第 23 卷 第 5 期
2002 年 10 月

焊 接 学 报
TRANSACTIONS OF THE CHINA WELDING INSTITUTION

Vol.23 No.5
October 2002

铝合金 TIG 焊熔池正面图像模式识别

王建军[1]， 杨学勤[2]， 林 涛[1]， 陈善本[1]
（1. 上海交通大学 材料科学与工程学院，上海 200030；
2. 上海航天局 800 研究所，上海 201600）

王建军

摘 要：将图像处理与模式识别技术用于铝合金 TIG 焊接过程信息提取过程，根据铝合金熔池图像随机噪声强的特点，采用加权中值滤波、统计灰度边缘检测、统计期望阈值法和投影方法对铝合金熔池图像进行了预处理。探索了将神经网络用于焊接熔池图像处理的方法，采用 BP 神经网络对二值化熔池图像进行边缘提取，取得了理想的效果。研究了大电流条件下铝合金熔池图像的对称性，通过单面图像，得到了完全的熔池边缘图像。

关键词：铝合金；TIG 焊；图像处理；模式识别
中图分类号：TN931.1 文献标识码：A 文章编号：0253－360X(2002)05－73－04

0 序 言

机器人焊接传感和焊接动态过程智能控制技术一直是人们研究的重点，而先进的传感方式是上述技术及其重要的组成部分。近年来，视觉传感技术在焊接中的应用取得了重要进展，人们利用视觉技术对焊接过程进行监测，实时了解焊接过程信息，进而对焊接过程进行控制[1]。目前，采用以工业 CCD 为视觉传感器件，以焊接弧光为光源的被动视觉传感技术和以激光为辅助光源的主动视觉传感技术在不锈钢材料的智能焊接过程的应用是研究的重点，然而对于铝合金而言，由于材料与焊接工艺的特殊性，采用被动视觉传感方式对焊接过程进行动态控制的研究尚未深入。在前期工作的基础上[2]，首次对铝合金 TIG 焊过程中采用被动视觉传感方法所取得的熔池图像处理和模式识别方法进行了系统的探讨。

1 熔池正面图像特征

采用工业 CCD 摄像机获取铝合金 TIG 焊熔池图像过程中，对取像位置的要求非常苛刻。文中图像是在与焊接方向平行的正侧面取得的。从图像可以看出（图 1），由于取像位置的特殊性，图像并不能反映熔池的全貌，而只是表现了熔池的大部分。从整体上看，图像可以分为喷嘴区，阴极雾化区，弧柱区，母材金属区，焊道边缘部分，熔池前部和熔池后部等区域。喷嘴区反射光线最弱，灰度值很低，呈现黑色。雾化区的灰度介于熔池最亮部分和图像背景之间。弧柱区发光强烈，因此在图像中这一部分的灰度值最高；熔池前部融化金属反光强烈，有类似平面镜的反光效果，表现为亮白色。在熔池后部填丝金属产生堆积，主要是由于添丝金属以及母材融化金属向后部流动所致，这使得电弧光产生漫反射，只有部分弧光进入取像系统，因此亮度较低。图像中熔池边缘清晰，而且可以观察到焊缝边界。

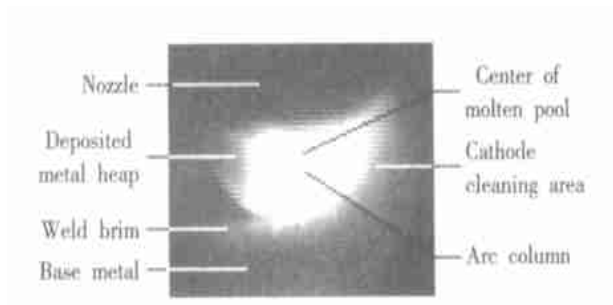

图 1 铝合金熔池图像

Fig. 1 Characteristics of front image of molten pool

2 图像处理

2.1 图像滤波

滤波的目的是去除噪声。从所获取的铝合金熔池图像可以看出，在图像的传输和转换过程中，已经产生了降质，在图像的行方向间隔引入了低频噪声，这主要是在图像的采集和量化过程中受到外界电磁干扰使得参考电源波动的结果。另外，图像中还含有一些高频噪声，是在图像的采集和传输过程中高频辐射造成的。针对上述图像特点，设计了加权中

收稿日期：2002－04－08

值滤波方法。它的原理是将当前像素点及其八邻域的像素灰度值按由小到大顺序排列，假设值的排列顺序为

$$\{a_{-4}, a_{-3}, a_{-2}, a_{-1}, a_0, a_1, a_2, a_3, a_4\}。\quad (1)$$

则当前像素点灰度值为

$$a_c=\frac{1}{4}a_1+\frac{1}{4}a_2+\frac{1}{2}a_3。\quad (2)$$

这种滤波方法有效降低了系统和随机噪声的影响。

2.2 边缘检测和二值化

边缘检测的目的是突出并得到图像的边缘信息，增强图像中物体的轮廓特征，以便于机器识别。从熔池图像特征可知，系统低频噪声在水平方向上弱化了像素间的灰度差别，如果采用水平与垂直方向作为检测方向，将得到虚假信息，所以采用 ±45° 方向作为梯度检测方向。离散化边缘检测输出算子为

$$G(x,y)=|f[i,j]-f[i+1,j+1]|+|f[i+1,j]+f[i,j+1]|。\quad (3)$$

如果只是应用上述公式进行边缘检测并不能得到理想的熔池边缘。因为熔池图像背景的灰度值很低，上述输出是求代数和的结果，这一结果可能与背景的灰度差别很小，这样反而弱化了边缘灰度特征。解决的方法是把阈值设定同统计增强结合起来。在检测过程中设定两个阈值 θ_1 和 θ_2，当 $G(x,y)>\theta_1$ 时把当前点作为边缘点，在输出时要考察周围像素点的灰度情况，如果周围八邻域像素灰度的平均值 $M>\theta_2$，则将 $G(x,y)$ 乘以一个大于 1 的系数 K 进行放大。实际边缘输出为

$$G(x,y)=\begin{cases}G(x,y) & M\leqslant\theta_2;\\ KG(x,y) & M>\theta_2。\end{cases}\quad (4)$$

在边缘检测的过程中也要伴随二值化的过程，目的是将物体从背景中分离出来，在这个过程中不可避免地要进行阈值选择。由于图像的灰度直方图中不存在明显的波峰之间的可用做阈值定位的波谷，传统的阈值选择方法不再适用。采用的方法是根据当前像素的灰度同周围邻域像素灰度的统计关系来进行操作。设当前像素与所在行方向、45° 方向和所在列方向的左右对称的四个像素的灰度平均值的差分别为 d_1、d_2 和 d_3，如果这三个值之中任意一个大于设定值 T，则将当前点看作边缘点。

2.3 投影法边缘细化

边缘细化的目的是将熔池图像的边缘处理成只用单个像素表示的曲线，为数据提取做准备。通常采用的边缘细化算法有 Hilditch 算法和 Rosenfeld 算法等[3]。但是这些算法的共同特点是只提供细化的方法，并不能反映边缘的准确位置。该文图像中熔池边缘并不在曲线区域中心，而是位于弧线的内侧，因此采用通用细化算法将引入误差。鉴于熔池图像的实际特征，建立了投影边缘细化算法。原理是将所有在列方向存在断点的像素群向列的断点上方水平直线 L 投影，在所有存在像素的列方向，只保留孤立点和连通区域中与直线 L 具有最小距离的像素点。原理如图 2，这一算法最大程度地保留了原图像边缘的连通性。需要说明的是在细化图像中采用了距离滤波技术将喷嘴边缘去除了。

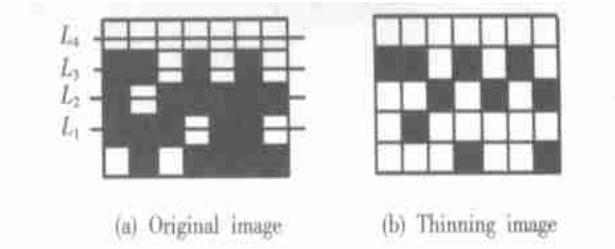

图 2 投影法原理

Fig. 2 Projection theory

2.4 神经网络边缘提取

经过上述处理得到的熔池边缘图像中仍然含有一些点状噪声和线条状假边缘，影响参数测量的准确性。而边缘与噪声是可以通过经验识别出来的，这正是神经网络的长处。由于神经网络在处理某些经验数据以及很难用数学解析规则或公式描述的问题上具有明显的优越性，人们很早就采用神经网络来处理灰度图像，如图像恢复等[4]。但是直接从灰度图像进行处理不仅需要大量的学习样本进行训练，而且网络的结构复杂，同时无法快速得到结果。比较而言，直接从二值图像进行边缘提取所需学习样本少，而且可以极大的提高处理速度。即是从二值图像开始利用 BP 网进行边缘提取的研究[5]。

2.4.1 网络结构

采用的网络结构见图 3。网络是含有输入层、隐层和输出层的 3 层网络结构。输入层含有 9 个节点，对应 3×3 的滤波窗口，输出层也含有 9 个节点，对应运算结果。隐层节点数目的确定没有明确的方法，中间隐层含有 7 个节点，这是试验的结果。输出层各节点的目标值是 0 或 1，隐层的激活函数采用 BP 算法最常用的 Sigmoid 函数

$$\tan\mathrm{sig}(x_i)=\frac{e^{x_i}-e^{-x_i}}{e^{x_i}+e^{-x_i}}。\quad (5)$$

输出层激活函数是 $f(x_i)=x_i$。 (6)

$$x_i=\sum_{j=1}^{n}w_{ij}P_j+b_i,\quad (7)$$

式中：x_i 是第 i 个节点的输入值；w_{ij} 表示第 j 个节点与第 i 节点之间的连接权值；b_i 是第 i 个节点的偏差。

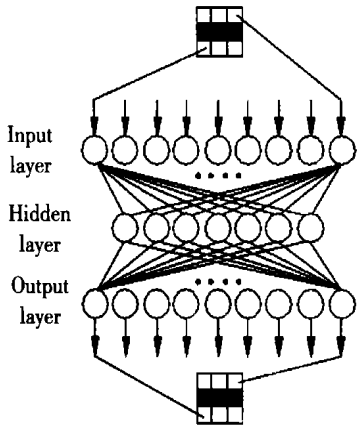

图 3　BP 网络结构

Fig. 3　BP network structure

2.4.2　网络学习算法

传统的 BP 算法的主要缺点是收敛速度慢, 如果参数选择不合理还容易陷入局部极小而不收敛。因此采用改进的 BP 算法。学习因子 η 不再取固定值, 而是利用局部最优动态得到。若某次权值修正后使得误差向减小的方向进行, η 就乘以一个大于 1 的系数, 否则就乘以一个小于 1 的系数。冲量常数 α 的确定要根据总误差来进行, 如果总误差没有减小, 就令 α 为 0, 否则 恢复为非零值。下一次的权值修改量为

$$\Delta W_{ij}(m+1)=\eta_i \sum (\delta_{qj} O_{qj})+\alpha \Delta W_{ji}(m), \quad (8)$$

式中: m 代表迭代次数; δ_{qj} 和 O_{qj} 分别是隐节点的误差信号和输出信号; q 是总样本对数。

2.4.3　网络学习方式

网络的输入采用当前点的 3×3 邻域所形成的窗口, 共有 9 个输入点, 可以检测当前点的上、下、左、右四个方向的拐角点和弧线。无论从理论还是实践上讲, 输入节点愈多, 则处理的结果将越精确。但网络结构将趋于复杂, 给训练和学习带来极大的困难。另外, 随着输入节点的增多, 为了保证完备性, 所需要的样本数目就会急剧增多, 给实际应用带来困难。图像处理过程中, 将 3×3 窗口在图像上移动, 当遇到当前点是黑色像素时, 对此像素及其八邻域的像素值作为网络输入样本, 网络输出是经过去除噪声的边缘点八邻域像素值。网络采用有教师学习方式, 输入样本与输出样本组成数据对。

图像中的噪声形式或形状可能是多种多样的, 比如条状、圆状、点状、拐点等等, 较多的学习样本对提高网络的容错性较为有利, 然而, 如果试图通过穷尽噪声的分布形式和形状来增加样本的数量, 是不可行的。经过试验, 选择了 7 组共 42 个学习样本, 既保证了系统误差收敛, 又保证了网络具有优良的容错性能。图 4 列举了全部的学习样本和教师样本。

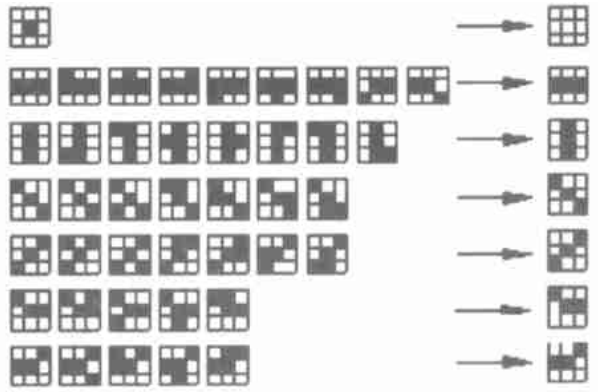

图 4　全部学习样本和教师信号

Fig. 4　Sets of learning patterns

将输入的 3×3 矩阵按行排列成 1×9 的单列, 送入训练完的网络, 将输出 1×9 的单列逆还原为 3×3 的阵列, 复制到原图像。第一组样本是为了去除个别的孤立噪声点, 其余几组样本结合各种噪声点的形式和连接特点, 从当前点的上、下、左、右四个方向和圆弧过渡处的连接特征来决定沿本形式。教师信号的判定是基于在图像中和当前点哪种方向和拐点出现的概率最大。在图5中列出了图像处理的

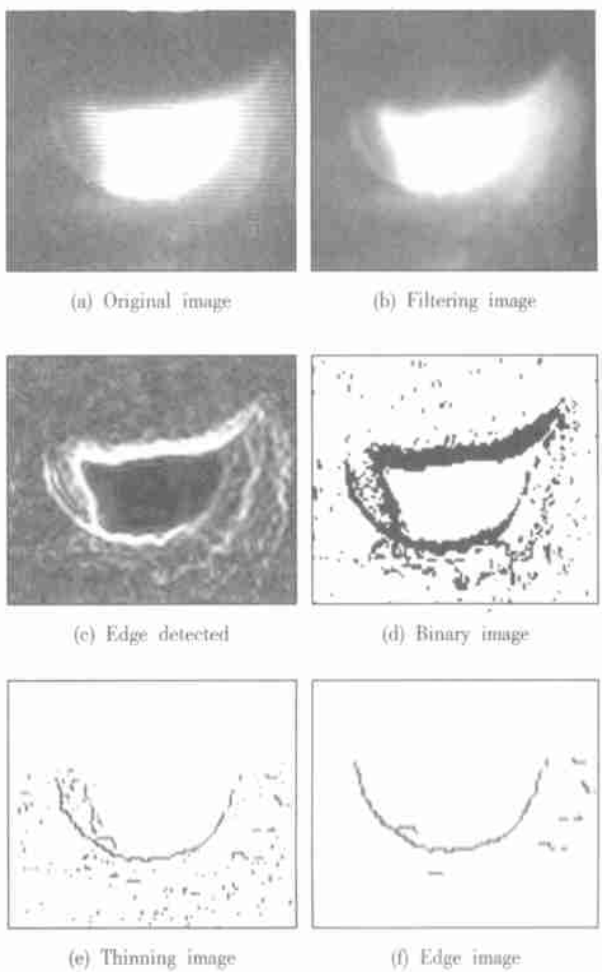

图 5　图像处理过程

Fig. 5　Flow of image processing

全过程。由结果可以看出神经网络能很好地保持图像的边缘特征，对离散点状和短条状加边缘进行有效识别。

3 熔池边缘整体恢复

上述图像处理过程得到的并不是整个熔池边缘图像，前文已述这是由于取像位置的限制无法观测到熔池的全貌。为了研究熔池的整体图像特性，采取了双摄象机在焊缝熔池两侧同时取像的方法，同时获取同一时刻熔池的两侧图像进行研究。采用前文所述的图像处理方法，将得到的边缘图像进行以焊缝为中心的复合（图 6）从结果可以看出，熔池图像具有良好的对称性。事实上，铝合金在大电流条件下，电弧挺度好，这正是熔池图像对称性强的原因。根据这一特性，只需将单侧图像进行以焊缝为中心的对称操作即可得到整个熔池边缘图像。

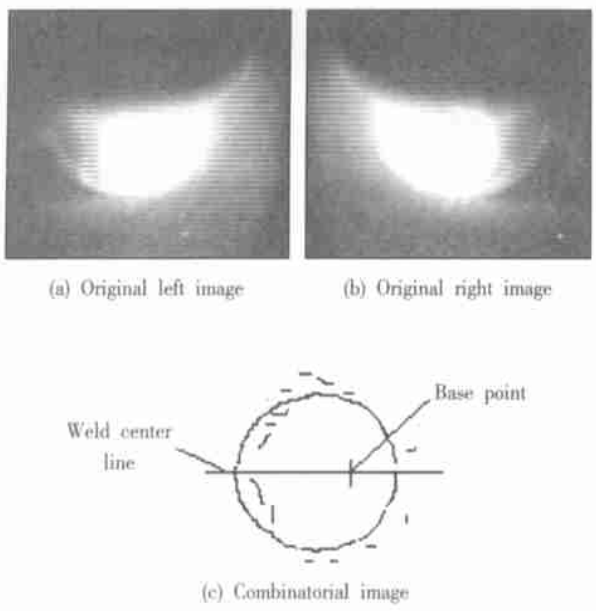

图 6 熔池边缘完整形状恢复

Fig. 6 Edge recovery of welding pool

4 结 论

（1）采用加权中值滤波、基于方向和统计阈值的检测及投影法边缘细化可以对铝合金 TIG 焊熔池图像进行良好的预处理。

（2）采用 BP 神经网络可以实现熔池边缘的准确识别，对断续边缘具有智能修补作用。

（3）铝合金熔池图像在大电流条件下具有良好的对称性，根据对称性可以由单面图像恢复整个熔池边缘图像。

说明：采用上述神经网络进行边缘识别时，对于较长的假边缘，效果不明显，这主要是由于网络的 3×3 输入窗口太小造成的，如果增多输入节点，效果会更好，这一问题，需要进一步研究。

参考文献：

[1] Chen S B, Zhao D B, Wu L, *et al*. Intelligent methodology for sensing, modeling and control of pulsed GEAW part2—butt joint welding[J]. Welding Journal, 2000, 79(6): 164～173.

[2] 王建军，林 涛，陈善本. 铝合金交流钨极氩弧焊熔池图像传感方法[J]. 上海交通大学学报，2002, 36(1): 5～8.

[3] 阮秋琦. 数字图像处理学[M]. 北京：电子工业出版社，2001.

[4] 沈 清，胡德文，时 春. 神经网络应用技术[M]. 长沙：国防科技大学出版社，1998. 130～139.

[5] Etemad K, Chellappa R. A neural network basededge detector [C]. IEEE International Conference on Neural Networks, 1993. 32～137.

作者简介：王建军，男，1971 年出生，上海交通大学博士研究生。主要从事焊接过程智能控制的研究，发表论文多篇。

Email：wangjj@mail.sjtu.edu.cn