



基于主客观评价相关性的多波段 融合图像评价方法

韩 泽, 蔺素珍*

中北大学计算机与控制工程学院,太原 030051



摘要:针对图像融合难以选择合适的评价指标的问题,通过分析主观评价与客观评价指标之间的相关性,提出一种合成评价指标方法对多波段图像融合结果进行评价。首先,从边缘清晰度、自然感、信息量及综合评价四个方面对多种方法融合结果进行主观评价;其次,用14种常用的客观评价指标对融合结果进行评价;然后,对主客观评价结果分别 归一化,并采用斯皮尔曼(Spearman)相关系数分析每种客观评价指标和四种主观评价之间的相关性;最后,根据该相 关性分别从四个评价方面,用14种客观指标合成一种综合指标。实验结果表明,综合指标与主观评价的相关性比单项 指标或其他合成指标更高。

关键词:图像融合;主观评价;客观评价;斯皮尔曼相关系数;综合指标中图分类号: TP751.1文献标志码: A

Multiband fusion image evaluation method based on correlation between subject and object evaluation

Ze Han and Suzhen Lin*

School of Computer and Control Engineering, North University of China, Taiyuan 030051, China

Abstract: It is difficult to select an appropriate evaluation index for current image fusion. In order to solve the problem, a synthesis evaluation index is proposed based on the correlation between subjective and objective evaluations. First, a variety of fusion results are evaluated subjectively from the edge of clarity, natural sense, information and comprehensive evaluation, respectively. Secondly, 14 commonly objective indexes are used to evaluate the fusion results. Then, the subjective and objective results are normalized, and the Spearman correlation coefficient is used to analyze the correlation between the four subjective evaluations and each objective evaluation. Finally, according to the correlation, a comprehensive index is constructed through 14 objective indexes in 4 aspects. The experimental results show that the synthesis index is more relevant to subjective evaluation than the individual evaluation index and any other comprehensive indexes.

Keywords: image fusion; subjective evaluation; objective evaluation; Spearman correlation coefficient; comprehensive index

DOI: 10.3969/j.issn.1003-501X.2017.09.006

Citation: Opto-Elec Eng, 2017, 44(9): 895–902

收稿日期:2017-05-30; 收到修改稿日期:2017-07-13 *E-mail:13835163417@163.com

1 引 言

多波段图像融合是信息融合的一个重要分支,是 将同一场景的多幅图像或序列探测图像合成一幅信息 更完整、全面的图像,在遥感探测、计算机视觉、目 标探测与识别等领域应用广泛[1]。在实际研究中,需 要对融合结果进行像质评价,但由于融合图像类型不 同,当前并无标准的评价方法,需研究者根据经验从 众多客观评价指标中选择合适的指标评价[2-8]。文献[9] 从信息量、图像特征、图像结构相似度、人类视觉感 知系统四方面分别选择互信息(mutual information, MI)、边缘信息保留值(edge information preservation values ,EIPV)、结构相似矩阵(structural similarity based metric, SSBM)和感知矩阵(human perception based metric, HPBM)作为评价指标文献[10], 采用光谱角度 映射图(spectral angle mapper, SAM)、结构相似度 (structural similarity, SSIM)^[11]、均方根误差(root mean square error, RMSE)、误差相对总量(erreur relative globale adimensionnelle de synthèse, ERGAS)对结果进 行评价。由于不同研究选择不同评价指标体系无法对 比,使得文献结果的说服力较差。当前的研究热点是 针对评价目的选择相关的评价指标,并将其合成一种 综合指标,通过指标间优势互补提高评价准确性,主 要有以下几类[12,13]:1)将所选择的指标按照一定的权 重相加合成一个指标[14,15],但需要依据经验来选择合 适的权重;2) D-S(dempster/shafer)证据理论、模糊集 理论和灰色关联分析理论在内的不确定性理论方法, 其中基于 D-S 证据理论的方法通过信任函数来对所选 指标进行综合,但在实际应用中很难确定基础概率; 基于模糊集理论的方法通过隶属度函数刻画评价指标 的重要程度来对指标进行合并,但该函数的建立当前 并无公认有效的方法;基于灰色关联分析理论[12,13]的 方法通过不同方案间的灰色关联系数取最小来对指标 合成,但需要计算每组评价图像的合适权重,不利于 工程化;3) 基于统计或机器学习的方法[16,17],通过从 大量样本数据中寻找合适的特征及权重来合成评价指 标,但其研究对象为可见光图像,由于对于多波段图 像融合研究缺少训练数据,很难获得有效的权重。综 上所述,尽管构成综合指标的理论众多,各有优势。 但无论选择何种理论,其目的都是希望合成的指标与 主观评价的吻合度高[13,18],然而这些研究在构建综合 指标时并未直接考虑吻合度,造成准确性降低的问题。

本文以多波段融合图像为对象,通过从边缘清晰

2017年,第44卷,第9期

度、自然感、信息量和综合评价四个评价方面对融合 图像进行主观评价,并计算多种客观评价指标分别与 主观评价的相关性,然后根据该相关性为每个评价方 面将这些客观评价指标按一定的权重合成一个综合指 标以提高客观评价结果准确性。



图1 本方法基本思路图.

Fig. 1 Basic idea of multiband fusion image evaluation method.

2 多波段融合图像评价方法

2.1 基本思路

多波段融合图像评价方法基本思路如图1所示。

1) 输入多波段图像集 *I*: {*I*₁, *I*₂,...,*I*_k}^K_{k=1},其中 *I*_k
 表示第 *k* 组待融合图像, 共 *K* 组;

2) 输入图像经融合方法集 $M: \{M_1, M_2, \dots, M_n\}_{n=1}^N$, 得到融合结果集 $F: \{F_{11}, F_{12}, \dots, F_{nk}\}_{n=1,k=1}^{N,K}$ 其中 M_n 表示 第n 种融合方法, 共N种, F_{nk} 表示第k 组图像第n 种融 合方法的融合结果;

3) 对融合结果进行主观评价, 评测员共 T人, 从
 U 个方面对融合结果集 F 进行主观评价, 得到
 T×K×N×U 的 评 价 结 果 矩 阵
 X = {X₁₁₁₁, X₁₁₁₂, ···, X_{iknu}^{T,K,N,U}_{i=1,k=1,n=1,u=1};

4) 平均并归一化主观评价结果集矩阵,得到 X',其形状为 K×N×U;

5) 采用 V 个客观评价指标对融合结果集 F 进行 评 价 , 得 到 $K \times N \times V$ 的 评 价 结 果 矩 阵 $Y = \{Y_{111}, Y_{112}, \dots, Y_{km}\}_{k=1,n=1,v=1}^{K,N,V}$,其中 Y_{knu} 表示第 k 个图 像的第 n 种方法的第 v 个指标的值;

6) 对 Y 进行归一化,得到 Y';

7) 计算 X' 和 Y' 之间的相关系数矩阵 $\overline{\theta}$;

8) 用相关系数矩阵 $\bar{\theta}$ 生成U 种主观评价对应的 综合指标的权重矩阵;

9) 计算所有测试图像每个评价方面对应的综合 指标;

10) 计算该综合指标与主观评价的相关性,进行 误差分析。

3 主观与客观评价实验

- 3.1 主观评价实验
- 3.1.1 评价指标

DOI: 10.3969/j.issn.1003-501X.2017.09.006

	Table 1 Subjective evaluation reference form.							
等级	分值	边缘清晰度	自然感	信息量	综合评价			
ħΖ	-	物体的边缘清晰锐利 ,纹理细节	图像干净无失真、光线分布均	场景中纹理清楚,细节信息	综合边缘清晰度、自然			
XI	5	丰富,很容易将各物体区分	匀,整体感觉自然舒服	多,一眼就能找到	感、信息量最好			
标忆	4	物体的边缘清晰 ,很容易将各物	没有明显块状、条状或点状失	场景中纹理清楚 ,细节信息较	综合边缘清晰度、自然			
ŦXXŢ	4	体区分开	真,不影响观 看	多,很容易就能找到	感、信息量较好			
რЛ	2	物体的边缘没有锯齿或失真 ,不	有少量失真,但画面过度自然,	场景部分纹理、 细节信息 , 不	综合边缘清晰度、自然			
DQ	3	影响观看	不影响观看	需费力就能找到	感、信息量一般			
林辛	2	物体的边缘是少量的锯齿或失	有少量块状、条状或点状失真 ,	场景中有部分纹理细节信息,	综合边缘清晰度、自然			
牧左	2	真,轻微影响观看	画面过度不自然 ,感觉轻微不适	需要费力才能找到	感、信息量较差			
兰	4	物体的边缘是有严重的锯齿或	有大量块状、条状或点状失真 ,	场景中完全没有纹理等细节	综合边缘清晰度、自然			
差	Ĩ	失真,严重影响观看	画面过度不自然 , 感觉不适	信息	感、信息量差			

表1 主观评价参考表.

图像的主观评价是直接利用观察者对被测图像的 主观反应来确定像质好坏的一种测试方法。常用的李克 特量表将图像质量分为五个层次,分别对应分数5、4、 3、2、1,表示"好"、"较好"、"一般"、"较差" 和"差"。但该方法仅对图像的整体感觉进行评价, 而且相邻级别之间的界定描述不清。本实验分别从边 缘清晰度、自然感、信息量和综合评价四个评价方面 对融合结果进行主观评价,并对每个层次之间的界定 做出详细规定,如表1所示。需要注意的是,综合评 价并不是将边缘清晰度、自然感、信息量的结果进行 平均 ,而是从主观角度综合图像边缘清晰度、自然感、 信息量进行评价。

3.1.2 实验条件

1) 实验图像。待评价图像为8组源图像的8种融 合方法的融合结果。8 组源图像包括 3 组可见光与红 外长波图像和 5 组可见光、红外短波和红外长波图像。 图像来自 TNO 图像融合库和 www.imagefusion.org。8 种融合方法涉及空间域、变换域、混合域和其他理论, 其中空间域的方法有主成分分析(principal component analysis, PCA)、加权平均(average, AG), 变换域的方 法有低通比金字塔(ratio of low-pass pyramid, RP)^[19]、 非下采样轮廓波变换(nonsubsampled contourlet transform, NSCT)^[19]、支持度变换(support value transform, SVT), 混合域的方法有拉普拉斯金字塔和 稀疏表示(Laplacian pyramid and sparse representation, LPSR)^[19],其他的方法有稀疏表示(sparse representation, SR)^[19]、深度卷积神经网络(deep convolutional neural network, DCNN)^[9]。各方法参数按文献中最优 情况设置。考虑到篇幅问题,后文给出两组待评测图 像(图 2、图 3),这两组融合图像都是三波段图像融合 结果。

2) 评价员。共 22 名(10 男 12 女),平均年龄 20~25 岁,视力(矫正后)正常,色视觉正常,多数为从事计 算机专业的硕士研究生,仅有少部分对多波段图像成



图 2 第 1 组待评价图像. (a) RP 融合结果图. (b) NSCT 融合结果图. (c) PCA 融合结果图. (d) SVT 融合结果图. (e) AVG 融 合结果图. (f) LPSR 融合结果图. (g) SR 融合结果图. (h) DCNN 融合结果图.

Fig. 2 Group 1 image to be evaluated. (a) RP fusion result. (b) NSCT fusion result. (c) PCA fusion result. (d) SVT fusion result. (e) AVG fusion result. (f) LPSR fusion result. (g) SR fusion result. (h) DCNN fusion result.

2017年,第44卷,第9期



图 3 第 2 组待评价图像. (a) RP 融合结果图. (b) NSCT 融合结果图. (c) PCA 融合结果图. (d) SVT 融合结果图. (e) AVG 融合结果图. (f) LPSR 融合结果图. (g) SR 融合结果图. (h) DCNN 融合结果图.

Fig. 3 Group 2 image to be evaluated. (a) RP fusion result. (b) NSCT fusion result. (c) PCA fusion result. (d) SVT fusion result. (e) AVG fusion result. (f) LPSR fusion result. (g) SR fusion result. (h) DCNN fusion result.

Iable 2 Subjective evaluation results.									
图像	评价方面	RP	NSCT	PCA	SVT	AVG	LPSR	SR	DCNN
	清晰度	3.47	3.85	2.76	3.85	1.52	3.38	2.90	2.80
笠 1 49	自然感	2.85	3.85	2.80	3.80	2.38	3.04	2.80	2.80
 第1组	信息量	3.57	4.04	3.09	3.47	1.90	3.33	2.90	2.76
	综合	3.23	3.88	2.84	3.57	2.28	2.93	2.90	2.66
	清晰度	3.90	3.95	2.95	4.04	3.00	3.95	1.47	2.04
笠り石	自然感	3.80	3.61	3.47	4.23	3.33	3.28	1.52	1.76
第 ∠组	信息量	4.42	4.09	2.90	3.80	2.71	3.71	1.76	1.52
	综合	3.79	3.90	3.20	3.82	2.88	3.79	1.73	1.69

表2 主观评价结果.

像和图像评价有研究背景。

3) 试验环境。实验在暗室中进行,显示设备为投 影仪,投影仪亮度为4000 lm,对比度为3000 1,标 准分辨率为1024 pixel×768 pixel,幕布大小为4m,在 实验前投影仪进行了1 h的预热,评测者与幕布距离 为5~8 m,坐在幕布正前方。

4) 实验过程。实验开始前向实验者详细讲解整个 实验的过程、多波段图像成像和图像评价的相关知识, 通过一组例子(不在最后评价图像中)讲解每种指标的 意义,并允许提问。实验开始后,依次投影待评价图像, 每组图像的8个融合结果同时存在,在规定的时间内对 每个图像的四种指标进行评分。整个实验在40 min 内 完成。两组图像评价员的平均结果如表 2 所示。

3.2 客观评价实验

图像融合领域常用的 14 种评价指标^[19-21]主要分为 三类,其中基于统计特性的评价指标包括标准差 (standard deviation, SD)、峰值信噪比(peak signal to noise ratio, PSNR)、相关系数(correlation coefficient, CC)、对比度(contrast,C)、平均梯度(average gradient, AG)和空间频率(spatial frequency,SF);基于信息量的 评价指标包括信息熵(information entropy,IE)和 MI; 基于人类视觉系统的评价指标包括 SSIM、EIPV、视觉 信息保真度(visual information fidelity for fission, VIFF)^[22]和 Piella 模型融合质量标准,包括图像融合质量 指数(image fusion quality index,IFQI)、加权融合质量 指数(weighted fusion quality index,WFQI)和边缘相关 融合质量指数(edge-dependent fusion quality index, EFQI)。其中 SD、C、AG、SF、IE 为无参考指标,其余 均为有参考指标,分别和源图像计算再取均值。评价指 标越大越好。

4 生成综合指标

在众多相关系数理论中,皮尔逊(Pearson)相关系 数和斯皮尔曼(Spearman)相关系数即简单又高效^[23], 但 Pearson 相关系数要求数据服从正态分布和同方差,

DOI: 10.3969/j.issn.1003-501X.2017.09.006

并且仅能分析线性关系,而 Spearman 相关系数却没有 以上限制,在像质评价相关性分析中应用相对广泛。 文献[18]用其分析高动态范围(high dynamic range imaging, HDR)视频的主客观评价之间的相关性,文 献[21]用其分析各指标之间的相关性,都证明其有效。 为此本文以 Spearman 相关系数分析主客观评价指标 之间的相关性。

在经过主客观评价实验后得到 64 幅图像的主观 评价结果矩阵 X 和客观评价结果矩阵 Y ,按以下步骤 生成综合指标:

1) 平均并归一化主观评价结果集矩阵,得到 X':

$$\left\{\frac{\sum_{t=1}^{T} X_{t111}}{5T}, \frac{\sum_{t=1}^{T} X_{t112}}{5T}, \dots, \frac{\sum_{t=1}^{T} X_{tknu}}{5T}\right\}_{k=1, n=1, u=1}^{K, N, U} 5 \text{ 为主观评价}$$

最高等级对应的分数;

2) 对 Y 进行归一化: 假设 Y^{kv}_{max} 表示第 k 行图像第
 v 个指标 N 种方法中的最大值, 归一化的结果 Y' 为

$$\left\{\frac{\boldsymbol{Y}_{111}}{\boldsymbol{Y}_{\max}^{1,1}}, \frac{\boldsymbol{Y}_{112}}{\boldsymbol{Y}_{\max}^{1,2}}, \cdots, \frac{\boldsymbol{Y}_{knv}}{\boldsymbol{Y}_{\max}^{k,v}}\right\}_{k=1,n=1,v=1}^{K,N,V};$$

3) 计算 X' 和 Y' 之间的相关系数,得到相关系数矩 阵 $\theta = \{\theta_{111}, \theta_{112}, \dots, \theta_{kuv}\}_{k=1,u=1,v=1}^{K,U,V}$, θ_{kuv} 为第k 组图像第u 个主观评价指标与第v个客观评价指标之间的相关系数:

$$\theta_{kuv} = 1 - \frac{6\sum_{n=1}^{N} d_n^2}{N(N^2 - 1)} , \qquad (1)$$

式中: $d_n = X''(n) - Y''(n)$, X'' 和 Y'' 分别为 X' 和 Y' 中元素在 X'_{sorted} 和 Y'_{sorted} 中的位置, X'_{sorted} 和 Y'_{sorted} 由 X' 和 Y' 按降序排列后得到;

4) 对相关系数矩阵 θ 进行平均,得:

 $\overline{\boldsymbol{\theta}} = \{\overline{\boldsymbol{\theta}}_{11}, \overline{\boldsymbol{\theta}}_{12}, \cdots, \overline{\boldsymbol{\theta}}_{uv}\}_{u=1, v=1}^{U, V};$

5) 用相关系数矩阵 $\bar{\boldsymbol{\theta}}$ 生成U种主观评价对应的 综合指标的权重矩阵 $\boldsymbol{\omega}=\{\boldsymbol{\omega}_{11}, \boldsymbol{\omega}_{12}, \cdots, \boldsymbol{\omega}_{nr}\}_{n=1}^{U,V}$,

$$\boldsymbol{\omega}_{uv} = \begin{cases} \lambda \frac{\overline{\boldsymbol{\theta}}_{uv}}{\left|\sum_{\nu=1}^{V} \min(0, \boldsymbol{\theta}_{u\nu})\right|}, & \overline{\boldsymbol{\theta}}_{uv} < 0\\ \frac{\overline{\boldsymbol{\theta}}_{uv}}{\left|\sum_{\nu=1}^{V} \max(0, \overline{\boldsymbol{\theta}}_{u\nu})\right|}, & \text{otherwise} \end{cases}$$
(2)

式中: $\sum_{v=1}^{\nu} \max(0, \overline{\theta}_{uv})$ 为 $\overline{\theta}_{uv}$ 中大于 0 的元素之和, $\left|\sum_{v=1}^{\nu} \min(0, \overline{\theta}_{uv})\right|$ 为 $\overline{\theta}_{uv}$ 中小于 0 的元素之和的绝对值, λ 为负权重占总权重的比重,依据经验 $\lambda = 0.5$ 。

表3为14种客观评价指标分别与主观评价之间的

表3 主客观评价指标的平均相关系数及权重表.

Table 3	Average correlation coefficients and weight table of subjective and objective evaluation.	

+6+=			权重					
相你	清晰度	自然感	信息量	综合	清晰度	自然感	信息量	综合
SD	-0.033	-0.241	-0.073	-0.174	-0.050	-0.131	-0.059	-0.114
IE	0.147	-0.062	0.099	0.004	0.044	-0.034	0.039	0.002
AG	0.471	0.284	0.546	0.392	0.140	0.133	0.218	0.174
SF	0.420	0.248	0.511	0.356	0.125	0.116	0.204	0.158
С	-0.017	-0.220	-0.090	-0.167	-0.025	-0.120	-0.073	-0.109
MI	-0.245	-0.395	-0.457	-0.426	-0.368	-0.215	-0.368	-0.278
PSNR	-0.038	0.229	0.086	0.176	-0.058	0.107	0.034	0.078
CC	0.128	0.274	0.017	0.152	0.038	0.128	0.007	0.067
SSIM	0.252	0.235	0.017	0.145	0.075	0.110	0.007	0.064
EIPV	0.295	0.129	0.099	0.130	0.088	0.060	0.039	0.057
VIFF	0.569	0.396	0.579	0.477	0.170	0.185	0.231	0.211
IFQI	0.298	0.044	0.158	0.083	0.089	0.021	0.063	0.037
WFQI	0.379	0.108	0.223	0.156	0.113	0.051	0.089	0.069
EFQI	0.393	0.189	0.170	0.183	0.117	0.088	0.068	0.081

平均相关系数及构建综合指标时所用的权重。

在得到权重矩阵后,即可根据该矩阵计算测试图 像的综合指标:

 1) 输入待测试的图像,分别计算图像的 14 种客 观评价指标,并归一化得 Y';

2) 根据权重矩阵 $\boldsymbol{\omega}$ 计算第 u 个评价方面对应的 综合指标 $\bar{y}_{u} = \sum_{\nu}^{\nu} \boldsymbol{\omega}_{uv} \times \boldsymbol{Y}_{v}^{\prime}$ 。

5 实验结果分析

为了说明本文方法的有效性,选择多层次合成分析 (MSA)^[12]和复数矩阵奇异值分解(CMSVD)^[13]进行对比实 验。CMSVD 通过复数矩阵奇异值分解对多种客观评价 指标综合,MSA 将评价指标分为"The bigger, the better", "The smaller, the better", "Moderate"三类,通过多 层次灰色关联法生成综合指标,二者都证明其比多种合 成方法效果好。实验中发现,按 MSA 所给出的分类体 系,"The bigger, the better"类的客观指标构建的综合 指标相比于其他两类与四种主观评价相关性都高,因 此, MSA 在进行实验时选择 14 种客观指标中"The bigger, the better"类的指标构建综合指标。CMSVD 在 构建综合指标时,分别选择与对应主观评价相关性最高 的5种客观评价指标。图4为三种方法主观评价与综合 指标的相关系数图,表4为三种方法的平均相关系数。

从结果可看出,本文方法构建的综合指标相比于 单项指标相关性高。相比于 MSA 和 CMSVD,本文方 法平均相关系数最高,MSA 次之,CMSVD 最差,原 因在于,1)评价对象不同,CMSVD 和 MSA 待融合 图像仅为两幅图像,且 CMSVD 和 MSA 所选的融合方 法都为多尺度变换方法,由于仅选择单类型融合方法 通用性差,因此本文选择不同类型的融合方法;2)主 观评价标准不同,CMSVD 和 MSA 都仅对融合结果做 了综合评价,本文对边缘清晰度、自然感、信息量和 综合评价四个方面进行了主观评价,更全面;3)客观 评价指标选择不同。在众多的指标中 VIFF、AG 和 SP 与四项主观评价相关性均比较高,原因在于,VIFF 反 映非扭曲信息和扭曲信息的比值,该比值可反映图像





图 4 三种综合指标与主观评价相关系数图. (a) 边缘清晰度相关系数图. (b) 自然感相关系数图. (c) 信息量相关系数 图. (d) 综合评价相关系数图.

Fig. 4 Correlation coefficient between three comprehensive indexes and subjective evaluation. (a) Clarity of edge correlation coefficient. (b) Natural sense correlation coefficient. (c) Information correlation coefficient. (d) Comprehensive evaluation correlation coefficient.

DOI: 10.3969/j.issn.1003-501X.2017.09.006



图 5 验证实验三种综合指标与主观评价相关系数图.(a)边缘清晰度相关系数图.(b)自然感相关系数图.(c)信息量 相关系数图.(d)综合评价相关系数图.

Fig. 5 Correlation coefficient between three comprehensive indexes and subjective evaluation of validation experiment. (a) Clarity of edge correlation coefficient. (b) Natural sense correlation coefficient. (c) Information correlation coefficient. (d) Comprehensive evaluation correlation coefficient.

的失真程度,而人在观察图像时在信号进入视觉系统 之前要先经过"失真通道",因此无论做任何评价, 如果图像有失真情况,人都不会感觉良好^[12,23],空间 频率与平均梯度可反映图像边缘的清晰程度,而人眼 对物体边缘敏感,因此二者相关性较高。在这些指标 中,MI都为负相关,原因在于互信息度量融合图像与 输入图像在灰度分布上的相似程度,而红外与可见光 图像的灰度差异很大,不同融合方法视觉差异很大, 这会造成主观评价结果与客观结果差异大的问题^[21]。

为了验证本方法的普适性,在相似的实验环境下进 行了验证实验。选择另外 8 组图像进行验证,其中 4 组 为双波段图像,其余 4 组为三波段图像。评价员为 12 人(6 男 6 女),年龄在 22~30 岁之间。实验前,他们会 了解到一些红外成像的专业知识,少部分人具有相关的 行业背景。评价员与图 4 实验人员不同。评价结束后, 利用表 3 的权重构建综合指标,并分别与主观评价进行 相关性分析。作为对比,MSA 选择"The bigger, the better"类的指标构建综合指标,CMSVD分别选择与对 应主观评价相关性最高的 5 种客观评价指标构建综合指标。实验结果如图 5 所示,三种方法的平均相关系数如 表 5 所示。从结果可以看出,本方法得到的综合指标相 比于 CMSVD、MSA 与主观评价的相关性依然更高。

表4 三种合成方法与主观评价平均相关系数.

Та	ıbl	e 4	1 A	Average	correlatio	on coeff	icient	of t	hree	meth	nods

	清晰度	自然感	信息量	综合
CMSVD	0.455	0.412	0.402	0.388
MSA	0.455	0.496	0.571	0.558
本文方法	0.634	0.630	0.737	0.661

表 5 验证实验三种合成方法与主观评价平均相关系数. Table 5 Validation experiment average correlation coefficient of three methods.

	清晰度	自然感	信息量	综合
CMSVD	0.258	0.342	0.217	0.318
MSA	0.138	0.078	0.019	0.113
本文方法	0.478	0.657	0.456	0.653

6 结 论

通过分析主观评价与客观评价指标之间的相关

性,可得出综合指标相比于单项指标与主观评价相关 性有明显提升,其中与清晰度、自然感、信息量和综 合评价的相关性分别为 0.634、0.630、0.737、0.661。 本文构建的综合指标相比CMSVD和MSA与主观评价 的相关性高,并且计算简单,方便实现。对不同的评 价方面,每个客观评价指标与主观评价的相关性不同, 但 AG、SF和 VIFF 与四种主观评价相关性都较高。

虽然本文方法所得的综合指标与主观评价相关性 高,但其结果依赖于主观评价的准确性,而主观评价 会受很多因素干扰,因此接下来考虑如何消除这些因 素的干扰。

基金项目

山西省应用基础研究项目(201701D121062);中北 大学第十三届研究生科技立项(20161354)。

参考文献

- Li Shutao, Kang Xudong, Fang Leyuan, *et al.* Pixel-level image fusion: a survey of the state of the art [J]. *Information Fusion*, 2017, **33**: 100–112.
- 2 Ding Li, Huang Hua, Zang Yu. Image quality assessment using directional anisotropy structure measurement [J]. *IEEE Transactions on Image Processing*, 2017, **26**(4): 1799–1809.
- 3 Krasula L, Le Callet P, Fliegel K, *et al.* Quality assessment of sharpened images: challenges, methodology, and objective metrics [J]. *IEEE Transactions on Image Processing*, 2017, 26(3): 1496–1508.
- 4 Vega M T, Mocanu D C, Stavrou S, et al. Predictive no-reference assessment of video quality [J]. Signal Processing: Image Communication, 2017, 52: 20–32.
- 5 Alaql O Ghazinour K, Lu Cheng Chang. Classification of image distortions for image quality assessment [C]// Proceedings of International Conference on Computational Science and Computational Intelligence. 2016: 653–658.
- 6 Yan Wen, Gong Fei, Zhou Ying, et al. Satellite cloud image fusion based on adaptive PCNN and NSST [J]. Opto-Electronic Engineering, 2016, 43(10): 70–76,83. 颜文, 龚飞, 周颖, 等. 基于 NSST 与自适应 PCNN 相结合的卫星
- 云图融合 [J]. 光电工程, 2016, 43(10): 70-76,83.
 7 Yin Ming, Duan Puhong, Chu Biao, et al. CT and MRI medical image fusion based on shift-invariant shearlet transform and compressed sensing [J]. Opto-Electronic Engineering, 2016,

43(8): 47-52. 殷明,段普宏,褚标,等. 结合 SIST 和压缩感知的 CT 与 MRI 图 像融合 [J]. 光电工程, 2016, **43**(8): 47-52.

8 Zhang Xuedian, Wang Hong, Jiang Minshan, et al. Applications of saliency analysis in focus image fusion [J]. Opto-Electronic

Engineering, 2017, 44(4): 435-441. 张学典, 汪泓, 江旻珊, 等. 显著性分析在对焦图像融合方面的应 用 [J]. 光电工程, 2017, 44(4): 435-441.

- 9 Liu Yu, Chen Xun, Peng Hu, et al. Multi-focus image fusion with a deep convolutional neural network [J]. Information Fusion, 2017, 36: 191–207.
- 10 Zhang Kai, Wang Min, Yang Shuyuan. Multispectral and hyperspectral image fusion based on group spectral embedding and low-rank factorization [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**(3): 1363–1371.
- 11 Wang Zhou, Bovik A C. A universal image quality index [J]. *IEEE* Signal Processing Letters, 2002, **9**(3): 81–84
- 12 He Guiqing, LiangFan, Xing Siyuan, *et al.* Study on algorithm evaluation of image fusion based on multi-hierarchical synthetic analysis[C]// Proceedings of 2016 IEEE International Conference on Signal Processing, Communications and Computing, 2016: 1–6.
- 13 Zhu Yahui. Research on quality evaluation methods of infrared and visible image fusion[D]. Xi'an: Northwestern Polytechnical University, 2015.

朱亚辉. 红外与可见光图像融合质量评价方法研究[D]. 西安: 西 北工业大学, 2015.

- 14 Xydeas C S, Petrovic V. Objective image fusion performance measure[J]. *Electronics Letters*, 2000, **36**(4): 308–309.
- 15 Piella G, Heijmans H. A new quality metric for image fusion [C]// Proceedings of 2003 International Conference on Image Processing, 2003, 2: III-173–176.
- 16 Nizami I F, Majid M, Khurshid K. Efficient feature selection for blind image quality assessment based on natural scene statistics [C]// Proceedings of 2017 14th International Bhurban Conference on Applied Sciences and Technology, 2017: 318–322.
- 17 Ding Yong, Zhao Yang, Zhao Xinyu. Image quality assessment based on multi-feature extraction and synthesis with support vector regression [J]. *Signal Processing: Image Communication*, 2017, **54**: 81–92.
- 18 Mukherjee R, Debattista K, Bashford-Rogers T, et al. Objective and subjective evaluation of high dynamic range video compression [J]. Signal Processing: Image Communication, 2016, 47: 426–437.
- 19 Liu Yu, Liu Shuping, Wang Zengfu. A general framework for image fusion based on multi-scale transform and sparse representation [J]. *Information Fusion*, 2015, **24**: 147–164.
- 20 Jagalingam P, Hegde A V. A Review of quality metrics for fused image [J]. *Aquatic Procedia*, 2015, **4**: 133–142.
- 21 Zhang Xiaoli, LI Xiongfei, LI Jun. Validation and correlation analysis of metrics for evaluating performance of image fusion[J]. Acta Automatica Sinica, 2014, 40(2): 306–315. 张小利,李雄飞,李军. 融合图像质量评价指标的相关性分析及性 能评估[J]. 自动化学报, 2014, 40(2): 306–315.
- 22 Han Yu, Cai Yunze, Cao Yin, *et al.* A new image fusion performance metric based on visual information fidelity[J]. *Information Fusion*, 2013, **14**(2): 127–135.
- 23 Warne R T. Testing Spearman's hypothesis with advanced placement examination data[J]. *Intelligence*, 2016, 57: 87–95.