

文章编号: 1672-2892(2011)02-0185-05

基于纹理特征的太阳能晶片计数研究

方超, 谭伟, 杜建洪

(复旦大学信息科学与工程学院, 上海 200433)

摘要: 针对工业生产中太阳能晶片人工计数的难以实施, 本文提出了运用数字图像处理技术的自动计数方法。从具体应用角度出发, 通过对晶片叠层图像的纹理分析, 先用一阶统计量法和边缘检测投影法对晶片区域进行定位, 然后通过极值点分析法进行计数。实验结果表明该算法的定位准确度可达到 94% 以上, 计数正确率可以达到 95%, 而且算法简单可行, 具有实用价值。

关键词: 纹理特征; 垂直投影; 边缘检测; 极值点分析

中图分类号: TN911.73; TP391.41

文献标识码: A

Solar energy wafer counting based on textural property

FANG Chao, TAN Wei, DU Jian-hong

(School of Information Science and Engineering, Fudan University, Shanghai 200433, China)

Abstract: This article presents an automatic counting method using digital image processing to solve difficulties in manual Solar Energy Wafer(SEW) counting in industry production. By analyzing the image texture of overlaid wafers, the author firstly locates wafer area employing first-order statistics method and edge detection projection, and then performs counting by extremum analysis in this practical application. The experiment results show that the algorithm is highly accurate with the locating accuracy up to 94% and the counting accuracy up to 95%, and it is convenient and applicable.

Key words: textural property; vertical projection; edge detection; extremum analysis

在太阳能晶片的大规模生产中, 晶片计数是生产包装销售等环节中不可或缺的部分。但由于太阳能晶片薄而易碎的特性, 人工数片方法已不能满足生产需求。本文提出了一种基于图像处理技术的太阳能晶片自动计数系统, 能快速有效地确定一叠晶片的块数。国内外针对这一特定应用领域的研究不多, 但对基于纹理特征的纸张计数有一定研究基础^[1-2]。文献[1]从纸张叠层图像的纹理特征出发, 通过 LOG 滤波等操作得到二值图像, 然后用像素投影算法和差分统计方法获得纸张数目。文献[2]从二维 Gabor 滤波和一维频率分析处理的角度入手, 通过提取纸张边缘来计数。但由于纸张叠层和太阳能晶片叠层纹理分布特征及应用场景的不同, 这两种算法都不能直接用于太阳能晶片计数。文献[3]从条纹分析的角度出发, 提出了基于条纹计数和定位的二值化、滤波和求条纹升降沿三者结合的综合方法, 能有效提高条纹图像的信噪比。文献[4]利用区域极值对极值区域和非极值区域进行标记得二值条纹, 并对二值条纹进行计数, 具有更强的抗毛边、抗油墨以及抗光照不均性能。文献[5]提出了一种基于数学形态学的纸张计数方法, 通过二值形态学开闭及膨胀运算填充纸张纹理中的孔洞, 能够满足实际生产的需要。目前纸张计数技术已用于实际生产^[6-7]。

本文从太阳能晶片计数这一具体应用出发, 提出了基于纹理特征的太阳能晶片计数算法。首先通过定位算法从采集样本图像提取太阳能晶片叠层区域的具体位置, 然后通过分析纵向灰度分布曲线来计数。

1 太阳能晶片叠层图像特点分析

1.1 样本采集

用如图 1 方式采集太阳能晶片叠层图像: 太阳能晶片为正六边形, 厚度为 $200\ \mu\text{m} \pm 20\ \mu\text{m}$ 。将太阳能晶片叠层排列整齐, 水平放置, 摄像头保持水平, 镜头与晶片叠层某一侧片垂直。拍摄时调整焦距, 保证纹理清晰可见。

由于晶片块数不同,导致晶片叠层高度不一致,因此采集图像不可能刚好只包含晶片叠层区域纹理,可能还包括部分环境及晶片上表面纹理(如图2所示)。这些因素都会对计数产生干扰,因此需要加入定位算法。

1.2 太阳能叠层图像与纹理模式的联系

观察图2的晶片叠层部分纹理特征可以发现:图像中水平或近水平的边棱两侧的灰度对比明显。这种高低对比分布于整个区域当中,具有明显的规律性,呈条纹状分布。该视觉效果是光照在晶片边棱附近留下的阴影造成的。这样,从纹理分析的角度考虑,晶片叠层图像可以近似看作1幅由不同深浅的灰斑条彼此重复排列组成的纹理图像。该纹理结构与纸张纹理结构类似,但也不尽相同。相同点有:

1) 条纹彼此基本平行,并且有很强的方向性;2) 黑白条纹周期性排列,具有重复性。不同点表现在:1) 纸张较薄,纸张边棱反映在图像中的部分不会带来较大的灰度变化,而且两侧空隙在适当光照条件下反映在图像上灰度分布呈渐变趋势;相反,太阳能晶片本身厚度较大,由于晶片材料的不均匀性,晶片边棱在光照条件下有不同的发散,表现在图像中呈现不同的灰度变化。2) 纸张柔软,通过适当的操作可以保证纸张间隔的均匀性,而太阳能晶片由于是刚性的,而且易碎,在实际操作中不能挤压,导致间隔与间隔之间,以及同一间隔的不同位置,间隔大小不一致。

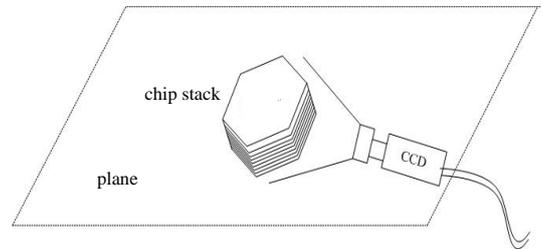


Fig.1 Solar wafers laminated image acquisition method
图1 太阳能晶片叠层图像摄取方法

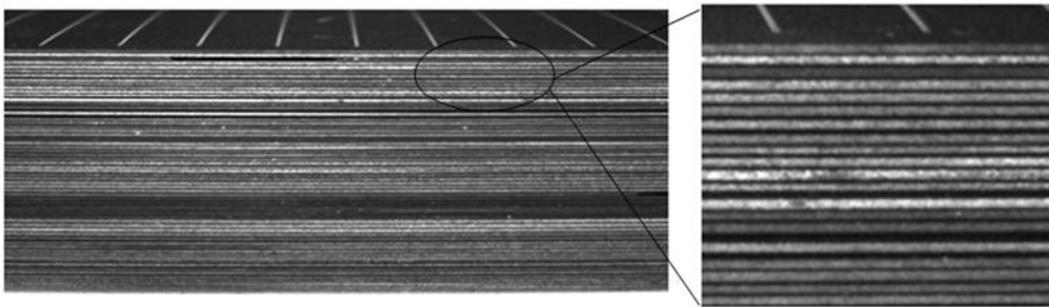


Fig.2 Chip stack image and its partial enlargement
图2 晶片叠层图像及其局部放大图像

2 基于纹理分析的计数算法

图3描述了算法总体设计流程。

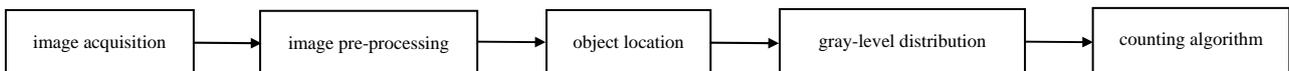


Fig.3 Flow of overall algorithm
图3 总体算法流程

2.1 图像预处理

在文献[1]和[2]中,图像预处理非常重要^[8],与计数正确率密切相关。文献[1]和[2]分别使用 LOG 滤波器和二维 Gabor 滤波器对纸张纹理进行滤波。本算法预处理采用简单易行的中值滤波即可。本文算法对实验环境如光照、拍摄角度等有一定要求,尤其是光照条件,在亮度较低的环境中,晶片纹理部分会由于对比度太低而无法正确定位,进而影响计数的准确性。本实验环境搭建了专门的实验平台,设计一暗箱,将摄像头置于暗箱内,暗箱有一矩形开口,可以通过该开口拍摄晶片叠层侧面图像。暗箱内摄像头前方上下左右4个方位可打入亮度可调的平行光,以保持适当光照。将待计数晶片整齐堆放于暗箱外矩形开口处,与暗箱内的摄像头垂直。在工业生产中,工作人员通过取放晶片堆来实现流水线操作。

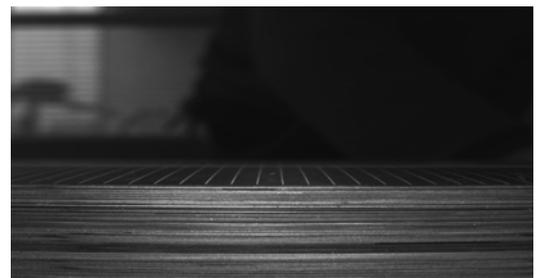


Fig.4 Sample image
图4 拍摄样本

2.2 定位算法

由于计数是通过统计叠层区域纵向灰度分布特性曲线的峰值点来计算,起始点和终止点的位置对实验结果影响很大,因此确定晶片叠层区域的上下边缘至关重要。定位算法就是从采集图像提取晶片叠层区域,主要是确定区域的上下边缘。如图 4 所示,拍摄样本图像主要包括晶片叠层区域、晶片上表面纹理,以及周围景物,如人影。这些图案的分布具有较强的方位性。经垂直投影图分析容易确定下边缘,因为下边缘处灰度变化明显,且下边缘以下灰度值基本接近 255(置于白色底物上)。主要考虑上边缘的确定。

本文采用一阶统计量方法和边缘检测投影法确定上边缘。如图 2 所示,晶片叠层与晶片上表面之间存在明显的分界。一阶统计量方法利用图像垂直投影图,如图 5(a),确定上边缘。从曲线原点出发,依次选取 N 个点,求这 N 点的方差。根据纹理区域的方差较大而纹理与上表面分界处方差很小的先验知识,从右(下边缘处)至左扫描所有的方差值,设置阈值 T ,当方差小于 T 时说明达到分界处,如图 5(b)所示。由于无法确定该点还是相邻的两个点更接近真实分界点,可能存在 $\pm N$ 个像素的误差。在 N 的选取上,如果 N 较大,将引起较大的定位误差;如果 N 较小,则定位时需要更多的约束条件,而且所选区域太小不能代表正确的分布特性,容易带来定位错误。 N 应该大于 1 片晶片的平均厚度(6~7pixels)。取 $N=10, T=1.0$ 。为了提高定位准确度,在以上步骤确定的分界点左右 10 个像素范围内进一步选择:依次扫描各点,计算 3 个像素点的方差,选取方差最大值对应坐标为最终上边缘。一阶统计量方法利用纹理分布特征,对不同的光照条件具有较强的适应性,但定位准确度不够。

边缘检测投影法具有较高的定位准确度,但受光照条件影响较大。首先原始图像进行边缘检测,然后对边缘检测后的二值化图像进行纵向投影统计(如图 6 所示),进而确定晶片区域的上边缘。理想拍摄条件可以弱化晶片上边缘的边缘信息及其他干扰信息,经边缘检测及垂直投影分析后,晶片上边缘区域的边缘强度明显低于晶片叠层区域强度,因此可以通过设定阈值的方法予以剔除。但该方法受光照影响较大。不同的光照条件下的样本,对应边缘检测的结果会有不同,对应投影分析后的边缘强度也有不同,影响定位正确性。本文结合两种方法,提出了一种具有较强适应能力的定位方法。具体算法如下:

1) 确定下边缘 $Lower_Boundary$ 。具体为:依次扫描图像各行,获得垂直投影曲线;设置分割阈值 T_d (本文选取为图像下半部分所有像素灰度的均值);从图像最后 1 行往前扫描,如果该行平均灰度大于 T_d ,进行标记,并记录对应行坐标;所有行坐标集合中最大值即对应为下边缘。

2) 一阶统计量方法进行粗定位。从第 1 行开始到 $\text{floor}(Lower_boundary / N) \times N$ (floor 为向下取整运算);对上一步得到的垂直投影曲线每隔 N 点进行一阶统计量分析,计算该区间灰度方差,得到如图 5(b)所示的曲线。从右至左依次扫描各点:如果对应方差小于 T ,说明该点到达分界位置。假设该点对应横坐标为 X ,则可以确定上边缘的大概分布范围 $X \times N \pm N$;在该范围内,依次扫描各点,计算 3 个像素点的方差,方差最大点对应坐标作为次优上边缘。

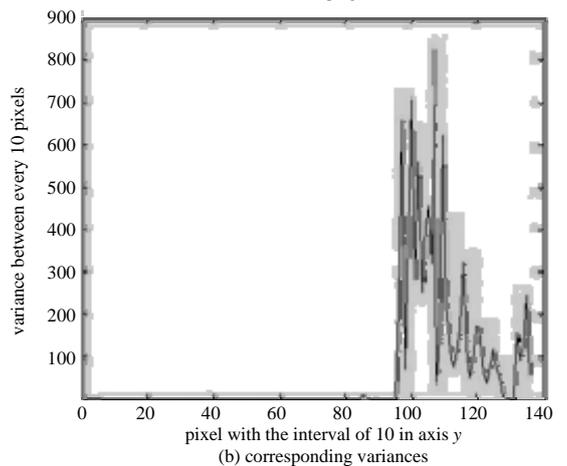
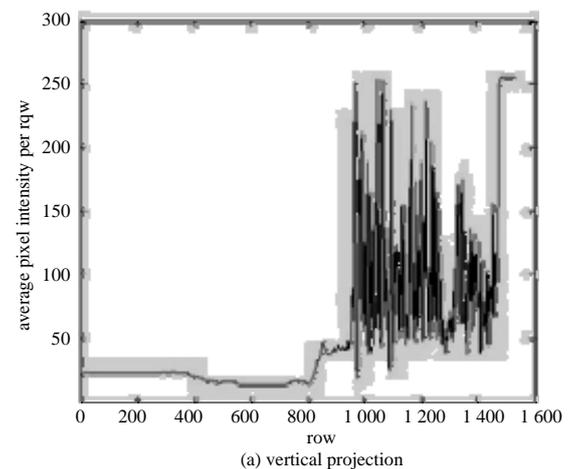


Fig.5 Vertical projection and the corresponding variances
图 5 垂直投影图及相应的方差分布

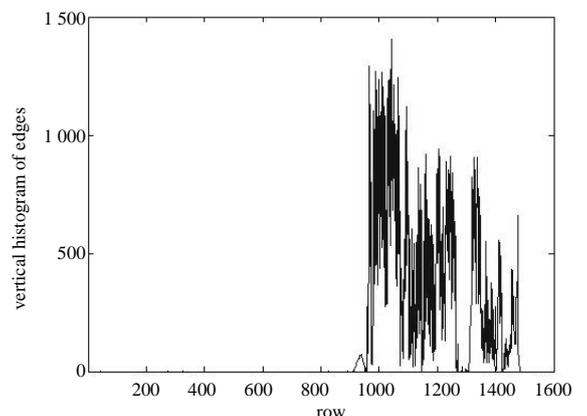


Fig.6 Edge detection
图 6 边缘检测投影图

3) 边缘检测投影法: 对原始图像用 Sobel 算子进行边缘检测。然后对边缘检测后的二值化图像进行纵向投影统计; 扫描得到最大边缘强度 $MaxEdge$, 设置阈值为 T_e 。经大量实验表明 T_e 取为 $0.3 \times MaxEdge$ 取得较好的效果; 再次扫描所有点, 记大于该阈值的点对应的横坐标为集合 Φ 。选取集合 Φ 中最小元素作为最优上边缘。如果该值在步骤 2) 所确定的范围 $X \times N \pm N$ 内, 取该结果作为最终结果。如果在范围之外, 则用步骤 2) 得到次优上边缘作为最终结果。

2.3 计数算法

定位后可以对晶片叠层进行计数。文献[1]在纸张计数中根据像素垂直投影图来计数, 受纸张图像质量影响较大, 而且计数结果与选取阈值相关。能采用这种方法与纸张纹理灰度分布均匀的特点有关; 而晶片纹理灰度变化不规则, 因此不能采用该方法。差分法受二值化效果和噪声影响较大, 需要较多的约束情况, 运算量较大, 实时处理性差。文献[2]从频域角度出发, 将某一列像素点的灰度序列进行傅里叶变换后对频谱计数, 能取得较好的效果, 但该方法局限性在于只能适应于纸张纹理的灰度分布特性, 对不规则的灰度分布缺乏处理能力。本文直接根据某一列像素的灰度分布特性, 通过求极值点的方法来确定晶片块数, 简单而实用。取某一列像素序列, 其灰度分布特性曲线如图 7 所示。

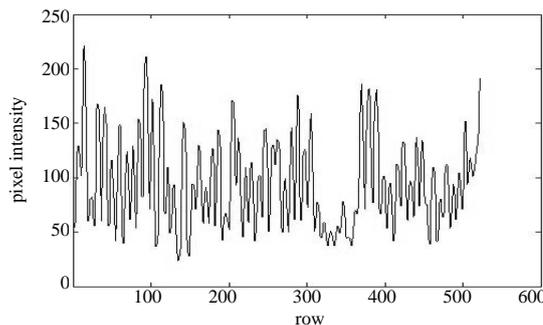


Fig.7 Gray level distribution of the twelve hundredth column
图 7 晶片叠层第 1200 列灰度分布

通过分析可知, 虽然不同明纹、不同暗纹的灰度相差很大(不适合二值化及边缘检测处理), 但相邻的明纹和暗纹之间灰度基本呈单调分布。虽然有时会有不规则如毛刺的出现, 但通过图像平滑如中值滤波能消除这一影响。因此, 计数转化为计算该段曲线的极值点个数(最大值或极小值)。同时, 为了减小误差, 在检测时如果发现两个极大值或极小值之间的距离小于某一阈值 T_c , 则认为该次计数是由于噪声引起, 最后计数结果减 1; 通过实验发现 $T_c=1$ 能取得较好结果。具体实现可以转化为求上升沿或下降沿的个数。

具体算法如下:

1) 从坐标原点出发, 沿 x 轴正向扫描。找到第 1 个上升沿。置计数器 $count=1$; 标记该位置 L_1 。

2) 依次扫描当前点 p 及后继点 p_{next} 。如果 p_{next} 灰度 g_{next} 不小于当前点灰度 g , 继续扫描 p_{next} 的后继点。依次类推, 直到某一后继点的灰度小于当前点灰度值(最近的下降沿), 标记为 p_{des} 。如果 p_{next} 小于 g , 说明 p_{next} 处于下降沿, 同样标记为 p_{des} 。

3) 依次扫描当前点 p_{des} 及后继点 q_{next} : 如果 q_{next} 灰度 g_{qnext} 不大于 p_{des} 灰度 g_{des} , 继续扫描 q_{next} 后继点, 并比较 q_{next} 与其后继点的灰度, 直到某一后继点的灰度大于当前点灰度(说明出现上升沿), 并标记该位置 L_2 ; 如果 $L_2 - L_1 \geq T_c$, 计数器 $count$ 加 1。

4) 重复 2) 和 3), 直到扫描完所有的点。

由于噪声及某些晶片边缘断裂所致, 个别列的计数结果不可避免存在误差, 所以采取隔列抽样统计的计数方法是十分合理的。图像最左和最右部分由于拍摄角度原因不能保持水平, 或纹理为相邻侧面非水平方向的纹理, 也不利于计数, 因此对图像中间部分隔列计数, 选取出现概率最大的结果作为最终的计数结果。

3 计数结果及分析

仿真在 MATLAB 下进行。实验样本包括多组实时拍摄晶片叠层图片。晶片块数分别为 30,50,60,85,100。通过调整拍摄角度和光照条件及背景, 包含不同条件下的各种样本共 50 张。

定位正确率如表 1 所示。定位是否正确的判断标准是检查分割出来的部分是否准确包括了第 1 片和最后 1 片, 如果分割位置在分界线 ± 5 pixels 范围内, 都判决为正确定位。其中定位错误的原因, 一张在于拍摄角度倾斜, 另两张拍摄环境很暗, 对比度太低。

表 1 定位结果统计

Table 1 Performance of positioning			
sample count	success to locate	failed to locate	ratio
50	47	3	94%

对于计数算法, 本文提出的算法与文献[1]的算法对比, 仿真结果如表 2 所示。由表 2 可知, 本算法适应性较强, 既可用于晶片计数, 又可用于纸张计数; 但文献[1]算法不能应用于晶片对象, 究其原因, 主要在于晶片

与纸张的纹理结构存在差异。正如在前文提到,纸张纹理灰度分布均匀,二值化后,通过像素、投影可以实现计数;而对于晶片,纹理复杂,二值化或边缘检测效果不好,导致不能正确计数。

对于纸张,本文算法计数正确率高出 3%;对于晶片,正确率也能达到 95%以上。其中,正确率为完全正确计数样本个数与整样本个数的比值;在错误计数的 5%中,计数偏差大部分为 $\pm 1 \sim \pm 2$ 片。

如果容许 1 片的误差,则计数正确率可以高达 98%;如果容许 2 片的误差,正确率可以高达 99%。对于实际应用, ± 1 片的误差一般为商家接受。

正如前文所述,定位错误是由于拍摄角度倾斜及拍摄环境偏暗造成的,因此,适当控制拍摄环境及拍摄角度,定位正确率可得到一定程度的提高。本文算法计数正确率可达 98%,当然,要完全满足工业生产需要,还需进一步优化算法,提高定位及计数算法的精度及适应能力。

4 结论

本文提出的太阳能晶片叠层计数算法,简单可行,计数正确率高。应用图像处理技术实现太阳能晶片的自动计数,可以降低人力、物力的投入,提高流水线的生产效率。今后需进一步研究倾斜图像的处理、算法复杂度的降低及处理速度的提高。

参考文献:

- [1] 苗良,平西建. 基于纹理特征的纸张计数算法研究[J]. 信息工程大学学报, 2005,6(4):47-50. (MIAO Liang,PING Xijian. Algorithm of Paper Counting Based on Texture Feature[J]. Journal of Information Engineering University, 2005,6(4):47-50.)
- [2] 李毅,阮秋琦. 应用纹理分析的纸张计数算法[J]. 中国图像图形学报, 2004,9(9):1042-1048. (LI Yi,RUAN Qiuqi, Algorithm of Paper Counting Based on Texture Analysis[J]. Journal of Image and Graphics, 2004,9(9):1042-1048.)
- [3] 尹志武,程维明,陈明仪. 条纹图的图像处理方法[J]. 光学精密工程, 1999,7(1):51-58. (YIN Zhiwu,CHENG WeiMing, CHEN MingYi. Image processing about Fringe pattern[J]. Optics and Precision Engineering, 1999,7(1):51-58.)
- [4] 王富治,黄大贵. 基于峰谷形态的纸张纹路分割算法[J]. 电子测量与仪器学报, 2009,23(6):103-107. (WANG Fuzhi, HUANG Dagui. Paper texture segmentation algorithm based on peak-valley morphological feature[J]. Journal of Electronic Measurement and Instrument, 2009,23(6):103-107.)
- [5] 郑光,陈幼平,余文勇. 基于数学形态学的纸张计数算法的研究[J]. 微计算机信息, 2007,23(21):214-216. (ZHENG Guang,CHEN Yaoping,YU Wenyong. The study of algorithm of paper counting based on mathematical morphology[J]. Microcomputer Information, 2007,23(21):214-216.)
- [6] 吴亚复,袁锦海,陈丽花. 一种基于非接触式视觉纸张计数方法和计数机:中国,CN200410049579.7[P]. 2005-12-21. (WU Yafu,YUAN Jinhai,CHEN Lihua. A Paper Counting Method based on Computer Vision:China,CN200410049579.7[P]. 2005-12-21.)
- [7] 修春波,徐勳,宁思光. 纸张计数器:中国,CN200910228936.9[P]. 2010-05-05. (XIU Chunbo,XU Meng,NING Siguang. Paper Couter:China,CN200910228936.9[P]. 2010-05-05.)
- [8] Rafael C, Richard E. 数字图像处理[M]. 2 版. 阮秋琦,译. 北京:电子工业出版社, 2006. (Rafael C, Richard E. Digital Image Processing[M]. 2nd ed. Translated by RUAN Qiuqi. Beijing:Publishing House of Electronics Industry, 2006.)

作者简介:



方 超(1985-), 男, 安徽省安庆市人, 在读硕士研究生, 主要研究方向为图像处理。
email:ttshmily@gmail.com.

谭 伟(1986-), 男, 湖南省沅江市人, 在读硕士研究生, 主要研究方向为图像处理。

杜建洪(1960-), 男, 江西省南昌市人, 博士, 教授, 研究方向为图像处理、无线通信。

表 2 计数算法比较

Table 2 Comparison of counting method		
method	success ratio for paper	success ratio for chip
reference [1]	96%	*
introduced in this article	99%	95%

*: the method introduced in reference [1] cannot be used to count chips