

文章编号:1004-4213(2011)02-0311-5

基于结构相似度与感兴趣区域的图像融合评价方法

张勇^{1,2}, 金伟其¹

(1 北京理工大学 光电学院 光电成像技术与系统教育部重点实验室, 北京 100081)

(2 军械工程学院 军械技术研究所, 石家庄 050000)

摘 要:针对红外和可见光图像融合效果评价问题,在分析图像结构相似度算法基础上,结合人眼视觉特性,提出了基于结构相似度与感兴趣区域的图像融合评价方法.利用红外和可见光传感器各自成像特性形成的不同图像特征,分别划分图像感兴趣区域和剩余区域.根据人眼对不同区域的重视程度,分别赋予不同的加权因子,较以往评价方法更突出了图像重要特征在评价结果中的影响,且评价过程无需标准参考图像,实验结果表明该方法能对融合图像质量进行正确评价,其客观评价结果与主观评价结果具有一致性.

关键词:图像融合;结构相似度;感兴趣区域;评价

中图分类号:TP391.4

文献标识码:A

doi:10.3788/gzxb20114002.0311

0 引言

图像融合可减少图像信息的不确定性,提高图像解析度.全面客观的图像融合评价方法不仅可比较各个融合算法的性能优劣,还可根据融合效果对算法进行改进和完善,具有重要的指导意义和参考价值^[1].融合图像主观评价方法虽然具有和人眼一致性好等特点,但评价过程费时费力,应用局限性大^[2].客观评价方法用可定量分析的数学模型表达人对图像的主观感受,再经过各种计算得到图像质量的科学指标.归纳近年来融合图像客观评价方法大致可分为四类.第一类是基于全像素统计理论的评测方法.该评测方法完全忽略人眼视觉特性对图像质量评测结果的影响,算法实现虽然简单,但评测结果常与人的主观感受相偏离^[3].第二类是基于信息理论的评测方法.该方法虽然克服了全像素统计理论中主观赋权的局限性,根据熵、交互信息量等评价指标的差异程度来修正权重,但仍和主观评价结果存在较大差异^[4].第三类是基于人眼视觉系统(Human Visual System, HVS)仿生理理论的评测方法.该方法模仿人眼对像素层绝对差值的认知机制,通过掩盖低于人眼感知能力的差异修正评测结果,从原理上较好地体现了人眼对图像质量的感知机制,但由于目前对人眼视觉感知机制认识的局限性,且仿生模型的实现又非常复杂,模型性能和上述两

种方法相比不具有明显的优越性^[5].第四类是以结构相似度(Structure Similarity, SSIM)为代表的基于特征参量的评测方法.该方法利用人眼视觉机制特点,通过设计和检测能典型代表图像特性的单个或多个特征参量来体现图像质量,算法实现简便,具有较好的应用前景^[6].本文在 SSIM 模型基础上,提出了一种基于 SSIM 模型和感兴趣区域(Region Of Interest, ROI)的图像融合评价方法(ROI and SSIM, R_SSIM),该方法根据红外与可见光传感器产生的不同图像特征,将输入图像划分为 ROI 和剩余区域(No Region Of Interest, NROI)^[7],利用不同权重因子模拟人眼对不同区域的重视程度,突出了图像重要特征在评价结果中的影响,主观评价结果也验证了方法的可行性.

1 结构相似度评价方法

SSIM 模型依据 HVS 高度适合于提取视觉场景中的结构信息,从而使测量结构信息的改变与感知图像质量的变化非常接近. SSIM 模型在计算时采用滑动窗口方法,窗口大小一般为 8×8 ,为避免“块效应”,也可采用一种标准差为 1.5 的 11×11 归一化圆对称高斯函数. X, Y 分别表示待比较两幅图像的窗口图像块,其结构相似度定义为

$$SSIM_{x,y}(i,j) = l_{x,y}(i,j) c_{x,y}(i,j) s_{x,y}(i,j) \quad (1)$$

式中, $l_{x,y}(i,j) = 2\mu_i\mu_j / (\mu_i^2 + \mu_j^2)$, $c_{x,y}(i,j) = 2\sigma_i\sigma_j /$

基金项目:总装备部军械技术研究所基金资助

第一作者:张勇(1974—),男,工程师,博士研究生,主要研究方向为图像融合评价. Email:bit10701159@163.com

导师:金伟其(1961—),男,教授,博导,主要研究方向为彩色夜视图像融合. Email:jwq@bit.edu.cn

收稿日期:2010-08-10;修回日期:2010-11-15

$(\sigma_i^2 + \sigma_j^2)$, $s_{x,y}(i,j) = \sigma_{i,j} / \sigma_i \sigma_j$. $\mu_i, \mu_j, \sigma_i^2, \sigma_j^2, \sigma_{i,j}$ 分别表示两图像块的均值、方差和协方差. 均值、方差分别可看作是对亮度、对比度的粗略估计, 协方差表示两图像块的非线性改变. 当 $(\mu_i^2 + \mu_j^2) / (\sigma_i^2 + \sigma_j^2) \neq 0$ 时, 图像块之间的结构相似度为

$$\text{SSIM}_{x,y}(i,j) = \frac{4\mu_i\mu_j\sigma_{i,j}}{(\mu_i^2 + \mu_j^2)(\sigma_i^2 + \sigma_j^2)} \quad (2)$$

当 $(\mu_i^2 + \mu_j^2)$ 或 $(\sigma_i^2 + \sigma_j^2)$ 接近 0 时, 特别是当图像中出现平坦区域时, 评价将变得不稳定, 为防止该情况的出现, 对式(2)做以下修正.

$$\text{SSIM}_{x,y}(i,j) = \frac{(2\mu_i\mu_j + C_1)(2\sigma_{i,j} + C_2)}{(\mu_i^2 + \mu_j^2 + C_1)(\sigma_i^2 + \sigma_j^2 + C_2)} \quad (3)$$

对所有图像块获得的 SSIM 评价值进行累加平均得到两幅图像的平均结构相似度指标

$$\text{MSSIM}(x,y) = \left(\frac{1}{NM} \right) \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \text{SSIM}_{x,y}(i,j) \quad (4)$$

2 R_SSIM 评价模型

2.1 模型的提出

WANG 提出的 SSIM 模型最初用于压缩图像质量评价, 而一般情况下图像融合效果评价并没有标准参考图像, 因此需对 SSIM 模型进行相应修改, 如文献[8]提出利用熵评价因子代替协方差评价因子, 文献[9]提出以结构信息与交互信息量乘积作为图像融合质量评价指标, 文献[10]提出对图像块及图像区域分别进行加权处理等. 本文根据实际应用的图像融合系统各自传感器的图像特征, 提出基于 SSIM 与 ROI 的图像融合评价方法.

2.2 R_SSIM 评价方法

本文提出的 R_SSIM 评价方法主要应用于红外和可见光图像融合系统, 设输入图像分别为 IR 和 VISIBLE, 构造的参考图像为 R, 融合图像为 F, 图像感兴趣区域和剩余区域分别定义为 ROI 和 NROI. 则参考图像块和融合图像块之间的结构相似度可定义为式(5)

$$\text{SSIM}_{R,F}(i,j) = l_{R,F}(i,j) c_{R,F}(i,j) s_{R,F}(i,j) \quad (5)$$

参考图像可由式(6)表示, λ_1 为加权因子

$$R(i,j) = \lambda_1 r_{\text{ROI}}(i,j) + (1-\lambda_1) r_{\text{NROI}}(i,j) \quad (6)$$

参考图像的兴趣区域可由式(7)表示, λ_2 为加权因子

$$R_{\text{ROI}}(i,j) = \lambda_2 r_{\text{IR_ROI}}(i,j) + (1-\lambda_2) \cdot r_{\text{VISIBLE_ROI}}(i,j) \quad (7)$$

参考图像剩余区域可由式(8)表示, λ_3 为加权因子

$$R_{\text{NROI}}(i,j) = \lambda_3 r_{\text{IR_NROI}}(i,j) + (1-\lambda_3) \cdot r_{\text{VISIBLE_NROI}}(i,j) \quad (8)$$

参考图像块和融合图像块的结构相似性可由式(9)

表示

$$\begin{aligned} \text{SSIM}_{r,f}(i,j) &= \lambda_1 \text{SSIM}_{\text{ROI},F}(i,j) + (1-\lambda_1) \cdot \\ &\text{SSIM}_{\text{NROI},F}(i,j) = \lambda_1 (\lambda_2 \text{SSIM}_{\text{IR_ROI},F}(i,j) + \\ &(1-\lambda_2) \text{SSIM}_{\text{VISIBLE_ROI},F}(i,j)) + (1-\lambda_1) \cdot \\ &(\lambda_3 \text{SSIM}_{\text{IR_NROI},F}(i,j) + (1-\lambda_3) \cdot \\ &\text{SSIM}_{\text{VISIBLE_NROI},F}(i,j)) \end{aligned} \quad (9)$$

则 R_SSIM 可由式(10)表示.

$$\text{MSSIM}(r,f) = \left(\frac{1}{NM} \right) \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} \text{SSIM}_{R,F}(i,j) \quad (10)$$

图 1 是 R_SSIM 评价模型实现框图.

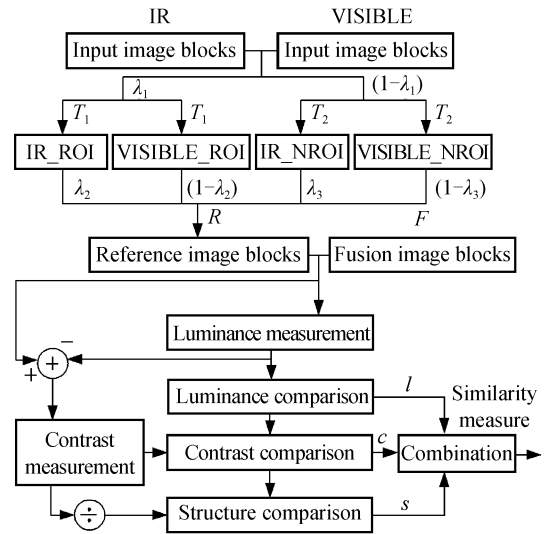


图 1 R_SSIM 评价模型实现框图

Fig. 1 Flow chart of R_SSIM assessment model

2.3 参量确定方法

R_SSIM 模型需确定的参量包括 $\lambda_1, \lambda_2, \lambda_3, T_1, T_2$. λ_1 表示实际应用的融合系统对于 ROI 的重视程度, λ_2 表示对红外图像或可见光图像 ROI 的重视程度, 而 λ_3 表示对红外图像或可见光图像 NROI 的重视程度. T_1, T_2 分别表示在对红外图像和可见光图像进行 ROI 和 NROI 图像分割时的分割阈值. 一般情况下, 对于 ROI 的重视程度应高于 NROI, 即 $\lambda_1 > 0.5$. 对红外图像 ROI 的重视程度应高于可见光图像, 即 $\lambda_2 > 0.5$. 对可见光图像 NROI 的重视程度应高于红外图像, 即 $\lambda_3 < 0.5$. 综合考虑红外图像和可见光图像的成像特性, 两种图像灰度值反差较大区域应为融合图像需重点突出的图像细节或重要特征, 式(11)为 T_1, T_2 的确定方法, 其中 T 为设定的阈值.

$$T_i = \text{abs}((\text{Gray_value}_{\text{IR}} - \text{Gray_value}_{\text{VISIBLE}})) > T \quad (11)$$

3 实验结果与分析

3.1 主观评价

利用麻省理工学院林肯实验室、荷兰 TNO 研究所、美国海军研究生院在网上提供的可见光和

红外图像,8个不同场景下包括车辆、人员、道路、房屋、舰船、营地、林地等不同类型目标.融合算法分别采用取大加权法(Maximum Weighing Method, MAX)、主成份分析方法(Principal Component Analysis, PCA)、拉普拉斯金字塔(Laplacian Pyramid, LAP)、梯度金字塔(Gradient Pyramid,

GRA)、离散小波变换(Discrete Wavelet Transform, DWT)、连续小波变换(Continuous Wavelet Transform, CWT).两种金字塔算法均采用4层结构、低通取平均、高通取大策略.图2是采用上述融合算法获得的融合图像(限于篇幅,其他场景的融合图像不再列出).

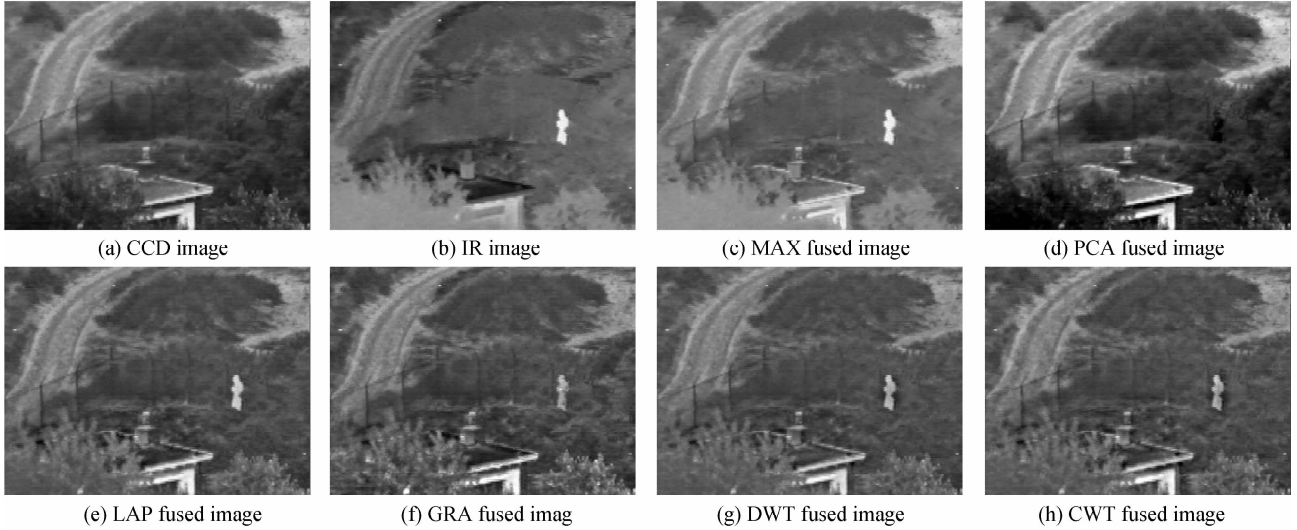


图2 采用不同融合算法获得的融合图像

Fig.2 Different fusion images by different fusion algorithms

主观评价测试人员由实验室研究生共6人组成,他们有正常或经过校正有正常的视力并了解有关图像专业知识和夜视成像的相关内容.打分采用相对评价规则,即由观察者将一批图像由好到坏进行分类并给出分数.在无法确定两幅图像质量优劣时,尽可能采用相同分数来避免分数的不确定性.表1是相对评价尺度与对应的分值.

表1 相对评价尺度与分值

Table 1 Relative assessment metric and score value

Subjective score	7	6	5	4	3	2	1
Image quality	Good		Fair		Poor		

对8幅输入图像分别采用MAX、PCA、LAP、GRA、DWT、CWT进行融合处理,每幅输入图像可

表2 不同算法获得的融合图像的主观综合打分

Table 2 Integrated subjective score to fusion images obtained by different algorithms

	MAX	PCA	LAP	GRA	DWT	CWT
Pic1	3.3	3.7	4.0	4.6	4.3	3.5
Pic2	2.0	2.7	4.3	5.0	5.0	3.7
Pic3	2.0	4.3	3.5	5.5	4.2	3.6
Pic4	3.3	2.8	3.7	5.3	4.0	4.3
Pic5	3.6	2.0	3.3	5.0	4.5	4.1
Pic6	1.8	2.8	3.8	5.2	4.0	4.8
Pic7	2.5	2.3	3.8	5.2	5.8	5.2
Pic8	3.5	2.0	3.7	5.0	4.8	4.7

获得6幅经过不同算法处理后的融合图像,将6名观察者对每幅图像的打分结果求均值,获得的主观综合打分列于表2.

3.2 R_SSIM 评价结果

图3是针对图2中的红外与可见光图像进行的ROI和NROI分割,分割阈值设定为100. λ_1 、 λ_2 均设定为0.75, λ_3 设定为0.25,利用(10)计算R_SSIM结果列于表3.由于图像分割时对剩余部分填充灰度值0,因此在计算R_SSIM时应减去图像剩余部分对结果的影响.

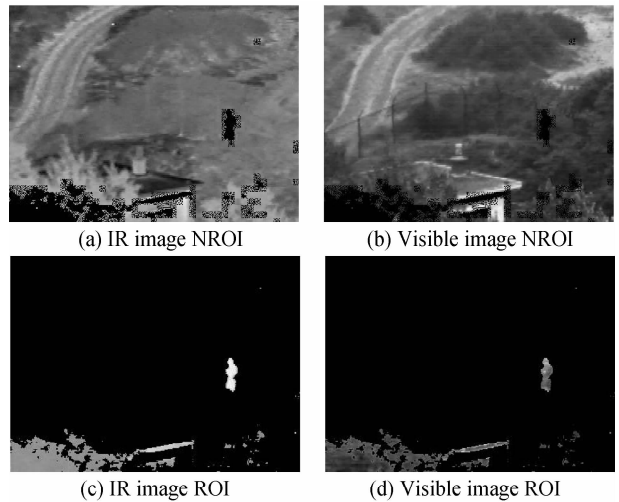


图3 红外与可见光图像感兴趣区域和剩余区域分割

Fig.3 Segmentation ROI and NROI to IR and visible images

表 3 基于 R_SSIM 的客观评价结果

Table 3 Objective assessment results based on R_SSIM

	MAX	PCA	LAP	GRA	DWT	CWT
Pic1	0.467 4	0.475 1	0.455 4	0.452 4	0.449 6	0.452 7
Pic2	0.549 7	0.617 9	0.531 3	0.333 9	0.497 5	0.510 9
Pic3	0.444 3	0.387 5	0.442 1	0.412 1	0.435 1	0.439 6
Pic4	0.464 1	0.470 5	0.499 5	0.472 3	0.491 4	0.495 3
Pic5	0.466 3	0.427 5	0.541 6	0.425 8	0.516 5	0.517 7
Pic6	0.445 7	0.505 4	0.514 0	0.481 8	0.511 3	0.516 0
Pic7	0.474 1	0.506 5	0.541 4	0.510 0	0.539 6	0.544 2
Pic8	0.481 4	0.403 5	0.483 6	0.428 1	0.469 7	0.478 6

3.3 结果分析

分析表 2 和表 3 可知,利用 R_SSIM 模型获得的评价结果和主观评价结果具有一致性.基于塔型结构的融合算法和基于小波变换的融合算法主观综合得分较高,验证了这两种融合算法针对不同场景的适应性较强,可在不同尺度、不同空间分辨率上有针对性的突出各种图像的重要特征和细节信息,提供给观察者更多的场景理解,而基于塔型结构的融合算法和基于小波变换的融合算法性能优劣很难依据平均得分做出客观准确的判断. MAX 和 PCA 算法针对不同场景获得的主观综合得分较低,同时针对不同场景的解释能力存在较大差异,说明两种算法针对不同场景的适应能力较差.值得关注的是,由于 MAX 融合算法能有效突出感兴趣区域的特征信息,同时算法实现简单,实时性好,因此在设计实际应用的融合系统时,不但要考虑不同成像方式和人类视觉因素的多样性,还要综合考虑融合系统的应用范围,根据任务需求选择合适的融合算法.需要说明的是,由于融合图像数据库数量有限,且主观评价过程受多方面因素影响,下步工作应着手构建具有典型代表性的主观评价数据库,同时 R_SSIM 模型的参量确定方法还需做进一步研究.

4 结论

图像融合质量评价方法是当前图像融合技术领域迫切需要解决的问题之一,构建适合于任务需求,结合主观和客观因素的图像融合评价方法是研究的重点.本文在分析现有客观评价方法基础上,提出的

基于结构相似度和感兴趣区域的融合图像质量评价方法,较好地利用了不同波段传感器输入图像的特征信息,且充分考虑了红外和可见光图像融合系统实际任务需求,实验结果也验证了其主观评价结果具有较好的一致性.

参考文献

- [1] TOET A, FRANKEN E M. Perceptual evaluation of different image fusion schemes[J]. *Displays*, 2003, **24**(1): 25-37.
- [2] PETROVIĆ V. Subjective tests for image fusion evaluation and objective metric validation [J]. *Information Fusion*, 2007, **8**(2): 208-216.
- [3] CUI Yan-mei, NI Guo-qiang, ZHONG Yan-li, *et al.* Analysis and evaluation of the effect of image fusion using statistics parameters[J]. *Journal of Beijing Institute of Technology*, 2000, **20**(1): 102-106.
崔岩梅, 倪国强, 钟堰利, 等. 利用统计特性进行图像融合效果分析及评价[J]. *北京理工大学学报*, 2000, **20**(1): 102-106.
- [4] QU Gui-hong, ZHANG Da-li, YAN Ping-fan. Information measure for performance of image fusion [J]. *Electronics Letters*, 2002, **38**(7): 313-315.
- [5] HOWELL C, MOORE R, BURKS S, *et al.* An evaluation of fusion algorithms using image fusion metrics and human identification performance [C]. *SPIE*, 2007: 65430V1-65430V11.
- [6] WANG Z, BOVIK A C, SHEIKH H R, *et al.* Image quality assessment: from error visibility to structural similarity[J]. *IEEE Transactions on Image Processing*, 2004, **13**(4): 600-612.
- [7] ZHANG Jing, SHEN Lan-sun, GAO Jing-jing. Regions of interest detection based on visual attention mechanism[J]. *Acta Photonica Sinica*, 2009, **38**(6): 1561-1565.
张菁, 沈兰荪, 高静静. 基于视觉注意机制的感兴趣区检测[J]. *光子学报*, 2009, **38**(6): 1561-1565.
- [8] XING Su-xia, CHEN Tian-hua. A new quality measure method for image fusion based on visual characteristics[J]. *Journal of Beijing Technology and Business University*, 2010, **28**(1): 60-63.
邢素霞, 陈天华. 基于视觉特性的图像融合质量评价新方法[J]. *北京工商大学学报*, 2010, **28**(1): 60-63.
- [9] DI Hong-wei, LIU Xian-feng. Image fusion quality assessment based on structural similarity [J]. *Acta Photonica Sinica*, 2006, **35**(5): 766-771.
狄红卫, 刘显峰. 基于结构相似度的图像融合质量评价[J]. *光子学报*, 2006, **35**(5): 766-771.
- [10] PIELLA G. A general framework for multiresolution image fusion: from pixels to regions [J]. *Information Fusion*, 2003, **4**(4): 259-280.

Image Fusion Assessment Method Based on Structural Similarity and Region of Interest

ZHANG Yong^{1,2}, JIN Wei-qi¹

(1 *MoE Key Laboratory of Photoelectronic Imaging Technology and System, School of Optoelectronics, Beijing Institute of Technology, Beijing 100081, China*)

(2 *Mechanical Engineering Institute, Mechanical Engineering College, Shijiazhuang 050000, China*)

Abstract: Aiming to infrared and visible image fusion effect evaluating question, an assessment method bonding structural similarity and region of interest was proposed based on structural similarity assessment method and human visual characteristics. The method divided region of interest and residual region from infrared and visible image respectively based on different image-forming feature, which indicate the different image characteristics. Then the different weight coefficients were endowed to different image regions based on the visual sensitive level. Compared with the convenient evaluating method, the proposed method needs no standard reference image participation and upgrades the weight value of image important characteristics in assessment result. Experiment results indicate that the proposed methods is more consistent with subjective evaluation results.

Key words: Image fusion; Structural similarity; Region Of Interest(ROI); Assessment