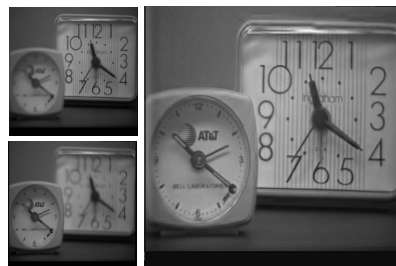




## 显著性分析在对焦图像融合方面的应用

张学典, 汪泓\*, 江旻珊, 蔡雨杏, 秦晓飞

上海理工大学光电信息与计算机工程学院, 上海 200093



**摘要:** 针对自动对焦技术中存在的全局对焦困难问题, 本文提出一种新的基于区域显著性分析的图像融合方法。首先用基于图论的显著性分析(GBVS)算法定位源图像中的聚焦区域, 然后使用分水岭和形态学方法进一步处理显著图的封闭区域以去除伪聚焦区域, 得到精确提取的聚焦区域; 离焦区域用剪切波变换处理后, 以 SML 算子选取有用的细节信息作为融合依据。最后将处理后的聚焦区域和离焦区域融合为全聚焦图像。实验证明, 所提出的方法融合图像边缘清晰, 细节丰富, 视觉效果最好, 并且在清晰度和融合度的评价指标上较传统方法提高 5% 以上。

**关键词:** 显著性分析; 图像融合; 对焦; GBVS; 剪切波变换

中图分类号: TP391

文献标志码: A

## Applications of saliency analysis in focus image fusion

Xuedian Zhang, Hong Wang\*, Minshan Jiang, Yuxing Cai and Xiaofei Qin

Institute of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China

**Abstract:** In the study of autofocus technology, we propose an image fusion method based on saliency analysis, which can solve the problem of all in focus. First, the focal area in the source image is positioned by the graph-based visual saliency (GBVS) algorithm, and then the watershed and morphological methods are used to obtain the closed area of the saliency graph and the pseudo-focus region is removed. The defocused region is processed by the Shearlet transform, and the SML operator is used to choose the fusion parts. Finally, the precise focused region and the processed defocused region are fused into an all in focus image. Experiments show that the fused image of our method is clear and rich in detail, which has the best visual effect, and improves more than 5% in definition and fusion compared with traditional methods.

**Keywords:** saliency analysis; image fusion; focus; GBVS; Shearlet transform

**DOI:** 10.3969/j.issn.1003-501X.2017.04.008

**Citation:** *Opto-Elec Eng*, 2017, **44**(4): 435-441

### 1 引言

光学系统受有限对焦范围的限制, 很难将视场中的目标全部清晰成像。为了得到清晰真实的场景, 许

多科学家提出了各种融合算法以获得包含特征信息和不同聚焦区域的全聚焦图像。多焦点图像融合技术的关键就在于如何从原图像中精确提取聚焦区域并去除伪聚焦区域<sup>[1,2]</sup>。现有的方法基本可以分为两类: 基于像素和基于分块。基于像素的融合方法虽能改善离焦图像的视觉效果, 但其对噪声和重合失调过于敏感, 容易丢失部分有用信息, 最后导致图像清晰度大大降

收稿日期: 2017-02-06; 收到修改稿日期: 2017-03-10

\*E-mail: 13162332050@163.com

低。基于分块的融合算法着重于精准识别原图像中的聚焦区域,保证了图像信息的完整,同时可以将图像本身的特征信息,如空间频率等作为融合参数。显著性分析就是一种基于分块的图像分析算法,它根据人类视觉显著性规律,利用图像颜色、强度等特征计算区域显著程度,区分前景和背景,检测不同深度物体的边缘,为精确分割聚焦区域提供有力帮助<sup>[3-7]</sup>。

Itti<sup>[8]</sup>提出了一种自底向上的显著性分析模型,并由此获得基于颜色、亮度和方向特征信息的多尺度显著性图。Harel<sup>[9]</sup>简化了 Itti 的模型并提出了一种基于马尔可夫随机链的图形视觉显著性模型(GBVS, graph-based visual saliency),该模型能在复杂背景中精确捕捉目标物的轮廓而获得显著性图。然而离焦区域中的高对比度部分可能会被误判为显著性区域,所以还需去除此类区域。考虑到空间频率值和图像的清晰度成正比,我们选择局部空间频率来检测并过滤掉伪聚焦区域<sup>[1,2]</sup>。

针对自动对焦技术中存在的全局对焦困难问题,本文提出一种新的基于区域显著性分析的图像融合方法,这种方法能尽可能地将原图像中的聚焦区域合并到融合图像中,利用显著性分析初步定位聚焦区域的边界,根据空间频率图过滤错误的边界轮廓,再用分水岭算法精确分割聚焦区域并去除伪聚焦区域,同时考虑到离焦区域所包含的纹理和方向信息,最后选择

剪切波变化来融合分割后余下的区域,在得到全聚焦图像的同时,极大程度上保留了原图像的细节信息。

## 2 方法与原理

本文提出的方法流程如图 1 所示。首先对原图像进行显著性分析,接着将显著性图二值化后,计算图像的空间频率以检测图像的模糊边缘,然后结合分水岭算法将图像划分为几个封闭区域并从中选取聚焦区域。为了移除伪聚焦区域,使用形态学方法确定聚焦区域的精确范围。源图像被分为聚焦区域和离焦区域后,聚焦区域直接融合,离焦区域经 Shearlet 变换后由 SML 算子提取融合系数,再与聚焦区域做最终的融合得到完整全聚焦图像。

### 2.1 显著性分析

根据多焦点图像的成像特点,离焦区域到聚焦区域的过渡是随景深的增加而逐渐过渡的,边界不明确,因此想要精确分割出聚焦区域相当困难<sup>[5]</sup>。显著性分析能够捕获显著区域的完整轮廓,用来检测聚焦区域。

显著性分析模拟人眼视觉注意机制,通过计算颜色、方向、亮度等特征信息得到图像的视觉显著区域。GBVS 引入了图形理论,图像的每个像素都被看作图形的节点。节点间的边缘信息呈现了不同节点间的颜色、纹理等差异。据此,GBVS 采用随机游走方法,从图上给定一点出发,随机地选择一个邻居结点并移

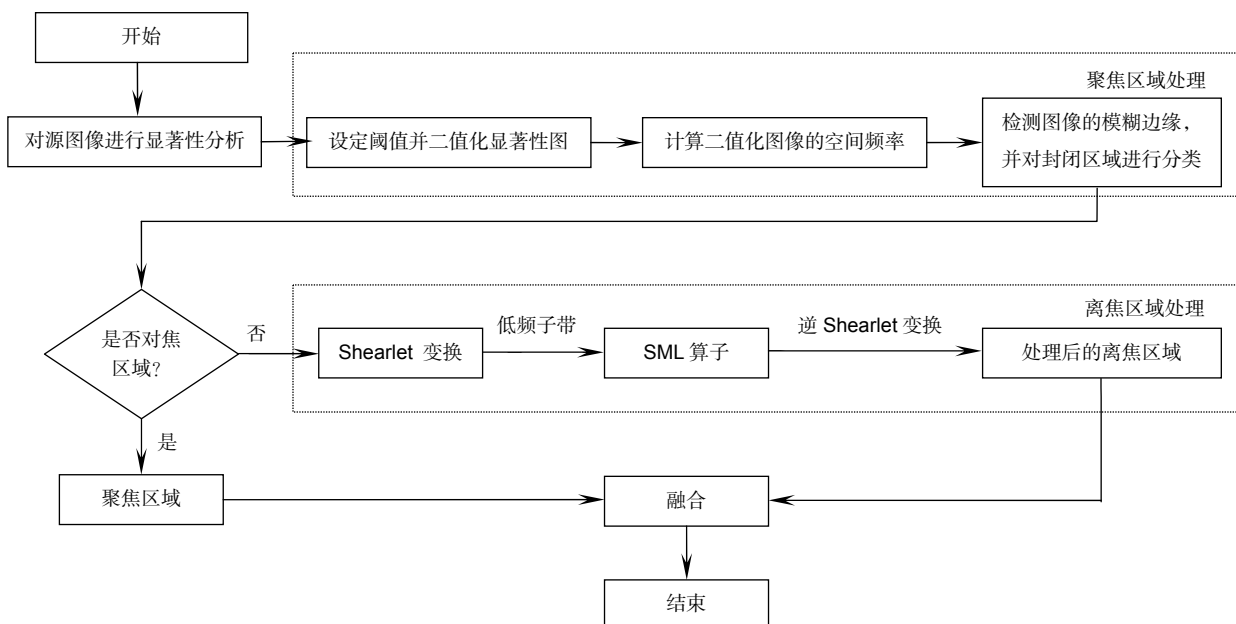


图 1 基于显著性分析的全局对焦方法流程图。

Fig. 1 The flow chart of autofocus method based on saliency analysis.

动到该点,然后把当前结点作为出发点,重复以上过程,此过程用于识别具有更大显著性值的节点,这样,GBVS 模型就能从全视场中定位重要目标并检测图像中的聚焦区域<sup>[1,6]</sup>。

为了生成特征映射并获取计算图像的平衡分布,我们通过计算特征向量,并构建马尔可夫链获得显著性图<sup>[1]</sup>。具体过程如下:

1) 用高斯金字塔低通滤波器和伽柏金字塔滤波器获得多尺度亮度和方向信息。

2) 在不同尺度和特征图中寻找马尔可夫平衡:设一幅特征图为  $M'$ , 由马尔可夫链可得:

$$dis(P, Q) \triangleq \left| \log_2 \frac{M'(P)}{M'(Q)} \right|. \quad (1)$$

$P$ 、 $Q$  是图  $M'$  上的点。根据式(1)可以计算任何两个像素之间的差,并建立像素之间的链接图。图上的每个顶点代表  $M'$  的一个像素,每条边表示两点之间的权重,由灰度差和点之间的距离决定。 $P(i_1, j_1)$  和  $Q(i_2, j_2)$  两点之间的权重为

$$w(P, Q) = dis(P, Q) \cdot \exp\left[-\frac{(i_1 - i_2)^2 + (j_1 - j_2)^2}{2\sigma^2}\right]. \quad (2)$$

通过式(2)能求解权重矩阵最大特征值的特征向量。

3) 根据不同特征图的平衡分布结果重新建立马尔可夫链,并且根据下式定义权重矩阵:

$$\bar{w}(P, Q) = bm(Q) \cdot \exp\left[-\frac{(i_1 - i_2)^2 + (j_1 - j_2)^2}{2\sigma^2}\right]. \quad (3)$$

式中  $bm(Q)$  是对应于平衡分布图的像素值<sup>[1,2]</sup>。

显著性分析的精确程度取决于原图像的亮度和方向信息,因此在分析显著性区域时,若离焦部分包含了明亮的平滑区域,就会导致离焦区域的部分图像被误认为是显著性区域<sup>[1,2]</sup>,此时就需要实现更精确的聚焦区域提取。

## 2.2 聚焦区域提取

在 2.1 节中提到视觉关注区域能由显著性分析获得,然而显著性区域的分布规律和图像的清晰度没有直接关系,并且离焦区域中的高对比度区域也会被误认为是显著性区域。不过,离焦光学成像系统可以看作一个低通滤波器,通过计算其清晰度就可以去除伪聚焦区域。空间频率对清晰区域十分敏感,可以通过灰度值的变化找到图像聚焦区域。空间频率图不仅包括聚焦区域,而且包括孤立的孔和其它均匀分布区域,根据聚焦成像方法,孤立的孔、小断裂和投影不应该

出现在多焦点图像中<sup>[3,8]</sup>。分水岭算法对灰度值的微妙变化十分敏感,可用于识别轻微的边界。通过分水岭算法作用后,几个封闭区域可以合并成封闭的均匀目标区域,结合分水岭算法和形态学算法,就能去除孤立的孔或断裂,获得更准确的聚焦区域<sup>[10-12]</sup>。

为了提取聚焦区域,首先计算显著性图的空间频率,定位显著性区域的边界;然后用分水岭算法分割出闭合区域;最后,根据聚焦区域没有内部孔和突起的原则,采用形态学方法去除小的伪聚焦区域以得到最终的聚焦区域。

## 2.3 离焦区域处理

传统的多分辨率分析方法,如 Wavelet 和 NSCT 等,只能捕获有限的方向信息,不能有效地沿轮廓方向提取边缘信息,在复杂结构和丰富纹理场景中描述高频特性的能力较弱。Labate<sup>[13]</sup>提出了 Shearlet 变换,它不仅能够利用图像的奇异性(奇异点、线和曲线)描绘全局特征,还可以自适应地跟踪奇异曲线的方向,从而在各种尺度和方向上表征图像特征。

假设分解层为 2,则 Shearlet 函数的虚拟变换系统如下:

$$LA(\psi) = \{\psi_{u,l,k}(x) = |\det LA|^{u/2} \psi(LB^l LA^u x - k)\}, \quad (4)$$

其中:  $u \in \mathbf{Z}$ ,  $k \in \mathbf{Z}^2$ ,  $LA$  和  $LB$  是  $2 \times 2$  的可逆矩阵,  $l$  和  $k$  分别是分解层和子带的数量,  $u$  是调整因子,

$$LA \in \left\{ LA \begin{vmatrix} LA_0 \\ LA_1 \end{vmatrix} = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}, LA_1 = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix} \right\}$$

是与尺度变换相关联的各向异性扩展矩阵;

$$LB \in \left\{ LB \begin{vmatrix} LB_0 \\ LB_1 \end{vmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, LB_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \right\}$$

是包含旋转或剪切变换的剪切矩阵,矩阵  $LA$  控制剪切变换的尺度,矩阵  $LB$  控制剪切变换的方向<sup>[14-16]</sup>。

离焦区域经过 Shearlet 变换后,得到的低频系数能反应图像的大多数信息。传统方法一般采用低频系数均值融合,算法虽然简单,但会丢失很多有用信息<sup>[15-16]</sup>,因此本文采取一种能够反映图像邻域聚焦程度的聚焦评价函数—改进拉普拉斯算子(SML 算子)作为低频系数的选择依据<sup>[14]</sup>,以最大程度提取图像聚焦部分的细节信息。形式如下:

$$ML(x, y) = |2f(x, y) - f(x - \Delta x, y) - f(x + \Delta x, y)| + |2f(x, y) - f(x, y - \Delta y) - f(x, y + \Delta y)|, \quad (5)$$

$$SML(i, j) = \sum_{i-N}^{i+N} \sum_{j-N}^{j+N} ML(i, j), \quad (6)$$

$SML(i, j)$  即点  $(x, y)$  的聚焦程度, 低频系数的选取规则为

$$F_0(i, j) = \begin{cases} A_0(i, j), SML^A(i, j) \geq SML^B(i, j) \\ B_0(i, j), SML^A(i, j) < SML^B(i, j) \end{cases} \quad (7)$$

### 3 实验结果与分析

为了有效评估本文方法的性能, 分别对四组多焦点图像(前三组为经典图像, 第四组为相机拍摄图像)进行融合实验。并将基于拉普拉斯金字塔变换的融合方法(Laplacian pyramid, LAP)、基于小波变换的融合方法(Wavelet)、基于非下采样轮廓波变换的融合方法(NSCT)与本文提出方法(基于 GBVS 的融合方法)进行融合性能对比, 融合规则相同。实验所用图像均已经过严格配准, 如图 2 所示。融合结果见图 3。

由融合效果图可以看出, 相对于其他方法, LAP 方法的融合图像模糊, 效果不理想; 小波变换融合图像整体亮度较暗, 对比度较低, 视觉效果比 LAP 方法有所提高; NSCT 变换融合图像虽然整体亮度和清晰度都提高了, 但边缘有波纹, 出现吉布斯现象; 而本文方法融合图像边缘清晰, 细节丰富, 视觉效果最好。

为了直观分析融合效果, 进一步采用信息熵(entropy, E)、 $Q^{AB/F}$  度量、交互信息(mutual information, MI)和空间频率(spatial frequency, SF)这四个定量评价指标对融合图像进行客观评价。信息熵反应图像细节信息的量;  $Q^{AB/F}$  衡量有多少边缘信息从源图像转移到了融合图像; MI 计算源图像有多少信息转移到了融合结果中; 而 SF 度量图像空间域的总体活跃度, 可反映图像对微小细节反差表达的能力。四个指标的值越大, 表明图像越清晰, 融合性能越好<sup>[17, 18]</sup>。

不同方法的客观评价指标如表 1 所示。基于 LAP 方法的融合图像质量在视觉上低于其他三种方法, 这一点在客观评价指标上也有体现。基于小波的方法和基于 NSCT 的方法在视觉上的差别不大, 从指标上看, 这两种方法的差别也是微乎其微, 但基于 NSCT 的方法略胜一筹。而基于 GBVS 方法的图像的信息熵较之基于 NSCT 方法高约 1%,  $Q^{AB/F}$  度量较之 NSCT 方法高约 4%, MI 较之 NSCT 方法高约 2%, SF 度量较之 NSCT 方法高约 2%, 虽然有个别实验显示评价参数略低, 但总体上看, 基于 GBVS 的方法在清晰度和融合度方面有明显优势。

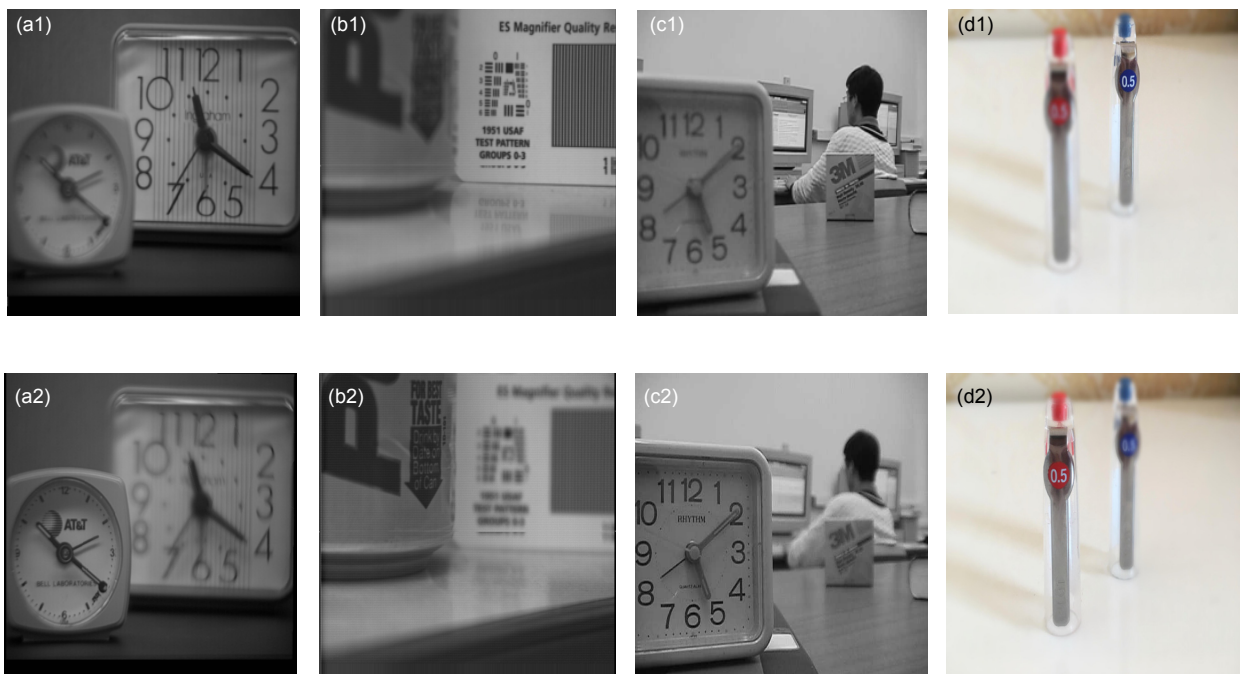


图 2 实验原图. (a1) Clock 右对焦. (a2) Clock 左对焦. (b1) Pepsi 右对焦. (b2) Pepsi 左对焦. (c1) Lab 右对焦. (c2) Lab 左对焦. (d1) Pen 右对焦. (d2) Pen 左对焦.

Fig. 2 The source image of the experiment. (a1) Clock right focus. (a2) Clock left focus. (b1) Pepsi right focus. (b2) Pepsi left focus. (c1) Lab right focus. (c2) Lab left focus. (d1) Pen right focus. (d2) Pen left focus.

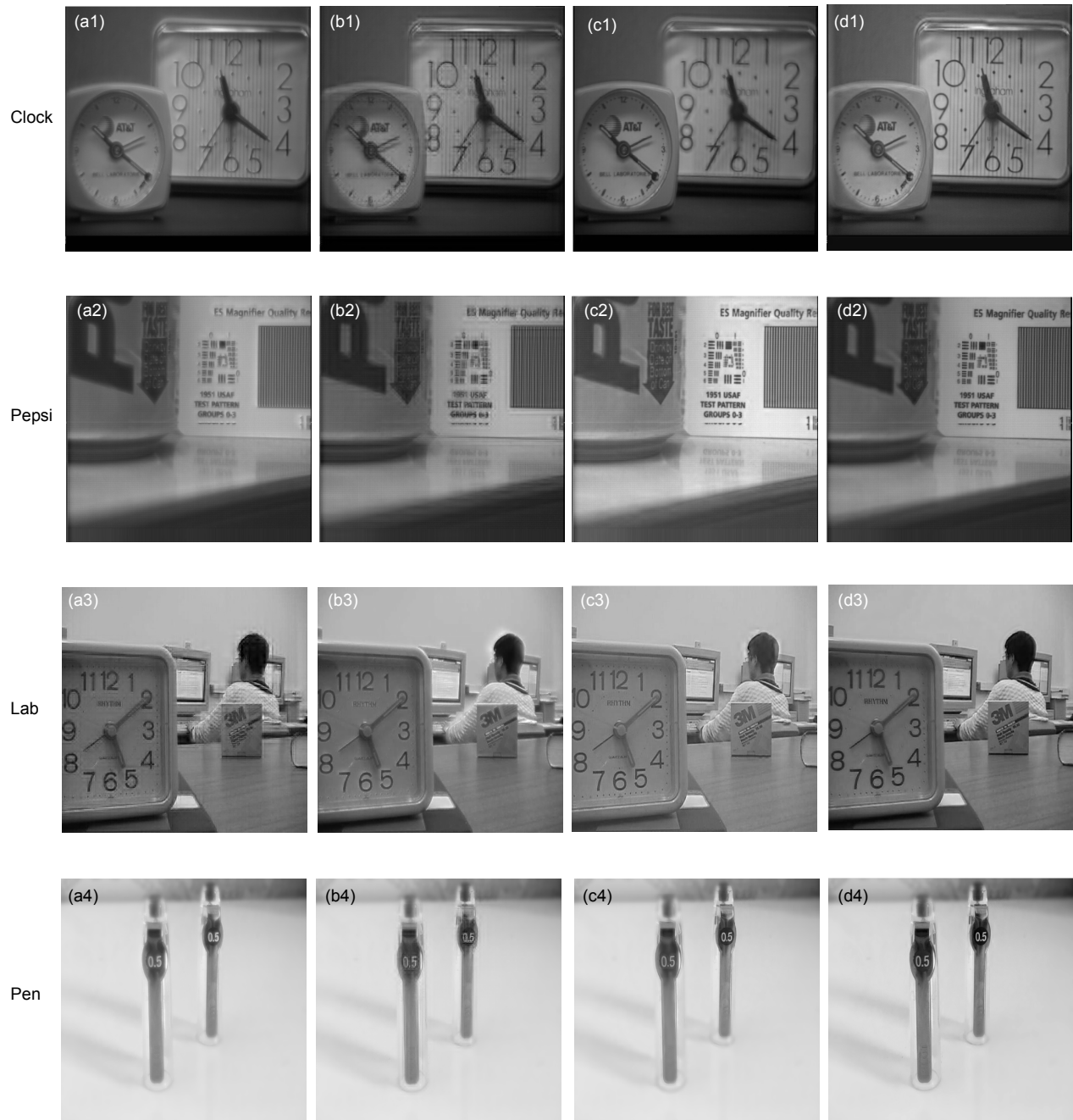


图3 四种融合方法的融合结果. (a1)~(a4) LAP. (b1)~(b4) Wavelet. (c1)~(c4) NSCT+PCNN. (d1)~(d4) GBVS.

Fig. 3 The results of 4 fusion methods. (a1)~(a4) LAP. (b1)~(b4) Wavelet. (c1)~(c4) NSCT+PCNN. (d1)~(d4) GBVS.

将表1中的数据以折线图的形式表现,可以更加直观地分析各个融合方法的性能优劣。各方法的融合质量对比如图4所示。

从图4可以看出,本文算法在大多数情况下都优于其余三种算法,仅clock图的信息熵和lab图的互信度低于NSCT方法,其余客观评价指标都高于NSCT方法以及其他两种方法,说明本文模型具有较强的鲁

棒性。主观视觉效果和客观评价指标的结果都有力证明了本文方法是一种有效的图像融合方法。

分析原因在于两点:首先,对原图像做了显著性分析,为之后的聚焦区域提取做了准备工作,大大提高了聚焦区域提取的精确度。其次,离焦区域中的有用信息被二次筛选,为融合图像保留了更多原图像的细节信息。

表 1 图像融合质量评价.

Table 1 Image fusion quality evaluation.

图像	实验方法	评价指标			
		E	$Q^{AB/F}$	MI	SF
Clock	LAP	2.8318	0.5899	6.9354	13.0092
	wavelet	2.8972	0.6837	6.9995	13.6347
	NSCT	2.9748	0.6849	7.0747	13.9053
	GBVS	2.9223	0.7168	7.2397	14.2661
Pepsi	LAP	2.8516	0.6477	6.8282	13.1893
	wavelet	2.9811	0.7695	6.8663	13.7388
	NSCT	2.9792	0.7673	6.9228	13.9192
	GBVS	2.9835	0.7826	7.0925	14.1657
Lab	LAP	2.8388	0.5947	7.0148	13.0805
	wavelet	2.8726	0.7014	7.0473	13.8231
	NSCT	2.9517	0.7002	7.1846	14.0397
	GBVS	2.9892	0.7301	7.1435	14.3055
Pen	LAP	2.8201	0.6845	7.0198	13.6592
	wavelet	2.8593	0.7113	7.0285	13.9003
	NSCT	2.9461	0.7164	7.0302	14.2635
	GBVS	2.9904	0.7901	7.0993	14.5338

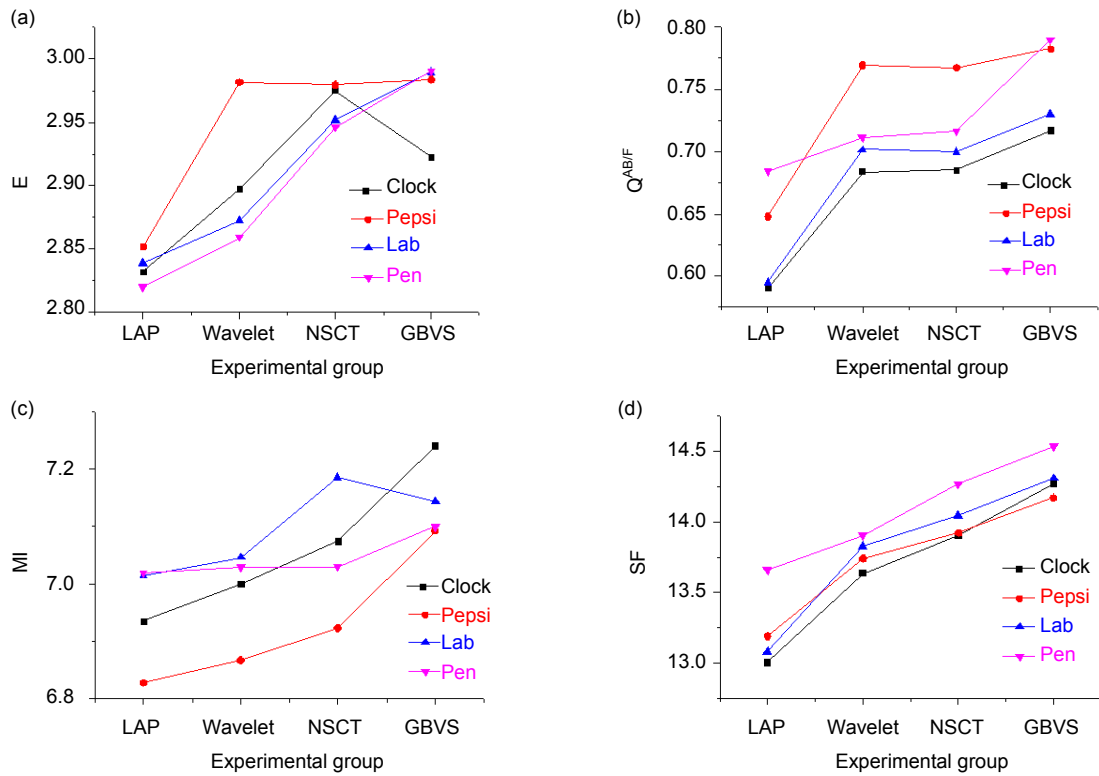


图 4 图像融合质量评价对比. (a) E. (b)  $Q^{AB/F}$ . (c) MI. (d) SF.

Fig. 4 Comparison of image fusion quality evaluation. (a) E. (b)  $Q^{AB/F}$ . (c) MI. (d) SF.

## 4 结论

本文从自动对焦技术中的全局对焦问题出发,提出了一种基于显著性分析的图像融合方法。首先用基于图论的显著性分析(GBVS)算法定位源图像中的聚焦区域,然后使用分水岭和形态学方法来获得封闭的聚焦区域并去除伪聚焦区域。最后,聚焦区域直接融合,离焦区域用剪切波变换处理,以 SML 算子选取融合系数后融合。方法提出后,在经典 clock、pepsi、lab 图像和普通相机拍摄的多聚焦图像上进行实验,并与传统方法做主观视觉和定量评价指标的比较,证实了此基于显著性分析的自动对焦方法的可行性和优越性。实验证明,所提出的方法可以精确提取聚焦区域,融合图像保留更多的特征,最终的融合图像清晰度和融合度较之传统方法提高了 5% 以上。但此方法的算法对彩色图像有限制,后续还要对彩色多焦点图像的融合做进一步的研究。

## 基金项目

国家重大仪器专项(2014YQ09070903、2013YQ03065104); 国家高技术研究发展计划(2015AA0200751)

## 参考文献

- Zhang Baohua, Lv Xiaoqi, Pei Haiquan, et al. Multi-focus image fusion algorithm based on focused region extraction[J]. *Neurocomputing*, 2016, **174**: 733–748.
- Zhang Baohua, Pei Haiquan, Lv Xiaoqi. Multi-focus image fusion based on saliency detection and sparse representation[J]. *Journal of Chinese Computer Systems*, 2016, **37**(7): 1604–1607. 张宝华, 裴海全, 吕晓琪. 基于显著性检测和稀疏表示的多聚焦图像融合算法[J]. *小型微型计算机系统*, 2016, **37**(7): 1604–1607.
- Yan Xiang, Qin Hanlin, Li Jia, et al. Multi-focus image fusion using a guided-filter-based difference image[J]. *Applied Optics*, 2016, **55**(9): 2230–2239.
- Duan Jiangyong, Meng Gaofeng, Xiang Shiming, et al. Multifocus image fusion via focus segmentation and region reconstruction[J]. *Neurocomputing*, 2014, **140**(22): 193–209.
- Zhang Likai. The research of more focus image fusion method[D]. Changchun: Jilin University, 2013: 1–59. 张立凯. 多焦点图像融合方法的研究[D]. 长春: 吉林大学, 2013: 1–59.
- Luo Xiaoyan, Zhang Jun, Dai Qionghai. A regional image fusion based on similarity characteristics[J]. *Signal Processing*, 2012, **92**(5): 1268–1280.
- Chai Yi, Li Huafeng, Li Zhaofei. Multifocus image fusion scheme using focused region detection and multiresolution[J]. *Optics Communications*, 2011, **284**(19): 4376–4389.
- Itti L, Koch C, Niebur E. A model of saliency-based visual attention for rapid scene analysis[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1998, **20**(11): 1254–1259.
- Harel J, Koch C, Perona P. Graph-based visual saliency[M]//Schölkopf B, Platt J, Hofmann T. *Advances in Neural Information Processing Systems*. Cambridge, MA: MIT Press, 2006: 545–552.
- Behura J, Wapenaar K, Snieder R. Autofocus imaging: image reconstruction based on inverse scattering theory[J]. *Geophysics*, 2014, **79**(3): A19–A26.
- Zhang Xuedian, Liu Zhaoqing, Jiang Minshan, et al. Fast and accurate auto-focusing algorithm based on the combination of depth from focus and improved depth from defocus[J]. *Optics Express*, 2014, **22**(25): 31237–31247.
- Li Huafeng, Chai Yi, Li Zhaofei. Multi-focus image fusion based on nonsubsampling contourlet transform and focused regions detection[J]. *Optik-International Journal for Light and Electron Optics*, 2013, **124**(1): 40–51.
- Guo Kanghui, Labate D. Optimally sparse multidimensional representation using shearlets[J]. *SIAM Journal on Mathematical Analysis*, 2007, **39**(1): 298–318.
- Wang Fei, Wang Yao, Shi Caicheng. Multi-focus image fusion using shearlet transform[J]. *Computer Engineering and Applications*, 2016, **52**(2): 205–208. 王飞, 王瑶, 史彩成. 采用 shearlet 变换的多聚焦图像融合[J]. *计算机工程与应用*, 2016, **52**(2): 205–208.
- Zheng Hong, Zheng Chen, Yan Xiusheng, et al. Visible and infrared image fusion algorithm based on shearlet transform[J]. *Chinese Journal of Scientific Instrument*. 2012, **33**(7): 1613–1619. 郑红, 郑晨, 闫秀生, 等. 基于剪切波变换的可见光与红外图像融合算法[J]. *仪器仪表学报*, 2012, **33**(7): 1613–1619.
- Liao Yong, Huang Wenlong, Shang Lin, et al. Image fusion based on Shearlet and improved PCNN[J]. *Computer Engineering and Applications*, 2014, **50**(2): 142–146. 廖勇, 黄文龙, 尚琳, 等. Shearlet 与改进 PCNN 相结合的图像融合[J]. *计算机工程与应用*, 2014, **50**(2): 142–146.
- Li Meili, Li Yanjun, Wang Hongmei, et al. Fusion algorithm of infrared and visible images based on NSCT and PCNN[J]. *Opto-Electronic Engineering*, 2010, **37**(6): 90–95. 李美丽, 李言俊, 王红梅, 等. 基于 NSCT 和 PCNN 的红外与可见光图像融合方法[J]. *光电工程*, 2010, **37**(6): 90–95.
- Hong Yuzhen, Ren Guoqiang, Sun Jian. Analysis and improvement on sharpness evaluation function of defocused image[J]. *Optics and Precision Engineering*, 2014, **22**(12): 3401–3408. 洪裕珍, 任国强, 孙健. 离焦模糊图像清晰度评价函数的分析与改进[J]. *光学精密工程*, 2014, **22**(12): 3401–3408.