feature; entropy weight method; fusion validity

## 0 引 言

红外与可见光视频在成像特点上有很大区别。可见光视频对比度和分辨率比较高,细节信息丰富,可视性好,但易受天气烟雾等遮挡影响,抗干扰能力差;而红外视频表征目标与场景的温度信息,可全天候工作,抗干扰能力强,但清晰度低,可视性差。因此,将红外与可见光视频进行融

合,可以充分利用其信息互补性,扩展系统探测的时空覆盖范围、空间分辨率、全天候工作能力以及后续目标检测和抗干扰能力,在航空、遥感、国防、医疗等诸多领域有重要应用<sup>[1]</sup>。

目前,红外与可见光视频融合主要分为两类:一类是静态逐帧视频融合方法,一般基于多尺度分解<sup>[2]</sup>、稀疏表示<sup>[3]</sup>、神经网络<sup>[4]</sup>、深度学习<sup>[5]</sup>或其他混合融合方法;另一

收稿日期:2021-07-31

\* 基金项目:国家自然科学基金项目(61672472)、山西省应用基础研究计划青年科技研究基金项目(201901D211233)、山西省回国留学人员科研项目(2021-110)、山西省研究生创新项目(2021Y612)资助

类是三维多尺度视频融合方法,把视频看作整体三维信号,采用时空能量匹配<sup>[6]</sup>、时空结构张量<sup>[7]</sup>、高阶奇异值分解<sup>[8]</sup>等手段进行融合。这两类视频融合方法都是依赖先验知识事先确定一种固定的融合算法,但在实际成像中,视频场景动态复杂多变,固定单一的融合算法不能保证在整个视频序列上始终都有较好的融合效果(见本文2.3节图7),因此需要根据场景变化动态调整选择合适的融合算法。

拟态融合是近年来一种比较新颖的仿生融合方法<sup>[9]</sup>,是指通过改变融合模型结构满足不同变化需求的一种融合新理论。文献[10]通过分析差异特征多个属性的差异变化,进而选择合适的融合算法自适应满足融合需求,取得了较好的融合效果。文献[11]研究了差异特征与融合算法的集值映射关系,为拟态变换的实现提供了依据。上述研究为实现动态场景下融合算法随差异特征自适应变化提供了新思路。因此,本文提出基于差异特征驱动的红外与可见光视频拟态融合方法,在提取视频帧间差异特征基础上利用改进的融合有效度公式,求得不同算法对3种差异特征的融合有效度,然后利用熵权法对融合有效度进行合成,进而得到多融合算法的决策评分,确定不同序列段上最优融合算法,实现了红外与可见光视频拟态融合。本文的主要创新工作如下:1)对基于余弦相似性测度的融合有效度公式进行了改进,并验证其有效性;2)提出了基于熵权法的最优融合算法选取方法;3)提出了一种视频拟态融合方法,为融合算法能随动态场景变化进行自适应变换提供了新思路。

## 1 本文的视频拟态融合方法

### 1.1 视频序列帧间差异特征的提取

红外与可见光传感器在成像机理上的不同决定了红外与可见光视频序列在空间分辨率、灰度差异以及边缘、纹理、像素相关性等特征方面都有很强互补性的特点。本文提取了亮度、边缘、纹理3种主要差异特征,并以灰度均值 Mean、边缘强度 ES、空间频率 SF 3个指标参数描述这3种差异特征<sup>[12]</sup>,它们的定义分别如下。

1)灰度均值定义为:

$$\text{Mean} = \sum_{k=0}^{L-1} z_k p(z_k) \quad (1)$$

$$p(z_k) = \frac{n_k}{MN} \quad (2)$$

其中,  $n_k$  是灰度  $z_k$  在图像中出现的次数,  $MN$  是像素总数,  $p(z_k)$  为灰度级  $z_k$  出现的概率。

2)边缘强度定义为:

$$ES(i, j) = \sqrt{\nabla_x f(i, j)^2 + \nabla_y f(i, j)^2} \quad (3)$$

$$\nabla_x f(i, j) = f(i, j) - f(i-1, j)$$

$$\nabla_y f(i, j) = f(i, j) - f(i, j-1) \quad (4)$$

3)空间频率定义为:

$$SF(i, j) = \sqrt{(RF)^2 + (CF)^2} \quad (5)$$

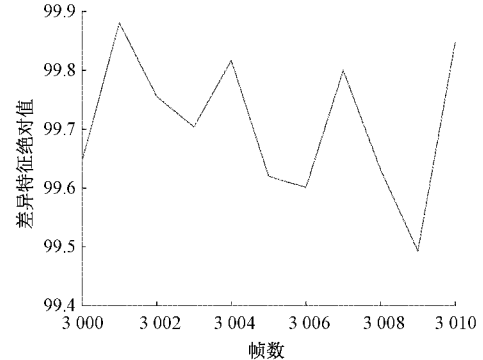
$$RF = \sqrt{\frac{1}{M \times N} \sum_{m=1}^M \sum_{n=2}^N [F(m, n) - F(m, n-1)]^2} \quad (6)$$

$$CF = \sqrt{\frac{1}{M \times N} \sum_{m=2}^M \sum_{n=1}^N [F(m, n) - F(m-1, n)]^2} \quad (7)$$

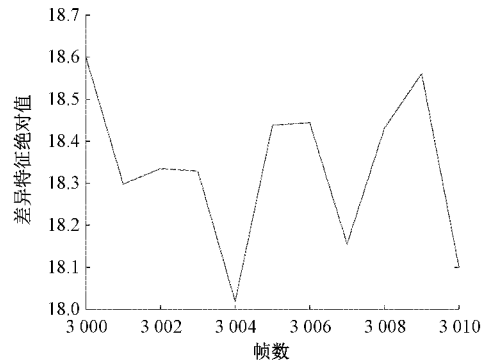
SF度量融合图像空间域上的整体活跃程度,可以反映图像对微小细节反差的描述能力, SF指标越大,融合图像越清晰。两类视频差异特征差的绝对值根据如下公式计算:

$$\Delta T_m(i) = |T_m V(i) - T_m I(i)| \quad (8)$$

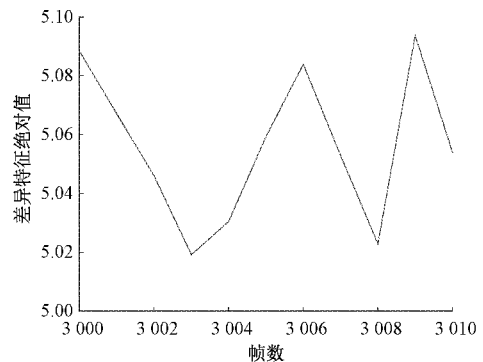
$i$ 表示视频序列第 $i$ 帧,  $T_m V(i)$ 和 $T_m I(i)$ 分别为可见光和红外视频序列第 $i$ 帧 $m$ 指标参数特征值,其中 $m = \{\text{Mean}, \text{ES}, \text{SF}\}$ ,  $\Delta T_m(i)$ 为可见光与红外视频序列第 $i$ 帧 $m$ 指标参数特征值差的绝对值。如图1所示为对测试数据集 Tree\_4 序列2提取3种差异特征的曲线分布结果。



(a) 亮度差异



(b) 边缘差异



(c) 纹理差异

图1 差异特征曲线分布

### 1.2 融合有效度公式的改进

融合有效度<sup>[13]</sup>是指在一定融合算法下,融合后图像特征对融合前两类图像特征的有效融合程度,表征了差异特征与融合算法之间的映射关系。将融合结果特征与源视频序列对应特征的接近程度归为距离或相似性测度进而构造融合有效度评价函数,3种常见的基于余弦相似性、兰氏距离、欧氏距离测度的融合有效度公式分别如下<sup>[14]</sup>:

$$\theta_1 = \frac{T_F \max(T_v, T_I)}{\sqrt{T_F^2 + \max(T_v, T_I)^2}} \quad (9)$$

$$\theta_2 = \frac{|T_F - \max(T_v, T_I)|}{T_F + \max(T_v, T_I)} \quad (10)$$

$$\theta_3 = |T_F - \max(T_v, T_I)| \quad (11)$$

其中,  $\theta$  表示融合有效度;  $T_F$  表示融合结果的差异特征表征;  $T_v, T_I$  分别为可见光和红外视频差异特征表征。在3种融合有效度函数中,余弦相似性注重方向上的差异,具有很好的空间旋转不变性,而且针对随机信息,是一种比较典型的相似度测量方式,但是余弦相似度只能分辨个体在低维上的差异,无法衡量多个维度数值的差异,因此,本文对余弦相似性公式做一些调整如式(12)所示。

$$\theta = \frac{(T_F - \bar{T}_v)(T_F - \bar{T}_I)}{\sqrt{(T_F - \bar{T}_v)^2} \sqrt{(T_F - \bar{T}_I)^2}} \quad (12)$$

其中,  $\theta$  表示融合有效度;  $T_F$  表示视频序列某帧某种算法融合结果的差异特征表征;  $T_v$  和  $T_I$  分别为可见光和红外视频差异特征表征;  $\bar{T}_v$  和  $\bar{T}_I$  分别表示源视频序列可见光与红外差异特征均值。其有效性验证见实验 2.2 节。

### 1.3 不同序列段最优融合算法的选取

首先,利用式(12)计算多融合算法对3种差异特征的融合有效度构成矩阵  $\mathbf{X}$ 。

然后,利用熵权法对不同算法对3种差异特征的融合有效度进行加权合成。熵权法<sup>[15]</sup>是一种客观赋权方法,当数据差异度越大,说明所含的信息量就越大,权重也就越高。本文基于熵权法对3种差异特征的加权合成计算步骤如下。

1)首先对输入的融合有效度矩阵  $\mathbf{X}$  进行正向化,然后对构成的正向化矩阵进行标准化得到矩阵  $\mathbf{Z}$ ,  $\mathbf{Z}$  中元素为正,如果  $\mathbf{Z}$  中存在负数,需要对矩阵  $\mathbf{Z}$  使用标准化方法:

$$z_{ij}^{\sim} = \frac{x_{ij} - \min\{x_{1j}, x_{2j}, \dots, x_{mj}\}}{\max\{x_{1j}, x_{2j}, \dots, x_{mj}\} - \min\{x_{1j}, x_{2j}, \dots, x_{mj}\}} \quad (13)$$

2)计算第  $j$  项差异特征下第  $i$  种融合算法的比重,将其看作相对熵计算概率:

$$p_{ij} = \frac{z_{ij}^{\sim}}{\sum_{i=1}^n z_{ij}^{\sim}} \quad (14)$$

3)计算每项差异特征的信息熵:

$$e_j = -\frac{1}{\ln n} \sum_{i=1}^n p_{ij} \ln(p_{ij}) \quad (15)$$

当  $e_j$  越大,第  $j$  项差异特征的信息熵越大,那么第  $j$  项

差异特征所包含的信息越少,如果所有的差异特征融合有效度值都相同,信息效用值最小。计算信息效用值为:

$$d_j = 1 - e_j \quad (16)$$

将信息效用值进行归一化,就能得到每种差异特征的权重:

$$W_j = d_j / \sum_{j=1}^m d_j \quad (17)$$

最后,计算每一种融合算法对3种差异特征融合有效度的决策评分:

$$P_i = \sum_{j=1}^m W_j z_{ij}^{\sim} \quad (18)$$

根据多融合算法在不同序列段上的排序结果选择最优融合算法。

## 2 实验仿真及结果分析

### 2.1 数据集及实验平台

本文用到的视频数据集有 Tree\_sequence、Nato\_camp\_sequence、INO\_ParkingEvening 以及 VLIRVDIF 数据库中的 Tree\_4 数据集,部分源视频序列如图 2~4 所示。图 4 所示为 Tree\_4 数据集源视频中两段不同场景的部分序列,左边 3 幅为序列 1,右边 3 幅为序列 2。融合算法选取了 5 种典型的低通比率金字塔 (ratio of low-pass pyramid, RP)、稳态小波变换 (steady-state wavelet transform, SWT)、形态学金字塔变换 (morphological difference pyramid, MOD)、拉普拉斯金字塔变换 (laplacian pyramid, LP)<sup>[16]</sup>、双树复小波变换 (dual-tree complex wavelet transform, DTCWT)<sup>[17]</sup> 融合源视频序列,通过本文所提方法选取不同序列上的最优融合算法,然后从主客观两个方面与固定单一的算法进行了对比分析。本节实验所使用的实验平台为 DELL Intel(R)Core(TM)i5-5200U 服务器(八核处理器主频 2.9 GHz),运行软件为 MATLAB 2014b。

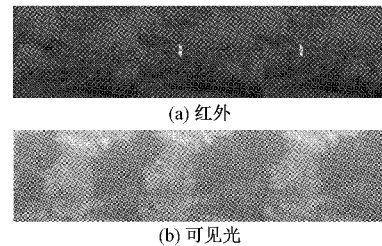


图 2 Tree\_sequence 数据集部分序列

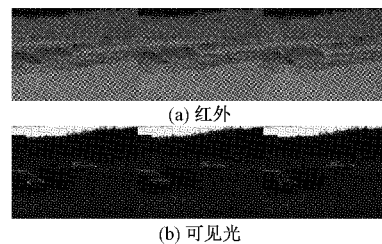


图 3 Nato\_camp\_sequence 数据集部分序列

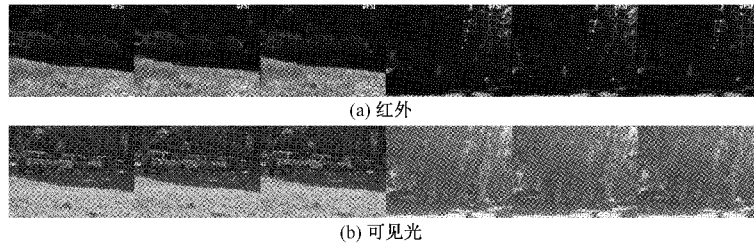


图4 Trcc\_4数据集部分序列

### 2.2 改进的融合有效度的有效性验证

为了验证本文提出的改进余弦相似性测度的融合有效度公式的有效性,Tree\_sequence、Nato\_camp\_sequence、INO\_ParkingEvening 多组数据集上进行了验证,这里仅给出 Tree\_sequence 数据集 1~10 帧的结果如图 5 所示。如图 6 所示为 3 种差异特征亮度、边缘、纹理对应的灰度均值

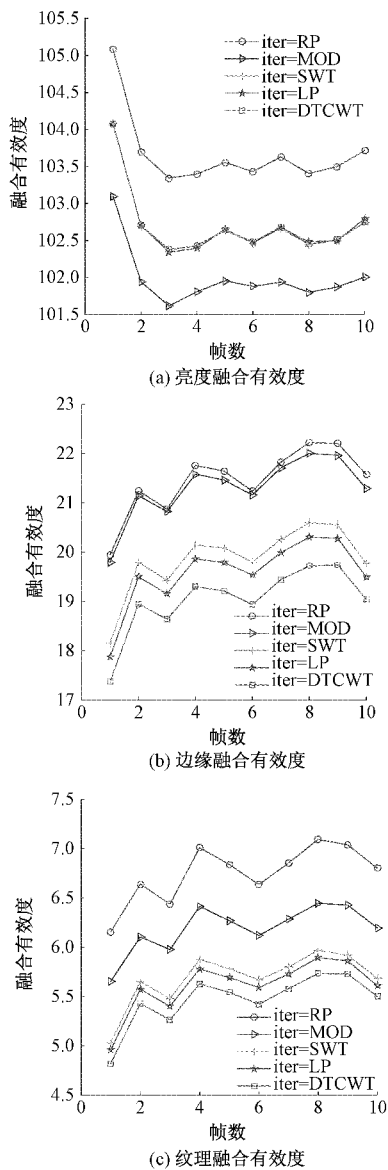


图5 差异特征融合有效度分布

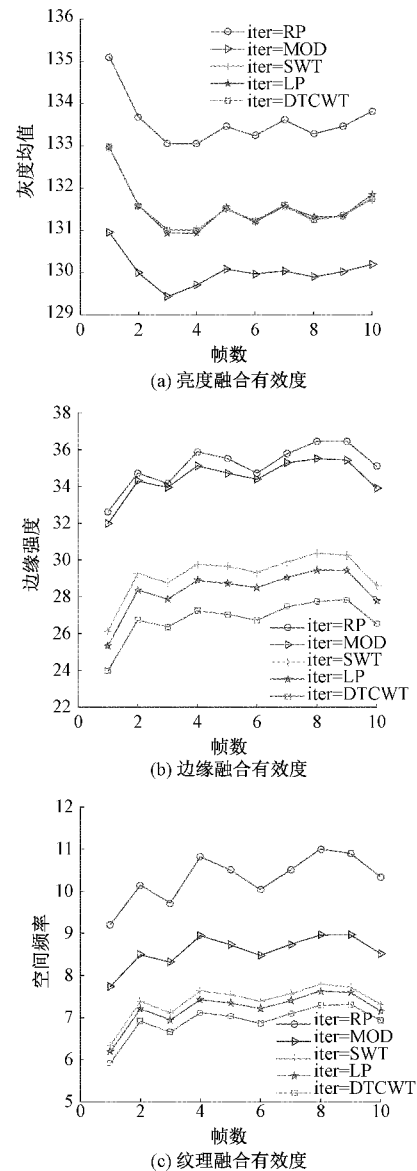


图6 各算法融合结果指标分布

值、边缘强度、空间频率分布曲线。可以看到,每种差异特征下多融合算法融合有效度分布和对应的客观指标分布曲线趋势和排序结果一致,验证了本文 1.2 节改进的余弦相似性测度的融合有效度公式作为融合有效度是有效的。



### 2.3 视频拟态融合结果的对比分析

#### 1) 多融合算法决策评分结果

本文选用的 VLIRVDIF 数据库中的 Tree\_4 数据集其中包含多种不同场景,包括道路、车辆、树林和人物目标等,满足动态场景的测试需求。在 Tree\_4 数据集上选取了

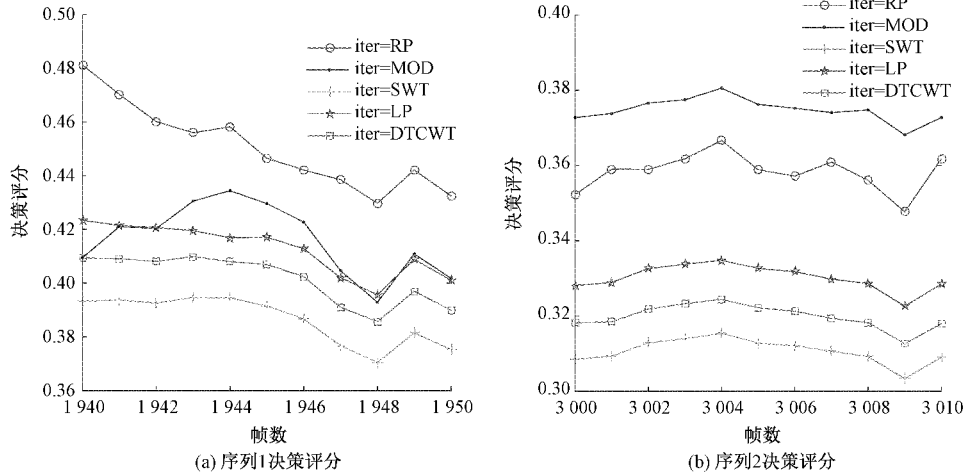


图 7 多融合算法决策评分

#### 2) 融合结果主观分析比较

如图 8 所示是 5 种固定单一融合算法和本文视频拟态融合方法对红外可见光源视频序列的融合结果部分序列,左边 3 列为序列 1 融合结果,右边 3 列为序列 2 融合结果,从上到下依次为 DTCWT、LP、SWT、MOD、RP 以及本文视频拟态融合方法融合结果。从主观视觉上,5 种固定融合算法都能在一定程度上融合源视频序列所提供信息的互补性,但仔细观察可知,当场景从序列 1 变化为序列 2

两段场景差异较大的序列,分别是第 1 940~1 950 帧(序列 1)和第 3 000~3 010 帧(序列 2)进行本文视频拟态融合方法的实验仿真。如图 7 所示为 5 种融合算法在序列 1 和序列 2 上的决策评分排序结果。由此可知,序列 1 应该选择 RP 算法,序列 2 应该选择 MOD 算法进行融合。

时,同一种算法的融合效果也会发生很大变化,比如原来在序列 1 上融合车辆细节纹理效果较好的 RP 算法在序列 2 上对树林和人物目标的融合效果不如 MOD 算法,这是因为场景变化导致差异特征随之变化,而每种融合算法对差异特征的融合优势各不相同。最后一行本文视频拟态融合方法能更好地适应这种场景变化,使不同算法在不同场景下发挥各自的优势,在整体上达到融合效果最优的目的。

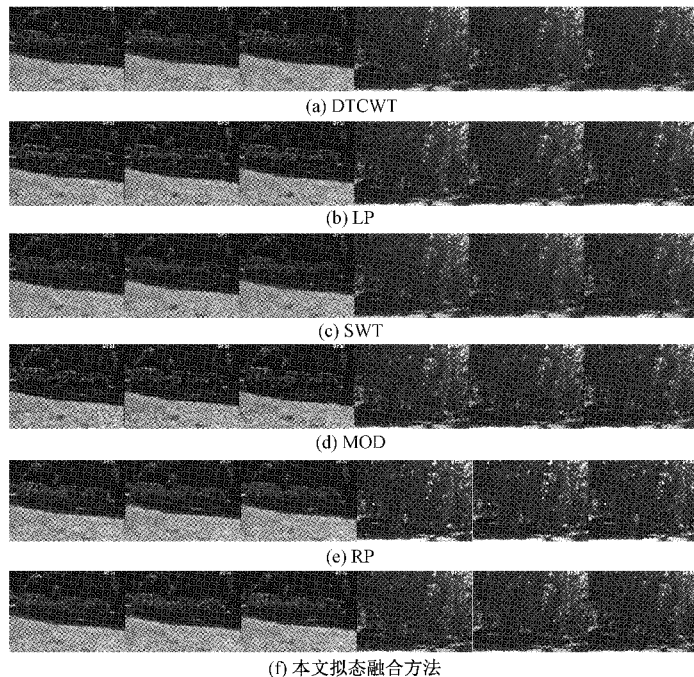


图 8 各算法融合结果部分序列

### 3) 融合结果客观指标评价

由于不同的客观评价指标只能从不同方面对融合结果作出评价,存在一定的片面性,为了提高融合算法综合性能评价结果的可靠性和准确性,本文用文献[18]的方式对4种客观评价指标(结构相似度 SSIM、图像质量因子  $Q_{AB/F}$ 、信息熵 EN、互信息指数 MI)进行构造,得到  $m$  种算法在  $n$  种客观指标下的综合客观评价指标,计算方法如下。

(1) 首先第  $i$  种算法在第  $j$  种客观指标下的相对量化指标为:

$$S_{ij} = \frac{E_{ij} - \min(E_{1j}, E_{2j}, \dots, E_{mj})}{\max(E_{1j}, E_{2j}, \dots, E_{mj}) - \min(E_{1j}, E_{2j}, \dots, E_{mj})} \quad (19)$$

其中,  $E_{ij}$  是算法  $A_i$  在评价指标  $X_j$  下的具体量化值。

(2) 算法  $A_i$  的综合客观评价指标为:

$$S_{wj} = \sum_{j=1}^n w_j S_{ij}, i = 1, 2, \dots, 5 \quad (20)$$

即综合客观评价指标是对上述4种客观单一指标进行综合加权以后的结果。这4种指标与图像融合质量均

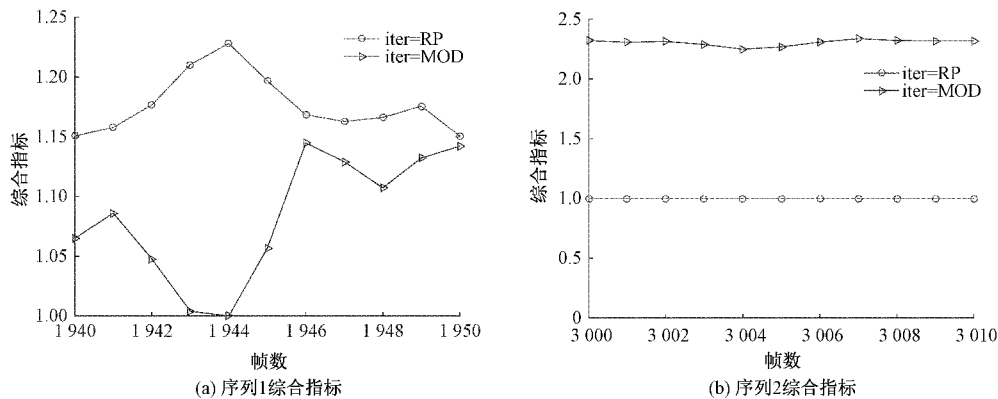


图9 优势算法综合指标评价结果

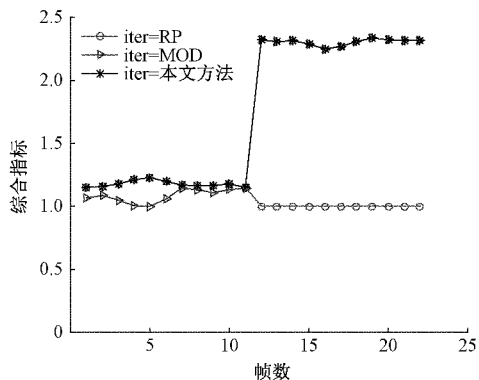


图10 本文拟态融合方法与固定算法对比结果

表1 视频拟态融合方法与固定算法综合指标对比

融合方法	平均值	提升百分比/%
固定 RP	1.088 3	59.925
固定 MOD	1.693 7	2.760 8
本文方法	1.740 5	

成正相关,数值越大,融合结果保留源图像细节、边缘等结构信息越丰富,图像融合效果也越好。因此,综合客观评价指标的数值越大,代表图像融合效果越好。

由于在2.3节中根据多融合算法决策评分排序结果选取了序列1和序列2上的最优融合算法分别是RP和MOD,本文仅绘制了如图9所示RP和MOD算法的综合指标评价结果,该结果与图7中的多融合算法决策评分排序结果所选出来的不同序列上的最优算法结果是一致的。为了更直观地说明本文视频拟态融合方法的综合优势,将序列1(第1940~1950帧)和序列2(第3000~3010帧)整合成一段序列(共计22帧,坐标数值从0~25),综合评价指标结果如图10所示。可以看到在整段序列上,本文视频拟态融合方法即随着动态场景图像差异特征变化动态变换融合算法相较于固定单一算法MOD和RP在综合客观指标上呈现最优,整段视频序列数值计算结果如表1所示。由此可见,本文视频拟态融合方法有效可行,能够在视频序列上根据动态场景中图像差异特征变化动态选取最优融合算法的目的。

## 3 结 论

本文提出了一种基于差异特征驱动的红外与可见光视频拟态融合方法,在提取视频帧间差异特征基础上利用改进的融合有效度公式求得不同融合算法对3种差异特征的融合有效度,然后利用熵权法得到多融合算法决策评分,进而确定不同序列段上的最优融合算法。实验结果表明,本文视频拟态融合方法对比单一固定的融合算法在综合指标评价下融合效果更好。同时本文给出了在不同场景序列动态选取优势融合算法的一种具体方法,为视频拟态融合及算法智能化选择提供一种新思路。

## 参考文献

- [1] 沈英,黄春红,黄峰,等. 红外与可见光图像融合技术的研究进展[J]. 红外与激光工程, 2021, 50(9): 152-169.
- [2] 冯贺,李立,赵凯. 基于拉普拉斯分解耦合亮度调节的

- 可见光与红外图像融合算法[J]. 电子测量与仪器学报, 2020, 34(10): 91-97.
- [3] XIA J, LU Y, TAN L, et al. Research of multimodal medical image fusion based on parameter-adaptive pulse-coupled neural network and convolutional sparse representation[J]. Computational and Mathematical Methods in Medicine, 2020, DOI: 10.1155/2020/3290136.
- [4] 罗迪, 王从庆, 周勇军. 一种基于生成对抗网络与注意力机制的可见光和红外图像融合方法[J]. 红外技术, 2021, 43(6): 566-574.
- [5] 李春艳, 孙韬, 谢俊峰. EMF深度学习可见光/红外图像融合算法[J]. 国外电子测量技术, 2020, 39(10): 25-32.
- [6] HU H M, WU J, BO L, 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.
- [7] WD A, SONG X A, JIE L B, et al. Fusion of hyperspectral and panchromatic images using structure tensor and matting model[J]. Neurocomputing, 2020, 399: 237-246.
- [8] ZHANG Q, WANG Y, MARTIN D LEVINE, et al. Multisensor video fusion based on higher order singular value decomposition[J]. Information Fusion, 2015, 24: 54-71.
- [9] 杨风暴. 红外偏振与光强图像的拟态融合原理和模型研究[J]. 中北大学学报(自然科学版), 2017, 38(1): 1-8.
- [10] 焦玉茜. 面向拟态融合的红外图像差异特征驱动方法研究[D]. 太原: 中北大学, 2019.
- [11] 杨风暴, 吉琳娜. 双模态红外图像差异特征多属性与融合算法间的深度集值映射研究[J]. 指挥控制与仿真, 2021, 43(2): 1-8.
- [12] 焦玉茜, 杨风暴, 吉琳娜, 等. 基于余弦值的红外光强与偏振图像特征差异度的度量方法[J]. 红外技术, 2019, 41(1): 27-34.
- [13] 张雅玲, 吉琳娜, 杨风暴, 等. 基于余弦相似性的双模态红外图像融合性能表征[J]. 光电工程, 2019, 46(10): 82-92.
- [14] 段汕, 王小凡, 张洪. 图像相似性度量方法的研究[J]. 中南民族大学学报(自然科学版), 2016, 35(4): 121-125.
- [15] ZHU Y, TIAN D, YAN F, et al. Effectiveness of entropy weight method in decision-making [J]. Mathematical Problems in Engineering, 2020, DOI: 10.1155/2020/3564835.
- [16] LIU F, CHEN L, LU L, et al. Medical image fusion method by using Laplacian pyramid and convolutional sparse representation[J]. Concurrency and Computation: Practice and Experience, 2020, DOI: 10.1002/cpe.5632.
- [17] BUDHIRAJA S, RUMMY I, AGRAWAL S, et al. Infrared and visible image fusion based on sparse representation and spatial frequency in DTCWT domain [J]. International Journal of Image and Graphics, 2020, DOI: 10.1142/S0219467821500170.
- [18] 陈立伟, 蒋勇. 图像融合算法的综合性能评价指标[J]. 计算机工程, 2015, 41(2): 219-223.

### 作者简介

李向燕, 硕士研究生, 主要研究方向为图像信息处理。

E-mail: 1102412001@qq.com

王肖霞, 博士, 副教授, 主要研究方向为信息处理与计算机成像。

E-mail: wangxiaoxia@nuc.edu.cn

杨风暴, 博士, 教授, 主要研究方向为信息融合。

E-mail: yfengb@163.com