

# AUTOMATIC GENERATION OF PEN-AND-INK DRAWINGS FROM PHOTOS

*Jiatao Song<sup>#\*</sup>, Zheru Chi<sup>#</sup>, Jilin Liu<sup>\*</sup>, and Hong Fu<sup>#</sup>*

<sup>#</sup> Center for Multimedia Signal Processing, Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong; Email: enzheru@eie.polyu.hk

<sup>\*</sup> Institute of Information and Communication Engineering, Zhejiang University, Hangzhou 310027, P. R. China; Email: sjt6612@163.com

## ABSTRACT

In this paper, a new algorithm for automatic generating pen-and-ink drawings from photos is presented. The basic processing step of our system is the edge extraction based on the multi-resolution property of Wavelet Transform. In order to enhance the variations in texture and tone, the proposed algorithm includes two binarization steps and a noise removing step. Experimental results show that the generated pen-and-ink drawings are of high quality, which accurately illustrate the scenic contents of the original photos.

## 1. INTRODUCTION

For last two decades, computer generation of pen-and-ink drawings has become a very interesting problem in computer graphics. The pen-and-ink drawing, which mainly uses various lines to concisely express the main outlines, poses, volume, motion, etc., is one of the art forms that have been loved by many people. The pen-and-ink drawing which is perfect to depict natural scenes including buildings, roads, grasses, trees, rivers, rocks, etc and various artificial objects [1,2] has been widely adopted in textbooks, product manuals, and other medias.

The pen-and-ink drawing has the following characteristics compared with digital images [3]: (1) Eliminating trivial details, i.e. a pen-and-ink drawing can vividly and concisely depict the main contour and shape of the objects in the scene; (2) There are only two colors (white and black) used in the pen-and-ink drawing and therefore a computer-generated pen-and-ink drawing should be a binary image. So, compared with grey scale images, pen-and-ink drawings use less space to store facilitating efficient transmission via networks; (3) The pen-and-ink drawing can use cheaper output devices such as a B/W printer to produce good outputs; (4) A pen-and-ink drawing consists of much more compact representation of lines and clusters of lines depicting various textures which are very useful for content-based image retrieval (CBIR). Therefore, automatic generation of a pen-and-ink drawing from a color or grey-scale image will provide an

alternative approach for the feature extraction and image representation for CBIR.

There have been two different mechanisms used in computer-generation of pen-and-ink drawings, one is based on geometrical information [3, 4] and the other is based on images [5-7]. The main difference between these two mechanisms is in the input to the system. The former makes use of the description of scene as an input and the latter directly processes a grey-scale image. The main advantage of geometry-based systems is that they can produce drawings whose strokes not only convey the tone and texture of the surfaces in the scene, but by placing strokes along the natural contours of surfaces, they can also convey the 3D forms of the surfaces because they have full access to the 3D geometry and viewing information. The image-based system offers the following advantages: (1) it greatly reduces the tasks of geometric modeling and of specifying surface reflectance properties, allowing much more complicated models, such as furry creatures and human faces, to be illustrated; (2) an image-based system provides the flexibility of using any type of physical photograph, computer-generated image, or arbitrary scalar, vector, or tensor field as input, allowing visualization of data that is not necessarily even physical in nature; (3) image-based systems offer more direct user control: the ability to much more easily modify tone, texture, or stroke orientation with an interactive digital-paint-style interface.

In this paper, a novel algorithm for the automatic generation of a pen-and-ink drawing from a gray-scale image is proposed. The major difference between our method and the conventional image-based method is that in the latter, the input image was often used only as a reference, whether or what lines, texture should be added to the pen-and-ink drawing is determined by the artists, and it was implemented through human-machine interaction, but in our method, the pen-and-ink drawing is fully automatically generated from the input image without human interaction. Experimental results on natural scenery and architectural scenery images show that the pen-and-ink drawings produced by our method are of good visual effects.

## 2. MAJOR FEATURES OF PEN-AND-INK DRAWING

The following is a brief description of some of the salient features of hand-drawn pen drawing. For more information, please refer to [1-2].

- The main elements in the pen-and-ink drawing are pen strokes.
- Texture in a drawing is the collective result of many pen strokes. Each individual stroke is not critical and need not be drawn precisely. A certain amount of irregularity in each stroke is desirable to keep the resulting texture from appearing too rigid or mechanical.
- There are black and white colors only in pen-and-ink drawings. The perceived grey level or tone in a drawing depends largely on how dense the strokes are in a region.
- Shapes in a drawing can be defined by outline strokes. These strokes are exceptional in that they may be long and individually significant. Often the outline is left implicit by a change in tone or texture.

Our system works as follows. As the basic descriptor of object shape, binary edges of an input image are first extracted using a Wavelet Transform (WT) based technique. In order to improve the quality of edges and enhance the effect of texture and tone, edges extracted at every single scale are merged and extracted again. Finally, a noise removing step is taken to make the resulting pen-and-ink drawings more concise.

## 3. THE PROPOSED METHOD

### 3.1. Extracting binary edges at single scale

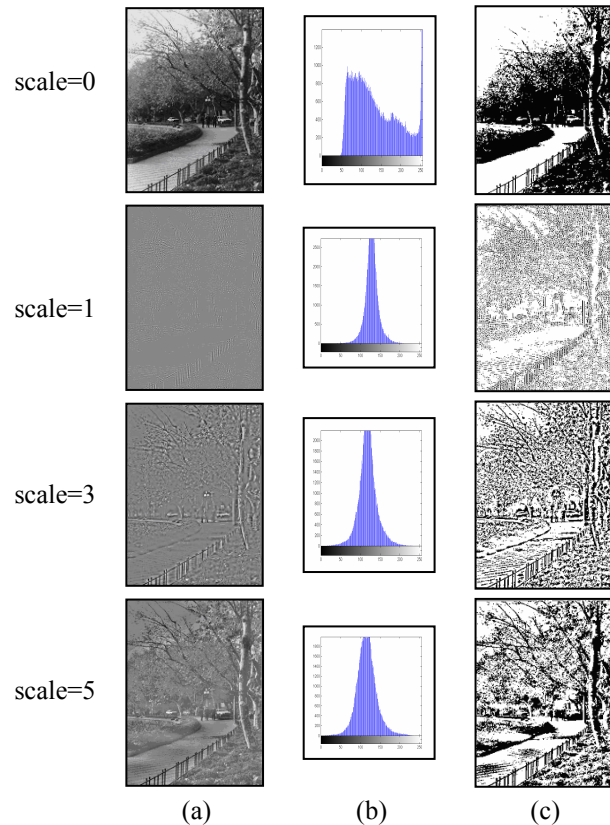
It is well known that WT can be considered as a microscope for signal analysis. From the frequency point of view, WT corresponds to a number of cascaded filter banks, each of which is composed of a low pass and a high pass filter. Using the multi-resolution property of WT, a signal can be decomposed into a series of frequency components, i.e,

$$S = \{A_V, D_V, D_{V-1}, D_{V-2}, \dots, D_2, D_1\} \quad (1)$$

where  $S$  is the original signal,  $A_V$  is the low frequency component at scale  $V$ ,  $D_i (i=1,2,\dots,V)$  is the high frequency component at scale  $i$ .

In order to extract the binary edges, in our method, the high frequency components are first employed to reconstruct the input image. The first column of Fig. 1 shows examples of the high frequency reconstructed images at scales 0 (the original image), 1, 3 and 5, respectively. The image reconstructed at scale  $V$  means that only frequency components  $D_V, D_{V-1}, \dots, D_1$  are exploited for image reconstruction, while the low

frequency component  $A_V$  is discarded. It is not difficult to find that compared with that at scale 0, the histograms at scale 1 and 3 have much smaller dynamic range of gray values, and each histogram has a strong peak in it. This suggests that the image reconstructed with high frequency components is more uniform, with only edges, such as the contour of a road, tree stems, and small tree branches. As the scale increases, the difference between the reconstructed image and the original image reduces.



**Figure 1. Edge extraction at single scale (a) high frequency reconstructed images; (b) histograms; (c) BEIs (scale 0 means the original image).**

By studying the high frequency reconstructed image and its histogram carefully, it can be found that the histogram is almost symmetric and pixels at edges has relatively lower grey values. So, if we select the gray level which is slightly smaller than the average value of the whole image as the threshold to binarize the high frequency reconstructed image, bilevel edges can be extracted. The extraction results, i.e. the binary edge images (BEIs), are shown in the third column of Fig. 1, forming the initial pen-and-ink drawings.

Note that the first row in Fig. 1 is the case in which there are very dense edges in the binarized edge image causing the loss of the depth information of original grayscale image. For example, some thick tree branches in the front are mixed with other trees behind it. This implies that simply binarizing the grey scale image using a global

threshold value is impossible to produce a high quality pen-and-ink drawing.

### 3.2. Edge Merging and Re-extraction

It can be observed from Fig. 1 that when the scale is small, edges are thin, but often disconnected. In other words, too many dots and short strokes have been kept. As the scale increases, more pixels are segmented as foreground pixels, so the edges become clearer and more connective. But if the scale is too large, some details of the objects in the scene may be blurred. In order to achieve a high-quality binary edge image, or obtain a pen-and-ink drawing with better strokes, textures and tone, we propose an algorithm for re-extracting edges. The algorithm includes the following two steps:

- (1) Merging all BEIs obtained at every single scale, i.e.,

$$X = \sum_{i=1}^{LS} X_i \quad (2)$$

where  $X_i$  is the BEI obtained at scale  $i$  and  $LS$  is the largest scale used.

- (2) Thresholding  $X$  with a fixed threshold value between 0 and  $LS$ . (In our experiments,  $LS=6$ , and 3 is selected as the optimal threshold value).

The achieved BEI is shown in Fig. 2(a).

### 3.3. Removing Noise from BEI

After zooming in Fig. 2, it is easy to observe that after the second edge extraction, the quality of a BEI has improved a lot. For example, the layer arrangement of the scene has been properly expressed and the illumination difference between two sides of the tree stem has been properly illustrated. But unfortunately, there are many isolated noise spots degrading the quality of the drawing. So, in order to further improve its performance, the Binary Edge Map (BEM) of the original grayscale image is used for removing noise.

The BEM is obtained using the LAT algorithm [8]. It includes the following steps: (1) Four directional Sobel operators and a low-pass filter template are applied to the original grayscale image  $f(x,y)$ , respectively. The results are denoted by  $Sobel_k(x,y)$  ( $k=1,2,3,4$ ) and  $g(x,y)$ ; (2) The  $LAT(x,y)$  for every image pixel is computed by

$$LAT(x,y) = \frac{\max_{k=1,2,3,4} \{|Sobel_k(x,y)|\}}{g(x,y)} \quad (3)$$

where  $1 \leq x \leq M$ ,  $1 \leq y \leq N$ ,  $M$  and  $N$  are the height and width of the image, respectively. (3) The following rule is applied to classify pixels:

$$pixel(x,y) = \begin{cases} 1 & \text{if } LAT(x,y) > 1 \\ 0 & \text{others} \end{cases} \quad (4)$$

The resulted BEM is shown in Fig. 2(b). It can be seen that compared with BEI, BEM mainly contains the contours of the objects. A main problem with BEM is that there are two edges for long-and-thin objects such as thin tree branches, detoured from the basis of the pen-and-ink drawing technique. The merit of BEM is that there is much less noise spots. Therefore, in our system, BEM is used to remove noise from BEI. It includes the following steps:

- (1) Multiplying BEI by BEM at each pixel, with result storing in matrix  $P$ , i.e.,

$$P(x,y) = BEI(x,y) * BEM(x,y) \quad (5)$$

- (2) Finding out all 8-connected components in BEI and labeling the clusters formed with integers 1, 2, ...,  $N_C$ , where  $N_C$  is the number of clusters. An  $L$  matrix is defined as

$$L(x,y) = \begin{cases} k & \text{if } (x,y) \in \text{cluster } k \\ 0 & \text{if } (x,y) \in \text{background} \end{cases} \quad (6)$$

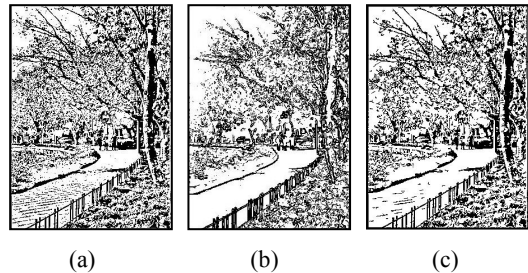
where  $1 \leq k \leq N_C$ .

- (3) Multiplying  $L$  by  $P$  using their corresponding elements and storing the result into matrix  $NI$  i.e.,

$$NI(x,y) = L(x,y) * P(x,y) \quad (7)$$

- (4) Processing the connected components in BEI one by one in the following way: if a nonzero element in  $L$  does not exist in  $NI$  it means that the corresponding foreground pixels in BEI are noise and should be removed.

After the processing, most isolated noise in BEI can be removed. But at the same time, some useful strokes or



**Figure 2. Noise removing of a BEI (a) BEI before noise removing; (b) BEM; (c) BEI after noise removing.**

texture may also be deleted. One solution to this problem is: analyzing all 8-connected components one by one in the difference image of BEIs before and after removing noise, if the aspect ratio  $r$  of the minimum bounding rectangle of a connected component meets:

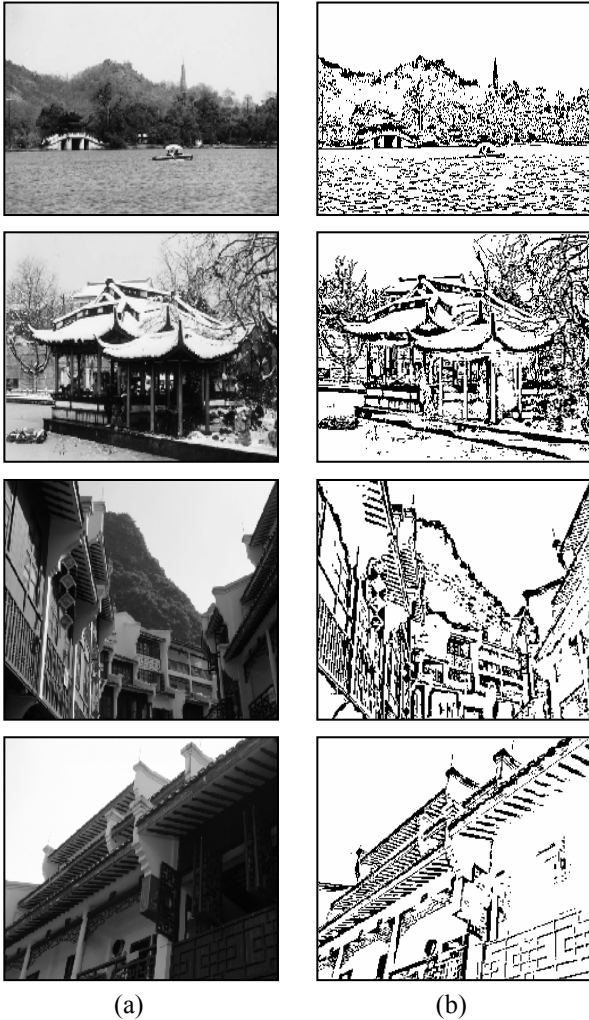
$$r < \frac{1}{3} \quad \text{or} \quad r > 3 \quad (8)$$

then this connected component should be retained.

The BEI after noise removing is shown in Fig. 2(c). It is also the final pen-and-ink drawing. Compared it with Fig. 2(a), most noise in it has been removed.

#### 4. EXPERIMENTAL RESULTS

More than 100 images containing natural scenery or architectures are used for our experiments. The wavelet employed for the edge extraction is *sym4*. Some examples, including the original gray scale image and the corresponding pen-and-ink drawings, are shown in Fig. 3. It can be seen that the generated pen-and-ink drawings have rich textures and tone, which precisely and vividly depict the main features, illumination, and layer information in the original scenery images.



**Figure 3. Examples of experimental results (a) grey scale images; (b) pen-and-ink drawings.**

#### 5. CONCLUSION

This paper presents a novel image-based method for the automatic generation of pen-and-ink drawings from grey scale images. The method is based on an improved wavelet transform edge extraction algorithm. Experimental results show that the generated pen-and-ink drawings are of high quality. The future work will include the human-

interaction facility in our system in order to produce more harmonic pen-and-ink illustration.

#### 6. ACKNOWLEDGEMENTS

The work described in this paper was partially supported by a research grant from the Hong Kong Polytechnic University (Project code: G-T768), Zhejiang Provincial Natural Science Foundation of China under the Grant 602118 and Ningbo Youth Scientific Research Foundation under the Grant 02J20120-02.

#### 7. REFERENCES

- [1] Qin Fan, Yuan Cheng. *Pen-and-ink Drawing*, Hubei Art Publishing, Wuhan, 2003 (in Chinese).
- [2] Gong Xiaobin. *Pen-and-ink drawings of Gardens*, Zhong Guo Wen Lian Publishing, Beijing, 2002 (in Chinese).
- [3] Georges Winkenbach and David H. Salesin. "Computer-generated pen-and-ink illustration". In Andrew Glassner, editor, *Proceedings of SIGGRAPH'94*, pp. 91-100. Orlando, Florida, July 1994.
- [4] John Lansdown and Simon Schofield. "Expressive rendering: A review of nonphotorealistic techniques". *IEEE Computer Graphics and Applications*, 15(3), pp. 29-37, May 1995.
- [5] Michael P. Salisbury, Sean E. Anderson, Ronen Barzel, et al. "Interactive pen-and-ink illustration". In Andrew glassner, editor, *Proceedings of SIGGRAPH'94*, pp. 101-108, Orlando, Florida, July 1994.
- [6] Michael P. Salisbury, Michael T. Wong, John F. Hughes, et al. "Orientable textures for image-based pen-and-ink illustration". In T. Whitted, Editor, *Proceedings of SIGGRAPH'97*, pp. 401-406, Los Angeles, California, 1997.
- [7] Oliver Deussen and Thomas Strothotte. "Computer-generated pen-and ink illustration of trees". In *Proceedings of SIGGRAPH 2000*, pp.13-18, New Orleans, Louisiana, July 2000.
- [8] S. Y. Lee, Y. K. Ham, R. H. Park. "Recognition of human faces using knowledge-based feature extraction and neuro-fuzzy algorithm". *Pattern Recognition*, 29(11), pp.1863-1876,1996.