ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

Robust Degraded Document Image Binarization Using Filter

Roshna K Jamal, Kavitha.N.Nair

Dept. of E.C.E, Mahatma Gandhi University College of Engineering, Thodupuzha, Kerala, India

Assistant Professor, Dept. of E.C.E., Mahatma Gandhi University College of Engineering, Thodupuzha, Kerala, India

ABSTRACT: Image binarization technique plays a key role for the extraction of text from degraded document images. In this paper a binarization technique is proposed that uses a guided filter in the preprocessing stage. An adaptive contrast map of the input image is constructed and text stroke edge pixels are detected from this map using otsu's global thresholding algorithm. Then a local threshold value is estimated for text extraction. Further postprocessing is done to enhance the result. The result shows that this method produces better PSNR, F-measure, P F-measure compared to other methods and produces better output image.

KEYWORDS: Document images; binarization; preprocessing; guided filtering

I. Introduction

A large number of degraded historical documents are stored in museums all over the world. These document images suffer from different types of degradations like the appearance of variable background intensity caused by non-uniform intensity, shadows, smear, smudge and low contrast. Extraction of text from these documents are done by using different types of binarization techniques.

Thresholding is an important step in binarization techniques. Thresholding is divided in to global and local thresholding techniques. Global thresholding is done for documents with uniform contrast delivery of background and foreground. For degraded documents where extensive background noise or difference in contrast and brightness exists local thresholding is done. So far different binarization techniques have been developed. Quality of thresholding method determines the quality of output image obtained.

In this paper we propose a binarization technique that addresses these issues. The technique is done in steps. First step is the preprocessing stage using guided filter which is an edge preserving filter. It removes noisy pixels. Then adaptive contrast map [1] is constructed. This map helps in the extraction of text stroke edge pixels by using otsu's global thresholding algorithm. Then a local threshold value is estimated for text extraction. Further postprocessing [1] is done to enhance the result.

The rest of this paper is organized as follows. Section II reviews the current binarization techniques. Robust document binarization technique using filter is described in section III. Then experimental results are shown in section IV to demonstrate the superior performance of the work. Finally, conclusions are presented in section V.

II. Related work

Many thresholding techniques [2] are available. For the global methods a single calculated threshold value is used to classify image pixels in to object or background classes, while for the local methods local area information guides the threshold value for each pixel. For document image binarization, global thresholding methods are not sufficient since document images usually are degraded and have poor quality including shadows, non-uniform illumination, low contrast, large signal-dependent noise, smear and strains.

The local image contrast and the local image gradient are very useful features for segmenting the text from the document background. They are very effective and have been used in many document image binarization techniques. Bernsen [3] described the local contrast as the difference of local maximum and minimum intensities as:

$$C(i, j) = I_{max}(i, j) - I_{min}(i, j) \quad \text{eq. (1)}$$

Where C(i, j) denotes the contrast of an image pixel (i, j), $I_{max}(i, j)$ and $I_{min}(i, j)$ denote the maximum and minimum intensities within a local neighborhood windows of *(i, j)*, respectively. If the local contrast C(i, j)is smaller

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

than a threshold, the pixel is set as background directly. Otherwise it will be classified into text or background by comparing with the mean of Imax (i, j) and Imin (i, j).It cannot work properly on degraded document images.

Later a novel document image binarization technique [4] introduced a normalization factor in denominator in the local image contrast equation as follows:

$$C(x,y) = \frac{fmax\ (x,y) - fmin\ (x,y)}{fmax\ (x,y) + fmin\ (x,y) + \varepsilon} \text{eq. (2)}$$

Where εis a positive but infinitely small number that is added in case the local maximum is equal to 0. The image contrast lowers the image background and brightness variation properly. In particular, the numerator captures the local image difference that is similar to the image gradient. The denominator acts as a normalization factor that lowers the effect of the image contrast and brightness variation. For image pixels within bright regions around the text stroke boundary, the denominator is large, which neutralizes the large numerator and results in a low image contrast. But for image pixels within dark regions around the text stroke boundary, the denominator is small, which compensates the small numerator and accordingly results in a high image contrast. But it's limitation is that it cannot handle bright text properly.

III. PROPOSED METHOD

To analyse the document, it's image is binarized before processing it. It is nothing but segmenting the document background and the foreground text. For the confirmation of the document processing task an accurate document image binarization technique is a must. Binarization technique using guided image filter is done using a sequence of steps as described in coming sections.

Input image is preprocessed using guided filter. Guided filter is an edge preserving filter. Adaptive contrast image map is constructed and text stroke edge pixels are detected by using otsu's global thresholding algorithm. Then local threshold value is estimated for text pixel detection. Postprocesing is done to improve the output image.

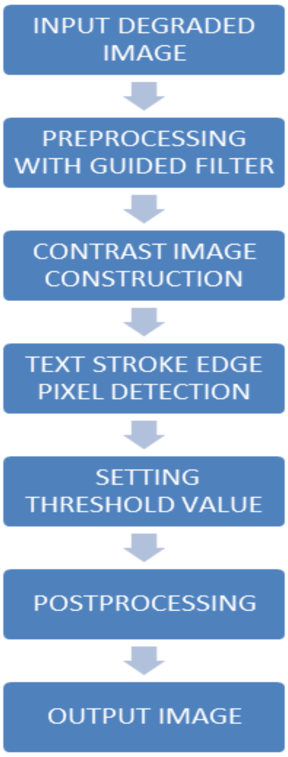

Fig. 1.Block diagram of binarization technique using guided filter

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

A. Preprocessingusing guided filter

Here the input image is taken. If the input image is a colour image it is converted to grayscale one. For degraded and poor quality documents, a preprocessing stage of the grayscale source image is essential for the elimination of noisy areas, smoothing of background texture and contrast enhancement between background and text areas. Guided filter [5] is used to smooth and restore the original image. It is a linear rotation variant explicit filter in which the filtered output is obtained by considering the guidance image which may be either input image itself. It preserves edges while smoothening. Window radius and regularization parameter are two important design functions to implement filtering operation.

B. Adaptive Contrast map Construction

The local image contrast is combined with the local image gradient and derive an adaptive local image contrast [1] as:

$$C\alpha(i,j) = \alpha C(i,j) + (1-\alpha)(\ Imax\ (i,j) - Imin(i,j)\ ) \qquad \text{eq. (3)}$$

Where C(i, j) denotes the local contrast and (Imax(i, j) − Imin(i, j)) refers to the local image gradient that is normalized to [0, 1]. The local windows size is set to 3 empirically. α is the weight between local contrast and local gradient that is controlled based on the document image statistical information. Ideally, the image contrast will be assigned with a high weight (i.e. large α) when the document image has significant intensity variation. So the proposed binarization technique depends more on the local image contrast that can capture the intensity variation well and hence produce good results. The binarization technique relies more on image gradient and avoid the over normalization problem of previous methods.

The mapping from document image intensity variation to α[1] is modelled by a power function as follows:

$$\alpha = \left(\frac{std}{128}\right)^{\gamma} \text{eq. (4)}$$

Where, Std denotes the document image intensity standard deviation γ is a pre-defined parameter. The power function has a nice property in that it monotonically and smoothly increases from 0 to 1 and its shape can be easily controlled by different γ. γ can be selected from [0,∞], where the power function becomes a linear function when γ = 1. Therefore, the local image gradient will play the major role when γ is large and the local image contrast will play the major role when γ is small.

C. Text Stroke Edge Pixel Detection

The constructed contrast image has a clear bi-modal pattern where the adaptive image contrast computed at text stroke edges is obviously larger than that computed within the document background. Detection of the text stroke edge pixel candidate is done by using Otsu's global thresholding method. For the contrast images shows a binary map by Otsu's algorithm that extracts the stroke edge pixels. The binary map can be further improved through the combination with the edges by Canny's edge detector [6], because CED has a good localization property that it can mark the edges close to real edge locations in the detecting image.

D. Local Threshold Value Estimation

The text can then be extracted from the document background pixels once the high contrast stroke edge pixels are detected. The document image text can thus be extracted based on the detected text stroke edge pixel [4] as follows:

$$R(x,y) = \begin{cases} 1, & I(x,y) \le Emean + Estd/2 \\ 0, & otherwise \end{cases} \qquad \text{eq. (5)}$$

Where Emean and Estd are the mean and standard deviation ofthe intensity of the detected text stroke edge pixels within a neighbourhood window W. The neighbourhood window should be at least larger than the stroke width in order to contain stroke edge pixels.

Since there is no need for a precise stroke width, calculate the most frequently occuring distance between two adjacent edge pixels (which denotes two sides edge of a stroke) in horizontal direction and use it as the estimated stroke width as in algorithm 1[1].

Algorithm 1: Edge Width Estimation

Require: The Input Document Image I and Corresponding Binary Text Stroke Edge Image Edg

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

Ensure: The Estimated Text Stroke Edge Width EW

1: Get the *width* and *height* of *I*

2: forEach Row *i*= 1 to *height* in *Edg*do

3: Scan from left to right to find edge pixels that meet the following criteria:

a) it's label is 0 (background);

b) the next pixel is labeled as 1(edge).

4: Examine the intensities in *I* of those pixels selected in Step 3, and remove those pixels that have a lower intensity than the following pixel next to it in the same row of *I*.

5: Match the remaining adjacent pixels in the same row into pairs, and calculate the distance between the two pixels in pair.

6: end for

7: Construct a histogram of those calculated distances.

8: Use the most frequently occurring distance as the estimated stroke edge width *EW*.

E. Postprocessing

It improves the binarization result. Single pixel artifacts along with text are filtered out. Postprocessing is done as in algorithm 2 [1].

Algorithm 2: Post-Processing Procedure

Require: The Input Document Image I, Initial Binary Result B and Corresponding Binary Text Stroke Edge Image Edg

Ensure: The Final Binary Result B f

1: Find out all the connect components of the stroke edge pixels in Edg.

2: Remove those pixels that do not connect with other pixels.

3: for each remaining edge pixels (i, j): do

4: Get its neighborhood pairs: (i − 1, j) and (i + 1, j); (i, j − 1) and (i, j + 1)

5: if the pixels in the same pairs belong to the same class (both text or background) then

6: Assign the pixel with lower intensity to foreground class (text), and the other to background class.

7: end if

8: end for

9: Remove single-pixel artifacts along the text stroke boundaries after the document thresholding.

10: Store the new binary result to Bf.

IV. SIMULATION RESULTS AND DISCUSSION

Binarization technique using guided filter is tested on different types of datasets. The algorithms are implemented using matlab. Datasets contain degraded document images. Contrast map of the images are constructed and text stroke edge pixels are detected by using canny edge detector. Then a local threshold value is estimated and text are extracted. Different degraded images are tested. These documents consist of degradations like water stains, bleed through, significantforeground text intensities. These are evaluated based on Peak Signal to Noise Ratio and F-measure values.

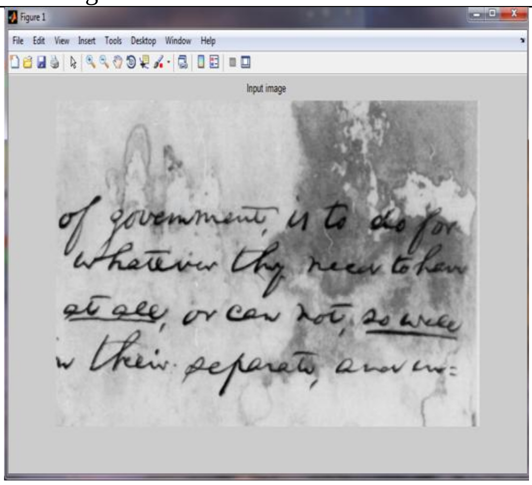

Fig. 2. Input degraded document image

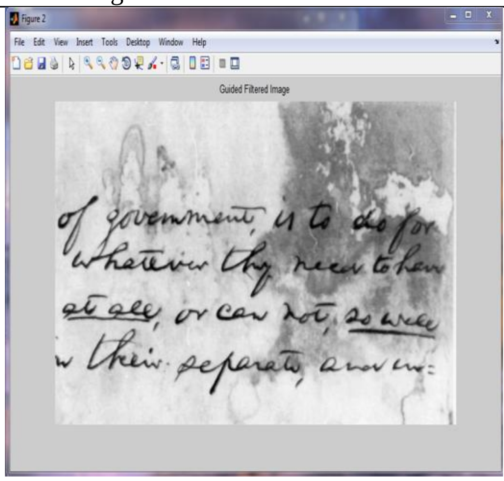

Fig. 3. Output after preprocessing with guided filter

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

The input degraded image is shown in Fig. 2. A preprocessing stage of the grayscale source image is essential for the elimination of noisy areas, smoothing of background texture and contrast enhancement between background and text areas. Guided filter is used to smooth and restore the original image. Output image obtained after preprocessing is shown in Fig. 3. The local image contrast is combined with the local image gradient and derive an adaptive local image contrast map of the input image as in Fig. 4. The contrast map in Fig. 4. helps to detect the stroke edge pixels of the document text properly.

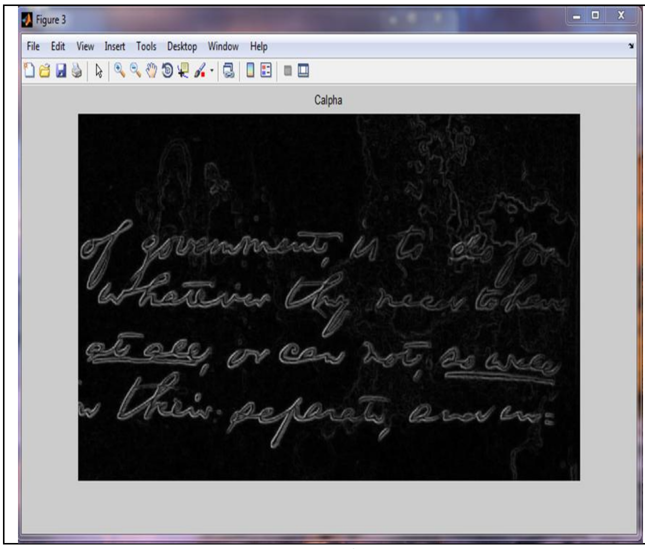

Fig. 4. Adaptive contrast image of input degraded image

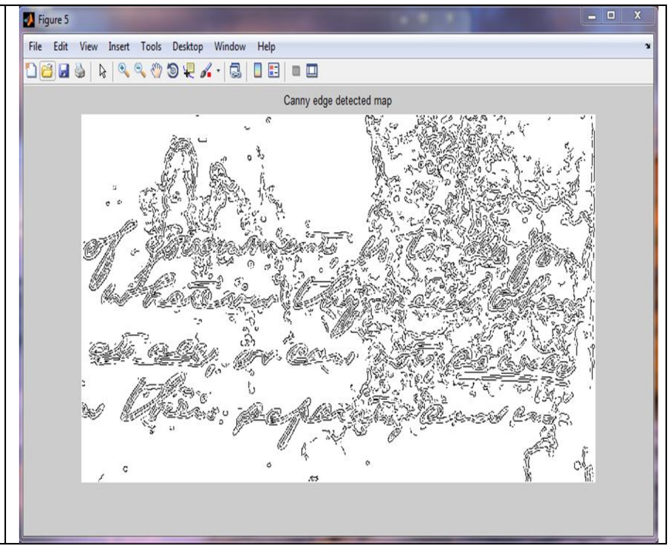

Fig. 5. Canny edge detector output image

In the combined map, keep only pixels that appear within both the high contrast image pixel map and canny edge map in Fig. 5. The combination helps to extract the text stroke edge pixels accurately as in Fig. 6. The stroke edge width *EW* can then be approximately estimated by using the most frequently occurring distances of the adjacent edge pixels as illustrated in Fig. 7.The binarization result can be further improved by postprocessing.Finally, some single-pixel artifacts along the text stroke boundaries are filtered out and an output as in Fig. 8.

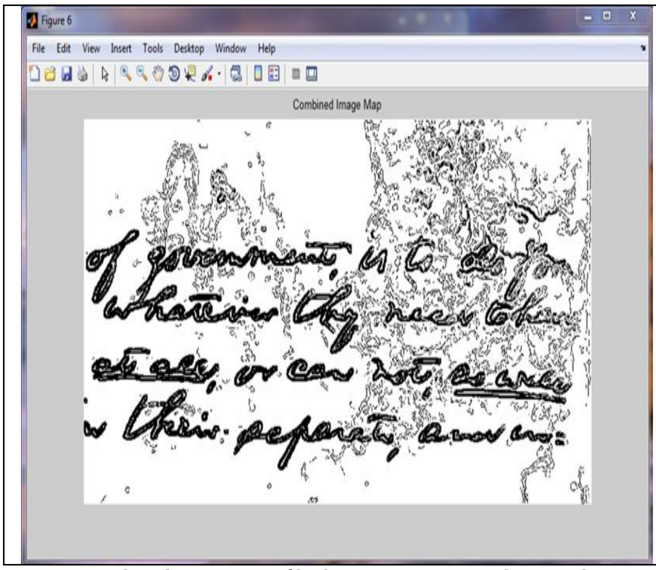

Fig. 6. Combined image map of high contrast image pixel map and

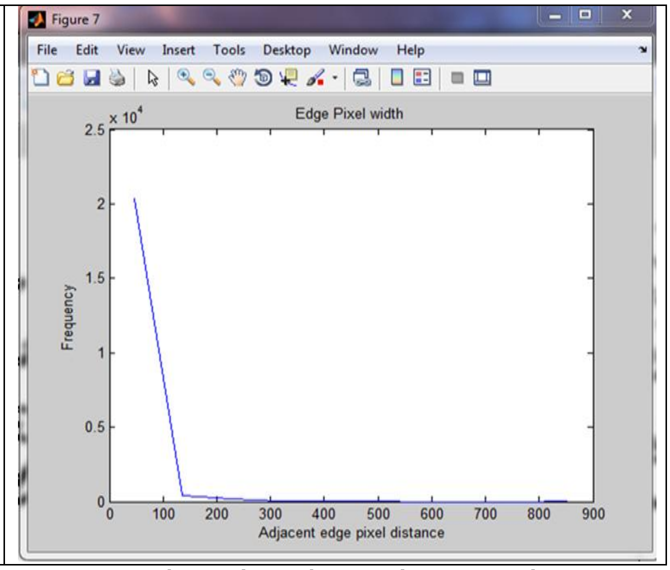

Fig. 7. Histogram showing distance between adjacent canny edge map edge pixels

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

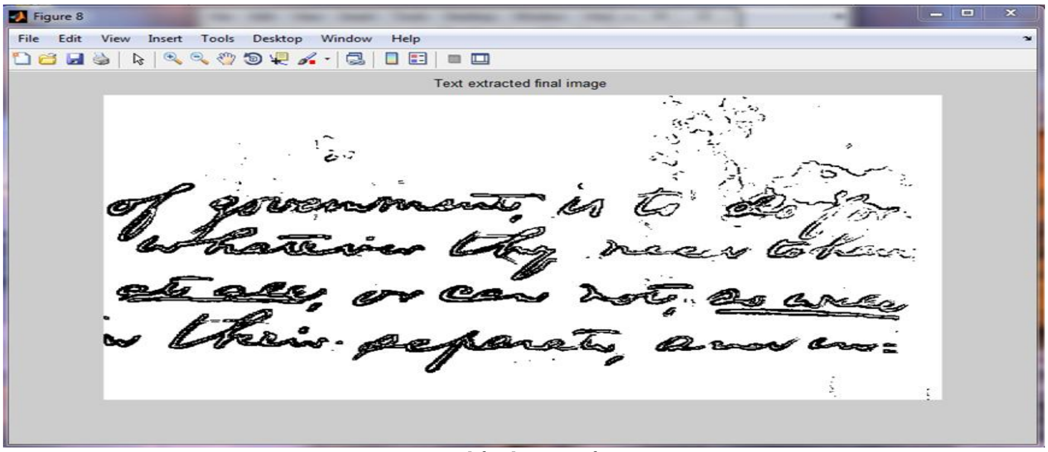

Fig. 8. Text extracted final image after postprocessing

For the evaluation purpose different measures that have been widely used for evaluation purposes are used. PSNR as in eq. (6) is a measure of how close is an image to another. Therefore, the higher the value of PSNR, the higher is the similarity of the two images. F-Measure as in eq. (7) is a measure of test's accuracy. It considers both precision p (number of correct positive results divided by the number of all positive results) and recall r (number of positive results divided by the number of positive results that should have been returned) of the test to compute the score. Pseudo F-Measure as in eq. (8) is introduced as an overall evaluation index. Compared with the typical F-Measure there exist a difference which concerns an alternate measure for recall, namely pseudo-Recall (p-Recall) which is based on the skeletonized ground truth image. Evaluation results of output image obtained by using guided image filter and without using guided image filter are listed in Table 1.

PSNR = $10 \log \left(\frac{C^2}{MSE}\right)$eq.(6)

Where MSE = $\frac{\sum_{x=1}^{M}\sum_{y=1}^{N} \left(I(x,y)-I'(x,y)\right)^2}{MN}$and C is the difference between foreground and background.

FM=$\frac{2\times Recall \times Precision}{Recall + Precision}$eq.(7)

Where Recall =$\frac{TP}{TP+FN}$, Precision = $\frac{TP}{TP+FP}$, TP, FP, FN denote the True positive, False positive and False Negative values, respectively.

P-FM = $\frac{2\times p-Recall \times Precision}{p-Recall + Precision}$eq.(8)

Table 1: Evaluation results of output image with / without using guided filter

	PSNR	F-MEASURE	P F-MEASURE
Without guided filter	14.2212	94.2252	70.1246
With guided filter	16.8875	85.8936	63.1612

From the evaluation table PSNR produces high value compared to the previous methods. F-Measure, P-F-Measure produces comparatively similar values. Thus this binarization technique outperforms the reported binarization methods. The technique is easy and robust, only few parameters are involved. It works well for different kinds of degraded document images. Higher value of PSNR shows the similarity between the two images. This clearly shows that binarization technique extracts texts clearly and removes the noise present in the image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

V. CONCLUSION

This paper presents a binarization technique is very useful for the presevation of historical documents used by scholars and researchers for research studies and for extraction of the text clearly from highly degraded documents. The results shows that the technique gives output image that produce better PSNR values compared to other methods. The use of guided image filter helps to remove noise in the preprocessing stage itself. This improves the accuracy of output image. The performance on Bickley diary dataset and some highlydegraded document images contents still needs to be improved, it can be improved by further studies in future.

REFERENCES

1. B. Su, S. Lu, and C. L. Tan, "Robust document image binarization for degraded document images", *Proc. Vis., Image SignalProcess.*, vol. 22, no. 4, pp. 1408–1417, Dec. 2005.
2. M. Sezgin and B. Sankur, “Survey over image thresholding techniques and quantitative performance evaluation,” J. Electron. Imag.,vol. 13, no. 1, pp. 146–165, Jan. 2004.
3. J. Bernsen, “Dynamic thresholding of gray-level images,” in Proc. Int. Conf. Pattern Recognit., pp. 1251–1255, Oct. 1986.
4. J. B. Su, S. Lu, and C. L. Tan, “Binarization of historical handwrittendocument images using local maximum and minimum filter,” in Proc. Int. Workshop Document Anal. Syst., pp. 159–166, Jun. 2010.
5. K. He, J. Sun and X. Tang, “Guided Image Filtering,” in Proc. IEEE Trans. Image Process., vol. 35, no. 10, pp. 918–921, 2013.
6. J. Canny, “A computational approach to edge detection,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679–698, Jan. 1986.