A Collaborative Nonlocal-Means Super-resolution Algorithm Using Zernike Moments

Lin Guo\(^1,2\)\(^,*\), Qinghu Chen\(^1\)

\(^1\) School of Electronic and Information, Wuhan University, Wuhan, 430074, China
\(^2\) School of Physics and Electronics, Hubei University, Wuhan, 430062, China

Email: gunglin@gmail.com, cqh@eis.whu.edu.cn

Abstract—Super-resolution (SR) with probabilistic motion estimation is a successful algorithm to circumvent the limitation of motion estimation upon conventional super-resolution methods. However, the algorithm can’t match similar patches with rotation or scale. This paper presents an efficient improved algorithm by introducing Zernike moments as representation of image invariant features into similarity measure. A collaborative strategy is proposed combining the moment based proximity and the bilateral proximity of nonlocal means (NL-means) algorithm for joint determination of weights. For the invariant property of Zernike moments, structure-similar pixels with rotation or scale can also be matched for computation of weights. Furthermore, the collaborative mechanism ensures higher accuracy of weights for a better estimation of each pixel in SR images. Experimental results indicate the proposed method is able to handle general video sequences with superior performance in SR reconstruction to the compared algorithms.

Index Terms—super resolution, Zernike moments, probabilistic motion estimation, nonlocal means, collaborative

I. INTRODUCTION

Super-Resolution (SR) technique is the fusion of a sequence of low-resolution noisy, blurred images to produce a higher resolution image or sequence overcoming the inherent resolution limitation of LR imaging systems. Since Huang and Tsai first proposed the concept of SR in 1984, the SR technique has attracted a lot of attention in the image processing community due to its wide variety of applications in image enhancement, medical imaging, high definition televisions and computer vision. A great deal of literature about SR can be found, and the representatives are referred in [1-4].

SR reconstruction is an ill-posed inverse problem. A widely used model for this problem is described as follows:

\[ y_t = DHM_{t,s} z_s + n. \]  (1)

Where the measurements \(y_t, t = 1, 2, \ldots, T\), are results of different motion, noise, blur, and decimation parameters from an original high resolution reference image \(z_s\). The matrix \(M_{t,s}\) indicates the geometric warp of \(y_t\) relative to the high resolution image \(z_s\). And \(H\) is the blur matrix. Both of them are assumed for simplicity to be linear space and time invariant. Similarly, \(D\) denotes the fixed spatial resolution decimation. Gaussian random noise \(n\) is assumed to be added to the measurements.

To recover \(z_s\) from \(y_t, M_{t,s}\) and \(H\) must be known or can be reliably estimated from inputs. Most of the existing SR methods are roughly based on an estimation of the motion between frames followed by the super-resolution fusion of inputs according to these motion vectors. As it is well know, motion estimation of sub-pixel precise between frames is indispensable and commonly regarded very critical for successful SR reconstruction.

However, it is a challenging task to obtain highly accurate motion estimation with an affordable computation load. In fact, it’s almost impossible to handle actual scenes with complex motion patterns or very low quality. Inaccurately registration often leads to deteriorated reconstruction results even compared to a simple interpolated version. So motion estimation has become the bottleneck for the conventional SR methods to get excellent performance.

In order to overcome the above problem, several recent articles [4-7] attempted to deliver SR methods avoiding explicit motion estimation apart from above conventional methods. The algorithm in [5] relies on extending their previous steerable kernel regression method to multi-frame super-resolution. The approach in [6] is based on the sparse 3D transform-domain collaborative filtering and iterative projection on the observation constrained subspace. The method in [7] develops the notion of probabilistic motion estimation into the classical SR formulation, which is regarded as a generalization of the very successful nonlocal means (NL-means) denoising method [8] to serve the super-resolution task [4]. The main idea of the NL-means is that the pixel is estimated as a weighted average of similar pixels in its nonlocal neighborhood, and the weights are computed according to the similarity between two pixels. It shows simple and robust to noise. However, the similarity measure only in intensity is crude to ensure the accuracy of weights for no any information about the underlying image features are considered. For example, structure-similar patches with rotation or various scales are unable to be matched. As a result, unsuitable weights are calculated and assigned to
pixels, and hence lead to the estimation value to deviate from the true one.

Several recent papers have tried to improve the NL-means algorithm in image denoising. Ref. [9] and [10] renders a similar approach by employing affine gray scale transformations to find patches at the same or different scales. Ref. [11] uses cross-scale (i.e., downsampled) neighborhoods in the NL-means filter [12]. Ref. [13] introduces SIFT as rigid invariant features to compute the similarity between different patches. SIFT features as local descriptors are suitable for image retrieval, affine registration etc., but they are not for denoising and SR. Ref. [14] develops Hu moment as rotationally invariant feature for better capabilities of invariant feature representation because of its orthogonal property. This motivates us to introduce Zernike moments as invariant descriptors of image shape features into the similarity measure to improve the super-resolution results. Here, we need to point out that the intuitive approach through independent interpolation of each frame followed by the Zernike moment images. For the invariant property of Zernike moment, the algorithm is enabled to match more similar patches not only with translation but also with rotation or scale. However, the images of super-resolution generally may contain complex degradation involving higher order moments that are more sensitive to the noise. So Zernike moments of SR images are usually unreliable to be the sole basis for computation of weights. To tackle the problem, a collaborative algorithm is designed combining the Zernike moment based proximity and intensity based bilateral proximity of NL-means algorithm for joint determination of weights.

The algorithm proposed in this paper has the following major features. Firstly, besides retaining the advantage of avoiding explicit motion estimation, our algorithm extends the notion of probabilistic motion estimation in [7] to include not only intensity-similar patches with translation but also structure-similar patches with rotation or scale. Secondly, the collaborative similarity measure strategy balances the influence of gray-level based proximity and invariant moment based proximity. Then weights with higher accuracy are computed for better estimation of a pixel.

The remainder of the paper is as follows. Section II presents the super-resolution framework with probabilistic motion estimation on which our method is based. Section III describes the proposed collaborative super-resolution algorithm using Zernike moments in details. A simplified numerical algorithm in iterative form is given at last in this section. Section IV shows experimental results on several general video sequences, followed by conclusion and discussion in Section V.

II. THE SR FRAMEWORK WITH PROBABILISTIC MOTION ESTIMATION

According the observation model (1), the Maximum-Likelihood (ML) estimation of high resolution image is expressed as

\[
\hat{z}_s = \arg \min_1^T 2 \sum_{r=1}^T \left\| \text{DHM}_{r,s} \hat{z}_s - y_r \right\|^2_{L_2}.
\]

The matrix \( \text{M}_{r,s} \) in classical SR methods denotes a one-to-one mapping between pixels in the s-th and the t-th image. And as such, it introduces sensitivity to errors. According to the idea of probabilistic motion estimation [7], the one-to-one mapping between pixels in classical SR methods is substituted to a probabilistic movement domain. That means every estimated pixel in the reference image with many possible correspondences in all the frames of the sequences (including itself). Each pixel inside the domain is assigned a value of weight to denote the probability of being correct. The movement domain is a spatial and temporal neighborhood centered at the estimated pixel in the reference image with radius \( R \) among all the sequences. For given \( s \) and \( t \), the displacement between the estimated pixel and every pixel inside the domain is written as \( (dx(n), dy(n)) \), \( n = 1, \ldots, N \), \( N, N = (2R+1)^2 \). The location relationship is described by a matrix \( \text{M}_{r,s} \) with size of \( S_1 \times S_2 \times N_1 \times S_2 \times N_2 \) and value of 1 in one position and 0 for others. \( S_1 \) and \( S_2 \) are sampling factors respectively in horizontal and vertical direction. For the pixel whose displacement is indicated by the 1 in \( \text{M}_{r,s} \), the corresponding weight is denoted by \( W_{d,s}^{r,t} \), a diagonal matrix with the same size as \( \text{M}_{r,s} \). Thus, we get the following equation:

\[
\hat{z}_s = \sum_{n=1}^N W_{d,s}^{r,t} \text{M}_{r,s} \hat{z}_s.
\]

According to (2) and (3), the probabilistic ML estimation of the high resolution image is formulated as follows:

\[
\hat{z}_{s, \text{PML}} = \arg \min_1^T 2 \sum_{r=1}^T \left\| \text{DHM}_{r,s} \hat{z}_s - y_r \right\|^2_{L_2}.
\]

where \( W_{d,s}^{r,t} \), with size of \( S_1 \times S_2 \times N_1 \times N_2 \), is the corresponding weight matrix in low resolution space by being downsampled from \( W_{d,s}^{r,t} \).

Since both \( \mathbf{H} \) and \( \mathbf{M} \), are space-invariant, they can be exchanged in position. Thus, defining \( \mathbf{x} = \mathbf{Hx} \), the task of SR is turned to be a two-step process: first estimation of the “blurry” high resolution image \( \hat{x} \) according to (5) and the subsequent acquirement of \( z \) from \( x \) by using existing deblurring algorithms.
\[ \hat{x}_{\text{PML}} = \arg \min \frac{1}{2} \sum_{n_i=1}^{N} \sum_{t=1}^{T} \left\| DM_{n_i} x_t - y_t \right\|_{L_2}^2. \]  

Minimization of (5) leads to a solution represented in pixel-wise as

\[ \hat{x}_{i,j} = \frac{\sum_{(k,l) \in N(i,j)} \sum_{t=1}^{T} W_{n_i,t} \cdot (k,l,i,j) y_t(k,l) + \sum_{(k,l) \in N(i,j)} \sum_{t=1}^{T} W_{n_i,t} \cdot (k,l,i,j)}{\sum_{(k,l) \in N(i,j)} \sum_{t=1}^{T} W_{n_i,t} \cdot (k,l,i,j)}, \]  

where \((i,j)\) is an arbitrary coordinate on high resolution grid. And \((k,l) \in N(i,j)\) denotes the \((k,l)\)-th pixel within the movement domain for pixel \((i,j)\), but is located on low resolution grid. That is, \((k,l)\) s.t. \((S_i k, S_i l) \in N(i,j)\), which ensures that the center pixel of the patch is on the decimation grid. The weight \(W_{n_i,t}\) is computed based on the bilateral proximity strategy according to

\[ W_{n_i,t} \cdot (k,l,i,j) = \exp \left( -\frac{\left\| R_{j}^{*} \cdot DM_{n_i} x_t - R_{j}^{*} y_t \right\|_{2}^2}{\sigma_{g}^2} \right) \times f \left( \sqrt{(dx(n))^2 + (dy(n))^2 + (s-t)^2} \right), \]  

where the function \(f\) may be an arbitrary monotonically non-increasing function, such as Gaussian or box form. The parameter \(\sigma_{g}\) controls the effect of the gray-level difference between two patches. \(R_{j}\) is an operator that extracts a patch of a fixed size centred at the \((i,j)\)-th pixel from an image. The square differences of all the pixels of two patches are accumulated. Both gray-level proximity and geometric proximity are considered for similarity measurement, which helps to enhance the effectiveness of the algorithm and robustness to noise. However, the similar patches in case of rotation and various scales can not be matched in this algorithm because the invariance property of a patch is not taken into account.

### III. The Proposed Method

#### A. Zernike Moment Based Image Representation

Moment based image feature representation has a very wide range of applications in the field of image processing and pattern recognition. Zernike moments are proved to be superior to other moments in noise sensitivity, redundancy and expression efficiency for the property of orthogonality and invariance. Zernike Moments with orthogonal basis functions can be used to represent image features by a set of mutually independent descriptors, with a near zero value of information redundancy [16]

The kernel of Zernike Moments is the set of orthogonal Zernike polynomials defined over a unit disk in the polar coordinate space. The Zernike basis function for order \(n\) and repetition \(m\) is

\[ V_{nm}(x,y) = V_{nm}(r,\theta) = R_{nm}(r) \exp(jm\theta), \]  

where \(n\) is a positive integer or zero, and \(m\) is an integer subject to the following constraints: \(n = |m| \) is even and \(|m| \leq n\). In addition, \(\theta\) and \(r\) is, respectively, the phase in polar coordinate space and the distance from point \((x, y)\) to the origin. And \(j = \sqrt{-1}\).

The radial polynomial \(R_{nm}\) is defined as

\[ R_{nm}(r) = \frac{(-1)^{n-m}}{\sqrt{2}} \frac{(-1)^{n-m}}{2^{n-m} \Gamma(n-m+1)} \int_{0}^{1} \int_{0}^{2\pi} f(x,y) \sqrt{1 - \rho^2}^{n-m} e^{-2\rho} \sin^{2\pi} (\frac{x}{2}) \sin^{2\pi} (\frac{y}{2}) \, dx \, dy, \]  

Given that \(f\) is a complex-valued function on the unit disk, the Zernike moment for \(f\) of order \(n\) and repetition \(m\) is

\[ Z_{nm} = \frac{n+1}{\pi} \int_{-1}^{1} \int_{-1}^{1} f(x,y) V_{nm}^{*}(x,y) \, dx \, dy, \]  

where \(V_{nm}^{*}\) is the complex conjugate of \(V_{nm}\).

When \(f\) is a digital image, the Zernike moment means the projection of image \(f(x,y)\) on above orthogonal bases. Then (10) becomes

\[ Z_{nm} = \frac{n+1}{\pi} \sum_{x,y} f(x,y) V_{nm}^{*}(x,y). \]  

To reckon the Zernike moments for \(f(x,y)\), the image (or a patch) is first mapped to the unit disk of polar coordinates, moving the origin of the unit disc to the centre of the image. In this paper, a square-to-circular mapping transformation [16] is used so that the polynomials \(R_{nm}(r)\) need be computed only once for all pixels mapped to the same circle. Furthermore, fast computation for \(R_{nm}(r)\) in [17] is adopted to speed up the calculation.

The magnitude of Zernike moment is rotation invariant as reflected in the mapping to the unit disc. The scale invariance can be achieved through normalizing the moments by the zero-order geometric moment [18].

In terms of (11), Zernike moments of different orders are calculated with varying \(n\); accordingly, for given \(n\) each moment of order \(n\) is computed with varying \(m\). And moments of different orders correspond to independent characteristics of the image, which constitutes a multi-level representation for describing various shape features of the image. The magnitudes of these moments can be presented as images [15]. Fig. 1 shows an image and its Zernike moment images up to the third order. Fig. 1 (a) is the image “lena”. Fig. (b)-(g) are the moment images of \(Z_{00}\), \(Z_{11}\), \(Z_{20}\), \(Z_{22}\), \(Z_{31}\) and \(Z_{33}\), respectively. It can be seen that the lowest order moment \(Z_{00}\) displays the main content of the image, the same as the average filtering result. And the higher moments deliver more detailed shape characteristics, but are also more sensitive to the noise. So Zernike moments of only up to third order are used in this paper.

#### B. Collaborative SR Algorithm

In this section, Zernike moments are first introduced as invariant features into the similarity measure strategy.
The Zernike moment based similarity measurement for SR reconstruction is proposed in this paper as

\[
W_z = \exp \left\{ \frac{\sum \| R_{i,j} M_{i,j} Z_s(x_i) - R_{i,j} Z_s(y_i) \|^2}{\sigma_z^2} \right\},
\]

where \(Z_s(x_i)\) and \(Z_s(y_i)\) mean the k-th moment image for high resolution image \(x_i\) and \(y_i\). \(Y_t\) is the interpolation result of the measurement \(y_i\), \(\sigma_z\) is the controlling parameter similar with \(\sigma_g\). They can be roughly decided by estimation of noise from inputs. So when the weights are calculated according to (12), a Zernike moment based SR algorithm can be executed through (6).

However, in the practical SR task, images may undergo complex motion and degradation, so their Zernike moments are usually not accurate enough to be the only basis for computation of weights, especially for the higher order moments that are more sensitive to the noise. Moreover, very inadequate information is rendered for the weak textures in the Zernike moment images, while the gray-levels of images contain the rich underlying details, and they are also more loyal to the original images. Thus, a collaborative algorithm is developed to combine Zernike moment features and the gray-level based proximity in (7) into the computation of weights.

The computation formula of the final weight for a searching pixel is designed as

\[
w = \frac{1}{2} w_g \cdot w_z + \frac{1}{4} (w_g + w_z),
\]

where \(w_g\) is an abbreviation for \(W_{n,i,j}\) in (7). \(w_g\) and \(w_z\) is calculated, respectively, according to (7) and (12). In the collaborative algorithm, the final weight is jointly determined by the moment images and bilateral proximity of gray-levels, which leads to a more accuracy calculation of weights in our method. This can be seen in Fig.2, where a comparison of the weight distribution with different algorithm is given. The 1st column in Fig. 2 is the original image without noise, the 2nd to the 4th columns are the weights distribution of the NLM algorithm, Zernike moment based algorithm and the collaborative algorithm. Differences between the right three columns in Fig. 2 show that the NLM algorithm strictly matches similar pixels only with translation. The moment based algorithm finds more pixels with similarity both in translation and rotation, but little difference is reflected for pixels with different similarity. The collaborative algorithm also can match similar pixels both with translation and rotation, but greater weights are assigned to pixels that are more similar. So the weights are more precise in comparison in the collaborative algorithm.
In addition, in order to improve the super-resolution algorithm, the following points are considered. Firstly, in our method whenever a pixel is SR estimated, the old value is replaced for the new one, which is closer to the true value than the old value. Hence, more accurate information is provided for computation of weights. That practically helps not only acquire more precise weight but also speed up the SR process. Secondly, in order to reinforce the reliability of Zernike moments in presence of noise, \( Y_t \) in (12) is processed by a NLM denoising before computation of moments. Finally, the image of moment \( Z_{\text{do}} \) is not necessary any more. Hence, we set \( Z = \{Z_k \mid k = 1, 2, 3, 4, 5 \} = \{Z_{11}, Z_{20}, Z_{22}, Z_{31}, Z_{32} \} \) in our experiments to decrease the computation.

Summarily, the proposed collaborative SR reconstruction for video sequences can be expressed in a simplified numerical algorithm. The iterative form of this numerical algorithm is represented as:

\[
X_{s+1}^t(i, j) = \frac{\sum_{(k,l)} w_{k,l}^t(i,j,k,l) y(k,l)}{\sum_{(k,l)} w_{k,l}^t(i,j,k,l)}
\]

(14)

\[
w_{k,l}^t(i,j,k,l) = \exp \left\{ -\frac{||1-S_{i,k} - j-S_{j,l}||}{\sigma_1} \right\}
\]

(15)

\[
w_{k,l}^{s+1}(i,j,k,l) = \exp \left\{ -\frac{\sum_{z=1}^{5} ||R_{i,j}Z_s(V_z) - R_{i+k,j+l}Z_s(V_z)||}{\sigma_2^2} \right\}
\]

(16)

where \( w_{k,l}^t \) in (14) is computed according to (13).

Especially, when \( n = 0, \{X_{s+1}^t\}_{t=1}^{T} \) in (15) are obtained by the bilinear interpolation of \( \{y_{i,j}\}_{i,j} \). And \( \{V_{s+1}^t\}_{t=1}^{T} \) in (16) are the denoised results of \( \{X_{s+1}^t\}_{t=1}^{T} \) by NLM algorithm [8]. Otherwise, when \( n > 0, V_{s+1}^t = X_{s+1}^t \). Both \( V_{s}^t \) and \( X_{s}^t \) are updated after each iteration.

IV. EXPERIMENTS

In this section, the performance of the proposed algorithm is validated. The obtained results of processing several real video sequences with a general motion pattern are presented. The comparison is provided with several methods: the bilinear and bicubic interpolation of single image as well as the state-of-the-art SR algorithm [4]. The results are evaluated from both the visual effects and the objective quality measure (PSNR = \( 10 \log_{10} \frac{255^2}{||X-X_s||_2^2} \) dB, where \( N \) is the number of pixels in the true image \( X \) or the constructed image \( X_s \)).

All the tests in this section were prepared in the following degradation: The input sequences were blurred using a \( 3 \times 3 \) uniform kernel, down-sampled with a factor of 1:3 in each axis, and then added by additive white zero-mean Gaussian noise with \( \text{std} = 3 \). All images were in the input range [0,255]. In processing all the sequences, all 30 frames took part in the iterative reconstruction of each image.

First, sequences “Miss America”, “Trevor” and “Foreman” are tested for evaluation of PSNR. Table 1 gives the average PSNR for each of the three test sequences, where two iterations were run for our method and the compared method in [4] with no additional deblurring followed. Fig. 3 illustrates the PSNR values frame by frame for every test sequence.

In the experiments, the parameter \( \sigma_1 \) and \( \sigma_2 \) was set manually to 2.5 and 2.4 for all the test sequences. The size of the patch used for calculating the weights \( w_{B} \) and \( w_{Z2} \) was equally selected as \( 7 \times 7 \) pixels (high resolution grid) for all sequences. The searching range for movement domain is \( 7 \times 7 \) pixels (high resolution grid) for sequence “Miss America” and “Trevor”, and \( 19 \times 19 \) pixels (high resolution grid) for sequence “Foreman”, which has greater displacements between frames.

The table shows that both the method in [4] and our method can handle sequences of arbitrary motion patterns and achieve effects of the state-of-the-art compared to the single image interpolation. And the proposed algorithm yields superior performance to all compared methods in PSNR.

Then, an example on sequence “Mobile” with 30 frames of \( 330 \times 264 \) pixels is to reveal the visual effects for the proposed method and the compared methods. Fig. 4 represents the selected two frames from the reconstructed results for bilinear interpolation, the method in [4] and the proposed method. The details are unfolded by their enlarged parts in Fig. 5. It can be seen that some numbers in the images, such as “15”, “16”, “18”, “19”, are obviously clearer in our results than others, due to that the invariant moments of Zernike are introduced, which improves the NLM algorithm.

V. CONCLUSION AND DISCUSSION

This paper proposes a SR algorithm using Zernike moments. The algorithm is based on the framework of probabilistic motion estimation and needs no explicit motion estimation. A collaborative similarity measure strategy is developed in our algorithm to combine the Zernike moment based proximity and the bilateral proximity of NL-means algorithm. As representation of image local invariant features, Zernike moments enable the algorithm to match more similar pixels not only with translation but also with rotation or scale. The collaborative mechanism ensures more suitable weights assigned to similar patches for better estimation.
TABLE I. MEAN-PSNR RESULTS FOR THREE TEST SEQUENCES WITH DIFFERENT METHODS.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>Bilinear</th>
<th>Bicubic</th>
<th>Protter et al. [4]</th>
<th>Our method (1st Iteration)</th>
<th>Our method (2nd Iteration)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miss America</td>
<td>33.91</td>
<td>34.31</td>
<td>35.31</td>
<td>35.65</td>
<td>35.87</td>
</tr>
<tr>
<td>Trevor</td>
<td>29.42</td>
<td>29.79</td>
<td>30.39</td>
<td>30.58</td>
<td>30.76</td>
</tr>
<tr>
<td>Foreman</td>
<td>28.38</td>
<td>28.89</td>
<td>29.57</td>
<td>29.65</td>
<td>29.84</td>
</tr>
</tbody>
</table>

Figure 3. The PSNR values of each frame reconstructed by different methods

(a) Miss America

(b) Trevor

(c) Foreman
Figure 4. Results for the 9th (the top row) and 14th (the bottom row) frames from “Mobile” sequence. From left to right: bilinear interpolation; the method in [4]; the proposed method.

Figure 5. Enlarged parts of images in Fig. 4.

of images. Experimental results demonstrate that the proposed algorithm is able to process real video sequences with general motion patterns with improvements both in PSNR and the visual effects compared to the state-of-the-art algorithm.

Several aspects of the proposed method may be further studied to improve the algorithm. Firstly, parameters \( \sigma_B \) and \( \sigma_Z \) reflecting the size of the noise are constant during the iterations. Since noise becomes smaller in the later reconstruction, it should be reasonable to decrease parameters \( \sigma_B \) and \( \sigma_Z \) appropriately with the increased iteration. Secondly, Zernike moment used in our method may be replaced by the Pseudo-Zernike moment, which is proved able to represent image details better with lower orders. And fast computation has been proposed to directly calculate arbitrary Pseudo-Zernike moment with high order without lower order moments first computed. Thus, to use Pseudo-Zernike moment instead may help to improve the performance without computation increased. Finally, several important parameters in our algorithm are manually selected in the experiments, such as \( \sigma_B \), \( \sigma_Z \), and the searching size for the motion domain. An adaptive selection strategy may be developed to improve the algorithm in the future work.

REFERENCES


Lin Guo was born in Hubei, China, in 1978. She is currently pursuing the Ph.D. degree in Wuhan University, Wuhan, China. Her research interests include super-resolution reconstruction, image processing and computer vision.

Qinghu Chen was born in Hubei, China, in 1957. He currently is a professor in School of Electronic Information at Wuhan University, Wuhan, China. His main research interests include image processing and intelligent recognition.