

图像融合质量评价方法的研究

胡良梅, 高 隼, 何柯峰

(合肥工业大学计算机与信息学院图像信息处理研究室, 安徽合肥 230009; 中科院合肥分院智能机械研究所, 安徽合肥 230031)

摘 要: 建立通用的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, 本文在分析现有图像融合质量评价方法特点的基础上, 讨论了一种新的结合主观和客观因素的图像融合质量评价方法. 该方法无需理想图像, 同时充分考虑了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. 对多类图像的不同融合算法的质量评价结果表明该方法是一种实用的、有效的、通用的图像融合质量评价方法.

关键词: 图像融合; 图像质量评价; 人类视觉系统

中图分类号: TP391 **文献标识码:** A **文章编号:** 0372-2112 (2004) 12A-218-04

Research on Quality Measures for Image Fusion

HU Liang-mei, GAO Jun, HE Ke-feng

(Institute of Computer and Information, Hefei University of Technology, Hefei, Anhui 230009, China;
Institute of Intelligent Machines, Chinese Academy of Sciences, Hefei, Anhui 230031, China)

Abstract: How to universally measure the quality of fused image is one of the most important problems in image fusion. This paper analyses the measures of fused image quality in existence, then presents a new evaluation of fused image, which combines objective and subjective factors. In addition, it considers the characteristics of human visual system without reference image and can be used to choose the best algorithm in image fusion applications. The experiments of different kinds of images and image fusion algorithms demonstrate that it is a practical, effective, and universal measure of quality of fused image.

Key words: image fusion; quality measure for image; human visual system

1 引言

图像融合是指把对同一目标或同一场景经多个传感器的成像或单一传感器的多次成像进行一定的处理, 生成一幅新的图像, 这幅新的图像能够准确地综合反映源图像的信息, 更适用于人眼观察或计算机处理. 图像融合能够提高系统的可靠性和信息的利用率, 已经广泛用于军事、遥感、机器视觉和医疗诊断等领域. 学术界在图像融合领域已取得了很大的成绩, 图像融合方法也各种各样. 但是, 图像融合技术的研究还刚刚开始, 有许多问题急需解决. 建立通用的图像融合质量评价标准就是其中的问题之一^[1-3].

当前图像融合效果的评价问题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, 对图像各项参数的要求不同, 导致选取的评价方法不同^[2, 14]. 目前评价图像融合效果的方法可分为两类, 即主观的评价方法和客观的评价方法. 在许多融合应用中, 最终的用户都是人, 人眼的视觉特性也是非常重要的考虑因素. 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. 同时, 由于图像融合往往作为特定任务的预处理部分, 因而融合性能的评价取决于能否提高后续任务的性能. 这就需要研究通用的、主观与客观因素相结合的图像融合质量的评价准则, 使计算机能够自动选取适合当前图像的、效果最佳的算法, 从而为不同场合下选择不同的算法提供依据.

本文第一部分分析比较了目前常用的图像融合质量的主观、客观评价方法, 第二部分讨论了一种新的结合主观、客观因素的图像融合评价方法, 第三部分对多类图像、多种融合算法的融合图像质量进行评价, 第四部分是对本文的一个总结.

2 目前常用的主观、客观评价方法

目前评价图像融合算法性能的方法可分为两类, 即主观的评价方法和客观的评价方法.

2.1 图像融合质量的主观评价方法

主观评价方法就是依靠人眼对融合图像效果进行主观判断. 例如, 找一些测试者, 让他们对用不同融合算法得到的融合图像中的特定目标进行识别, 测量出识别时间并统计出识别的正确率, 从而判断图像融合算法性能的优劣. 主观评价方

法简单方便,在一些特定应用中十分可行.在美国国防部高级研究计划局资助的先进夜视系统开发计划中,研究者就是用主观评价方法来比较两种假彩色图像融合方法的好坏.由于这套系统是用来提高飞行员的夜视能力,所以主观评价方法不失为一种最佳的选择^[8].

但是在人为评价的过程中有很多主观因素影响评价结果,这就需要给出客观的评价方法.

2.2 图像融合质量的客观评价方法

设源图像为 A 和 B , R 为理想融合图像, F 为实际融合图像.目前常用的客观评价指标主要包括以下几种^[1,4~7,9~12].

2.2.1 均方根误差

融合图像 F 与理想图像 R 的均方根误差为:

$$RMSE = \sqrt{\frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N [R(i,j) - F(i,j)]^2} \quad (1)$$

其中 M 和 N 为图像的尺寸,均方根误差越小,说明融合图像与理想图像越接近.

2.2.2 平均误差

两幅图像的灰度平均误差为:

$$= \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N |R(i,j) - F(i,j)| \quad (2)$$

各变量的定义与公式(1)中相同.

2.2.3 灰度标准差

设一幅图像的灰度分布为 $P = \{p(0), p(1), \dots, p(g), \dots, p(L-1)\}$, $p(g)$ 为灰度等于 g 的像素数与图像总的像素数的比值, L 为灰度级数,且 $\sum_{g=0}^{L-1} p(g) = 1$, 则该图像的灰度统计均值为:

$\bar{g} = \sum_{g=0}^{L-1} g \cdot p(g)$, 其灰度标准差的定义为:

$$s = \sqrt{\sum_{g=0}^{L-1} (g - \bar{g})^2 \cdot p(g)} \quad (3)$$

灰度标准差反映了相对灰度均值的离散状况,标准差越大,灰度分布越分散.

2.2.4 熵

图像的熵的大小反映了图像携带信息的多少,其定义为:

$$EN = - \sum_{g=0}^{L-1} p(g) \log_2 p(g) \quad (4)$$

2.2.5 熵差

两幅图像的熵差体现了图像信息量上的差异,其定义为:

$$EN = |EN_R - EN_F| \quad (5)$$

其中 EN_F 和 EN_R 分别为融合图像与理想图像的熵.

2.2.6 交叉熵

设理想图像和融合图像的概率分布分别为: $P = \{p(0), p(1), \dots, p(g), \dots, p(L-1)\}$, $Q = \{q(0), q(1), \dots, q(g), \dots, q(L-1)\}$ 交叉熵度量它们之间的信息差异,定义为:

$$CEN(P, Q) = \sum_{g=0}^{L-1} p(g) \log_2 \frac{p(g)}{q(g)} \quad (6)$$

2.2.7 互信息

图像间的互信息定义为:

$$MI = \sum_{i=1}^L \sum_{j=1}^L h_{R,F}(i,j) \log_2 \frac{h_{R,F}(i,j)}{h_R(i) + h_F(j)} \quad (7)$$

其中 $h_{R,F}(i,j)$ 为理想图像 R 和融合图像 F 间的归一化联合灰度直方图分布, $h_R(i)$ 和 $h_F(j)$ 分别为两幅源图像的归一化边缘直方图分布, L 为灰度级数.

分析上述图像融合质量的客观评价方法,可以看出:

(1) 很多指标的计算都需要理想图像,而在实际的图像融合应用问题中,往往无法获得理想图像.

(2) 上述指标都是针对特定融合目的和算法的.如均方根误差、平均误差、灰度标准差三个指标都是基于统计特性的,因而比较适合提高分辨率的图像融合算法中.而熵、熵差、交叉熵、互信息四个指标都是基于信息量的,比较适合评价提高信息量的图像融合算法,如图像传输和图像特征提取等应用中.

(3) 在许多融合应用中,最终的用户都是人,人眼的视觉特性也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.

这就需要给出通用的,主观与客观因素相结合的图像融合质量的评价准则,从而使计算机能够自动选取适合当前图像的、效果最佳的算法,为不同场合下选择不同的算法提供依据.下面就讨论一种新的结合主观和客观因素的图像融合质量评价方法.

3 通用的主观和客观因素相结合的图像融合评价方法

参文[13]提出了一种通用的图像质量评价指标,并比较了几种不同失真情况下的指标评价结果,指出该指标由于能够度量两幅图像结构上的失真,因而优于传统的均方根误差评价指标,且具有通用性,适用于评价不同的图像处理过程.

如果两幅图像 a, b 的大小都为 $M * N$, 用 \bar{a} 表示 a 的均值, σ_a^2 表示 a 的方差, σ_{ab} 表示 a 和 b 的协方差,即

$$\sigma_a^2 = \frac{1}{MN - 1} \sum_{m=1}^M \sum_{n=1}^N (a(m,n) - \bar{a})^2 \quad (8)$$

$$\sigma_{ab} = \frac{1}{MN - 1} \sum_{m=1}^M \sum_{n=1}^N (a(m,n) - \bar{a})(b(m,n) - \bar{b}) \quad (9)$$

则可以定义评价指标

$$Q_0(a, b) = 4 \frac{\sigma_{ab}}{((\sigma_a^2 + \sigma_b^2)(\frac{\sigma_a}{\sigma_a} + \frac{\sigma_b}{\sigma_b}))} \quad (10)$$

$Q_0(a, b)$ 可分解为三项:

$$Q_0(a, b) = \frac{\sigma_{ab}}{\sigma_a + \sigma_b} \frac{2\sigma_{ab}}{\sigma_a^2 + \sigma_b^2} \frac{2}{\frac{\sigma_a}{\sigma_a} + \frac{\sigma_b}{\sigma_b}} \quad (11)$$

实际上, $Q_0(a, b)$ 是图像 a 和 b 之间的结构化相似度的一种度量,其值在 0 和 1 之间.公式(11)中的第一项是 a 和 b 的相关系数,第二项是平均照度失真,其值在 0 和 1 之间,第三项是对比度失真,其值也在 0 和 1 之间.

在图像处理中,通常先分块计算局部区域的 Q_0 值,再合成一个总体指标.如采用滑动窗口的方法,先计算图像 a 和 b 中对应窗口中的指标 $Q_0(a, b|_w)$, 总体的指标就是这些局部指标的均值:

$$Q_0(a, b) = |W|^{-1} \sum_w Q_0(a, b|_w) \quad (12)$$

其中 w 是所有窗口的总和, $|W|$ 是 w 的集的势。

然而在图像融合问题中,由于融合图像往往是通过两幅源图像得到的,进行融合图像指标评价时,要考虑源图像 a, b 以及融合图像 f 这三者之间的关系,可以构造函数 $Q(a, b, f)$ 来评价融合图像的质量^[3],

$$Q(a, b, f) = aQ_0(a, f) + bQ_0(b, f) \quad (13)$$

$$a(\cdot) = \frac{s(|a|)}{s(|a|) + s(|b|)} \quad (14)$$

$$b(\cdot) = 1 - a(\cdot) \quad (15)$$

$s(|a|)$ 是源图像 a 中该窗口的某些显著特征,如方差、边缘信息等,即

$$Q(a, b, f) = \frac{1}{|W|} \left(a(\cdot)Q_0(a, f) + b(\cdot)Q_0(b, f) \right) \quad (16)$$

公式(16)可以综合评价源图像 a, b 以及融合图像 f 之间的结构相似度。然而在 $Q(a, b, f)$ 中源图像 a 和 b 中每个窗口的贡献是对等的,考虑到人类视觉系统比较重视图像中视觉感知上比较显著的区域,可以对每个窗口赋予不同的权值 $c(\cdot) = \max\{s(|a|), s(|b|)\}$,从而定义加权的融合质量评价指标:

$$Q_w(a, b, f) = \frac{c(\cdot)}{w} \left(a(\cdot)Q_0(a, f) + b(\cdot)Q_0(b, f) \right) \quad (17)$$

公式(17)中, $c(\cdot) = C(\cdot) / (C(\cdot))$ 。

另外还可以考虑人类视觉系统的其它特性,如对边缘信息的敏感性,用梯度范数 a 代替源图像中的灰度值 a ,相应的得到 $Q_w(a, b, f)$,从而定义基于边缘的评价指标:

$$Q_E(a, b, f) = Q_w(a, b, f)^{1-\alpha} + Q_w(a, b, f) \quad (18)$$

公式(18)中 α 表示边缘图像对原始图像的贡献,其值越大,边缘图像的贡献越大。

以上三个指标 Q, Q_w, Q_E 的值都在 $[-1, 1]$ 之间,越接近于 1 表示融合图像的质量越高。

4 实验结果

我们用 Q, Q_w, Q_E 这三个指标对多类图像的不同融合算法得到的融合图像质量进行了评价。

4.1 多聚焦图像的不同融合算法的质量评价结果比较

对两幅多聚焦的时钟图像(如图 1(a)和(b)所示)对分别采用均值、laplace 金字塔、对比度金字塔和离散小波变换四种融合算法得到的融合图像进行质量评价。其中 laplace 金字塔、对比度金字塔和离散小波变换都采用四层分解,各

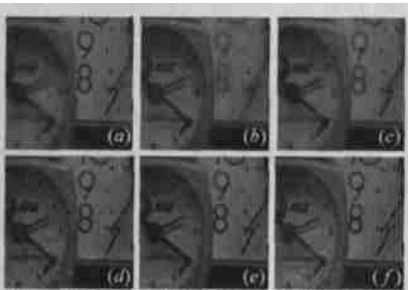


图 1 时钟图像的多融合方法结果比较:(a)(b)两幅源图像;(c)-(f)分别采用均值、laplace 金字塔、对比度金字塔、离散小波变换方法得到的融合图像

层采用绝对值最大的融合规则。

采用 Q, Q_w, Q_E 三个指标对各种融合算法的图像质量评价结果如表 1 所示。可以看出, Laplace 金字塔、对比度金字塔以及离散小波变换的融合图像的评价指标明显比均值方法的指标高,这与对图 1(c)(d)(e)(f)的主观判断评价结果一致。另外,加权的评价指标 Q_w 和基于边缘的评价指标 Q_E 由于强调了人眼的视觉特性,因而比评价指标 Q 更具区分性。

表 1 时钟图像不同融合算法的融合质量评价结果

融合算法	均值	Laplace 金字塔	对比度金字塔	离散小波变换
Q	0.8367	0.8558	0.8542	0.8170
Q_w	0.8664	0.9377	0.9368	0.9324
Q_E	0.6827	0.8424	0.8473	0.8326

4.2 医学图像的不同融合算法质量评价结果比较

图 2(a)和(b)分别为 CT 和核磁共振的医学图像。图 2(c)(d)(e)(f)分别为均值、laplace 金字塔、对比度金字塔和离散小波变换方法得到的融合图像。其中 laplace 金字塔、对比度金字塔和离散小波变换都采用四层分解,各层采用绝对值最大的融合规则。

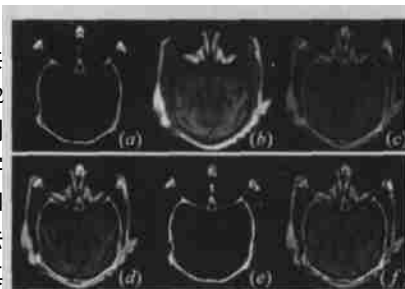


图 2 医学图像的多融合方法结果比较:(a)(b)CT 和核磁共振图像;(c)-(f)分别采用均值、laplace 金字塔、对比度金字塔、离散小波变换方法得到的融合图像

表 2 医学图像不同融合算法的融合质量评价结果

融合方法	均值	Laplace 金字塔	对比度金字塔	离散小波变换
Q	0.7885	0.6562	0.1358	0.6150
Q_w	0.8340	0.8276	0.4651	0.7861
Q_E	0.6745	0.6790	0.3397	0.6056

表 2 是采用以上三个指标对不同融合算法的融合图像质量评价的结果。从中可以看出,对于 CT 和核磁共振这两幅医学图像,采用最简单的均值算法得到的图像质量最高,而对比度金字塔的方法得到的图像质量最差,显然不适用于这两类图像的融合,这与对图 2(c)-(f)的融合图像的主观评价结果一致。

5 总结

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。本文分析了现有图像融合质量的主观评价方法和客观评价方法,讨论了一个新的结合主观和客观因素的图像融合质量的评价方法,该方法无需理想图像,同时充分考虑了人类的视觉特性,能够定量评价融合图像中保留了多少显著信息(如方差、对比度及边缘信息等),因而与主观评价结果一致。以均值、laplace 金字塔、对比度金字塔、

离散小波变换四种常用的融合算法为例进行了多类图像的融合质量评价,实验结果表明该方法得到的客观评价指标主观评价的结果是一致的,是一种实用的、有效的、通用的融合图像质量评价方法,可以为不同场合下选择不同的算法提供依据。

参考文献:

- [1] 李树涛,王耀南,张昌凡.多传感器图像融合的客观评价与分析[J].仪器仪表学报,2002,23(6):651-654.
- [2] 阳方林,郭红阳,杨风暴.像素级图像融合效果的评价方法研究[J].测试技术学报,2002,16(4):276-279.
- [3] Gemma Piella. New quality measures for image fusion[A]. The 7th International Conference on Information Fusion[C]. Stockholm, Sweden, June 28 to July 1, 2004. 542-546.
- [4] 李玲玲,周成平,丁明跃,张天序.基于易操纵金字塔的多传感器图像融合[J].计算机工程与应用,2003,22:24-26+41.
- [5] 姜庆娟,谭景信.像素级图像融合方法与选择[J].计算机工程与应用,2003,25:116-120.
- [6] 钟伟才,刘静,刘芳,焦李成.基于形态学 top-hat 算子的多传感器图像融合[J].电子学报,2003,31(9):1415-1417.
- [7] 洪靖云,张桂林,洪汉玉,杜峥.基于高频通道的图像融合算法[J].红外与激光工程,2003,32(5):509-512.
- [8] 毛士艺,赵巍.多传感器图像融合技术综述[J].北京航空航天大学学报,2002,28(5):512-518.
- [9] 夏明革,何友,唐小明,夏仕昌.多传感器图像融合综述[J].电光与控制,2002,9(4):1-7.
- [10] 吉书鹏,丁晓青.可见光与红外图像增强融合算法研究[J].红外与激光工程,2002,31(6):518-521+544.
- [11] 李玲玲,周成平,丁明跃,张天序.一种基于多尺度分解的多传感器图像融合方法[J].华中科技大学学报(自然科学版),2003,31(4):82-84.
- [12] 夏明革,何友,唐小明,关欣.像素级图像融合方法分类与比较[J].火力与指挥控制,2002,27(3):1-4.
- [13] Zhou Wang, Bovik A C. A universal image quality index[J]. IEEE Signal Processing Letters, 2002, 9(3):81-84.
- [14] Xydaes C, Petrovi V. Objective image fusion performance measure[J]. Electronic Letters, 2000, 36(4):308-309.

作者简介:

胡良梅 女,1974年4月出生于安徽合肥,合肥工业大学博士研究生,讲师,主要研究方向:信息融合、图像理解、智能信息处理. E-mail:liangmeihu@sohu.com.

高隽 男,1963年10月出生于安徽淮南,中国科学技术大学博士,合肥工业大学教授,博士生导师,主要研究方向:图像处理与分析、模式识别与人工智能、神经网络理论及应用、光电信息处理、智能信息处理.

(上接第 250 页)

- [8] Lu R, Jin Z. Domain Modeling-Based Software Engineering[M]. Kluwer Academic Publishers, 2001.
- [9] A Finkelstein J, Kramer, B Nuseibeh, L Finkelstein, M Gøedicke. Viewpoints: A framework for integrating multiple perspectives in system development [J]. International Journal of Software Engineering and Knowledge Engineering, 1992, 2(1), 31-58.
- [10] A Silva. Requirements, Domain and specifications: A viewpoint-based approach to requirements engineering[A]. Proceedings of ICSE02[C]. Orlando, Florida, USA, May, 2002. 94-104.
- [11] S Easterbrook, M Chechik. A framework for multivalued reasoning over inconsistent viewpoints[A]. In Proceedings of International Conference on Software Engineering (ICSE '01) [C]. Toronto, Canada, May 2001. 411-420.
- [12] 牟克典,金芝,陆汝钊.视点合成中重叠需求的不一致优先级处理[J].计算机学报,2004,27(10):1379-1387.
- Kedian Mu, Zhi Jin, Ruqian Lu. Handling inconsistent priorities of overlapped requirements in merging viewpoints[J]. Chinese Journal of Computers, 2004, 27(10):1379-1387.
- [13] B Nuseibeh, S Easterbrook, A Russo. Leveraging inconsistency in software development[J]. IEEE Computer, 2000, 33(4):24-29.
- [14] B Nuseibeh, S Easterbrook, A Russo. Making inconsistency respectable in software development[J]. Journal of Systems and Software, 2001, 58(2):171-180.
- [15] A Finkelstein, D Gabbay, A Hunter, J Kramer, B Nuseibeh. Inconsistency handling in multi-perspective specifications[J]. Transactions on Software Engineering, 1994, 20(8):569-578.
- [16] G Spanoudakis, A Zisman. Inconsistency management in software engineering: Survey and open research issues[A]. S K Chang, editor. Handbook of Software Engineering and Knowledge Engineering [C]. World Scientific Publishing Co., 2001. 329-380.
- [17] A Hunter, B Nuseibeh. Managing inconsistent specification [J]. ACM Transactions on Software Engineering and Methodology, 1998, 7(4):335-367.
- [18] Kedian Mu, Zhi Jin, Ruqian Lu. A logical approach to testing requirements[A]. to appear in Proceedings of IASTED conference on SE05 [C]. Zurich, Switzerland ACTA Press, 2005.