

# Analysis on Multiresolution Mosaic Images

Ming-Shing Su, Wen-Liang Hwang, and Kuo-Young Cheng

**Abstract**—Image mosaicing is the act of combining two or more images and is used in many applications in computer vision, image processing, and computer graphics. It aims to combine images such that no obstructive boundaries exist around overlapped regions and to create a mosaic image that exhibits as little distortion as possible from the original images. In the proposed technique, the to-be-combined images are first projected into wavelet subspaces. The images projected into the same wavelet space are then blended. Our blending function is derived from an energy minimization model which balances the smoothness around the overlapped region and the fidelity of the blended image to the original images. Experiment results and subjective comparison with other methods are given.

## I. INTRODUCTION

WHEN TWO or more images are overlapped to form a single mixed image, finding an ideal image combination can be difficult. An image mosaic processing technique can be applied to greatly reduce this difficulty. To mosaic an image is to combine overlapped images so that the mixed image contains no obstructive boundaries in the transition region while preserving the general appearance of the original images. An image mosaic is typically completed in two stages. In the first stage, the corresponding points in the two to-be-combined images are identified and registered. This stage is usually referred to as image registration [1]. In the second stage, the intensities of the images are blended after the corresponding points have been registered. Not all applications of image mosaicing require registration, such as in movie special effects. In these cases, unrelated pictures are blended to produce visual effects. In this paper, we assume that the to-be-combined images have either been well registered or require no registration. In other words, we focus on blending image intensities of arbitrary images.

Image combination problems can be found in many fields. Typical examples include photogrammetry, computer vision, image processing, image synthesis, and computer graphics [7], [11], [16]. One significant application of image mosaicing is in the generation of panoramic views from a series of overlapped

snapshot images, such as pictures taken by the Explorer satellite [19]. Another application of image mosaicing in the area of computer graphics is the generation of three-dimensional (3-D) texture mappings. Our study was motivated by the current need for a 3-D model of person's head generated from two-dimensional (2-D) orthogonal views. This works by mapping 2-D snapshot views onto a conformal mapped 3-D head model. Understanding the characteristics of image mosaicing itself is important, and a deep exploration of the underlying theoretical foundation supporting each image mosaicing technique is necessary. Our study on mosaic image analysis aims to develop a theory that supports different image mosaicing techniques and to develop new image mosaic processing techniques based upon the results.

It is justifiable to adopt a weighted average to blend two overlapping images so that the values of pixels within a transition zone are evaluated by a weighted average of the corresponding pixel values in each image. In order to properly blend image features at different resolutions, multiresolution analysis is introduced. In the multiresolution approach, the to-be-combined images are then decomposed into different resolution levels so that they can be blended at each resolution level. Theoretically, in the multiresolution representation of the to-be-combined images, detailed features (such as wrinkles) and large area features (such as skin color) are separated into different resolution levels. Different weighted average functions can be applied to the different resolution levels without interfering with each other. Hence, in a multiresolution image mosaic approach, two main issues are of concern: effectively decomposing images into multiresolution levels and smoothly blending images at each multiresolution level.

Of previous works on multiresolution image mosaicing, the method proposed by Burt and Adelson [2], [3] is the most popular. Their method uses spline functions for blending subband coefficients based on multiresolution pyramidal representation. Their method consists of two major steps. First, two to-be-combined images are decomposed into subband coefficients by means of pyramidal Laplacian operators. Next, spline functions are applied to blend the images at overlap regions where the subband coefficients will be combined. The advantage of this method is that it is easily computed and, in most cases, yields a satisfactory result in terms of human perception. However, no objective measure is mentioned in either of their papers, and, thus, there is no way to determine how effective the method is.

Image blending based on multiresolution pyramidal representation leads to the possibility of using a wavelet transform to combine images in wavelet subspaces. However, direct application of the wavelet approach using the same multiresolution pyramidal concept is problematic in that, if image combination is done by blending wavelet coefficients, the mosaic

Manuscript received February 19, 2003; revised December 10, 2003. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Truong Q. Nguyen.

M.-S. Su is with AnCAD, Inc., Taipei, Taiwan, R.O.C., and he is also with the Institute of Information Science, Academia Sinica, Taiwan, R.O.C., and also with the Department of Information Science and Computer Engineering, National Taiwan University, Taipei, Taiwan, R.O.C. (e-mail: mingshing@ancad.com).

W.-L. Hwang is with the Institute of Information Science, Academia Sinica, Taiwan, R.O.C. (e-mail: whwang@iis.sinica.edu.tw).

K.-Y. Cheng is with the Institute of Information Science, Academia Sinica, Taiwan, R.O.C., and also with the Department of Information Science and Computer Engineering, National Taiwan University, Taipei, Taiwan, R.O.C. (e-mail: kycheng@iis.sinica.edu.tw).

Digital Object Identifier 10.1109/TIP.2004.828416

image will inevitably be effected by the synthesis filter in the inverse wavelet transform. One typical example is found in the wavelet-based image mosaicing method proposed by Hsu and Wu [10]. Their approach is conceptually similar to the multiresolution pyramidal approach proposed by Burt and Adelson; however, the multiresolution Laplacian operators are replaced by wavelet analysis filters.

During our study, we observed that recognition of the quality of a mosaic image is subjective. We generally expect a smooth transition when combining two images of the sky. In some applications, such as combining IC layout images which contain several boxes, we prefer not to over-smooth the sharp edges that exist in the original images. Hence, it is important to build a model for mosaic images that has an objective measure that incorporates certain user parameters so that users can control or improve the behavior of the resultant images.

Energy minimization models have been widely used in combining low-level image properties with higher-level knowledge [12]. We use the minimization of a blending energy function as our model. Within our blending energy function, two variation terms, *image value variation* and *first derivative variation*, are measured and minimized. Image value variation measures the difference between corresponding pixel values of the mosaic image and the to-be-combined image. First derivative variation measures the difference between the first derivative values of the mosaic image and the blended values of each respective first derivative. Our mosaic image can be effectively obtained by minimizing the blending energy function. The quality of the resultant image can be controlled and improved with an additional parameter  $\lambda$  that balances the two variation terms of the blending energy function. This model can be extended to other customized measurements. For example, perceptual measurement could have been included to measure mosaic images. However, proposing such a perceptual measurement is difficult due to its subjectivity. Therefore, the blending energy function in the following discussion is restricted to the image value variation and the first derivative variation.

## II. MULTIREOLUTION IMAGE MOSAICS

A 2-D biorthogonal multiresolution analysis can be obtained by means of a tensor product, from two one-dimensional (1-D) biorthogonal multiresolution analyses, which obtains the scaling functions  $\Phi_{j;k,m}(x,y)$  and  $\tilde{\Phi}_{j;k,m}(x,y)$  and the wavelet functions  $\Psi_{j;k,m}^p(x,y)$  and  $\tilde{\Psi}_{j;k,m}^p(x,y)$  with  $p \in \{H, V, D\}$ , respectively, where

$$\Phi_{j;k,m}(x) = 2^{-j}\phi(2^{-j}x - k)\phi(2^{-j}y - m)$$

and

$$\begin{aligned}\Psi_{j;k,m}^H(x) &= 2^{-j}\phi(2^{-j}x - k)\psi(2^{-j}y - m) \\ \Psi_{j;k,m}^V(x) &= 2^{-j}\psi(2^{-j}x - k)\phi(2^{-j}y - m) \\ \Psi_{j;k,m}^D(x) &= 2^{-j}\psi(2^{-j}x - k)\psi(2^{-j}y - m).\end{aligned}$$

Based on wavelet theory [6], an image  $f(x,y)$  in the  $N$ th multiresolution analysis space can be projected into subspace images, shown in (1) at the bottom of the page, where

$$\langle f(x,y), \Phi_{N;k,m}(x,y) \rangle = \int \int_{-\infty}^{\infty} f(x,y) \overline{\tilde{\Phi}_{N;k,m}(x,y)} dx dy$$

and

$$\langle f(x,y), \Psi_{j;k,m}^p(x,y) \rangle = \int \int_{-\infty}^{\infty} f(x,y) \overline{\tilde{\Psi}_{j;k,m}^p(x,y)} dx dy$$

with  $p \in \{H, V, D\}$ . The original image can then be reconstructed by summing all of the subspace images, i.e.

$$f(x,y) = \sum_{j=0}^N f_j(x,y). \quad (2)$$

Let  $l(x,y)$  and  $r(x,y)$  be the left and right images to be combined along a seam line  $L$  and  $f(x,y)$  be the resultant mosaic image. For simplicity, we assume that  $L$  is  $x = 0$  in the following discussion. Let  $l_j(x,y)$  and  $r_j(x,y)$  be the images obtained by projecting  $l(x,y)$  and  $r(x,y)$  into the wavelet subspace at scale  $2^j$  according to (1). The notation  $f_j(x,y) = \mathcal{M}_{\epsilon_j}(l_j(x,y), r_j(x,y))$  indicates that the subspace image  $f_j(x,y)$  is obtained by blending the left part of  $l_j(x,y)$  and the right part of  $r_j(x,y)$  from seam line  $L$ .  $\mathcal{M}_{\epsilon_j}$  indicates the blending procedure. The parameter  $\epsilon_j$  indicates that the overlapped region of  $l_j(x,y)$  and  $r_j(x,y)$  is within a distance of  $\epsilon_j$  from the seam line  $L$ . Note that, instead of being blended in subband coefficients, images are combined in their wavelet subspaces. As a result,  $f(x,y)$  can be obtained by directly summing the blended subspace image  $f_j(x,y)$  by (2) without being affected by the synthesis filters.

In this paper, we assume that the blending operator  $\mathcal{M}_{\epsilon_j}$  can be expressed as the following polynomial weighting formula to merge the corresponding points in each image:

$$\begin{aligned}\mathcal{M}_{\epsilon_j}(l_j(x,y), r_j(x,y)) &= c_{j,0}(x,y) + c_{j,11}(x,y)l_j(x,y) + c_{j,12}(x,y)r_j(x,y) \\ &\quad + c_{j,21}(x,y)l_j^2(x,y) + c_{j,22}(x,y)r_j^2(x,y) \\ &\quad + c_{j,23}(x,y)l_j(x,y)r_j(x,y) + \text{HigherOrderTerms}(x,y)\end{aligned} \quad (3)$$

$$f_j(x,y) = \begin{cases} \sum_{k,m} \langle f(x,y), \tilde{\Phi}_{N;k,m}(x,y) \rangle \Phi_{N;k,m}(x,y) & j = 0 \\ \sum_{k,m} [\langle f(x,y), \tilde{\Psi}_{j;k,m}^H(x,y) \rangle \Psi_{j;k,m}^H(x,y) + \langle f(x,y), \tilde{\Psi}_{j;k,m}^V(x,y) \rangle \Psi_{j;k,m}^V(x,y) \\ \quad + \langle f(x,y), \tilde{\Psi}_{j;k,m}^D(x,y) \rangle \Psi_{j;k,m}^D(x,y)] & j = 1 \cdots N \end{cases} \quad (1)$$

where  $c_{(\cdot)}(x, y)$  is a weighting coefficient function for each term. In special cases when  $l_j(x, y)$  and  $r_j(x, y)$  are identical, i.e.,  $l_j(x, y) = r_j(x, y) = \check{f}_j(x, y)$ , then the resultant subspace image  $f_j(x, y)$  should be identical to the input subspace image  $\check{f}_j(x, y)$  after applying  $\mathcal{M}_{\epsilon_j}$  to input subspace images. That is

$$\begin{aligned} \check{f}_j(x, y) &= c_{j,0}(x, y) + c_{j,11}(x, y)\check{f}_j(x, y) \\ &\quad + c_{j,12}(x, y)\check{f}_j(x, y) + c_{j,21}(x, y)\check{f}_j^2(x, y) \\ &\quad + c_{j,22}(x, y)\check{f}_j^2(x, y) + c_{j,23}(x, y)\check{f}_j^2(x, y) \\ &\quad + \text{Higher Order Terms}(x, y). \end{aligned} \quad (4)$$

Because (4) must hold for any arbitrary input image  $\check{f}_j(x, y)$ , it is necessary that

$$\begin{aligned} c_{j,0}(x, y) &= 0 \quad c_{j,11}(x, y) + c_{j,12}(x, y) = 1 \\ c_{j,21}(x, y) &= c_{j,22}(x, y) = c_{j,23}(x, y) = 0 \\ \text{Higher Order Terms}(x, y) &= 0. \end{aligned}$$

Now, letting  $w_j(x, y) = c_{j,11}(x, y)$ , operator  $\mathcal{M}_{\epsilon_j}$  becomes

$$\begin{aligned} \mathcal{M}_{\epsilon_j}(l_j(x, y), r_j(x, y)) \\ = w_j(x, y)l_j(x, y) + (1 - w_j(x, y))r_j(x, y). \end{aligned} \quad (5)$$

In the following discussion, we derive an optimal  $\mathcal{M}_{\epsilon_j}$  from an energy minimization model in each wavelet subspace  $2^j$ . The final mosaic image  $f(x, y)$  is obtained from

$$f(x, y) = \sum_{j=0}^N f_j(x, y) = \sum_{j=0}^N \mathcal{M}_{\epsilon_j}(l_j(x, y), r_j(x, y)). \quad (6)$$

### III. BLENDING MODEL

We propose to build a blending model for mosaic images that has an objective measure which incorporates certain user intervention parameters which control the behavior of resultant images according to user requirements. Using an energy function to solve a low-level image processing problem through user intervention was first proposed in [12]. Our blending model which derives the operator  $\mathcal{M}_{\epsilon_j}$  is based upon the minimization of an energy function which measures the variation between the mosaic image and the to-be-combined images. Image intensity and its derivatives are key low-level features. To minimize the image value variation, we impose a constraint that allows the pixel values of a blended image to be as close as possible to the corresponding pixel values of the to-be-combined images. To minimize the first derivative variation we impose the constraint that requires the first derivative of the mosaic images to consistently agree with that of the to-be-combined image.

We formulate our energy functional at scale  $2^j$  as

$$\begin{aligned} E_{\text{blend}}^*(f_j, l_j, r_j) \\ = \int_{-\epsilon_j}^{\epsilon_j} E_{\text{blend}}(f_j, l_j, r_j) dx \\ = \int_{-\epsilon_j}^{\epsilon_j} (E_{\text{image}}(f_j, l_j, r_j) + \lambda_j^2 E_{\text{deriv}}(f_j, l_j, r_j)) dx \end{aligned} \quad (7)$$

where parameter  $\epsilon_j$  indicates that the effective overlapped region of  $l_j(x, y)$  and  $r_j(x, y)$  is within a distance  $\epsilon_j$  from the

seam line  $L$ .  $E_{\text{image}}$  is the image value variation,  $E_{\text{deriv}}$  is the first derivative variation, and  $\lambda_j$  balances these two variation terms at scale  $2^j$ . In practice, we let  $\epsilon_j = 2^j \epsilon$  and  $\lambda_j = 2^j \lambda$ , where  $\epsilon$  and  $\lambda$  are chosen parameters. A larger  $\epsilon$  means that a larger area of one image is influenced by the other image in the blending processing.

In general, the left side of a mosaic image should be similar to the left to-be-combined image, and the right side to the right to-be-combined image. A left cutoff operation  $P^{l, \epsilon_j}$  for the left image at scale  $2^j$ ,  $l_j(x, y)$ , is defined as

$$P^{l, \epsilon_j} l_j(x, y) = \begin{cases} l_j(x, y), & \text{if } x < -\epsilon_j \\ p_{\epsilon_j}(x, y)l_j(x, y), & \text{if } -\epsilon_j \leq x \leq \epsilon_j \\ 0, & \text{if } x > \epsilon_j \end{cases} \quad (8)$$

where  $p_{\epsilon_j}(x, y)$  is a left-to-right monotonically decreasing function satisfying  $p_{\epsilon_j}(x, y) + p_{\epsilon_j}(-x, y) = 1$ . After applying a left cutoff operation to the variation, the right part is cut off and the left part is preserved. The right cutoff operation  $P^{r, \epsilon_j}$  is defined similarly which takes the right side of the right image  $r_j(x, y)$  at scale  $2^j$  as

$$P^{r, \epsilon_j} r_j(x, y) = \begin{cases} r_j(x, y), & \text{if } x > \epsilon_j \\ (1 - p_{\epsilon_j}(x, y))r_j(x, y), & \text{if } -\epsilon_j \leq x \leq \epsilon_j \\ 0, & \text{if } x < -\epsilon_j. \end{cases} \quad (9)$$

Then, we can write the term of the image value variation as

$$\begin{aligned} E_{\text{image}}(f_j, l_j, r_j) &= [P^{l, \epsilon_j} f_j(x, y) - P^{l, \epsilon_j} l_j(x, y)]^2 \\ &\quad + [P^{r, \epsilon_j} f_j(x, y) - P^{r, \epsilon_j} r_j(x, y)]^2. \end{aligned} \quad (10)$$

The minimization of the first derivative variation  $E_{\text{deriv}}$  tends to preserve the consistency of the first derivative between the blended image and the to-be-combined images. Two derivative maps are considered, one being the first derivative values of the blended image and the other being the map obtained from blending the first derivative values of the two to-be-combined images. Then, the first derivative variation is defined so as to measure the difference between these two maps, i.e.,  $E_{\text{deriv}}$  is

$$\begin{aligned} E_{\text{deriv}}(f_j, l_j, r_j) \\ = \left[ \mathcal{M}_{\epsilon_j} \left( \frac{\partial}{\partial x} l_j, \frac{\partial}{\partial x} r_j \right) - \frac{\partial}{\partial x} \mathcal{M}_{\epsilon_j}(l_j, r_j) \right]^2. \end{aligned} \quad (11)$$

From (7), (10), and (11), we get the blending energy function of the blended image  $f_j(x, y)$  at scale  $2^j$  [shown in (12), at the bottom of the next page] where  $w_j$ ,  $p_j$ ,  $l_j$ , and  $r_j$  are abbreviations for  $w_j(x, y)$ ,  $p_{\epsilon_j}(x, y)$ ,  $l_j(x, y)$ , and  $r_j(x, y)$ , respectively. The blending operator  $\mathcal{M}_{\epsilon_j}$  is found by minimizing  $E_{\text{blend}}^*$  in (12) for images  $l_j(x, y)$  and  $r_j(x, y)$ . In summary, the blending problem becomes an optimization problem in our blending model.

### IV. MINIMIZING THE BLENDING ENERGY FUNCTION

The objective energy function of the above (12) must determine functions  $p_j$  and  $w_j$ . The cutoff function  $p_j$  divides the mosaic image to left and right sections in order to compare these overlapped sections with the two original images at scale  $2^j$ . The weighting function  $w_j$  determines how the left and right

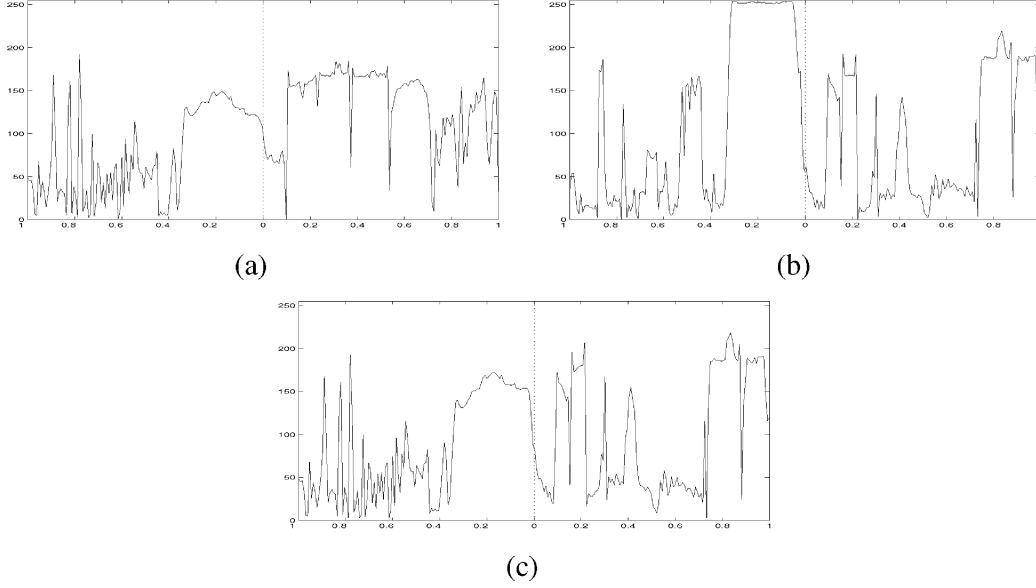


Fig. 1. (a) Scan line of 1024 points from a row in the left image. (b) Scan-line from the same row of the right image. (c) Resultant mosaic signal with  $\epsilon = 1.0$  point width,  $\lambda = 0.5$ , and  $N = 7$  levels.

sections are blended. These two functions are dependent on each other and on the images. The objective function can be minimized by a variational calculus approach. However, finding the optimal solution of the objective function by variational calculus is time-consuming because a partial differential equation with two boundary constraints must be solved. We apply the finite element method [15] to find the solution in a more effective way. With the finite element method, an unknown function is expressed as a linear combination of selected basis functions called finite elements. An optimal solution can be obtained by solving a set of linear equations.

According to the finite element method, we first choose a set of B-spline basis functions  $\mathbf{u}_j(x, y) = [u_j^1(x, y) \cdots u_j^m(x, y)]^T$  called the finite elements [8]. We restrict  $w_j(x, y)$  and  $p_j(x, y)$  to lie within the spaces spanned by a linear combination of these finite elements

$$\begin{aligned} w_j(x, y) &= \sum_{k=1}^m \alpha_j^k u_j^k(x, y) = \boldsymbol{\alpha}_j^T \mathbf{u}_j(x, y) \quad \text{and} \\ p_j(x, y) &= \sum_{k=1}^m \beta_j^k u_j^k(x, y) = \boldsymbol{\beta}_j^T \mathbf{u}_j(x, y) \end{aligned} \quad (13)$$

where  $\boldsymbol{\alpha}_j = [\alpha_j^1 \cdots \alpha_j^m]^T$  and  $\boldsymbol{\beta}_j = [\beta_j^1 \cdots \beta_j^m]^T$ .

TABLE I  
AVERAGE SCORE OF OUR SUBJECTIVE TEST FOR IMAGE MOSAICS

	$\lambda = 0.1$	$\lambda = 0.2$	$\lambda = 0.5$	$\lambda = 1.0$	(1)	(2)
Tower	2.9	2.8	4.3	<b>4.9</b>	3.6	1.8
Butterfly	<b>3.5</b>	3.3	3.3	3.3	3.3	1.6
Landscape	3.8	3.9	<b>4.0</b>	3.9	3.8	2.8
Arena	3.5	4.0	<b>4.4</b>	3.4	4.0	3.9

(1) Burt's method and (2) Hsu's method

Bold numbers give the highest score of each image.

Our intention is to find  $w_j(x, y)$  and  $p_j(x, y)$  that together minimize (12) using the following boundary conditions:

$$\begin{aligned} w_j(x, y)|_{x=0} &= 0.5, w_j(x, y)|_{x=\epsilon_j} = 0, \\ p_j(x, y)|_{x=0} &= 0.5, \text{ and } p_j(x, y)|_{x=\epsilon_j} = 0. \end{aligned} \quad (14)$$

At the middle of the overlapped region ( $x = 0$ ), we give equal weights to both left and right to-be-combined images. The condition at  $x = \epsilon_j$  indicates the diminishing influence of the left image on the right section of the mosaic image at scale  $2^j$ . We first substitute (13) into the linear constraints given in (14) and obtain

$$\begin{bmatrix} \mathbf{u}_j^T(0, y) \\ \mathbf{u}_j^T(\epsilon_j, y) \end{bmatrix} \boldsymbol{\alpha}_j = \begin{bmatrix} \mathbf{u}_j^T(0, y) \\ \mathbf{u}_j^T(\epsilon_j, y) \end{bmatrix} \boldsymbol{\beta}_j = \begin{bmatrix} 0.5 \\ 0 \end{bmatrix}. \quad (15)$$

$$\begin{aligned} E_{\text{blend}}^*(f_j, l_j, r_j) &= \int_{-\epsilon_j}^{\epsilon_j} \left\{ [P^{r, \epsilon_j} f_j(x, y) - P^{r, \epsilon_j} r_j(x, y)]^2 \right. \\ &\quad \left. + [P^{l, \epsilon_j} f_j(x, y) - P^{l, \epsilon_j} l_j(x, y)]^2 + \lambda_j^2 \left[ \mathcal{M}_{\epsilon_j} \left( \frac{\partial}{\partial x} l_j, \frac{\partial}{\partial x} r_j \right) - \frac{\partial}{\partial x} \mathcal{M}_{\epsilon_j}(l_j, r_j) \right]^2 \right\} dx \\ &= \int_{-\epsilon_j}^{\epsilon_j} \left\{ [p_j(1 - w_j)(l_j - r_j)]^2 + [(1 - p_j)w_j(l_j - r_j)]^2 + \lambda_j^2 \left[ \left( \frac{\partial}{\partial x} w_j \right) (l_j - r_j) \right]^2 \right\} dx \end{aligned} \quad (12)$$

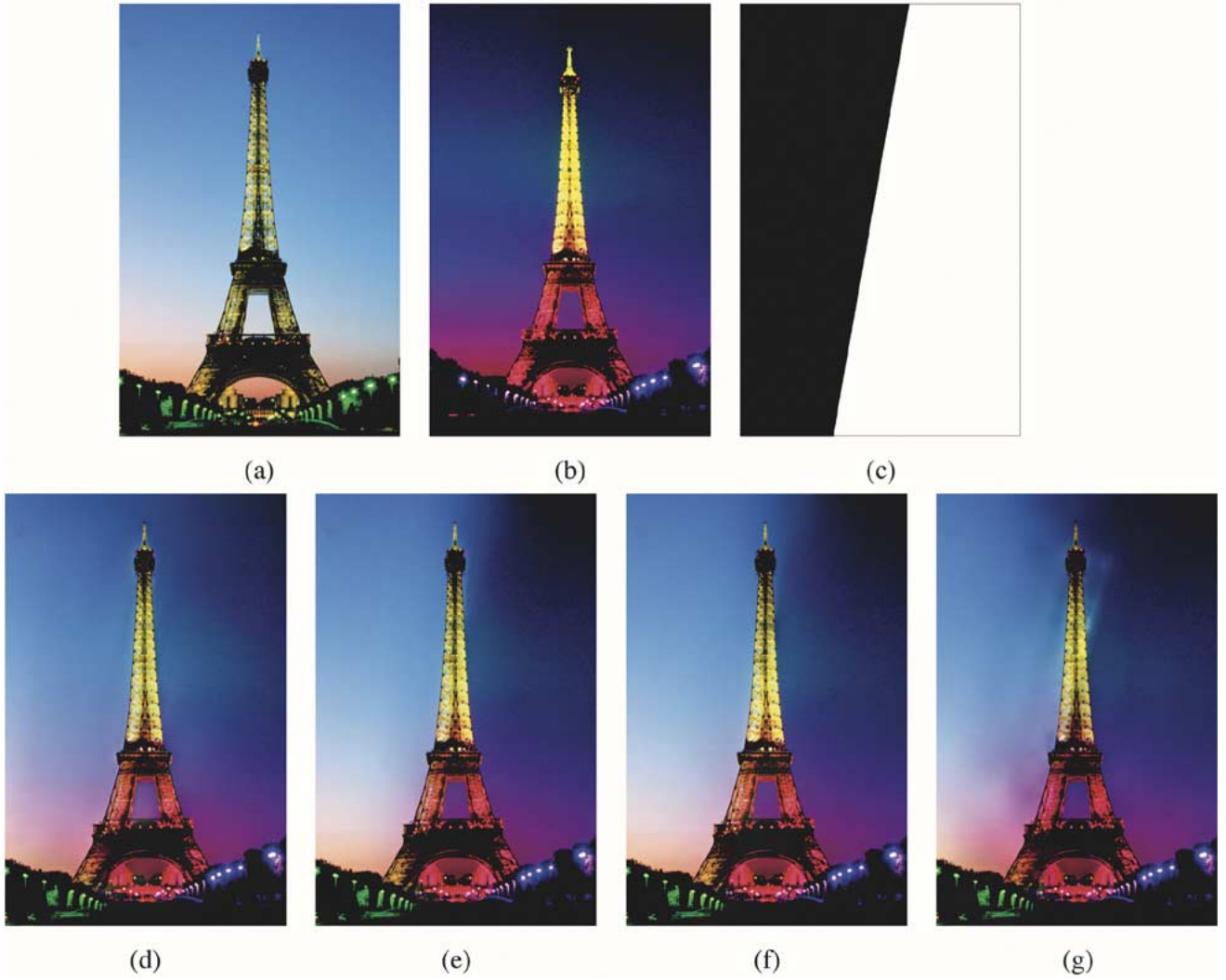


Fig. 2. Images (a) and (b) are the original tower images. (c) shows the seam line along which images (a) and (b) were blended. (d) is our mosaic image obtained with  $\lambda = 1.0$ , which produced the image that the subjects preferred. (e) is our mosaic image obtained with  $\lambda = 0.2$ . (f) and (g) are the mosaic images obtained respectively by Burt's method and Hsu's method.

Equation (13) is then substituted into (12), and the resulting equation is then solved by alternating between finding the optimal alpha vector for  $V_j^\alpha$  and finding the optimal beta vector for  $V_j^\beta$ , where  $V_j^\alpha$  depends only on the alpha vector, and  $V_j^\beta$  depends only on the beta vector. The functions  $V_j^\alpha$  and  $V_j^\beta$  are given as

$$V_j^\alpha(f_j; l_j, r_j, L) = \frac{1}{2} \alpha_j^T \mathbf{H}_j^\alpha \alpha_j + \mathbf{g}_j^\alpha \alpha_j + C_j^\alpha \quad \text{and} \quad (16)$$

$$V_j^\beta(f_j; l_j, r_j, L) = \frac{1}{2} \beta_j^T \mathbf{H}_j^\beta \beta_j + \mathbf{g}_j^\beta \beta_j + C_j^\beta \quad (17)$$

where the *Hessian matrices*  $\mathbf{H}_j^\alpha$ ,  $\mathbf{H}_j^\beta$ ,  $\mathbf{g}_j^\alpha$ ,  $\mathbf{g}_j^\beta$ ,  $C_j^\alpha$ , and  $C_j^\beta$  are

$$\begin{aligned} \mathbf{H}_j^\alpha &= 2 \int_0^{\epsilon_j} \mathbf{u}_j(x, y) \mathbf{u}_j^T(x, y) (2p_j^2(x, y) - 2p_j(x, y) + 1) J_j(x, y) dx \\ &+ 2\lambda_j^2 \int_0^{\epsilon_j} \frac{\partial}{\partial x} \mathbf{u}_j(x, y) \frac{\partial}{\partial x} \mathbf{u}_j^T(x, y) J_j(x, y) dx \end{aligned}$$

$$\begin{aligned} \mathbf{H}_j^\beta &= 2 \int_0^{\epsilon_j} \mathbf{u}_j(x, y) \mathbf{u}_j^T(x, y) (2w_j^2(x, y) - 2w_j(x, y) + 1) \\ &\times J_j(x, y) dx \\ \mathbf{g}_j^\alpha &= -2 \int_0^{\epsilon_j} \mathbf{u}_j^T(x, y) (4p_j(x, y)) J_j(x, y) dx \\ \mathbf{g}_j^\beta &= -2 \int_0^{\epsilon_j} \mathbf{u}_j^T(x, y) (4w_j(x, y)) J_j(x, y) dx \\ C_j^\alpha &= \int_0^{\epsilon_j} (2p_j^2(x, y) + 2) J_j(x, y) dx \\ C_j^\beta &= \int_0^{\epsilon_j} (2w_j^2(x, y) + 2) J_j(x, y) dx \quad \text{and} \\ J_j(x, y) &= [(l_j(x, y) - r_j(x, y))^2 + (l_j(-x, y) - r_j(-x, y))^2]. \end{aligned}$$

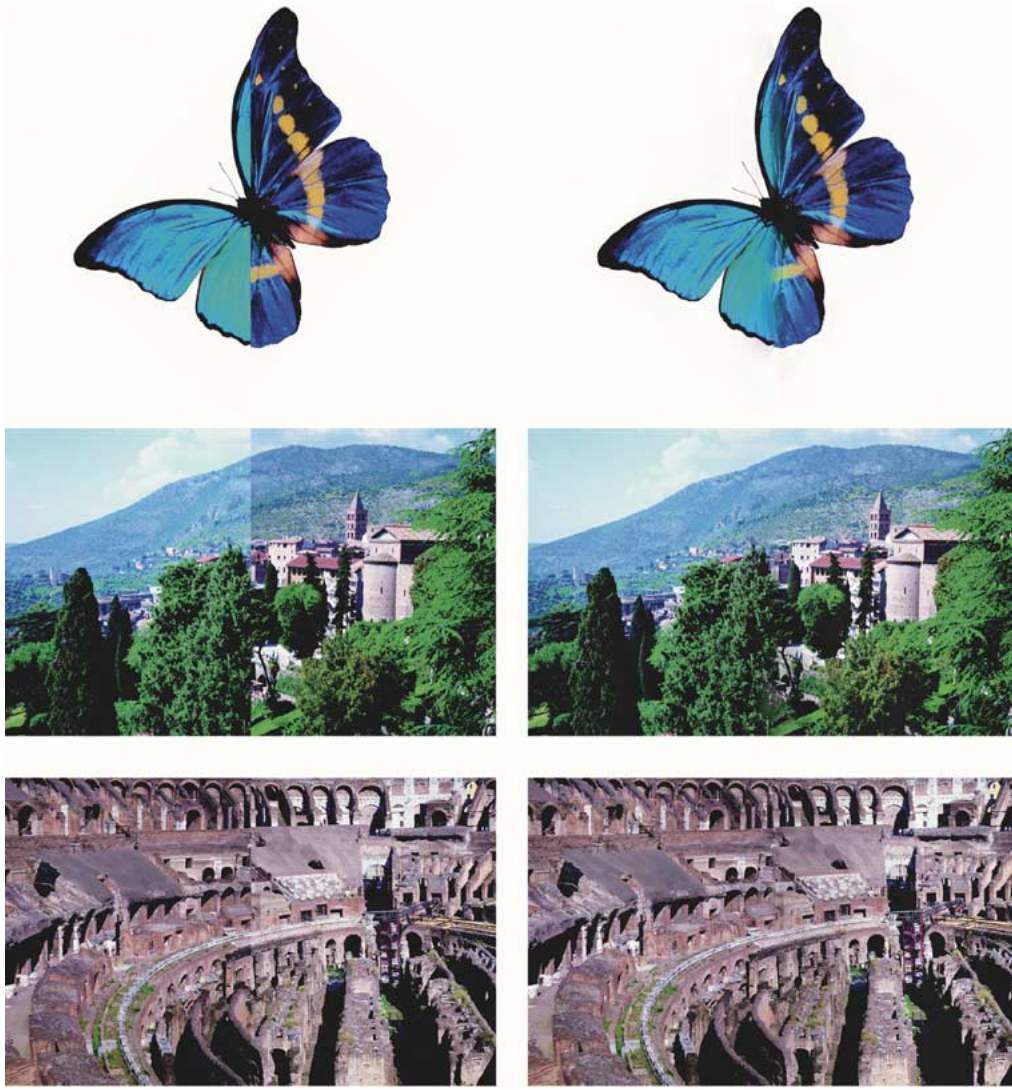


Fig. 3. Left: mosaic images of butterfly, landscape, and arena obtained through direct combination along the mid-line. One can clearly see the discontinuity in intensity along the mid-line. Right: mosaic images obtained using our algorithm where  $\epsilon = 1.0$  and  $\lambda = 0.1, 0.5$  and  $0.5$ , respectively, for butterfly, landscape, and arena. The  $\lambda$ s are given according to the best results of a subjective test. The discontinuity along the mid-line no longer exists.

The minimization of  $V_j^\alpha(f_j; l_j, r_j, L)$  and  $V_j^\beta(f_j; l_j, r_j, L)$  is quadratic programming with linear constraints. *Lagrange multipliers* can be used to obtain optimal solutions by solving a linear system. This alternative minimization method is stable and fast [4]. Once the alpha and beta vectors are obtained, we use (13) to find the corresponding functions  $w_j(x, y)$  and  $p_j(x, y)$ . In our experiments,  $w_j(x, y)$  and  $p_j(x, y)$  converged on average in four iterations. An example demonstrating our blending algorithm on 1-D signals is given in Fig. 1.

## V. EXPERIMENTAL RESULTS

Because human recognition of a mosaic image tends to be subjective, an ideal mosaic image is unknowable. Thus, a quantitative comparison of mosaic images obtained using various methods is impossible. Instead, a simple subjective test was designed in which ten subjects were asked to grade images by their perception of image quality. The to-be-combined images were first shown to each subject, and then a sequence of mosaic images was displayed in random order. Subjects were instructed to

assign a score to each image from five, representing the range from excellent quality, to one, representing poor quality. As shown in Table I, an averaged score was given to each image. In this test,  $\lambda$  has values 0.1, 0.2, 0.5, and 1.0. The finite element basis function is B spline with  $m = 7$  in (13) in our algorithm. In Burt and Adelson's method, we use their suggested weighting function using  $a = 0.375$ . This value is chosen according to the appearance of the resultant images by performing experiments over a range of  $a \in [0.25, 0.5]$ . In Hsu's method, we use eight levels of wavelet decomposition with the 9-7 wavelet [17].

Fig. 2 shows the tower image. Fig. 2(a) and (b) show the original to-be-combined images on the left and right, respectively. In Fig. 2(c), the seam line where the images were combined is shown, and Fig. 2(d) shows the mosaic image with  $\lambda = 1.0$  along the seam line shown in Fig. 2(c). The black region of Fig. 2(c) was taken from Fig. 2(a) and the white region was taken from Fig. 2(b). The images butterfly, landscape, and arena are shown in Fig. 3. The bold numbers given in Table I show that our algorithm, with different values of  $\lambda$ , yields the best perceptual quality for each image. The advantage of the proposed



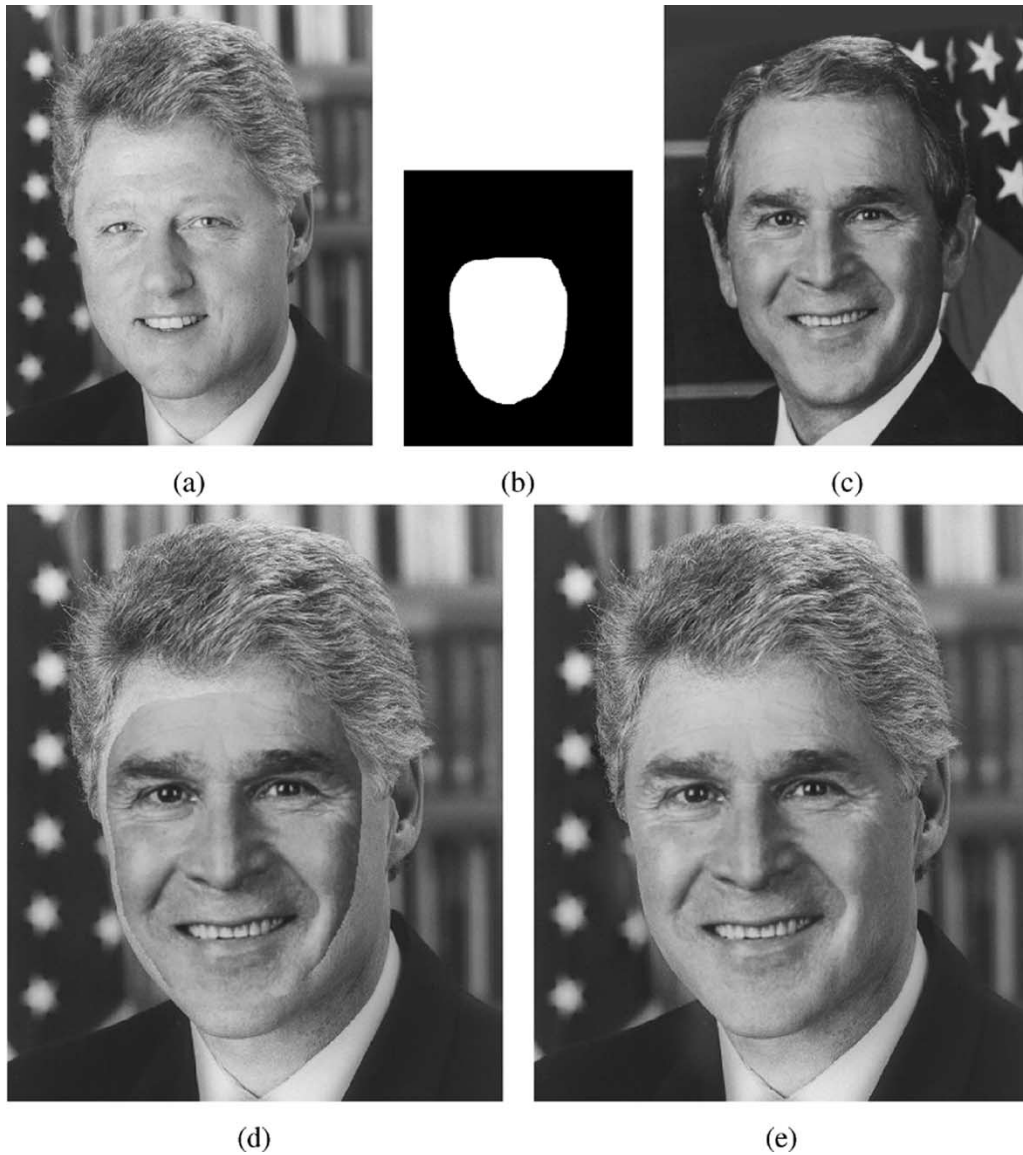


Fig. 4. Image mosaic created with a curved boundary. (a) and (c) are the to-be-combined images. (b) is the curved boundary where (a) and (c) will be combined. (d) is the mosaic image created through direct combination. (e) is our mosaic image created with  $\lambda = 1.0$ .

algorithm lies in the fact that the user can obtain his or her desired image by tuning  $\lambda$ . For example, our mosaic image with  $\lambda = 1.0$  scored best for the tower image, and with  $\lambda = 0.1$ , it scored best for the butterfly image. This flexibility can be useful when a specific level of image quality is required in custom design work.

## VI. BLENDING ALONG A CIRCULAR BOUNDARY

So far, we have only discussed image blending along a straight line. Extending the proposed method to a curved boundary can be achieved by approximating the curve using segments of circles with varying curvature. This idea is illustrated with a simple example of image blending along a circular boundary  $C$ .

An image pixel at  $(x, y)$  can be represented in polar coordinates  $(r, \theta)$  by

$$x = r \cos \theta \text{ and } y = r \sin \theta.$$

Let a circle  $C$  be centered at  $(0,0)$  with radius  $r_0$ . The circle  $C$  will become a line at  $r = r_0$  with the coordinate  $(r, \theta)$ . The images in each subspace can be blended by first representing them as  $(r, \theta)$  and then blending along the seam line with  $r = r_0$  using the previously mentioned algorithm. Finally, the blended image can be transformed back to the Cartesian coordinates  $(x, y)$ . Fig. 4 shows an example in which Fig. 4(e) is the mosaic image obtained by combining Fig. 4(a) and (c) using a circular boundary enclosing the mask shown in Fig. 4(b). The result of direct image combination is shown in Fig. 4(d).

## VII. CONCLUSION

Subjective comparison has shown that the subjective image quality achieved by the proposed method is better than that achieved by other methods. However, it has a longer computing time than other methods (it is about six to eight times slower than the Burt and Adelson method in Matlab [14]). Our model can be extended by adding some higher-order derivative variations.

However, experiments show that the perception difference between the mosaic images obtained by the proposed method and those obtained by a higher-order derivative variation method (with a second derivative variation) is not easily distinguished.

Humans recognize a mosaic image subjectively while computer vision algorithms measure a mosaic image objectively. Therefore, we do not try to find an ideal mosaic image. Instead, we obtain an objectively optimal mosaic image by minimizing the blending energy function that measures two variation terms: image pixel variation and first derivative variation. The two variation terms are balanced by an additional parameter  $\lambda$ . By adjusting this parameter, we can improve the subjective quality of the mosaic image. Although image intensity and its derivatives are important low-level features; however, the two variation terms in our energy function do not include perceptual measurement of the mosaic image. A criterion that measures the appropriate perceptual quality can be incorporated in the model in future works.

## REFERENCES

- [1] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–639, 1992.
- [2] P. J. Burt and E. H. Adelson, "A multiresolution spline with application to image mosaics," *ACM Trans. Graphics*, vol. 2, pp. 217–236, 1983.
- [3] —, "The Laplacian pyramid as a compact image code," *IEEE Trans. Commun.*, vol. COM-31, pp. 532–540, Apr., 1983.
- [4] B. S. Gottfried and J. Weisman, *Introduction to Optimization Theory*. Englewood Cliffs, NJ: Prentice-Hall, 1973.
- [5] S. E. Chen, "QuickTime VR – An image-based approach to virtual environment navigation," in *SIGGRAPH*, 1995, pp. 29–38.
- [6] I. Daubechies, "Ten lectures on wavelets," *SIAM*, 1994.
- [7] P. E. Debevec, C. J. Taylor, J. Malik, G. Levin, G. Borshukov, and Y. Yu, "Image-based modeling and rendering of architecture with interactive photogrammetry and view-dependent texture mapping," in *Proc. IEEE Int. Symp. Circuits and Systems*, vol. 5, 1998, pp. 514–517.
- [8] A. Goshtasby, F. Chen, and B. Barsky, "B-spline curves and surfaces viewed as digital filters," *Comput. Vis. Graphics Image Process.*, vol. 54, pp. 264–275, 1990.
- [9] F. B. Hilderand, *Methods of Applied Mathematics*. Englewood Cliffs, NJ: Prentice-Hall, 1965.
- [10] C. T. Hsu and J. L. Wu, "Multiresolution mosaic," *IEEE Trans. Consumer Electron.*, vol. 42, pp. 981–990, Aug., 1996.
- [11] S. B. Kang, "A survey of image-based rendering techniques," Digital Equipment Corporation, Cambridge Res. Lab., DEC Tech. Rep. 97/4, 1997.
- [12] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *Int. J. Comput. Vis.*, pp. 321–331, 1988.
- [13] W. S. Lee and N. M. Thalmann, "Fast head modeling for animation," *Image Vis. Comput.*, vol. 18, no. 4, pp. 355–364, 2000.
- [14] Matlab, Trademark of MathWorks, Inc., Natick, MA [Online]. Available: <http://www.mathworks.com/>
- [15] J. N. Reddy, *Energy and Variational Methods in Applied Mechanics — With an Introduction to the Finite Element Method*. New York: Wiley, 1984.
- [16] H. Y. Shum and R. Szeliski, "System and experiment paper: Construction of panoramic image mosaics with global and local alignment," *Int. J. Comput. Vis.*, vol. 36, no. 2, pp. 101–130, 2000.
- [17] G. Strang and T. Nguyen, *Wavelets and Filter Banks*. Cambridge, MA: Wellesley-Cambridge, 1996.
- [18] M. S. Su, W. L. Hwang, and K. Y. Cheng, "A variational calculus approach to multiresolution image mosaic," in *Int. Conf. Image Processing*, 2001.
- [19] P. R. Wolf, *Elements of Photogrammetry*, 2nd ed. New York: McGraw-Hill, 1983.



**Ming-Shing Su** received the B.S. and M.S. degrees from the Department of Electrical Engineering and the Ph.D. degree in computer science and information engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1993, 1995, and 2003, respectively.

He is currently with AnCAD, Inc., Taipei, as a Chief Technology Officer. His research interests include scientific visualization, image analysis, and computer graphics.



**Wen-Liang Hwang** received the B.S. degree from the Nuclear Engineering Department, National Tsing-Hua University, Taiwan, R.O.C., the M.S. degree from the Electrical Engineering Department, Polytechnic Institute of New York, Brooklyn, and the Ph.D. degree from the Computer Science Department, New York University, New York, in 1993.

He was a Postdoctoral Researcher in the Mathematics Department, University of California, Irvine, in 1994. He became a member of the Institute of Information Science, Academia Sinica, Taiwan, in January 1995, where he is currently an Associated Research Fellow. His research interests include wavelet analysis, signal and image processing, and multimedia transmission.



**Kuo-Young Cheng** received the B.S. degree in electrical engineering from the National Taiwan University, Taipei, Taiwan, R.O.C., in 1964, the M.S. degree in electrical engineering from North Dakota University, Fargo, in 1967, and the Ph.D. degree in applied mathematics from the State University of New York, Stony Brook, in 1972.

He joined the Institute of Information Science, Taiwan, in 1978, where he has been a Research Fellow since 1979.