

基于稀疏表示的失真卫星立体图像全参考质量评价

熊义明, 邵枫*, 孟祥超

宁波大学信息科学与工程学院, 浙江 宁波 315211

摘要 针对建筑物检测特定应用, 提出一种基于稀疏表示的失真立体图像全参考质量评价方法。首先构建了一种新的失真卫星立体图像数据库, 使用角点检测和数字表面模型的高程信息进行建筑物检测, 并根据失真图像检测角点变化, 提出检测准确率指标来表示图像的失真程度; 然后提出一种基于稀疏表示的客观评价模型, 其分别提取原始图像和失真图像的尺度不变特征转换和二进制稳健不变尺度特征进行字典学习; 利用稀疏表示测量原始图像和失真图像之间的相似性, 得到 4 个质量分数; 最后通过支持向量回归融合 4 个质量分数得到最终的客观评价价值。在构建的数据库上进行测试, 实验结果表明, 皮尔逊线性相关系数值高于 0.90, 斯皮尔曼等级相关系数值高于 0.87, 与现有的评价方法相比, 所提方法能更好地反映卫星立体图像的质量。

关键词 图像处理; 全参考质量评价; 卫星立体图像; 建筑物检测; 角点检测; 数字表面模型高程信息; 稀疏表示
中图分类号 TP751.1 **文献标识码** A **doi**: 10.3788/AOS201838.1210002

Sparse Representation-Based Full-Reference Quality Assessment of Distorted Satellite Stereo Images

Xiong Yiming, Shao Feng*, Meng Xiangchao

Faculty of Information Science and Engineering, Ningbo University, Ningbo, Zhejiang 315211, China

Abstract Aiming at the specific application of building detection, a sparse representation-based full-reference quality assessment method for distorted satellite stereoscopic images is proposed. First, a new distorted satellite stereo image database is constructed, in which the corner detection and the digital surface model elevation information are used for building detection. And a detection accuracy index is proposed to represent the degree of distortion based on the change of the detected corners. Then, an objective evaluation model based on sparse representation is proposed, which extracts scale-invariant feature transforms and binary robust invariant scalable key points of the original and the distorted images for dictionary learning. Four quality scores are obtained using sparse representation to measure the similarity between the original and the distorted images. Finally, the final objective assessment value is obtained by fusing the four quality scores using support vector regression. The test is carried out on the constructed database. The test results on the constructed database show that the Pearson linear correlation coefficient is higher than 0.90, and the Spearman rank correlation coefficient is higher than 0.87. Compared with the existing assessment methods, the proposed objective evaluation method can better reflect the quality of satellite stereo images.

Key words image processing; full reference quality assessment; satellite stereo image; building detection; corner detection; digital surface model elevation information; sparse representation

OCIS codes 100.2000; 100.6890; 110.2970

1 引 言

近些年, 遥感卫星技术在国民经济各部门得到广泛应用。遥感卫星技术带动了相关产业的发展, 促进了遥感系统的正常运行, 并带来了巨大的经济效益和社会效益, 也为国家军事提供重要保障, 为人民提

供了更便利的生活^[1]。目前, 土地资源、海洋资源、农业、林业、水利、地震监测、石油、天然气应用和环境等部门都依赖于遥感卫星影像^[2]。同时, 随着 ZY1-02C、ZY-3、GF-1 等高分辨率遥感卫星的成功发射^[3], 高分辨率遥感卫星立体影像的数量急剧增加, 数据采集速度加快, 数据更新周期变短。但与采

收稿日期: 2018-07-10; 修回日期: 2018-08-06; 录用日期: 2018-08-13

基金项目: 国家自然科学基金(61622109)

* E-mail: shaofeng@nbu.edu.cn

集能力相比,评价遥感卫星立体数据能力相对较低,这将对从事于遥感技术相关工作者带来极大不便。因此,建立高效的遥感卫星立体图像质量评价模型十分必要。

目前,图像质量评价的研究主要集中于自然图像^[4-5]。尽管遥感传感器采集图像的技术不断升级,但在实际应用中,遥感图像质量评价还依赖遥感工作者主观筛选,这将很难去判断图像在特定应用中的质量。Hasanlou 等^[6]提出几何指标(RG)预测融合后的图像质量,用多种方法融合多光谱图像和全色图像,提取融合后的图像光谱信息和空间分量,构成 RG 来评价融合后的图像质量。Kwan 等^[7]用 4 种融合方法生成高分辨率短波红外(HR-SWIR)图像,并提出一种无参考广义质量(GQNR)评价模型,提取融合后图像各个波段的自然场景特征,定义新的评价度量指标,来评价融合后的 HR-SWIR 图像的质量。Yang 等^[8]提出了一种基于质量感知特征提取的无参考图像质量评价方法,提取多个图像质量感知的统计特征;接着在图像中划分不同的局部块,并在每个块上拟合多变量高斯(MVG)模型;然后计算每个重建图像的块 MVG 模型与基准 MVG 模型之间的修正巴氏距离,得到客观质量分数。Yang 等^[9]提出了一种基于尺度不变特征变换(SIFT)特征提取的无参考质量评价方法,首先构建遥感卫星主观数据库,然后提取图像的 SIFT 特征构成字典,最后使用支持向量机(SVM)来获得无参考质量评价模型。以上几种客观评价方法,都需要人工对卫星图像进行主观评价,而不是根据卫星图像的应用来衡量图像质量,这大大降低了遥感卫星图像的利用率。因此,有必要根据特定应用场景,如根据建筑物检测准确率为评判图像质量标准,设计一种客观评价模型对遥感卫星图像进行评价,这将有效地提高遥感卫星图像评价的效率。

本文在分析遥感卫星图像与自然图像差别基础上,从特定应用场景和客观角度,对卫星立体图像的质量评价问题进行深入研究。为此,构建了一个新的失真卫星立体图像数据库,共组成 400 幅失真图像,基于数据库的构建,提出了一种基于稀疏表示的客观评价方法,主要贡献包括:1) 建立了失真卫星立体图像数据库,包含 20 幅原始图像,失真类型分别有高斯模糊和高斯噪声,每幅原始图像失真等级为 10 级,共组成 400 幅失真图像;2) 针对建筑物检测特定应用,使用角点检测和数字表面模型(DSM)高程信息的建筑物检测方法,根据失真图像检测的

角点变化,提出检测准确率来表示图像的失真程度,检测准确率将用于评估客观评价模型的性能;3) 提取原始图像和失真图像的 SIFT 特征和二进制稳健不变尺度特征(BRISK)特征,分别进行字典学习,将失真图像和原始图像字典组合为联合字典,基于稀疏表示的相似性度量,构成质量分数,最终经过支持向量回归(SVR)融合 4 个质量分数,得到最终的客观质量预测分数。实验结果表明,所提客观模型更适合预测卫星立体图像的质量。

2 失真卫星立体图像数据库

为了研究失真卫星立体图像的质量评价,构建了一种新的失真卫星立体图像数据库,包含 20 幅原始图像,失真类型分别有高斯模糊和高斯噪声,每幅原始图像失真等级为 10 级,共组成 400 幅失真图像。

2.1 失真卫星立体图像数据库构建

卫星立体图像来自 Digitalglobe 公司提供的数据集,将数据集图像进行预处理,裁剪出 20 幅不同分辨率的卫星立体图像和对应的 DSM。这些立体图像分别是 2012 年 9 月 21 日 WorldView-1 卫星拍摄的美国华盛顿哥伦比亚特区(Washington D.C., USA)和 2016 年 2 月 16 日 WorldView-1 卫星拍摄的瑞典斯德哥尔摩地区(Stockholm, Sweden)的卫星立体图像。图 1 为数据库原始卫星立体图像的左图像,数据库图像有大小形状不同的建筑物,如剧院、居住区、商场、学校、高层小区等不同类型的建筑物。数据库图像分别有高斯模糊和高斯噪声两种失真,每幅原始图像的失真等级为 10 级,数据库共 400 幅失真图像。

2.2 建筑物检测

本文的失真卫星立体图像数据库是结合卫星立体图像的角点检测和 DSM 高程信息的建筑物检测构建而成的。Özcan 等^[10]提出角点检测和 DSM 成信息可以有效地检测出建筑物,但其仅仅对于较小、较为单一的建筑物。本文在其基础上改进,可检测出不同大小和类型的建筑物。图 2 为建筑物检测框图,首先卫星立体图像经过预处理,裁剪得到卫星立体图像和对应的 DSM,其中卫星立体图像的左图像用于角点检测,DSM 由立体左右图像立体匹配得到;然后,将角点信息与 DSM 高程信息结合,最终可检测出建筑物。

2.2.1 Harris 角点检测

角点是图像重要的特征,对图像图形的理解和分析有很重要的作用。角点在保留图像重要特征的同时,可以有效地减少信息的数据量,使其信息的

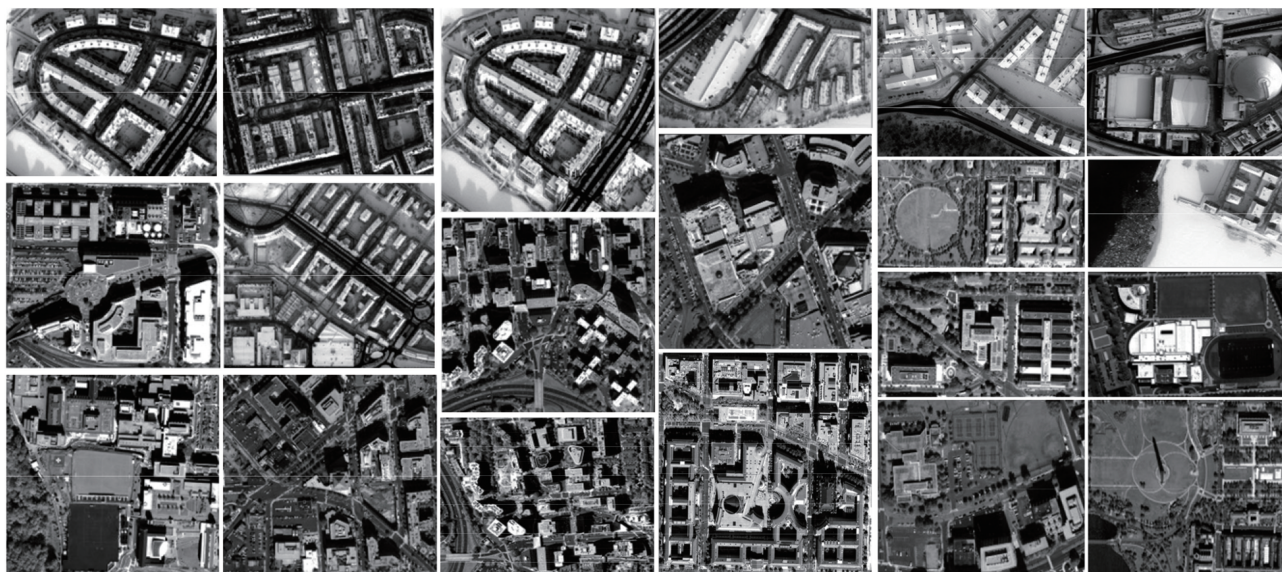


图 1 数据库原始卫星立体图像的左图像

Fig. 1 Left images of the original satellite stereo images in the database

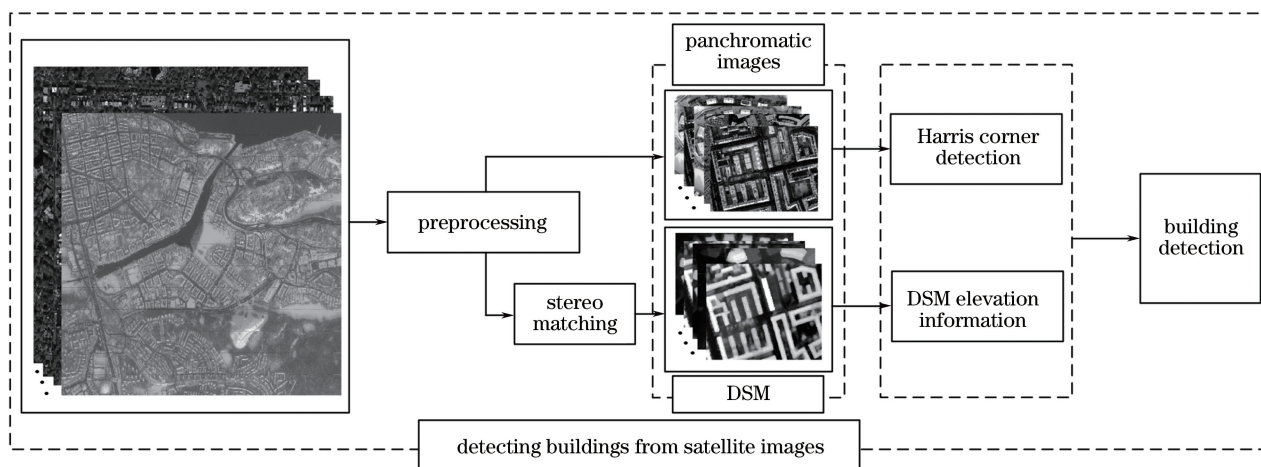


图 2 建筑物检测框图

Fig. 2 Block diagram of building detection

含量很高,有效地提高了计算的速度,有利于图像的可靠匹配和实时处理。Schmid 等^[11]对通用图像处理应用中的不同角点检测器进行评价,得出最优的是 Harris 角点检测^[12]。因此,采用 Harris 角点检测来提取图像的局部特征。

Harris 角点检测器分为三步:梯度计算,矩阵形成和特征值计算。首先,计算 x 和 y 方向上的平滑(使用高斯函数)梯度来检测给定灰度图像 $I(x, y)$ 中的角点,表示为

$$g_x(x, y) = \frac{-x}{2\pi\tau_g^4} \exp\left(-\frac{x^2 + y^2}{2\tau_g^2}\right), \quad (1)$$

$$g_y(x, y) = \frac{-y}{2\pi\tau_g^4} \exp\left(-\frac{x^2 + y^2}{2\tau_g^2}\right), \quad (2)$$

式中 τ_g 是平滑参数。

计算图像 $I(x, y)$ 的平滑梯度为

$$I_x = g_x(x, y) \otimes I(x, y), \quad (3)$$

$$I_y = g_y(x, y) \otimes I(x, y), \quad (4)$$

式中“ \otimes ”表示二维卷积运算。

Harris 角点检测器依赖于计算一个矩阵(与自相关函数有关),即

$$\mathbf{A}(x, y) = \begin{pmatrix} a_{xx} & a_{xy} \\ a_{xy} & a_{yy} \end{pmatrix}, \quad (5)$$

其中,

$$a_{xx} = \sum_{x_i \in W} \sum_{y_i \in W} I_x^2(x_i, y_i), \quad (6)$$

$$a_{yy} = \sum_{x_i \in W} \sum_{y_i \in W} I_y^2(x_i, y_i), \quad (7)$$

$$a_{xy} = \sum_{x_i \in W} \sum_{y_i \in W} I_x(x_i, y_i) I_y(x_i, y_i). \quad (8)$$

(6)~(8)式可以看出, a_{xx} 、 a_{yy} 和 a_{xy} 表示平均梯度幅值, 矩阵 \mathbf{A} 的特征值提供关于给定位置的边缘的信息。如果给定位置的矩阵的特征值都很大, 那么大部分区域均为角点。Harris 等^[12]通过计算响应函数可以避免精确的特征值计算, 即

$$R(\mathbf{A}) = |\mathbf{A}| - k \cdot [\text{tr}(\mathbf{A})]^2, \quad (9)$$

式中 k 是可调参数, 一般设置为 0.04~0.15, 本文设置 $k=0.06$ 。通过判定 R 大小来判断像素点是否为角点, 对于角点, $|R|$ 的值很大; 而平坦的区域, $|R|$ 的值很小。根据(9)式, 可以得到图像中的角点 $I_c(x, y)$ 的坐标位置。

2.2.2 DSM 高程信息

DSM 包含有关地形表面, 建筑物和其他物体的高程信息。DSM 可由卫星拍摄的左右图像立体匹配得到。Tian 等^[13]提出了基于立体图像和立体匹配生成 DSM 的方法, 提取 DSM 高程信息, 评估海地震后建筑物损伤情况。Tian 等^[14]利用 DSM 高程信息, 检测同一地点, 不同时期的建筑物高度变化情况。本文 DSM 数据由 Digitalglobe 公司的数据集提供。DSM 高程信息用于建筑物检测中。

2.2.3 建筑物检测分析

卫星立体图像的左图像经过角点检测的结果如

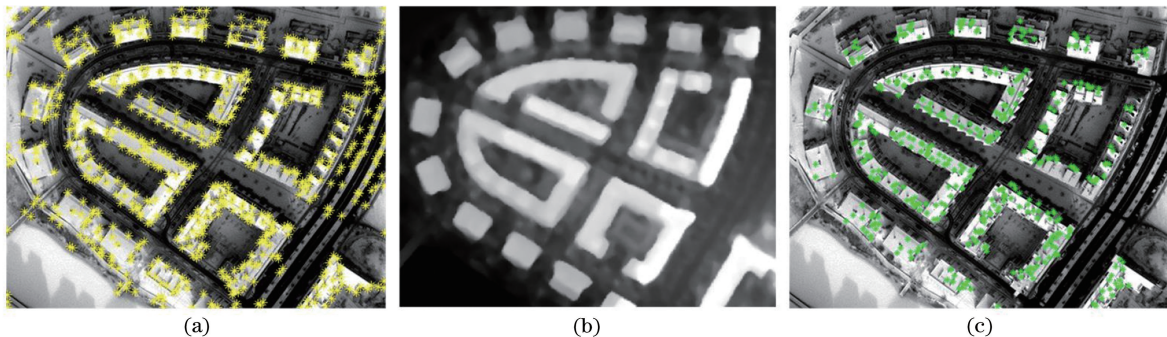


图 3 原始左图像的建筑物检测。(a)原始图像角点检测结果;(b)图(a)对应的 DSM;(c)建筑物检测结果
Fig. 3 Original left image of building detection. (a) Corner detection results of original image; (b) DSM corresponding to Fig. (a); (c) building detection results

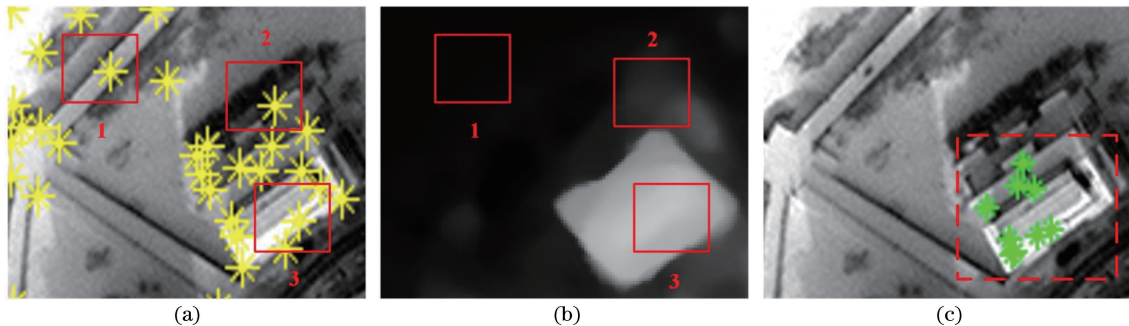


图 4 部分区域建筑物检测。(a)角点检测结果;(b)图(a)对应的 DSM;(c)建筑物检测结果
Fig. 4 Building detection in some areas. (a) Corner detection results; (b) DSM corresponding to Fig. (a); (c) building test results

图 3(a)所示。为了便于分析, 对图 3(a)的部分区域进行分析, 图 4 为部分区域检测结果, 设每个角点在对应 DSM 上, 以每个角点位置 $I_c(x, y)$ 为中心, 扩展大小为 $w \times w$ 的窗口, $I_s(x, y)$ 为窗口内的所有点, 其中, 窗口 w 为可调参数, 本文设置 $w=5$ 。窗口内设置 DSM 高度阈值 T , 来消除非建筑物区域的角点, 定义函数 $H(\cdot)$, 表示角点对应的 DSM 高度。最终, 建筑物检测的角点 $J(x, y)$ 表示为

$$J(x, y) = \begin{cases} 1, & \text{if } \max\{H[I_s(x, y)]\} > T \\ 0, & \text{else} \end{cases}, \quad (10)$$

式中: $I_s(x, y)$ 是以角点 $I_c(x, y)$ 为中心, 5×5 窗口内的所有点; T 为设置的高度阈值; $H[I_s(x, y)]$ 为窗口内每个点对应的 DSM 高度; $J(x, y)$ 为最终建筑物检测的角点。例如, 图 4(b)中 1 和 2 矩形窗口分别对应于图 4(a)中 1 和 2 矩形窗口, 窗口内所有点的最大值不在设定的阈值范围内, 则将角点消除。这样, 遍历图中所有的角点, 即可检测出建筑物。图 5 为数据库中部分失真图像检测结果, 从图 5 可以看出, 不同失真图像, 建筑物检测的结果有所不同。

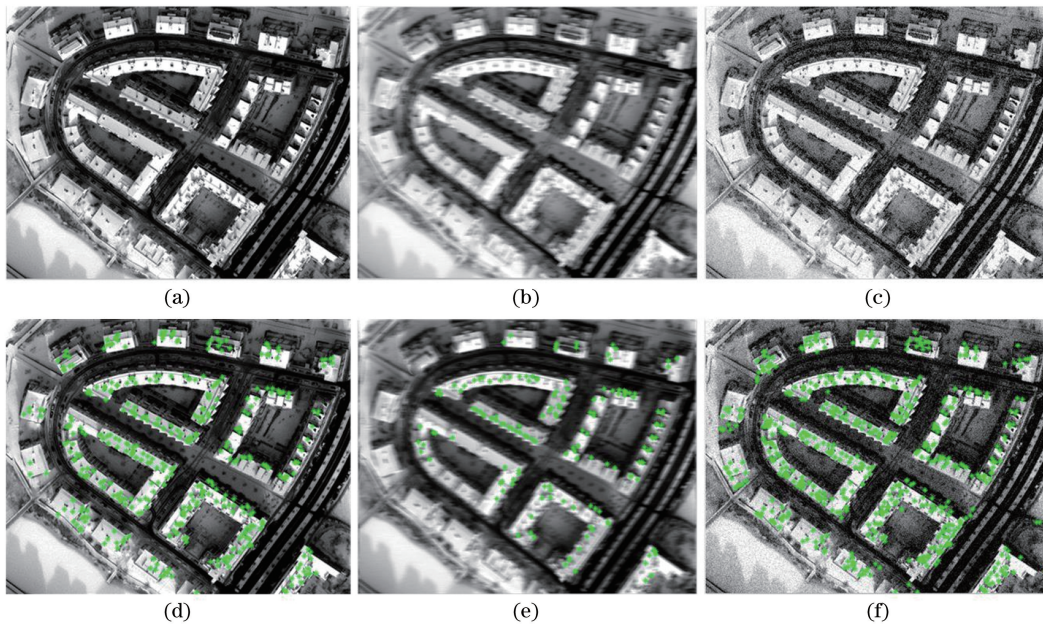


图 5 建筑物检测结果。(a)原始图;(b)模糊失真图;(c)噪声失真图;(d)原始图建筑物检测结果;(e)模糊失真后建筑物检测结果;(f)噪声失真后建筑物检测结果

Fig. 5 Building detection results. (a) Original image; (b) blur distortion image; (c) noise distortion image; (d) building detection of original image; (e) building detection result after blur distortion; (f) building detection result after noise distortion

2.3 图像质量度量

如 2.2.3 节所述,不同失真图像,建筑物检测结果将有所不同,为了去衡量不同失真图像建筑物检测的影响,本文提出检测准确率(即失真图像与原始图像之间的角点变化关系),定义为

$$R_C = \frac{R_{TP}}{R_{TP} + R_{FN} + R_{FP}}, \quad (11)$$

式中: R_{TP} 为失真图检测角点与原始图检测角点相同的数量; R_{FN} 为原始图像角点检测到,而失真图像未检测到角点的数量; R_{FP} 为失真图像角点检测到,而原始图未检测到角点的数量。

角点数量轻微变化都会直接影响建筑物检测准确率,根据实验得出,对于大部分失真图像,角点检测后,有些角点将落在建筑物区域,应视作为 R_{TP} ,

为进一步提高描述检测准确率的精度,以失真图像建筑物检测的角点为中心,扩展一个 $m \times m$ 窗口,窗口内的点如果与原始图像角点相同,即为 R_{TP} 。窗口 m 为可调参数,本文设置 $m=3$ 。以失真图建筑物检测得到的一个角点 $J_S(x, y)$ 为中心,扩展一个 $m \times m$ 窗口,则 $J_S^m(x, y)$ 为窗口内所有点,则 R_{TP} 为 $J_S^m(x, y)$ 与原始图像检测角点 $J_O(x, y)$ 相同点的个数。图 6(a)为原始图建筑物检测的角点;图 6(b)矩形窗 1 为失真图检测的角点处 $m \times m$ 窗口内的点,与原始检测的角点相同,则记为 R_{TP} ;图 6(b)中圆形窗 2 检测无角点,而图 6(a)中圆形窗 2 检测有角点,其不同角点的个数,记为 R_{FN} ;图 6(c)中圆形窗 2 检测有角点,而图 6(a)中圆形窗 3 检测无角点,其不同角点的个数,记为 R_{FP} 。

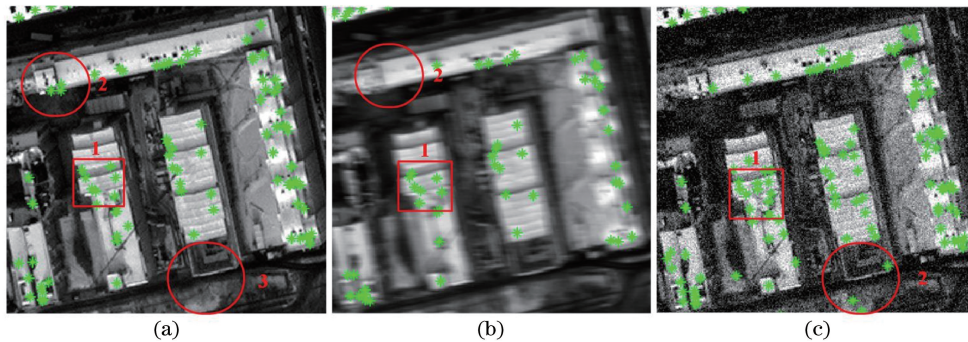


图 6 建筑物检测结果。(a)原始图建筑物检测结果;(b)模糊失真后建筑物检测结果;(c)噪声失真后建筑物检测结果
Fig. 6 Building detection results. (a) Building detection results of original image; (b) building detection results after blur distortion; (c) building detection results after noise distortion

失真卫星立体图像数据库检测准确率分布情况如图 7 所示,检测准确率分布在 50~100 之间,检测准确率值越高,表明图像对建筑物检测的效果越好,检测准确率表明图像的失真程度。检测准确率将用于评估客观评价模型的性能。

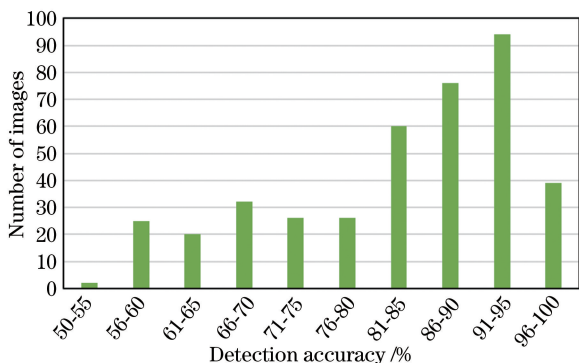


图 7 数据库中检测准确率直方图

Fig. 7 Histogram of detection accuracy in the database

3 失真卫星立体图像质量评价方法

根据特定应用,使用角点检测和DSM高程信

息进行建筑物检测,为了更客观、精确地预测卫星立体图像的质量,本文提出一种客观质量评价模型,图 8 为提出的客观质量评价框图,首先,分别提取原始图像左图像和失真图像左图像的 SIFT 和 BRISK 特征,分别构建字典;然后,基于稀疏表示的相似性度量,构建对应的质量分数;最后,经过支持向量回归融合 4 个质量分数,得到最终的客观评价价值。由于获取的卫星立体图像的特征信息具有较强的稳定性,并且能够较好地反映卫星立体图像的质量变化情况,提高了客观评价模型与检测准确率之间的相关性。

3.1 特征提取

3.1.1 SIFT 特征提取

提取一幅图像 SIFT 特征^[15],每个 SIFT 特征点用 128 维特征向量表示。通常可以从图像中检测出数千个特征点。图 9 为原图像和失真图像的 SIFT 特征点检测结果。从图 9 中可以看出,SIFT 特征点检测结果对于原始图像和失真图像是不同的。实验结果表明,不同失真程度的图像提取的

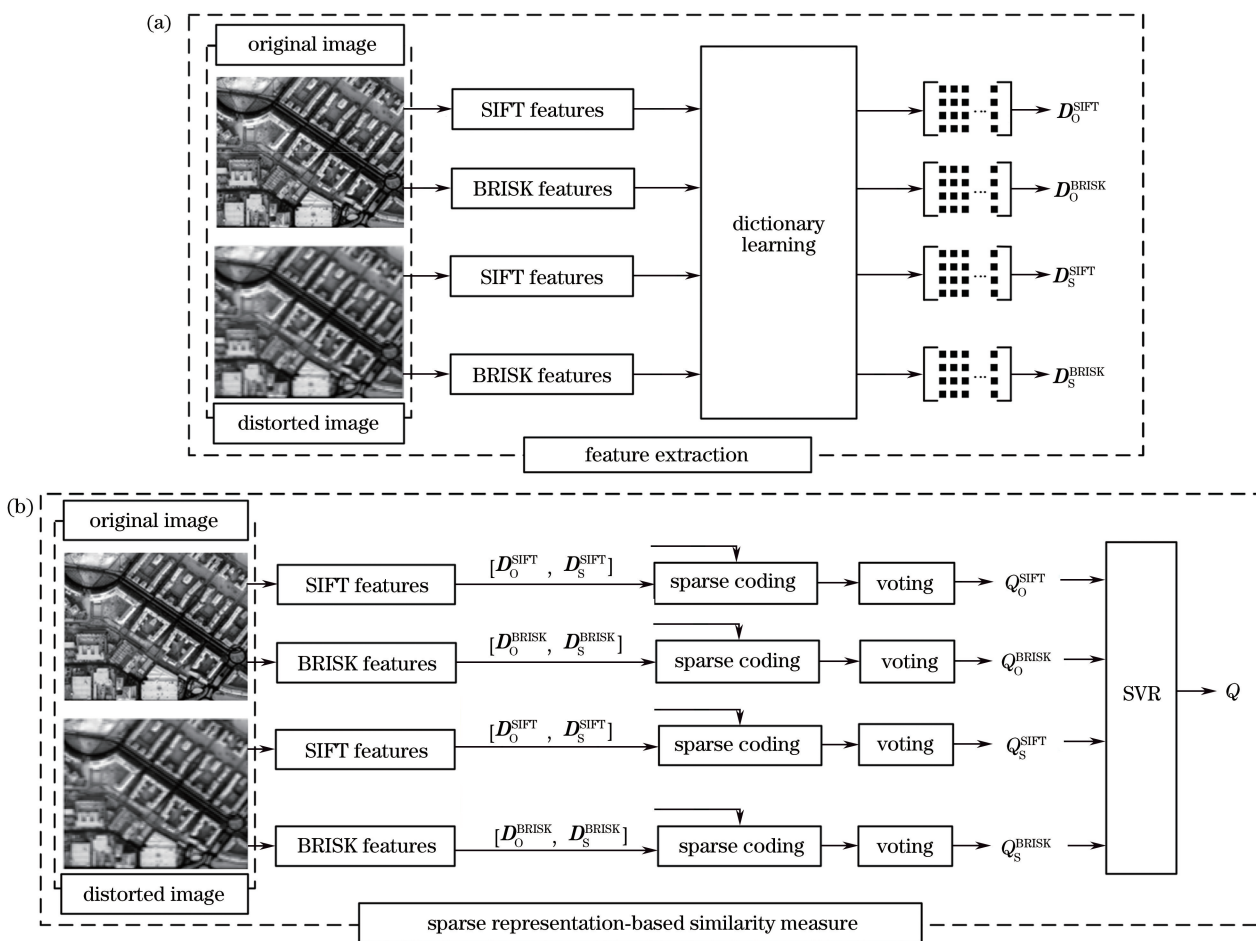


图 8 客观质量评价框图。(a)特征提取;(b)基于稀疏表示的相似性度量

Fig. 8 Block diagram of objective quality evaluation. (a) Feature extraction; (b) sparse representation-based similarity measure

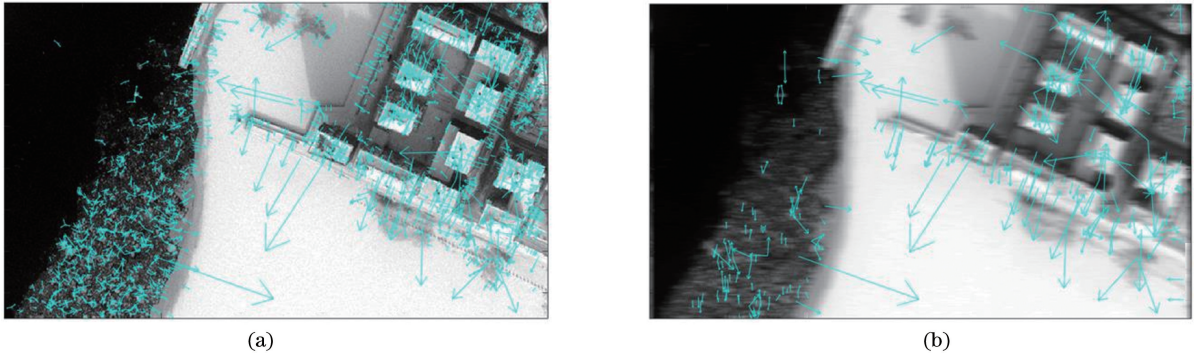


图 9 SIFT 特征提取。(a)原始图像 SIFT 特征提取;(b)失真图像 SIFT 特征提取

Fig. 9 SIFT features extraction. (a) SIFT features extraction of the original image; (b) SIFT features extraction of distorted image

SIFT 特征将有所不同,这对预测图像质量有很大的帮助。例如,给定一幅图像,设 $\{f_i^{\text{SIFT}}\}_{i=1}^N$ 表示 SIFT 特征点,每列 $f_i^{\text{SIFT}} \in \mathbb{R}^{128 \times 1}$ 代表第 i 个特征点的 SIFT 特征向量, N 是 SIFT 特征点的数量。所有提取的 SIFT 特征向量将用于字典学习。

3.1.2 BRISK 特征提取

BRISK 算法^[16]用于特征点检测,是描述和匹配图像的一种新方法。BRISK 的优点在于速度快,速度的关键在于一种新的尺度空间的应用。针对 BRISK 特征的优越性,将提取 BRISK 特征用于评价模型的一个预测值。实验结果表明,SIFT 和 BRISK 特征将有很好的互补性,经过对应的特征质量分数融合,将极大地提高评价模型的性能。例如,给定一幅图像,设 $\{f_i^{\text{BRISK}}\}_{i=1}^K$ 表示 BRISK 特征点,每列 $f_i^{\text{BRISK}} \in \mathbb{R}^{64 \times 1}$ 表示第 i 个特征点的 BRISK 特征向量, K 是 BRISK 特征点的数量。图像中提取的所有 BRISK 特征点将用于字典学习。

3.2 过完备字典学习

稀疏表示分类技术的性能依赖于过完备字典的构建。目的是通过在特定稀疏约束 τ 下寻找每个特征向量的稀疏表示,同时学习 M 个基本原子 $\mathbf{D}^{\text{SIFT}} = [d_1^{\text{SIFT}}, d_2^{\text{SIFT}}, \dots, d_M^{\text{SIFT}}] \in \mathbb{R}^{128 \times M}$ 及其对应的稀疏系数矩阵 $\mathbf{x}^{\text{SIFT}} = \{\mathbf{x}_i^{\text{SIFT}}\}_{i=1}^N \in \mathbb{R}^{M \times N}$ 的过完备字典。形式上,这个问题可以表述为

$$\begin{aligned} \{\mathbf{D}^{\text{SIFT}}, \mathbf{X}^{\text{SIFT}}\} = \min \sum_{i=1}^N \|\mathbf{f}_i^{\text{SIFT}} - \mathbf{D}^{\text{SIFT}} \mathbf{x}_i^{\text{SIFT}}\|_2^2, \\ \text{s.t. } \forall_i, \|\mathbf{x}_i^{\text{SIFT}}\|_0 \leq \tau \end{aligned} \quad (12)$$

式中 $\mathbf{x}_i^{\text{SIFT}}$ 是 $\mathbf{f}_i^{\text{SIFT}}$ 的稀疏系数向量; $\|\cdot\|_2$ 是 l_2 范数, $\|\cdot\|_0$ 是 l_0 范数,表示向量中非零元素的数量,并且 τ 是预定义期望数量的非零系数^[17]。本文使

用高效的 K-SVD 算法^[18-19]解决上述问题。以 SIFT 提取一幅图像的特征构建字典为例,(12)式中, $\mathbf{D}^{\text{SIFT}} \in \mathbb{R}^{128 \times M}$ 为过完备字典, M 为原子数; $\{\mathbf{f}_i^{\text{SIFT}}\}_{i=1}^N \in \mathbb{R}^{128 \times N}$ 为 SIFT 特征向量, N 为检测一幅图像的特征点数量; $\mathbf{f}_i^{\text{SIFT}} \in \mathbb{R}^{128 \times 1}$ 表示第 i 个特征点的 128 维特征向量, $\{\mathbf{x}_i^{\text{SIFT}}\}_{i=1}^N \in \mathbb{R}^{M \times N}$ 为稀疏系数矩阵。

3.3 特征相似性度量

3.3.1 基于联合字典稀疏表示

将原图像和失真图像之间的特征相似性度量值转换为一个图像中提取的特征在另一个图像中有多少相同的特征。为此,通过使用联合字典 $\mathbf{D}_{\text{S.O}}^{\text{SIFT}}$,表示每个 SIFT 特征向量 $\mathbf{f}_{\text{S},i}^{\text{SIFT}}$ 时的稀疏表示为

$$\hat{\mathbf{x}}_{\text{S},i}^{\text{SIFT}} = \operatorname{argmin} \left[\sum_{i=1}^{N_1} \|\mathbf{f}_{\text{S},i}^{\text{SIFT}} - \mathbf{D}_{\text{S.O}}^{\text{SIFT}} \mathbf{x}_{\text{S},i}^{\text{SIFT}}\|_2^2 + \lambda \|\mathbf{x}_{\text{S},i}^{\text{SIFT}}\|_1 \right], \quad (13)$$

式中 $\mathbf{D}_{\text{S.O}}^{\text{SIFT}} = [\mathbf{D}_0^{\text{SIFT}}, \mathbf{D}_\text{S}^{\text{SIFT}}] \in \mathbb{R}^{128 \times (M_1 + M_2)}$ 为原图像和失真图像的联合字典, $\hat{\mathbf{x}}_{\text{S},i}^{\text{SIFT}} \in \mathbb{R}^{(M_1 + M_2)}$ 为 $\mathbf{D}_{\text{S.O}}^{\text{SIFT}}$ 表示 $\mathbf{f}_{\text{S},i}^{\text{SIFT}}$ 的稀疏系数向量。本文采用批量正交匹配追踪(batch-OMP)算法^[19]来解决上述问题。

3.3.2 稀疏残差重构

1) 重构失真图

通过使用 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_\text{S}^{\text{SIFT}}$ 中的原子来重构 $\mathbf{f}_{\text{S},i}^{\text{SIFT}}$ 来计算残差,以区分哪个子字典更适合于表示 $\mathbf{f}_{\text{S},i}^{\text{SIFT}}$ 。为此,定义了一个函数 $\xi(\cdot)$,它只保留相关的子字典的稀疏系数不变,而将另一个字典所有元素设为零。例如,给定 $\hat{\mathbf{x}}_{\text{S},i}^{\text{SIFT}}$ 作为输入, $\xi_\text{S}(\hat{\mathbf{x}}_{\text{S},i}^{\text{SIFT}})$ 为新的系数向量,其中只有与 $\mathbf{D}_\text{S}^{\text{SIFT}}$ 相关的系数保持不变而其他系数设为零,而 $\xi_0(\hat{\mathbf{x}}_{\text{S},i}^{\text{SIFT}})$ 产生新的系数向量,其中仅与 $\mathbf{D}_0^{\text{SIFT}}$ 相关的系数不变,其他系数均设

为零。因此,使用 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 的原子来重构 $\mathbf{f}_{s,i}^{\text{SIFT}}$ 的残差为

$$E_s(\mathbf{f}_{s,i}^{\text{SIFT}}) = \|\mathbf{f}_{s,i}^{\text{SIFT}} - \mathbf{D}_{s,0}^{\text{SIFT}} \xi_s(\hat{\mathbf{x}}_{s,i}^{\text{SIFT}})\|_2, \quad (14)$$

$$E_o(\mathbf{f}_{s,i}^{\text{SIFT}}) = \|\mathbf{f}_{s,i}^{\text{SIFT}} - \mathbf{D}_{s,0}^{\text{SIFT}} \xi_o(\hat{\mathbf{x}}_{s,i}^{\text{SIFT}})\|_2, \quad (15)$$

式中 $E_s(\mathbf{f}_{s,i}^{\text{SIFT}})$ 和 $E_o(\mathbf{f}_{s,i}^{\text{SIFT}})$ 分别表示与 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 相关的稀疏重构残差。然后,对一幅失真图像所有的 SIFT 特征的重构残差进行投票^[20-21],计算与 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 对应的投票的概率为

$$V_{s,s}^{\text{SIFT}} = \frac{1}{N_1} \sum_{i=1}^{N_1} V_s(\mathbf{f}_{s,i}^{\text{SIFT}}), \quad (16)$$

$$V_{s,o}^{\text{SIFT}} = \frac{1}{N_1} \sum_{i=1}^{N_1} V_o(\mathbf{f}_{s,i}^{\text{SIFT}}), \quad (17)$$

式中: $V_{s,s}^{\text{SIFT}}$ 和 $V_{s,o}^{\text{SIFT}}$ 分别表示对应 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 投票的概率, $0 < V_{s,s}^{\text{SIFT}}, V_{s,o}^{\text{SIFT}} < 1, V_{s,s}^{\text{SIFT}} + V_{s,o}^{\text{SIFT}} = 1$; N_1 为失真图 SIFT 特征点的数量; $V_s(\mathbf{f}_{s,i}^{\text{SIFT}})$ 和 $V_o(\mathbf{f}_{s,i}^{\text{SIFT}})$ 为一幅失真图像的第 i 个 SIFT 特征点的投票,表示为

$$V_s(\mathbf{f}_{s,i}^{\text{SIFT}}) = \begin{cases} 1, & \text{if } E_o(\mathbf{f}_{s,i}^{\text{SIFT}}) \geq E_s(\mathbf{f}_{s,i}^{\text{SIFT}}) \\ 0, & \text{else} \end{cases}, \quad (18)$$

$$V_o(\mathbf{f}_{s,i}^{\text{SIFT}}) = \begin{cases} 1, & \text{if } E_o(\mathbf{f}_{s,i}^{\text{SIFT}}) < E_s(\mathbf{f}_{s,i}^{\text{SIFT}}) \\ 0, & \text{else} \end{cases}. \quad (19)$$

根据投票策略^[21],将原图像和失真图像之间的相似性度量(即质量分数)定义为

$$Q_s^{\text{SIFT}} = \frac{V_{s,o}^{\text{SIFT}} - V_{s,s}^{\text{SIFT}} + 1}{2}, \quad (20)$$

式中 $Q_s^{\text{SIFT}} \in [0, 1]$ 为质量分数,表示在失真图像中提取的特征能够从其原图像中提取多少。

2) 重构原图

为了实现更稳健的相似性度量,对原图像重复执行上述实施的稀疏编码和投票处理,以获得另一质量分数, $Q_o^{\text{SIFT}} \in [0, 1]$, 表示原图像提取的特征,可以从其失真图像中提取多少。在这种情况下, $\mathbf{D}_s^{\text{SIFT}}$ 应该在字典学习阶段被强制为比 $\mathbf{D}_0^{\text{SIFT}}$ 学习的更好。使用 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 的原子来重构 $\mathbf{f}_{o,j}^{\text{SIFT}}$ 的残差为

$$E_s(\mathbf{f}_{o,j}^{\text{SIFT}}) = \|\mathbf{f}_{o,j}^{\text{SIFT}} - \mathbf{D}_{s,0}^{\text{SIFT}} \xi_s(\hat{\mathbf{x}}_{o,j}^{\text{SIFT}})\|_2, \quad (21)$$

$$E_o(\mathbf{f}_{o,j}^{\text{SIFT}}) = \|\mathbf{f}_{o,j}^{\text{SIFT}} - \mathbf{D}_{s,0}^{\text{SIFT}} \xi_o(\hat{\mathbf{x}}_{o,j}^{\text{SIFT}})\|_2, \quad (22)$$

式中 $E_s(\mathbf{f}_{o,j}^{\text{SIFT}})$ 和 $E_o(\mathbf{f}_{o,j}^{\text{SIFT}})$ 分别表示与 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 相关的稀疏重构残差;对一幅原始图像所有的 SIFT 特征的重构残差进行投票,计算与 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 对应投票的概率为

$$V_{o,o}^{\text{SIFT}} = \frac{1}{N_2} \sum_{j=1}^{N_2} V_o(\mathbf{f}_{o,j}^{\text{SIFT}}), \quad (23)$$

$$V_{o,s}^{\text{SIFT}} = \frac{1}{N_2} \sum_{j=1}^{N_2} V_s(\mathbf{f}_{o,j}^{\text{SIFT}}), \quad (24)$$

式中: $V_{o,o}^{\text{SIFT}}$ 和 $V_{o,s}^{\text{SIFT}}$ 分别表示对应于 $\mathbf{D}_0^{\text{SIFT}}$ 和 $\mathbf{D}_s^{\text{SIFT}}$ 的投票的概率, $0 < V_{o,o}^{\text{SIFT}}, V_{o,s}^{\text{SIFT}} < 1, V_{o,o}^{\text{SIFT}} + V_{o,s}^{\text{SIFT}} = 1$; N_2 为原始图像 SIFT 特征点的数量; $V_s(\mathbf{f}_{o,j}^{\text{SIFT}})$ 和 $V_o(\mathbf{f}_{o,j}^{\text{SIFT}})$ 为原始图像的第 j 个 SIFT 特征点的投票,即

$$V_s(\mathbf{f}_{o,j}^{\text{SIFT}}) = \begin{cases} 1, & \text{if } E_o(\mathbf{f}_{o,j}^{\text{SIFT}}) \geq E_s(\mathbf{f}_{o,j}^{\text{SIFT}}) \\ 0, & \text{else} \end{cases}, \quad (25)$$

$$V_o(\mathbf{f}_{o,j}^{\text{SIFT}}) = \begin{cases} 1, & \text{if } E_o(\mathbf{f}_{o,j}^{\text{SIFT}}) < E_s(\mathbf{f}_{o,j}^{\text{SIFT}}) \\ 0, & \text{else} \end{cases}. \quad (26)$$

基于这种投票策略^[21],将原图像和失真图像之间的相似性度量(即质量分数)定义为

$$Q_o^{\text{SIFT}} = \frac{V_{o,s}^{\text{SIFT}} - V_{o,o}^{\text{SIFT}} + 1}{2}, \quad (27)$$

式中 $Q_o^{\text{SIFT}} \in [0, 1]$ 为质量分数,表示在原始图像中提取的特征能够从失真图像中提取多少。

同理,对于 BRISK 特征提取方法可得出质量分数 Q_s^{BRISK} 和 Q_o^{BRISK} 分别为

$$Q_s^{\text{BRISK}} = \frac{V_{s,o}^{\text{BRISK}} - V_{s,s}^{\text{BRISK}} + 1}{2}, \quad (28)$$

$$Q_o^{\text{BRISK}} = \frac{V_{o,s}^{\text{BRISK}} - V_{o,o}^{\text{BRISK}} + 1}{2}. \quad (29)$$

3.4 质量分数融合

分别提取原始图像和失真图像 SIFT 特征和 BRISK 特征;然后,基于稀疏表示的相似性度量得到 4 个质量分数分别为 $Q_s^{\text{SIFT}}, Q_o^{\text{SIFT}}, Q_s^{\text{BRISK}}, Q_o^{\text{BRISK}}$, 将 4 个质量分数融合得到最终的客观预测值。训练预测函数 f 来预测失真图像的最终质量分数。最终质量值表示为

$$Q_f = f(Q), \quad (30)$$

式中预测函数 f 通过 ϵ -SVR^[22] 训练得到, $Q = [Q_s^{\text{SIFT}}, Q_o^{\text{SIFT}}, Q_s^{\text{BRISK}}, Q_o^{\text{BRISK}}]$ 作为输入并经过 SVR 得到输出 Q_f 作为最终质量分数。在实验中,使用径向基函数(RBF)的 SVR 性能最佳。

4 实验结果与分析

4.1 模型性能分析

将客观评价模型在构建的数据库上进行测试,采用目前常用的皮尔逊线性相关系数(PLCC)、斯

皮尔曼等级相关系数(SROCC)、肯德尔等级相关系数(KROCC)和均方误差(RMSE)4个客观参量指标来量化对比结果。SROCC和KROCC反映客观评价模型的单调性,PLCC和RMSE反映客观评价模型的准确性。PLCC,SROCC和KROCC的值越接近1,RMSE的值越接近0,说明客观评价模型越好。值得注意的是,PLCC和RMSE两个参数需要经过逻辑函数的非线性映射后再进行计算。

表1所示为质量分数和经过SVR融合后最终预测值的各项性能指标,从表1可以看出,SIFT特征较BRISK特征更适合稀疏表示失真图像,而BRISK特征较SIFT更适合稀疏表示原始图像,最终,经过SVR融合得到的质量分数 $Q = [Q_s^{SIFT}, Q_o^{SIFT}, Q_s^{BRISK}, Q_o^{BRISK}]$ 的性能指标有较大的提升,PLCC值超过0.9013,SROCC值超过0.8772,说明本文提出的特征描述方法和融合质量分数的方法有效。

表1 各个质量值的性能比较

Table 1 Performance comparison of individual quality values

Criteria	Q_s^{SIFT}	Q_o^{SIFT}	Q_s^{BRISK}	Q_o^{BRISK}	Q_f
PLCC	0.8213	0.7203	0.8204	0.8469	0.9013
SROCC	0.7951	0.7048	0.8024	0.8024	0.8772
KROCC	0.5976	0.5279	0.5755	0.6065	0.7043
RMSE	6.6405	8.2312	6.7287	6.3688	5.0103

4.2 模型性能比较分析

为了验证本文提出客观评价模型的有效性,将本文方法与现有的8种全参考图像质量评价进行比较。其方法分别为多尺度结构相似性(MS-SSIM)^[23]、结构相似性(SSIM)^[24]、信息保真度(IFC)^[25]、视觉质量(VIF)^[26]、文献[27]方法、特征相似性指数(FSIM)^[28]、文献[29]方法和梯度相似

性(GSM)^[30]。表2给出8种全参考图像质量评价方法在本文构建的失真卫星立体图像库上的各项性能指标。从表2可以看出,本文方法的PLCC超过0.9013,SROCC超过0.8772,各指标均优于其他评价方法。图10(a)~(i)所示为不同客观评价模型预测得到的评价值和检测准确率的散点图,散点图越集中,说明客观评价模型与检测准确率的一致性越好。从图10中可以看出,采用本文方法得到的客观预测值与检测准确率的散点图较为集中。表3~4所示为8种全参考评价方法和本文方法在不同失真类型下的PLCC和SROCC的性能指标值,从表3可以看出,高斯模糊(Gblur)失真类型下,本文方法的PLCC值和SROCC值,非常接近对比方法的最优值。从表4可以看出,高斯白噪声(WN)失真类型下,本文方法的PLCC值和SROCC值均优于其他方法。总体来说,本文提出的全参考评价方法对于建筑物检测特定应用下的卫星立体图像的客观质量评价更为有效。

表2 不同评价方法的总体性能比较

Table 2 Comparison of overall performance of different evaluation methods

Model	PLCC	SROCC	KROCC	RMSE
MS-SSIM	0.8146	0.7938	0.6288	6.9891
SSIM	0.6551	0.6135	0.4311	8.8133
IFC	0.8285	0.8154	0.6705	6.6789
VIF	0.8890	0.8722	0.6839	6.3872
Model in Ref.[27]	0.7929	0.7445	0.5477	7.0585
FSIM	0.6971	0.6623	0.4735	8.2684
Model in Ref.[29]	0.6891	0.6750	0.4830	8.4530
GSM	0.6610	0.6916	0.4987	8.6132
Proposed	0.9013	0.8772	0.7043	5.0103

表3 不同失真类型的PLCC值

Table 3 PLCC values of different distortion types

Distortion	Model									
	MS-SSIM	SSIM	IFC	VIF	Model in Ref.[27]	FSIM	Model in Ref.[29]	GSM	Proposed	
Gblur	0.8624	0.6651	0.8342	0.9152	0.8374	0.6494	0.6475	0.7013	0.8901	
WN	0.8617	0.5321	0.8360	0.8823	0.7632	0.5836	0.5952	0.5579	0.9272	

表4 不同失真类型的SROCC值

Table 4 SROCC values of different distortion types

Distortion	Model									
	MS-SSIM	SSIM	IFC	VIF	Model in Ref.[27]	FSIM	Model in Ref.[29]	GSM	Proposed	
Gblur	0.9384	0.7079	0.8714	0.8629	0.7774	0.7074	0.7147	0.7384	0.8567	
WN	0.8967	0.6719	0.8463	0.8513	0.7577	0.6824	0.6904	0.7013	0.9076	

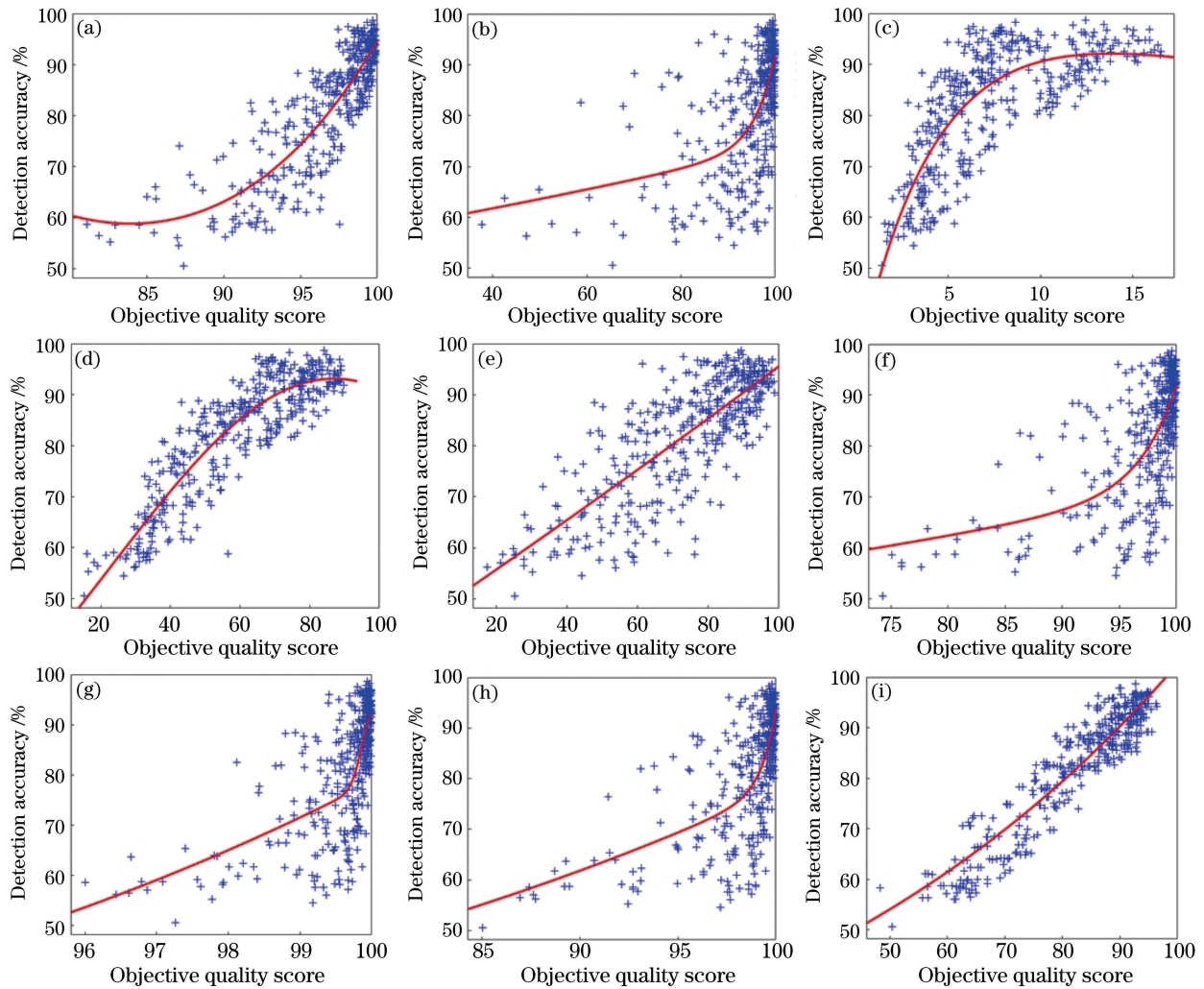


图 10 不同评价方法得到的评价预测值与检测准确率的散点图。(a) MS-SSIM; (b) SSIM; (c) IFC; (d) VIF; (e)文献[27]方法; (f) FSIM; (g) GSM; (h)文献[29]方法; (i)本文方法

Fig. 10 Scatter plots of evaluation prediction values and detection accuracy rates obtained by different evaluation methods.

(a) MS-SSIM; (b) SSIM; (c) IFC; (d) VIF; (e) model in Ref. [27]; (f) FSIM; (g) GSM; (h) model in Ref. [29]; (i) proposed method

5 结 论

针对建筑物检测特定应用,提出了一种基于稀疏表示的失真立体图像全参考质量评价方法。首先构建了一种失真卫星立体图像数据库,并使用角点检测和 DSM 高程信息的建筑物检测方法,提出了检测准确率来表示失真图像的质量。接着,分别提取原始图像和失真图像的特征,进行字典学习;然后,基于稀疏表示的相似性度量得到 4 个质量分数,最后通过支持向量回归得到最终预测卫星立体图像的客观评价价值。实验结果表明,所提出的客观评价模型与检测准确率具有较高的一致性,能够较为准确的反映卫星立体图像的质量。在此基础上,未来将考虑建立无参考质量评价模型,更为精确地预测

卫星立体图像的质量。

参 考 文 献

- [1] Zhang Y J, Zheng M T, Xiong X D, *et al.* Multistrip bundle block adjustment of ZY-3 satellite imagery by rigorous sensor model without ground control point[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(4): 865-869.
- [2] Scales K L, Hazen E L, Jacox M G, *et al.* Scale of inference: on the sensitivity of habitat models for wide-ranging marine predators to the resolution of environmental data [J]. Ecography, 2017, 40(1): 210-220.
- [3] Jiang Y H, Zhang G, Tang X M, *et al.* Improvement and assessment of the geometric accuracy of Chinese high-resolution optical satellites

- [J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2015, 8(10): 4841-4852.
- [4] Li S M, Chang Y L, Duan Z C. Objective assessment of stereoscopic image comfort based on convolutional neural network [J]. Acta Optica Sinica, 2018, 38(6): 0610003.
李素梅, 常永莉, 段志成. 基于卷积神经网络的立体图像舒适度客观评价 [J]. 光学学报, 2018, 38(6): 0610003.
- [5] Hou C P, Ma T T, Yue G H, *et al.* Multiply-distorted image quality assessment based on high-order phase congruency [J]. Laser & Optoelectronics Progress, 2017, 54(7): 071001.
侯春萍, 马彤彤, 岳广辉, 等. 基于高阶相位一致性的混合失真图像质量评价 [J]. 激光与光电子学进展, 2017, 54(7): 071001.
- [6] Hasanlou M, Saradjian M R. Quality assessment of Pan-sharpening methods in high-resolution satellite images using radiometric and geometric index [J]. Arabian Journal of Geosciences, 2016, 9(1): 1-10.
- [7] Kwan C, Budavari B, Bovik A C, *et al.* Blind quality assessment of fused WorldView-3 images by using the combinations of pansharpening and hypersharpening paradigms [J]. IEEE Geoscience and Remote Sensing Letters, 2017, 14(10): 1835-1839.
- [8] Yang J X, Zhao Y Q, Yi C, *et al.* No-reference hyperspectral image quality assessment via quality-sensitive features learning [J]. Remote Sensing, 2017, 9(4): 305-329.
- [9] Yang X C, Sun Q S, Wang T S. A usability-based subjective remote sensing image quality assessment database [J]. Signal, Image and Video Processing, 2017, 11(4): 697-704.
- [10] Özcan A H, Ünsalan C, Reinartz P. Building detection using local features and DSM data [C]//2013 6th International Conference on Recent Advances in Space Technologies (RAST), 2013: 139-143.
- [11] Schmid C, Mohr R, Bauckhage C. Evaluation of interest point detectors [J]. International Journal of Computer Vision, 2000, 37(2): 151-172.
- [12] Harris C, Stephens M. A combined corner and edge detector [C]//Proceedings of the Alvey Vision Conference, 1988.
- [13] Tian J J, Nielsen A A, Reinartz P. Building damage assessment after the earthquake in Haiti using two post-event satellite stereo imagery and DSMs [J]. International Journal of Image and Data Fusion, 2015, 6(2): 155-169.
- [14] Tian J J, Cui S Y, Reinartz P. Building change detection based on satellite stereo imagery and digital surface models [J]. IEEE Transactions on Geoscience and Remote Sensing, 2014, 52(1): 406-417.
- [15] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [16] Leutenegger S, Chli M, Siegwart R Y. BRISK: binary robust invariant scalable keypoints [C]//2011 International Conference on Computer Vision, 2011: 2548-2555.
- [17] Tropp J A, Gilbert A C. Signal recovery from random measurements via orthogonal matching pursuit [J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655-4666.
- [18] Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation [J]. IEEE Transactions on Signal Processing, 2006, 54(11): 4311-4322.
- [19] Rubinstein R, Zibulevsky M, Elad M. Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit [J]. Computer Science Department, 2008, 40(8): 1-15.
- [20] Kang L W, Hsu C Y, Chen H W, *et al.* Feature-based sparse representation for image similarity assessment [J]. IEEE Transactions on Multimedia, 2011, 13(5): 1019-1030.
- [21] Jiang Q P, Shao F, Lin W S, *et al.* Learning sparse representation for objective image retargeting quality assessment [J]. IEEE Transactions on Cybernetics, 2018, 48(4): 1276-1289.
- [22] Chang C C, Lin C J. LIBSVM [J]. ACM Transactions on Intelligent Systems and Technology, 2011, 2(3): 1-27.
- [23] Wang Z, Simoncelli E P, Bovik A C. Multiscale structural similarity for image quality assessment [C]//The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers, 2003: 1398-1402.
- [24] 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.
- [25] Sheikh H R, Bovik A C, de Veciana G. An information fidelity criterion for image quality assessment using natural scene statistics [J]. IEEE Transactions on Image Processing, 2005, 14(12): 2117-2128.
- [26] Sheikh H R, Bovik A C. Image information and visual quality [J]. IEEE Transactions on Image Processing, 2006, 15(2): 430-444.
- [27] Zhang L, Zhang L, Mou X Q. RFSIM: a feature based image quality assessment metric using Riesz

-
- transforms[C]//2010 IEEE International Conference on Image Processing, 2010: 321-324.
- [28] Zhang L, Zhang L, Mou X Q, *et al.* FSIM: a feature similarity index for image quality assessment [J]. IEEE Transactions on Image Processing, 2011, 20(8): 2378-2386.
- [29] Zhang L, Li H Y. SR-SIM: a fast and high performance IQA index based on spectral residual [C]//2012 19th IEEE International Conference on Image Processing, 2012: 1473-1476.
- [30] Liu A M, Lin W S, Narwaria M. Image quality assessment based on gradient similarity [J]. IEEE Transactions on Image Processing, 2012, 21(4): 1500-1512.