

基于结构相似度的图像融合质量评价*

狄红卫 刘显峰

(暨南大学光电工程研究所, 广州 510632)

摘 要 在分析现有图像融合质量评价方法特点的基础上, 提出了新型的、基于结构相似度的图像融合质量评价方法. 针对不同情况, 分别采用平均结构相似度、加权平均结构相似度、结构信息与交互信息量之乘积作为图像融合质量客观评价标准. 该方法充分考虑了图像的结构信息和人类视觉系统的特性, 可以为不同场合下选择不同的算法提供依据. 对不同融合算法的质量评价结果表明, 该方法是一种有效的图像融合质量评价方法.

关键词 图像融合; 结构相似度; 图像质量评价; 人眼视觉系统

中图分类号 TP391 **文献标识码** A

0 引言

信息融合理论和技术正成为信息和信号处理领域中的研究热点, 而图像融合作为信息融合的一个重要领域, 已经广泛应用于遥感、计算机视觉、医学、军事目标探测和识别等领域. 所谓图像融合是指综合两个或多个源图像的信息, 以获取对同一场景的更为精确、全面、可靠的图像描述.

近年来, 有不少学者提出了各种图像融合的方法^[1,2], 如主成分分析方法(PCA)、基于金字塔分解的算法、小波变换方法等, 然而对图像融合效果的评价问题却一直没有得到很好的解决. 原因是同一融合算法对不同类型的图像, 其融合效果不同; 同一融合算法, 对同一图像观察者, 感兴趣的部分不同, 则认为效果不同; 不同的应用方面, 对图像各项参量的要求不同, 导致选取的评价方法不同. 目前评价图像融合效果的方法主要可以分为两类^[3,4], 即主观评价方法和客观评价方法. 主观评价方法是由判读员直接对图像进行质量评价, 具有简单直观等特点, 然而在人为评价融合方法的过程中, 会有很多主观因素影响评价结果. 现有常用的客观评价方法最大问题在于没有充分考虑人眼视觉系统的特点, 这对于以人眼为最终信宿的图像来说是很不合理的. 因此, 研究基于人眼视觉系统的图像融合评价标准就显得极为重要. 本文在比较目前常用的图像融合质量评价方法的基础上, 充分考虑到图像的结构信息和人类视觉系统的特性, 提出了一种新型的、基于结构相似度的图像融合质量评价准则. 对不同融合算法的质量评价结果表明该方法是一种有效的图像融合质量评价方法.

1 图像融合质量的客观评价

为了客观、定量地评价融合图像的效果, 人们提出了一些不受人为影响的客观评价方法^[5~7]. 根据评定方法需要条件的不同, 目前通常采用的客观评价法主要分为三大类: 1) 根据单个图像统计特征的评定方法; 如信息熵、标准差; 2) 根据融合图像与标准参考图像关系的评定方法, 如均方误差、峰值信噪比; 3) 根据融合图像与源图像关系的评定方法, 如交叉熵、交互信息量. 在下面的讨论中, 设待融合的源图像为 A 、 B , 标准参考图像为 R , 实际融合图像为 F , 图像大小为 $M \times N$, L 为图像总的灰度级.

1.1 信息熵

融合后图像的熵值大小反映了融合图像所包含的平均信息量的多少. 图像的熵定义为

$$H = - \sum_{i=0}^{L-1} p_i \log_2 p_i \quad (1)$$

式中, H 为图像的熵, p_i 表示灰度值为 i 的像素数 N_i 与图像总像素数 N 之比. 如果融合图像的熵越大, 表示融合图像的信息量增加, 融合图像所含的信息越丰富, 融合质量越好.

1.2 标准差

标准差反映了图像灰度相对于灰度平均值的离散情况, 在某种程度上, 标准差也可用来评价图像反差的大小. 若标准差大, 则图像灰度级分布分散, 图像的反差大, 可以看出更多的信息. 标准差小, 图像反差小, 对比度不大, 色调单一均匀, 看不出太多的信息. 图像的标准差定义为

$$\sigma = \sqrt{\frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [F(x, y) - \bar{F}]^2}{M \times N}} \quad (2)$$

式中 \bar{F} 为融合图像的均值.

1.3 均方误差

均方误差用来评价融合图像与标准参考图像之

* 广东省自然科学基金(04010465)资助项目

Tel: 020-85220484-412 Email: tdhw@jnu.edu.cn

收稿日期: 2006-01-04

间的差异程度. 如果差异小, 则表明融合的效果好. 均方误差定义为

$$\text{MSE} = \frac{\sum_{x=0}^{M-1} \sum_{y=0}^{N-1} [F(x, y) - R(x, y)]^2}{M \times N} \quad (3)$$

1.4 峰值信噪比

峰值信噪比定义为

$$\text{PSNR} = 10 \lg \frac{255^2}{\text{MSE}} \quad (4)$$

峰值信噪比越高, 说明融合效果和质量越好.

1.5 交叉熵

交叉熵亦称相对熵, 用来度量两幅图像间的差异. 交叉熵越小, 则融合图像从源图像提取的信息量越多, 说明融合后图像与标准参考图像间的差异越小, 即融合效果越好. 设源图像和融合图像的灰度分布分别为: $p = \{p_0, p_1, \dots, p_i, \dots, p_{L-1}\}$ 和 $q = \{q_0, q_1, \dots, q_i, \dots, q_{L-1}\}$, 则交叉熵定义为

$$C = \sum_{i=0}^{L-1} p_i \log_2 \frac{p_i}{q_i} \quad (5)$$

1.6 交互信息量

交互信息量是信息论中的一个重要概念, 它可作为两个变量之间相关性的量度, 或一个变量包含另一个变量的信息量的量度. 这里利用这个概念来衡量融合图像与源图像的交互信息, 从而评价融合的效果. F 与 A 、 B 的交互信息量分别表示为

$$MI_{FA} = \sum_{k=0}^{L-1} \sum_{i=0}^{L-1} p_{FA}(k, i) \log_2 \frac{p_{FA}(k, i)}{p_F(k) p_A(i)} \quad (6)$$

$$MI_{FB} = \sum_{k=0}^{L-1} \sum_{j=0}^{L-1} p_{FB}(k, j) \log_2 \frac{p_{FB}(k, j)}{p_F(k) p_B(j)} \quad (7)$$

式中, p_A 、 p_B 和 p_F 分别是 A 、 B 、 F 的灰度直方图; $p_{FA}(k, i)$ 和 $p_{FB}(i, j)$ 分别代表两组图像的归一化联合灰度直方图. 综合考虑这两个数值量, 可以取

$$MI_F^{AB} = MI_{FA} + MI_{FB} \quad (8)$$

来表示融合图像 F 包含源图像 A 和 B 的交互信息量总和. 交互信息量是反应融合效果的一种客观指标, 它的值越大, 表示融合图像从源图像中获取的信息越丰富, 融合效果越好. 它可以更准确地评价各种融合方法的优劣.

上述图像融合质量评价方法或者是基于像素点间的误差, 或者是依据图像的统计特性, 都没有考虑图像的整体结构. 许多融合应用中, 最终的用户都是人, 人眼的视觉特性是非常重要的考虑因素, 而上述指标往往无法反映人的视觉特性. 因此, 建立新型的、结合主观与客观因素的图像融合质量评价准则对于图像融合的进一步研究非常重要.

2 图像结构相似度

传统的图像质量客观评价方法都是基于像素点

间的误差, 即基于误差敏感的图像质量评价, 如 MSE、PSNR. 误差敏感评价方法在图像质量的定义上直接把像素误差看成质量的亏损. 而事实上, 人眼是一个复杂的具有自适应调节能力的光学神经系统, 它响应外界光信号刺激形成视觉, 对图像信号的处理并不是逐点进行的. 同时, 人的视觉感知特点与统计意义上的信息分布并不一致, 即在统计上需要更多信息量才能表征的特征对视觉感知可能并不重要.

图像具有很强的结构性. 像素之间存在很强的相关性, 而且这些相关性携带着图像结构的重要信息. Wang Zhou 和 Bovik 等人在多年对图像处理、图像压缩及图像的视觉质量评价的研究工作基础上, 在 2002 年首次提出了结构信息的概念^[8]. 作者认为: 人眼视觉的主要功能是提取背景中的结构信息, 而且人眼视觉系统能高度自适应地实现这一目标, 因此对图像的结构失真的度量应是图像感知质量的最好近似. 在此基础上给出了一种符合人眼视觉系统特性的图像质量客观评判标准: 结构相似度^[9].

设 X 、 Y 表示待比较的两图像, 则它们之间的结构相似度 SSIM(X , Y) 定义为

$$\text{SSIM}(X, Y) = l(X, Y) \cdot c(X, Y) \cdot s(X, Y) = \frac{(2\mu_X\mu_Y + C_1)(2\sigma_{XY} + C_2)}{(\mu_X^2 + \mu_Y^2 + C_1)(\sigma_X^2 + \sigma_Y^2 + C_2)} \quad (9)$$

式中

$$l(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}$$

$$c(X, Y) = \frac{2\sigma_{XY} + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}$$

$$s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3}$$

式中, μ_X 、 μ_Y 、 σ_X^2 、 σ_Y^2 、 σ_{XY} 分别表示图像 X 、 Y 的均值、方差、协方差. $l(X, Y)$ 、 $c(X, Y)$ 、 $s(X, Y)$ 分别是亮度比较、对比度比较、结构比较. C_1 、 C_2 、 $C_3 = C_2/2$ 为小的常量, 以避免式(9)中分母为零而出现的的不稳定.

图 1 是对图像质量进行评价的对比实验. (a) 为原始图像, (b) 为叠加高斯噪声的图像, (c) 为模糊处理的图像, 这里 MSSIM 利用式(10)计算. 如果采用 MSE 作为图像质量评价准则, (b)、(c) 效果一样, 而采用结构相似度就可以区分二者之间的差别. 由此可见, 结构相似度可以更好地和主观评判标准相吻合.

人眼在观察图像时, 实际上是提取了图像的结构信息, 而不是图像像素点间的误差, 结构失真是图像质量评估中至关重要的因素. 这种观点无疑为目前的图像质量研究指出了一个新的方向. 应该说, 像点对像点的误差是最基本的失真, 而点与点组成

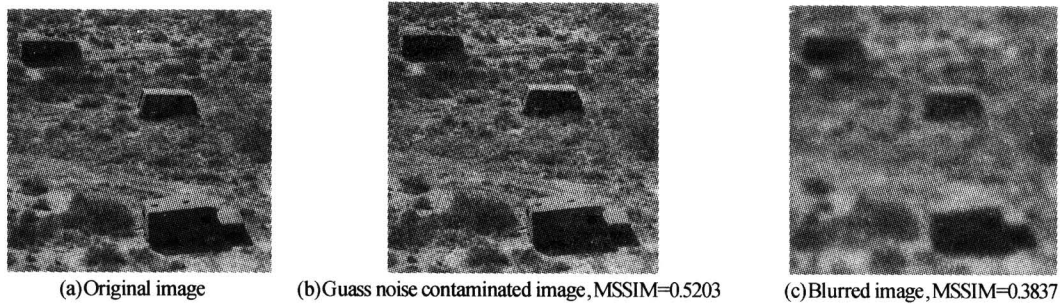


图1 不同失真对比实验(MSE=311)

Fig. 1 Comparison experiment with different distortions (MSE=311)

的结构或者说图像内容则是更高层次的东西,将涉及到图像理解、识别等研究分支。

3 基于结构相似度的图像融合质量评价

图像融合的目的是为了更好地理解、识别图像,而结构信息直接与图像内容相关,因此,本文提出以结构相似度作为图像融合质量评价标准。

3.1 根据融合图像与标准参考图像关系的评价方法

该方法考虑存在标准参考图像,多聚焦图像融合一般适用这一评价方法。

3.1.1 基于平均结构相似度的图像融合质量评价

首先利用滑动窗口操作将标准参考图像 R 、融合后图像 F 以相同大小的窗口进行互不重叠的分块,令分块总数为 T ,利用式(9)分别计算对应块的结构相似度 $SSIM(X_j, Y_j)$,最后按式(10)进行平均求和得到融合图像质量的评价值 $MSSIM$ 。

$$MSSIM(F, R) = \frac{1}{T} \sum_{j=1}^T SSIM(X_j, Y_j) \quad (10)$$

结构信息是图像中客观存在的特性,是图像中各像素值之间的相互关系,它更加注重参考图像与融合图像内在规律的相似性,这是对人眼系统特性的一个更高层次的理解。

3.1.2 基于加权平均结构相似度的图像融合质量评价

式(10)计算平均结构相似度时每个窗口的贡献是对等的,这跟人眼的主观视觉结果有出入。事实上,人眼对图像的不同部分敏感度是不同的。对人眼视觉系统的研究表明,人眼对边缘及条状结构比较敏感,尤其对边缘、条状结构的位置和方向特别敏感。人眼视觉的主要作用是感知形状,人眼通过追踪未知物体的轮廓来扫视一个未知的物体。因此,对每个窗口赋予不同的权值,从而定义基于加权平均结构相似度的图像融合质量评价指标。

如评价方法一,先对标准参考图像 R 、融合后图像 F 分块,计算每一块的结构相似度 $SSIM(X_j, Y_j)$ 。

然后计算标准参考图像 R 中每一块的方差 D_j ,选取权系数为 $w_j = D_j / (\sum_{j=1}^T D_j)$ 。这样,可以得到基于加权平均结构相似度的图像融合质量评价价值 $WMSSIM$ 。

$$WMSSIM(F, R) = \frac{1}{T} \sum_{j=1}^T [w_j \cdot SSIM(X_j, Y_j)] \quad (11)$$

3.2 根据融合图像与源图像关系的评定方法

该方法考虑没有标准参考图像的融合评价,这是更一般的情形。

在图像融合问题中,许多情况无法得到标准参考图像,融合图像往往是通过两幅或以上源图像获取。这时,评价图像融合效果的一个重要指标是看融合图像从源图像中获得了多少信息。交互信息量作为信息论中的一个重要概念,可作为两个变量之间相关性的量度,或一个变量包含另一个变量的信息量的量度。在评价提高信息量的图像融合算法中,交互信息量是很适合的。然而,作为一个基于统计信息的参量,交互信息量丢失了图像的重要信息:结构信息。

为此,提出以结构信息与交互信息量之乘积作为图像融合质量评价指标。这样,源图像 A 、 B 与融合图像 F 之间的质量评价价值定义为

$$MMS_{FA} = MI_{FA} \cdot MSSIM(F, A) \quad (12)$$

$$MMS_{FB} = MI_{FB} \cdot MSSIM(F, B) \quad (13)$$

综合考虑这两个数值量,可以取

$$MMS_F^{AB} = MMS_{FA} + MMS_{FB} \quad (14)$$

式(12)、(13)中,如果采用加权平均结构相似度代替平均结构相似度,则可以得到 $MWMS_{FA}$ 、 $MWMS_{FB}$ 、 $MWMS_F^{AB}$ 。

4 实验结果与讨论

为了验证本文提出的图像融合质量评价方法的正确性和有效性,进行了仿真实验。图2是多聚焦图像的融合实验。其中(a)为标准参考图像,对(a)进行处理,模拟摄影时的离焦现象,得到前景清晰、远景模糊的图像(b)和前景模糊、远景清晰的图像(c)。对(b)、(c)分别采用不同的融合算法得到融合

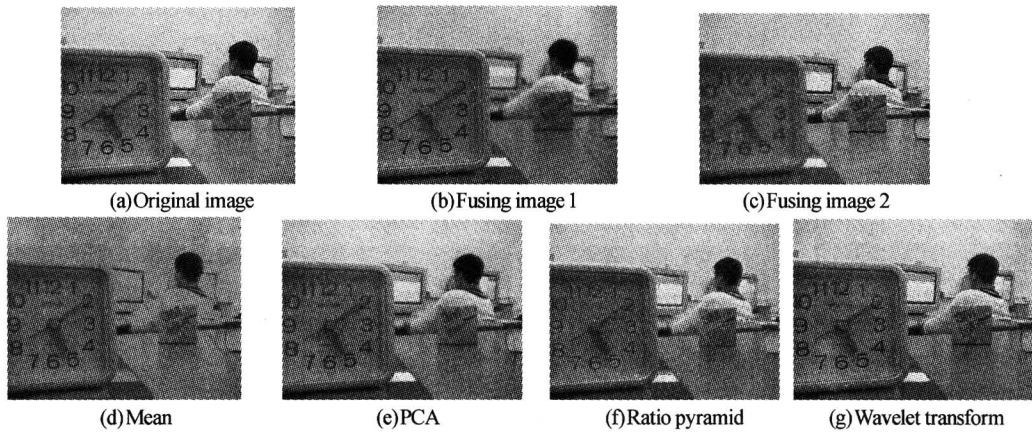


图2 多聚焦图像融合实验结果

Fig. 2 Experimental results with different fusion methods of multi-focus images

结果. (d)、(e)、(f)、(g) 分别为均值方法、PCA 算法、基于比率金字塔算法和基于小波变换方法的融合图像.

对各种融合算法的图像质量评价结果如表 1 和表 2. 其中在式(9)中, $C_1 = (k_1 \times L)^2$, $C_2 = (k_2 \times L)^2$, $k_1 = 0.01$, $k_2 = 0.03$, $L = 255$.

表 1 不同算法的融合质量评价结果 1(存在标准参考图像)

	Mean	PCA	Ratio Pyramid	Wavelet Transform
MSSIM	0.8737	0.9436	0.9206	0.9736
WMSSIM	0.6666	0.8753	0.8178	0.9803
PSNR	18.1639	30.9933	27.0273	34.8419

表 2 不同算法的融合质量评价结果 2 (不存在标准参考图像)

	Mean	PCA	Ratio Pyramid	Wavelet Transform
MI_F^{AB}	4.9517	7.1224	7.0124	7.0627
$WMSSIM_F^{AB}$	1.3557	1.7202	1.6129	1.8530
$MWMS_F^{AB}$	3.3558	6.1264	5.6553	6.5430

表 1 为考虑存在标准参考图像的情况. 分别用峰值信噪比 PSNR、平均结构相似度 MSSIM、加权平均结构相似度 WMSSIM 进行评价. 总体来说, 三种评价指标结果一致, 说明基于结构相似度评价方法的有效性. 同时, 图 1 的实验结果可知, 由于考虑了图像的结构信息和人眼的视觉特性, 基于结构相似度的评价方法是更可靠的评价标准. WMSSIM 由于强调了人眼对图像的不同部分敏感度的不同, 因而比评价指标 MSSIM 更具可区分性, 如对于均值融合方法和 PCA 方法, MSSIM 评价中, 评价值之差为 0.0699, 而 WMSSIM 评价中, 评价值之差为 0.2087.

表 2 为考虑没有标准参考图像的情况. 分别用交互信息量 MI_F^{AB} 、融合图像与两源图像结构相似度

之和 $WMSSIM_F^{AB}$ 、结构信息与交互信息量之乘积 $MWMS_F^{AB}$ 进行评价. 从 PCA 方法与小波变换方法看, MI_F^{AB} 的结果与表 1 的结果不一致, 这是由于交互信息量评价方法是基于统计信息的评价, 丢失了图像的整体结构信息, 并且没有考虑人眼的视觉特性. 而结构信息与交互信息量之乘积兼顾了图像的结构信息、人眼的视觉特性以及融合图像包含源图像的信息量程度, 因而可以更有效地评价融合图像的质量.

图 3 是多传感器图像的融合结果. 其中(a)、(b)为待融合的源图像, (a)为可见光图像, (b)为红外图像, 图中右边人身上藏有一手枪. 对(a)、(b)分别采用不同的融合算法得到融合结果. (c)、(d)、(e)、(f)分别为 PCA 算法、基于拉普拉斯金字塔算法、基于比率金字塔算法和基于小波变换方法的融合图像. 表 3 是对各种融合算法的图像质量评价结果.

表 3 不同算法的融合质量评价结果 3(不存在标准参考图像)

	PCA	Laplace Pyramid	Ratio Pyramid	Wavelet Transform
MI_F^{AB}	1.7539	1.6123	1.4263	2.1849
$WMSSIM_F^{AB}$	1.2022	1.7072	1.4208	1.4864
$MWMS_F^{AB}$	1.3732	1.3784	1.0132	1.6171

从图 3 可以看出, 基于比率金字塔融合算法受噪声的影响大, 融合质量差. PCA 算法融合结果包含(b)的信息很少, 隐藏的手枪几乎看不到. 基于小波变换方法的融合结果最好. 然而, 从表 3 来看, PCA 融合算法得到的交互信息量 MI_F^{AB} 很大, 基于拉普拉斯金字塔融合算法得到的结构相似度 $WMSSIM_F^{AB}$ 最大, 这与主观评价不一致. 而结构信息与交互信息量之乘积 $MWMS_F^{AB}$ 正确反应了图像的主观质量.

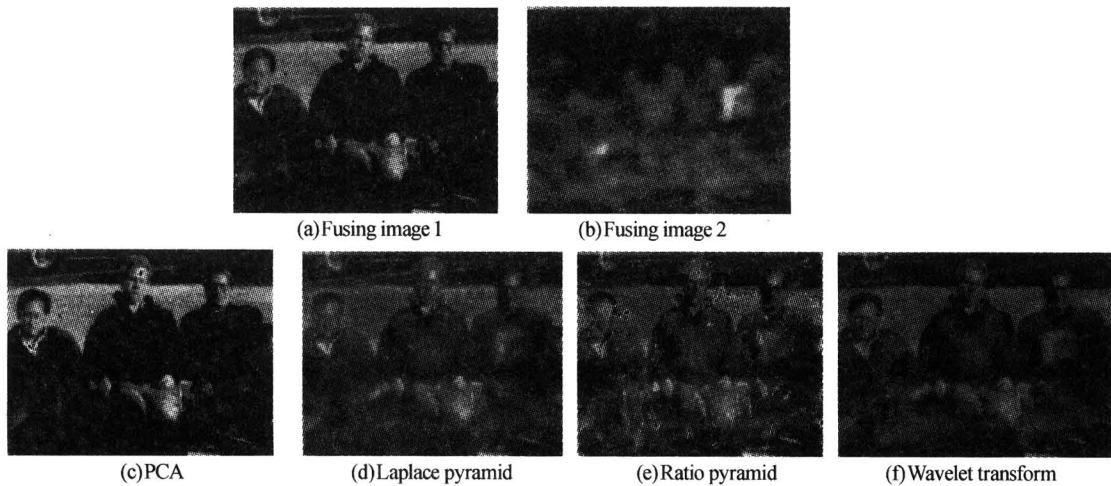


图3 多传感器图像融合实验结果

Fig. 3 Experimental results with different fusion methods of multi-sensor images

5 结论

建立图像融合质量评价标准是目前图像融合中急需解决的问题之一,通用的、结合主观和客观因素的图像融合质量的评价方法是未来研究的重点. 本文在分析现有图像融合质量评价方法的基础上,通过引入图像结构信息,提出基于结构相似度的图像融合质量评价方法. 这种新颖的评价方法为目前的图像融合质量研究提供了一个新的方向. 针对是否存在标准参考图像,分别提出平均结构相似度、加权平均结构相似度、结构信息与交互信息量之乘积作为图像融合质量客观评价标准. 提出的方法充分考虑了人类的视觉特性和图像之间的相关性,因而与主观评价结果一致. 实验结果表明提出的方法是一种有效的、正确的融合图像质量评价方法.

参考文献

- 1 那彦,史林,杨万海. 小波包变换与遥感图像融合. 光子学报,2004,**33**(6):736~738
Na Y, Shi L, Yang W H. *Acta Photonica Sinica*, 2004, **33**(6):736~738
- 2 李晖晖,郭雷,刘航. 基于区域分割的遥感图像融合方法. 光子学报,2005,**34**(12):1901~1905
Li H H, Guo L, Liu H. *Acta Photonica Sinica*, 2005, **34**(12):1901~1905
- 3 刘贵喜,杨万海. 基于小波分解的图像融合方法及性能评价. 自动化学报,2002,**28**(6):927~934
Liu G X, Yang W H. *Acta Automatica Sinica*, 2002, **28**(6):927~934
- 4 狄红卫,陈木生. 一种自适应的多光谱图像与全光图像融合新方法. 光子学报,2005,**34**(3):452~454
Di H W, Chen M S. *Acta Photonica Sinica*, 2005, **34**(3):452~454
- 5 王海晖,彭嘉雄,吴巍. 采用交互信息量评价遥感图像融合结果的方法. 华中科技大学学报(自然科学版),2003,**31**(12):32~34
Wang H H, Peng J X, Wu W. *J Huazhong Univ of Sci & Tech. (Nature Science Edition)*, 2003, **31**(12):32~34
- 6 杨静,王岩飞,刘波. 一种新的非抽取提升结构小波变换图象融合算法. 光子学报,2004,**33**(6):728~731
Yang J, Wang Y F, Liu B. *Acta Photonica Sinica*, 2004, **33**(6):728~731
- 7 胡良梅,高隽,何柯峰. 图像融合质量评价方法的研究. 电子学报,2004,**32**(12A):218~221
Hu L M, Gao J, He K F. *Acta Electronica Sinica*, 2004, **32**(12A):218~221
- 8 Wang Zhou, Bovik A C. A universal image quality index. *IEEE Signal Processing Letters*, 2002, **9**(3):81~84
- 9 Wang Zhou, Bovik A C, Sheikh H R, et al. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 2004, **13**(4):600~612

Image Fusion Quality Assessment Based on Structural Similarity

Di Hongwei, Liu Xianfeng

Institute of Optoelectronic Engineering, Jinan University, Guangzhou 510632

Received date: 2006-01-04

Abstract How to assess the image quality is one of the most important problems in image fusion. By analyzing the existent assessments of image fusion quality, a novel assessment of image fusion quality based on structural similarity is presented. According to different cases, average structural similarity, weighted average structural similarity, product of structural similarity and mutual information are introduced to assess image fusion quality as the objective methods. This assessment considers the structural information of image and the characteristics of human visual system, and can be used to choose the best algorithm in image fusion applications. The experiments of different kinds of image fusion algorithms demonstrate that it is an effective assessment of image fusion quality.

Keywords Image fusion; Structural similarity; Image quality assessment; Human visual system



Di Hongwei was born in 1969. He received M. S. degree in Jinan University in 1994 and Ph. D. degree in South China University of Technology in 1999. Now he is an associate professor in Jinan University. His research interests include image processing and wavelet transform.