Two-Dimensional Polar Harmonic Transforms for Invariant Image Representation

Pew-Thian Yap, Xudong Jiang, *Senior Member*, *IEEE*, and Alex Chichung Kot, *Fellow*, *IEEE*

Abstract—This paper introduces a set of 2D transforms, based on a set of orthogonal projection bases, to generate a set of features which are invariant to rotation. We call these transforms *Polar Harmonic Transforms* (PHTs). Unlike the well-known Zernike and pseudo-Zernike moments, the kernel computation of PHTs is extremely simple and has no numerical stability issue whatsoever. This implies that PHTs encompass the orthogonality and invariance advantages of Zernike and pseudo-Zernike moments, but are free from their inherent limitations. This also means that PHTs are well suited for application where maximal discriminant information is needed. Furthermore, PHTs make available a large set of features for further feature selection in the process of seeking for the best discriminative or representative features for a particular application.

Index Terms—Polar harmonic transforms, harmonic kernels, rotation invariance, Zernike moments, pseudo-Zernike moments, orthogonal moments.

1 Introduction

Humans, often unfailingly and effortlessly, are able to recognize a wide variety of objects irrespective of their rotations. Man-made perception systems are often designed to mimic this basic capability. One way of achieving this is by training a certain set of classifiers to recognize the objects by working in parallel to cater for a finite set of different angles. Another way is to directly devise a set of features which are invariant to the image orientation. In this paper, we follow the latter approach.

A number of rotation-invariant features have been proposed in the literature. Among them are the popular Zernike moments (ZMs) [1] and also the often mentioned together pseudo-Zernike moments (PSMs) [2]. These moments have been very successfully applied in a variety of contexts including, but not limited to, modeling of corneal surface [3], watermarking [4], face recognition [5], character recognition [6], multispectral texture classification [7], and edge detection [8]. Despite their usefulness, these two sets of moments are often faced with the problem of computation difficulty, especially when the high-order moments are concerned. The computational problem of these moments is inherently related to the fact that many factorial terms are involved in the process of calculating the moment kernels.

Other options of rotation-invariant features include rotational moments (RMs) [2] and complex moments (CMs) [9], [10]. However, these moments are not orthogonal, as are ZMs and PZMs. Nonorthgonality implies lack of information compactness in each of the computed moments. On the other hand, orthogonality of the kernels means that an image is projected onto a set of pairwise orthogonal axes, and the classifier can hence be relatively simple. The kernels of ZMs and PZMs can be shown to be the outcome of Gram Schmidt orthogonalization on the CM kernels [9]. Refer to [11], [12], [13], [14] for a comprehensive survey on moments.

In this paper, we introduce a set of transforms, called *Polar Harmonic Transforms*[1] (PHTs), which can be used to generate rotation-invariant features. The computation of the PHT kernels is significantly simpler compared with that of ZMs and PZMs, and can hence be performed at a much higher speed. With PHTs, there is also no numerical instability issue, as with ZMs and PZMs, which often limits their practical usefulness. A large part of the computation of the PHT kernels can be precomputed and stored. In the end, for each pixel, as little as three multiplications, one addition operation, and one cosine and/or sine evaluation are needed to obtain the final kernel value. In this paper, three different transforms will be introduced, namely, Polar Complex Exponential Transform (PCET), Polar Cosine Transform (PCT), and Polar Sine Transform (PST). We have grouped them under the name Polar Harmonic Transforms as the kernels of these transforms are harmonic in nature, that is, they are basic waves.

Another additional advantage of our framework is that it offers the possibility of constructing limitless number of features. In order to be able to recognize objects effectively, apart from the issue of constructing invariant object descriptors, one also has the problem of generating a sufficiently large number of features, which is often required by many recognition tasks. The more classes one has to discriminate, the more features may be necessary. Therefore, a mechanism is needed, capable of generating a large number of invariant features, which do not necessarily need to have physical or geometric meaning. The transform that

• *The authors are with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore. E-mail: ptyap@med.unc.edu, {exdjiang, eackot}@ntu.edu.sg.*

Manuscript received 6 Aug. 2008; revised 1 Mar. 2009; accepted 5 May 2009; published online 15 May 2009.
Recommended for acceptance by J. Matas.
For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number TPAMI-2008-08-0469.
Digital Object Identifier no. 10.1109/TPAMI.2009.119.

1. We have refrained from using the term Polar Harmonic Moments considering the fact that moments are essentially formulated with monomials/polynomials as their basis.

we propose in the next section has exactly this property. The ease of computation of the PHT kernels causes the evaluation of the higher order PHT coefficients to be more feasible compared to ZMs and PZMs, and hence, a large pool of features can be made available for later feature selection or discriminant analysis.

To the best of our knowledge, the work most similar to ours is Ren et al.'s radial-harmonic-Fourier moments [15] (RHFMs), in which trigonometric functions are also used to form the radial kernels. While, on the surface, their moments might be very similar to the transforms, PCT and PST in particular, that we are going to propose, it is not difficult to observe that our formulation is fundamentally different from theirs in the following ways: 1) Ren et al.'s formulation does not have a DC term, the existence of which might be useful for some applications; PCET and PCT have the DC term, 2) their formulation is problematic at $r = 0$ since one needs to compute $1/\sqrt{r}$, and 3) PCT and PST have more variations in terms of frequency components since the kernels of PCT and PST have a base frequency of π instead of 2π as in Ren et al.'s formulation.

In Section 2, we will define the Polar Harmonic Transforms, properties of which will be further discussed in Section 3. We will then, in Section 4, describe a few experiments which we have performed to gauge the performance. Section 5 concludes this paper.

2 Polar Harmonic Transforms

The Polar Complex Exponential Transform of order n with repetition l, $|n| = |l| = 0, 1, \ldots, \infty$, is defined as

$$M_{nl} = \frac{1}{\pi} \int_0^{2\pi} \int_0^1 [H_{nl}(r,\theta)]^* f(r,\theta) r dr d\theta, \tag{1}$$

where $[\cdot]^*$ denotes the complex conjugate and the basis $H_{nl}(r,\theta)$ can be decomposed into radial and circular components:

$$H_{nl}(r,\theta) = R_n(r) e^{il\theta}, \tag{2}$$

with the radial kernel being a complex exponential in the radial direction:

$$R_n(r) = e^{i2\pi n r^2} \tag{3}$$

and satisfying orthogonality condition:

$$\int_0^1 R_n(r)[R_{n'}(r)]^* r dr = \frac{1}{2}\delta_{nn'}, \tag{4}$$

and also:

$$\int_0^{2\pi} \int_0^1 H_{nl}(r,\theta)[H_{n'l'}(r,\theta)]^* r dr d\theta = \pi \delta_{nn'} \delta_{ll'}. \tag{5}$$

Some visual illustrations of the kernels are given in Fig. 1. The factor $1/\pi$ of (1) is due to the norm of the kernels and can be discarded if the kernels are orthonormalized, i.e., by letting:

$$\tilde{H}_{nl}(r,\theta) = \frac{1}{\sqrt{\pi}} H_{nl}(r,\theta), \tag{6}$$

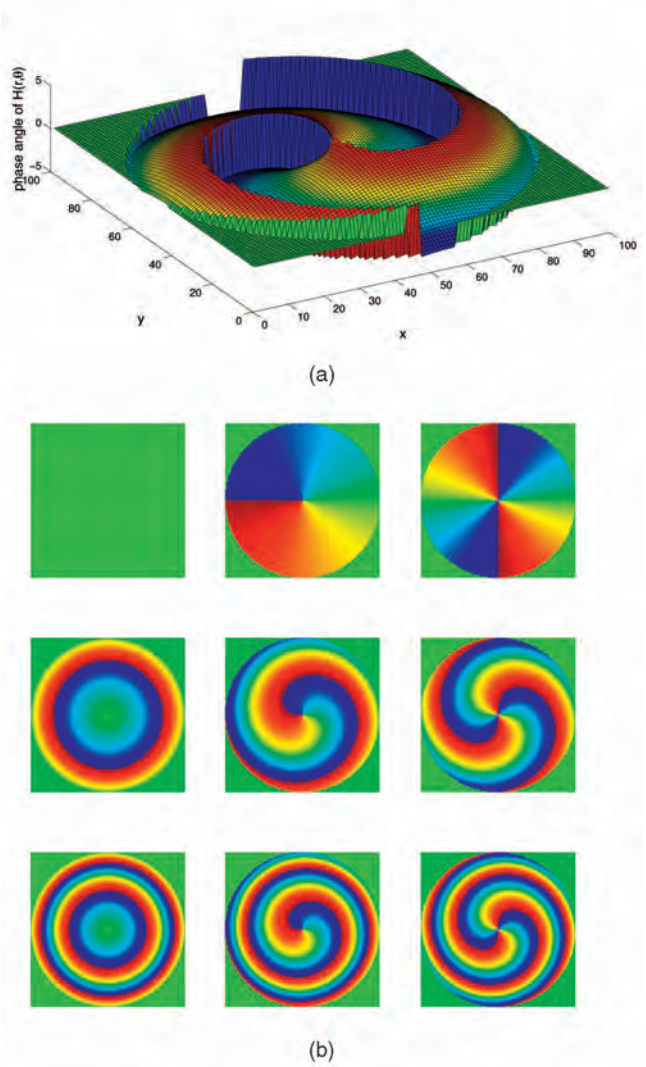

Fig. 1. Some examples of the PCET kernels. (a) 3D view of the phase angle of H_{23}. (b) 2D views of different H_{nl}s: the different rows show the views for $n = 0, 1, 2$ and the columns $l = 0, 1, 2$.

and hence:

$$\int_0^1 \tilde{H}_{nl}(r,\theta)\left[\tilde{H}_{n'l'}(r,\theta)\right]^* r dr d\theta = \delta_{nn'}\delta_{ll'}. \tag{7}$$

The PCET coefficients $\{M_{nl}\}$ can be shown to be bounded via Bessel's inequality:

$$\frac{1}{\pi}\sum_{\forall n,l} M_{nl}^2 \le \int_0^{2\pi}\int_0^1 [f(r,\theta)]^2 r dr d\theta. \tag{8}$$

Similar to the form of PCET, we can define another two sets of harmonic transforms, i.e., Polar Cosine Transform:

$$M_{nl}^{\mathrm{C}} = \Omega_n \int_0^{2\pi}\int_0^1 \left[H_{nl}^{\mathrm{C}}(r,\theta)\right]^* f(r,\theta) r dr d\theta \quad n, |l| = 0, 1, \ldots, \infty \tag{9}$$

and Polar Sine Transform:

$$M_{nl}^{\mathrm{S}} = \Omega_n \int_0^{2\pi} \int_0^1 \left[H_{nl}^{\mathrm{S}}(r,\theta)\right]^* f(r,\theta) r dr d\theta \quad n = 1,\ldots,\infty, |l| = 0,1,\ldots,\infty \tag{10}$$

with

$$H_{nl}^{\mathrm{C}}(r,\theta) = R_n^{\mathrm{C}}(r)e^{il\theta} = \cos(\pi n r^2)e^{il\theta}, \quad H_{nl}^{\mathrm{S}}(r,\theta) = R_n^{\mathrm{S}}(r)e^{il\theta} = \sin(\pi n r^2)e^{il\theta}, \tag{11}$$

and

$$\Omega_n = \begin{cases} \frac{1}{\pi}, & n = 0, \\ \frac{2}{\pi}, & n \neq 0. \end{cases} \tag{12}$$

The polar portions of the above transforms have very similar definitions with Fourier Cosine and Sine Transforms on a finite interval, which cater for even and odd functions, respectively. They are, hence, also very similar to the definition of Discrete Cosine Transform (DCT) and Discrete Sine Transform (DST). However, it should be pointed out that the PCT/PST is fundamentally different from 2D-DCT/2D-DST since the latter is defined in the Cartesian coordinates and the former in the polar coordinates. There is also no equivalent rotation invariance for 2D-DCT/2D-DST as compared to PCT/PST. The distinction between PCT/PST and PCET is that the former uses only cosine/sine functions, while the latter uses both cosines and sines (in the form of complex exponentials). The other difference that should be noted is that the range of angles covered by PCT/PST in its radial kernel, i.e., $[0,\pi]$, is half of that of PCET, i.e., $[0,2\pi]$. In fact, the kernels of PCET and PCT/PST are all *harmonic* in the sense that they are all basic waves. However, we have named the former PCET and the latter PCT/PST for differentiation. PCET, PCT, and PST, though similar in form, capture different image information, as will be evident from the experimental results that will be shown in a later section. In the following sections of the paper, we will mainly focus our discussions on PCET with an occasional foray into PCT/PST only when necessary. This is to avoid unnecessary repetition, as PCT/PST essentially has many properties identical to PCET.

Bhatia and Wolf [16] have shown that a polynomial[2] that is invariant in form for any rotation of axes about the origin must be of the form:

$$V(r\cos\theta, r\sin\theta) = R_n(r)\exp(jm\theta), \tag{13}$$

where $R_n(r)$ is a radial polynomial in r of degree n. There are a few sets of moments with their kernels complying with this form, namely, Zernike moments, with their radial kernels defined as [1]:

$$R_{nl}(r) = \sum_{s=0}^{\frac{(n-|l|)}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+|l|}{2}-s\right)!\left(\frac{n-|l|}{2}-s\right)!} r^{n-2s}. \tag{14}$$

Pseudo-Zernike moments [2], a variation of Zernike moments, with their radial kernels defined as:

$$R_{nl}(r) = \sum_{s=0}^{(n-|l|)} (-1)^s \frac{(2n+1-s)!}{s!(n+|l|-s)!(n-|l|+1-s)!} r^{n-s}, \tag{15}$$

orthogonal Fourier-Mellin Moments [17]:

$$R_n(r) = Q_n(r) = \sum_{s=0}^{n} \alpha_{ns} r^s, \tag{16}$$

$$\alpha_{ns} = (-1)^{n+s} \frac{(n+s+1)!}{(n-s)!s!(s+1)!}, \tag{17}$$

Fourier-Mellin descriptors [18]:

$$R_s(r) = r^{s-1}, \tag{18}$$

rotational moments[3] [2]:

$$R_n(r) = r^n, \tag{19}$$

and radial-harmonic-Fourier Moments [15]:

$$R_n(r) = \begin{cases} \frac{1}{\sqrt{r}}, & n = 0, \\ \sqrt{\frac{2}{r}}\cos(\pi n r), & n \quad \text{even}, \\ \sqrt{\frac{2}{r}}\sin(\pi(n+1)r), & n \quad \text{odd}. \end{cases} \tag{20}$$

Among the moments listed, Zernike moments, pseudo-Zernike moments, Orthogonal Fourier-Mellin moments (OFMMs), radial-harmonic-Fourier moments (RHFMs), and also Polar Harmonic Transform are orthogonal in the sense that their kernels satisfy orthogonality conditions similar to (4). Fourier-Mellin descriptors (FMDs) and rotational moments, on the other hand, are nonorthogonal and do not satisfy the orthogonality condition. Features based on orthogonal kernels are more efficient in terms of information compactness since the image is projected onto a set of axes which are pairwise orthogonal, and hence, the overlapping of information is minimal. Note also that the kernel computation of ZM, PZM, and OFMM involves computation of a number of factorial terms, which inevitably cause the numerical stability of these moments. Recurrence equations [19] are often used in place of these direct formulas, but although recurrence formulas do help to push the maximal order of computable moment higher, they eventually will suffer from numerical instability. See [19], [20] for some ZM-computation-related issues. A summary of comparison of the various moment kernels and that of PHTs is given in Table 1.

3 Properties of Polar Harmonic Transforms

3.1 Computation Complexity of Polar Harmonic Transform

The computation of the kernels of PHTs is significantly easier when compared to that of ZMs, PZMs, and OFMMs.

2. Complex exponential, cosine, and sine functions used to define PCET, PCT, and PST are technically not polynomials, but they can be expressed in the form of power series which are essentially polynomials of infinite degrees.

3. We have left out Complex Moments [9], [10] because they are very similar to rotational moments.

TABLE 1
Comparison of Different Moments and Transforms—Orthogonality and Numerical Stability

MOMENTS	RADIAL KERNEL	ORTHOGONALITY	NUM. STAB.
ZMs [1]	$R_{nl}(r) = \sum_{s=0}^{(n-\lvert l\rvert)/2}(-1)^s \frac{(n-s)!}{s!((n+\lvert l\rvert)/2-s)!((n-\lvert l\rvert)/2-s)!} r^{n-2s}$	✓	×
PZMs [2]	$R_{nl}(r) = \sum_{s=0}^{(n-\lvert l\rvert)}(-1)^s \frac{(2n+1-s)!}{s!(n+\lvert l\rvert-s)!(n-\lvert l\rvert+1-s)!} r^{n-s}$	✓	×
OFMMs [17]	$Q_n(r) = \sum_{s=0}^{n} \alpha_{ns} r^s$, $\alpha_{ns} = (-1)^{n+s} \frac{(n+s+1)!}{(n-s)!s!(s+1)!}$	✓	×
RHFMs [15]	$R_n = \begin{cases} \frac{1}{\sqrt{r}} & n = 0 \\ \sqrt{\frac{2}{r}}\cos(\pi n r) & n \text{ even} \\ \sqrt{\frac{2}{r}}\sin(\pi(n+1)r) & n \text{ odd} \end{cases}$	✓	×
PCET	$R_n(r) = e^{i2\pi n r^2}$	✓	✓
PCT	$R_n^{\mathrm{C}}(r) = \cos(\pi n r^2)$	✓	✓
PST	$R_n^{\mathrm{S}}(r) = \sin(\pi n r^2)$	✓	✓
FMDs [18]	$R_s(r) = r^{s-1}$	×	✓
RMs [2]	$R_n(r) = r^n$	×	✓

Based on (2) and (3), we can rewrite the definition of the PCET kernels in the following form:

$$H_{nl}(r,\theta) = e^{i2\pi n r^2} e^{il\theta} = e^{i(2\pi n r^2 + l\theta)}. \quad (21)$$

We can see that in this form the computation is not complex at all: The transformed image coordinates (r,θ) need to be evaluated only once and then stored in the form of (r^2,θ), which is just an extra step from the former. The rest is just multiplications of these stored values with n and l and followed finally by an evaluation of the complex exponential function which can be written in the form of trigonometric functions, i.e., $e^{iz} = \cos(z) + i\sin(z)$. Hence, for each pixel, only three multiplications (assuming the value of 2π is precalculated and stored), one addition, and one cosine and sine (complex exponential) evaluation are needed to obtain the final kernel value. This implies PHTs have the added advantage of offering the possibility of constructing thousands of features. Indeed, in order to be able to recognize objects, apart from the issue of invariant object descriptors, one also has the problem of generating a large number of features that are necessary. The more classes one has to discriminate, the more features may be necessary for recognition tasks. Therefore, a mechanism is needed, capable of generating a large number of invariant features which do not necessarily have physical or geometric meaning.

3.2 Storage

The ease of computation of PHT kernels also makes it possible to reduce the storage space required to store the kernels. For the case of ZMs, if N terms of ZMs are needed, we need to typically store N different ZM kernels if the repetitive computation of the computationally expensive ZM kernels is to be avoided. But, on the other hand, we see that for PHTs, there is no need to precalculate and store the kernels since the computation needed to arrive at the kernels is very small. Based on our discussion in Section 3.1, we merely need to store the image-mapped coordinates (r^2,θ), and kernels of any orders can be computed without much hassle thereon. This means N times more economical in terms of storage space.

3.3 Information Extraction and the Number of Zeros

The radial kernels of ZMs and PZMs are polynomials and the number of zeros corresponds to the capability of the polynomials to describe high-spatial-frequency components of an image. It is easy to see that the radial kernel of ZM has $(n-m)/2$ duplicated roots in the interior of the interval $0 \le r \le 1$, apart from the trivial case at $r = 0$. On the other hand, for the radial kernel of PCET, writing it in the form of:

$$R_n(r) = \cos(2\pi n r^2) + i\sin(2\pi n r^2), \quad (22)$$

we can observe that the real and imaginary parts have $2n$ and $2n+1$ (inclusive of $r = 0$) zeros, respectively. Hence, for one to have the same number n_0 of zeros, the degree of the ZM radial polynomial has to be $2n_0 + m$, much higher than that of PCET, which is around $n_0/2$.

3.4 Suppression Problem

The zeros of the radial kernels of ZMs are located in the region of large radial distance r from the origin. On the other hand, the zeros of the kernels of PCET are distributed nearly uniformly over the interval $0 \le r \le 1$. Abu-Mostafa and Psaltis [9] in their paper term the nonuniform distribution of zeros of complex moment kernels as the *suppression problem*. We use the same term here to describe a similar problem faced by ZMs. Suppression problem causes unnecessary emphasis on certain part of the image and negligence on the rest. This is illustrated in Fig. 2.

3.5 Rotation Invariance

The coefficients of PHTs have the inherent property of rotation invariance. If image $g_x(x,y)$ is rotated through a

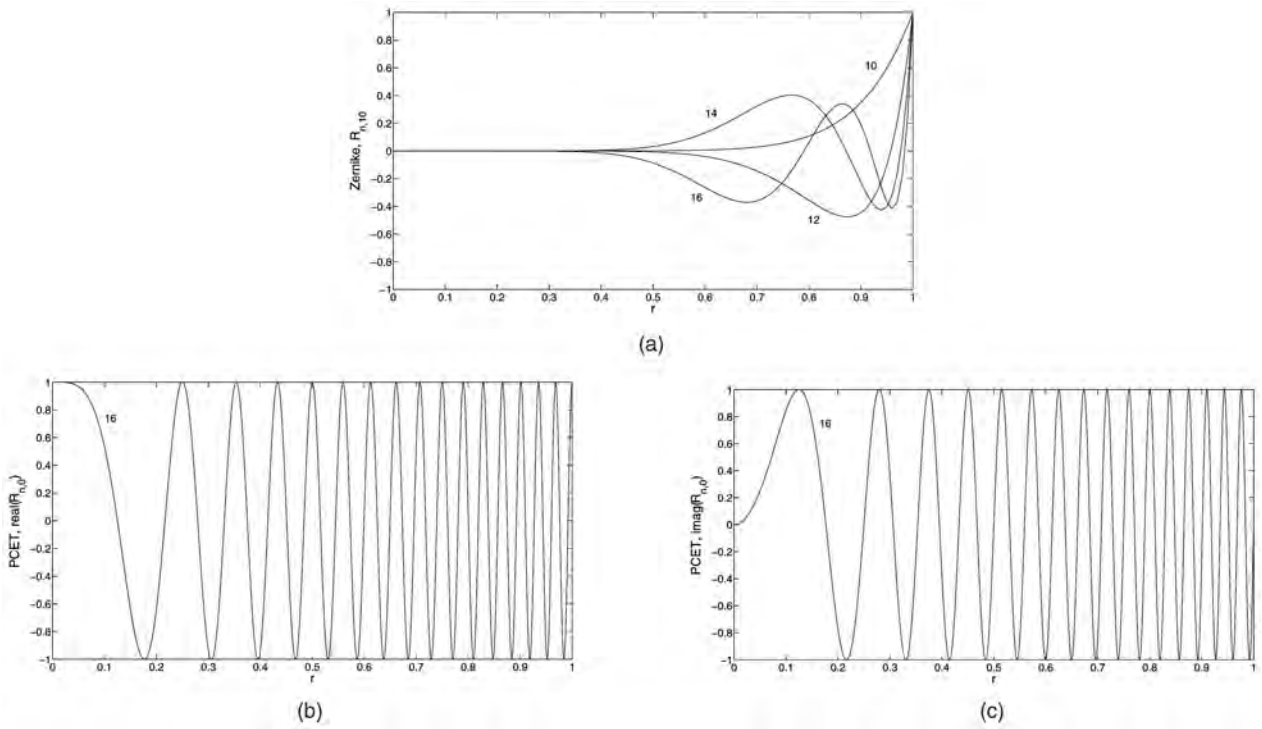

Fig. 2. Information suppression. The kernels of ZMs (a) put too much emphasis on certain portions of the signal. The same problem does not happen to the kernels of PCET. (b) Real and (c) imaginary components of the PCET kernel. The numbers alongside the curves indicate the orders n of the radial components of the kernels.

clockwise angle ϕ to become $g_2(x, y)$, more formally if $g_2(x_2, y_2) = g_1(x_1, y_1)$, where:

$$\begin{bmatrix} x_1 \\ y_1 \end{bmatrix} = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \end{bmatrix}, \tag{23}$$

then the new set of PCET coefficients, namely, $\{M'_{nl}\}$, is related to the old set $\{M_{nl}\}$ by the relation:

$$M'_{nl} = M_{nl} e^{-il\phi}. \tag{24}$$

This equation is the basis for derivation of invariants from the PCET coefficients. If we cancel out the exponential factor in (24), we get *absolute* invariants. For instance, $|M'_{nl}| = |M_{nl}|$ or $M'_{nl}[M'_{nl}]^* = M_{nl}[M_{nl}]^*$. To avoid discarding too much information from the set of PCET coefficients, invariants can be obtained by multiplication of appropriate coefficient terms [21]. Specifically, let $N \geq 1$ and let k_i and n_i be nonnegative integers and l_i be integer, where $i = 1, \ldots, N$, such that:

$$\sum_{i=1}^{N} k_i l_i = 0. \tag{25}$$

Then, any product

$$I = \prod_{i}^{N} M_{n_i l_i}^{k_i} \tag{26}$$

is invariant to rotation. The invariants, as defined by (26), are, in general, complex valued. If real-valued features are needed, the real and imaginary parts (or equivalently, the magnitude and phase) of each of the invariants can be taken separately. A basis $\mathcal{B}$ for the invariants generated using (26) is:

$$\mathcal{B} = \left\{ M_{n,l}\left[M^{l}_{n_0,l_0}\right]^*, \quad n, l \geq 0, \quad M_{n_0,l_0} \neq 0 \right\}. \tag{27}$$

Set $\mathcal{B}$ is *complete* (see the Appendix for proof) as it can be shown that invariants generated using (26) can be expressed in terms of the elements in $\mathcal{B}$ using only operations consisting of addition/subtraction, multiplication, involution with a positive/negative integer exponent, and complex conjugation. The above form of invariance can be similarly applied to PCT and PST.

3.6 Orientation Estimation

In the last section, we have shown how to negate the effects of the rotation angle by forming invariance from the PHT coefficients. But it is also possible, on the other hand, to use the PHT coefficients to estimate the orientation of an image by not discarding the rotation information. The formulation is quite straightforward as we can rewrite (24) to obtain the angle of orientation shift:

$$\phi = -\frac{1}{l} \arg \frac{M'_{nl}}{M_{nl}}. \tag{28}$$

Hence, if we define the orientation of an image to be ω, the rotated image has an orientation of $\omega - \phi$. In cases of noise or unintentional image variations, a more accurate estimate of ϕ can be obtained by taking the average of the ϕ values returned by various PHT coefficients with different values of ns and ls.

3.7 Image Reconstruction

The kernels of PHTs are pairwise orthogonal, and hence, any image can be expressed in terms of the PHT coefficients and the kernels in the following form:

$$f(r,\theta)=\sum_{n=-\infty}^{\infty}\sum_{l=-\infty}^{\infty}M_{nl}H_{nl}(r,\theta). \tag{29}$$

If only a subset $(n,l)\in\mathcal{S}'\subset\mathcal{S}=\{(n,l)||n|,|l|=0,1,\ldots,\infty\}$ of the PHT coefficients is available, an approximation of the image function, denoted as $\hat{f}(r,\theta)$, can be obtained by:

$$\hat{f}(r,\theta)=\sum_{(n,l)\in\mathcal{S}'}M_{nl}H_{nl}(r,\theta). \tag{30}$$

The reconstruction error:

$$\epsilon^2=\int_0^{2\pi}\int_0^1[f(r,\theta)-\hat{f}(r,\theta)]^2 r dr d\theta \tag{31}$$

can be easily proven, for the case of PCET, to be:

$$\epsilon^2=\frac{1}{\pi}\sum_{(n,l)\in\mathcal{S}\setminus\mathcal{S}'}M_{nl}^2. \tag{32}$$

It can be shown that the partial sum, given by relation (29), converges to the image when the image is nonzero in a limited region, is continuous, or has a finite number of finite jumps so that $f(r,\theta)$ and $|f(r,\theta)|^2$ are integrable and $R_n(r)$ are bounded over $(0,1)$ as n becomes infinite [22]. These conditions are, in general, satisfied.

3.8 Relation with Rotational Moments

PCET coefficients can be expressed in terms of rotational moments:

$$D_{nl}=\int_0^{2\pi}\int_0^1 r^n e^{-il\theta}f(r,\theta)r dr d\theta \tag{33}$$

in the following form:

$$M_{nl}=\sum_{k=0}^{\infty}\frac{(i2\pi n)^k}{k!}D_{2k,l}. \tag{34}$$

From the above equation, it can be observed that the information captured by the PCET coefficients is two times the frequency of that of rotational moments.

3.9 Fourier Series Interpretation

PCET coefficients can be expressed in the terms of Fourier series by observing that:

$$M_{nl}=\frac{1}{2\pi}\int_0^{2\pi}\left[\frac{1}{2\pi}\int_0^{2\pi}g(r',\theta)e^{-inr'}dr'\right]e^{-il\theta}d\theta, \tag{35}$$

where $r'=2\pi r^2$ and $g(r',\theta)\equiv f(\sqrt{r'/2\pi},\theta)$. PCET coefficients are hence 2D fourier coefficients as defined above—first in the radial direction, followed by the circular direction.

3.10 3D Formulation

Defining (r,θ,φ) as the radial distance, zenith, and azimuth, respectively, we can form 3D PHTs with similar properties of their 2D versions. The kernels of 3D PHTs are defined as:

$$H_{nml}(r,\theta,\varphi)=\frac{1}{r}R_n(r)Y_{ml}(\theta,\varphi). \tag{36}$$

The spherical harmonics [23], [24], [25] $Y_{ml}(\theta,\varphi)$ are given by:

$$Y_{ml}(\theta,\varphi)=N_{ml}P_{ml}(\cos\theta)e^{im\varphi}, \tag{37}$$

where N_{ml} is a normalization factor:

$$N_{ml}=\sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}} \tag{38}$$

and P_{ml} denotes the associated Legendre functions. 3D PHT coefficients can be computed using the above kernels to perform volumetric integrals over a 3D object. For more information on 3D invariants, please refer to [24] and [26]. A 3D formulation as such is important for 3D object (ranging from sparse point sets to dense volumetric MRI or PET) recognition, registration, segmentation, etc.

3.11 Discrete Implementation

Up to this point, the discussion has been limited to the case where the image is assumed to be defined on a continuous domain. In practice, however, this is not the case and images are defined on a discrete domain. As such, (1) can be written in Cartesian coordinates as:

$$M_{nl}=\frac{1}{\pi}\int\int_{x^2+y^2\le 1}[H'_{nl}(x,y)]^* f'(x,y)dxdy, \tag{39}$$

where $H'_{nl}(x,y)=H'_{nl}(r\cos\theta,r\sin\theta)\equiv H_{nl}(r,\theta)$ and

$$f'(x,y)=f(r\cos\theta,r\sin\theta)\equiv f(r,\theta).$$

Given an image defined on a discrete domain $g[k,l]$, where $k=0,\ldots,M-1$ and $l=0,\ldots,N-1$, we map the image to a domain of $(x_k,y_l)\in[-1,1]\times[-1,1]$ with:

$$x_k=\frac{k-M/2}{M/2},\quad y_l=\frac{l-N/2}{N/2} \tag{40}$$

and we have:

$$\begin{aligned}M_{nl}&=\frac{1}{\pi}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}[H'_{nl}(x_k,y_l)]^* f'(x_k,y_l)\Delta x\Delta y\\&=\frac{4}{\pi MN}\sum_{k=0}^{M-1}\sum_{l=0}^{N-1}[H'_{nl}(x_k,y_l)]^* f'(x_k,y_l)\end{aligned} \tag{41}$$

subject to $x_k^2+y_l^2\le 1$ (only the center portion of the image is considered in computing the transform coefficients), where $f'(x_k,y_l)=g[k,l]$, and note that $\Delta x=2/M$, $\Delta y=2/N$. A more detailed discussion about this form of discretization and mapping can be found in [27].

3.12 Feature Selection

Features generated might or might not be useful, depending on the application involved. Hence, feature selection is a crucial step to sieve through all available features and to pick out those which are really useful for the task at hand. For this purpose, a general good choice of feature selector and classifier constructer is the adaptive boosting algorithm (AdaBoost) [28]. In essence, the learning algorithm is used to boost the classification performance of a set of simple (sometimes called weak) classifiers to form a strong classifier. There are a number of formal guarantees provided by the AdaBoost learning procedure. Schapire et al. [29] proved that the training error of the strong classifier approaches zero exponentially in a number of rounds. More importantly, a number of results were later proved about generalization performance. The key insight

TABLE 2
Kernel Computation Time (s)

Moments	Computation Time (ms/moment)
ZM	17.47
PZM	19.78
PCET	2.55
PCT	2.66
PST	2.53
RHFM	2.60
RM	4.49

is that generalization performance is related to the margin of the examples, and that AdaBoost achieves large margins rapidly. We have shown in [30] how AdaBoost can be used to boost the performance of complex moments.

4 Experiments

PHTs are generally more feasible than ZMs/PZMs when computationally economical kernels are needed. Moreover, the ease of computation of PHTs makes the computation of higher order coefficients more tractable when compared to that of ZMs/PZMs. In this section, we will show the results from a number of experiments which are performed to gauge the performance of PHTs.

4.1 Kernel Computation Complexity

Here, we show in terms of computation time how much less complex the computation of PHT kernels is when compared to those of ZMs and PZMs. Table 2 shows a comparison of the computation time needed for computing the ZM, PZM, and PHT kernels. Values shown are time in seconds per moments, averaged over 10 trials for kernels computed up to $0 \leq n, l \leq 50$ for a 100×100 image. We have used the coefficient method [19] for computing the ZM and PZM kernels. The PC used for testing has a 2.50 GHz Processor with 2 GB RAM. The test is done using MATLAB version 7.4. As can be observed, the computation time needed for PHT kernels is much less than those of ZMs and PZMs.

4.2 Image Reconstruction and Numerical Stability

The most direct method of testing the image representation capability of the transforms is via image reconstruction. Image reconstruction also gives an indication of when the numerical stability of certain feature breaks down. In this experiment, finite sets of ZMs, PZMs,[4] and PHT coefficients of the 20 test images, shown in Fig. 4, are first calculated, and then, from these sets of moments/coefficients, the images are reconstructed according to (30). The difference between the reconstructed image and the original image is then measured using the Root-Mean-Square Error (RMSE)[5] as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{x'^2+y'^2\leq 1}[f(x,y) - \hat{f}(x,y)]^2}{|\{(x',y')|x'^2+y'^2 \leq 1\}|}}. \quad (42)$$

4. We have used the coefficient methods [19] for computing the kernels of ZMs and PZMs for all the experiments.

5. The reconstructed images are clipped at 0 and 255 so that the pixel values are maintained in the range of $[0, 255]$.

Let K be some constant. We have limited the number of moments/coefficients used in reconstruction based on the following:

- ZMs: $n - |l| = \text{even}$, $|l| \leq n \leq K$,
- PZMs: $|l| \leq n \leq K$,
- PCET: $|l| + |n| \leq K$,
- PCT: $|l| + n \leq K$,
- PST: $|l| + n \leq K$,
- RHFMs: $|l| + n \leq K$.

These conditions are selected so that the moments/coefficients capturing the lowest frequency information are used for the reconstruction process. Some examples of the reconstructed images are shown in Fig. 3. As more moments are added to the reconstruction process, the reconstructed images get closer to the original images. As can be observed from the reconstructed images, PCT, PST, and PCET capture the image information, especially the edges, better than ZMs and PZMs. For the same maximum order K of moments used, the images reconstructed using ZMs are less sharp judging from the edges. As for PZMs, the reconstructed images degrade quite swiftly toward the end when the number of moments is further increased. In Fig. 5, we show plots of the average RMSE values. It can be observed that the numerical stability of PZMs breaks down when the number of moments is increased up to a certain point, as evident from the sudden upturn of the respective RMSE curve. ZMs break down later in the curve (not shown in the figure). The performance of PST is better than ZMs and is very close to PZMs. PCT gives better reconstruction results when compared to ZMs and PZMs. We have also included the results for RHFMs here and it is as per prediction that RHFMs have some inherent numerical stability problem. Special care has to be taken so that $r = 0$ is avoided while generating the kernels of RHFMs.

4.3 Pattern Recognition Using Simulated Data

In this experiment, we gauge the noise robustness of ZMs, PZMs, PCET, PCT, and PST under noisy conditions via a simple pattern recognition experiment. The images utilized in this experiment are taken from the COREL photograph data set, which was used in [31]. The test set consists of 100 selected images which are initially stored in JPEG format with size 384×256 or 256×384. Each image is converted to gray scale and scaled to a standard size of 128×128 before its features are extracted. Some sample images are shown in Fig. 6. These images are the training images, moments/coefficients of which will be taken as the ground truth for comparison with that of the testing images. The testing images comprise the rotated and noise-contaminated versions of the 100 images. The images are first rotated at angles $\theta = 0, 45, \ldots, 315$ degree and are then added with Gaussian noise of variance[6] $\sigma^2 = 0.00, 0.05, 0.10, 0.20$. These noise-contaminated images are then classified according to their moments/coeffcients by comparing their distance (euclidean distance) to that of the training images. The average accuracies for three trials are shown in Table 3. It can be seen that PCT gives performance comparable to that of PZMs, with PCET and PST following closely behind but still better than ZMs and RMs. The

6. The variances are normalized values. They correspond to image intensities ranging from 0 to 1.

Fig. 3. Some samples of reconstructed images. $K = 0, 1, \ldots, 29$. (a) Polar complex exponential transform, (b) polar cosine transform, (c) polar sine transform, (d) Zernike moments, and (e) pseudo-Zernike moments.

results of rotational moments are included here to show that nonorthogonal moments are often less effective when compared to the orthogonal ones.

In this following experiment, we evaluate the performance of rotational invariants formed with (26): More specifically, we will only consider (27) since it forms the basis of, and hence, encapsulates the same information as (26). We compare our results with Flusser and Suk's moment rotation invariants, which are based on complex moments [9], [10]:

$$C_{pq} = \int_0^{2\pi} \int_0^{\infty} r^{p+q} e^{i(p-q)\theta} f(r, \theta) r dr d\theta. \tag{43}$$

For comparison, we set $n = p + q$, $l = p - q$ for PCT and PST and $n = \text{floor}((p + q)/2)$, $l = p - q$ for PCET. The number n for PCET is halved because the kernels of PCET have two times the number of zeros when compared to those of PCT and PST. Hence, to match their spatial frequencies, the orders of the radial components of PCET are set to half of that of PCT and PST. For the same set of images used in Section 4.3, but scaled to 32×32 for faster computation, the results are shown in Table 4. The number of features used are determined by the constant K, where $0 \leq p + q \leq K$, which is equivalent to $T = (K + 1)(K + 2)/2$ features.[7] From the table, it is evident that PCET, PCT, and PST yield significant improvement when compared to the rotation invariants generated based on complex moments in the majority of cases.

Fig. 4. Test images. Each image is resized to 64×64 before performing the experiment. The range of pixel values is $[0, 255]$.

7. PST has one less feature because $M^s_{0,0} = 0$, which happens when $p = q = 0$.

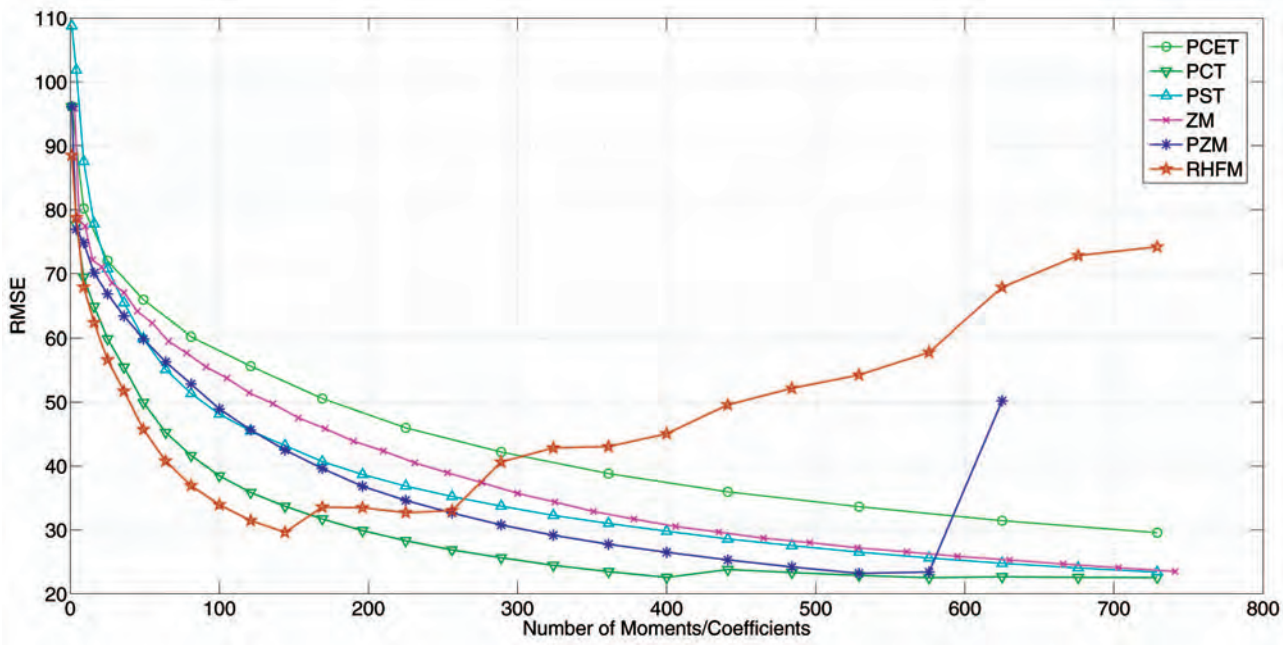

Fig. 5. Average RMSE values yielded by the various methods in the image reconstruction experiment.

4.4 Character Recognition as a Real-World Problem

We evaluate the performance of PHT using the well-known MNIST database [32] of handwritten digits. MNIST consists of 60,000 training and 10,000 testing images, each of which has been size-normalized and centered in a 28×28 box (784 pixels). Some samples are shown in Fig. 7. The euclidean distance nearest neighbor classifier using raw pixels is used as the baseline for comparison. To study the effect of the number of pixels on the classification accuracy, we downsample the images by factors of 2 and 3 (corresponding to $14 \times 14 = 196$ and $10 \times 10 = 100$ pixels, respectively) before feeding them into the classifier for training and testing. Due to the large memory requirement and computation cost, out of the 70,000 samples, only 7,000 are used for this experiment: 6,000 for training and 1,000 for testing. For PHTs, we use a similar nearest neighbor classifier with the coefficients normalized to prevent the domination of a subgroup of features. The normalization is performed by subtracting off the mean and dividing by the standard deviation of the training samples. Following the convention in Section 4.3 for invariants of the form (26), we choose $K = 3, 4, 5$ (corresponding to 60, 90, and 126 features, respectively), which we shall see later gives results close to that of the baseline method. Instead of using the PCET, PCT, and PST coefficients as individual features, we combine them into a single feature vector. Since PCET, PCT, and PST features capture different characteristics of an image, combining all PHT coefficients into a feature vector makes it more discriminant. The classification results are shown in Table 5 (Original),[8] where the performance of the baseline method, with no pixel downsampling, matches that reported in [32]. It can be seen that for a lesser amount of features, PHTs outperform the baseline method in all three cases. To further demonstrate the benefit of PHTs, we randomly rotate the images in the testing set by -45 to 45 degrees, and the classification accuracy of these rotated images is shown in Table 5 (rotated). It clearly shows that in such a case the PHTs maintain relatively high classification accuracy, while the baseline method unsurprisingly fails. One should keep in mind that PHTs in general provide plenty of features sufficient for image representation. However, PHTs are, at the current stage, not specifically designed for classification in the sense that some PHT features might not be discriminative and some might even be unstable. In light of this, a more principled scheme of feature selection and also a more intelligent form of classifier which properly weights each feature need to be devised so as to take full advantage of the rich representation capability of PHTs. In this sense, this experiment is not designed to sufficiently demonstrate the classification performance of PHT-based features, since there is neither feature selection nor dedicated classifier training mechanisms involved. Rather, it illustrates that the PHT features extracted from the image raw pixels do contain adequate information for pattern classification.

Fig. 6. Some samples of test images for the pattern recognition experiment.

5 Conclusion

In conclusion, we have proposed a set of transforms, namely, Polar Complex Exponential Transform, Polar

8. The digits "6" and "9" are treated the same since the invariance property of PHTs makes them indistinguishable.

TABLE 3
Average Classification Rates η (Percent) for Images Under Different Degrees of Gaussian Noise

K	σ^2	RM	ZM	PZM	PCET	PCT	PST
5	0.00	98.50	99.00	100.00	100.00	100.00	100.00
	0.05	19.42	48.79	77.67	72.92	77.63	23.71
	0.10	11.54	19.54	36.42	30.08	33.96	11.96
	0.15	9.29	12.04	20.50	18.63	17.58	7.71
	0.20	7.67	8.25	14.42	13.08	12.42	4.54
10	0.00	97.50	100.00	100.00	100.00	100.00	100.00
	0.05	23.50	83.54	98.29	91.33	98.08	88.79
	0.10	13.21	38.75	59.33	47.96	61.50	41.71
	0.15	11.46	24.75	34.17	22.25	28.63	20.79
	0.20	8.21	16.33	20.96	16.58	19.79	17.17
15	0.00	97.50	100.00	100.00	100.00	100.00	100.00
	0.05	24.67	98.17	99.92	96.96	100.00	98.58
	0.10	14.08	53.63	73.29	60.96	76.75	65.58
	0.15	11.46	28.96	40.17	30.67	36.17	29.50
	0.20	9.13	20.21	26.38	20.33	22.08	20.67
20	0.00	97.00	100.00	100.00	100.00	100.00	100.00
	0.05	26.92	98.88	100.00	99.42	100.00	100.00
	0.10	14.63	61.21	79.08	69.92	81.92	76.38
	0.15	12.54	32.33	46.92	37.92	44.92	38.63
	0.20	9.46	20.96	29.38	22.75	25.17	23.33
25	0.00	97.00	100.00	100.00	100.00	100.00	100.00
	0.05	26.54	99.96	100.00	99.92	100.00	100.00
	0.10	14.92	70.00	82.75	74.63	86.75	81.50
	0.15	12.13	37.71	49.96	42.96	50.63	46.50
	0.20	9.50	25.67	32.88	23.79	27.33	25.79
30	0.00	96.50	100.00	100.00	100.00	100.00	100.00
	0.05	27.75	100.00	100.00	100.00	100.00	100.00
	0.10	15.54	76.33	85.67	76.71	90.13	87.42
	0.15	12.08	40.71	55.21	45.38	54.63	51.92
	0.20	9.21	26.54	35.08	26.50	31.17	30.54

Cosine Transform, and Polar Sine Transform, for rotation-invariant image representation. In comparison to ZMs/PZMs and PCET, the proposed PCT and PST are much less complex in terms of kernel generation. This is in addition to the fact that, for the latter, numerical stability issue is nonexistent. This is apparently a desirable property judging from the number of papers devoted to the mitigation of the numerical problem of ZMs/PZMs. Some other advantages of the set of Polar Harmonic Transforms are economical kernel storage space, tractable higher order coefficients, and easier implementation. Experiments conducted show that the proposed transforms exhibit performance which is comparable to that of the popular ZMs and PZMs. Possible applications of PHTs, just to name a few, are image retrieval, image data mining, biomedical imaging such as MRI or PET scans (using the 3D formulation), face detection, corneal surface modeling, and texture classification.

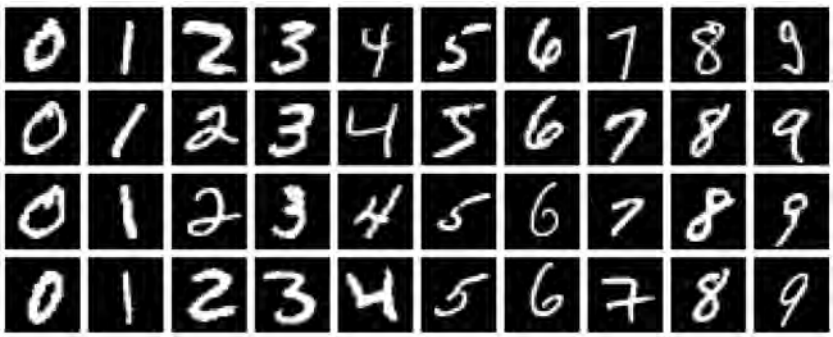

Fig. 7. Sample images from the MNIST handwritten digits database.

TABLE 4
Average Classification Rates η (Percent) for Images Under Different Degrees of Gaussian Noise

K	σ^2	PCET	PCT	PST	CM
5	0.00	100.00	100.00	100.00	100.00
	0.01	99.96	99.83	99.46	99.38
	0.02	99.46	98.58	96.50	93.13
	0.03	97.92	91.29	92.79	80.54
	0.04	92.58	81.79	86.92	67.33
	0.05	86.17	70.38	82.13	56.13
10	0.00	99.50	100.00	98.50	100.00
	0.01	99.08	98.83	91.75	93.46
	0.02	98.08	96.50	86.46	78.88
	0.03	96.54	91.33	82.00	64.50
	0.04	92.96	82.38	75.17	52.58
	0.05	86.88	68.38	69.33	43.71
15	0.00	98.00	88.50	75.00	97.50
	0.01	96.54	87.25	69.50	82.04
	0.02	94.88	84.50	64.08	67.25
	0.03	93.25	79.88	59.54	55.50
	0.04	89.58	72.79	55.88	46.88
	0.05	84.83	64.17	52.75	37.88
20	0.00	95.00	87.00	71.00	86.50
	0.01	94.21	84.17	61.67	70.58
	0.02	92.17	80.75	57.92	61.04
	0.03	90.58	76.58	54.21	50.96
	0.04	86.71	70.46	50.71	42.38
	0.05	81.58	61.17	48.29	34.54

APPENDIX

PROOF OF COMPLETENESS

Equation (26) can be written as:

$$I = \prod_{i}^{N} M_{n_i l_i}^{k_i} \left[M_{n_0,l_0}^{\sum_{j=1}^{N} k_j l_j} \right]^*$$

since:

$$\sum_{j=1}^{N} k_j l_j = 0.$$

TABLE 5
Classification Using the MNIST Data Set

BASELINE			
Pixels	100	196	784
Original	86.80%	92.40%	94.30%
Rotated	65.90%	72.90%	77.80%

PHTs			
Features	60	90	126
Original	93.40%	96.30%	100.00%
Rotated	92.30%	95.20%	95.80%

Grouping terms with $i = j$, we have

$$I = \prod_{i}^{N} \left\{ M_{n_i l_i} \left[M_{n_0,l_0}^{l_i} \right]^* \right\}^{k_i}.$$

Hence,

$$\mathcal{B} = \left\{ M_{n,l} \left[M_{n_0,l_0}^{l} \right]^*, \quad n, l \geq 0, \quad M_{n_0,l_0} \neq 0 \right\}$$

is the complete basis for I.

ACKNOWLEDGMENTS

This work was supported by Singapore A*Star SERC Research Grant No: 0621300056. This work was performed when Pew-Thian Yap was with the Nanyang Technological University.

REFERENCES

[1] M. Teague, "Image Analysis via the General Theory of Moments," *J. Optical Soc. Am.*, vol. 70, no. 8, pp. 920-930, Aug. 1980.
[2] C. Teh and R. Chin, "On Image Analysis by the Method of Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 10, no. 4, pp. 496-513, July 1988.
[3] D. Iskander, M. Collins, and B. Davis, "Optimal Modeling of Corneal Surfaces with Zernike Polynomials," *IEEE Trans. Biomedical Eng.*, vol. 48, no. 1, pp. 87-95, Jan. 2001.
[4] H.S. Kim and H.-K. Lee, "Invariant Image Watermark Using Zernike Moments," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 13, no. 8, pp. 766-775, Aug. 2003.

[5] N.H. Foon, Y.-H. Pang, A.T.B. Jin, and D.N.C. Ling, "An Efficient Method for Human Face Recognition Using Wavelet Transform and Zernike Moments" *Proc. Int'l Conf. Computer Graphics, Imaging, and Visualization,* pp. 65-69, July 2004.
[6] C. Kan and M.D. Srinath, "Invariant Character Recognition with Zernike and Orthogonal Fourier-Mellin Moments," *Pattern Recognition,* vol. 35, no. 1, pp. 143-154, Jan. 2002.
[7] L. Wang and G. Healey, "Using Zernike Moments for the Illumination and Geometry Invariant Classification of Multispectral Texture," *IEEE Trans. Image Processing,* vol. 7, no. 2, pp. 196-203, Feb. 1998.
[8] S. Ghosal and R. Mehrotra, "Detection of Composite Edges," *IEEE Trans. Image Processing,* vol. 3, no. 1, pp. 14-25, Jan. 1994.
[9] Y. Abu-Mostafa and D. Psaltis, "Recognitive Aspects of Moment Invariants," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 6, no. 6, pp. 698-706, Nov. 1984.
[10] Y. Abu-Mostafa and D. Psaltis, "Image Normalization by Complex Moments," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 7, no. 1, pp. 46-55, Jan. 1985.
[11] H. Shu, L. Luo, and J.-L. Coatrieux, "Moment-Based Approaches in Imaging. Part 1: Basic Features," *IEEE Eng. in Medicine and Biology Magazine,* vol. 26, no. 5, pp. 70-74, Sept./Oct. 2007.
[12] J.-L. Coatrieux, "Moment-Based Approaches in Imaging. Part 2: Invariance," *IEEE Eng. in Medicine and Biology Magazine,* vol. 27, no. 1, pp. 81-83, Jan./Feb. 2008.
[13] J.-L. Coatrieux, "Moment-Based Approaches in Imaging. Part 3: Computational Considerations," *IEEE Eng. in Medicine and Biology Magazine,* vol. 27, no. 3, pp. 89-91, Aug. 2008.
[14] R. Mukundan and K. Ramakrishnan, *Moment Functions in Image Analysis—Theory and Applications.* World Scientific, 1998.
[15] H. Ren, A. Liu, J. Zou, D. Bai, and Z. Ping, "Character Reconstruction with Radial-Harmonic-Fourier Moments," *Proc. Fourth Int'l Conf. Fuzzy Systems and Knowledge Discovery,* vol. 3, pp. 307-310, Aug. 2007.
[16] A. Bhatia and E. Wolf, "On the Circle Polynomials of Zernike and Related Orthogonal Sets," *Proc. Cambridge Philosophical Soc.,* vol. 50, pp. 40-48, 1954.
[17] Y. Sheng and L. Shen, "Orthogonal Fourier-Mellin Moments for Invariant Pattern Recognition," *J. Optical Soc. Am. A,* vol. 11, no. 6, pp. 1748-1757, June 1994.
[18] Y. Sheng and J. Duvernoy, "Circular Fourier-Radial Mellin Descriptors for Pattern Recognition," *J. Optical Soc. Am. A,* vol. 3, no. 6, pp. 885-888, June 1986.
[19] C.-W. Chong and R.P.R. Mukundan, "A Comparative Analysis of Algorithms for Fast Computation of Zernike Moments," *Pattern Recognition,* vol. 36, no. 3, pp. 731-742, 2003.
[20] M. Al Rawi, "Fast Zernike Moments," *J. Real-Time Image Processing,* vol. 3, nos. 1/2, pp. 89-96, Mar. 2008.
[21] J. Flusser and T. Suk, "Construction of Complete and Independent Systems of Rotation Moment Invariants," *Computer Analysis of Images and Patterns,* pp. 41-48, Springer, 2003.
[22] D. Jackson, *Fourier Series and Orthogonal Polynomials,* ch. VII and XI. Math. Assoc. Am., 1941.
[23] M. Novotni and R. Klein, "3D Zernike Descriptors for Content Based Shape Retrieval," *Proc. Eighth ACM Symp. Solid Modeling and Applications,* June 2003.
[24] N. Canterakis, "3D Zernike Moments and Zernike Affine Invariants for 3D Image Analysis and Recognition," *Proc. 11th Scandinavian Conf. Image Analysis,* pp. 85-93, 1999.
[25] T. Funkhouser, P. Min, M. Kazhdan, J. Chen, A. Halderman, D. Dobkin, and D. Jacobs, "A Search Engine for 3D Models," *ACM Trans. Graphics,* vol. 22, no. 1, pp. 83-105, 2003.
[26] C.H. Lo and H.S. Don, "3-D Moment Forms: Their Construction and Application to Object Identification and Positioning," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 11, no. 10, pp. 1053-1064, Oct. 1989.
[27] S. Liao and M. Pawlak, "On the Accuracy of Zernike Moments for Image Analysis," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 20, no. 12, pp. 1358-1364, Dec. 1998.
[28] Y. Freund and R.E. Schapire, "1997, A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting," *Proc. European Conf. Computational Learning Theory,* pp. 23-37, 1995.
[29] R.E. Schapire, P.B.Y. Freund, and W.S. Lee, "Boosting the Margin: A New Explanation for the Effectiveness of Voting Methods," *Proc. 14th Int'l Conf. Machine Learning,* pp. 322-330, 1997.
[30] P.-T. Yap, X.D. Jiang, and A. Kot, "Boosted Complex Moments for Discriminant Rotation Invariant Object Recognition," *Proc. Int'l Conf. Pattern Recognition,* pp. 1-4, 2008.
[31] J.Z. Wang, J. Li, and G. Wiederhold, "SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries," *IEEE Trans. Pattern Analysis and Machine Intelligence,* vol. 23, no. 9, pp. 947-963, Sept. 2001.
[32] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-Based Learning Applied to Document Recognition," *Proc. IEEE,* vol. 86, no. 11, pp. 2278-2324, Nov. 1998.

Pew-Thian Yap received the PhD degree in image processing from the University of Malaya in 2007. He was a postdoctoral fellow with Nanyang Technological University (NTU), Singapore before joining the University of North Carolina at Chapel Hill as a postdoctoral research associate. His current research interests include registration, segmentation, and statistical analysis of medical images, particularly magnetic resonance images.

Xudong Jiang received the BEng and MEng degrees from the University of Electronic Science and Technology of China (UESTC) in 1983 and 1986, respectively, and the PhD degree from Helmut Schmidt University, Hamburg, Germany, in 1997, all in electrical and electronic engineering. From 1986 to 1993, he was a lecturer at UESTC, where he received two Science and Technology Awards from the Ministry for Electronic Industry of China. From 1993 to 1997, he was with Helmut Schmidt University, Hamburg, as a scientific assistant. From 1998 to 2002, he was with Nanyang Technological University (NTU), Singapore, as a senior research fellow, where he developed a fingerprint verification algorithm that achieved the most efficient and the second most accurate fingerprint verification at the International Fingerprint Verification Competition (FVC '00). From 2002 to 2004, he was a lead scientist and the head of the Biometrics Laboratory at the Institute for Infocomm Research, Singapore. He joined NTU as a faculty member in 2004. Currently, he serves as the director of the Centre for Information Security, School of Electrical and Electronic Engineering, NTU, Singapore. His research interests include pattern recognition, signal and image processing, computer vision, and biometrics. He is a senior member of the IEEE.

Alex Chichung Kot has been with Nanyang Technological University, Singapore, since 1991. He headed the Division of Information Engineering at the School of Electrical and Electronic Engineering for eight years and served as associate chair/research, vice dean research for the School of Electrical and Electronic Engineering. He is currently a professor and associate dean for College of Engineering. He has published extensively in the areas of signal processing for communication, biometrics, image forensics, and information security. He served as an associate editor for the *IEEE Transactions on Signal Processing*, the *IEEE Transactions on Circuits and Systems for Video Technology*, the *IEEE Transactions on Circuits and Systems I and II*, and the *EURASIP Journal of Advanced Signal Processing*. He also served as a guest editor for two IEEE journals. He is currently an associate editor of the *IEEE Transactions on Multimedia*, *IEEE Signal Processing Letters*, and an editorial board member of *IEEE Signal Processing Magazine*. He is a member of the Visual Signal Processing and Communication Technical Committee and the Image and Multidimensional Signal Processing Technical Committee. He has served the IEEE in various capacities such as the general cochair for the 2004 IEEE International Conference on Image Processing. He served as an IEEE distinguished lecturer and is a fellow of the IEEE and the IES.