

## 基于视觉注意模型的红外图像分级压缩方法

苏娟, 杨罗, 卢俊

(第二炮兵工程大学 907 室, 陕西 西安 710025)

**摘要:** 针对红外成像末制导武器人在回路模式中回传图像高压缩比与高保真度之间的矛盾, 提出了一种基于视觉注意模型的红外图像分级压缩方法。在采用视觉注意模型提取得到红外图像显著区域的基础上, 按照显著区域、背景区域、过渡区域对图像进行三级划分, 然后根据不同区域图像中信息的重要程度采取不同的压缩方式, 进而实现对整幅图像的压缩。实验结果表明: 该方法在保留目标重要信息的同时, 大大减少了图像传输的数据量, 与传统的整图压缩方法相比, 更适用于人在回路的应用场合。

**关键词:** 人在回路; 视觉注意模型; 显著区域; 图像压缩

**中图分类号:** TP391    **文献标志码:** A    **文章编号:** 1007-2276(2014)06-2040-06

## Hierarchical infrared image compression method based on visual attention model

Su Juan, Yang Luo, Lu Jun

(Section 907, The Second Artillery Engineering University, Xi'an 710025, China)

**Abstract:** Aiming at the contradiction between high compression ratio and high fidelity of the return image in man-in-the-loop (MITL) of infrared imaging guidance weapons, a hierarchical infrared image compression algorithm based on visual attention model was proposed. On the basis of salient regions of infrared image extracted by visual attention model, salient regions, background regions and transition regions were divided and compressed by different compression methods according to the importance of information of different image regions, in this way the compression of the whole image was realized. The experimental result shows that the proposed algorithm can not only preserve the important information of typical targets, but also decrease the data size in image transmission, and it is more suitable for MITL operation mode than the traditional full image compression method.

**Key words:** man-in-the-loop; visual attention model; salient region; image compression

收稿日期: 2013-10-14; 修订日期: 2013-11-17

基金项目: 国家自然科学基金(61132008, 61302195)

作者简介: 苏娟(1973-), 女, 博士, 副教授, 主要从事遥感图像处理和目标识别方面的研究。Email: suj04@mails.tsinghua.edu.cn

## 0 引言

人在回路(Man-in-the-loop)是目前精确制导技术的重要研究方向,在提高精导武器的突防能力、命中精度、对抗能力、增加打击与毁伤效果等方面发挥着重要作用<sup>[1]</sup>。人在回路是指飞行器在飞行末段将导引头获取的目标实时图像通过卫星数据链回传到地面指控中心,再由地面操作人员根据目标图像对目标进行识别和锁定,形成目标指示信息和制导指令,再通过数据链回传给飞行器,从而实现复杂环境下的目标识别与指示。为尽可能减小导引头获取目标图像时刻和导弹获取目标指示信息时刻之间的时延,提高人在回路制导精度,必须减小数据上下行传输时间。在数据链路传输速度一定的情况下,一种可行的解决方法是在飞行器上对导引头获取的目标实时图像进行压缩,减少数据链路的传输时间<sup>[2]</sup>。

常用的图像压缩方法通过对全图进行压缩得到高压缩比的图像,这种方法针对性不强,在减少数据量的同时造成了图像质量的下降,使得目标不易区分,对地面操作手的观察和判断造成影响。根据视觉认知理论,人类的视觉注意系统能够快速地从场景中注视到最显著的区域,目标常常位于某一显著区域之中<sup>[3]</sup>。因此,如果在回传图像时预先提取出图像中的显著区域,对目标区域与背景区域分别采用无损压缩和有损压缩,则能在保留目标重要信息的同时降低数据量,有利于地面操作手迅速准确地定位目标。目前最有影响力的视觉注意机制模型是1998年Itti<sup>[4-5]</sup>等人提出的Itti模型,该算法简单且计算速度较快,可以有效检测并提取图像中的显著目标区域。

鉴于此,文中针对红外成像末制导的人在回路工作模式,提出了一种基于视觉注意模型的分级图像压缩方法,首先利用改进的Itti模型提取显著区域,然后根据图像中各区域不同的重要性,采用不同的压缩比来压缩图像,从而实现对整个图像的压缩。

## 1 文中算法

文中算法流程如图1所示,主要分为基于视觉注意模型的显著图生成、基于显著度的区域划分和分级图像压缩三部分。

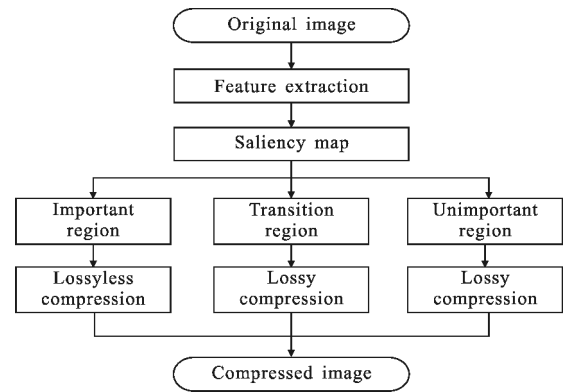


图1 文中算法流程

Fig.1 Flowchart of the proposed algorithm

### 1.1 基于视觉注意模型的显著图生成

文中采用Itti模型进行视觉显著图生成。Itti模型所采取的特征包括图像的亮度、方向和颜色特征。文中所处理的对象为红外图像,缺乏丰富的颜色信息,若直接采用Itti模型来生成显著图像将不能得到较好的效果。因此,根据红外图像的特点和典型目标的红外特性,在Itti模型框架下引入局部熵<sup>[6-7]</sup>特征替换颜色特征加以改进。局部熵反映了图像能量在空间分布的统计特性,局部熵越大,像素灰度分布的无序程度就越高,信息量就越多;反之,信息量越少。

局部熵定义为<sup>[8]</sup>:

$$H_{D,R_x} = - \sum_i P_{D,R_x}(d_i) \log_2 P_{D,R_x}(d_i) \quad (1)$$

式中: $P_{D,R_x}(d_i)$ 表示灰度值等于 $d_i$ 的像素数与 $R_x$ 局部区域总像素数之比, $d_i \in [0, 255]$ , $D = \{d_1, d_2, \dots, d_r\}$ 。

图2给出了基于视觉注意模型的显著图生成示意图。其步骤如下<sup>[9]</sup>:

(1) 分别构造亮度、方向、局部熵特征的高斯金字塔,通过对逐层进行低通滤波处理和降采样,使每个特征通道产生9个尺度的金字塔;

(2) 利用中央-周边差计算各特征通道的差异图,该方法与人类视觉细胞类似,具有对感受中央信息敏感、对感受周边信息抑制的响应特性;

(3) 通过跨尺度融合将各尺度的差异图像进行融合,得到亮度、方向、局部熵的特征图,并对各特征通道的特征图进行归一化处理;

(4) 对归一化后的各特征通道的特征图进行线性融合,生成最终的显著图。

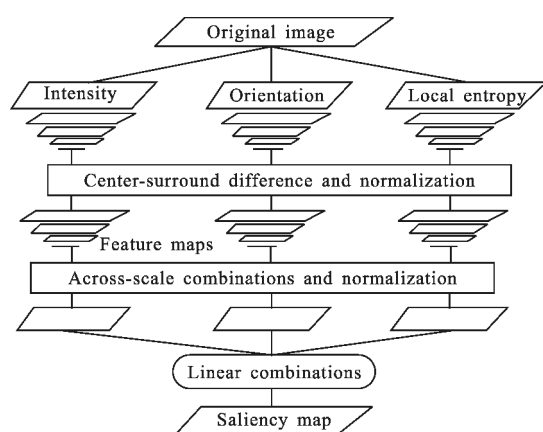
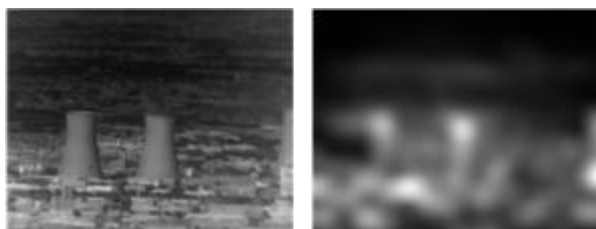


图 2 显著图生成算法流程

Fig.2 Flowchart of saliency map generation

图 3 显示了融入局部熵的改进模型和 Itti 模型在同一幅红外图像上生成的显著图比较。可以看出,融入局部熵的改进模型生成的显著图更加符合人类视觉的注意区域。由于局部熵特征是区域信息复杂程度的体现,红外图像中典型人造目标(如电厂冷却塔、雷达天线罩等)区域内部的灰度值分布比较均匀,与周边区域相比,其熵值偏低,在进行中央-周边差异后能够突出目标区域,因此,局部熵特征与亮度和方向特征相结合能够很好地反映目标区域的信息,从而得到很好的显著图生成效果。



(a) 原图

(b) Itti 模型

(a) Original image

(b) Itti model



(c) 改进模型

(c) Improved model

图 3 显著图生成

Fig.3 Saliency map generation

## 1.2 基于显著度的区域划分

为解决人在回路模式中图像压缩时存在的高压缩

比与高保真度之间的矛盾,文中在确保包含目标的重要区域观测数据信息不损失的前提下,对不同重要程度的区域进行有选择性的压缩。因此,文中通过对视觉显著图的分析,以显著度作为重要程度的指标,根据各区域的重要程度,将图像区域分为以下三类区域:

(1) 重要区域。即可能包含目标信息的区域,文中将采用视觉注意模型提取得到的显著区域作为重要区域;

(2) 非重要区域。即不包含目标信息的区域,文中将显著度值较低的、目标所处的背景区域作为非重要区域;

(3) 过渡区域。即显著区域与背景区域之间的区域。设置过渡区域的原因在于,一方面该类区域的重要性无法确定,另一方面避免解压重构后图像的目标区域与背景区域之间因压缩比差距过大导致的边界效应。

在确定重要区域时,在生成的显著图上,根据赢家全胜和返回抑制机制,实现注意焦点的选择与转移,注意焦点即为显著图中当前显著度值最大的位置,显著区域即为以该焦点为圆心,原图像宽度与高度两者最小值的  $1/6$  为半径的圆。根据经验,设定注意焦点数为 5,即场景中显著目标区域为 5 个,这在一般情况下是不会漏失目标的。为方便图像压缩,在压缩阶段将显著区域用矩形表示。

对每一个显著区域,其对应的过渡区域按照如下方法确定:如图 4 所示,以显著区域边界为内边界,设定一个外边界向外扩展的矩形框(白色矩形框)。计算该显著区域的平均显著度,然后计算白色矩形框的平均显著度,当白色矩形框的平均显著度为显著区域平均显著度的  $2/3$  时,停止向外扩展外边界。此时的白色矩形框即为该显著区域的过渡区域。对所考虑的尺寸为  $320 \times 240$  的红外图像,过渡区域的最小尺寸为 10 个像素。

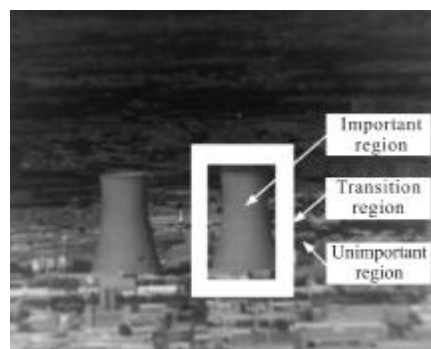


图 4 区域划分

Fig.4 Region division

图 4 给出了三类区域的划分示意图。划分时,图像中存在多个显著区域,因此区域之间可能出现重叠交叉的现象,区域类别按照以下优先顺序确定:重要区域、过渡区域和非重要区域。

### 1.3 分级图像压缩

经过区域划分之后,对不同的区域采用分级压缩的策略:对重要区域(即显著区域)采取无损压缩,其压缩比较低,图像保真度较高;对非重要区域(即背景区域)进行有损压缩,其压缩比较高,图像保真度较低;对于过渡区域则采用介于前两者的压缩比,目的在于消除因压缩比差距过大导致显著区域与背景区域之间产生边界效应。

文中对于显著区域采用改进的 Huffman 编码<sup>[9]</sup>的方式进行无损压缩,Huffman 编码是消除编码冗余的最常用技术,当对独立信息源的符号进行编码的时候,Huffman 编码对每个信源符号生成可能的最小数量的编码符号。对背景区域和过渡区域均利用 EZW 编码算法<sup>[10]</sup>进行有损压缩,该算法是一种小波系数量化编码方法,具有较好的均方误差性能和适中的编码复杂度,且具有嵌入式比特流的优点。考虑到各显著区域间可能存在重叠部分,对于相同的位置区域只压缩一次。

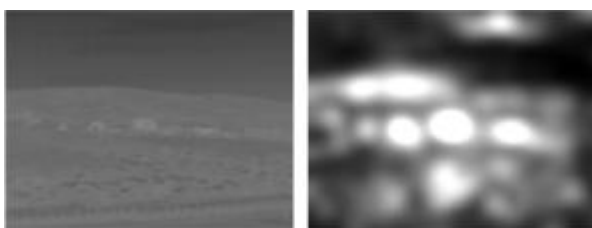
## 2 实验结果与分析

通过实验对显著区域提取和分级图像压缩的性能进行了验证。

### (1) 显著区域提取

采用上述显著区域提取算法对包含多种不同类型目标的红外图像进行处理,均取得了较好效果。文中给出两组比较有代表性的实验。

图 5 为一组建筑物的红外图像,该组建筑物位于山腰,对比度较低,目标较小,但文中算法能够准确提取出建筑物目标所在的显著区域。



(a) 原始图像 (a) Original image (b) 显著图 (b) Saliency map



(c) 显著区域 (c) Saliency regions

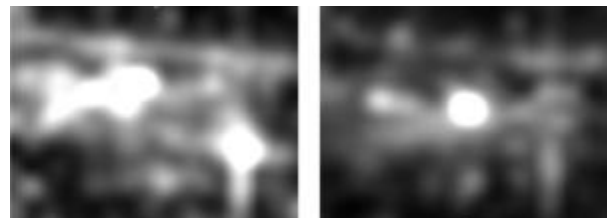
图 5 建筑物目标的显著区域提取

Fig.5 Saliency region extraction of buildings

图 6 为一组行进中的车辆目标的红外图像,随着目标由远及近,成像距离的改变导致目标在图像的尺寸越来越大,可以看出,文中算法在这组序列图像上取得了较好的显著区域提取效果。



(a) 原始图像 (a) Original image



(b) 显著图 (b) Saliency map



(c) 显著区域 (c) Saliency regions

图 6 车辆目标的显著区域提取

Fig.6 Saliency region extraction of vehicle

### (2) 分级图像压缩

文中采用上述分级图像压缩算法对一组机载红外图像进行实验,与传统的全图压缩方法进行比较,



实验图像中包含两个热电厂冷却塔目标,如图 7 所示,图(a)为原始红外图像,在图(b)中,由视觉注意模型生成的显著区域用方框的形式标记出来(图中标记为 1、2、3、4、5 的区域)。在压缩时,显著区域均采用无损压缩方式,背景区域采用有损压缩方式,过渡区域的压缩比介于前两者之间。文中将显著区域的压缩比设为 7:1,背景区域设为 23:1,过渡区域设为 11:1。图 7(c)为直接对全图进行有损压缩的解压后重构图像,图 7(d)为采用文中提出的分级压缩的解压后重构图像。

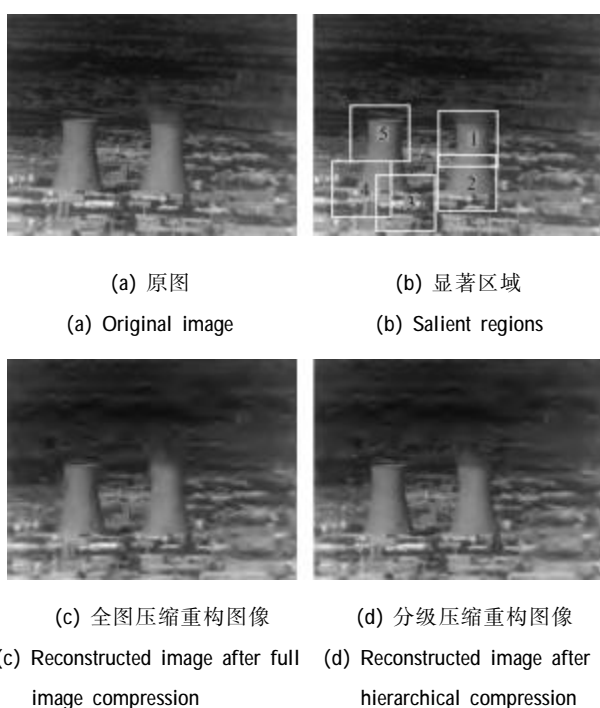


图 7 图像压缩结果

Fig.7 Image compression results

通过主观观察比较可以发现,文中方法得到的重构图像效果明显优于全图压缩方法,目标区域保真度较高,例如重构图像中冷却塔边缘特征保持较好。但是图像质量的主观评价方法操作复杂,且在实时传输的背景下难以实施,在实际应用中,更广泛使用的是客观保真度准则。均方误差(Mean Square Error, MSE)和峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)是两个常用的评价指标,如公式(2)和公式(3)所示:

$$MSE = \frac{1}{M \times N} \sum_{m=1}^M \sum_{n=1}^N (f(m, n) - g(m, n))^2 \quad (2)$$

$$PSNR = 10 \log_{10} \left( \frac{(\max_{1 \leq m \leq M, 1 \leq n \leq N} (f(m, n)) - \min_{1 \leq m \leq M, 1 \leq n \leq N} (f(m, n)))^2}{MSE} \right) \quad (3)$$

式中: $f(m, n)$ 和 $g(m, n)$ 分别表示解压后的重构图像和原始图像,尺寸为 $M \times N$ 。对于解压后重构图像,MSE 越小越好,PSNR 越大越好。

采用上述指标对图 7(c)和图 7(d)进行评价,表 1 给出了相同压缩程度下两种压缩方法的客观评价结果。可以看出在同等压缩率的情况下,全图压缩方法的 MSE 远大于文中方法,且 PSNR 也小于文中方法。这说明文中方法在性能上优于全图压缩方法,解压后重构图像具有较高的保真度。

表 1 压缩图像客观评价

Tab.1 Objective evaluation of compressed image

	Proposed method	Full image compression
MSE/pixel	37.44	66.72
PSNR/dB	32.39	29.88
Compression	22.47%	22.47%

主观观察和客观评价均表明,文中方法解压后重构图像具有较高的保真度和较好的可视化效果。这是因为显著区域是最有可能携带图像重要信息的区域,全图压缩的方法并未区别对待,造成重要目标信息的损失,分级压缩的方法对显著区域采用无损压缩方式,能够尽量保持其质量;而背景区域基本不包含目标,采用高压缩比的有损压缩方式压缩图像可以大幅压缩图像数据量而不影响目标区域的质量;对过渡区域采用介于前两者之间的压缩比能够实现较好的图像平滑,保证解压后图像整体的可视性。综上所述,文中方法实现了具有针对性和区分性的图像压缩,在减少数据量的同时较好地保持了图像质量,保证图像中目标区域的可视性,有利于地面操作手迅速、准确地捕获目标。

### 3 结论

针对人在回路模式中图像压缩时存在的高压缩比与高保真度之间的矛盾,文中提出基于视觉注意模型的分级图像压缩方法。首先采用改进的视觉注意模型提取显著性区域,然后将图像分为不同重要性等级的区域进行不同的压缩。与传统的全图压缩方法相比,该方法在减少图像数据量的同时保留了重要目标信息,解压后重构图像的质量明显提高,能够保证目标区域数据的完整无损性。算法的编码效

率与要求无损的区域大小有关,当图像中背景区域比例较大时,可以达到很高的压缩比。

#### 参考文献:

- [1] Song Fuzhi. Selection between ATR and man-in-the-loop techniques[J]. Tactical Missile Technology, 2006(2): 59-62. (in Chinese)  
宋福志. ATR与人在回路的选择[J]. 战术导弹技术, 2006(2): 59-62.
- [2] Su Juan. Research on key techniques for automatic recognition and tracking of FLIR targets [R]. Xi'an: The Second Artillery Engineering University, 2012. (in Chinese)  
苏娟. 前视红外目标识别与跟踪关键技术研究 [R]. 西安: 第二炮兵工程大学, 2012. (in Chinese)
- [3] Sang Nong, Li Zhenglong, Zhang Tianxu. Applications of human visual attention mechanisms in object detection [J]. Infrared and Laser Engineering, 2004, 33(1): 38-42. (in Chinese)  
桑农, 李正龙, 张天序. 人类视觉注意机制在目标检测中的应用[J]. 红外与激光工程, 2004, 33(1): 38-42.
- [4] Itti L, Koch C, Niebur. 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.
- [5] Navalpakkam V, Itti L. An integrated model of top-down and bottom-up attention for optimizing detection speed[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 2006, 2049-2056.
- [6] Rosin P L. A simple method for detecting salient regions[J]. Pattern Recognition, 2009, 42(11): 2363-2371.
- [7] Kadir T, Brady M. Saliency, scale and image description[J]. International Journal of Computer Vision, 2001, 45(2): 83-105.
- [8] Yuewei Lin, Bin Fang, Yuanyan Tang. A computational model for saliency maps by using local entropy [C]// Proceedings of the Twenty-Fourth AAAI Conference on Artificial Intelligence, 2010: 967-973.
- [9] Wu Shanyu, Yan Zhenming. Modified Huffman code and its applications [J]. Communications Technology, 2009, 42(1): 309-311. (in Chinese)  
武善玉, 宴振鸣. 改进的 Huffman 编码及其应用 [J]. 通信技术, 2009, 42(1): 309-311.
- [10] Shapiro J M. Embedding image coding using zero trees of wavelet coefficients [J]. IEEE Transactions on Signal Processing, 1993, 41(12): 3445-3462.