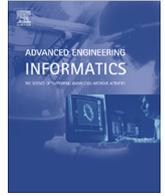




Contents lists available at ScienceDirect

Advanced Engineering Informatics

journal homepage: www.elsevier.com/locate/aei

Defect detection in multi-crystal solar cells using clustering with uniformity measures [☆]

Du-Ming Tsai ^{a,*}, Guan-Nan Li ^b, Wei-Chen Li ^a, Wei-Yao Chiu ^a

^a Department of Industrial Engineering and Management, Yuan-Ze University, Taiwan, ROC

^b College of Mechanical Science and Engineering, Jilin University, Changchun, China

ARTICLE INFO

Article history:

Received 28 April 2014

Received in revised form 25 January 2015

Accepted 28 January 2015

Available online xxx

Keywords:

Defect detection
Surface inspection
Clustering
Fuzzy C-means
Solar cells

ABSTRACT

Solar cells that convert sunlight into electrical energy are the main component of a solar power system. Quality inspection of solar cells ensures high energy conversion efficiency of the product. The surface of a multi-crystal solar wafer shows multiple crystal grains of random shapes and sizes. It creates an inhomogeneous texture in the surface, and makes the defect inspection task extremely difficult. This paper proposes an automatic defect detection scheme based on Haar-like feature extraction and a new clustering technique. Only defect-free images are used as training samples. In the training process, a binary-tree clustering method is proposed to partition defect-free samples that involve tens of groups. A uniformity measure based on principal component analysis is evaluated for each cluster. In each partition level, the current cluster with the worst uniformity of inter-sample distances is separated into two new clusters using the Fuzzy C-means. In the inspection process, the distance from a test data point to each individual cluster centroid is computed to measure the evidence of a defect. Experimental results have shown that the proposed method is effective and efficient to detect various defects in solar cells. It has shown a very good detection rate, and the computation time is only 0.1 s for a 550×550 image.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Solar cells are very important in recent years as an attractive alternative of energy resources. Solar cells are mainly based on crystalline silicon in the photovoltaic industry. Compared to monocrystalline solar cells, multicrystalline solar cells dominate the current market shares due to lower material and manufacturing costs.

Since defects in solar cells degrade the conversion efficiency and usable lifespan, the inspection of solar cells is very important in the manufacturing process. The surface of a multicrystalline solar wafer shows multiple crystal grains of random shapes and sizes, as seen in Fig. 1(a). Fig. 1(b) presents the CCD-captured images of a multicrystalline solar cell. The vertical thin metal lines are finger electrodes. They supply current to the two horizontal bus bars. The crystal grains appear randomly in the surface. The solar cell images thus involve inhomogeneous textures.

Texture analysis techniques [1] in image processing have been used for defect detection of various material surfaces. They mainly aim at uniform or homogeneous texture surfaces. Local textural features or descriptors are extracted either from the spatial domain [2–5] or from the spectral domain [6–10] of a texture image. Discriminant classifiers are then applied to separate local defects from the homogeneous background. They cannot be directly extended to the inspection of heterogeneous surfaces.

Defect detection using machine vision methods has been studied extensively in the literature. However, most of the methods can only handle uniform surfaces, or textured surfaces with homogeneous/repetitive patterns. The target object in this study is multi-crystal solar cells that contain heterogeneous textures. A defect-free solar cell surface may involve various texture patterns in different regions. This makes the inspection task extremely difficult. Image processing techniques have also been applied to the inspection of solar wafers and solar cells. Fu et al. [11] developed a machine vision algorithm to detect cracks in solar cells. The method can only identify cracks that show distinct high-contrast gray levels in the cell edges. Pilla et al. [12] used the thermographic technique to intensify cracks in solar cells. A simple thresholding can then separate the defects from the uniform surface. Warta

[☆] Handled by C.-H. Chen.

* Corresponding author at: Department of Industrial Engineering & Management, Yuan-Ze University, 135 Yuan-Tung Road, Nei-Li, Tao-Yuan, Taiwan, ROC. Fax: +886 (03) 463 8907.

E-mail addresses: iedmtsai@saturn.yzu.edu.tw (D.-M. Tsai), lign09@mails.jlu.edu.cn (G.-N. Li), jeecool@itri.org.tw (W.-C. Li), willchiu@itri.org.tw (W.-Y. Chiu).

[13] reviewed different sensing techniques to intensify defects and impurity in solar cells.

Tsai et al. [14] presented an anisotropic diffusion algorithm for detecting micro-cracks in multicrystalline solar wafers. The diffusion process removes the crystal-grain background, and retains only the crack. The method is very effective and computationally fast, but it can detect only micro-cracks. It cannot be directly used for other defect types. Li and Tsai [15] proposed a machine vision algorithm to detect saw-mark defects in the solar wafer cutting process. The Fourier transform is used to remove the crystal grain background. A Hough-like line detection process is then applied line by line in the filtered image to detect possible defect points. It can detect only the specific saw-mark defects. Chiou and Liu [16] used a near infrared imaging system to highlight micro-cracks in solar wafers. The method works only when the sensed micro-crack is significantly darker than the crystal grains. It may falsely detect a dark, thin defect-free crystal grain as a defect.

A typical CCD camera cannot effectively capture fatal defects such as micro-cracks and subtle finger interruptions. The electroluminescence (EL) imaging technique [17,18] has been introduced to the photovoltaic industry to intensify the deficiencies of a solar cell. The solar cell is first excited with voltage in the EL imaging system. This causes the solar cell to emit the infrared light. A cooled Si-CCD camera then captures the infrared light. Silicon areas with high conversion efficiency present brighter luminescence in the sensed image. Deficiencies of a solar cell appear as dark regions in the EL image. Fig. 2(a) presents the EL image of a defect-free solar cell. Fig. 2(b)–(d) shows respectively three EL images with micro-cracks, breaks, and finger interruptions. The defect areas are inactive and do not emit light well. They thus generate dark regions in the EL image. Although the EL image can visually present defects as dark objects, the background also shows dark grain boundaries of the random crystal grains. Automatic visual inspection of solar cell defects in the EL image becomes very difficult. The currently available inspection machine for solar modules in the manufacturing process can automatically acquire the EL image and display it on the monitor. However, it still requires a human operator to visually identify defects from the EL image. It took the operator a few seconds per solar cell to complete the inspection. The distinctly visible defects, such as large-size breaks, can be easily identified by human eyes. However, subtle defects, such as small thin cracks, could be carelessly ignored by the operator.

In this paper, we propose a machine vision scheme to automatically detect micro-cracks, breaks, and finger interruptions of multi-crystal solar cells in EL images. Those defects are the main sources that reduce the conversion efficiency of solar cells. They

could occur during module assembly and material handling. It has been shown [19] that the breakage rate (breakage cells/total cells) accounts for 2% from manufacturing to transportation in the photovoltaic industry. The high defective rate could be a serious problem of a solar cell manufacturer.

The proposed method is based on the clustering technique. Clustering is an unsupervised classification method to separate a set of multivariate data points into meaningful groups. All members within a partitioned group present similar characteristics. Data points in different groups are distinct from each other. The fuzzy C-means (FCM) algorithm [20] is one of the most popular techniques used for clustering. Since the surface of a clear solar cell contains random crystal grains with a variety of grain patterns, the discriminative features extracted from individual crystal-grain patterns will then show a huge number of clusters in the feature space. It may require several tens of clusters to describe all possible grain patterns in a defect-free solar cell. The conventional Fuzzy C-means methods work well to partition a dataset in high-dimensional space into a few clusters. Its performance degrades as the required number of clusters increases.

In this study, we present a binary-tree partition procedure to cluster crystal grain patterns of defect-free solar cells into groups. Given a set of high-dimensional data points, the proposed clustering method first divides the dataset into two groups using the FCM. A uniformity measure of inter-sample distances in a cluster is then calculated for all current clusters, and the one with the worst uniformity is further divided into two small groups using the FCM. The partition process is repeated until the total number of clusters meets a preset value. The proposed binary-tree partition procedure can more accurately cluster a dataset involving a high number of clusters, compared to the conventional FCM.

Since a solar cell involves distinct crystal grain patterns, it is difficult to use a binary classifier to separate samples into defective and defect-free classes. Instead, we train only defect-free samples and group them into multiple clusters using the binary-tree partition procedure. Any test sample that does not show an acceptable small distance from at least one of the trained cluster centroids is then identified as a defect. In order to characterize the local properties in a solar cell image, we design a set of Haar-like features so that each pixel point defined in a small neighborhood window can be represented by the discriminative feature vector. It is expected that the feature vector of a defect point will result in distinctly large distances from all trained cluster centroids.

This paper is organized as follows: Section 2 first describes briefly the conventional FCM method. The proposed binary-tree partition procedure for defect inspection is then described in

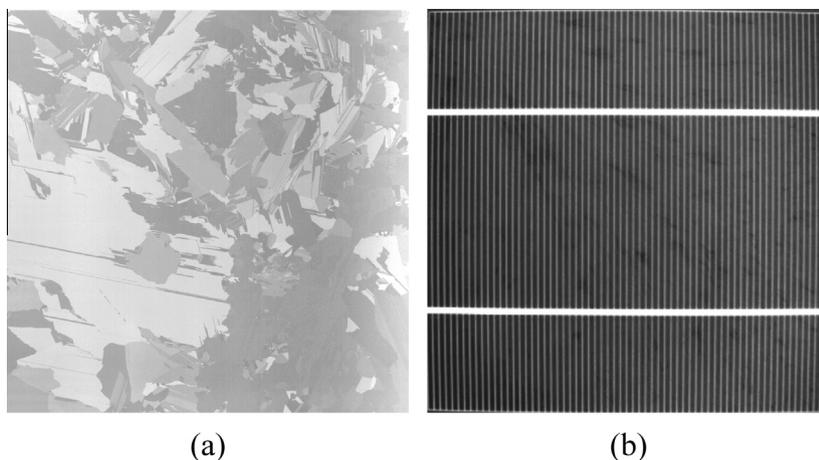


Fig. 1. CCD-captured image of a multi-crystal solar cell: (a) solar wafer and (b) solar cell.

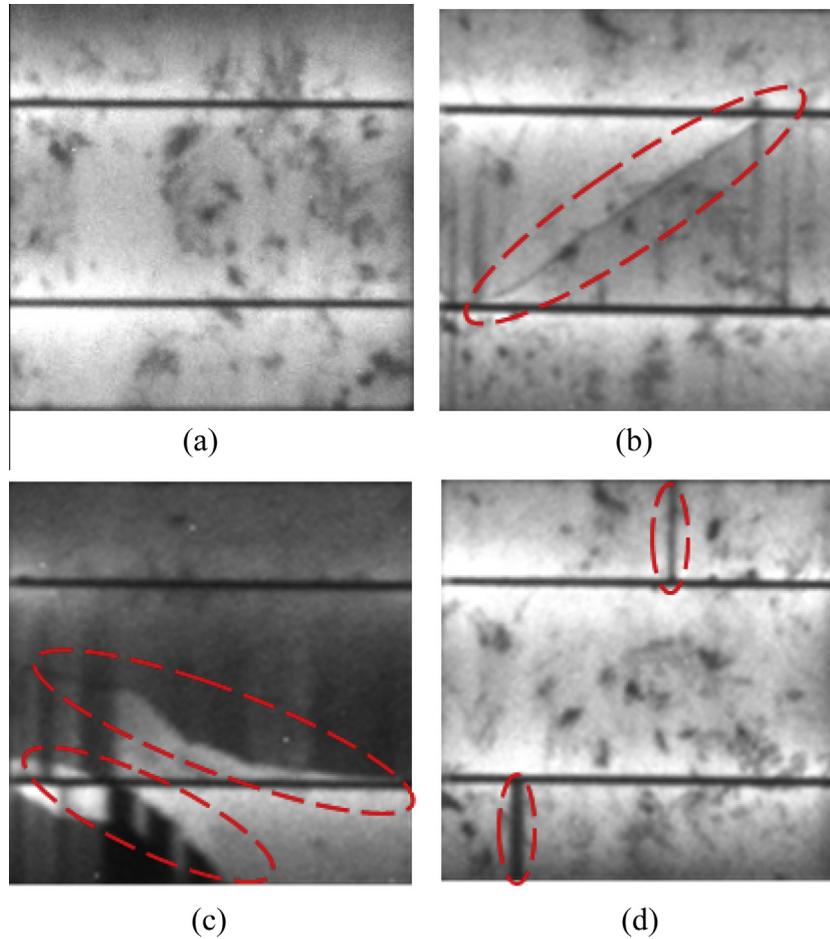


Fig. 2. Demonstrative EL images of solar cells: (a) defect-free sample; (b) micro-crack; (c) break; and (d) finger-interruption.

detail. The Haar-like representation for feature extraction of solar cell images is finally described. Section 3 analyzes the effect of changes in parameter values of the proposed algorithms and presents the experimental results. The paper is concluded in Section 4.

2. Clustering with uniformity measures

The target object studied in this paper is multi-crystal solar cells that involve inhomogeneous texture surfaces. Each individual solar cell surface presents crystal grains of random shapes and sizes. No two solar wafers have the same crystal-grain pattern. In a single solar cell, different local regions show distinct grain patterns and thus different feature values. It means a single class cannot effectively describe the cluttering samples of all defect-free solar cells. If we take different defect-free samples as one single class, the intra-variance of the class will be extremely large. Since defect-free samples have distinct features from cell to cell (and from region to region in a solar cell), a binary clustering or a binary classifier that divides samples into defect-free and defective classes fails to find a clear boundary to separate the two classes.

A defect-free solar cell in the EL image presents distinct texture features in different local regions. For any two arbitrary local regions in a defect-free solar cell, it is difficult to manually determine if these two regions have the similar grain pattern and, thus, have the same class label. Using the traditional binary clustering or classification methods to identify unexpected defects from inhomogeneous texture surfaces becomes impossible. In this section,

we present a new binary-tree partition procedure based on the fuzzy C-means. The conventional FCM and its formulation are first overviewed in Section 2.1. Section 2.2 then describes the proposed clustering procedure for defect detection in EL images of solar cells. Section 2.3 presents the extraction of the Haar-like features.

2.1. Fuzzy C-means clustering

The FCM separates a data set \mathbf{X} into C clusters by minimizing the errors in terms of the weighted distance of each data point \mathbf{x}_i to all centroids \mathbf{v}_c 's of the C clusters. That is,

$$\text{Min } J_{\text{FCM}} = \sum_{c=1}^C \sum_{i=1}^N w_{ic}^p \|\mathbf{x}_i - \mathbf{v}_c\|^2$$

s.t.

$$\sum_{c=1}^C w_{ic} = 1, \quad i = 1, 2, \dots, N$$

where p is the exponent.

The weight w_{ic} is a normalized distance from a sample point i to the cluster c , and is considered as a membership value of the point to the cluster. The weight w_{ic} and the centroid \mathbf{v}_c can be updated by the expectation-maximization (E-M) algorithm:

E-step:

$$w_{ic} = 1 / \sum_{j=1}^C \left(\frac{d_{ic}^2}{d_{ij}^2} \right)^{\frac{1}{p-1}} \quad \text{for } i = 1, 2, \dots, N \text{ and } c = 1, 2, \dots, C$$

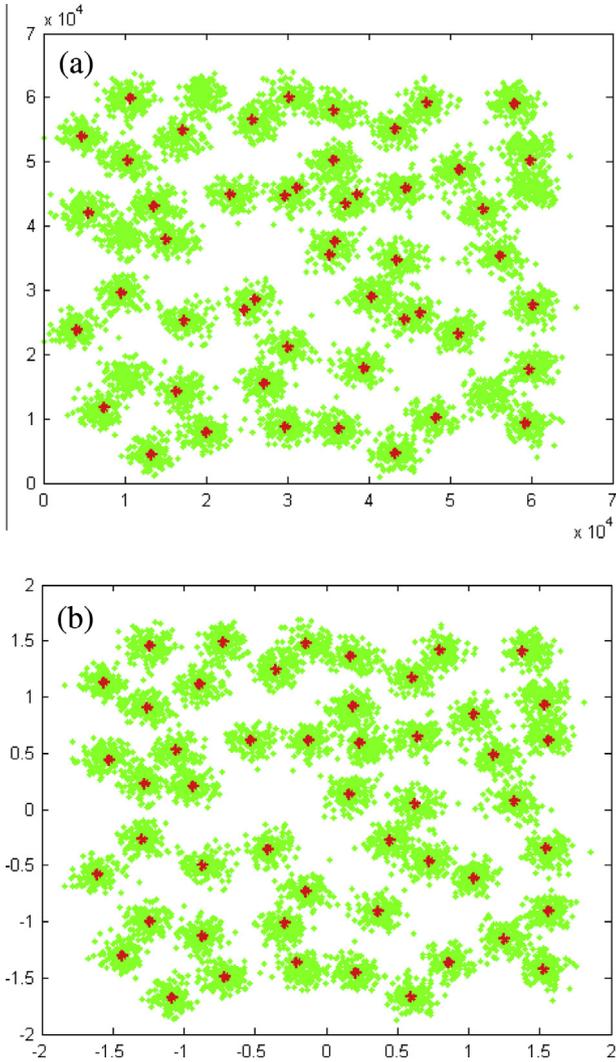


Fig. 3. Clustering of 7500 data points in 50 groups: (a) clustering result from the conventional FCM and (b) clustering result from the proposed method.

where

$$d_{ic}^2 = \|\mathbf{x}_i - \mathbf{v}_c\|^2$$

M-step:

$$\mathbf{v}_c = \frac{\sum_{j=1}^N w_{jc}^p \cdot \mathbf{x}_j}{\sum_{j=1}^N w_{jc}^p} \quad \text{for } c = 1, 2, \dots, C.$$

2.2. Proposed FCM clustering procedure

The conventional model of FCM clustering algorithm can be theoretically used for any dataset that has arbitrary number of clusters. However, the clustering performance is dramatically degraded when the number of clusters is too large. For defect detection in EL images of solar cells, the possible crystal-grain patterns may involve 30 or more clusters. An effective clustering algorithm is required for a high number of multi-group samples.

In order to improve the effectiveness of the FCM for non-spherical, sparsely-distributed datasets, we adopt the distance metric proposed by Tsai and Lin [21]. It incorporates the distance variation in a cluster to regularize the distance between a data point and the cluster centroid. The distance metric is defined as:

$$\hat{d}_{ic}^2 = \frac{\|\mathbf{x}_i - \mathbf{v}_c\|^2}{\sigma_c} \quad (1)$$

where σ_c is the weighted mean distance in cluster c , and is given by

$$\sigma_c = \left\{ \frac{\sum_{j=1}^N w_{jc}^p \cdot \|\mathbf{x}_j - \mathbf{v}_c\|^2}{\sum_{j=1}^N w_{jc}^p} \right\}^{1/2} \quad (2)$$

The distance measure normalizes the distance based on the spread of data points from the centroid in a cluster. The new FCM algorithm, named FCM- σ , searches for C clusters by minimizing the objective:

$$\text{Min } J_{\text{FCM-}\sigma} = \sum_{c=1}^C \sum_{i=1}^N w_{ic}^p \cdot \frac{\|\mathbf{x}_i - \mathbf{v}_c\|^2}{\sigma_c}$$

s.t.

$$\sum_{c=1}^C w_{ic} = 1, \quad i = 1, 2, \dots, N$$

The E-M algorithm is also iteratively used to solve for the weights w_{ic} and the centroids \mathbf{v}_c :

E-step:

$$w_{ic} = 1 / \sum_{j=1}^C \left(\frac{\hat{d}_{ic}^2}{\hat{d}_{ij}^2} \right)^{\frac{1}{p-1}} \quad \text{for } i = 1, 2, \dots, N \text{ and } c = 1, 2, \dots, C$$

where $\hat{d}_{ic}^2 = \frac{\|\mathbf{x}_i - \mathbf{v}_c\|^2}{\sigma_c}$, and σ_c is given by Eq. (2)

M-step:

$$\mathbf{v}_c = \frac{\sum_{j=1}^N w_{jc}^p \cdot \mathbf{x}_j}{\sum_{j=1}^N w_{jc}^p}$$

The defect-free solar cell surfaces present a huge variety of random crystal grain patterns. We therefore propose a new binary-partition clustering based on the uniformity measure of inter-sample distances in a cluster, and apply it for high multi-group clustering. A poor uniformity of a cluster indicates that the cluster contains heterogeneous samples. In a partition level, the uniformity is measured for all current clusters, and the one with the worst uniformity value is further partitioned into two smaller clusters using the FCM- σ . The partition procedure is repeated until the number of clusters meets the predetermined cap.

The uniformity measure of inter-sample distances in a cluster is derived from the principal component analysis (PCA) of the covariance matrix of all inter-samples in a cluster. Given a cluster $\mathbf{X}_c = \{\mathbf{x}_{c,i}, i = 1, 2, \dots, n_c\}$ that contains n_c data points, the data matrix is constructed as

$$\mathbf{A}_c = [\Delta \mathbf{x}_{c,1} \quad \Delta \mathbf{x}_{c,2} \quad \dots \quad \Delta \mathbf{x}_{c,n_c}]^T$$

where $\Delta \mathbf{x}_{c,i} = \mathbf{x}_{c,i} - \bar{\mathbf{x}}_c$ and $\bar{\mathbf{x}}_c = \frac{1}{n_c} \sum_{i=1}^{n_c} \mathbf{x}_{c,i}$.

The covariance matrix of inter-samples in a cluster c is formed by

$$\mathbf{M}_c = \mathbf{A}_c \cdot \mathbf{A}_c^T$$

Let $\lambda = [\lambda_1 \quad \lambda_2 \quad \dots \quad \lambda_{n_c}]$ be the n_c eigenvalues of the covariance matrix \mathbf{M}_c . It is expected that the variance of these n_c eigenvalues should be very small if all the data points are equally spread in the cluster. The variance of λ 's is thus a good indicator of uniformity of inter-sample distances. If the cluster contains heterogeneous data points, the resulting eigenvalues will not present equal distance variances in all coordinates. The variance of eigenvalues will then be distinctly large.

Since high-density and low-density clusters should have the same uniformity magnitude if they present a homogeneous spread in the space, the standard deviation of λ 's normalized with respect to the mean eigenvalue is used as the uniformity measure for data spread in a cluster. That is,

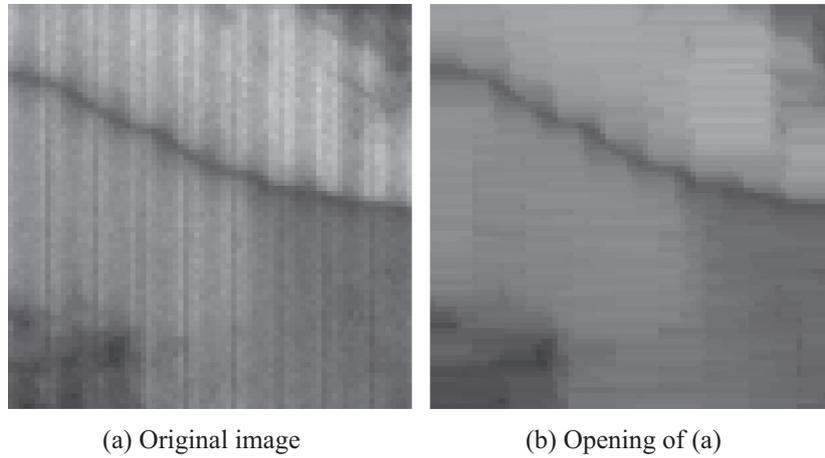


Fig. 4. Morphological opening process for a solar cell image: (a) original image and (b) result of the opening process in (a).

$$u_i = \frac{\sigma_\lambda}{\bar{\lambda}} \quad (3)$$

where $\bar{\lambda}$ and σ_λ are the mean and the standard deviation of λ 's, i.e.

$$\bar{\lambda} = \frac{1}{n_c} \sum_{i=1}^{n_c} \lambda_i$$

$$\sigma_\lambda = \sqrt{\frac{1}{n_c} \sum_{i=1}^{n_c} (\lambda_i - \bar{\lambda})^2}$$

The uniformity measure of inter-samples can well indicate the clusters that need to be further divided.

The binary-tree partition procedure for multi-group clustering is summarized as follows:

Input dataset $\mathbf{X} = \{\mathbf{x}_i, i = 1, 2, \dots, N\}$, and the expected number of clusters C .

Let \mathcal{B} be the collection of clusters. Initially, $\mathcal{B} = \phi$.

Step 1: Apply the FCM- σ to partition \mathbf{X} into two clusters \mathbf{B}_1 and \mathbf{B}_2 , and $\mathcal{B} = \mathbf{B}_1 \cup \mathbf{B}_2$.

Step 2: Compute the uniformity measure using Eq. (3) for every new cluster \mathbf{B}_i , and $\mathbf{B}_i \in \mathcal{B}$. Let $u_i(\mathbf{B}_i)$ be the measured uniformity value of cluster \mathbf{B}_i . Find the cluster in \mathcal{B} with the worst uniformity value, i.e.

$$\mathbf{B}_i^* = \arg \max_{\mathbf{B}_i \in \mathcal{B}} \{u_i(\mathbf{B}_i)\}$$

Partition the cluster \mathbf{B}_i^* into two clusters $\mathbf{B}_{i^*,1}$ and $\mathbf{B}_{i^*,2}$ using the FCM- σ , and let $\mathcal{B} = (\mathcal{B} - \mathbf{B}_i^*) \cup \mathbf{B}_{i^*,1} \cup \mathbf{B}_{i^*,2}$.

Step 3: Let $|\mathcal{B}|$ represent the total number of clusters in \mathcal{B} . Repeat Step 2 if $|\mathcal{B}| < C$. Else, deliver the resulting clusters \mathcal{B} .

In order to evaluate the performance of multi-group clustering with uniformity measures, the public benchmark dataset [22] that contains 7500 two-dimensional data points in 50 groups is used for the test. When the conventional FCM clustering algorithm is applied, the objective function will fall into a local minimum in the iterative process. It then leads to the wrong clustering results as showed in Fig. 3(a). The crosshairs in the figure mark the detected centroids of individual clusters. Some of the homogeneous groups are wrongly assigned to two clusters (i.e., a group with two centroids), and some groups are falsely detected by assigning the centroid in the middle of two clusters.

Fig. 3(b) shows the clustering results of the same test dataset from the proposed partition procedure. All 50 groups are correctly

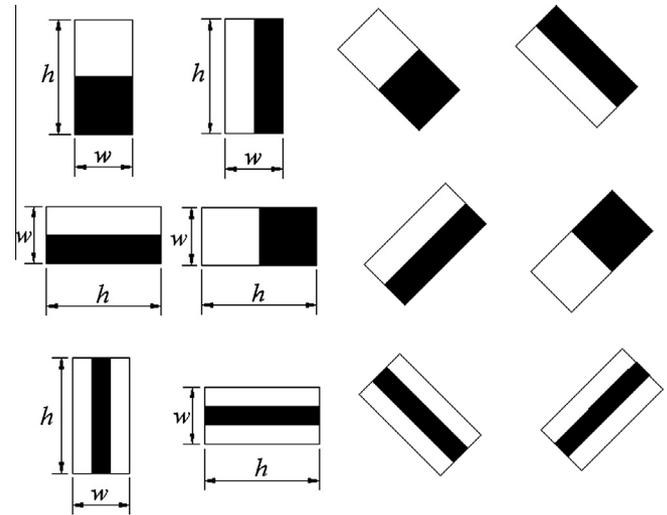


Fig. 5. Twelve prototypes of the Haar-like features used in this study.

identified. Note that the proposed method can reliably partition the overlapping groups into separated clusters. Numerous benchmark datasets have also been evaluated. The experiments consistently show that the proposed clustering with uniformity measures generates similar results as the conventional FCM when the datasets involve only a few groups. The proposed method distinctly outperforms the conventional FCM for the datasets containing tens of groups.

To apply the proposed clustering, the number of clusters C must be predetermined. If the true number of patterns is less than a preset C , some of the resulting clusters may contain only a few samples (e.g. 3 or less) in a cluster. In this case, the number of clusters C can be reduced and the clustering process repeats until each cluster contains a sufficient number of samples. Since the clustering process can be carried out off-line, an appropriate number of clusters C can be always determined prior to the inspection process. In addition, the clustering results are used to determine if a test sample is close enough to at least one of the clusters of defect-free samples. The detection results are generally not sensitive to a small change of the number of clusters C .

2.3. Feature extraction

As seen in Fig. 2, the finger electrodes present repetitive vertical lines in the EL image. The thin finger electrodes could be

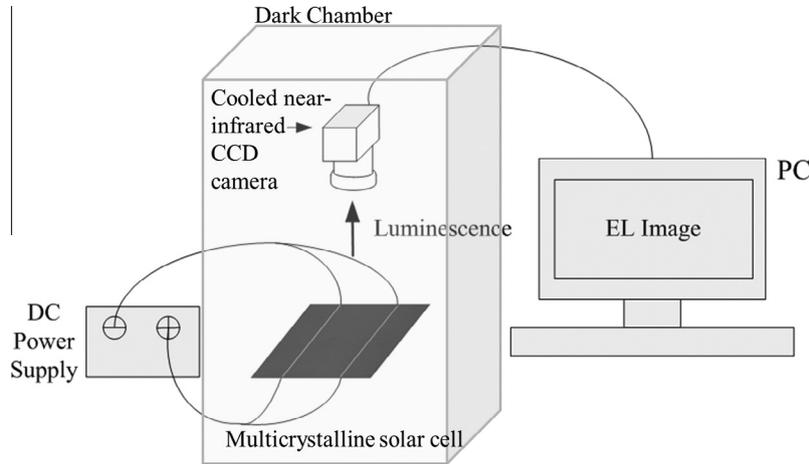


Fig. 6. Configuration of the EL imaging system.

mistakenly identified as line-shaped defects. Morphology opening is thus applied first to the gray-valued EL image to remove the thin finger lines. The pixel value is first replaced with the maximum gray-level in the window of a structuring element. It is then replaced with the minimum gray-level in the same window. The structuring element for gray-level opening is given by a 5×1 row window in this study. The opening operation can remove all the pixels in regions which are too small to contain the 5×1 window. Fig. 4(a) shows the original EL image that contains vertical finger lines. Fig. 4(b) presents the result of the opening operation. The finger lines are removed while the shape of the break defect is well retained in the smoothed image.

For saving the computation time and improving the detection accuracy, only edge points in the smoothed EL image are evaluated for possible defects. The Canny edge operator [23] is used to extract the edges in the EL image of a solar cell. In the training stage of the proposed method, only edge points are taken as the training samples for clustering. In the inspection stage, only edge points of a test image are evaluated.

In order to distinguish various defects of micro-cracks, breaks and finger interruptions from a randomly patterned background, effective discriminative features must be designed and extracted for solar cells in EL images. As seen in Fig. 2(b)–(d), defects are line- or bar-shaped and contain dark gray levels from their surroundings. In this study, the Haar-like features, which are originally proposed by Viola and Jones [24] for face recognition, are extracted for each edge point in the EL image. Fig. 5 shows 12 windows in horizontal, vertical and diagonal directions for the extraction of the Haar-like features, where each window is divided into black and white rectangular regions. The resulting feature value of a given window is the difference of the gray-level sums in the black and white regions. The windows in the first two rows are used to extract line edges and, especially, for breaks. The windows in the last row are used to extract thin lines or strips. All these 8 windows can well capture the characteristics of line- and bar-shaped defects in arbitrary directions. To adaptively detect different defect sizes, a variety of Haar window sizes for each of the 12 window patterns are also generated. Let $h \times w$ be the size of a rectangular window. We set 7 varying values for the width with $w = 6, 12, 18, 24, 30, 36$ and 42 pixels. The height h is set at a fixed value of 18 for possible defect lengths. There are 12 Haar window patterns, each of 7 window sizes, used to extract the discriminative features. Thus, a total of 84 (12×7) features are used to represent the characteristics of each edge point in the image. All the required computation for the 84 Haar-like features can be efficiently implemented with the integral image technique [25]. They can be efficiently applied for on-line, real-time inspection of solar cells in a manufacturing process.

2.4. Defect detection

In this study, only the edge points are used as training samples. They are randomly collected from defect-free solar cell images in the training stage. These training samples are grouped into a pre-determined number of clusters using the proposed binary-tree clustering method described previously.

Let $\mathcal{B} = \{\mathbf{B}_c, c = 1, 2, \dots, C\}$ be the resulting clusters, and $\mathbf{B}_c = \{\mathbf{x}_{c,1}, \mathbf{x}_{c,2}, \dots, \mathbf{x}_{c,n_c}\}$ the cluster c that contains n_c members. In order to detect possible defect points in a test image, an adaptive distance threshold for each resulting cluster is formulated. Because individual clusters may have different densities and the distances of inter-samples may be different from cluster to cluster, the distance threshold must be separately determined for each cluster. The distance between a test edge point and the centroid of a cluster \mathbf{B}_c in \mathcal{B} is calculated. It is then compared with the distance threshold. If there exists at least one cluster that gives the distance less than the threshold, the test edge point is classified as a defect-free point. If the distance violates the distance restriction for all clusters in \mathcal{B} , the test edge point is then identified as a defect point. The adaptive distance threshold T_c of a cluster \mathbf{B}_c is given as follows.

The Euclidean distance from a point \mathbf{x}_i to the cluster centroid \mathbf{v}_c of a cluster \mathbf{B}_c is calculated by

$$d_{c,i} = \|\mathbf{x}_i - \mathbf{v}_c\|, \quad \mathbf{x}_i \in \mathbf{B}_c \quad (4)$$

Let μ_c and σ_c be the mean and standard deviation of distances $d_{c,i}$ for all members in cluster \mathbf{B}_c . The distance threshold for cluster \mathbf{B}_c is adaptively defined by the control limit:

$$T_c = \mu_c + t \cdot \sigma_c, \quad c = 1, 2, \dots, C \quad (5)$$

where t is the control constant.

In the inspection stage, the distance from a test sample \mathbf{x}' to the centroid of cluster c is

$$d_c = \|\mathbf{x}' - \mathbf{v}_c\|, \quad \forall c = 1, 2, \dots, C$$

If the distance is greater than the threshold T_c , the edge point \mathbf{x}' in the test image is marked with 1 for cluster c . Else, it is marked with 0. That is,

$$d_c(\mathbf{x}') = \begin{cases} 1, & \text{if } d_c(\mathbf{x}') > T_c \\ 0, & \text{otherwise} \end{cases}$$

\mathbf{x}' is finally identified as a defect point if its distance is larger than the threshold for every cluster in \mathcal{B} , i.e.

$$\prod_{c=1}^C d_c(\mathbf{x}') = 1$$

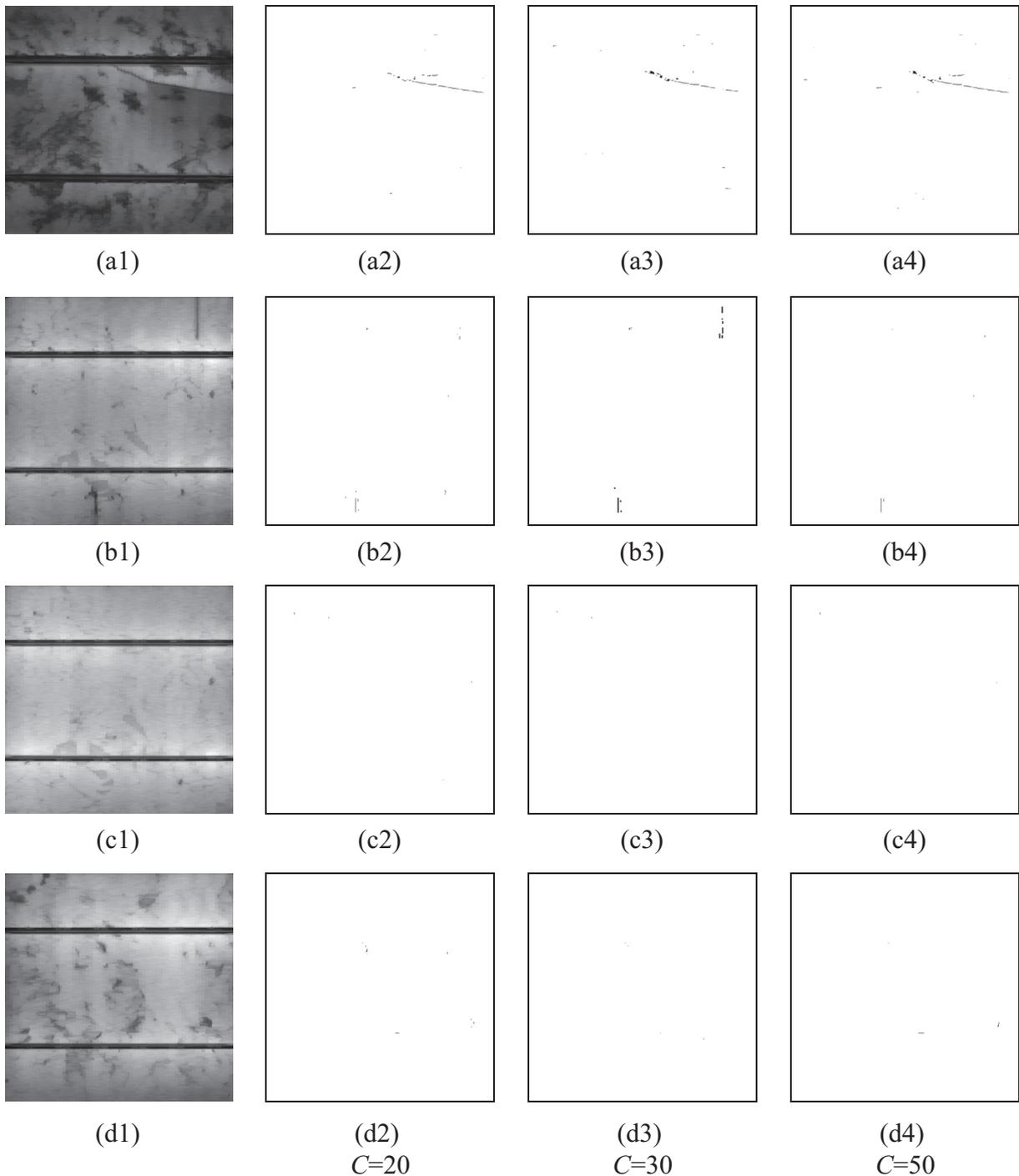


Fig. 7. Effect of changes in the number of clusters with $C = 20, 30$ and 50 , given the control constant $t = 0.02$: (a1) break defect image; (b1) finger-interruption defect image; (c1), (d1) two defect-free images; (a2)–(d2) respective detection results with $C = 20$; (a3)–(d3) respective detection results with $C = 30$; and (a4)–(d4) respective detection results with $C = 50$.

3. Experimental results

The EL imaging system used in the experiment consisted of a dark chamber, in which the solar module was placed. A DC power supply provided voltage to excite the solar module. A cooled near-infrared CCD camera then captured the infrared light emitting from the excited module. The configuration of the EL imaging system is illustrated in Fig. 6. The image size of a solar cell in the experiment was 550×550 pixels with 8-bit gray-levels. The proposed detection algorithms were coded using MATLAB and implemented on

a Pentium Core 2 Duo, 3.0 GHz personal computer. The proposed algorithm currently takes 0.1 s to process a 550×550 image. With a more computationally efficient programming language, such as C# and a high-end personal computer, the computation time of the proposed method can be further improved. For on-line inspection, a solar module comprises 6×6 (i.e. 36) solar cells. The system thus requires 36 images of size 550×550 pixels to complete the inspection of the whole module surface. Only 3.6 s are needed to inspect the entire solar module surface. It is fast enough for on-line and real-time solar module inspection.

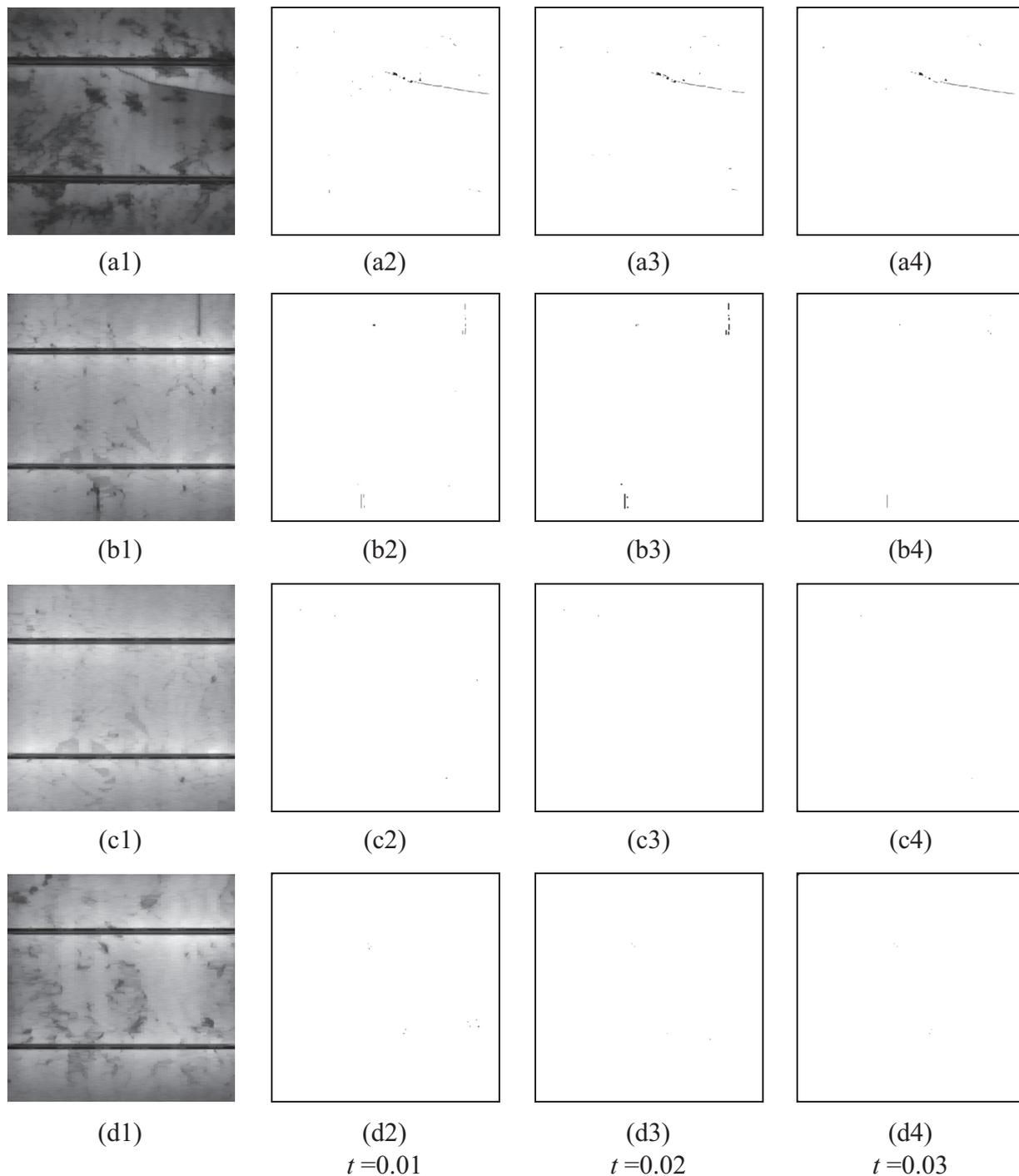


Fig. 8. Effect of changes in control constant with $t = 0.01, 0.02$ and 0.03 , given the number of clusters $C = 30$: (a1) break defect image; (b1) finger-interruption defect image; (c1), (d1) two defect-free images; (a2)–(d2) respective detection results of $t = 0.01$; (a3)–(d3) respective detection results of $t = 0.02$; and (a4)–(d4) respective detection results of $t = 0.03$.

In this study, crack, break and finger-interruption are line- or bar-type defects in solar cells. The image size used in the experiment was 550×550 pixels. Given a spatial resolution of 0.28 mm per pixel for a solar cell of $156 \times 156 \text{ mm}^2$ in its physical size, the minimum detectable size is about 0.84 mm (3 pixels) in width of a defect.

The total number of edge points selected from defect-free solar cell images for clustering in the training stage is around 40,000. The control constant t for the distance threshold of each trained cluster is set at 0.02 . The total number of clusters C is 30. The

height h of the Haar window is fixed at 18. The effects of changes in parameter values of C (number of clusters) and t (control constant) are first presented in Section 3.1. The detection results of the test samples with various defect types are discussed in Section 3.2.

3.1. Effects of changes in parameter values of C and t

The proposed defect detection algorithms have two main parameters: the number of clusters C and the control constant t in Eq.

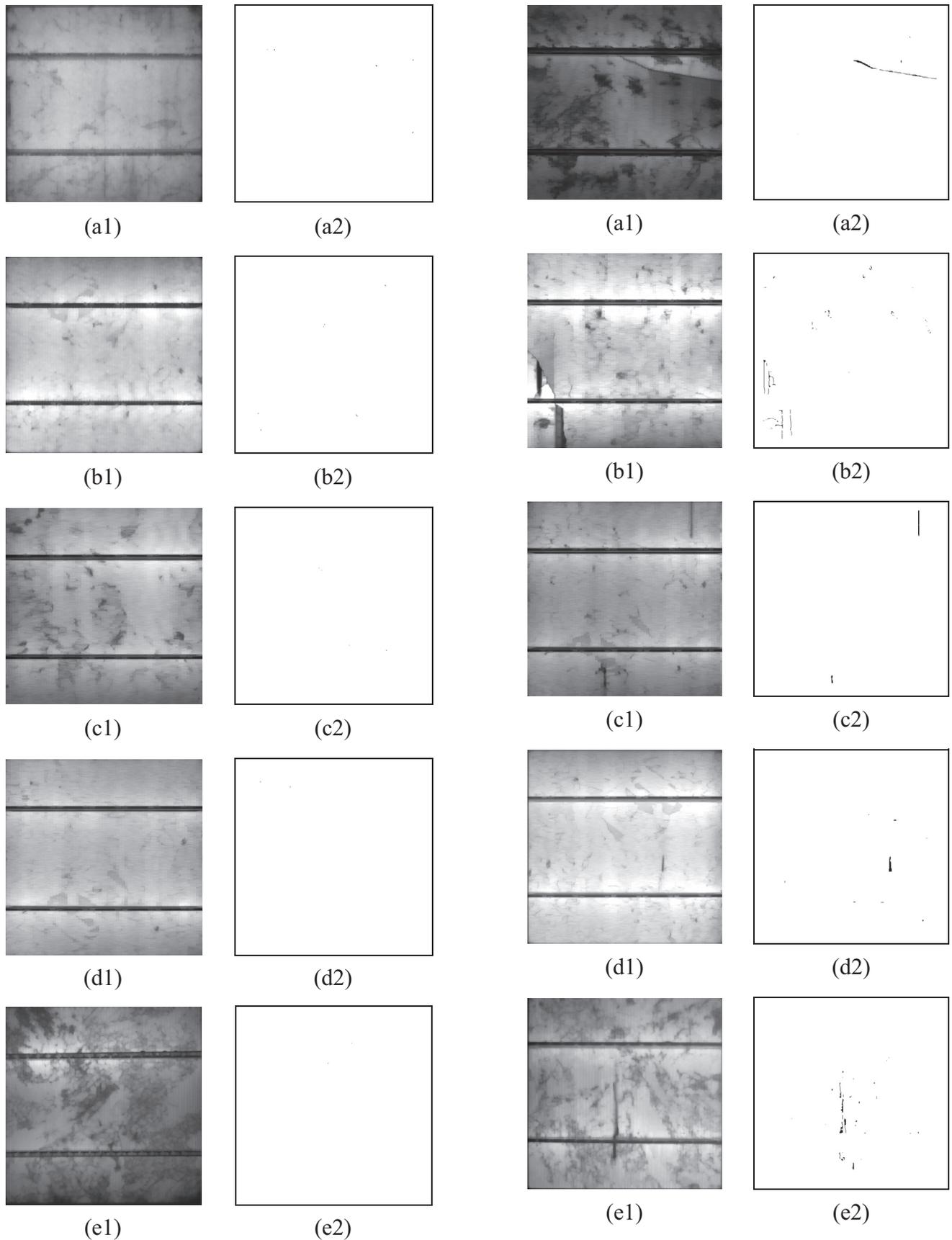


Fig. 9. Detection results of defect-free solar cell images: (a1)–(e1) five faultless samples and (a2)–(e2) respective detection results of the proposed inspection scheme, where no defects are declared.

Fig. 10. Detection results of defective solar cell images: (a1) break defect image; (b1) micro-crack/finger-interruption image; (c1)–(e1) finger-interruption images; and (a2)–(e2) detection results of the proposed inspection scheme.

(5). In order to evaluate effects of changes in parameter values for solar cell inspection, Fig. 7 first presents the detection results with varying numbers of clusters $C = 20, 30,$ and 50 , given that the control constant $t = 0.02$. In the first column of the figure, the test samples sequentially are break defect image (a1), finger-interruption defect image (b1), and two defect-free images (c1)–(d1). The detected defect points are shown in black and the defect-free points are presented in white in the resulting binary images, as seen in Fig. 7(a2)–(d2) with $C = 20$, Fig. 7(a3)–(d3) with $C = 30$, and Fig. 7(a4)–(d4) with $C = 50$. It can be seen from Fig. 7(a2)–(a4) that the break defect can be reliably detected, regardless of the changes in C values. In this study, an insufficient or an excessive number of clusters may miss the detection of subtle defects. As seen in Fig. 7(b2), the results show that the finger-interruption defects are not well detected when the number of clusters is only 20. When the number of clusters is too large, such as 50, the upper-right finger-interruption in the image is also not detected, as seen in Fig. 7(b4). With $C = 30$, the break and finger-interruption defects can be well detected. In consideration of computational efficiency and detection effectiveness, 30 clusters are therefore selected in the subsequent experiments and are recommended for the implementation.

The parameter t sets up the adaptive control limit (threshold) to segment the defect points from the background. A very small t value gives a tight control limit and may result in severe noise. Conversely, an excessively large t value yields a loose control limit and may fail to detect a subtle defect. Fig. 8 shows the detection results as binary images with three control constants $t = 0.01, 0.02$ and 0.03 , given the number of clusters $C = 30$. The first column of the figure shows the same test images as those in Fig. 7. When the control constant is tightly selected, such as $t = 0.01$, the break (as seen in (a2)) can be still detected, and the finger-interruption (as seen in (b2)) can also be identified. There are some minor noisy points detected in the defect-free image, as seen in Fig. 8(d2). The random noise generated in the binary images of defect-free solar cells can be easily removed by simple morphological closing operations. Given a loose control limit, for example $t = 0.03$, the upper-right finger-interruption in the image cannot be identified, as seen in Fig. 8(b4). A control constant $t = 0.02$ gives the best overall performance, where the random noise is removed and the small defects are well detected.

3.2. Detection results

In the experiment, the number of clusters $C = 30$ and control constant $t = 0.02$ are applied to all test samples. Figs. 9 and 10 show the detection results of the proposed method for defect-free and defective images, respectively. In Fig. 9, the first column ((a1)–(e1)) presents five defect-free images for the test. The second column ((a2)–(e2)) presents the detection results of the proposed inspection scheme. In this experiment, we used the morphological closing operation with a 3×3 structuring element to remove noise in the binary images. In the five defect-free test samples, each EL image contains a unique background pattern. The proposed

defect-detection scheme can reliably ignore the random grain boundaries in the detection process and results in clear surfaces in the final binary images with the same parameter setting. Fig. 10(a1)–(e1) shows five defective solar cell images containing break, micro-crack, and finger-interruption defects for the test. It can be seen from Fig. 10(a1) and (e1) that the break and the finger-interruption are embedded within grain boundaries. It makes the inspection task very difficult. Fig. 10(a2)–(e2) presents the detection results. All of the defects are well detected.

In this study, only edge points in the EL image are used for evaluation. Extraction of edges using the Canny edge detector is generally not affected by illumination changes. Also, the dark objects (defects and crystal grain boundaries) in the EL image show relatively high contrast with respect to their surrounding background. The proposed detection method is thus robust to illumination changes. As seen in Fig. 9, the image in (e1) is significantly darker than those in (a1)–(d1). Similarly, the illumination of the image in Fig. 10(a1) is distinctly lower than the bright image in Fig. 10(d1). The proposed method can reliably detect defects under various illuminations.

In order to further verify the detection performance of the proposed method, a total of 50 solar cell images are also evaluated. In the 50 test images, 31 are defect-free and 19 are defective. All the parameter values are the same as those used in the previous experiment. The 19 defect images involve 6 break samples, 3 micro-crack samples, and 10 finger-interruption samples. The detection statistics are summarized in Table 1. In this study, we used a 3×3 structuring element to remove noise. Thus, any black object larger than 3×3 pixels in the final segmented (binarized) image is declared as defect. As long as parts of defects are detected, the system alarms. All of the 19 defective solar cell images are correctly identified, and the false negative rate is zero. Only one of the 31 defect-free solar cell images is falsely detected as a defect, and

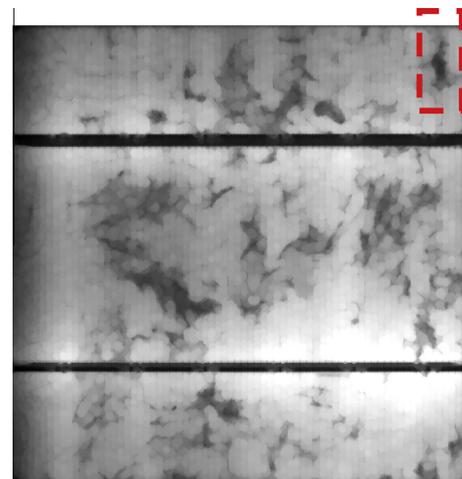


Fig. 11. A small dark object similar to the finger-interruption.

Table 1

Detection results of the proposed method on the 50 test samples.

Sample types	Number of samples	Detection results	
		Defective	Defect-free
Defect-free	31	1	30
<i>Defective</i>			
Break	6	6	0
Micro-crack	3	3	0
Finger-interruption	10	10	0

Table 2

Detection results of the FCM on the 50 test samples.

Sample types	Number of samples	Detection results	
		Defective	Defect-free
Defect-free	31	13	18
<i>Defective</i>			
Break	6	6	0
Micro-crack	3	2	1
Finger-interruption	10	10	0

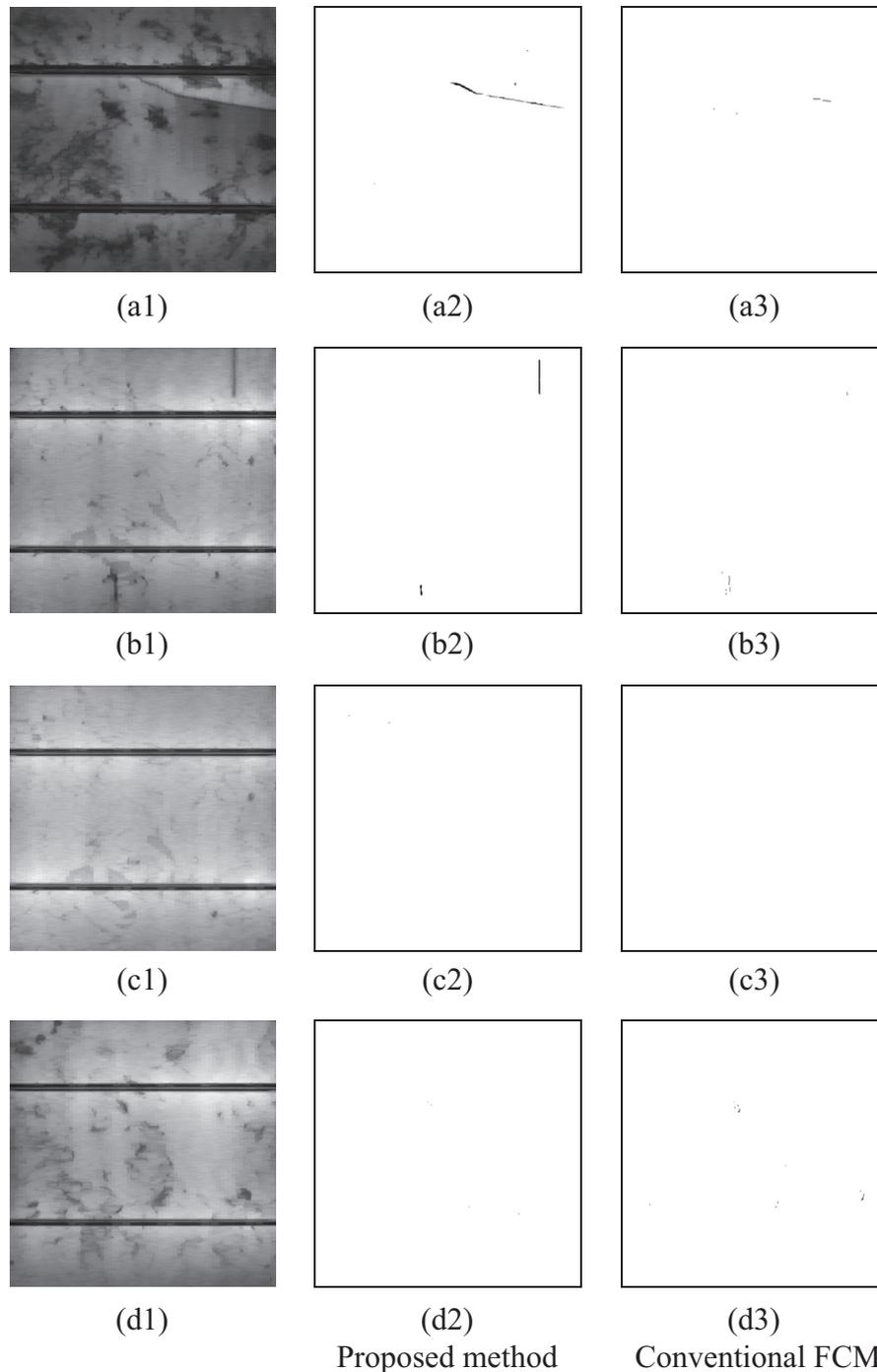


Fig. 12. Comparison of proposed method and conventional FCM with the same parameter settings: (a1) break defect image; (b1) finger-interruption defect image; (c1), (d1) two defect-free images; (a2)–(d2) detection results of the proposed method; and (a3)–(d3) detection results of the conventional FCM.

the false positive rate is 3.2% for the test images in the experiment. The falsely-detected defect-free sample is a dark object similar to the finger-interruption, as seen in the upper-right corner of the EL image in Fig. 11.

For comparing the detection performance, the FCM method is also used to evaluate the same 50 solar cell images. The parameter values of the FCM are selected so that most of the true defects can be identified. For the FCM method in Table 2, one of the 19 defective solar cell images is falsely detected as defect-free, and the false negative rate is 5.2%. However, thirteen of the 31 defect-free solar cell images are wrongly detected as defects. The false positive rate is as high as 41.9%. As for the processing time, the tradition FCM

method is faster than the proposed method for off-line clustering. However, the processing time of on-line inspection is the same, and is invariant to the clustering methods.

Fig. 12 visually demonstrates the comparison of the proposed clustering method and the conventional FCM under the same parameter settings. Fig. 12(a1) is a break defect image. The detection results show that only the proposed clustering procedure can well detect the thin break, as seen in Fig. 12(a2) and (a3). Fig. 12(b1) is a finger-interruption defect image. The conventional FCM method cannot detect the finger interruption, as seen in Fig. 12(b3). Fig. 12(c1) and (d1) are two defect-free images. The conventional FCM shows some noise in the detection results for defect-free

images, as seen in Fig. 12(d3). The proposed clustering procedure declares no defect points for the two defect-free images.

4. Conclusions

This paper has proposed a machine vision method for solar cell inspection in electroluminescence images. The Haar-like features are designed and extracted to represent the characteristics of local crystal-grain patterns. The improved clustering procedure can effectively group a dataset containing tens of clusters by evaluating the uniformity of inter-sample distances in each cluster. In the training process, only defect-free images are taken as training samples, and a simple distance threshold is automatically determined for each cluster. The adaptive distance threshold is easily determined based on the mean and standard deviation of distances from sample points to the cluster centroid for each individual cluster. In the inspection process, the distance from the testing data to individual cluster centroids can be easily computed and compared with the distance threshold. The proposed method effectively detects solar cell defects in electroluminescence images, and gives the location of a detected defect. With the implementation of integral image techniques, the computation time in the inspection process is extremely fast. The proposed method is thus practical for one-line, real-time inspection in the solar cell manufacturing process.

The proposed method mainly focuses on the detection of line- and bar-type defects, including crack, break and finger interruption, which are serious and commonly-occurred in the solar cell manufacturing process. Other types of defects, such as circular-shaped stains and contaminations in solar cells, may need different discriminative features for better inspection results. But, the proposed clustering mechanism can be still applied for the grouping of defect-free samples.

The proposed method in its present form requires an appropriate setting of two critical parameters, the number of clusters C and the control constant t for distance threshold. They are empirically determined in this paper. It is worthy of further investigation for automatic parameter setting in the future.

References

- [1] X. Xie, A review of recent advances in surface defect detection using texture analysis techniques, *Electron. Lett. Comp. Vis. Image Anal.* 7 (2008) 1–22.
- [2] K.V. Ramana, B. Ramamoorthy, Statistical methods to compare the texture features of machined surfaces, *Pattern Recogn.* 29 (1996) 1447–1459.
- [3] J. Iivarinen, K. Keikkinen, J. Rauhamaa, P. Vuorimaa, A. Visa, A defect detection scheme for web surface inspection, *Int. J. Pattern Recognit. Artif. Intell.* 14 (2000) 735–755.
- [4] R. Pan, W. Gao, J. Jiu, H. Wang, Automatic recognition of woven fabric pattern based on image processing and BP neural network, *J. Text. Inst.* 102 (2011) 19–30.
- [5] J. Li, Y. Lu, B. Pu, Y. Xie, J. Qin, W.-M. Pang, P.-A. Heng, An automated cotton contamination detection system based on co-occurrence matrix contrast information, in: *IEEE International Conference on Intelligent Computing and Intelligent System*, vol. 4, 2009, pp. 517–521.
- [6] S.S. Liu, M.E. Jernigan, Texture analysis and discrimination in additive noise, *Comp. Vis., Graph., Image Process.* 49 (1990) 52–67.
- [7] C.-H. Chan, G.K.H. Pang, Fabric defect detection by Fourier analysis, *IEEE Trans. Indust. Appl.* 36 (2000) 1267–1276.
- [8] A. Kumar, Computer-vision-based fabric defect detection: a survey, *IEEE Trans. Indust. Electron.* 55 (2008) 348–363.
- [9] S.G. Mallat, A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Trans. Pattern Anal. Mach. Intell.* 11 (1989) 674–693.
- [10] X. Yang, G. Pang, N. Yung, Robust fabric defect detection and classification using multiple adaptive wavelets, *IEE Proc. Vis., Image Process.* 152 (2005) 715–723.
- [11] Z. Fu, Y. Zhao, Y. Liu, Q. Cao, M. Chen, J. Zhang, J. Lee, Solar cell crack inspection by image processing, in: *International Conference on the Business of Electronic Product Reliability and Liability*, 2004, pp. 77–80.
- [12] M. Pilla, F. Galmiche, X. Maldague, Thermographic inspection of cracked solar cells, in: *Proceedings of SPIE*, 2002, pp. 699–703.
- [13] W. Warta, Defect and impurity diagnostics and process monitoring, *Sol. Energy Mater. Sol. Cells* 72 (2002) 389–401.
- [14] D.-M. Tsai, C.-C. Chang, S.-M. Chao, Micro-crack inspection in heterogeneously textured solar wafers using anisotropic diffusion, *Image Vis. Comput.* 28 (2010) 491–501.
- [15] W.-C. Li, D.-M. Tsai, Automatic saw-mark detection in multicrystalline solar wafer images, *Sol. Energy Mater. Sol. Cells* 95 (2011) 2206–2220.
- [16] Y.-C. Chiou, F.-Z. Liu, Micro crack detection of multi-crystalline silicon solar wafer using machine vision techniques, *Sensor Rev.* 31 (2011) 154–165.
- [17] T. Fuyuki, H. Kondo, T. Yamazaki, Y. Takahashi, Y. Uraoka, Photographic surveying of minority carrier diffusion length in polycrystalline silicon solar cells by electroluminescence, *Appl. Phys. Lett.* 86 (2005) 262108.
- [18] T. Fuyuki, A. Kitiyanan, Photographic diagnosis of crystalline silicon solar cells utilizing electroluminescence, *Appl. Phys. A* 96 (2009) 189–196.
- [19] S.-T. Hsu, Y.-S. Long, Y.-T. Li, Characterization of solar cells in transportation, in: *IEEE International Conference on Photovoltaic Specialist*, 2014, pp. 2592–2594.
- [20] J.C. Bezdek, *Pattern Recognition with Fuzzy Objective Function Algorithm*, Plenum Press, New York, NY, 1981.
- [21] D.M. Tsai, C.C. Lin, Fuzzy C-means based clustering for linearly and nonlinearly separable data, *Pattern Recogn.* 44 (2011) 1750–1760.
- [22] Clustering Datasets <<http://cs.joensuu.fi/sipu/datasets/>>.
- [23] R.C. Gonzalez, R.E. Woods, *Digital Image Processing*, Pearson, Upper Saddle River, NJ, 2008.
- [24] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: *Proceedings of IEEE Computer Vision and Pattern Recognition*, vol. 1, 2001, pp. 511–518.
- [25] F. Crow, Summed-area tables for texture mapping, *Comp. Graph.* 18 (1984) 207–212.