

基于深度玻尔兹曼模型的红外与可见光图像融合*

冯鑫[†] 李川 胡开群

(重庆工商大学机械工程学院, 制造装备机构设计与控制重庆市重点实验室, 重庆 400067)

(2014年2月19日收到; 2014年5月1日收到修改稿)

为了克服红外与可见光图像融合时噪声干扰及易产生伪影导致目标轮廓不鲜明、对比度低的缺点, 提出一种基于深度模型分割的图像融合方法. 首先, 采用深度玻尔兹曼机器学习红外与可见光的目标和背景轮廓先验, 构建轮廓的深度分割模型, 通过 Split Bregman 迭代算法获取最优能量分割后的红外与可见光图像轮廓; 然后再使用非下采样轮廓波变换对源图像进行分解, 并针对所分割的背景轮廓采用结构相似度的规则进行系数组合; 最后进行非下采样轮廓波反变换重构出融合图像. 数值试验证明, 该算法可以有效获取目标和背景轮廓均清晰的融合图像, 融合结果不但具有较高的对比度, 还能抑制噪声影响, 具有有效性.

关键词: 深度模型, 深度玻尔兹曼机, 非下采样轮廓波变换, 图像融合

PACS: 42.30.Va, 42.30.Wb

DOI: 10.7498/aps.63.184202

1 引言

图像融合是利用各种成像传感器所具有的不同成像方式提供互补信息、增加图像的信息量来获得更加可靠和准确的图像信息以供观察和进一步处理, 它在地理信息系统、机器视觉以及生物医学工程等相关领域有广泛应用. 可见光与红外线成像机理不同, 前者主要依靠物体的光谱反射达到成像目的, 而后者则根据物体的热辐射来成像. 因此, 通常可见光具有丰富的背景信息, 能较好地描述场景中的环境信息, 而红外光能够给出较好的目标特性. 红外与可见光的融合就是利用这两种图像所具有的互补信息, 将可见光的背景信息与红外光的目标特征进行有机结合, 提高人或者机器的目标识别能力和环境释义能力^[1-4].

但是, 红外与可见光图像具有不同的边缘特性. 首先, 可见光边缘要比红外光边缘陡得多, 并且对应同一景物来说它们的边缘存在一定程度的缺失和偏移现象; 其次, 红外光图像中低频成分较多, 其相邻像素之间的相关度也相对较低, 这

些因素严重影响红外与可见光融合的精度. 目前, 国内对红外与可见光图像融合研究主要使用基于各种小波基的融合方法. 郑红等^[5]提出一种基于剪切波变换的红外光与可见光图像融合方法, 在主观视觉效果与客观量化指标性能上有一定的提高; 王珺等^[6]提出一种基于非下采样轮廓波 (non-subsampled contourlet transform, NSCT) 与稀疏表示的红外与可见光图像融合方法, 针对不同稀疏度的系数制定不同的融合规则, 取得了比较好的效果; 甘甜等^[7]提出一种基于分块离散余弦变换 (discrete cosine transformation, DCT) 编码小波的图像融合方法. Kong 和 Lei^[8]基于 NSCT 变换和色度-饱和度-强度 (intensity-hue-saturation, IHS) 变换提出一种改进的红外与可见光融合方法. 沈瑜等^[9]基于一种比较新的 Tetrolet 变换来进行红外与可见光图像的融合, 通过优化算法对各种规则融合后的稀疏系数进行恢复以获取融合图像, 有效减少了融合数据量. 但是这些方法在目标和背景的边缘保留上效果并不好, 并且很可能出现伪影.

本文提出一种基于深度模型分割的红外与可

* 国家自然科学基金 (批准号: 51375517)、重庆高校创新团队项目 (批准号: KJTD201313)、重庆工商大学校内青年博士基金 (批准号: 1352007) 和重庆市教委自然科学基金 (批准号: KJ1400628) 资助的课题.

[†] 通讯作者. E-mail: 149495263@qq.com

见光融合算法, 其流程如图 1 所示. 采用先验学习的方法自适应获得目标和背景的清晰轮廓, 能够在克服噪声干扰的同时有效保留目标与背景的边缘清晰度, 并分别对目标轮廓和背景轮廓采用相应的融合规则进行融合, 实验结果证明了本文算法的有效性.

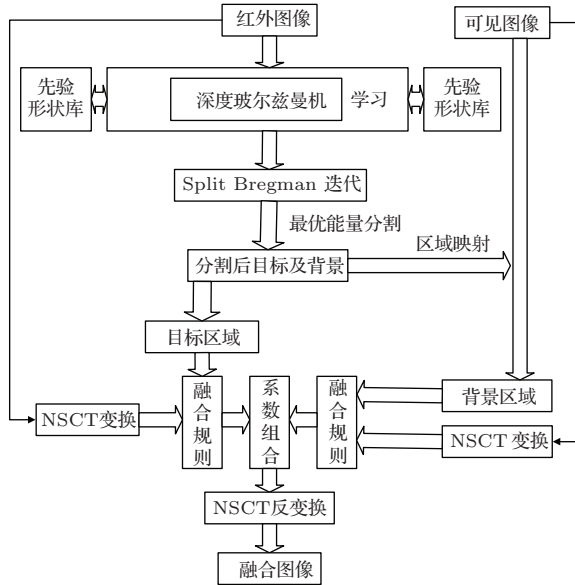


图 1 本文方法的流程图

2 深度模型的建立

深度学习模型由于具有对复杂数据的卓越学习能力而越来越受国内外学者的关注 [10]. 深度学习模型具有多层感知单元, 能通过组合底层特征形成更加高层的抽象数据、更好的表示数据 (图像、声音和文本) 的分布结构. 其中, 深度玻尔兹曼机 (deep Boltzmann machine, DBM) [11] 作为其中的一种结构而迅速发展. 在文献 [12] 中, Eslami 等提出了基于 DBM 的目标形状模板, 该方法不仅能够产生与实际情况相符合的目标形状, 还可以推广用于生成不同于先验库中的形状. 本节主要思路是通过先验形状的深度学习建立高层形状或是轮廓的估计, 并以能量的方式将其有效整合为底层变分分割模型, 建立能量方程. 最后通过对能量的优化获取所需轮廓.

2.1 深度玻尔兹曼机轮廓先验学习

深度玻尔兹曼机是受限玻尔兹曼神经网络的扩展, 包含了多个抽象的隐含层、层间全连接、层内无连接. 通常轮廓也可以描述为两层表示: 高层的

特征 (局部目标或者整体目标) 以及底层的特征 (边缘或角点等). 底层特征具有不变性, 而高层特征具有全局描述性, 并能够在噪声或者形变等外扰情况下很好地对目标轮廓进行理解. 本文中采用三层深度玻尔兹曼机来进行先验学习. 其模型与普通 BM 模型如图 2 所示.

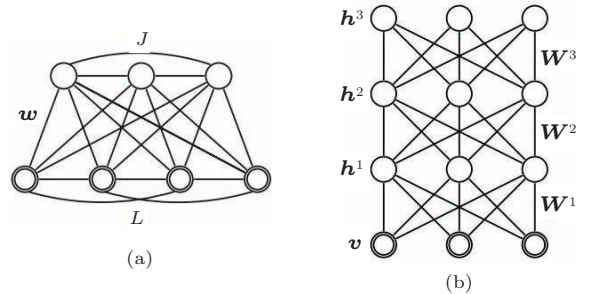


图 2 普通 BM 与深度 BM 模型 (a) 普通 BM 模型; (b) 本文使用 3 层 DBM 模型

假如用一个二维二值的可见单元向量 \mathbf{v} 表示一个任意形状, \mathbf{h}^1 与 \mathbf{h}^2 是底层与高层的二值隐含单元向量. 则深度玻尔兹曼机关于状态 $\{\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2\}$ 的能量可以定义为

$$E^{\text{DBM}}(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta) = -\mathbf{v}^T \mathbf{W}^1 \mathbf{h}^1 - \mathbf{h}^{1T} \mathbf{W}^2 \mathbf{h}^1 - a^{1T} \mathbf{h}^1 - a^{2T} \mathbf{h}^2 - b^T \mathbf{v}, \quad (1)$$

其中, $\theta = \{\mathbf{W}^1, \mathbf{W}^2, a^1, a^2, b\}$ 为模型参数, \mathbf{W}^1 与 \mathbf{W}^2 分别为从可见单元到隐含单元以及隐含单元到隐含单元的权值矩阵 (对称连接项), a^1 和 a^2 分别表示隐含单元的偏置 (自连接项), b 表示可见单元的自连接项. 则可见单位向量 \mathbf{v} 的概率为

$$P(\mathbf{v}, \theta) = \frac{1}{Z(\theta)} \sum_{\mathbf{h}^1, \mathbf{h}^2} \exp(-E^{\text{DBM}}(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta)), \quad (2)$$

其中, 常数项定义为

$$Z(\theta) = \sum_{\mathbf{v}} \sum_{\mathbf{h}^1, \mathbf{h}^2} \exp\{-E^{\text{DBM}}(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta)\}. \quad (3)$$

对于所给定对齐排列的训练形状 $\{v_1, v_2, \dots, v_n\}$ 深度玻尔兹曼机学习用来确定 (1) 式中的权值矩阵和自连接项.

2.2 分割模型的建立

在上一小节中, 通过深度玻尔兹曼机学习得到权值和偏置项, 则隐含确定了关于训练库中的概率

分布, 可以表示为

$$P(\mathbf{v}, \mathbf{h}^1, \mathbf{h}^2; \theta) \propto \exp(-E^{\text{DBM}}). \quad (4)$$

这三层学习能够有效捕捉先验形状层次结构、底层检测形状的简单局部特征然后反馈给高层, 使得高层可以捕捉更加复杂的全局性特征. 一旦隐含层状态确定, 则可以通过条件概率推理产生一个轮廓. 由于轮廓产生状态是以概率形式表示, 所以这里采用概率的方式^[13]来表示轮廓, 并将二维可见单元的向量 \mathbf{v} 用由概率定义的形状 q 进行定义. q 为轮廓概率表示 $q: \Omega \rightarrow [0, 1]$.

定义高层轮廓约束项:

$$E(q, \mathbf{h}^1, \mathbf{h}^2) = E_{\text{DBM}}(q, \mathbf{h}^1, \mathbf{h}^2, \theta), \quad (5)$$

其中 $\theta = \{\mathbf{W}^1, \mathbf{W}^2, a^1, a^2, b\}$ 为上节学习所得参数模型, 且隐含层状态 \mathbf{h}^1 和 \mathbf{h}^2 状态可以由逻辑函数

估计为

$$P(\mathbf{v}_i = 1 | \mathbf{h}^1) = \sigma \left(\sum_j \mathbf{W}_{i,j}^1 \mathbf{h}_j^1 + b_i \right), \quad (6)$$

其中, $\sigma(x)$ 为逻辑函数, 定义为

$$\sigma(x) = 1 / (1 + \exp(-x)). \quad (7)$$

定义底层模型能量公式为

$$E_i(q) = \int_{\Omega} r_o(x)q(x)dx + \int_{\Omega} r_b(x)(1-q(x))dx + \int_{\Omega} r_e(x)|\nabla q(x)|dx, \quad (8)$$

这里的 r_o 和 r_b 分别用来进行可见光与红外光图像目标与背景的区域描述, r_e 用来进行边缘检测. 给定一初始目标区域, 最小化 (8) 式能量函数可以得到目标的基于底层灰度区域和边界轮廓信息的分割结果. 结合 (5) 和 (8) 式建立总的能量方程:

$$E(q, \mathbf{h}^1, \mathbf{h}^2; \theta) = \underbrace{\|\nabla q\|_e + \alpha_1 q^T r}_{\text{data term}} - \alpha_2 \underbrace{(q^T \mathbf{W}^1 \mathbf{h}^1 + \mathbf{h}^{1T} \mathbf{W}^2 \mathbf{h}^2 + a^{1T} \mathbf{h}^1 + a^{2T} \mathbf{h}^2 + q^T b)}_{\text{shape term}}, \quad (9)$$

其中, α_1 和 α_2 为正常数, 数据项为 (8) 式的简化;

$$\|\nabla q\|_e = \int_{\Omega} r_e(x)|\nabla q(x)|dx$$

是一个赋权的全变分模, $r = r_o - r_b$. 这种三层结构的轮廓先验作为高层信息来指导底层的数据驱动, 并通过最小化能量函数以获得分割结果.

由于能量方程关于 q 为凸函数, 则可以采用 Split Bregman 算法^[12]对参数 q , \mathbf{h}^1 与 \mathbf{h}^2 进行优化, 迭代算法为:

1) 首先定义 $z^k = (c_1^k - u)^2 - (c_2^k - u)^2 - \alpha(\mathbf{W}^1 \mathbf{h}^1 + b)$;

2) 计算

$$(q^{k+1}, \mathbf{d}^{k+1}) = \arg \min \|\mathbf{d}\|_e + \alpha_1 q^T z^k + \frac{\lambda}{2} \|\mathbf{d} - \nabla q - \mathbf{e}^k\|^2;$$

3) $\mathbf{d}^{k+1} = \text{shrink}_q(\mathbf{e}^k + \nabla q^{k+1}, \lambda)$;

4) $\mathbf{e}^{k+1} = \mathbf{e}^k + \nabla q^{k+1} - \mathbf{d}^{k+1}$;

5) 计算 $\Omega_{\tau}^k = \{x : q^{k+1}(x) > \tau\}$;

6) 更新 $c_1^{k+1} = \int_{\Omega_{\tau}^k} u dx$ 和 $c_2^{k+1} = \int_{\Omega/\Omega_{\tau}^k} u dx$.

由步骤 1 到 6 进行迭代, 直至满足条件 $\|q^{k+1} - q^k\|^2 < \varepsilon$ 为止.

其中, u 为待分割图像, c_1 与 c_2 分别表示待分割图像 u 中待分割区域内部与外部的平均灰度值, 类似于文献^[14], \mathbf{d} 为引入的替换变量.

训练数据库选择 MPEG7-CE-Shape 数据库^[15], 从中选择 100 类目标 300 张来进行训练. 玻尔兹曼机中隐含层单元数量设置为 1600 和 800, 迭代次数设置为 500 和 200 次, 全局训练迭代次数 1000 次, 并在 Matlab 环境下进行仿真实验.

对一张有干扰的图像轮廓进行拟合, 结果如图 3 所示. 图中依此显示了迭代过程中图像的几个样本: 第一行中曲线最后精确匹配目标的边界; 第二行显示相应的轮廓概率表示, 通过 50 次迭代以后获得了几乎与轮廓一样的分割形状.

3 融合方法

3.1 区域映射

对于分割后的区域定义其属于目标区域的信任度:

$$C_i = \left[1 + e^{-\lambda_1(\mu_f - \mu_1)} \right]^{-1} \times \left[1 + e^{-\lambda_2(\mu_f - \mu_b - \mu_2)} \right]^{-1} \quad (10)$$

其中, μ_t 与 μ_b 分别表示第 i 个分割区域目标的前景区域灰度均值和背景区域灰度均值; λ_1 和 λ_2 控制 Sigmoid (S 函数) 函数的形状; μ_1 和 μ_2 为 S 函数的位移 (偏移量). 如果某一区域的灰度均值比背景区域的灰度均值高, 则该区域属于目标区域的信任度为 1, 否则该区域属于目标区域的信任度为 0.

然后, 将红外光图像的目标区域与背景区域映射到可见光图像中, 实现对原图像的区域分割.

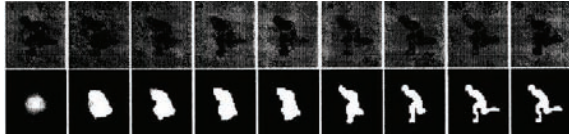


图3 分割算法的迭代过程

3.2 NSCT 分解

NSCT 变换^[16]在图像的分解和重构过程中取消了对图像的上采样与下采样, 使得其不仅具有多尺度性、局部性和方向性, 还具有平移不变特性以及子带图像之间具有相同尺寸大小等特性.

NSCT 变换主要由非下采样金字塔 (nonsub-sampled pyramid, NSP) 和非下采样方向滤波器组 (nonsub-sampled directional filter banks, NSDFB) 构成. 其中, NSP 滤波器满足 Bezout 恒等式 $H_0(z)G_0(z) + H_1(z)G_1(z) = 1$, 其中 $H_0(z)$ 与 $H_1(z)$ 分别为低通与高通分解滤波器, $G_0(z)$ 与 $G_1(z)$ 为对应合成滤波器; 与 NSP 结构类似, NSDFB 中分解滤波器 $U_0(z)$, $U_1(z)$ 与合成滤波器 $V_0(z)$, $V_1(z)$ 同样满足 Bezout 恒等式 $U_0(z)V_0(z) + U_1(z)V_1(z) = 1$.

本文采用 NSCT 变换对原红外图像进行分解, 即通过 NSP 以及 NSDFB 进行分解, 获得原红外图像在高频部分的子带系数 $C_{J,r}$ 以及低频部分的子带系数 C_J .

3.3 融合规则制定

对于目标区域来说, 为了尽可能保持红外图像的目标特性, 制定以下规则:

$$C^F(m, n) = C^I(m, n), (m, n) \in R_T, \quad (11)$$

其中, R_T 表示目标区域, 且 F 与 I 分别表示融合后的红外图像与源红外图像; $C(m, n)$ 为经 NSCT 分解后的高频子带和低频子带系数.

对于背景区域来说, 融合目的就是提取更多的背景细节信息. 由于红外与可见光图像在对应背景部分灰度值存在很大的差异, 所以针对其相应背景区域定义结构相似度^[17]:

$$S_{SSIM}(I_v, I_i) = \frac{(2\bar{I}_v \cdot \bar{I}_i + C_1)}{(\bar{I}_v^2 + \bar{I}_i^2 + C_1)} \frac{(2\sigma_{v,i} + C_2)}{(\sigma_v^2 + \sigma_i^2 + C_2)}, \quad (12)$$

其中, \bar{I}_v 与 \bar{I}_i 分别表示红外光图像与可见光图像所对应区域的均值; σ_i^2 和 σ_v^2 分别为对应区域的方差; $\sigma_{v,i}$ 为对应区域的协方差; 结构相似度中的参数 C_1 与 C_2 计算分别为 $(k_1 l)^2$ 与 $(k_2 l)^2$, l 为 255 (八位灰度图像像素值范围), k_1 与 k_2 分别为极小值常数 ($k_1 \ll 1, k_2 \ll 1$), 本文为计算方便取特殊情况时的值, 即 $C_1 = C_2 = 0$.

在系数选择的设定上, 本文设定一阈值 (取经验阈值 0.65), 定义以下的融合规则:

$$C_{J,r}^F(m, n) = \begin{cases} \underbrace{C_J^F(m, n) = (C_J^v(m, n) + C_J^i(m, n))/2}_{\text{low coefficient}} \\ \underbrace{\begin{cases} C_J^v(m, n) & \text{if } E_{J,r}^v(m, n) \geq E_{J,r}^i(m, n) \\ C_J^i(m, n) & \text{if } E_{J,r}^v(m, n) < E_{J,r}^i(m, n) \end{cases}}_{\text{high coefficient}} \\ \text{when } S_{SSIM}(v, i) \geq \varepsilon \\ \underbrace{\begin{cases} C_J^v(m, n) & \text{if } X^v(R) > X^i(R) \\ C_J^i(m, n) & \text{if } X^v(R) \leq X^i(R) \end{cases}}_{\text{high coefficient}} \\ \text{when } S_{SSIM}(v, i) < \varepsilon \end{cases}, \quad (13)$$

式中, $C_J(m, n)$ 与 $C_{j,r}(m, n)$ 分别为低频子带系数与 j 尺度 r 方向上的高频子带系数. 区域能量 $E_{J,r}(m, n)$ 与区域显著性 $X(R)$ 分别定义为

$$E_{J,r}(m, n) = \sum_{x=-(M-1)/2}^{(M-1)/2} \sum_{y=-(N-1)/2}^{(N-1)/2} \times |C_{J,r}(m+x, n+y)|^2, \quad (14)$$

$$X(R) = \frac{1}{|R|} \sum_{\forall(m,n) \in R} |C_{J,r}(m, n)|^2 \times \log |C_{J,r}(m, n)|^2, \quad (15)$$

其中, $|R|$ 为区域 R 中的像素个数, $M \times N$ 表示局部区域尺寸大小, 一般取 3×3 或者 5×5 . 为了算法的时效性, 本文算法中取值 5×5 .

经过系数组合规则处理以后, 对所获系数进行 NSCT 反变换, 得到最终的融合图像.

4 实验结果与分析

4.1 区域分割结果

图 4 为基于深度玻尔兹曼机模型的区域分割结果, 其中图 4(a), (c) 与 4(e) 分别为本文采用的三组红外与可见光图像的红外光图像; 而图 4(b),

4(d) 与 4(f) 分别为经过本文方法区域分割后的结果. 可以看出, 采用本文的深度玻尔兹曼机模型进行学习分割的算法能够有效获取准确的目标和背景轮廓, 有利于后面 NSCT 变换后的目标背景区域确定和制定融合规则.

4.2 融合结果及评价

通过三组可见光与红外光图像验证本文融合方法的性能, 并且用文献 [18] 中的 NSCT 融合方法、文献 [5] 中的 Shearlet 融合方法以及本文区域分割部分加上文献 [7] 中离散小波变换 (discrete wavelet transform, DWT) 方法 (DBM-DWT) 进行对比试验, 验证本文方法的有效性. NSCT 融合方法采用文献 [18] 中的参数设置标准, 分解层数设置为 3 层, 尺度分解滤波器采用 “maxflat” 滤波器, 方向分解滤波器采用 “dmaxflat” 滤波器, 分解级数由粗到细的设置为 2, 3, 3; Shearlet 融合方法中取 3 层尺度分解; 本文 NSCT 变换滤波器分解尺度数取 1 层分解, 其他设置与文献 [18] 一样; DWT 分解部分参数采用文献 [7] 参数设置. 试验所用的计算机配置为: Windows XP 的操作系统, Pentium(R) Dual-Core E5400@ 2.70 GHz 2.69 GHz CPU, 2G 内存; 算法编程所用平台为 Matlab7.0.1.

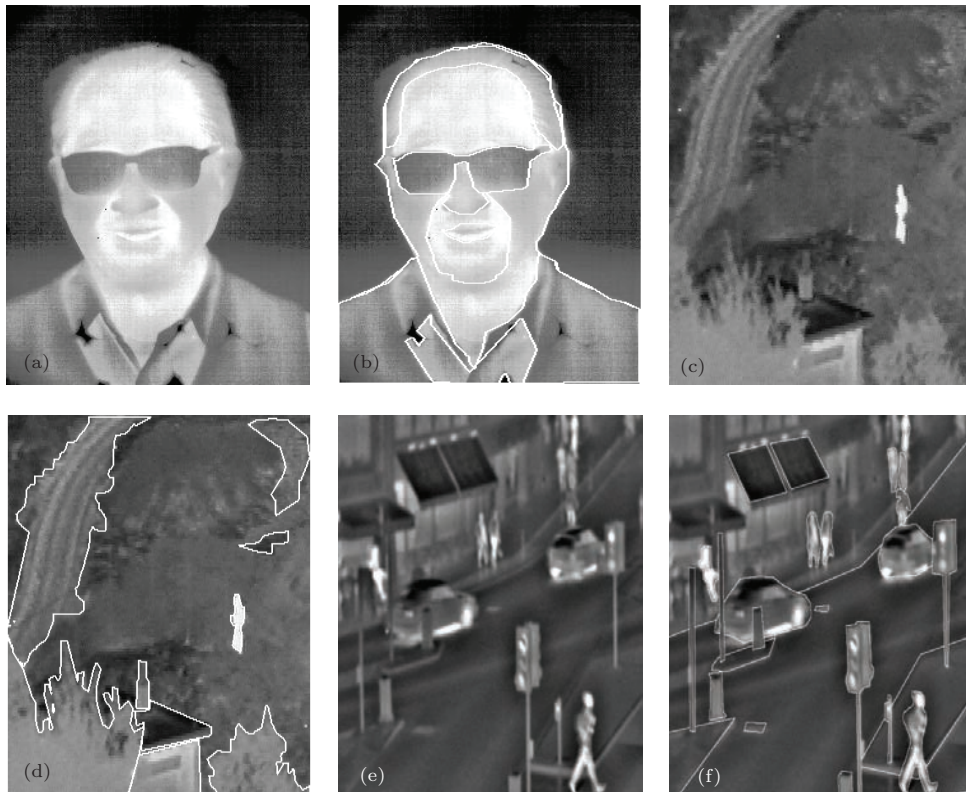


图 4 图像区域分割后的结果 (a) 红外图像 1; (b) 区域分割后的结果; (c) 红外图像 2; (d) 区域分割后的结果; (e) 红外图像 3; (f) 区域分割后的结果

第一组可见光与红外光图像的源图像为已配准 Equinox faces 图像^[19]. 其中, 图 5(a) 为可见光图像; 图 5(b) 为红外光图像; 图 5(c) 为 NSCT 融合方法结果, 基本上不具有重影和块状现象, 具有较高的清晰度, 体现了 NSCT 变换对于边缘信息的捕捉能力; 图 5(d) 为 Shearlet 方法融合结果, 其视觉效果稍好过前者, 细节更清晰; 图 5(e) 为 DBM-DWT 方法融合的结果, 分割区域部分内部对比度相对较高, 但是分割边缘部分有少许重影和块状现象出现; 图 5(f) 为本文融合方法, 相比而言对比度是最高的, 甚至可以捕捉到少量的纹理信息, 轮廓也相当完整, 充分说明了 NSCT 变换由于其平移不变性以及多方向表示性在图像融合中的应用优势是 DWT 变换不能替代的.

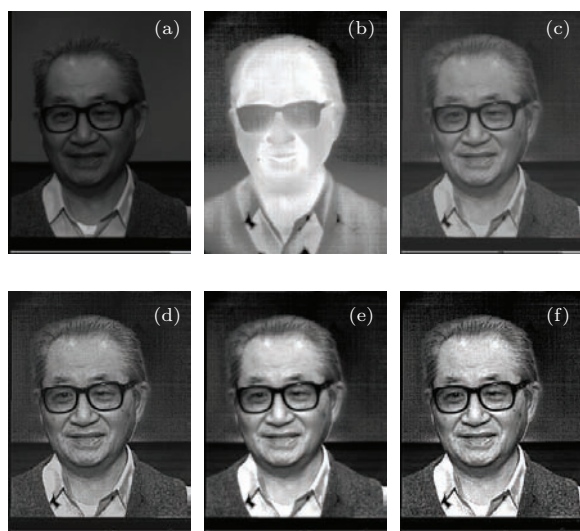


图 5 几种方法融合对比 (a) 可见图像; (b) 红外图像; (c) NSCT 方法; (d) Shearlet 方法; (e) DBM-DWT 方法; (f) 本文方法

第二组为配准后的 TNO UN Camp 图像^[19], 如图 6 所示. 图 6(a) 为原始红外光图像, 图 6(b) 为原始可见光图像. 通过对图 6(c)—(f) 四种融合方法结果进行对比可以发现, 本文方法融合结果所得对比度相对最高, 并保持了比较完整的图像边缘轮廓和纹理细节信息.

第三组为配准后的 Bristol Queen's Road 图像^[19], 如图 7 所示. 图 7(a) 为原始可见光图像, 图 7(b) 为原始红外光图像. 为了更进一步验证本文算法在噪声抑制上的优越性, 在原始图像上分别加上均值为 0 方差为 0.01 的高斯白噪声, 加噪声后图像如图 7(c) 和 (d). 由图 7(e)—(h) 可以明显看出, 由加噪融合结果图 7 和放大结果图 8 可以看

出, 经过本文 DBM 方法学习后的融合方法对于噪声的抑制效果最好. 说明基于深度玻尔兹曼机的分割方法具有的学习能力能够学习和分割出准确的目标和背景轮廓, 并且与纯粹的 NSCT 与 Shearlet 方法相比能够较好地抑制噪声. NSCT 变换和 Shearlet 变换由于融合时无法区分, 并且去除高频子带部分的噪声奇异点受噪声影响也很大, 相比于 DBM-DWT 方法本文方法对比度更高, 具有更高的清晰度.

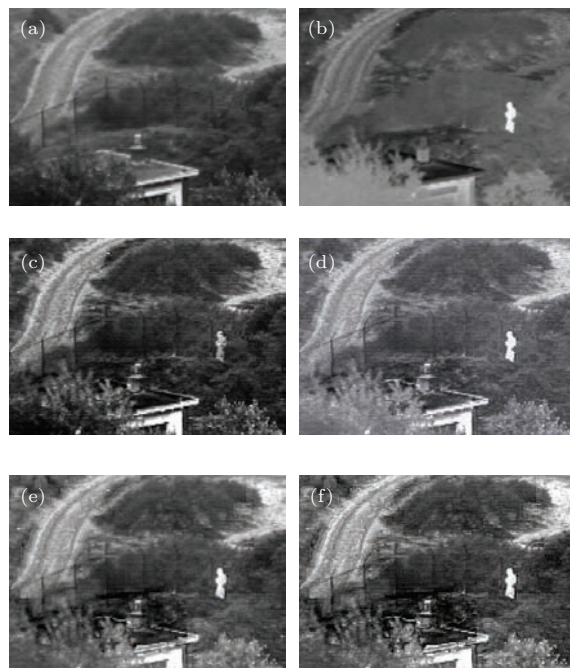


图 6 各种方法融合效果对比 (a) 可见图像; (b) 红外图像; (c) NSCT 方法; (d) Shearlet 方法; (e) DBM-DWT 方法; (f) 本文方法

图 8 为图 7 所有融合方法的图像进行局部区域放大后的对比, 其中图 8(a) 为文献 [18] 中 NSCT 方法的融合结果, 其轮廓更加清晰, 边缘特性保留较好, 且基本没有重影, 但是含有大量噪声, 说明对噪声没有抑制作用. 图 8(b) 为文献 [5] 中提出的基于改进的 Shearlet 变换融合方法的结果, 其边缘更加平滑, 失真小, 对比度更高, 但是同样没有噪声抑制作用. 图 8(c) 为 DBM+DWT 方法的融合结果, 可以看出, 由于其对噪声的抑制, 清晰度相对于前两种方法来说更高, 但是对比度相对较低, 区域边缘有少许失真. 图 8(d) 为本文方法的融合结果, 可以明显看出, 图像在保留了边缘特性的同时, 也克服了边缘震荡, 视觉效果好于前三种方法, 并且有效抑制了噪声干扰.

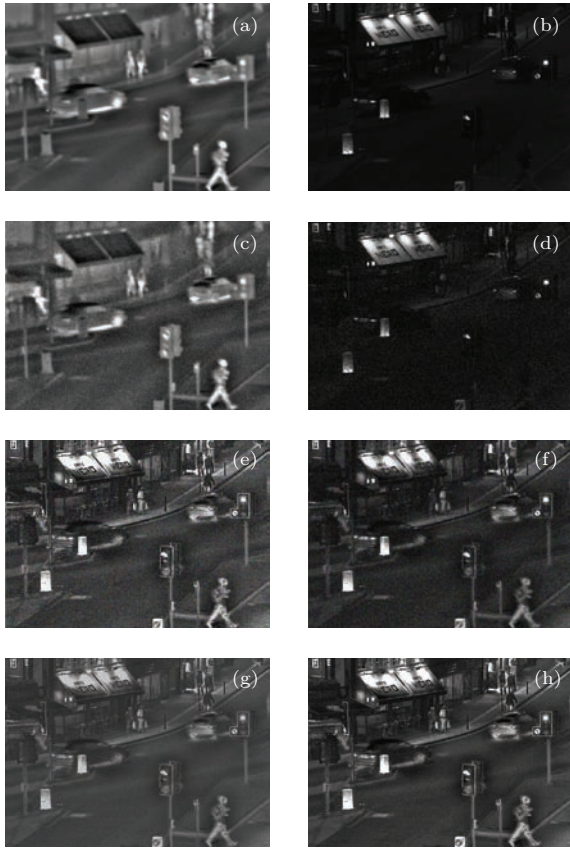


图7 各种方法融合效果对比 (a) 红外图像; (b) 可见图像; (c) 加噪红外图像; (d) 加噪可见图像; (e) NSCT方法; (f) Shearlet方法; (g) DBM-DWT方法; (h) 本文方法

为了客观地评价融合性能, 采用文献[20]中的五种图像评价指标来量化并进行比较, 分别为: IE (信息熵)、 AG (平均梯度)、 $PSNR$ (峰值信噪比)、 Q

(边缘保持度)、 T (图像恢复时间). 本文实验采用的各种融合方法在红外线与可见光图像上的融合定量评价指标见表1.

由表1可以看出, DBM-DWT方法除了在边缘保持度上稍低于NSCT方法与Shearlet方法以外, 其他指标与前两种方法相差不大, 有些指标甚至更好; 本文方法除了运行时间耗费比DBM-DWT方法、Shearlet方法和NSCT方法稍多之外, 其他指标相对于其他三种融合方法来说更好. 由第三组数据可以看出, 本文算法与DMB-DWT方法在噪声影响情况下比其余三种方法具有明显的抑制效果, 所以就综合性能来说本文算法具有一定的优势.

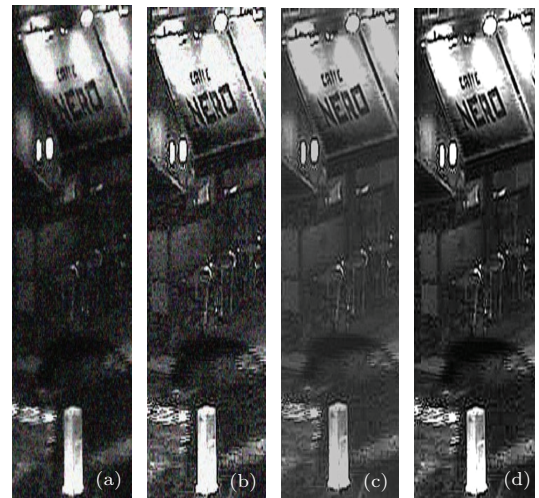


图8 局部区域融合效果对比 (a) NSCT方法; (b) Shearlet方法; (c) DBM-DWT方法; (d) 本文方法

表1 融合评价指标

图像	融合方法	IE	AG	$PSNR$	$Q_{AB/F}$	T/s
第一组可见与红外图像	NSCT	7.68	7.45	39.55	0.79	101
	Shearlet	7.88	7.61	38.26	0.76	289
	DBM-DWT	7.76	7.85	38.89	0.71	267
	本文方法	7.96	8.05	39.83	0.86	351
第二组可见与红外图像	NSCT	6.74	6.43	30.56	0.68	485
	Shearlet	6.53	6.15	32.47	0.71	541
	DBM-DWT	6.59	6.41	32.12	0.64	447
	本文方法	6.89	6.93	34.19	0.79	470
第三组可见与红外图像	NSCT	5.23	4.22	19.23	0.51	452
	Shearlet	5.44	4.86	21.23	0.55	561
	DBM-DWT	6.17	4.18	30.12	0.61	621
	本文方法	6.36	5.91	32.54	0.70	786

5 结 论

本文提出一种基于深度模型分割的红外与可见光融合算法,采用学习的方法自适应获得目标和背景的清晰轮廓,这样能够有效保留目标与背景的边缘清晰度并有效抑制噪声,然后分别对目标轮廓和背景轮廓采用相应的融合规则进行融合.本方法能够有效克服红外与可见光图像融合时容易受噪声干扰并产生伪影以致目标轮廓不鲜明、对比度低的缺点,最后通过实验证明了本文算法的有效性.

参考文献

- [1] Zhang C, Bai L F, Zhang Y 2007 *Acta Phys. Sin.* **56** 3227 (in Chinese) [张闯, 柏连发, 张毅 2007 物理学报 **56** 3227]
- [2] Zhao L Y, Ma Q L, Li X R 2012 *Acta Phys. Sin.* **61** 194204 (in Chinese) [赵辽英, 马启良, 厉小润 2012 物理学报 **61** 194204]
- [3] Ma J F, Hou K, Bao S L, Chen C 2011 *Chin. Phys. B* **20** 028701
- [4] Wang X Y, Wang Y X, Yun J J 2011 *Chin. Phys. B* **20** 104202
- [5] Zheng H, Zheng C, Yan X S 2012 *Chin. J. Sci. Instrum.* **33** 1613 (in Chinese) [郑虹, 郑晨, 闫秀生 2012 仪器仪表学报 **33** 1613]
- [6] Wang J, Peng J Y, He G Q 2013 *Acta Armament.* **34** 815 (in Chinese) [王珺, 彭进业, 何贵青 2013 兵工学报 **34** 815]
- [7] Gan T, Feng S T, Nie S P 2011 *Acta Phys. Sin.* **60** 114205 (in Chinese) [甘甜, 冯少彤, 聂守平 2011 物理学报 **60** 114205]
- [8] Kong W, Lei Y 2011 *IET Signal Processing* **5** 75
- [9] Shen Y, Dang J W, Feng X 2013 *Spectroscopy and Spectral Analysis* **33** 1506 (in Chinese) [沈瑜, 党建武, 冯鑫 2013 光谱学与光谱分析 **33** 1506]
- [10] Mohamed A, Dahl G, Hinton G 2011 *IEEE Trans. Audio, Speech, and Language Process.* **12** 134
- [11] Tang Y 2010 *In NIPS Workshop on Transfer Learning by Learning Rich Generative Models* Vancouver, B C, Canada, December 6, 2010 p2202
- [12] Eslami S M, Heess N, Winn J 2012 *Computer Vision and Pattern Recognition* Rhode Island Provine, America, 16 June, 2012 p983
- [13] Wang Z, Bovik A C 2004 *IEEE Trans. Image Process.* **13** 600
- [14] Goldstein T, Bresson X, Osher S 2010 *J. Sci. Comput.* **45** 272
- [15] Latech L, Lakamper R, Eckhardt U 2000 *IEEE Conference on Computer Vision and Pattern Recognition* Hilton Head, June 13, USA, 2000 p424
- [16] da Cunha A L, Zhou J P, Do M N 2006 *IEEE Trans. Image Process.* **15** 3089
- [17] Le Q V, Ngiam J, Coates A, Ng A Y 2011 *28th Int. Conf. Machine Learning Bellevue* Washington, USA, June 28, 2011 p1209
- [18] Zhang Q, Guo B L 2007 *J. Infrared Millim. Waves* **26** 185 (in Chinese) [张强, 郭宝龙 2007 红外与毫米波学报 **26** 185]
- [19] <http://www.imagefusion.org/.ftpquota>
- [20] Li X, Qin Y 2011 *IET Image Process* **5** 141

Infrared and visible image fusion based on deep Boltzmann model*

Feng Xin[†] Li Chuan Hu Kai-Qun

(College of Mechanical Engineering, Key Laboratory of Manufacturing Equipment Mechanism Design and Control of Chongqing, Chongqing Technology and Business University, Chongqing 400067, China)

(Received 19 February 2014; revised manuscript received 1 May 2014)

Abstract

In the infrared and visible light image fusion, the noise interference always exists. There is also the disadvantage that image fusion is easy to produce artifacts which cause blurred edge and low contrast. In order to solve these problems, in this study we propose an image fusion method based on deep model segmentation. First of all, deep Boltzmann machine is adopted to learn prior target and background contour and construct a contour deep segmentation model. After the optimal energy segmentation, Split Bregman iteration is used to obtain the infrared and visible image contour. Then non-subsampled contourlet transform is adopted to decompose the source images. The segmented background contour coefficients are fused by the structure similarity rule. Finally, the fused image is reconstructed by the non-subsampled contourlet inverse transform. The experimental results show that this algorithm can effectively obtain fused images with clear target contour and background contour. The fused images also have high contrast and low noise. The results show that it is an effective method of achieving the infrared and visible image fusion.

Keywords: deep model, deep Boltzmann machine, non-subsampled contourlet transform, image fusion

PACS: 42.30.Va, 42.30.Wb

DOI: [10.7498/aps.63.184202](https://doi.org/10.7498/aps.63.184202)

* Project supported by the National Natural Science Foundation of China (Grant No. 51375517), Chongqing University Innovation Team Project, China (Grant No. KJTD201313), the Dr Campus youth fund of Chongqing Technology and Business University, China (Grant No. 1352007), and the Natural Science Foundation of Chongqing City Board of Education, China (Grant No. KJ1400628).

[†] Corresponding author. E-mail: 149495263@qq.com