Pattern Recognition Letters xxx (2011) xxx–xxx

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Rotation-invariant texture features from the steered Hermite transform

Alfonso Estudillo-Romero *, Boris Escalante-Ramirez

Universidad Nacional Autonoma de Mexico, Fac. de Ingenieria, Edif. de Posgrado e Investigacion, Ciudad Universitaria, C.P. 04510, Mexico, D.F., Mexico

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Texture analysis
Hermite transform
Rotation-invariant texture features

ABSTRACT

We propose the steered Hermite transform to analyze and capture visual patterns from textures regardless their orientation. Visual texture information is locally described as one dimensional patterns by steering the Cartesian Hermite coefficients according to the energy direction; therefore, no predefined orientation selective filters are required. We evaluate classification accuracy of some texture features individually. During the training stage, a filter selection strategy based on the augmented variance ratio analysis of the training features is employed in order to determine the filters that provide better classification accuracy and reduce computational costs during the classification stage.

1. Introduction

Texture can be described as a set of repetitive patterns within a spatial neighborhood. The extent of such spatial neighborhood depends upon the scale or resolution at which the texture is perceived. Many research areas and even more industrial applications use texture features to segment (Mao and Jain, 1992), retrieve (Manjunath and Ma, 1996), classify (Mao and Jain, 1992; Ojala et al., 2002) and evaluate quality requirements of some products (Arivazhagan et al., 2006).

In texture analysis it is often desired to obtain rotationally invariant texture features because in most cases textures are presented with orientation variations. Early methods realizing the importance of rotation invariance used polarograms (Davis, 1981). Model based methods proposed a circular symmetric autoregressive model (Kashyap and Khotanzad, 1986), Gaussian Markov random field models (GMRF) (Cohen et al., 1991) and combination of quadrature mirror filters with hidden Markov model (Chen and Kundu, 1994).

Although wavelets by nature are not rotation invariant, wavelet-based methods have recently performed rotation-invariant texture analysis based on preprocessing stages, by transforming the input texture to a rotation invariant form through polar (Pun and Lee, 2003) or Radon transformations (Jafari-Khouzani and Soltanian-Zadeh, 2005). These transformations involve many free parameters whose optimal values need to be found. Gabor wavelets have also been proposed. In this case, several techniques have been proposed in order to obtain rotation-invariant features. These include post-processing stages such as circular shifts over the feature map according to a dominant orientation (Rallabandi and Rallabandi, 2008; Montoya-Zegarra et al., 2008) and addition of the different directional coefficients at each scale of analysis (Han and Ma, 2007). Other wavelet-based methods use combination of phase and magnitude information obtained from complex Daubechies wavelets (Chu and Chan, 2009) and show high accuracy after a feature selection procedure.

Other methods propose texture descriptions invariant not only to rotation but scale and perspective changes (Varma and Zisserman, 2003; Lazebnik et al., 2005; Zhang et al., 2007). The method presented in (Varma and Zisserman, 2003) assumes that a pixel depends only on its neighborhood and is independent of the rest of the image. Neighborhood descriptors based on Markov random fields are clustered using the *k*-means algorithm, thus producing a set of representative textons per texture class. These textons are recorded in a texton dictionary which records all the cluster centers (textons) for all samples within the dataset. In order to build a representation of a new texture, the texton dictionary is used to label images. The methods presented in (Lazebnik et al., 2005; Zhang et al., 2007) use features extracted from a sparse set of keypoint locations or patches by first extracting salient image structures using local region detectors such as Harris and Laplace and rotationally invariant descriptors such as spin images and RIFT. Testing challenging datasets with the above methods show high accuracy in retrieval and classification experiments. However, these are highly computational demanding methods since the free parameters may include the patches size, histogram computations (e.g. 100-dimensional and 32-dimensional histogram descriptors per image patch (Lazebnik et al., 2005)) and a clustering algorithm to generate a texture signature represented by the number of clusters and their weights.

The Cartesian Hermite transform was firstly introduced in digital image processing by Martens (1990). It is a local

* Corresponding author.
E-mail addresses: aestudillor@uxmcc2.iimas.unam.mx (A. Estudillo-Romero), boris@servidor.unam.mx (B. Escalante-Ramirez).

0167-8655/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.patrec.2011.06.014

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

decomposition technique that expands an image into orthogonal polynomials with respect to a Gaussian window. One of the advantages of the Hermite transform over other wavelet-based methods is that its analysis functions are similar to Gaussian derivatives. Psychovisual evidence suggests that the Gaussian derivatives fit the receptive field profiles of mammalian (RFP) visual systems (Young, 1978). We can take advantage of the Hermite analysis functions to extract visual details from textures using different analysis orders on multi-scale or multi-resolution schemes. Fast implementation techniques based on the Zak transform have also been proposed for computing expansion coefficients relative to the frame formed from Hermite functions at different spatial frequencies (Gertner and Geri, 1994).

Research on the steerability of the Cartesian Hermite coefficients has found that the steered Hermite transform (van Dijk and Martens, 1997; Silvan-Cardenas and Escalante-Ramirez, 2006; Martens, 2006), is an efficient way to compactly describe image features into a smaller number of coefficients than the Cartesian Hermite transform. Moreover, since the steering direction depends on the local maximum energy, it is possible to obtain image descriptors invariant to the image orientation. This method presents an advantage over classical filter bank design because in the latter a fixed number of orientations for the analysis have to be selected.

Previous works related to the use of the Hermite transform to extract texture features have proposed a Gabor-like Hermite model (Rivero-Moreno and Bres, 2003). It has been tested in denoising and indexing experiments (Bres et al., 2005; Eglin et al., 2007) using the Hermite analysis filters tuned at fixed number of orientations. Others have used the steered Hermite transform for texture retrieval (Estudillo-Romero and Escalante-Ramirez, 2009). However, in these works there is no discussion on the classification effects in vector dimensionality reduction, nor on the kind of texture features that better classify textures.

In our proposal, visual image details are extracted regardless the orientation by steering the Cartesian Hermite coefficients on the direction of local maximum energy. We evaluate the classification accuracy of some of the different texture features reported in the literature. Dimensionality reduction of the feature set is also a main issue when implementing filter banks. A disadvantage of the feature selection strategy in (Chu and Chan, 2009) is that this is performed after the classification stage by selecting the features that give the highest correct classification rate (CCR). The method employed in (Celik and Tjahjadi, 2011) for dimensionality reduction of the feature vectors is based on principal components analysis (PCA). However, every principal component obtained from the PCA is a linear combination of the whole elements in the feature vector. This represents a disadvantage since frequently filter based methods require to reduce the number of filtering operations during the classification stage. Moreover, after a PCA transformation, the feature elements in the original feature space which better discriminate a given texture are still unknown. The proposed strategy in the present method is to obtain the most discriminant filter indexes from the training dataset using the *augmented variance ratio* (AVR) and then use these filters to perform testing with the texture features obtained with the selected filters. This represents an advantage in real vision system implementations. Evaluation of the CCR was performed using two distance metrics.

Section 2 summarizes the Hermite transform theory for one and two dimensional signals. In Section 3 the steered Hermite transform is presented. We analyze and compare with the Gabor model some common rotation-invariant features in Section 4. In Section 5 we explain the pipeline of our proposal. Experimental setups and results are reported in Section 6 and finally, conclusions are given in Section 7.

2. Cartesian Hermite transform

2.1. One dimensional case

For the one dimensional case, a polynomial transform $L_n(x)$ is a local decomposition technique in which an input signal $L(x)$ is localized through a window $V(x)$ and then expanded into orthogonal polynomials $G_n(x)$ at every window position (Martens, 1990):

$$L_n(x_0) = \int_x L(x) G_n(x_0 - x) V_n^2(x_0 - x) dx \tag{1}$$

The Hermite transform arises when G_n are the Hermite polynomials $H_n(x)$, given by Rodrigues' formula (Abramowitz and Stegun, 1965):

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n e^{-x^2}}{dx^n}, \quad n = 0, 1, 2, \ldots \tag{2}$$

and the orthogonal window corresponds to a Gaussian window (a Gaussian-weighted window can be found in (Yang and Reeves, 1995)):

$$V(x) = \frac{1}{\sqrt{\sqrt{\pi}\sigma}} \cdot e^{-x^2/2\sigma^2} \tag{3}$$

Following (1), the expansion coefficients $L_n(x)$ can be derived by convolution of the signal $L(x)$ with the Hermite analysis functions $d_n(x)$, which are given in terms of the window and Hermite polynomials as:

$$d_n(x) = \frac{(-1)^n}{\sqrt{2^n n!}} \cdot \frac{1}{\sigma\sqrt{\pi}} H_n\left(\frac{x}{\sigma}\right) e^{-x^2/\sigma^2} \tag{4}$$

2.2. Two dimensional case

Generalization of the Hermite analysis functions to two dimensions can be easily extended, since the analysis functions have the property of being both spatially separable and rotationally symmetric. We then can write the two dimensional analysis functions as:

$$d_{n-m,m}(x, y) = d_{n-m}(x) d_m(y) \tag{5}$$

where $n - m$ and m denote the analysis order in x and y direction respectively. As a result, we can expand a given input image $L(x,y)$ into the basis $d_{n-m,m}(x,y)$ as:

$$L_{n-m,m}(x_0, y_0) = \int_x \int_y L(x, y) \cdot d_{n-m,m}(x_0 - x, y_0 - y) dx dy \tag{6}$$

for $n = 0, 1, \ldots, \infty$ and $m = 0, \ldots, n$.

2.3. Frequency domain

The Hermite analysis functions in the Fourier domain can be written as:

$$\hat{d}_{n-m,m}(\omega_x, \omega_y) = \frac{(-j\omega_x\sigma)^{n-m}(-j\omega_y\sigma)^m}{\sqrt{2^n (n-m)! m!}} \cdot \hat{d}_{0,0}(\omega_x, \omega_y) \tag{7}$$

where

$$\hat{d}_{0,0}(\omega_x, \omega_y) = e^{-\sigma^2(\omega_x^2 + \omega_y^2)/4} \tag{8}$$

corresponds to a 2D Gaussian with scale parameter σ. The center frequency and the frequency bandwidth are related to the derivative order and the scale of analysis of the Hermite analysis function respectively. From Eq. (7) we see that the analysis of the image at different frequencies can be performed by increasing the order of the Hermite analysis function. Also from Eq. (8) we observe that increasing of the spread σ on the spatial domain reduces the spatial

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

frequency coverage of 2D Gaussian on the Fourier domain which is suitable for analysis of low frequencies and vice versa. In fact, the relationship is given by $(\omega\sigma)^2 = 2n$. Therefore, the frequency bandwidth and the spatial scale parameter are inversely proportional. In this way, the behavior of an image decomposition with a set of Hermite analysis functions of increasing order and at different scales resembles the systematic analysis used in common wavelet schemes. This property can be used to detect patterns at a given scale of analysis. Such patterns give information about the characteristics of the texture under analysis.

3. Steered Hermite transform

A steerable filter is described as a class of filters in which a filter of arbitrary orientation is synthesized as a linear combination of a set of *basis filters* (Freeman and Adelson, 1991). Since all Hermite analysis filters are polynomials times a radially symmetric window function, rotated versions of a filter of order n can be constructed by taking linear combinations of the original filters of order n. In this way, a more general expression of the original $L_{n-m,m}$ Cartesian Hermite coefficients can be written in terms of the orientation selectivity θ (Silvan-Cardenas and Escalante-Ramirez, 2006):

$$L^{\theta}_{n-m,m}(x_0, y_0, \theta) = \sum_{k=0}^{n} L_{n-k,k}(x_0, y_0)\alpha_{n-k,k}(\theta) \tag{9}$$

which has been named the steered Hermite transform in (van Dijk and Martens, 1997). The terms $\alpha_{n-m,m}(\theta)$ are the Cartesian angular functions of order n which give such orientation selectivity and are defined as:

$$\alpha_{n-m,m}(\theta) = \sqrt{C^m_n}\cos^{n-m}(\theta)\sin^m(\theta) \tag{10}$$

where C corresponds to a binomial window which approximates the discrete Gaussian window.

We can write the local energy in terms of the steered Hermite coefficients as:

$$E_N = \sum_{n=0}^{N}\sum_{m=0}^{n}[L_{n-m,m}]^2 = \sum_{n=0}^{N}\sum_{m=0}^{n}[L^{\theta}_{n-m,m}]^2 \tag{11}$$

for all $N \geqslant 0$.

In natural images, many of the image details that are of prime importance, such as edges and lines, can be locally described as one-dimensional patterns, that is, patterns that vary only in one direction (and are constant along the orthogonal direction). One may distinguish 1D local energy terms and 2D local energy terms. Thus, we can split local energy of (11) up to order N as:

$$E_N = [L_{0,0}]^2 + E^{1D}_N + E^{2D}_N \tag{12}$$

where $L_{0,0}$ represents the DC Hermite coefficient and

$$E^{1D}_N = \sum_{n=1}^{N}\left[L^{\theta}_{n,0}\right]^2 \tag{13}$$

$$E^{2D}_N = \sum_{n=1}^{N}\sum_{m=1}^{n}\left[L^{\theta}_{n-m,m}\right]^2 \tag{14}$$

One of the objectives when steering coefficients is to maximize detection of patterns along a given local direction θ. In this way, (van Dijk and Martens, 1997; Silvan-Cardenas and Escalante-Ramirez, 2006; Martens, 2006) propose strategies in which θ is selected such that E^{1D}_N is maximized. As a consequence, compaction of the energy of the Cartesian Hermite coefficients can be efficiently achieved. One of such strategies is to take the orientation angle between the vertical and horizontal first order Cartesian Hermite coefficients which approximates the direction of the maximum energy on the decomposition. In our proposal a multi-resolution decomposition requires θ to be computed at each resolution level. In this way θ contains directionality of coarse structures at low resolution levels.

4. Rotation-invariant features

As stated before, the orientation angle between the vertical and horizontal first order Cartesian Hermite coefficients approximates the direction of local maximum energy on the decomposition. Moreover, an auxiliary function named angular function that is based on such directional information is used to weight the Hermite Cartesian coefficients in order to obtain a rotation-invariant coefficients set. Therefore, the steered Hermite transform presented in the previous section is suitable to extract rotation-invariant texture features such as energy and second order statistics.

In this section we present visual and numerical examples of such rotation-invariant texture features and compare with a common approach using Gabor wavelets. In fact, both models have been used to describe the RFP. These models involve Gaussian functions in some form to perform smoothing and effectively localize a signal (Gertner and Geri, 1994).

We considered rotated versions of an input image. Note that the original information can not be completely preserved due to

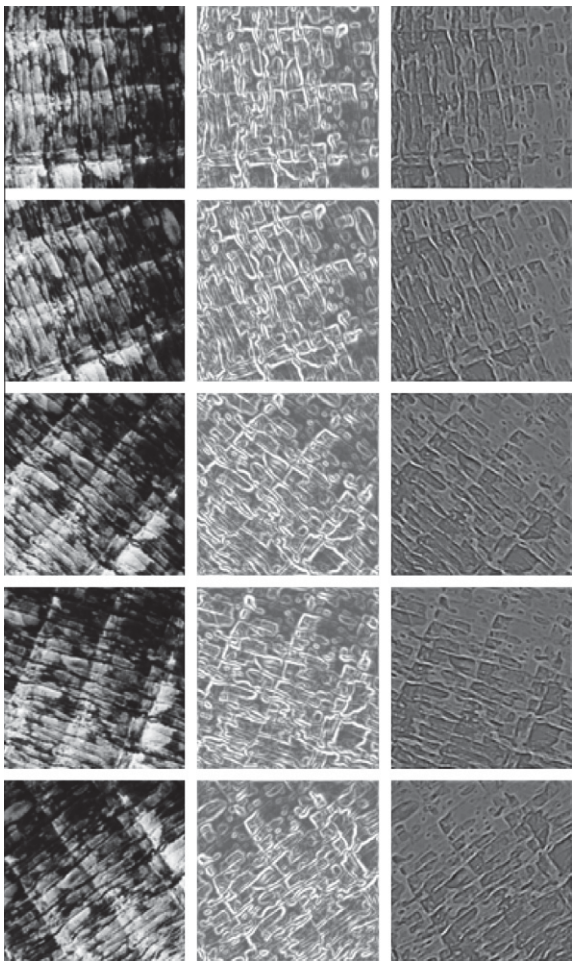

Fig. 1. Rotation of an input texture and its 1D steered Hermite coefficients for $n = 1$ and $n = 2$. Input texture corresponds to a region from the VisTex texture *Bark0000* VisTex, 2002. The input texture has been rotated to show that the same visual information is preserved and obtained by means of the steered Hermite transform. From the top to the bottom counterclockwise rotating angles are 0°, 17°, 45°, 65° and 120°.

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

cropping and the method that was used to rotate the input image. We have assumed that illumination conditions were preserved during the rotation process.We aim to show the intra-variability of the texture features between the various rotated versions of the same input texture that were computed from the steered Hermite coefficients.

Fig. 1 shows the input texture and its first and second order steered Hermite coefficients. From the top to the bottom counterclockwise rotating angles are 0°, 17°, 45°, 65° and 120°. First and second order coefficients, shown in the middle and right columns respectively, have the same perceptual information content in all the rotated versions of the texture.

Another method that has been popular for texture analysis involves the use of Gabor wavelets. Similar to the Hermite transform, the Gabor wavelet is a biologically-inspired vision model. A Gabor function on the spatial domain is a 2D Gaussian modulated sinusoid with aspect ratio of σ_x/σ_y and radial center frequency F:

$$G(x,y,\sigma_x,\sigma_y,F,\theta)=\frac{1}{2\pi\sigma_x\sigma_y}e^{-\frac{1}{2}\left(\frac{x_r^2}{\sigma_x^2}+\frac{y_r^2}{\sigma_y^2}\right)}\cdot e^{(j2\pi Fx_r)} \tag{15}$$

where the Gaussian function and the complex sine grating share the same orientation (Bovik et al., 1990) given by:

$$\begin{bmatrix}x_r\\ y_r\end{bmatrix}=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}\cdot\begin{bmatrix}x\\ y\end{bmatrix} \tag{16}$$

To simplify notation, from now on we will refer to a Gabor function Eq. (15) as $G(x,y,\theta)$. The design of a Gabor family is constrained to parameters such as: central frequency, orientation selectivity, frequency and angular bandwidths. A comprehensive filter design can be found in (Clausi and Ed Jernigan, 2000).

A common approach in texture analysis is to filter an input texture with a family of Gabor functions whose filters at each analysis frequency of increasing frequency bandwidths are tuned to different orientations. In order to cover the spatial frequency plane, the analysis is performed in octave steps and with minimal degree of overlap between the spatial frequency coverage of the filters. In order to obtain rotation invariance the filtered images with the different oriented filters are then added at each octave thus obtaining a more compact set of filtered images:

$$F_{im}^{\theta}(x,y)=\sum_{k=0}^{K-1}\mathrm{Im}[(L(x,y)\star G(x,y,\theta_k))] \tag{17}$$

$$F_{re}^{\theta}(x,y)=\sum\nolimits_{k=0}^{K-1}\mathrm{Re}[(L(x,y)\star G(x,y,\theta_k))] \tag{18}$$

where K is the number of discrete orientations of the filter bank and $\star$ means convolution of the input image $L(x,y)$ with the filter $G(x,y,\theta_k)$. Typical texture features such as energy, mean and standard deviation are obtained from the compact filtered images for each octave and for both imaginary and real parts of the filtered images to form the feature vector.

Fig. 2 shows the Gabor expansion coefficients of an input texture that has been rotated for several angles. From the top to the bottom counterclockwise rotating angles are 0°, 17°, 45°, 65° and 120°. The imaginary Gabor filtered textures (in the middle) and the real Gabor filtered textures (right column), were obtained from addition of their corresponding oriented bands. A first visual inspection of the filtered outputs suggests that it is not possible to extract the same information provided by the steered Hermite transform through this scheme.

Numerical computation of commonly used features such as energy, mean, and standard deviation were also performed for both models. Tables 1 and 2 show the features computed through the Hermite and Gabor models respectively. The variation coefficient for each feature (energy, mean and standard deviation), computed

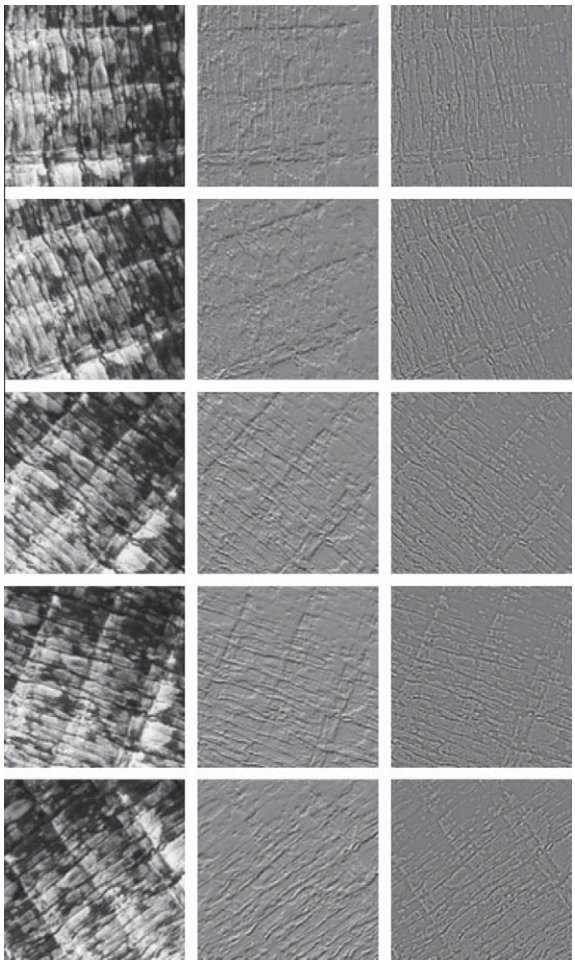

Fig. 2. Rotation of an input texture and its imaginary and real Gabor expansion coefficients. Input texture corresponds to a region from the VisTex texture *Bark0000* VisTex, 2002. The input texture has been rotated and filtered with four Gabor filters. From the top to the bottom counterclockwise rotating angles are 0°, 17°, 45°, 65° and 120°.

Table 1
Rotation-invariant features extracted from steered Hermite Coefficients. Energy feature values are given $\times 10^3$. Last column shows the betweenness class variation coefficient (c_v).

Feature	0°	17°	45°	65°	120°	c_v
$E_{L_1^\theta}$	0.7315	0.7267	0.7335	0.7530	0.7214	0.0164
$E_{L_2^\theta}$	0.1596	0.1586	0.1581	0.1651	0.1587	0.0182
$\mu_{L_1^\theta}$	0.0650	0.0651	0.0656	0.0663	0.0649	0.0087
$\mu_{L_2^\theta}$	0.0020	0.0021	0.0019	0.0021	0.0019	0.0302
$\sigma_{L_1^\theta}$	0.0458	0.0452	0.0452	0.0460	0.0450	0.0098
$\sigma_{L_2^\theta}$	0.0371	0.0370	0.0369	0.0377	0.0370	0.0090

Table 2
Rotation-invariant features extracted from Gabor coefficients. Energy feature values are given$\times 10^3$. Last column shows the betweenness class variation coefficient (c_v).

Feature	0°	17°	45°	65°	120°	c_v
$E_{F_{im}^\theta}$	3.2578	2.8436	4.1155	5.4514	5.9513	0.3122
$E_{F_{re}^\theta}$	3.2817	3.2435	3.2347	3.3301	3.2703	0.0115
$\mu_{F_{im}^\theta}$	0.0003	0.0000	0.0003	0.0019	0.0028	1.1389
$\mu_{F_{re}^\theta}$	0.0005	0.0007	0.0002	0.0003	0.0002	0.5476
$\sigma_{F_{im}^\theta}$	0.1679	0.1568	0.1887	0.2171	0.2269	0.1583
$\sigma_{F_{re}^\theta}$	0.1685	0.1675	0.1673	0.1697	0.1682	0.0057

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

from the steered coefficients of the rotated versions of the same texture, represents the betweenness class variation and is defined as:

$$c_v = \frac{\sigma_f}{\mu_f} \tag{19}$$

where σ_f and μ_f are the standard deviation and mean for each feature f respectively. The variation coefficient is shown in the right most column on both Tables 1 and 2. The slight numerical variation between the same features for all rotated versions of the input texture indicates that the extracted features from steered Hermite coefficients are robust to the texture orientation. These results suggest that the steered Hermite model better extracts and preserves not only visual but numerical information compared to the features extracted using Gabor wavelets. It has been shown in previous works that the steered Hermite transform-based features outperformed average retrieval rates of Gabor-based texture features (Estudillo-Romero and Escalante-Ramirez, 2009).

5. Proposed methodology

The pipeline of the proposed methodology is presented in Fig. 3. Our methodology first includes estimation during the training stage of those steered Hermite analysis filters which offer better discrimination power. The purpose is to obtain a similar or even better classification performance by using a small set of steered Hermite analysis filters during the testing set. Classification is performed with the k-NN classifier. We evaluate classification performance using two distance metrics.

5.1. Features and filter selection

We evaluated our methodology with several features found in the literature:

1. Mean, μ
2. Standard deviation, σ
3. Energy features

$$E_0 = \sum_{x=1}^{M}\sum_{y=1}^{N}[L_{n,0}^{\theta}]^2 \tag{20}$$

$$E_1 = \frac{E_0}{M \times N} \tag{21}$$

$$E_2 = \frac{1}{M \times N}\sum_{x=1}^{M}\sum_{y=1}^{N}|L_{n,0}^{\theta}| \tag{22}$$

$$E_3 = \frac{1}{M \times N}\sum_{x=1}^{M}\sum_{y=1}^{N}\sqrt{|L_{n,0}^{\theta}|} \tag{23}$$

Because testing datasets are bigger than training datasets, by finding *optimal* filters during the training stage we can obtain the same or even improve discrimination performance and reduce computational cost during the testing stage at the same time that the feature dimension of our descriptor decreases. This strategy may be useful in an on-line real-time system performing texture classification tasks.

Since the texture characteristics are unknown, we propose to extract texture features using an overcomplete decomposition during the training stage. In order to select a subset of the *optimal* filters it is necessary to process each training texture with a full Hermite decomposition followed by a steering transformation of the Cartesian Hermite coefficients. In this work we extract several texture features from the steered Hermite coefficients to generate a feature vector for every single feature.

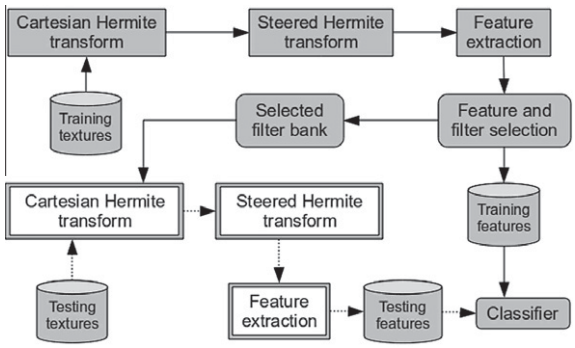

Fig. 3. Texture analysis scheme using the steered Hermite transform and dimensionality reduction via feature selection.

A texture feature vector $\mathbf{x}_i$ is obtained by concatenating the feature that is extracted from each steered Hermite coefficient $1 \leqslant n \leqslant N$ up to order N at every scale of analysis s, where $0 \leqslant s \leqslant S-1$ and S represents the number of scales:

$$\mathbf{x}_i = \left[f_1^{(0)}, f_2^{(0)}, \ldots, f_1^{(1)}, f_2^{(1)}, \ldots, f_N^{(S-1)}\right] \tag{24}$$

Once the feature vectors (with dimension of $S \times N$) have been obtained for each texture in the training texture set, these are arranged in a feature matrix as:

$$\mathbf{F_X} = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,n} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,n} \\ \vdots & \vdots & \vdots & \vdots \\ x_{m,1} & x_{m,2} & \cdots & x_{m,n} \end{bmatrix} \tag{25}$$

where $x_{m,n}$ represents the feature n of the vector $\mathbf{x}_m$. The *augmented variance ratio* (AVR) (Liu et al., 2004) has shown to provide a quantitative basis to separate non-discriminative features before feature subset selection. The AVR is defined as:

$$AVR(F) = \frac{Var(S_F)}{\frac{1}{C}\sum_{i=1..C}\frac{Var_i(S_F)}{\min_{i\neq j}(|mean_i(S_F) - mean_j(S_F)|)}} \tag{26}$$

where $Var(S_F)$ is the cross-class variance of feature F, $Var_i(S_F)$ and $mean_i(S_F)$ are the within-class variance and mean of the feature F for class i out of C distinct classes. Similar to Fisher criterion, AVR represents the ratio of cross-class variance of the feature over within-class variance, with added penalty to features that have close inter-class means.

By ranking the elements of the feature vectors with high AVR ratios it is possible to know the indexes of the filters that were used to obtain such features. This represents a way to obtain a similar or even better classification accuracy with a reduced number of feature elements. Filter selection to extract testing texture features is a straightforward procedure.

5.2. Classifier

The final stage in our pipeline (Fig. 3) is classification. In this paper we evaluated classification accuracy based on the k-NN classifier using two distance metrics. The k-NN classifier consists in assigning a class label to a sample vector based on the mode of its k nearest neighbors vectors. Given two feature vectors $\mathbf{x}$ and $\mathbf{y}$ of n features, the Euclidean distance is defined as:

$$d_E(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{j=1}^{n}(x_j - y_j)^2} \tag{27}$$

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

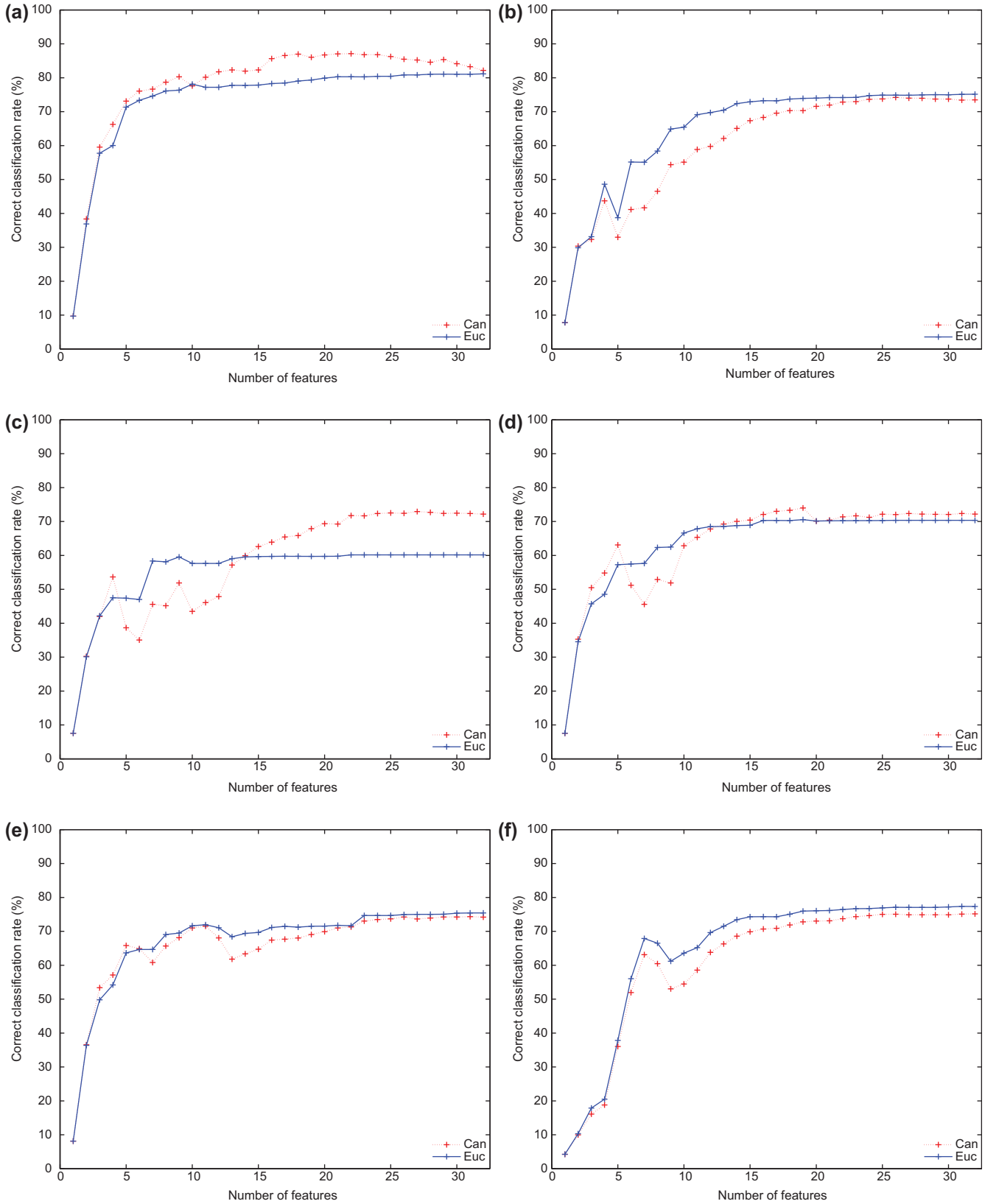

Fig. 4. CCR using (a) Mean, (b) Standard deviation, (c), E_0, (d) E_1, (e) E_2, (f) E_3 for experiment Section 6.2.

and the Canberra distance as:

$$d_C(\mathbf{x},\mathbf{y}) = \sum_{j=1}^{n} \frac{|x_j - y_j|}{|x_j| + |y_j|} \tag{28}$$

6. Experimental setups and results

In order to investigate the classification performance using a reduced set of texture features obtained via AVR, we individually

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

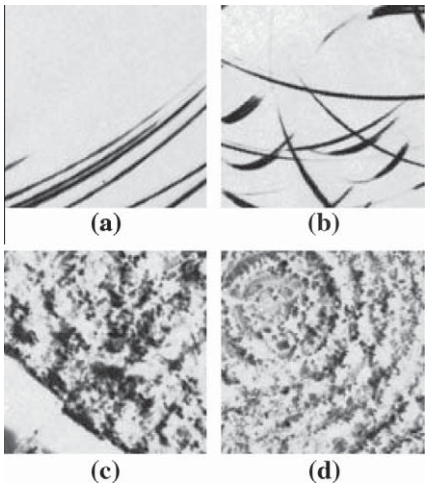

Fig. 5. Texture samples from the Brodatz database used to illustrate inter and intra-class variations. Textures (a) and (b) belong to different texture classes whereas (c) and (d) belong to the same texture class.

computed the CCR of several features presented in Section 5.1. For all the experiments the training textures were analyzed at $S = 4$ decomposition levels with the steered Hermite transform up to order $N = 8$, yielding a feature vector of 32 arranged elements as shown in Eq. (24). AVR produced a rearranged vector of the feature elements ranked in order of importance. Evaluation of the effects in feature dimensionality was performed by computing the CCR by successive increments of the vector dimension.

6.1. Datasets

1. Brodatz (1966) database. The Brodatz database contains little intra-class variability. However, some textures appear very similar up to a scale change. Most images in the Brodatz dataset have a constant texture pattern in the entire image area (except some textures such as: D7, D43, D44, D45, D97). In our experiments we include such textures to increase the intra-class variability and decrease the inter-class variability.

2. VisTex VisTex (2002) database. This database contains real world texture images, some of them with different objects in the scene generating intra-class variation.
3. The UIUC database (Lazebnik et al., 2005) contains photographic textures with varying local affine transformations, caused by perspective transformations and no homogeneous areas due to defocusing. This characteristics generate a texture database with high intra-class variations. It consists of 25 classes, each one containing 40 image samples.

6.2. Experimental setup I

We used 112 texture images of 512×512 pixels from the Brodatz texture album. A training texture set was formed from 16 sub-images of 128×128 pixels without overlap and rotated 0 radians which gives 16 training textures per texture class. For the testing texture set, each original texture of 512×512 pixels was first artificially rotated from $\pi/18$ to $16\pi/18$ radians with incremental step size of $\pi/18$ radians. For each rotated texture 4 sub-images of 128×128 without overlap were selected in such a way that the sub-images had minimal overlap between the different rotated versions. Finally, the training and the testing texture sets consisted of 400 images and 1600 images respectively (20% for training and 80% for testing, (Jafari-Khouzani and Soltanian-Zadeh, 2005)).

Classification results for each texture feature are depicted in Fig. 4. It was found that the average classification performance, once an optimal rate has been reached, tends to be constant despite the number of features. Both distance metrics performed well in average for all the features. Better results were obtained with the features μ, σ, E_2 and E_3. Low CCRs were observed due to the high intra-class variability that was introduced by regions of textures containing no information or that were not homogeneous over the entire texture images and also by little inter-class variability. Fig. 5 shows texture samples belonging to different classes ((a) and (b)) and to the same class ((c) and (d)).

6.3. Experimental setup II

In this experiment we used the fifty-four 512×512 gray-scaled texture classes from VisTex database (VisTex, 2002) that were used in (Pan et al., 2008). Each original image was divided into 16 128×128 non-overlapping sub-images, comprising a training database of 864 (54×16) images. To create the testing set, each original texture image was rotated at angles ranging from 10° to 160° and 10° increments. From the center of each rotated image, four non-overlapping sub-images were extracted. Therefore, a total of 3840 ($60 \times 16 \times 4$) testing images were obtained.

We performed classification experiments using only one kind of texture feature at a time. Classification results for each texture feature are depicted in Fig. 6. Similar to Fig. 4, classification performance remains constant after the optimal rate was found, regardless the features number. Measuring 16-dimension feature vector similarities with the Euclidean distance showed that σ, E_2 and E_3 have better CCR than the μ, E_0 and E_1. However, similarity measures using the Canberra distance with the same number of vector elements showed in average (84.81%) better CCR for all the features than using the Euclidean distance (80.76%). It is interesting to note the high CCR obtained with only 16 feature elements. Other methods tested with the same dataset and experimental setup reported correct classification percentages of 98.9% with correlation distance of the Radon projections (Wang et al., 2010) and 96.82% with the Ridgelet transform (Pan et al., 2008). They employ high dimensional feature vectors and involve many free parameters.

6.4. Experimental setup III

The classification experiment consists of removing a number N_{tr} of training samples from the databases and then using these to classify the remaining textures. In this experiment, the UIUC database was used. For each texture class we randomly selected 80 sets of N_{tr} training samples. Evaluation with many randomly selected training sets are required to minimize the classification dependance on the training set. In order to evaluate the CCR as a function of the number of training samples, N_{tr} ranges from 1 to 20 training samples. For each N_{tr} we computed the average CCR from all the classification scores obtained from the 80 random training set selections.

Classification results using the Canberra similarity measure are depicted in Fig. 7. We obtained the CCR for every single texture feature. We also evaluated classification accuracy as a function of the vector feature dimension. In all cases, except by the features μ and E_0, the dimension increment of the feature vector slightly improves the CCR. The best CCR was obtained with the μ feature. The CCRs using the Euclidean similarity measure are shown in Fig. 8. In almost all the cases, better performances were obtained with the Canberra distance based measure.

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

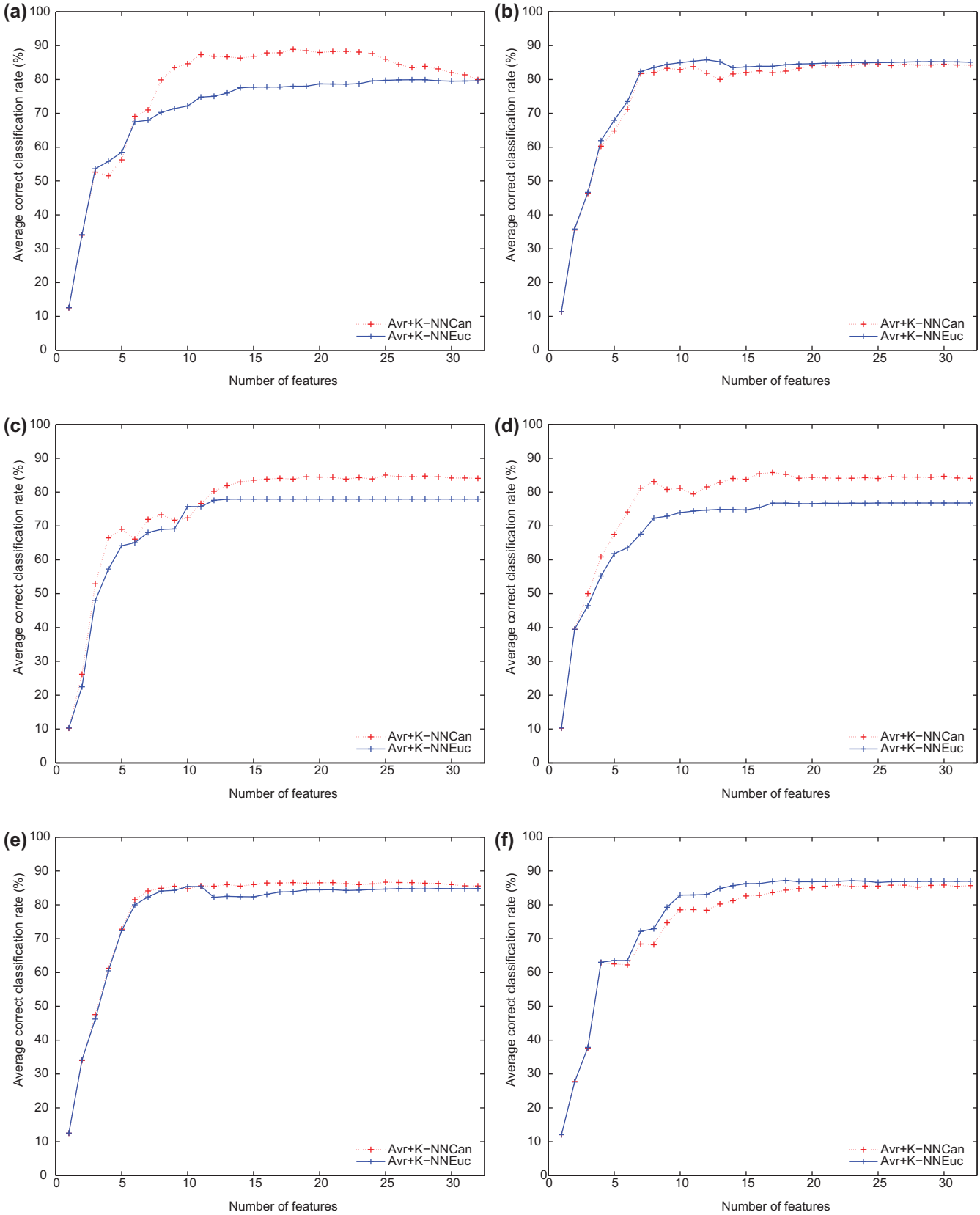

Fig. 6. CCR using (a) Mean, (b) Standard deviation, (c), E_0, (d) E_1, (e) E_2, (f) E_3 for experiment Section 6.3.

6.5. Experimental setup IV

In this experiment the same experimental setup from Experiment Section 6.4 was used. In previous experiments we found that the texture features based on the mean, standard deviation and energy (E_3) of the steered Hermite coefficients performed better than the remaining texture features. In order to improve classification accuracy we combined $[\mu,\sigma]$, $[\mu,E_3]$, $[\sigma,E_3]$ features for the $n = 8$,

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

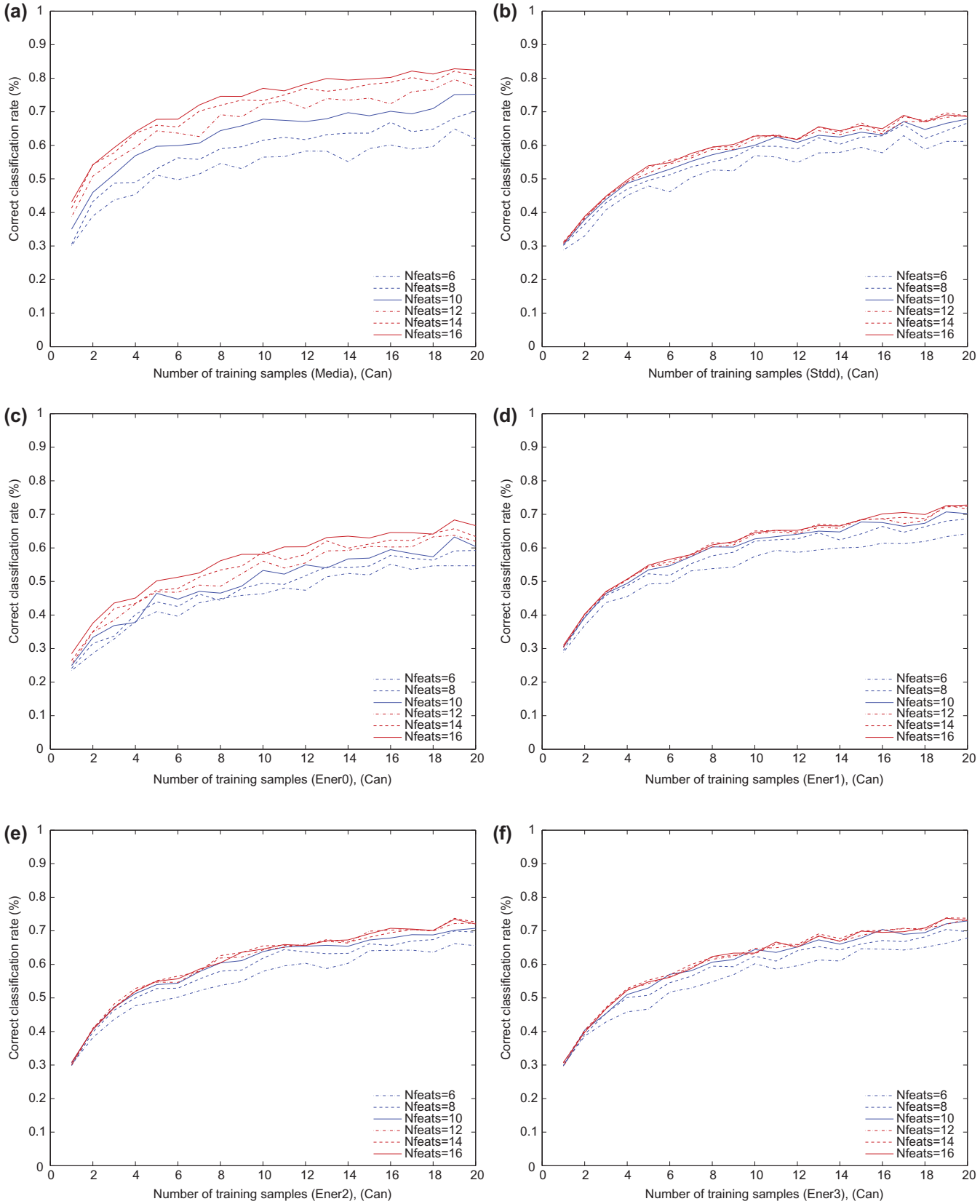

Fig. 7. CCR using (a) Mean, (b) Standard deviation, (c), E_0, (d) E_1, (e) E_2, (f) E_3 and the Canberra based distance for experiment Section 6.4.

16 and 32 most important features obtained during the training stage by means of the AVR (see Section 5.1). The mixed feature vectors were of 16, 32 and 64-dimension. Classification results are depicted in Figs. 9–11. In order to combine texture features, since one of the objectives is to reduce the number of filters used during the testing stage, selection of filters is performed by analysing the

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

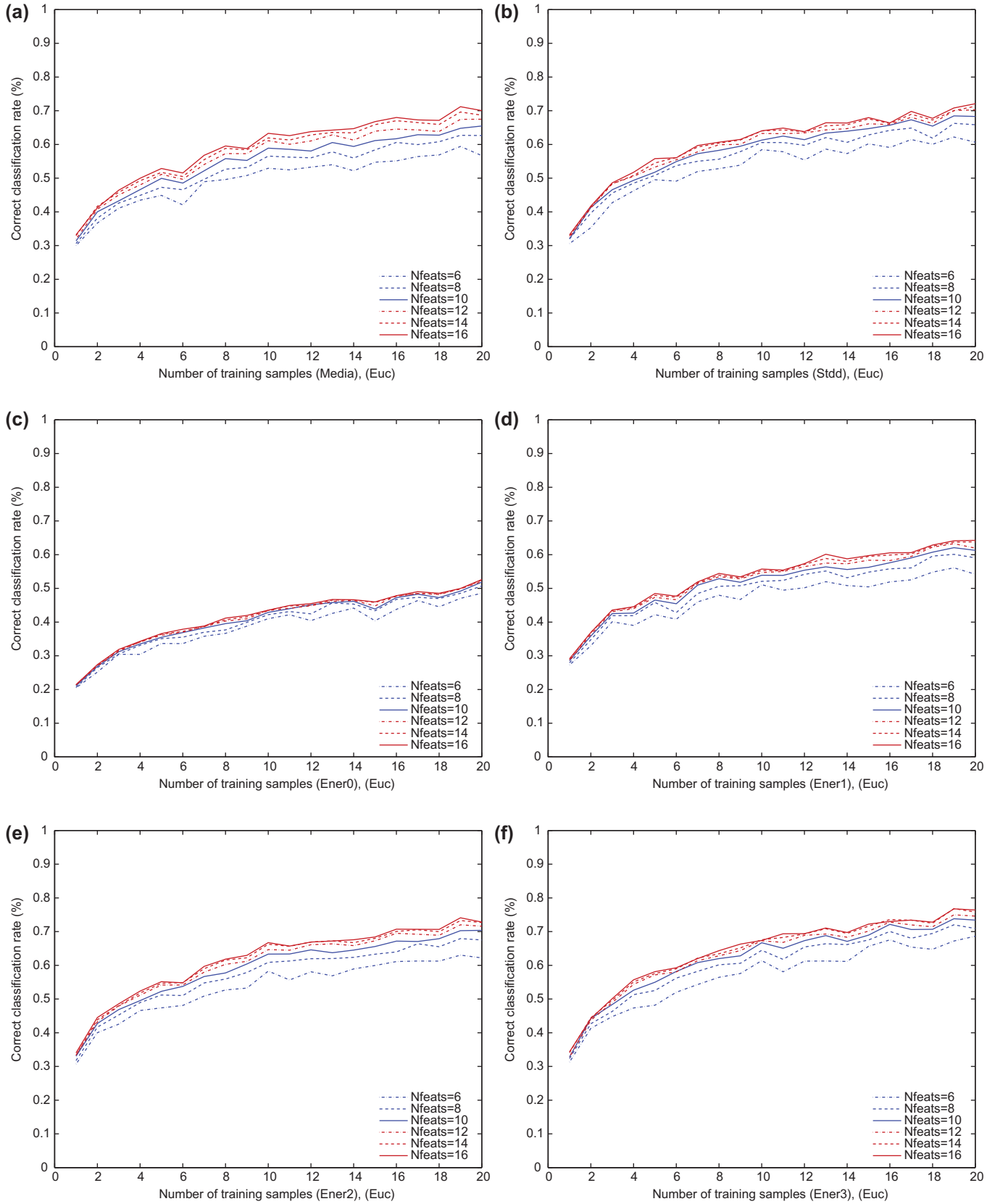

Fig. 8. CCR using (a) Mean, (b) Standard deviation, (c), E_0, (d) E_1, (e) E_2, (f) E_3 and the Euclidean based distance for experiment Section 6.4.

most significant common features between both AVR-ranked feature vectors.

From the experimental results we observed that best classification performance was achieved by combining the texture features

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

μ and σ. The second best classification performance was observed by combining μ and E_3. The Canberra distance measure gives the best classification performances for 32 and 64-dimension feature vectors. Moreover, it was observed that good classification performances can be obtained by using only 16 steered Hermite coefficients (note that the mixed 32-dimension feature vector was generated from two 16-dimension feature vectors). The combination σ and E_3 showed the worst classification performances and little differences were found using the Canberra and Euclidean distances for all the 16, 32 and 64-dimension feature vectors.

It was also interesting to show the filter indexes for mixed 16 and 32-dimension feature vectors that were used during the testing stage. We present these indexes as histograms generated from all the random training sets for the three combination cases. Fig. 12 shows the histograms of the filter indexes. The filter indexes are ordered from 1 to 32 possible filters ($S = 4$ scales of analysis and $N = 8$ is the maximum order of the Hermite decomposition) such that a filter index = 11 represents a steered Hermite coefficient with parameters $s = 2$ and $n = 3$. A steered Hermite coefficient with parameters $s = 4$ and $n = 7$ is represented by a filter index = 31. As shown in the histograms of Fig. 12, the firsts orders of the steered Hermite coefficients at each scale of analysis give the most important features with few exceptions especially when the dimension of the mixed feature vector is increased. Preference for the first steered Hermite coefficients indicates that the frequency filter responses are overlapped. This result is supported by recent research which suggests that a certain amount of superposition between the frequency filter responses could improve classification (Bianconi and Fernandez, 2007).

Table 3 shows CCR comparisons of different methods on the UIUC database using only $N_{tr} = 12$ training images. One of the most robust methods reported in the literature (Zhang et al., 2007) obtained an average CCR about 97% whereas by combining the texture features μ and σ we obtained from 16 and 32 steered Hermite analysis filters obtained an average CCR about 85%, see Table 3. A fair direct comparison is not possible yet since authors in (Zhang et al., 2007) have made intensive experiments in order to find optimal values for their free parameters including: invariance levels on the datasets, selection of the best detector/descriptor combination, evaluation of SVM with different kernels, different signature sizes. Moreover, one of the main differences between both methods is that ours provides a texture descriptor whereas

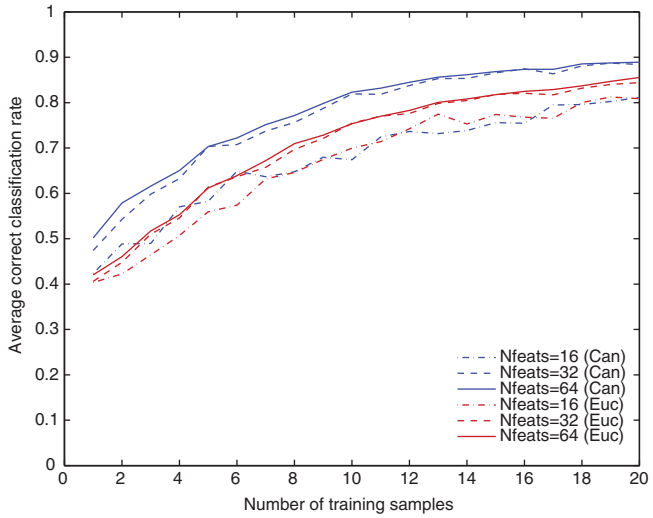

Fig. 9. Average CCR in experiment Section 6.5 using the mixed feature vectors $[\mu, \sigma]$. Average CCR using the Canberra and Euclidean based distance measures are also shown.

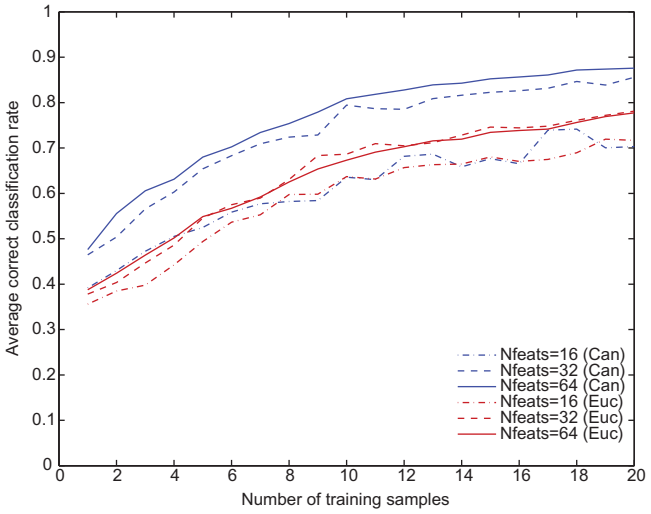

Fig. 10. Average CCR in experiment Section 6.5 using the mixed feature vectors $[\mu, E_3]$. Average CCR using the Canberra and Euclidean based distance measures are also shown.

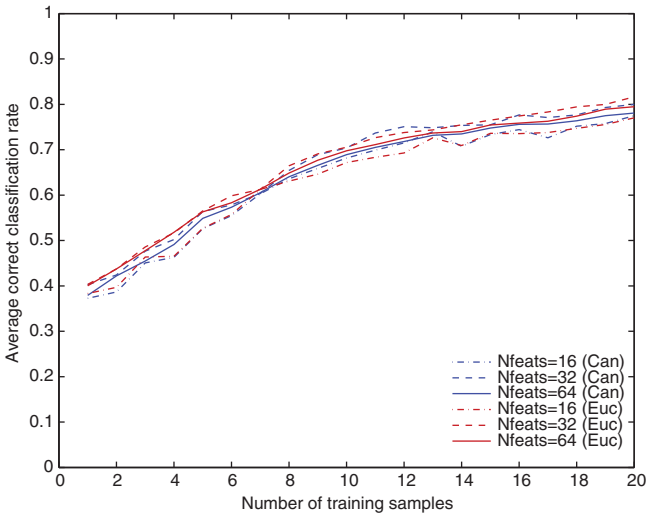

Fig. 11. Average CCR in experiment Section 6.5 using the mixed feature vectors $[\sigma, E_3]$. Average CCR using the Canberra and Euclidean based distance measures are also shown.

the method in (Zhang et al., 2007) provides an image descriptor since it combines detectors and texture descriptors. Indeed, the presented method avoids exhaustive experiments in order to find many optimal parameters influenced by a particular dataset, instead, it only estimates optimal filter indexes during the training stage and use small feature vectors. Nevertheless, we obtained better CCRs compared with previous methods involving Gabor wavelets (Manjunath and Ma, 1996) and MRF models (Varma and Zisserman, 2003).

7. Conclusions

In this paper we presented and exploited some of the properties of the steered Hermite transform to analyze and then classify a texture image regardless its orientation. Texture feature extraction was performed by considering visual information by means of the analysis functions of the Hermite transform. Visual details were then locally described as one-dimensional patterns by steering

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

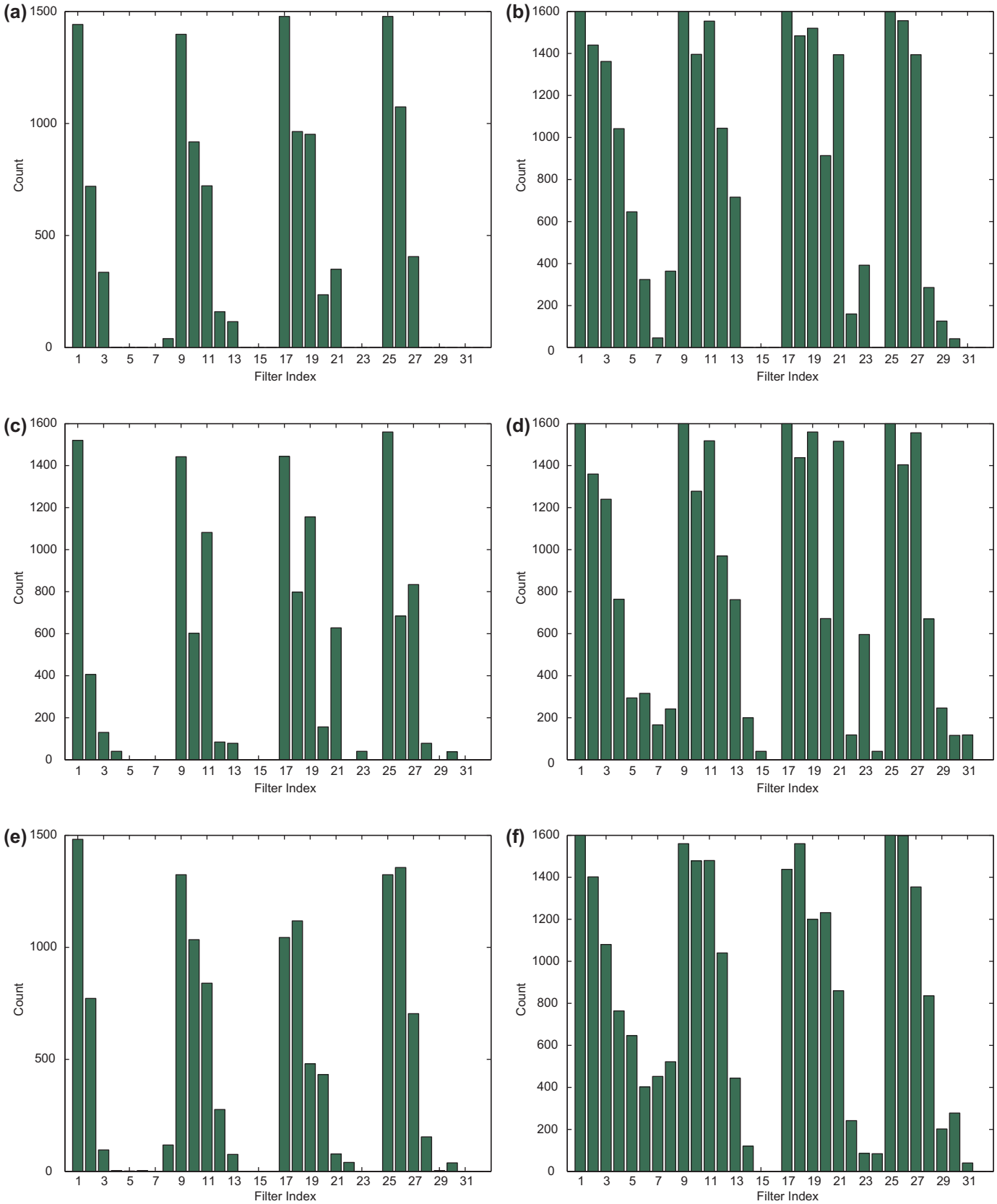

Fig. 12. Histograms of the filter indexes. (a) and (b) for a mixed 16 and 32-dimension feature vector respectively combining μ and σ. (c) and (d) for a mixed 16 and 32-dimension feature vector respectively combining μ and E_3. (e) and (f) for a mixed 16 and 32-dimension feature vector respectively combining σ and E_3.

the Cartesian Hermite coefficients. In this method, rotation invariance was achieved by steering the basis coefficients to locally maximize the energy and then represent such visual features with less number of coefficients.

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014

Table 3
Comparison of different methods on the UIUC database using only 12 training images.

Method	CCR (%)
Gabor [μ,σ] Manjunath and Ma (1996)	57
VZ-Joint Varma and Zisserman (2003)	72
Ours [μ,σ]	85
Lazebnik et al. (2005)	94
Zhang et al. (2007)	97

We evaluated classification accuracy with several kinds of texture features that were computed from each steered Hermite coefficient. Moreover, we showed that in most cases, the number of analysis filters can be significantly reduced and obtain almost the same Average CCRs.

An interesting finding in this work is that the feature dimension reduction based on the AVR showed that most of the textural discrimination power using the steered Hermite transform is in the first Hermite analysis orders. In other words, the overcomplete image representation obtained with the first successive orders of the Hermite transform increases classification accuracy.

We also observed an increase in the CCR when mixing features. However, further research is required on the combination of texture features and on filter selection strategies. The selection of appropriate distance metrics (every feature represents a different energy measure) to compare similarities is also an exploration area in future works.

Acknowledgements

This work was supported by UNAM grants IN113611 and IX100610.

References

Abramowitz, M., Stegun, I., 1965. Handbook of Mathematical Functions. Dover.
Arivazhagan, S., Ganesan, L., Bama, S., 2006. Fault segmentation in fabric images using Gabor wavelet transform. Mach. Vis. Appl. 16 (6), 356–363.
Bianconi, F., Fernandez, A., 2007. Evaluation of the effects of Gabor filter parameters on texture classification. Pattern Recognit. 40 (12), 3325–3335.
Bovik, A..C., Clark, M., Geisler, W.S., 1990. Multichannel texture analysis using localized spatial filters. IEEE Trans. PAMI 12 (1), 55–73.
Bres, S., Eglin, V., Rivero, C., 2005. Handwriting documents denoising and indexing using Hermite transform. In: Pattern Recognition and Data Mining, vol. 3686 of LNCS. pp. 664–673.
Brodatz, P., 1966. Texture: A Photographic Album for Artists and Designers. Dover, New York.
Celik, T., Tjahjadi, T., 2011. Bayesian texture classification and retrieval based on multiscale feature vector. Pattern Recognition Lett. 32 (2), 159–167.
Chen, J.L., Kundu, A., 1994. Rotation and gray scale transform invariant texture identification using wavelet decomposition and hidden Markov model. IEEE Trans. PAMI 16 (2), 208–214.
Chu, X., Chan, K., 2009. Rotation and scale invariant texture analysis with tunable Gabor filter banks. In: Advances in Image and Video Technology, vol. 5414 of LNCS. pp. 83–93.
Clausi, D.A., Ed Jernigan, M., 2000. Designing Gabor filters for optimal texture separability. Pattern Recognit. 33 (11), 1835–1849.
Cohen, F., Fan, Z., Patel, M., 1991. Classification of rotated and scaled textured images using Gaussian Markov random field models. IEEE Trans. PAMI 13 (2), 192–202.
Davis, L.S., 1981. Polarograms: A new tool for image texture analysis. Pattern Recognit. 13 (3), 219–223.
Eglin, V., Bres, S., Rivero, C., 2007. Hermite and Gabor transforms for noise reduction and handwriting classification in ancient manuscripts. Internat. J. Document Anal. Recognit. 9, 101–122.
Estudillo-Romero, A., Escalante-Ramirez, B., 2009. Advances in rotation-invariant texture analysis. In: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, vol. 5856 of LNCS. pp. 145–152.
Freeman, W., Adelson, E., 1991. The design and use of steerable filters. IEEE Trans. PAMI 13 (9), 891–906.
Gertner, I., Geri, G., 1994. Image representation using Hermite functions. Biol. Cybernet. 71, 147–151.
Han, J., Ma, K.K., 2007. Rotation-invariant and scale-invariant Gabor features for texture image retrieval. Image Vision Comput. 25 (9), 1474–1481.
Jafari-Khouzani, K., Soltanian-Zadeh, H., 2005. Rotation-invariant multiresolution texture analysis using radon and wavelet transforms. IEEE Trans. Image Process. 14 (6), 783–795.
Kashyap, R.L., Khotanzad, A., 1986. A model-based method for rotation invariant texture classification. IEEE Trans. PAMI 8 (4), 472–481.
Lazebnik, S., Schmid, C., Ponce, J., 2005. A sparse texture representation using local affine regions. IEEE Trans. PAMI 27 (8), 1265–1278.
Liu, Y., Teverovskiy, L., Carmichael, O., Kikinis, R., Shenton, M., Carter, C.S., Stenger, V.A., Davis, S., Aizenstein, H., Becker, J.T., Lopez, O.L., Meltzer, C.C., 2004. Discriminative MR image feature analysis for automatic schizophrenia and alzheimers disease classification. In: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2004, vol. 3216 of LNCS. pp. 393–401.
Manjunath, B.S., Ma, W.Y., 1996. Texture features for browsing and retrieval of image data. IEEE Trans. PAMI 18 (8), 837–842.
Mao, J., Jain, A.K., 1992. Texture classification and segmentation using multiresolution simultaneous autoregressive models. Pattern Recognit. 25 (2), 173–188.
Martens, J.B., 1990. The Hermite transform-theory. IEEE Trans. Acoust. Speech Signal Process. 38 (9), 1595–1606.
Martens, J.B., 2006. The Hermite transform: A survey. EURASIP J. Appl. Signal Process. 26145, 1110–8657.
Montoya-Zegarra, J.A., Papa, J.P., Leite, N.J., da Silva Torres, R., Falcão, A.X., 2008. Learning how to extract rotation-invariant and scale-invariant features from texture images. EURASIP J. Adv. Signal Process. 691924, 1–15.
Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. PAMI 24 (7), 971–987.
Pan, W., Bui, T.D., Suen, C.Y., 2008. Rotation invariant texture classification by ridgelet transform and frequency-orientation space decomposition. Signal Process. 88 (1), 189–199.
Pun, C.M., Lee, M.C., 2003. Log-polar wavelet energy signatures for rotation and scale invariant texture classification. IEEE Trans. PAMI 25 (5), 590–603.
Rallabandi, V.R., Rallabandi, V.S., 2008. Rotation-invariant texture retrieval using wavelet-based hidden Markov trees. Signal Process. 88 (10), 2593–2598.
Rivero-Moreno, C.J., Bres, S., 2003. Conditions of similarity between Hermite and Gabor filters as models of the human visual system. In: Computer Analysis of Images and Patterns, vol. 2756 of LNCS. pp. 762–769.
Silvan-Cardenas, J., Escalante-Ramirez, B., 2006. The multiscale Hermite transform for local orientation analysis. IEEE Trans. Image Process. 15 (5), 1236–1253.
van Dijk, A.M., Martens, J.B., 1997. Image representation and compression with steered Hermite transforms. Signal Process. 56 (1), 1–16.
Varma, M., Zisserman, A., 2003. Texture classification: are filter banks necessary? In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, vol. 2. pp. 691–698.
VisTex, 2002. Mit media lab vision textures. url<http://vismod.media.mit.edu>.
Wang, X., xia Guo, F., Xiao, B., feng Ma, J., 2010. Rotation invariant analysis and orientation estimation method for texture classification based on radon transform and correlation analysis. J. Vis. Commun. Image Represent. 21, 29–32.
Yang, J., Reeves, A., 1995. Bottom-up visual image processing probed with weighted Hermite polynomials. Neural Netw. 8 (5), 669–691.
Young, R.A., 1978. Orthogonal basis functions for form vision derived from eigenvector analysis. In: ARVO Abstracts. Association for Research in Vision and Ophthalmology, Abstract, Sarasota, FL, p. 22.
Zhang, J., Marszalek, M., Lazebnik, S., Schmid, C., 2007. Local features and kernels for classification of texture and object categories: A comprehensive study. Int. J. Comput. Vision 73, 213–238.

Please cite this article in press as: Estudillo-Romero, A., Escalante-Ramirez, B. Rotation-invariant texture features from the steered Hermite transform. Pattern Recognition Lett. (2011), doi:10.1016/j.patrec.2011.06.014