Normalization-Cooperated Gradient Feature Extraction for Handwritten Character Recognition

Cheng-Lin Liu, *Senior Member, IEEE*

Abstract—The gradient direction histogram feature has shown superior performance in character recognition. To alleviate the effect of stroke direction distortion caused by shape normalization and provide higher recognition accuracies, we propose a new feature extraction approach, called normalization-cooperated gradient feature (NCGF) extraction, which maps the gradient direction elements of original image to direction planes without generating the normalized image and can be combined with various normalization methods. Experiments on handwritten Japanese and Chinese character databases show that, compared to normalization-based gradient feature, the NCGF reduces the recognition error rate by factors ranging from 8.63 percent to 14.97 percent with high confidence of significance when combined with pseudo-two-dimensional normalization.

Index Terms—Character recognition, feature extraction, normalization-cooperated gradient feature (NCGF).

---◆---

1 INTRODUCTION

OPTICAL character recognition (OCR) techniques have found numerous applications, including business card recognition, bank-check recognition, tax form processing, mail address reading, etc. Higher recognition accuracy is always desired in applications, especially for unconstrained handwriting and degraded document images. This paper proposes a new feature extraction approach to improve the accuracy of handwritten character recognition.

Previous studies have shown the superiority of gradient direction histogram feature (gradient direction feature or gradient feature in brief) in character recognition [1]. As the contour direction feature (chaincode feature) does, the gradient feature grasps the local distribution of stroke direction. Both the chaincode feature and the gradient feature can be computed in moderate CPU resource and result in moderate dimensionality. The gradient feature is more robust against contour noise and applies to gray-scale images as well as binary images. A recent comparison also shows that the gradient feature performs comparably well or better than the popular Gabor filter feature [2]. The gradient direction feature is also demonstrated effective in face detection [3] and has the potential of applying to face recognition and other image recognition problems.

The recognition performance also depends on the approach of character normalization. Line-density-based nonlinear normalization [4], [5] has been commonly used in handwritten Chinese and Japanese character recognition. Curve-fitting-based normalization [6] and pseudo-two-dimensional (P2D) normalization techniques [7] have been proposed to improve either the computation efficiency or the recognition accuracy. With nonlinear and P2D normalization, however, the recognition performance of chaincode and gradient features may be traded off by stroke shape distortion. To alleviate the effect of distortion, a normalization-cooperated feature extraction (NCFE) approach [8], and its improved version called continuous NCFE [9], compute the chaincode direction feature of original image incorporating coordinate normalization. However, it is not straightforward to extend NCFE to gradient feature.

- *The author is with the National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, PO Box 2728, Beijing 100080, P.R. China. E-mail: liucl@nlpr.ia.ac.cn.*

To provide higher recognition accuracies in handwritten character recognition, we propose a new feature extraction approach, called normalization-cooperated gradient feature (NCGF) extraction. We view both the original image and the normalized image as functional in continuous 2D space and associate them by coordinate normalization (coordinate mapping). The normalized gradient can then be computed from the original image without generating the normalized image explicitly. Replacing the direction of normalized gradient with that in the original image while maintaining the magnitude of the normalized gradient, the resulting gradient feature thus records the undistorted stroke directions and keeps stroke-width invariance.

We combined gradient feature extraction with various coordinate normalization methods in experiments of handwritten character recognition on two large databases. The recognition results show that the proposed NCGF outperforms both the normalization-based gradient feature (NBGF) and the normalization-cooperated chaincode feature (NCCF) in most cases. When combined with P2D normalization methods, the NCGF performs significantly better than the NBGF and provides the overall best performance.

Since the feature extraction process is combined with character normalization, we first review the normalization methods in Section 2. The proposed NCGF extraction approach is then described in Section 3. Experimental results are presented in Section 4 and concluding remarks are offered in Section 5.

2 REVIEW OF CHARACTER NORMALIZATION

A character image $f(i,j)$ has digital (binary or gray-scale) pixel values sampled from a continuous image plane $f(x,y)$. We view a digital image as a functional in continuous 2D space with each pixel defined by a unit square (Fig. 1a). Specifically, a pixel $f(i,j)$ is defined in domain $x \in [i, i+1)$, $y \in [j, j+1)$, and $f(x,y)$ is constant in this domain. By this definition, we can treat $f(x,y)$ and $f(i,j)$ equivalently.

An original image $f(x,y)$ is transformed to a normalized image $f'(x',y')$ of standard size. Normalization is realized by mapping the pixel coordinates from the original image to a normalized image plane:

$$x' = u(x,y), \; y' = v(x,y), \text{ and } f'(u,v) = f(x,y). \qquad (1)$$

In implementation, each unit square in the original image, $U_{ij} = ((i,j),(i+1,j),(i,j+1),(i+1,j+1))$, is mapped to a quadrilateral Q_{ij} with four vertices $(u(i,j),v(i,j))$, $(u(i+1,j),v(i+1,j))$, $(u(i,j+1),v(i,j+1))$, and $(u(i+1,j+1),v(i+1,j+1))$. Accordingly, the square grid of original image is transformed to a net overlapping on a square grid of normalized plane (Fig. 1b). The normalized image is obtained by assigning $f(i,j)$ to the pixels of normalized image overlapping with the quadrilateral, each pixel multiplied with the overlapping area [7].

The shape of normalized image thus depends on the coordinate mapping functions $u(x,y)$ and $v(x,y)$. For facilitating implementation and restricting the distortion of normalized shape, the coordinate functions are often simplified to one-dimensional (1D): $x' = u(x)$, $y' = v(y)$. Pseudo-two-dimensional (P2D) normalization methods form 2D coordinate functions by combining row-wise or column-wise 1D coordinate functions. An early and computationally expensive method, computes row-wise/column-wise coordinate functions after line density smoothing [10]. Our recent study shows that combining only three 1D coordinate functions performs sufficiently [7]. Specifically, for computing $u(x,y)$, three 1D coordinate functions and their weight functions, $u^i(x)$ and $w^i(y)$, $i = 1, 2, 3$, are computed from three soft strips of the original image. The 1D functions are combined by

$$u(x,y) = \begin{cases} w^1(y)u^1(x) + w^2(y)u^2(x), & y < y_c, \\ w^3(y)u^3(x) + w^2(y)u^2(x), & y \geq y_c. \end{cases} \qquad (2)$$
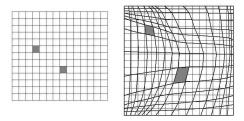
Fig. 1. Grids for (a) original image 13×14 and (b) normalized image 16×16 pixels and the coordinate-normalized grid in thick lines.

(x_c, y_c) is the centroid of the original image. $v(x, y)$ is computed similarly.

We consider five 1D normalization methods: linear normalization (LN), line-density-based nonlinear normalization (NLN) [4], moment normalization (MN) which aligns image centroid and rescales according to second-order moments, bimoment normalization (BMN), and modified centroid-boundary alignment (MCBA). The coordinate mapping of moment normalization is linear, and those of NLN, BMN, and MCBA are nonlinear. The P2D extension of NLN is called line density projection interpolation (LDPI). Similarly, the MN, BMN, MCBA methods are extended to P2D versions, called P2DMN, P2DBMN, and P2DCBA, respectively. The details of all these normalization methods can be found in [7]. The methods except NLN and LDPI are applicable to both binary and gray-scale images. Some examples of character normalization are shown in Fig. 2.

The normalization-cooperated gradient feature (NCGF), proposed as follows, can be combined with various normalization methods, which are uniformly formulated by coordinate functions as in (1).

3 NORMALIZATION-COOPERATED GRADIENT FEATURE EXTRACTION

Gradient direction feature is flexible for application to machine-print/handwriting, binary/gray scale, and low-resolution images [11]. A gradient direction decomposition technique, originally proposed in online character recognition [12], was shown to give high recognition performance [1]. Specifying a number of standard directions (very often, eight chaincode directions), a gradient vector of arbitrary direction is decomposed into two components coincident with the two neighboring standard directions (Fig. 3). The components are assigned to the corresponding direction planes. On decomposing all the gradient vectors, a number of feature values is extracted from each direction plane by Gaussian filtering (blurring) and down-sampling [1].

Conventionally, the gradient is computed on each pixel of the normalized image. Since generating the normalized image consumes computing resource, and pixel discretization/interpolation may lose precision, it is desirable to compute the normalized
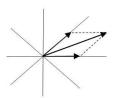

Fig. 3. Direction decomposition of gradient vector.

gradient directly from the original image. This falls in the normalization-cooperated feature extraction (NCFE) strategy.

3.1 Gradient Computation and Assignment

Recall the relation (1) between original image and normalized image. It is essential for normalization that the local order of pixels remains, i.e., the coordinate mapping functions satisfy $\frac{\partial u}{\partial x} > 0$ and $\frac{\partial v}{\partial y} > 0$. For P2D normalization methods, it is reasonable to assume $\frac{\partial u}{\partial x} \gg \frac{\partial u}{\partial y}$ and $\frac{\partial v}{\partial y} \gg \frac{\partial v}{\partial x}$. This can be verified from (2): $u^1(x)$, $u^2(x)$, and $u^3(x)$ do not differ significantly in practice, $w^1(y) + w^2(y)$ and $w^3(y) + w^2(y)$ are constant (see [7]). Consider the inverse functions, we similarly assume $\frac{\partial x}{\partial u} \gg \frac{\partial x}{\partial v}$ and $\frac{\partial y}{\partial v} \gg \frac{\partial y}{\partial u}$. In 1D normalization, it is true that $\frac{\partial u}{\partial y} = 0$, $\frac{\partial v}{\partial x} = 0$, $\frac{\partial x}{\partial v} = 0$, and $\frac{\partial y}{\partial u} = 0$.

The gradient of normalized image $f'(u, v)$, mapped from the original image $f(x, y)$, is defined by $\mathbf{g}' = (\frac{\partial f'}{\partial u}, \frac{\partial f'}{\partial v})$. Using the chain rule and incorporating the above assumptions,

$$\frac{\partial f'}{\partial u} = \frac{\partial f'}{\partial x} \cdot \frac{\partial x}{\partial u} + \frac{\partial f'}{\partial y} \cdot \frac{\partial y}{\partial u}$$
$$\simeq \frac{\partial f'}{\partial x} \cdot \frac{\partial x}{\partial u} = \frac{\partial f}{\partial x} / \frac{\partial u}{\partial x}, \tag{3}$$

where $\frac{\partial f'}{\partial x} = \frac{\partial f}{\partial x}$ is from $f'(u, v) = f(x, y)$. Similarly,

$$\frac{\partial f'}{\partial v} \simeq \frac{\partial f}{\partial y} / \frac{\partial v}{\partial y}. \tag{4}$$

It is seen from (3) and (4) that the normalized gradient $\mathbf{g}' = (\frac{\partial f'}{\partial u}, \frac{\partial f'}{\partial v})$ is computed from the orginal gradient $\mathbf{g} = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$ incorporating coordinate mapping. In implementation, $\mathbf{g}$ is computed for each pixel in the original image, converted to $\mathbf{g}'$ and assigned to direction planes.

Before introducing the numerical computation of $\frac{\partial f'}{\partial u}$ and $\frac{\partial f'}{\partial v}$, we explain how to map the gradient $\mathbf{g}'$ to direction planes. As in Fig. 3, $\mathbf{g}'$ is decomposed into two components in two directions with lengths l_1 and l_2. Since $\mathbf{g}'$ is computed for each pixel (i, j) of the original image, which is mapped to a quadrilateral Q_{ij} in the normalized image, it represents the gradient of all points in Q_{ij}. Accordingly, the two gradient components are assigned to the pixels of the two corresponding direction planes that overlap with Q_{ij}: Each pixel assigned a value equal to the overlapping area times the gradient component length. In doing this, we obtain normalized
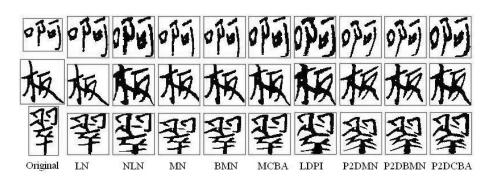

Fig. 2. Character images: original (leftmost) and normalized by nine methods.
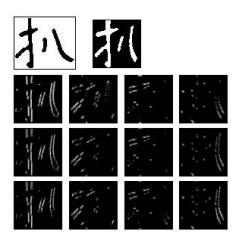
Fig. 4. Gradient orientation planes of NBGF (second row), nNCGF (third row), and NCGF (last row). The first row shows the original image and the gray-scale normalized image.

Fig. 5. Sample images of CASIA database.

gradient direction feature, without generating the normalized image explicitly. We call this *normalized* normalization-cooperated gradient feature (nNCGF), in which the gradient direction is normalized.

For extracting features from the original gradient direction, the gradient of original image at pixel (i, j), $\mathbf{g} = (\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$, is directly assigned to direction planes. As assigning $\mathbf{g}'$, $\mathbf{g}$ is decomposed into two components of two standard directions. In one direction plane, the pixels overlapping with the quadrilateral Q_{ij} are assigned values equal to the overlapping area times the gradient component length times a variable ratio. The variable ratio is based on the following reason. While the normalized gradient $\mathbf{g}'$ characterizes the edgeness of normalized image and the product of its magnitude with the quadrilateral area is approximately stroke-width invariant, the product of quadrilateral area with the magnitude of $\mathbf{g}$ is not. To maintain stroke-width invariance, the multiplier is set equal to the ratio $\|\mathbf{g}'\|/\|\mathbf{g}\|$. Thus, we take the direction of the original gradient and the magnitude of the normalized gradient in NCGF.

3.2 Numerical Implementation

For numerically computing the gradient in original and normalized images, we use the Sobel operator. At pixel (i, j) of original image,

$$
\begin{aligned}
\frac{\partial f}{\partial x} = & [f(i+1, j-1) + 2f(i+1, j) + f(i+1, j+1) \\
& - f(i-1, j-1) - 2f(i-1, j) - f(i-1, j+1)]/8, \\
\frac{\partial f}{\partial y} = & [f(i-1, j+1) + 2f(i, j+1) + f(i+1, j+1) \\
& - f(i-1, j-1) - 2f(i, j-1) - f(i+1, j-1)]/8.
\end{aligned}
\tag{5}
$$

The partial derivatives of coordinate mapping functions are replaced by numerical differences

$$
\begin{aligned}
\frac{\partial u}{\partial x} &= [u(i+1, j) - u(i-1, j)]/2, \\
\frac{\partial v}{\partial y} &= [v(i, j+1) - v(i, j-1)]/2.
\end{aligned}
\tag{6}
$$

The normalized gradient $\mathbf{g}' = (\frac{\partial f'}{\partial u}, \frac{\partial f'}{\partial v})$ is then computed according to (3) and (4).

For either NCGF or nNCGF, the gradient or normalized gradient is computed for each pixel (i, j) of the original image and its components are assigned to the pixels in direction planes overlapping with the quadrilateral Q_{ij}. Feature values are then extracted from direction planes by Gaussian blurring and down-sampling.

Assigning values to the pixels overlapping with a quadrilateral can be implemented by decomposing the quadrilateral into trapezoids falling in unit squares of pixels (quadrilateral decomposition [7]). This procedure is somewhat complicated. Actually, the precision of assigning gradient magnitude to overlapping pixels is not necessarily high because the direction planes are finally blurred. To simplify, we just assign the gradient magnitude (component length or length times a ratio) times the quadrilateral area to one single pixel (unit square) that the point $(u(i, j), v(i.j))$ falls in. And further, the quadrilateral area is approximated by $(u(i+1, j) - u(i, j)) \times (v(i, j+1) - v(i, j))$ (this is exact in the case of 1D coordinate normalization). We will show in experiments that this simplified *point-assignment* procedure performs comparably well with quadrilateral decomposition.

Fig. 4 shows the gradient orientation planes merged from eight direction planes, obtained by three approaches: normalization-based gradient feature (NBGF), nNCGF, and NCGF. We can see evident difference in the left-most column (corresponding to edges of vertical strokes) and the right-most column (edges of left-diagonal strokes), due to the orientation change of the right-most stroke.

4   EXPERIMENTAL RESULTS

To evaluate the recognition performance of the proposed NCGF, we experimented on two databases of handwritten character images. The ETL9B database[1] has been widely tested in previous works (e.g., [13], [14]). It contains binary images of 3,036 character classes (71 hiragana characters and 2,965 Kanji characters), 200 samples per class. The CASIA database[2] contains binary images of 3,755 characters (in level-1 set of GB2312-80), 300 samples per class. Some sample images of CASIA database are shown in Fig. 5.

For the ETL9B database, we used the first 20 and the last 20 samples of each class for testing, and the remaining 160 samples of each class for training classifiers. For the CASIA database, we used the first 250 samples of each class for training, and the remaining 50 samples of each class for testing. Recognition error rates are reported on the test sets, with 121,440 samples and 187,750 samples for ETL9B and CASIA, respectively. We did not partition the datasets in multiple ways (e.g., cross validation) because the number of test samples is large enough.

From each character image, we extracted 8-direction gradient features by three approaches: NCGF, normalized NCGF (nNCGF), and normalization-based gradient feature (NBGF, based on gray-scale normalized image). The size of normalized image and direction planes was set to 64×64 pixels. From each of eight direction planes, we extracted 8×8 feature values. As result, the dimensionality of feature vector is 512. We also evaluated the

---

1. Collected by the Electro-Technical Laboratory, which was reorganized into the National Institute of Advanced Industrial Science and Technology (AIST), Japan.

2. Collected by the Institute of Automation, Chinese Academy of Sciences (CASIA).

TABLE 1
Error Rates and Feature Extraction Time on Hiragana Characters

Normal.	Error rate (%)				CPU time (ms)			
	qradri.	point	NBGF	NCCF	quadri.	point	NBGF	NCCF
LN	2.46	**2.44**	2.55	2.53	1.32	0.78	0.90	0.50
NLN	**1.83**	1.85	2.12	2.04	1.54	0.89	1.03	0.60
MN	1.67	**1.58**	1.70	1.85	1.26	0.77	0.91	0.49
BMN	1.59	**1.56**	1.68	1.75	1.29	0.79	0.95	0.53
MCBA	**1.65**	1.72	1.98	1.85	1.36	0.83	0.98	0.52
LDPI	**0.88**	0.97	1.01	1.15	2.22	0.95	1.96	0.65
P2DMN	1.50	**1.31**	1.94	1.65	2.01	0.90	1.79	0.58
P2DBMN	1.20	**1.14**	1.64	1.22	2.04	0.91	1.81	0.63
P2DCBA	**1.04**	1.10	1.35	1.32	2.09	0.96	1.71	0.66

"Quadri." and "point" refer to quadrilateral decomposition and point-assignment for NCGF.

TABLE 2
Error Rates and Ratio of Error Reduction on ETL9B Database

Normal.	Error rate (%)				NCGF vs NBGF		NCGF vs NCCF	
	NCGF	nNCGF	NBGF	NCCF	reduce(%)	z	reduce(%)	z
LN	2.27	**2.26**	**2.26**	2.51	-0.44	-0.17	9.56	3.87
NLN	**0.80**	0.91	0.91	0.89	12.09	2.95	10.11	2.43
MN	0.88	**0.87**	0.90	0.94	2.22	0.53	6.38	1.56
BMN	**0.85**	0.86	0.86	0.93	1.16	0.27	8.60	2.10
MCBA	**0.87**	0.90	0.91	0.91	4.40	1.05	4.40	1.05
LDPI	**0.64**	0.74	0.74	0.73	13.51	2.98	12.33	2.69
P2DMN	**0.71**	0.74	0.79	0.75	10.13	2.28	5.33	1.16
P2DBMN	**0.68**	0.73	0.76	0.74	10.53	2.33	8.11	1.76
P2DCBA	**0.73**	0.75	0.80	0.79	8.75	1.98	7.59	1.70

chaincode direction feature extracted by continuous NCFE, referred to as normalization-cooperated chaincode feature (NCCF).

The normalization methods experimented are those mentioned in Section 2, namely, linear normalization (LN), nonlinear normalization (NLN), moment normalization (MN), bimoment normalization (BMN), modified centroid-boundary alignment (MCBA), line density projection interpolation (LDPI), P2D moment normalization (P2DMN), P2D bimoment normalization (P2DBMN), and P2D MCBA (P2DCBA).

For classification, the dimensionality of feature vector is first reduced from 512 to 160 by Fisher linear discriminant analysis [15], with the transformation matrix estimated on the training data set. The reduced feature vector is classified using the modified quadratic discriminant function (MQDF2) [16], which has been proven very effective in handwritten Chinese character recognition. The MQDF2 assumes that the probability density of each class is the combination of a multivariate Gaussian in principal linear subspace and a spherical Gaussian in the minor subspace. We use 40 principal eigenvectors for each class, make the minor eigenvalue class-independent and optimize it via holdout cross-validation on the training data set.

Before evaluating the performance of large character set recognition, we compared two implementation techniques of NCGF: quadrilateral decomposition and point-assignment, in respect of recognition accuracy and computation time on the 71 hiragana characters in ETL9B database. Here each sample is represented by 288 feature values (36 values from each of eight direction planes of size 48×48). MQDF2 with 40 principal eigenvectors per class is applied to the 288-dimensional vector for classification. Because the total number of hiragana samples (14,200) is small, we estimated the error rates by 5-fold cross-validation, i.e. the data of each class is partitioned into five parts, each part is tested in rotation with the classifier trained on the remaining parts. The error rates and the CPU times (on Pentium 4-3 GHz) of feature extraction per sample are shown in Table 1. We do not compare nNCGF here because its computation time is almost the same as

NCGF. We can see that the two implementation techniques of NCGF yield very similar accuracies (for eight of the nine normalization methods, the difference of error rate is less than 0.1 percent). The computation of point-assignment is much faster than quadrilateral decomposition, and is intermediate between NBGF and NCCF.

In 3,036-class recognition of ETL9B database and 3,755-class recognition of CASIA database, we compare the error rates of four feature extraction techniques: NCGF, nNCGF, NBGF, and NCCF, and evaluate the improvement of NCGF versus NBGF and NCGF versus NCCF in respect of the ratio of error reduction and the level of significance.

Denote the error rates of two recognizers by p_1 and p_2 ($p_1 < p_2$), the error reduction ratio is $r = (p_2 - p_1)/p_2$. The level of significance is measured by z-test [17]. Take the null hypothesis that two error rates p_1 and p_2 do not differ significantly, the distribution of variable $p_2 - p_1$ can be approximated as normal with zero mean and variance $\sigma^2 = 2p(1-p)/n$, where p is the average of p_1 and p_2 and n is the number of test samples. Accordingly, $z = (p_2 - p_1)/\sigma$ has a standard normal distribution. When the value of z satisfies $|z| > 1.96$, the null hypothesis can be rejected (i.e., the two error rates are judged to be different) with confidence higher than 0.95. Hence, we use the value of z as a measure of significance.

The error rates and measures of improvement on ETL9B database are shown in Table 2 and those on CASIA database are shown in Table 3.

We have some common observations from the results of both ETL9B and CASIA. Comparing the feature extraction techniques using the same normalization method, the lowest error rate is mostly (in 13 of 18 cases) given by NCGF. With 1D normalization, the error rates of nNCGF and NBGF are comparable and mostly (in six of 10 cases) lower than that of NCCF. With P2D normalization, the error rates of nNCGF and NCCF are comparable and mostly (in seven of eight cases) lower than that of NBGF. The error rate of NCGF is evidently lower than that of NBGF when combined with one-dimensional NLN or P2D normalization methods.

The superiority of NBGF to NCCF with 1D normalization is due to the inherent advantage of gradient feature over chaincode

TABLE 3
Error Rates and Ratio of Error Reduction on CASIA Database

Normal.	Error rate (%)				NCGF vs NBGF		NCGF vs NCCF	
	NCGF	nNCGF	NBGF	NCCF	reduced(%)	z	reduced(%)	z
LN	3.92	3.89	**3.80**	4.18	-3.16	-1.91	6.22	4.04
NLN	**1.88**	2.04	2.07	2.00	9.18	4.19	6.00	2.67
MN	2.23	2.21	**2.20**	2.39	-1.36	-0.63	6.69	3.27
BMN	2.05	2.05	**2.03**	2.19	-0.99	-0.43	6.39	2.98
MCBA	**2.05**	2.11	2.10	2.08	2.38	1.07	1.44	0.64
LDPI	**1.59**	1.75	1.87	1.72	14.97	6.58	7.56	3.13
P2DMN	**1.91**	1.93	2.18	2.01	12.39	5.85	4.98	2.21
P2DBMN	**1.80**	1.85	1.99	1.88	9.55	4.27	4.26	1.82
P2DCBA	**1.80**	1.83	1.97	1.91	8.63	3.83	5.76	2.50

feature. With P2D normalization, the distortion of stroke shapes is more influential to NBGF than to NCCF, thus obscures the advantage of NBGF. The nNCGF is also less sensitive to shape distortion because the normalized gradient is computed from the original image. The NCGF performs even better because it takes the direction of the original gradient. The improvement level of NCGF versus NBGF depends on the normalization methods, which have variable smoothness of coordinate functions and extent of shape distortion. The line-density-based NLN and P2D normalization methods yield large shape distortion.

With NLN and P2D normalization methods, the error reduction ratio of NCGF versus NBGF ranges from 8.63 percent to 14.97 percent with high level of significance ($z > 1.96$). With the other normalization methods, the NCGF and the NBGF perform comparably ($|z| < 1.96$).

With 1D normalization, the NCGF outperforms the NCCF significantly ($z > 1.96$) in seven of 10 cases; while with P2D normalization, the NCGF outperforms the NCCF significantly only in four of eight cases. This is because both the NCGF and the NCCF take advantage of the shape restoration capability of P2D normalization and overcome the effect of shape distortion.

Though the NCGF does not outperform the NBGF and the NCCF in all cases, it is overall the best choice because it outperforms the NBGF significantly in 10 of 18 cases, outperforms the NCCF significantly in 11 of 18 cases and is never beat significantly. Most important, when combined with good normalization methods (e.g., P2D methods), the NCGF yields highest recognition accuracies. In addition, gradient features (including the NCGF) apply to both binary and gray-scale images, while chaincode features (including the NCCF) apply to binary images only.

5 CONCLUSION

As the storage capacity and computing power of personal computers constantly increase, it is a trend to process gray-scale images instead of binary images in document analysis, particularly at the shape identification level, for achieving higher accuracy. The gradient direction feature, performing superiorly on both binary and gray-scale images, is among the top choices for feature extraction of character recognition. While nonlinear and pseudo-2D (P2D) shape normalization methods are efficient to improve handwritten character recognition, the accuracy is traded off by distorted stroke direction in normalization. The proposed normalization-cooperated gradient feature (NCGF) extraction approach, can alleviate the effect of this distortion and significantly improve the recognition accuracy, especially with P2D normalization methods.

ACKNOWLEDGMENTS

REFERENCES

[1]   C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten Digit Recognition: Benchmarking of State-of-the-Art Techniques," *Pattern Recognition,* vol. 36, no. 10, pp. 2271-2285, 2003.
[2]   C.-L. Liu, M. Koga, and H. Fujisawa, "Gabor Feature Extraction for Character Recognition: Comparison with Gradient Feature," *Proc. Eighth Int'l Conf. Document Analysis and Recognition,* pp. 121-125, 2005.
[3]   L.-L. Huang, A. Shimizu, Y. Hagihara, and H. Kobatake, "Gradient Feature Extraction for Classification-Based Face Detection," *Pattern Recognition,* vol. 36, no. 11, pp. 2501-2511, 2003.
[4]   J. Tsukumo and H. Tanaka, "Classification of Handprinted Chinese Characters Using Non-Linear Normalization and Correlation Methods," *Proc. Nineth Int'l Conf. Pattern Recognition,* pp. 168-171, 1988.
[5]   H. Yamada, K. Yamamoto, and T. Saito, "A Nonlinear Normalization Method for Hanprinted Kanji Character Recognition—Line Density Equalization," *Pattern Recognition,* vol. 23, no. 9, pp. 1023-1029, 1990.
[6]   C.-L. Liu, H. Sako, and H. Fujisawa, "Handwritten Chinese Character Recognition: Alternatives to Nonlinear Normalization," *Proc. Seventh Int'l Conf. Document Analysis and Recognition,* pp. 524-528, 2003.
[7]   C.-L. Liu and K. Marukawa, "Pseudo Two-Dimensional Shape Normalization Methods for Handwritten Chinese Character Recognition," *Pattern Recognition,* vol. 38, no. 12, pp. 2242-2255, 2005.
[8]   M. Hamanaka, K. Yamada, and J. Tsukumo, "Normalization-Cooperated Feature Extraction Method for Handprinted Kanji Character Recognition," *Proc. Third Int'l Workshop Frontiers of Handwriting Recognition,* pp. 343-348, 1993.
[9]   C.-L. Liu, K. Nakashima, H. Sako, and H. Fujisawa, "Handwritten Digit Recognition: Investigation of Normalization and Feature Extraction Techniques," *Pattern Recognition,* vol. 37, no. 2, pp. 265-279, 2004.
[10]   T. Horiuchi, R. Haruki, H. Yamada, and K. Yamamoto, "Two-Dimensional Extension of Nonlinear Normalization Method Using Line Density for Character Recognition," *Proc. Fourth Int'l Conf. Document Analysis and Recognition,* pp. 511-514, 1997.
[11]   G. Srikantan, S.W. Lam, and S.N. Srihari, "Gradient-Based Contour Encoder for Character Recognition," *Pattern Recognition,* vol. 29, no. 7, pp. 1147-1160, 1996.
[12]   A. Kawamura, K. Yura, T. Hayama, Y. Hidai, T. Minamikawa, A. Tanaka, and S. Masuda, "On-Line Recognition of Freely Handwritten Japanese Characters Using Directional Feature Densities," *Proc. 11th Int'l Conf. Pattern Recognition,* vol. 2, pp. 183-186, 1992.
[13]   F. Kimura, T. Wakabayashi, S. Tsuruoka, and Y. Miyake, "Improvement of Handwritten Japanese Character Recognition Using Weighted Direction Code Histogram," *Pattern Recognition,* vol. 30, no. 8, pp. 1329-1337, 1997.
[14]   N. Kato, M. Suzuki, S. Omachi, H. Aso, and Y. Nemoto, "A Handwritten Character Recognition System Using Directional Element Feature and Asymmetric Mahalanobis Distance," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 21, no. 3, pp. 258-262, Mar. 1999.
[15]   K. Fukunaga, *Introduction to Statistical Pattern Recognition,* second ed. Academic Press, 1990.
[16]   F. Kimura, K. Takashina, S. Tsuruoka, and Y. Miyake, "Modified Quadratic Discriminant Functions and the Application to Chinese Character Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 9, no. 1, pp. 149-153, 1987.
[17]   T.G. Dietterich, "Approximate Statistical Test for Comparing Supervised Classification Learning Algorithms," *Neural Computation,* vol. 7, no. 10, pp. 1895-1924, 1998.