Comparison of texture features based on Gabor filters

P. Kruizinga, N. Petkov and S.E. Grigorescu
Institute of Mathematics and Computing Science, University of Groningen
P.O. Box 800, 9700 AV Groningen, The Netherlands
Email: peterkr@cs.rug.nl, petkov@cs.rug.nl, simona@cs.rug.nl

Abstract

The performance of a number of texture feature operators is evaluated. The features are all based on the local spectrum which is obtained by a bank of Gabor filters. The comparison is made using a quantitative method which is based on Fisher’s criterion. It is shown that, in general, the discrimination effectiveness of the features increases with the amount of post-Gabor processing.

1. Introduction

Features related to the local spectrum have been proposed in the literature and used in one way or another for the purpose of texture classification and/or segmentation. In most of these studies the relation to the local spectrum is established through features which are obtained by filtering with a set of two-dimensional Gabor filters. Such a filter is linear and local and is characterised by a preferred orientation and a preferred spatial frequency. Roughly speaking, it acts as a local band-pass filter with certain optimal joint localisation properties in both the spatial domain and the spatial frequency domain [5]. Typically, a multi-channel filtering scheme is used: an image is filtered with a set of Gabor filters with different preferred orientations and spatial frequencies, which cover appropriately the spatial frequency domain, and the features which are obtained form a feature vector field which is used further.

Gabor feature vectors can be used directly as input to a classification or segmentation operator or they can first be transformed into new feature vectors which are then used as such an input. In references [3, 7, 18, 19], for example, pairs of Gabor features, which correspond to the same preferred orientation and spatial frequency but differ in the value of a phase parameter, are combined, yielding the so-called Gabor-energy quantity. In references [1, 16] so-called complex moments are derived from Gabor features. Finally, in references [10, 11, 12, 15] so-called grating cell operator features are computed using Gabor features.

Since the type of ‘post-Gabor’ processing in the above mentioned methods is different, it is interesting to evaluate the effect of the different types of post-processing on the usefulness of the resulting features regarding texture classification.

At this point the question arises of how to measure the usefulness of different features in this respect. Several authors have made a comparison of the performance of various operators and features for texture segmentation. Most of these studies are based on the so-called Classification Result Comparison (CRC) [4]. In this method a segmentation algorithm is applied to a feature vector field and the number of misclassified pixels is used to evaluate the segmentation performance and suitability of the features.

For a quantitative comparison of various post-Gabor processing schemes and their related features we do not use the CRC method that is used in most previous studies because this method characterises the joint performance of a feature operator and a subsequent classifier. We rather use a new method which we proposed elsewhere [11, 12]. This method can be used to compare the features only, regardless of any subsequent classification or segmentation operations. It is based on a statistical approach to evaluate the capability of a feature operator to discriminate two textures by quantifying the distance between the corresponding clusters of points in the feature space according to Fisher’s criterion.

The rest of this paper is organised as follows: in Section 2 we review the linear Gabor filter. Various post-processing operators and features are introduced in Section 3. The properties of these operators with respect to texture classification are compared in Section 4 in a series of computational experiments. In Section 5 we summarise the results of the study and draw conclusions.

2. Gabor filters

A number of authors used a bank of Gabor filters to extract local image features [3, 7, 18, 19]. An input image \(I(x, y), x, y \in \Omega\) (\(\Omega\) - the set of image points), is convolved with a two-dimensional Gabor function \(g(x, y), x, y \in \Omega\),
to obtain a Gabor feature image \( r(x, y) \) as follows:

\[
 r(x, y) = \int_{\Omega} I(\xi, \eta) g(x - \xi, y - \eta) d\xi d\eta \tag{1}
\]

We use the following family of Gabor functions:

\[
g_{\lambda, \Theta, \varphi}(x, y) = e^{-\frac{(x^2 + y^2)}{2\sigma^2}} \cos(2\pi \frac{x'}{\lambda} + \varphi) \tag{2}
\]

\[
x' = x\cos\Theta + y\sin\Theta
\]

\[
y' = -x\sin\Theta + y\cos\Theta,
\]

The standard deviation \( \sigma \) of the Gaussian factor determines the effective size of the surrounding of a pixel in which weighted summation takes place. The eccentricity of the Gaussian and here with the eccentricity of the convolution kernel \( g \) is determined by the parameter \( \gamma \), called the spatial aspect ratio. The value \( \gamma = 0.5 \) is used in our experiments [15]. Since this value is constant, the parameter \( \gamma \) is not used to index a Gabor filter in the following.

Figure 1. Two 2-dimensional Gabor functions with the same standard deviation \( \sigma \) but with different values of the ratio \( \frac{x}{\lambda} \) and consequently different preferred spatial frequencies and spatial frequency bandwidths.

The parameter \( \lambda \) is the wavelength and \( \frac{1}{\lambda} \) the spatial frequency of the harmonic factor \( \cos(2\pi \frac{x'}{\lambda} + \varphi) \). Since the spatial frequency tuning curve of a filter with an impulse response \( g \) has a maximum at \( \frac{1}{\lambda} \), we refer to \( \frac{1}{\lambda} \) as the preferred spatial frequency of the Gabor filter. The ratio \( \frac{x}{\lambda} \) determines the spatial frequency bandwidth of the Gabor filters (see Figure 1). The half-response spatial frequency bandwidth \( \beta \) (in octaves) and the ratio \( \frac{x}{\lambda} \) are related as follows:

\[
\beta = \log_2 \left( \frac{\frac{x}{\lambda}}{\frac{x}{\lambda} - \frac{1}{2\pi}} \right) \tag{3}
\]

The angle parameter \( \Theta (\Theta \in [0, \pi]) \) specifies the orientation of the normal to the parallel positive and negative lobes of the Gabor filters (this normal is the axis \( x' \) in eq.2). Since a filter based on the function \( g \) will respond most strongly to a bar, edge or grating, the normal to which coincides with \( \Theta \), the orientation specified by \( \Theta \) is referred to in the following as the preferred orientation.

Finally, the parameter \( \varphi \), which is a phase offset in the argument of the harmonic factor \( \cos(2\pi \frac{x'}{\lambda} + \varphi) \), determines the symmetry of the function \( g_{\lambda, \Theta, \varphi}(x, y) \): for \( \varphi = 0 \) and \( \varphi = \pi \) it is symmetric, or even, with respect to the centre point \((0, 0)\); for \( \varphi = -\frac{1}{2}\pi \) and \( \varphi = \frac{1}{2}\pi \), \( g_{\lambda, \Theta, \varphi}(x, y) \) is antisymmetric, or odd.

Figure 2. Power-spectra of two 2-dimensional Gabor functions.

Figure 2 shows the power spectra of two Gabor functions with different parameter settings. The light areas indicate spatial frequencies and wavevector orientations which will pass the corresponding filters. In this way Gabor filters act as local bandpass filters.

3. Texture features based on Gabor filters

3.1. Linear Gabor features

The filter responses that result from the application of a filter bank of Gabor filters can be used directly as texture features, though none of the approaches described in the literature employ such texture features. In this study, linear Gabor features are used only for comparison. In our experiments we used two filter banks, one with symmetric and one with antisymmetric Gabor filters.

The ratio \( \frac{x}{\lambda} \) which is used, is constant (\( \frac{x}{\lambda} = 0.56 \)) for all filters in the bank and corresponds to a half-response spatial frequency bandwidth of one octave. This choice is motivated by the properties of simple cells in the visual cortex which can be modelled by the Gabor filter. The spatial frequency bandwidth and the spatial aspect ratio determine the orientation bandwidth of the filter which is about 19° at half response and is also constant for all filters in the bank used.

Three different preferred spatial frequencies and eight different preferred orientations were used, resulting in a bank of 24 Gabor filters (Figure 3). The application of such a filter bank results in a 24-dimensional feature vector in each point of the image, i.e. a 24-dimensional vector field for the whole image.

3.2. Thresholded Gabor features

In contrast to the linear features described above, most Gabor filter related texture features are obtained by apply-
ing non-linear post-processing on the vector field of linear Gabor features. The specific type of nonlinearity varies from method to method.

Several authors have proposed the application of a threshold on the Gabor filter results [9, 13], in analogy to the function of simple cells, which can be modelled by a linear weighted spatial summation, characterised by a Gabor weighting function, followed by a half-wave rectification [15]. The thresholded Gabor features are computed as follows:

\[\chi(z) = 0 \quad \text{for} \quad z < 0, \quad \chi(z) = z \quad \text{for} \quad z \geq 0 \quad \text{and} \quad r_{\lambda, \Theta, \varphi}(x, y)\]

is the filter response of a Gabor filter with a convolution kernel \(g_{\lambda, \Theta, \varphi}\).

Only filters with symmetric Gabor functions have been used in previous studies reported in the literature. For completeness, we use filters with both symmetric and antisymmetric Gabor functions in our experiments. Two banks, each of 24 filters, are used, one comprising the symmetric and the other the antisymmetric filters. They are related to the linear filter banks described in the previous subsection and have the same coverage of the spatial frequency domain.

3.3. Gabor-energy features

The filter results of a symmetric and an antisymmetric filter can be combined in a single quantity which is called the Gabor-energy. This feature is related to a model of so-called complex cells in the primary visual cortex [17] and is defined in the following way:

\[e_{\lambda, \Theta}(x, y) = \sqrt{r_{\lambda, \Theta, 0}^2(x, y) + r_{\lambda, \Theta, \varphi}^2(x, y)}\]  

where \(r_{\lambda, \Theta, 0}^2(x, y)\) and \(r_{\lambda, \Theta, \varphi}^2(x, y)\) are the responses of the linear symmetric and antisymmetric Gabor filters, respectively. Combining the symmetric and antisymmetric filter banks described in Subsection 3.1 results in a new, non-linear filter bank of 24 channels with the same coverage of the spatial frequency domain.

The Gabor-energy is closely related to the local power spectrum. The local power spectrum associated with a pixel in an image is defined as the squared modulus of the Fourier transform of the product between the image and a window function. This window function has the role of choosing a neighbourhood of the pixel of interest; a 2D Gaussian is used in most cases.

Let us consider the quantity

\[p_{\lambda, \Theta}(x, y) = r_{\lambda, \Theta, 0}^2(x, y) + r_{\lambda, \Theta, \varphi}^2(x, y)\]  

Taking into account eq. 1 and eq. 2, it is clear that \(p_{\lambda, \Theta}(x, y)\) is the local power spectrum of the input image \(I\) at point \((x, y)\) using a Gaussian windowing function which appears as a factor on the right hand side of eq. 2.

3.4. Complex moments features

In [1], the real and imaginary parts of the complex moments of the local power spectrum are proposed as features. These features are translation invariant inside homogeneous texture regions and give information about the presence or absence of dominant orientations in the texture.

The complex moments of the local power spectrum are defined as follows:

\[C_{mn}(x, y) = \iint (u + i v)^m (u - i v)^n p_{u, v}(x, y) \, du \, dv\]  

where:

\[u = \frac{1}{\lambda} \cos \Theta, \quad v = \frac{1}{\lambda} \sin \Theta, \quad p_{u, v}(x, y) = p_{\lambda, \Theta}(x, y)\]

The sum \(m + n \quad (m, n \in \mathbb{N})\) is called the order of the complex moment; it is related to the number of dominant orientations in the texture. In [2] it is proven that a complex moment of even order \(m + n\) has the ability to discriminate textures with \(m+n\) dominant orientations. For example, moments of order two \((C_{20} \text{ and } C_{11})\) are able to detect textures with a single dominant orientation. The moduli of the complex moments give information about the presence or absence of dominant orientations while their arguments specify which orientations are dominant. In [2], the advantages are discussed of considering the real and imaginary parts of the complex moments as features instead of considering the moduli and the arguments of the complex moments.
In our experiments, we use a feature vector that has as elements the non-zero real and imaginary parts of the complex moments of the local power spectrum. It can be proven that complex moments of odd order are zero and that all complex moments of even order are zero and that all complex moments up to order 8 resulting in 24-dimensional feature vectors. For computing the local power spectrum the same filter bank (with 8 orientations and 3 spatial frequencies) was used as in the computations of the Gabor-energy features.

3.5. Grating cell operator features

A different type of nonlinearity is applied in the so-called grating cell operator [10, 11, 15]. This operator is based on a computational model of a particular type of neuron which is found in areas V1 and V2 of macaque monkeys [20]. The cells differ from the majority of cells found in those areas of the visual system in that they do not react to a single line or edge. They only respond when a system of at least three bars is present in their receptive field. The grating cell operator reproduces this property of grating cells by employing an AND-type nonlinearity to combine the responses of a number of bar detectors. The operator signals periodicity with a certain spatial frequency and orientation in an image.

The grating cell operator consists of two stages [15]. In the first stage, the responses of so-called grating subunits are computed using as input the outputs of simple cell operators (see [14, 15] for further details). The grating subunit stage is conceived in such a way that the unit is activated by a set of three bars with appropriate periodicity, orientation and position. In the second stage, the outputs of grating subunits of a given preferred orientation and periodicity are summed together within a certain area to compute the output of the grating cell operator. This is next explained in more detail:

A quantity \( q_{\Theta, \lambda}(x, y) \), called the activity of a grating subunit with position \((x, y)\), preferred orientation \(\Theta(\Theta \in [0, \pi])\) and preferred grating periodicity \(\lambda\), is computed as follows:

\[
q_{\Theta, \lambda}(x, y) = \begin{cases} 
1 & \text{if } \forall n, M_{\Theta, \lambda, n}(x, y) \geq \rho M_{\Theta, \lambda}(x, y) \\
0 & \text{if } \exists n, M_{\Theta, \lambda, n}(x, y) < \rho M_{\Theta, \lambda}(x, y) 
\end{cases}
\]

where

\[ n \in \{-3 \ldots 2\} \]

and \(\rho\) is a threshold parameter with a value smaller than but near 1 (e.g. \(\rho = 0.9\)) and the auxiliary quantities \(M_{\Theta, \lambda, n}(x, y)\) and \(M_{\Theta, \lambda}(x, y)\) are computed as follows:

\[
M_{\Theta, \lambda, n}(x, y) = \max\{s_{\Theta, \lambda, \varphi_n}(x', y') \mid x', y' : \frac{n}{2}\cos\Theta \leq (x' - x) < \frac{(n + 1)}{2}\cos\Theta, \\
\frac{n}{2}\sin\Theta \leq (y' - y) < \frac{(n + 1)}{2}\sin\Theta, \\
\varphi_n = \begin{cases} 
0 & n = -3, -1, 1 \\
\pi & n = -2, 0, 2 \end{cases} \}
\]

(9)

where \(s_{\Theta, \lambda, \varphi_n}(x', y')\) is the output of a simple cell operator of preferred orientation \(\Theta\) and periodicity \(\lambda\) at position \((x', y')\), similar to the thresholded Gabor feature \(t_{\Theta, \lambda, \varphi_n}(x', y')\) and

\[
M_{\Theta, \lambda}(x, y) = \max\{M_{\Theta, \lambda, n}(x, y) \mid n = -3 \ldots 2\}
\]

(10)

The quantities \(M_{\Theta, \lambda, n}(x, y), n = -3 \ldots 2\), are related to the activities of simple cells with symmetric receptive fields along a line segment of length \(3\lambda\) passing through point \((x, y)\) in orientation \(\Theta\). This segment is divided in intervals of length \(\frac{1}{2}\lambda\) and the maximum activity of one sort of simple cell operator, centre-on (\(\varphi_n = 0\)) or centre-off (\(\varphi_n = \pi\)), is determined in each interval. \(M_{\Theta, \lambda, -3}(x, y)\), for instance, is the maximum activity of centre-on simple cell operators in the corresponding interval of length \(\frac{1}{2}\lambda\); \(M_{\Theta, \lambda, -2}(x, y)\) is the maximum activity of centre-off simple cells in the adjacent interval, etc. Centre-on and centre-off simple cell activities are alternately used in consecutive intervals. \(M_{\Theta, \lambda}(x, y)\) is the maximum among the above interval maxima.

Roughly speaking, the concerned grating cell subunit will be activated if centre-on and centre-off cells of the same preferred orientation \(\Theta\) and spatial frequency \(\frac{1}{2}\lambda\) are alternately activated in intervals of length \(\frac{1}{4}\lambda\) along a line segment of length \(3\lambda\) centred on point \((x, y)\) and passing in direction \(\Theta\). This will, for instance, be the case if three parallel bars with spacing \(\lambda\) and orientation \(\Theta\) of the normal to them are encountered. In contrast, the condition is not fulfilled by the simple cell activity pattern caused by a single bar or two bars, only.

In the next, second stage of the model, the response \(w_{\Theta, \lambda}(x, y)\) of a grating cell operator is computed by weighted summation of the outputs of the grating subunits.

\[
w_{\xi, \eta, \Theta, \lambda} = \int_{\Omega} e^{-\frac{(x' - \xi)^2 + (y' - \eta)^2}{2\beta^2}} (q_{\xi', \eta', \Theta, \lambda} + q_{\xi', \eta', \Theta + \pi, \lambda}) \, d\xi' \, d\eta', \\
\Theta \in [0, \pi]
\]

(11)

The parameter \(\beta\) determines the size of the area over which effective summation takes place. A value of \(\beta = 5\) is used
in our experiments. For further details on the operator we refer to [15].

In our experiments we use a set of grating cell operators with eight different preferred orientations and three preferred spatial-frequencies, yielding a vector of 24 features in each point of the image. This new non-linear filter bank is derived from the thresholded symmetric Gabor filter bank and has the same coverage of the spatial-frequency domain.

4. Comparison of texture features

The feature vectors computed in different points of a texture image using a given operator are not identical; they rather form a cluster in the multi-dimensional feature space. The larger the distance between two clusters which correspond to two different types of texture, the better the discrimination properties of the texture operator that produced the feature vectors. This distance has, however, to be considered in relation to the size of the clusters. In order to determine the distance between two clusters of feature vectors, it is sufficient to look at their projection on a line, under the assumption that this projection maximises the separability of the clusters in the one-dimensional space. A linear transformation that realizes this projection was first introduced by Fisher [6] and is called the Fisher’s linear discriminant function. It has the following form: 

\[ y = (\mu_1 - \mu_2)^T S^{-1} \tilde{x} \]

where \( \mu_1 \) and \( \mu_2 \) are the means of the two clusters and \( S^{-1} \) is the inverse of the pooled covariance matrix.

The projection of the feature vectors onto the projection line maximises the so-called Fisher criterion (see e.g. [8])

\[ f = \frac{1}{\sigma^2} \ln \frac{\sigma_1}{\sigma_2} \]

where \( \sigma_1 \) and \( \sigma_2 \) are the standard deviations of the distributions of the projected feature vectors of the respective clusters and \( \eta_1 \) and \( \eta_2 \) are the projections of the means \( \mu_1 \) and \( \mu_2 \), respectively. The Fisher criterion thus expresses the distance between two clusters relative to their compactness in one single quantity.

We evaluated the performance of the various operators presented in the previous section according to the Fisher criterion by looking at the separability of nine test textures (Fig.4). The separability was measured in the following way: the pooled covariance matrix was calculated for each pair of images using 1000 sample feature vectors from each image. Then the feature vectors were projected on a line using Fisher’s linear discriminant function and the Fisher criterion was evaluated in the projection space.

Essential statistics, such as minimum, maximum and average, of the Fisher criterion values computed for the different types of features are given in Table 1.

As can be seen from the table, the values for the symmetric and the antisymmetric features are approximately the same. The features do not distinguish between the texture pairs very well: the projected feature vector distributions overlap extensively, as can be concluded from the small values of the Fisher criterion (maximum 0.20).

With a mean value of 1.22 and 1.25 for the symmetric and antisymmetric thresholded Gabor features, respectively, the discrimination properties of the nonlinear features are better than those of the linear Gabor features, though the projected clusters still overlap.

In case of the Gabor-energy features, the values are all much higher as compared to the linear and the thresholded Gabor features. The maximum value listed is 12.89 meaning that there is virtually no overlap between the two corresponding clusters.

As evidenced by the values in Table 1, the separability which is achieved by the complex moment features is smaller than the one achieved on the basis of the Gabor-energy features. This can be explained by the fact that calculation of the complex moments is a linear transformation of the local power spectrum. The Fisher criterion values will therefore be equal to or smaller than the values which are directly based on the local power spectrum itself.

For any pair of texture images, the inter-cluster distance computed using the grating cell operator features is considerably greater than the inter-cluster distance computed with any of the other operators. The minimum value of the Fisher criterion is 5.44 (achieved for the pair of textures T3 and T7); two Gaussian distributions of equal standard deviation for which a Fisher criterion of 5.44 is computed, overlap for 0.01%. This means that the probability of misclassification is negligibly small.
5. Summary and conclusions

In this paper, a quantitative comparison was made of a number of texture features which are all based on the use of Gabor filters with a different type and degree of ‘post-Gabor’ processing. The following features were compared: linear Gabor features (both symmetric and antisymmetric), thresholded Gabor features (again both symmetric and antisymmetric), Gabor-energy features, complex moments features, and grating cell operator features.

To assess the discrimination properties of the various features, a quantitative method was used which is based on the Fisher criterion and the Fisher linear discriminant. This method differs from the often used CRC method in that it does not mix the discrimination properties of the concerned features with the performance of a subsequent feature vector classifier which is employed in the CRC method.

The values of the Fisher criterion show clear differences between the various features with respect to the discrimination properties. In general, the discrimination effectiveness increases with the amount of post-Gabor processing, as can be seen in Table 1. The only exception from this rule are the complex moment features, which perform worse than the Gabor-energy features, from which they are derived. The best performance is achieved with the grating cell operator features. This result is in accordance with the results previously obtained in a comparison of the grating cell operator features with commonly used texture features, such as those extracted from cooccurrence matrices [11, 12].

Another interesting finding is that, while only thresholded Gabor features which are based on symmetric Gabor functions were used previously in the literature, filters based on antisymmetric Gabor functions show comparable performance.

References


<table>
<thead>
<tr>
<th>Feature type</th>
<th>Avg</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear sym. Gabor</td>
<td>0.16</td>
<td>0.11</td>
<td>0.20</td>
</tr>
<tr>
<td>Linear asym. Gabor</td>
<td>0.16</td>
<td>0.14</td>
<td>0.20</td>
</tr>
<tr>
<td>Thresh. sym. Gabor</td>
<td>1.22</td>
<td>0.60</td>
<td>1.83</td>
</tr>
<tr>
<td>Thresh. asym. Gabor</td>
<td>1.25</td>
<td>0.64</td>
<td>1.88</td>
</tr>
<tr>
<td>Gabor energy</td>
<td>6.33</td>
<td>2.35</td>
<td>12.89</td>
</tr>
<tr>
<td>Complex moments</td>
<td>2.69</td>
<td>0.65</td>
<td>4.84</td>
</tr>
<tr>
<td>Grating cell operator</td>
<td>14.02</td>
<td>5.44</td>
<td>31.62</td>
</tr>
</tbody>
</table>

Table 1. Statistics of the Fisher criterion values.