

聚类分析提取 PCA 变换后的 火焰图像颜色特征

燕梅,林剑,郭小玉,王静

(湖南科技大学 知识处理与网络化制造湖南省普通高校重点实验室,湖南 湘潭 411201)

摘要:炉口火焰颜色特征是判断炼钢终点碳及终点温度的重要特征.为实现实时控制,通常利用 PCA 变换缩短图像特征提取时间,由于 PCA 变换存在误差,难以在变换后的图像中提取独立颜色,文中首先在原始火焰图像中选取各类颜色样本,计算颜色变换前后的类内距和类间距,发现图像变换前后颜色的聚类性能具有一致性,然后利用聚类方法确定变换后的颜色窗口大小,通过颜色窗口提取序列图像颜色像素数量随时间变换特征,采用 A,B 两炉冶炼同一型号钢种图像数据进行实验,实验结果表明:通过聚类分析的方法所提取的颜色特征能较好地反映转炉炼钢物理化学过程.

关键词:火焰图像;颜色特征;聚类分析;炼钢终点控制

中图分类号:TP751.1 **文献标志码:**A **文章编号:**1672-9102(2014)03-0064-05

Extracting the flame image color feature after PCA transformation based on clustering analysis

YAN Mei, LIN Jian, GUO Xiaoyu, WANG Jing

(Key Laboratory Knowledge Processing and Networked Manufacturing, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract: Combustion flame plays an important role in predicting endpoint contents of carbon and steel temperature's in a basic oxygen furnace (BOF). Since the results of principal component analysis (PCA) exist errors, it was difficult to extract independent colors from the converted images. Different kinds of color swatch were selected from original flame images. By calculating the colors' within-class distance and class distance between before and after transformation, the colors' clustering performance was found that it keep consistent. Clustering method was used to determine color's window size after PCA transformation. Through colors' window, the characteristic of pixel number of sequence images change with time were extracted. Image data from two furnace A and B, smelted same type of steel, was used to make experiment. Results show that the color features, extracted by clustering analysis, is able to much better reflect the physical and chemical processes in steel-making.

Key words: flame image; color feature; clustering analysis; steel-making endpoint control

转炉炼钢终点控制主要是控制转炉吹炼结束时钢水的碳含量和温度,其控制水平直接影响生产效率和产品质量.实现转炉炼钢自动控制的动态数据源主要有利用副枪直接获取的钢水温度和钢水碳含量数据^[1],炉气分析仪获取的炉口溢出的炉气成分数据^[2],利用光学仪器获取的炉口火焰光

谱数据或炉口火焰图像^[3-4].由于利用副枪成本高,目前一般只是在大型转炉中采用,在中小型转炉炼钢最为常见的还是依赖有经验的工人通过观察炉口火焰标志颜色及变化快慢判断炼钢的终点,自动提取炉口火焰图像颜色及其变化特征是实现炉口火焰的自动判断的基础^[5].

从炉口火焰图像中提取控制信息主要有图像纹理特征^[6-7]、区域边缘特征^[8]、图像中不同区域的光谱强度等特征^[9-10],转炉炼钢的过程是一个复杂的多相物理化学过程,不同钢水成分、不同温度所产生的光谱波长及变化频率不同,为实现有效的控制,所提取的特征越全面、越具有代表性,越有利于提高控制的鲁棒性^[11-12],对于火焰图像尽可能地利用不同波段的图像提取多种能够反映钢水的特征,但所提取特征越多所花费时间越长.为使图像特征提取所花费的时间满足实时控制的要求,通常需把三维彩色图像转换为二维数字图像,缩短图像特征提取时间,目前应用最为广泛的是多变量图像分析(Multivariate Image Analysis, MIA)方法, MIA 方法的核心是利用主元变换(Principal - Component Analysis, PCA)把三维的图像数据转换为二维^[13-14],在二维平面设计多个固定大小的窗口,利用直方图提取窗口内总亮度、平均亮度、总的颜色、平均颜色等特征^[15-17].在固定窗口内的二维数据由多种三维颜色数据经 PCA 转换而来,由于 PCA 转换存在误差,转换后的二维数据与三维 R \ G \ B 数据不是一一对应关系,采用目前的方法难以获取炼钢过程中直接反映炉内温度与碳含量标志颜色信息,如淡黄、白色、橘红等.

本文采用区域聚类分析方法确定窗口的大小,直接提取火焰标志颜色特征,采用冶炼同一类型炉种的不同炉次进行对比分析说明方法的有效性.

1 原理与方法

1.1 序列图像 MIA 方法

如果不考虑图像像元的空间坐标关系,图像矩阵 I 进行 PCA 变换有:

$$I = \sum_{k=1}^k t_k p_k^T + E. \quad (1)$$

式中, k 为主成分个数, t_k 为得分矢量,对于三维图像 PCA 变换 k 取值通常为 2,在此为变换后图像二维矩阵, p_k 为载荷矢量, E 为残差.对于序列图像, p_k 由平均协方差矩阵 $\sum_j I_j^T I_j$ 求出, j 为序列图像数,那么 $t_k = I p_k$,对 t_k 进行归一化,使其数值处于 0 ~ 255 之间

$$s_{k,i} = \text{Round} \left(\frac{t_{k,i} - t_{k,\min}}{t_{k,\max} - t_{k,\min}} \times 255 \right) \quad i = 1, \dots, N, k = 1, 2. \quad (2)$$

$s_{k,i}$ 为归一化后二维图像数据, N 为图像像元数, $t_{k,\min}$ 和 $t_{k,\max}$ 分别为所有序列图像 t_k 中最小值和最大值, MIA 方法在二维 $s_{k,i}$ 中提取各种火焰图像特征.对于图像,经转换后存在多个像元具有相同的 s_1, s_2 值,为描述二维空间,定义一个 $256 \times$

256 的直方图 T 统计 s_1, s_2 值相同的像元个数, T 为 256×256 的矩阵并可由下式计算

$$T_{m,n} = \sum_i 1 \quad \forall i, s_{1i} = m, s_{2i} = n, m, n = 0, \dots, 255. \quad (3)$$

定义二元矩阵 M ,当二维平面中的点属于某一窗口, M 值为 1,反之为 0,若窗口的点均属于某种颜色的二维数据(该窗口称之为颜色窗口),那么该种颜色在火焰图像中的像元数量 CN 为

$$CN = \sum_{m,n} T_{m,n} \quad \forall (m,n), M_{m,n} = 1. \quad (4)$$

MIA 方法选择多个窗口,通过 T 和 M 提取各类图像特征,其窗口的选择主要通过实验或迭代的方法选择,使窗口集中于图像像元数密集区域.

1.2 聚类分析确定颜色窗口方法

1) 方法的基本思想

要确定二维颜色窗口边界,首先须确定三维 R \ G \ B 颜色边界,建立三维区域与二维的对应关系.通常彩色可用 R \ G \ B 矩形、彩色三角形或彩色立体图等方式表示,由于视觉感觉颜色与刺激强度、刺激面积、颜色对比、亮度对比等因素有关,数值上难以建立严格的 R \ G \ B 颜色边界,事实上,自然界也难以存在单一波长的纯颜色,感觉的每种纯颜色区域必然包含其他颜色成分^[18],另一方面,三维图像在 PCA 变换过程中存在误差,因此,难以通过严格的数值计算建立三维颜色空间与二维颜色区域的对应关系.但是,在彩色图中视觉感觉的某种颜色区域,必然存在 R \ G \ B 数值接近的像元占整个区域的绝大多数,由于三维 R \ G \ B 数据的 PCA 变换,其实质为二维正交轴的投影,多数像元经变换后在正交轴上的点数仍为多数,如果转换前后数据的聚类性能相似,那么在 R \ G \ B 图像可区分的颜色在变换后也是可分的.因此,可以通过聚类的方法确定二维颜色窗口的边界.

2) 聚类分析方法

根据上述基本思想,聚类分析方法确定颜色窗口的方法有如下步骤^[9]:

①计算不同类型颜色的类内距和类间距,分析 PCA 变换前后颜色聚类性能.在炉口火焰图像中 C 种不同类型颜色区域分别选取一定像元,设这些像元 R \ G \ B 值组成的三维模式集为 $\{x_1, x_2, \dots, x_N\}$,其中 ω 类的模式集为 $\{x_l^{(\omega)}; \omega = 1, \dots, C; l = 1, \dots, n_\omega\}$, ω 表示类别, l 表示颜色类别模式序号;

$\sum_{\omega=1}^C n_\omega = N$,类内距离 L_ω 为

$$L_\omega = \sum_{l=1}^{n_\omega} x_l^{(\omega)} - m_\omega^2. \quad (5)$$

式中, m_ω 表示 ω 类的均值矢量,

类间距离 $L_{b\omega}$ 为

$$L_{b\omega} = (\mathbf{m}_{\omega} - \mathbf{m}_{\omega'})^T (\mathbf{m}_{\omega} - \mathbf{m}_{\omega'}). \quad (6)$$

式中, \mathbf{m}_{ω} 表示 ω 类平均矢量, $\omega' = 1, \dots, C, \omega' \neq \omega$.

三维模式集 $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ 经式(1), 式(2)变换后可得二维矢量集 $\{s_1, s_2, \dots, s_N\}$, 同样采用式(5)和式(6)分别计算二维矢量集的内距和类间距, 分析三维矢量集和二维矢量集类内距和类间距, 可以判断变换前后颜色类型的聚类特征是否具有 consistency.

②对变换后同一颜色类型的二维矢量进行聚类, 求取元素最多的二级类别, 确定二维矢量与三维颜色的对应关系. 由于目视方法选取某种颜色区域的像元不为纯色, 但其选定的颜色必占多数(如黄色区域会存在红色、棕色等颜色像元, 但黄色像元必占多数), 对单一颜色区域选定的像元进行聚类, 当聚类类数越多, 所形成的二级类别类内距越小, 颜色越纯, 若 PCA 变换前后的类内距、类间距具有一致性, 变换前后元素最多的二级类别具有对应关系. 那么, 对 ω 类中 n_{ω} 个二维矢量集进行聚类, 形成二级类别模式集为 $\{s_g^{(\lambda_{\omega})}; \lambda_{\omega} = 1, \dots, c; g = 1, \dots, n_{\lambda_{\omega}}\}$, c 为二级类别聚类数, λ_{ω} 表示 ω 类的二级类别, $n_{\lambda_{\omega}}$ 表

示 λ_{ω} 类的元素个数, $\sum_{\lambda_{\omega}=1}^c n_{\lambda_{\omega}} = n_{\omega}$.

③求取元素最多二级类别中元素的最大最小值, 确定颜色窗口边界. 假设 λ 为 c 类中元素最多的类, λ 即为纯颜色类别, 其类中元素个数为 n_{λ} , 计算 λ 类中矢量元素分量的最大最小值作为颜色窗口对角点, 设 $s_{k, \max}, s_{k, \min}$ 分别表示 λ 类中第 k 个主成分分量的最大值和最小值, 有 $s_{k, \max} = \max(s_{k, g}^{\lambda}), s_{k, \min} = \min(s_{k, g}^{\lambda})$, 其中 $g = 1, \dots, n_{\lambda}, k = 1, 2$, 那么颜色窗口的对角顶点为 $(s_{1, \min}, s_{2, \min})$ 和 $(s_{1, \max}, s_{2, \max})$.

2 实验结果分析

2.1 实验数据及图像

实验摄像速率为 2 张/s, 共采集 65 炉火焰图像, 涉及 15 种冶炼钢类型, 文中任选同一转炉冶炼同种类型钢种的 A, B 两炉图像数据(如表 1), 图 1 为不同吹炼时刻具有代表性的图像.

表 1 实验数据概况

	图像开始时刻	图像终止时刻	终点温度/°C	终点碳含量/%
A 炉	13:38'28"	13:42'37"	1 680	0.0 9
B 炉	14:15'48"	14:20'13"	1 664	0.0 7

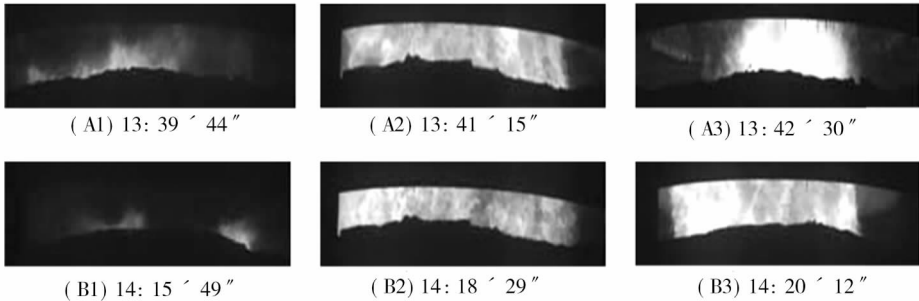


图 1 不同炉次不同吹炼时刻炉口火焰原始图像

从图 1 中可见不同吹氧时刻炉口火焰图像有较大的差异, 在吹炼的早期以棕褐色为主、烟雾较大, 中期以橘黄、棕红为主, 后期图像亮度较大以淡黄、白色为主.

2.2 变换前后聚类分析

在炉口火焰图像的白色、黄色、橘红和棕色区域中分别提取一定数量的像元, 形成上述颜色的 R \ G \ B 三维模式集, 经 PCA 变换后可得到相应的变换后的二维模式集, 分别求取变换前后模式集的类内距和类间距, 由于变换前后模式维数不同, 在此采用归一化的类内距和类间距, 归一化的类内距为

$$\frac{L_{\omega}}{c}, c \text{ 为类别数, 归一化类间距与此同, 归一化}$$

的其类内距如表 2, 类间距如表 3.

表 2 变换前后标志颜色类别的类内距

标志颜色	变换前	变换后	标志颜色	变换前	变换后
白色	0.270 5	0.270 4	橘红	0.154 5	0.153 5
黄色	0.256 7	0.258 3	棕色	0.318 3	0.317 8

表 3 变换前后标志颜色的类间距

后 \ 前	白色	黄色	橘红	棕色
白色	-	0.070 8	0.184 1	0.257 6
黄色	0.110 7	-	0.129 4	0.203 4
橘红	0.145 3	0.101 0	-	0.154 7
棕色	0.238 4	0.217 4	0.186 8	-

从表 2, 表 3 可见: 1) 变换前后的归一化类内距具有很好的一致性; 2) 变换前后有变大的, 也有变小的, 总体变化幅度不大; 3) 变换前后类内距和类间距具有很好的一致性; 即, PCA 变换不会影响

颜色的聚类性能.

2.3 火焰序列图像标志颜色特征分析

采用 K - 均值法分别对 PCA 变换后的不同颜色样本的二维矢量进行聚类,初始类别数为 10 类(用不同的初始类别实验,元素最多的二级类别趋于稳定为准),求取元素最多的二级类别中元素数值的最大最小值,作为颜色窗口的对角顶点如表 4.

表 4 标志颜色窗口对角顶点

	$s_{1,min}$	$s_{2,min}$	$s_{1,max}$	$s_{2,max}$
白色	30	105	40	131
黄色	58	161	87	177
橘红	60	151	131	197
棕色	145	37	195	115

根据式(4)计算不同冶炼时刻炉口火焰图像标志颜色数量,选择冶炼结束前 300 s 的火焰图像进行分析,以纵坐标为标志颜色像元数量,横坐标为冶炼时刻,形成不同冶炼时刻图像标志颜色数量变化曲线(如图 2).

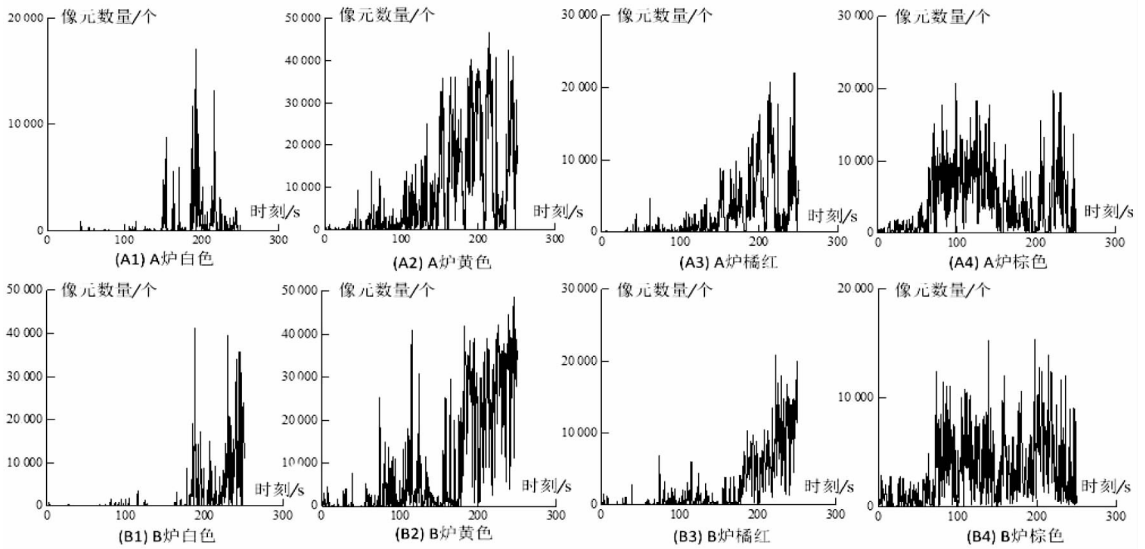


图 2 A,B 两炉不同颜色像元数与冶炼时刻曲线

图 3 为 MIA 方法所提取的颜色总数与冶炼时刻曲线,MIA 方法不区分单一颜色,它在炉口火焰图像的明亮区域采用多个固定窗口,窗口根据火焰图像变换后二维平面点分布确定,在此选取 10 个窗口.

对比图 2 和图 3 可见:

1) 图 3 中的曲线形态是各标志颜色的综合反映,尤其与主要明亮颜色黄色和橘红的曲线形态相似,如图 3(A) 与图 2(A2) 和图 2(A3) 相似,如图 3(B) 与图 2(B2) 和图 2(B3) 相似;

2) 图 3 在频率特征方面同样能反映冶炼后期

从图 2 可见:

1) 不同的标志颜色在冶炼过程中具有明显的变化特征;同一种颜色具有类似的特征;

2) 冶炼早期白色较少,橘红和棕色成分较多;随着冶炼的时间增加各种颜色成分均逐渐增加,且曲线跳动剧烈;到冶炼后期白色、黄色和橘红等浅色成分由峰值逐步变少,棕色相对较多,且各类曲线跳动变缓;

3) 在冶炼的后期,B 炉各类标志颜色曲线比 A 炉跳动更小,这与 B 炉的终点含量及炉口温度较 A 炉偏低,具有较好的对应性.

在吹炼前期熔池温度较低,碳氧化少,炉口火焰颜色暗,冶炼中期暗红色逐渐变红,浓度变淡黄色成分增多,当剧烈氧化时,爆裂程度越大,白光增加,火焰弹跳有力,随着碳含量的减少白色变淡,火焰收缩显柔软.上述图像特征与冶炼过程具有较好的一致性.

变化特征,即 B 炉比 A 炉跳动更小;

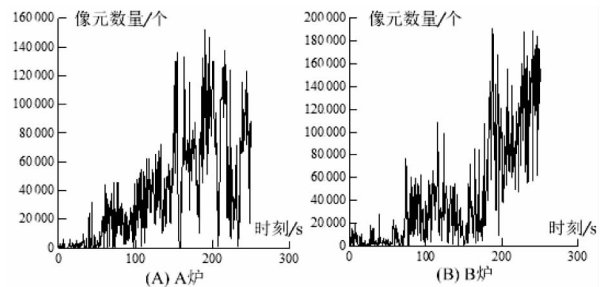


图 3 A,B 两炉 MIA 方法所提取的颜色像元总数与冶炼时刻曲线

3)图3中白色和棕色的反映不明显,白色是炉口火焰终点判断的主要标志颜色,在图2(B4)中棕色跳动趋于平稳也是判断终点的重要依据。

3 结论

转炉吹氧炼钢火焰颜色与其变化快慢是确定冶炼终点判断的重要标志,通过聚类分析确定火焰标志颜色窗口提取火焰颜色变化有如下结论:

1)三维的彩色图像数据经PCA后,其变换前后的颜色聚类性能具有一致性,在变换后的二维数据中分别提取标志颜色特征是可行的;

2)利用聚类分析的方法确定标志颜色窗口,通过颜色窗口提取标志颜色特征能反映冶炼过程火焰变化;

3)不同标志颜色特征反映冶炼过程中不同的物理化学过程,体现钢水中不同成分的变化情况,分别提取标志颜色特征比MIA方法提取综合特征具有更为丰富的信息量,有利于利用火焰图像控制炼钢的终点。

参考文献:

[1] 左康林,邹俊苏,孙晓辉,等.转炉副枪测量与成分预报技术[J].炼钢,2009,25(2):59-61.

[2] 邢曼华,袁守谦,赵田丽,等.转炉终点控制技术的发展[J].炼钢工程学,2010,4(2):54-59.

[3] 徐正光,赵英杰.转炉炉口序列火焰图像的特征提取[J].微计算机技术,2007,27(12-3):89-93.

[4] 温宏愿,陈延如,赵琦,等.利用Direct Show提取转炉炉口火焰图像特征信息[J].华中科技大学学报(自然科学版),2007,35(1):246-248.

[5] 刘辉,张云生,张印辉,等.基于火焰图像特征与GRNN的转炉吹炼状态识别[J].计算机工程与应用,2011,47(26):7-10.

[6] 尤佳,王绍纯,李希胜,等.通过转炉炉口火焰纹理分析判断氧气顶吹转炉吹炼终点[J].北京科技大学学报,2000,22(6):21-26.

[7] 刘辉,张云生,张印辉,等.基于灰度差分统计的火焰图

像纹理特征提取[J].控制工程,2013,2(6):67-73.

[8] 张进华,庄健,杜海峰,等.一种基于视频多特征融合的火焰识别算法[J].西安交通大学学报,2006,40(7):811-814.

[9] 何涛焘,田陆,文华北,等.基于炉口火焰信息的转炉炼钢终点预报系统[J].中国冶金,2013,23(2):40-44.

[10] Huang H W, Wang Q, Tang H J, et al. Characterisation of external acoustic excitation on diffusion flames using digital colour image processing[J]. Fuel, 2012, 94(5): 102-109.

[11] MacGregor J F. Image - based endpoint Carbon prediction for a basic Oxygen furnace [J]. Control Theory & Applications, 2011, 28(3): 21-25.

[12] Geladi P, Grahn H. Multivariate image analysis [D]. Chichester, UK: Wiley, 1996.

[13] Nomikos P, MacGregor J F. Monitoring batch processes using multiway principal component analysis[J]. AIChE Journal, 1994, 40(8): 1361-1375.

[14] Yu H L, John F. MacGregor. Monitoring flames in an industrial boiler using multivariate image analysis [J]. AIChE Journal, 2004, 50(7): 1474-1483.

[15] Theodora K, Paul N, John F M. Analysis, monitoring and fault diagnosis of batch Processes using multiblock and multiway PLS[J]. Journal Process Control. 1995, 5(4): 277-284.

[16] Graham K J, Rishnapisharody K K, Irons G A. Ladle eye area measurement using multivariate image analysis [J]. Canadian Metallurgical Quarterly, 2007, 46(4): 397-405.

[17] David C, Rawlings B C, Edgar T F. Multivariate image analysis (MIA) for industrial flare combustion control [J]. Industrial & Engineering Chemistry Research, 2012, 51(39): 12642-12652.

[18] 王向阳,李东明,杨红颖.基于Zernike色度分布矩的彩色图像检索算法[J].模式识别与人工智能,2012,25(2):313-317.

[19] 周江,王国华,赵跃龙.一种基于聚类预测模型的动态多目标进化算法[J].湖南师范大学自然科学学报,2014,37(2):56-61.