Fast and accurate computation of orthogonal moments for texture analysis

Cecilia Di Ruberto\textsuperscript{1}, Lorenzo Putzu\textsuperscript{2} and Giuseppe Rodriguez\textsuperscript{1}

\textsuperscript{1}Department of Mathematics and Computer Science, University of Cagliari, via Ospedale 72, 09124 Cagliari, Italy
\textsuperscript{2}Department of Electrical and Electronic Engineering, University of Cagliari, piazza d’Armi, 09123 Cagliari, Italy

May 3, 2018

Abstract

In this work we describe a fast and stable algorithm for the computation of the orthogonal moments of an image. Indeed, orthogonal moments are characterized by a high discriminative power, but some of their possible formulations are characterized by a large computational complexity, which limits their real-time application. This paper describes in detail an approach based on recurrence relations, and proposes an optimized Matlab implementation of the corresponding computational procedure, aiming to solve the above limitations and put at the community’s disposal an efficient and easy to use software. In our experiments we evaluate the effectiveness of the recurrence formulation, as well as its performance for the reconstruction task, in comparison to the closed form representation, often used in the literature. The results show a sensible reduction in the computational complexity, together with a greater accuracy in reconstruction. In order to assess and compare the accuracy of the computed moments in texture analysis, we perform classification experiments on six well-known databases of texture images. Again, the recurrence formulation performs better in classification than the closed form representation. More importantly, if computed from the GLCM of the image using the proposed stable procedure, the orthogonal moments outperform in some situations some of the most diffused state-of-the-art descriptors for texture classification.
1 Introduction

Many texture feature extraction methods have been reported in the literature. Among them, the statistical methods compute simple statistical properties (e.g., first order parameters) or sophisticated properties (e.g., co-occurrence matrices [12, 13, 31], or gray level run length [6]) of images. Other texture feature extraction approaches are based on models (Markov random field and fractals, or local binary pattern [28]), mathematical morphology [35], or classical transform methods (Fourier, Gabor, wavelet transforms [24]).

The application of a texture analysis method consists of extracting a set of parameters from an image, in order to characterize the texture it contains [7]. Each parameter expresses a special property, such as coarseness, homogeneity, or the local contrast. The texture feature extraction can be carried out either at the pixel level, by calculating the textural parameters on a very small neighborhood, or at the level of a region of interest defined by the user, which may correspond to a large number of pixels.

Moments are statistical measures which can be used to obtain relevant information on an object. Since Hu introduced them in image analysis [17], moment functions have been widely used as discriminative descriptors in image processing and pattern classification applications. In 1994, Tuceryan discussed texture feature extraction and texture analysis based on geometric moments [40]. Begum applied complex moments to texture segmentation [3]. However, both geometric and complex moments contain redundant information and are sensitive to noise. This is due to the fact that the associated kernel polynomials are not orthogonal. Many authors proposed the use of orthogonal moments: Mukundan adopted the discrete Chebyshev moments [26], Yap studied another set of discrete moments, known as Krawtchouk moments [43], Teague suggested the use of Legendre and Zernike moments [39].

Orthogonal moments are shown to be less sensitive to noise and have an efficient capability of feature representation. They allow to reconstruct the image intensity function analytically, from a finite set of moments, using the inverse moment transform. Legendre and Zernike moments are most widely used because of their minimum redundancy. Indeed, they can represent the properties of an image with no redundancy or overlap of information between the moments [18]. Because of these important features, Zernike moments have been widely used in different types of applications [9, 21]. They have
been utilized in shape-based image retrieval [10], edge detection [22], and as a feature set in pattern recognition [14]. In [16], the authors apply Zernike moment features to retrieve binary and gray level images from MPEG-7 and COIL-20 dataset, respectively, showing their suitability for image retrieval, due to their rotation invariance. A mixture of ternary patterns and Zernike moments has been proposed to provide an efficient image classification which is robust against illumination and rotation [27].

Unfortunately, these approaches are often characterized by a large computational complexity, making them unsuitable for real-time applications. This lead many researchers to develop faster algorithms for computing Zernike moments. In [38], the authors propose a CAD system for the diagnosis of breast masses in mammography images which uses Zernike moments for extracting the shape and margin properties of the masses. In [30], an image annotation system, based on different type of moments, has been developed to allow searching image databases. The experimental results showed that the annotation system coupling Legendre moments to Bayesian networks gives good results for images that are well and properly segmented.

In [42], a moment-based approach is proposed for texture analysis of medical images, namely, CT liver scan and prostate ultrasound. The neighborhood of a texture pixel is calculated by different moments for texture feature extraction. After being verified on Brodatz textures, the moment-based texture analysis method is applied to CT liver scan classification and prostate ultrasound segmentation. A support vector machine and a multi-channel active contour model are used in this application. The results show that the proposed method is promising, but still have some limitations. Local binary pattern and Legendre moments have been used as features for CT liver images classification [11]. In [23], a new set of rotation and scale invariants of Legendre moments is introduced, achieving good results in image classification experiments. Lakshmi et al. [20] used Legendre moments for palm print authentication with a very good prediction accuracy.

The computation of Legendre moments is in general a time consuming process. In many references [30, 42, 11, 23], their computation has been performed using closed form representations for orthogonal polynomials, and taking little care to the accuracy of the quadrature formulas used to approximate integrals. In this work, we describe a fast and stable algorithm for the computation of the orthogonal moments of an image with respect to both a continuous and a discrete inner product, based on classical recurrence relations for orthonormal polynomials. Despite their recurrent structure, such relations can be coded without recursive calls, allowing us to develop an efficient and accurate Matlab toolbox. We propose this software as a standard tool for orthogonal moments computation and we make it freely available to
the scientific community. The reconstruction performance of the proposed methods and their discriminative ability in texture classification are investigated in comparison to orthogonal moments computed from a closed form representation, and to other classification methods. We finally discuss the effect of weighted orthogonal moments on the processing of two particular datasets.

Our numerical experiments confirm the well known fact that, while recurrence relations for orthogonal polynomials produce reliable results, closed form representations should be avoided for moments computation. Moreover, computing moments from Gray Level Co-occurrence Matrices (GLCM) produces a more accurate classification and, at least in some situation, weighted moments may lead to a slight performance improvement. Orthogonal moments from GLCM appear to be competitive with state-of-the-art approaches based on Convolutional Neural Networks (CNN). We give a possible interpretation for the different performance of orthogonal moments and CNN when applied to datasets with specific features.

The paper is organized as follows. In Section 2 we introduce the definition and properties of Legendre moments. The algorithms for the computation of orthogonal moments contained in the Matlab toolbox are described in Section 3. Section 4 presents the results of numerical experiments assessing the reconstruction performance and the discriminative accuracy in the texture analysis of six dataset collections. Finally, in Section 5 we summarize the content of the paper and describe our plans for future work.

2 Orthogonal moments

The use of moments for image analysis and pattern recognition was inspired by Hu [17]. Among all types of moments, orthogonal moments present the peculiar property of being characterized both by small information redundancy and by high discriminative power. A representative family of orthogonal moments is the well known Legendre moments. They were first introduced in image analysis by Teague [39] and have been exhaustively used in many pattern recognition and features extraction applications, due to their invariance to scale, rotation, and reflection change [5, 23].

Here, we briefly recall some definitions and properties of orthogonal polynomials; see [11]. Let us consider the space $L^2_{\omega}[-1,1]$ endowed with the weighted inner product

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)w(x) \, dx \quad (2.1)$$
and the induced norm \( \|f\| = \sqrt{\langle f, f \rangle} \), being \( w(x) \geq 0 \) a weight function. Two functions \( f, g \in L^2_w[-1, 1] \) are said to be orthogonal if \( \langle f, g \rangle = 0 \); they are orthonormal if, additionally, both functions have unitary norm.

A fundamental theorem states that there exists a unique infinite sequence of orthogonal polynomials, that is, polynomials \( p_k(x) \) of degree \( k \), with \( k = 0, 1, \ldots \), such that \( \langle p_k, p_\ell \rangle = 0 \) whenever \( k \neq \ell \). Orthogonal polynomials may be scaled in different ways, i.e., so that they are monic (the monomial of maximum degree has unitary coefficient) or so that they are orthonormal. We will adopt the latter normalization.

Monic orthogonal polynomials associated to inner product (2.1) are defined by the following three-term recurrence relation

\[
\begin{cases}
\tilde{p}_{k+1}(x) = (x - \alpha_k)\tilde{p}_k(x) - \beta_k\tilde{p}_{k-1}(x), \\
\tilde{p}_0(x) = 1, \quad \tilde{p}_{-1}(x) = 0,
\end{cases}
\]

for \( k = 0, 1, \ldots \), where \( \tilde{p}_{-1}(x) \) serves the only purpose of starting the recursion,

\[
\alpha_k = \frac{\langle x\tilde{p}_k, \tilde{p}_k \rangle}{\langle \tilde{p}_k, \tilde{p}_k \rangle}, \quad k = 0, 1, \ldots, \quad \beta_k = \frac{\langle \tilde{p}_k, \tilde{p}_k \rangle}{\langle \tilde{p}_{k-1}, \tilde{p}_{k-1} \rangle}, \quad k = 1, 2, \ldots,
\]

and \( \beta_0 = \langle \tilde{p}_0, \tilde{p}_0 \rangle \). This process is often referred to as the Stieltjes procedure [11]. Polynomials which are orthonormal with respect to (2.1), that is \( p_k(x) = \tilde{p}_k(x)/\|\tilde{p}_k(x)\| \) for \( k = 0, 1, \ldots \), satisfy the three-term recurrence

\[
\begin{cases}
\sqrt{\beta_{k+1}}p_{k+1}(x) = (x - \alpha_k)p_k(x) - \sqrt{\beta_k}p_{k-1}(x), \\
p_0(x) = \sqrt{\beta_0}, \quad p_{-1}(x) = 0,
\end{cases}
\]

with the same coefficients \( \alpha_k \) and \( \beta_k \) used in (2.2).

Legendre polynomials correspond to the weight function \( w(x) = 1 \). In this case,

\[
\alpha_k = 0, \quad k = 0, 1, \ldots, \quad \beta_k = \frac{k^2}{4k^2 - 1}, \quad k = 1, 2, \ldots, \quad \beta_0 = 2.
\]

Second kind Chebyshev polynomials are obtained by setting \( w(x) = \sqrt{1 - x^2} \). This weight function emphasizes the importance of the central part of the interval \([-1, 1]\) with respect to a neighborhood of the endpoints. The corresponding recurrence coefficients are

\[
\alpha_k = 0, \quad k = 0, 1, \ldots, \quad \beta_k = \frac{1}{4}, \quad k = 1, 2, \ldots, \quad \beta_0 = \frac{\pi}{2}.
\]
An image can be represented as the discretization of an intensity function $f(x, y)$ on the continuous domain $[-1, 1]^2 := [-1, 1] \times [-1, 1]$. In such case, one may define bivariate polynomials as products of univariate orthogonal polynomials

$$P_{k, \ell}(x, y) = p_k(x)p_\ell(y).$$

Such polynomials are themselves orthonormal, in the sense that

$$\langle P_{k, \ell}, P_{r, s} \rangle = \int_{-1}^{1} \int_{-1}^{1} P_{k, \ell}(x, y)P_{r, s}(x, y)w(x)w(y)\,dx\,dy = \delta_{kr}\delta_{\ell s},$$

where $\delta_{ij}$ is the Kronecker symbol, that takes the value 1 when $i = j$ and 0 when $i \neq j$.

The orthogonal weighted moments of order $q$ of an intensity function $f(x, y)$ are defined as

$$\mu_{i, q-i} = \int_{-1}^{1} \int_{-1}^{1} p_i(x)p_{q-i}(y)f(x, y)w(x)w(y)\,dx\,dy, \quad i = 0, 1, \ldots, q. \quad (2.5)$$

The orthogonality of the basis implies that there is no redundancy or overlapping of information between moments. The knowledge of the moments, under suitable regularity conditions on the intensity function, allows one to approximate $f(x, y)$ in the least squares sense \[33\] in the form

$$f_{mn}(x, y) = \sum_{i=0}^{m} \sum_{j=0}^{n} \mu_{ij} p_i(x)p_j(y). \quad (2.6)$$

This means that

$$\|f - f_{mn}\|_2 = \min_{p \in \Pi_{m,x} \otimes \Pi_{n,y}} \|f - p\|_2,$$

where $\Pi_{m,x}$ is the vector space of polynomials in the variable $x$ having degree lesser or equal than $m$, and $\otimes$ denotes the tensor product of linear spaces.

Orthogonal moments \[2.5\] must be computed by a suitably accurate quadrature formula. Given the large number of pixels usually available in an image, we employed a simple Cartesian product rule based on Simpson’s composite formula \[36\]. Assuming the image has size $M \times N$, with both $M$ and $N$ odd, we denote by

$$x_k = \frac{2k - (M + 1)}{M - 1}, \quad y_\ell = \frac{2\ell - (N + 1)}{N - 1}, \quad (2.7)$$
(\(k = 1, \ldots, M\) and \(\ell = 1, \ldots, N\)) the normalized pixel coordinates in the
interval \([-1, 1]\) and by \(f(x_k, y_\ell)\) the intensity of the pixel at row \(k\) and column
\(\ell\). Then, the moments are approximated as follows

\[
\mu_{ij} \approx K \sum_{k=1}^{M} \sum_{\ell=1}^{N} c_{M,k} c_{N,\ell} p_i(x_k)p_j(y_\ell)f(x_k, y_\ell),
\tag{2.8}
\]

where \(K = 4/(9MN)\) and the quadrature weights are defined by

\[
c_{M,i} = \begin{cases} 
1, & i = 1, M, \\
4, & i = 2, 4, \ldots, M - 1, \\
2, & i = 3, 5, \ldots, M - 2.
\end{cases}
\]

In some situations, it may be preferable to substitute (2.1) by the following
discrete inner product

\[
\langle f, g \rangle = \sum_{k=1}^{M} f(t_k)g(t_k),
\tag{2.9}
\]

where \(\{t_k\}\) is a discretization of a given interval \([a, b]\). If the interval \([0, M - 1]\)
is discretized by the points \(t_k = k - 1, k = 1, \ldots, M\), the polynomials \(p_i^{(M)}(x)\),
ortonormal on \([0, M - 1]\) with respect to (2.9), are generally referred to as
discrete Chebyshev polynomials. They can be obtained by substituting in
(2.3) the recurrence coefficients

\[
\alpha_k = \frac{M - 1}{2}, \quad \beta_k = \frac{M^2}{4} \left(1 - \frac{k}{M}\right)^2, \quad k = 1, 2, \ldots, \quad \beta_0 = M.
\]

In such a case, all the theory is unchanged, except the continuous moments
(2.5) are substituted by the discrete moments

\[
\mu_{i,q-i} = \sum_{k=1}^{M} \sum_{\ell=1}^{N} f_{k\ell} p_i^{(M)}(k - 1)p_{q-i}^{(N)}(\ell - 1), \quad i = 0, 1, \ldots, q,
\tag{2.10}
\]

where \(f_{k\ell}\) is the intensity of the pixel at position \((k, \ell)\).

3 Computation of orthogonal moments

The generation of orthogonal moments for an image can be a time consuming
process. Indeed, formula (2.8) must be computed for each required moment.
The summation involves order $MN$ floating-point operations, once both the polynomial bases have been evaluated at the points (2.7). If the image is square ($M = N$) and the user sets $m = n$, only one basis suffices, nevertheless the moments computation speed is strongly influenced by the efficiency in the polynomial bases evaluation. Using recurrence relations (2.3), the number of operations for the evaluation of each basis is linear in the number of pixels on each side of the image.

Additionally, it is essential to guarantee a sufficient accuracy in the computation, since errors in the approximation of the moments may affect their discriminative ability. In some previous work, the moments have been approximated by considering in (2.8) constant quadrature weights

$$c_{M,k} = c_{N,\ell} = 1 \quad \text{and} \quad K = \frac{(2i - 1)(2j - 1)}{MN}. \quad (3.1)$$

This simplification greatly degrades the quality of the approximation of the moments, without any advantage in terms of complexity.

There are various possible definitions of orthogonal polynomials. Rodrigues’ formula expresses them as derivatives of simple polynomials, which is not useful in this setting. Some authors adopted explicit expansions, like the following which holds for Legendre polynomials,

$$p_i(x) = 2^i \sum_{k=0}^{i} \binom{i}{k} \left(\frac{i+k-1}{2i}\right)x^k. \quad (3.2)$$

Such representations should be avoided because, as it is well known, the canonical polynomial basis $\{1, x, x^2, \ldots, x^n\}$ is extremely ill-conditioned [33], in the sense that small perturbations in the coefficients produce large errors in the polynomial evaluation. Moreover, the coefficients in (3.2) have alternating signs, causing cancellation and error propagation when the computation is performed on a computer, and their implementation requires a large complexity, compared to other approaches.

The Stieltjes procedure (2.3), despite being an implicit definition, is the most effective from the computational point of view, as we will show in the following.

Algorithm 1 implements a simple function opcoef which returns the recurrence coefficients of a few families of orthogonal polynomials. At the moment, only the families mentioned in Section 2 are supported, but the function can be easily extended. The operations involving vectors (displayed as bold lower case letters) in lines 4 and 10 of Algorithm 1 are to be performed component by component. This implies, in Matlab, the use of the array operators .*, ./, and .^, in place of the usual matrix operators.
Algorithm 1 \([\alpha, \beta] = \text{opcoef}(\text{type}, n, M)\) recurrence coefficients of orthogonal polynomials

**Require:** orthogonal polynomial type, maximum degree \(n\), \(M\) number of discretization points (only for discrete Chebyshev polynomials)

**Ensure:** \(\alpha, \beta \in \mathbb{R}^{n+1}\) vectors of coefficients for the orthogonal polynomials

1. \(v = (1, 2, \ldots, n+1)\)
2. if type='Legendre' then
   3. \(\alpha = (0, \ldots, 0)\)
   4. \(\beta = \left(2, v^2/(4v^2 - 1)\right)\) (componentwise)
   5. else if type='Chebyshev second kind' then
      6. \(\alpha = (0, \ldots, 0)\)
      7. \(\beta = \left(\pi^2/4, \frac{1}{4}, \ldots, \frac{1}{4}\right)\)
   8. else if type='discrete Chebyshev' then
      9. \(\alpha = \frac{M-1}{2}(1, \ldots, 1)\)
      10. \(\beta = \left(M, \frac{M^2}{4}(1 - (v/M)^2)/(4 - 1/v^2)\right)\) (componentwise)
  11. end if

Algorithm 2 \(P = \text{opevmat}(\alpha, \beta, x)\) evaluation of orthogonal polynomials

**Require:** \(\alpha, \beta \in \mathbb{R}^{n+1}\) vectors of coefficients for the orthogonal polynomials, \(x \in \mathbb{R}^m\) vector of evaluation points

**Ensure:** \(P \in \mathbb{R}^{m \times (n+1)}\) matrix whose column \(P_{:,j}\) is the evaluation of the \((j-1)\)th orthogonal polynomial at the components of \(x\)

1. \(P_{:,1} = (1, \ldots, 1)^T/\sqrt{\beta_1}\)
2. \(P_{:,2} = (x - \alpha_1) \ast P_{:,1}\) (componentwise)
3. \(P_{:,2} = P_{:,2}/\sqrt{\beta_2}\)
4. for \(k = 2, \ldots, n\) do
   5. \(P_{:,k+1} = (x - \alpha_k) \ast P_{:,k} - \sqrt{\beta_k} P_{:,k-1}\) (componentwise)
   6. \(P_{:,k+1} = P_{:,k+1}/\sqrt{\beta_{k+1}}\)
5. end for
Once the recursion coefficients have been assigned, the evaluation of an orthogonal polynomials basis at a vector of points is performed by the function \texttt{opevmat} in Algorithm 2 which employs the Stieltjes procedure. Again, vector multiplications in lines 2 and 5 are intended component by component.

**Algorithm 3** $M = \text{legmoms}(F,q)$ computation of Legendre moments

**Require:** $F \in \mathbb{R}^{M \times N}$ grayscale digital image, $q$ order of the moments to be computed

**Ensure:** $M \in \mathbb{R}^{(q+1) \times (q+1)}$ matrix of Legendre moments

1: \textbf{if} $M$ is even \textbf{then} $M = M - 1$ \textbf{end if}
2: \textbf{if} $N$ is even \textbf{then} $N = N - 1$ \textbf{end if}
3: $h_x = 2/M, \; x = (-1, -1 + h_x, -1 + 2h_x, \ldots, 1)^T$
4: $h_y = 2/N, \; y = (-1, -1 + h_y, -1 + 2h_y, \ldots, 1)^T$
5: \hspace{0.5cm} $[\alpha, \beta] = \text{opcoef}(\text{’Legendre’,}q)$
6: \hspace{0.5cm} $P_1 = \text{opevmat}(\alpha, \beta, x)$
7: \hspace{0.5cm} $P_2 = \text{opevmat}(\alpha, \beta, y)$
8: \hspace{0.5cm} \textbf{for} $i = 1, \ldots, q + 1$ \textbf{do}
9: \hspace{1cm} \hspace{0.5cm} \textbf{for} $j = 1, \ldots, q + 1$ \textbf{do}
10: \hspace{1.5cm} compute $M_{i-1,j-1}$ by applying Simpson’s rule \textbf{2.8} to integral \textbf{2.5}
11: \hspace{1cm} \hspace{0.5cm} \textbf{end for}
12: \hspace{0.5cm} \textbf{end for}

Algorithm 3 describes the function \texttt{legmoms} which approximates the orthogonal moments \textbf{2.5} with respect to Legendre polynomials by the Cartesian product Simpson’s rule \textbf{2.8}. The approximation of the orthogonal moments with respect to Chebyshev polynomials of the second kind is performed by a very similar function \texttt{cheb2moms}, which we do not report here for the sake of brevity.

**Algorithm 4** $M = \text{dchebmoms}(F,q)$ comp. of discrete Chebyshev moments

**Require:** $F \in \mathbb{R}^{M \times N}$ digital grayscale image, $q$ order of the moments to be computed

**Ensure:** $M \in \mathbb{R}^{(q+1) \times (q+1)}$ matrix of Legendre moments

1: $x = (0, 1, \ldots, M - 1)^T$
2: $y = (0, 1, \ldots, N - 1)^T$
3: $[\alpha_1, \beta_1] = \text{opcoef}(\text{’discrete Chebyshev’,}q,M)$
4: $[\alpha_2, \beta_2] = \text{opcoef}(\text{’discrete Chebyshev’,}q,N)$
5: $P_1 = \text{opevmat}(\alpha_1, \beta_1, x)$
6: $P_2 = \text{opevmat}(\alpha_2, \beta_2, y)$
7: $M = P_1^T * F * P_2$ (matrix product)
Finally, the exact computation of the discrete orthogonal moments (2.10) is described in the function \texttt{dchebmoms} of Algorithm 4.

All the above algorithms have been implemented in the Matlab programming language and are available from the authors’ web sites; see the \texttt{orthomoms} package at \url{http://bugs.unica.it/cana/software}. The software have been verified to be compatible with Octave.

4 Experimental results

In this section we present two sets of numerical experiments carried out in order to verify the effectiveness of the proposed algorithm. The experiments were performed on a dual Xeon CPU E5-2620 system (12 cores), running the Debian GNU/Linux operating system and Matlab 9.2.

In the first experiment, we investigate the reconstruction performance of the above described approach for computing Legendre moments. Such an experiment is useful to ascertain the correctness of the computation and the advantages, both in terms of execution time and of accuracy, of the recurrence relations with respect to the closed form representation of Legendre polynomials.

In the second experiment, the descriptive ability of the computed moments is tested in texture classification problems.

4.1 Image reconstruction

In order to verify the correctness of the computation on a synthetic dataset, we consider the following smooth model intensity function

\[
f(x, y) = \frac{1}{2} e^x \sin(\pi x) \sin(\pi y), \quad (x, y) \in [-1, 1]^2.
\]

The function is sampled on a regular grid of 1023 \times 1023 points on the square \([-1, 1]^2\), then its values are translated and rescaled so that they are contained in the interval \([0, 1]\). The graph of the function and the resulting grayscale image \(F\) are displayed in Figure 1.

We first evaluate the moments \(\mu_{ij}, i, j = 0, \ldots, n\), by Simpson’s rule (2.8), where the orthogonal polynomials are computed using the three-term recursion (2.3). Then the least-squares approximation \(f_{nn}(x, y)\) of the model intensity function \(f(x, y)\) is evaluated by formula (2.6) on the grid (2.7), as well as the relative error

\[
E_n(f) = \frac{\max_{k,\ell} |f(x_k, y_\ell) - f_{nn}(x_k, y_\ell)|}{\max_{k,\ell} |f(x_k, y_\ell)|}.
\]
We perform different tests, letting the number of moments $n = 5, 10, \ldots, 50$. In Figure 2 we compare the execution time and the relative error in the reconstruction, obtained by the above approach, to the results obtained by evaluating the Legendre polynomials using the closed form representation (3.2) and approximating the moments by employing in (2.8) the quadrature weights (3.1).

The graph on the left of Figure 2 shows the great reduction in the computational complexity of the approach based on recurrence relations. The reconstruction error is displayed on the right. The error deriving from the closed form representation increases with the number of moments, because of error propagation and lack of precision in the approximation of the integrals.
The computation performed by recursion formulas appears to be much more stable. The accuracy improves up to 15 moments, then it degrades, but the relative error stays below $10^{-4}$. This degradation is probably due to the fact that Simpson’s rule is not able to approximate with sufficient accuracy the integral of fast oscillating functions, like large degree orthogonal polynomials are.

The same figure displays also the results obtained by discrete Chebyshev polynomials and the corresponding discrete moments (2.10). The execution time is equivalent to Legendre polynomials, indeed the two approaches use the same computational scheme with different recursion coefficients. The accuracy improves up to 20 moments, where it reaches machine accuracy. This is somehow expected, as in this case the moments are computed exactly and not obtained by approximating an integral by a quadrature formula. Nevertheless, the fact that the accuracy stays at the same level when $n > 20$ clearly testifies that error propagation is negligible when recursion formulas (2.3) are used.

The two families of orthogonal polynomials used in the experiments are depicted in Figure 3.

### 4.2 Texture analysis

To evaluate and compare the performance in texture analysis of the analyzed algorithms for computing orthogonal moments, we used six different databases of texture images: Brodatz, Mondial Marmi, Outex, Vectorial, Kylberg Sintorn and ALOT; see Figure 4. They present various materials and textures representing different classification problems, and include
Figure 4: Each column displays images from the six texture databases used in the experiments. Starting from the left: Brodatz, Mondial Marmi, Outex, Vectorial, Kylberg Sintorn, and ALOT.

hardware-rotated images taken at 4–9 different orientations, making them the most suitable for our experiments.

The only database that contains software rotated images is Vectorial, a collection of 20 artificial texture classes proposed by Bianconi [2]. As it can be guessed from the name, these images are not raster, thus the rotation does not affect the image structure. They have been rotated with angular steps of 10° (0°, 10°, . . . , 90°) and then converted into raster with a resolution of 300 dpi. Finally, each image has been subdivided into 16 sub-images of size 255 × 255, resulting in 16 samples per class and 3200 total images. We did not consider other databases which include software rotated images, since
this operation may modify the original image structure, leading to wrong results.

The other five databases present real texture images. The first one belongs to one of the most diffused collection of textures, that is, Brodatz’s album. We did not use the whole collection, since the original images included in the album are not rotated, but just a subset of 13 textures. This subset has been proposed by Bianconi [2], who acquired hardware-rotated images directly from the original book. The images have been acquired with angular steps of 10°, like in the previous case. Then, each image has been subdivided into 16 sub-images of size 205 × 205, resulting in 2080 total samples. Brodatz’s album is one of the oldest texture database. Indeed, among the tested datasets it is the only one composed by gray level images.

The remaining databases comprise color images, thus they have been converted into gray scale before the feature extraction step. Mondial Marmi is a free image database of granite tiles for color and texture analysis, that includes 12 granite classes. Each texture is captured in a 24-bit RGB image of size 544 × 544 using nine rotation angles (0°, 5°, 10°, 15°, 30°, 45°, 60°, 75°, 90°). The database presents 4 images for each class and for each angle, so the original images have been subdivided into four non-overlapping sub-images of size 272 × 272 for a total image count of 1780.

The Outex database is a collection of 320 textures (both macrotextures and microtextures) acquired with well defined variations in terms of illumination, rotation, and spatial resolution. Each texture is captured in a 24-bit RGB image of size 538 × 746, using three different simulated illuminations, six spatial resolutions (100, 120, 300, 360, 500, and 600 dpi), and nine rotation angles (0°, 5°, 10°, 15°, 30°, 45°, 60°, 75°, 90°). Hence the current texture database includes 51840 images. Given the considerable size of this database, we focused our experiments on a smaller subset, taking inspiration by the test suite proposed by the author of the Outex database himself, called OUTEX00045, that uses 45 texture classes. Following the instructions proposed in Outex, the original images have been divided in 20 non overlapping sub-images with size 128 × 128, for a total count of 8100 images.

The Kylberg Sintorn database is a collection of 25 textural classes of materials, such as fabric, grains, sugar, rice, etc. As in the other databases, the images are provided with nine rotation angles, but in this case the images have been rotated with angular steps of 40° (0°, 40°, . . . , 320°). The original images (one for each class) are 24-bit RGB with a resolution of 5184 × 3456 pixels, but they have been provided also in small subsets for texture classification, presenting 400 images for each angle and thus 16 samples per class. The final dataset contains, therefore, 3600 images.

The last database, ALOT, is very different from the previous ones. In-
Indeed, it presents images acquired with just four rotation angles with steps of 60° (0°, 60°, 120°, 180°). The original image database contains 250 textures, each one with 100 images obtained under different illumination conditions. For our experiments we considered a subset of the original dataset containing 80 textures. The original images have been divided into 16 sub-images of size 181 × 181, for a total of 5120 images.

To compare the recurrent formulation of orthogonal moments to the one based on a closed form expansion, we performed a set of image classification experiments, evaluating both performance and robustness of the proposed descriptors against image rotation. For each dataset, we executed 100 experiments; for each of them, both the training and the test sets are represented by half of the original samples. Each dataset was divided by a stratified sampling, which guarantees that each class is properly represented in both the training and the testing set. The classification performances have been evaluated by the accuracy index

\[ A = \frac{TP + TN}{TP + TN + FP + FN}, \]  

(4.1)

where TP (TN) and FP (FN) represent the number of true positive (negative) classifications, and false positive (negative) classifications, respectively. This index gives a meaningful indication of the performance, since it considers each class of equal importance.

The classification accuracy has been estimated by a k-Nearest Neighbor (k-NN) classifier, with \( k = 1 \), computed using the Euclidean distance. The k-NN has been preferred to a more complex classifier in order to document the effectiveness of the moments as features descriptors, rather than assessing the performance of the classifier itself. To better study the effects of image rotation, the classifier is always trained with features extracted from images acquired at orientation 0° and then tested with feature extracted from images acquired at other orientations.

The results of this experiment are reported in Figures 5 and 6. The graphs on the left display the accuracies for the different algorithms with a fixed number of moments (q=3) and varying the orientation of the image; the graphs on the right report the same quantity versus the variation of the number of moments. As it can be observed, the results obtained using the Legendre polynomials and the discrete Chebyshev polynomials computed by recursion formulas present roughly the same accuracy values, and appear to be much more stable than the closed form representation, for what concerns the orientation. Indeed, the computational scheme used by the two approaches based on recurrence relations leads to a better performance for the classification task. The same trend can be observed also when the num-
Figure 5: Brodatz, Mondial Marmi, and Outex databases: accuracy index \((4.1)\) versus image orientation (left) and number of moments (right).

ber of moments varies; see the graphs on the right in the same figures. In this case, we reported the average accuracy value resulting from all the orientations. The computing time is reported in Figure 7 showing again the superiority of the proposed approach in terms of computational complexity.
Figures 5 and 6 also show that, unlike the reconstruction phase, the performance in classification decreases when the number of moments grows and, therefore, as the related descriptors specialize. Even though the recurrent formulation produces a better performance in classification than the closed
form, it is still insufficient for accurate texture classification. For this reason, we performed a further experiment using an approach already proposed in [8], where the moments were computed starting from a different representation of the images. Indeed, higher level features were created using the Gray Level Co-occurrence Matrices (GLCM) \cite{1,4,15,25} and the Local Binary Pattern (LBP) \cite{28}, since it was observed that using these approaches (in particular, the first one) the invariant moments are more discriminative.

Thus, we computed the moments starting from the GLCM generated with a distance value $d = 1$ and angles $0^\circ, 45^\circ, 90^\circ,$ and $135^\circ$. The final feature vector was created by concatenating the resulting moments. Since the GLCM has been computed from 4 angles, the final feature vector is 4 time larger than the previous one. Indeed, now the feature vector size ranges from 40 to 4324, with a number of moments ranging from 3 to 45, while the size of the feature vector obtained by extracting the moments directly from the images ranges from 10, when the number of moments is 3, to 66, when the number of moments is 10. The results of this comparison are reported in Figures 8 and 9.

The execution time for the whole process of feature extraction is obviously larger, since it also includes the time needed to compute the GLCMs. Anyway, the execution time just increases by a constant value, since the GLCMs have a fixed size, typically $256 \times 256$. Thus, the trend remains essentially the same as depicted in Figure 7. At the same time, the performance in classification greatly improves and the classification accuracy increases with the number of moments. This can be observed in Figures 8 and 9 where

Figure 7: Execution time when the number of moments takes the values $n = 3, 4, \ldots, 10$ (left plot), and with a fixed number of moments ($n = 10$, right plot) and image size varying from $200 \times 200$ up to $2000 \times 2000$. 
Figure 8: Brodatz, Mondial Marmi, and Outex databases: accuracy index versus image orientation (left) and number of moments (right), when the moments are extracted from the GLCM.

we compare the effects on accuracy of image rotation with a fixed number of moments (q=45) (graphs on the left) and of the number of moments (on the right).
Figure 9: Vectorial, Kylberg Sintorn, and ALOT databases: accuracy index \([4.1]\) versus image orientation (left) and number of moments (right), when the moments are extracted from the GLCM.

Second kind Chebyshev (Cheb2) polynomials \([2.4]\) are orthogonal with respect to the inner product \([2.1]\), with \(w(x) = \sqrt{1 - x^2}\). This weight function achieves its maximum at the center of \([-1, 1]\), while it drops to zero
at the endpoints. This means that the computation of the moments (2.5) emphasizes the central part of the image and reduces the importance of its border area. This effect may be useful when classifying images having a particular structure. This fact is outlined in Figure 10, where we display the accuracy index (4.1) versus the number of moments used for classification, for two datasets. The results displayed in the left column are obtained by extracting the moments directly from the images. In this case, the accuracy index for the Brodatz database improves when Cheb2 polynomials are employed, while their performance is much worse than Legendre and discrete Chebyshev polynomials for the Vectorial database. On the contrary, when the moments are extracted from GLCMs, Cheb2 polynomials overall
performance is the best for Vectorial, while it is not very good for Brodatz. We believe that weighted orthogonal moments may have a great potential in image classification, but a much deeper analysis is needed in order to understand when they are preferable to unweighted moments. This will be the subject of future work.

To further highlight the performance of the approaches studied in this paper, we compared them to some widely used state-of-the-art descriptors for texture classification. We computed the rotation invariant GLCM features as proposed in [32], the rotation invariant LBP (LBP-RI) [29] with a $3 \times 3$ pixels neighborhood and distance 1, and the Convolutional Neural Network (CNN) features from three different well known network architectures. The first architecture is AlexNet [19] that gained popularity for its good performance in many classification tasks. It consists of 8 layers, and we extracted the features from the activations of the second last fully connected layer that produces a feature vector of size 4096. The second architecture is the Vgg19Net [34], which is quite similar to AlexNet except for the number of layers being 19. Also in this case, we extracted the features from the activations of the second last fully connected layer that produces a feature vector of size 4096. The third architecture, GoogLeNet [37], is quite different. It presents a stack of 22 layers but, since it uses inception layers, the total number of layers rises to 100. We extracted the features from the activations of the last fully connected layer (the only one that can be used for feature extraction) that leads to a feature vector of size 1000.

As in the previous comparison, we evaluated the effect of image rotation by training the k-NN classifier with features extracted from images acquired at orientation $0^\circ$, and then we tested with features extracted from images acquired at other orientations. To better illustrate the overall performance, we computed the average results for all the angles. The outcome of this experiment is reported in Table 1, where we compare the accuracy of the algorithms investigated in this paper to the state-of-the-art approaches. The best results for each dataset are displayed in bold face.

As it can be observed, the proposed formulation outperforms the state-of-the-art approaches in 3 datasets out of 6 (Brodatz, Mondial, and Outex), and produces essentially equivalent results for the Kylberg Sintorn dataset. On the contrary, the CNN features descriptors perform sensibly better on the Vectorial and ALOT datasets. This is probably due to the fact that these datasets present a very large number of classes and, since the CNN features are extracted from networks that learn from millions of images, in this case they are more effective in describing the image content.

There is another possible interpretation for the poor performance of the orthogonal moments on the Vectorial and ALOT datasets. These two col-
Table 1: Comparison of classical invariant moments, and invariant moments extracted from GLCM, to three state-of-the-art approaches: rotation invariant GLCM Features, rotation invariant LBP, and CNN Features from AlexNet, Vgg19Net, and GoogLeNet.

| Texture descriptor | Dataset          | Brodatz | Mondial | Outex | Vectorial | Kyberg | ALOT |
|--------------------|------------------|---------|---------|-------|-----------|--------|------|
| Legendre-closed form |                 | 32.8    | 48.3    | 26.0  | 32.6      | 46.2   | 18.6 |
| Legendre-recursion  |                 | 35.9    | 54.2    | 25.3  | 37.4      | 47.0   | 19.7 |
| Discrete Chebyshev |                 | 35.9    | 54.1    | 25.3  | 37.5      | 47.1   | 19.6 |
| Legendre-c.f. GLCM |                 | 90.6    | 88.4    | 83.3  | 65.5      | 92.9   | 64.8 |
| Legendre-rec. GLCM |                 | 92.1    | 89.6    | 83.2  | 83.6      | 93.9   | 66.2 |
| Discr. Cheb. GLCM  |                 | 89.9    | 87.4    | 84.9  | 73.4      | 94.1   | 64.4 |
| GLCM Features      |                 | 66.2    | 65.6    | 61.9  | 50.5      | 85.7   | 40.7 |
| LBP-RI             |                 | 83.3    | 82.5    | 74.6  | 70.2      | 89.9   | 70.6 |
| AlexNet Features   |                 | 85.7    | 80.5    | 78.5  | 90.5      | 91.8   | 80.7 |
| Vgg19Net Features  |                 | 86.9    | 76.7    | 81.3  | 96.3      | 95.1   | 91.0 |
| GoogLeNet Features |                 | 87.6    | 70.8    | 72.2  | 95.0      | 94.0   | 82.6 |

lections of images, as it can be observed in Figure 4, exhibit rather regular shapes and smooth textures. The other datasets, instead, contain fine textures, characterized by the presence of high frequencies. It is well known that when a function is approximated in the least squares sense by a sum of orthogonal functions, the decay speed of the expansion coefficients increases in correspondence to an increase in the function regularity. So, there is the possibility that the decay in the moments extracted from the Vectorial and ALOT datasets makes it difficult to perform an effective moment matching during classification. This fact needs a deeper investigation and understanding, which may lead to an effective technique to improve the classification performance of orthogonal moments. This will be a further development of our research.

We also remark that in our experiments we use GLCM computed with distance 1, because our aim is to show how the invariant moments can be more discriminative if extracted from a different image representation, rather than obtaining perfect texture classification results. Indeed, in a recent work [32] we demonstrated that a larger distance may lead to better classification performances. Also, GLCM computed with distance 1 are not well suited to characterize coarse textures or even objects, both present on Vectorial and ALOT datasets (see Figure 4), but performs better in describing fine textures and close patterns.
5 Conclusions

Among all types of moments, orthogonal moments present the peculiar property of being characterized both by small information redundancy and by high discriminative power. For these reasons, they are widely used in different applications: shape-based image retrieval, edge detection, and as a feature set in pattern recognition and in biomedical image analysis.

In this work we compute orthogonal moments by employing recurrence relations for Legendre and discrete Chebyshev polynomials, taking into account not only speed, but also accuracy in computation. Indeed, errors in moments approximation may affect their discriminative ability in classification. As it can be expected, using the described approach we achieve a great reduction in the computational complexity and a more stable computation, with respect to closed form representation.

We present the results of numerical experiments to assess the discriminative power of orthogonal moments in texture analysis tasks. To this end, we use six different databases of texture images. The Legendre polynomials and the discrete Chebyshev polynomials computed by recursion formulas appear to be much more effective than the Legendre polynomials computed by the closed form representation, when the moments are extracted from the input image, and their accuracy slowly decrease when the number of moments increases. Additionally, the moments computed from the GLCM of the input image achieve a high accuracy in classification, which improves as the number of moments gets larger. Finally, we found that the use of weighted moments produces better results in some situations; this will be the object of future work.

To further highlight the performance obtained by this approach, we compare these results with some widely used state-of-the-art descriptors for texture classification: rotation invariant GLCM, rotation invariant LBP, and CNN features. Outperforming the alternative approaches in 3 datasets out of 6, and producing equivalent results on a fourth collection of images, the orthogonal moments reveal themselves as powerful and competitive descriptors for texture analysis and classification.

References

[1] M. Benčo and R. Hudec. Novel method for color textures features extraction based on GLCM. Radioengineering, 4(16):64–67, 2007.
[2] F. Bianconi and A. Fernández. Rotation invariant co-occurrence features based on digital circles and discrete Fourier transform. *Pattern Recognit. Lett.*, 48(Supplement C):34–41, 2014.

[3] J. Bigun. N-folded symmetries by complex moments in Gabor space and their application to unsupervised texture segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 16(1):80–87, 1994.

[4] S. Chen, C. Wu, D. Chen, and W. Tan. Scene classification based on gray level-gradient co-occurrence matrix in the neighborhood of interest points. In *IEEE International Conference ICIS on Intelligent Computing and Intelligent Systems*, volume 4, pages 482–485, Nov 2009.

[5] C.-W. Chong, P. Raveendran, and R. Mukundan. Translation and scale invariants of Legendre moments. *Pattern Recognition*, 37(1):119–129, 2004.

[6] B. V. Dasarathy and E. B. Holder. Image characterizations based on joint gray levelrun length distributions. *Pattern Recognit. Lett.*, 12(8):497–502, 1991.

[7] C. Di Ruberto, G. Fodde, and L. Putzu. Comparison of statistical features for medical colour image classification. In *International Conference ICVS on Computer Vision Systems*, volume 9163 of *Lecture Notes in Computer Science*, pages 3–13. Springer International Publishing, 2015.

[8] C. Di Ruberto, A. Loddo, and L. Putzu. Histological image analysis by invariant descriptors. In *International Conference ICIAP on Image Analysis and Processing*, volume 10484 of *Lecture Notes in Computer Science*, pages 345–356. Springer International Publishing, 2017.

[9] C. Di Ruberto and A. Morgera. A comparison of 2-D moment-based description techniques. In F. Roli and S. Vitulano, editors, *Proceedings of International Conference on Image Analysis and Processing (ICIAP)*, pages 212–219. Springer Berlin Heidelberg, 2005.

[10] C. Di Ruberto and A. Morgera. Moment-based techniques for image retrieval. In *2008 19th International Workshop on Database and Expert Systems Applications*, pages 155–159, 2008.

[11] W. Gautschi. *Orthogonal Polynomials: Computation and Approximation*. Oxford University Press, 2004.
[12] A. Gelzinis, A. Verikas, and M. Bacauskiene. Increasing the discrimination power of the co-occurrence matrix-based features. *Pattern Recognit.*, 40(9):2367–2372, 2007.

[13] R. Gong and H. Wang. Steganalysis for GIF images based on color-gradient co-occurrence matrix. *Opt. Commun.*, 285:4961–4965, November 2012.

[14] J. Haddadnia, M. Ahmadi, and K. Faez. An efficient feature extraction method with pseudo-Zernike moment in RBF neural network-based human face recognition system. *EURASIP J. Adv. Signal Process.*, 2003(9):890901, 2003.

[15] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *IEEE Trans. Syst. Man Cybern.*, SMC-3(6):610–621, 1973.

[16] M. S. Hitam, S. Abu Bakar, W. N. Jawahir, and W. Yussof. Content-based image retrieval using Zernike moments for binary and grayscale images. In *Moments and Moment Invariants - Theory and Applications*, volume 1, pages 275–288, 2014.

[17] M. K. Hu. Visual pattern recognition by moment invariants. *IRE Trans. Inf. Theory*, 8(2):179–187, 1962.

[18] S.-K. Hwang and W.-Y. Kim. A novel approach to the fast computation of Zernike moments. *Pattern Recognit.*, 39(11):2065–2076, 2006.

[19] A. Krizhevsky, I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1*, volume 1 of *NIPS’12*, pages 1097–1105, 2012.

[20] C. Lakshmi Deepika, A. Kandaswamy, C. Vimal, and B. Satish. Palmprint authentication using modified Legendre moments. *Procedia Comput. Sci.*, 2(Supplement C):164–172, 2010. Proceedings of the International Conference and Exhibition on Biometrics Technology.

[21] S. Li, M. C. Lee, and C. M. Pun. Complex zernike moments features for shape-based image retrieval. *IEEE Trans. Syst. Man Cybern.*, 39(1):227–237, 2009.

[22] X. Li and A. Song. A new edge detection method using Gaussian-Zernike moment operator. In *IEEE 2nd International Asia Conference...*
on Informatics in Control, Automation and Robotics (CAR), volume 1, pages 276–279, 2010.

[23] Z. Ma, B. Kang, and J. Ma. Translation and scale invariant of Legendre moments for images retrieval. *J. Inf. Comput. Sci.*, 8(11):2221–2229, 2011.

[24] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. Pattern Anal. Mach. Intell.*, 11(7):674–693, 1989.

[25] D. Mitrea, P. Mitrea, S. Nedevschi, R. Badea, M. Lupsor, M. Socaciu, A. Golea, C. Hagi, and L. Ciobanu. Abdominal tumor characterization and recognition using superior-order co-occurrence matrices, based on ultrasound images. *Comput. Math. Method Med.*, 2012, 2012.

[26] R. Mukundan, S. H. Ong, and P. A. Lee. Image analysis by Tchebichef moments. *IEEE Trans. Image Process.*, 10(9):1357–1364, 2001.

[27] R. R. Obulakonda, R. B. Eswara, and R. E. Keshava. Ternary patterns and moment invariants for texture classification. *ICTACT J. Image Video Process.*, 7(1):1295–1298, 2016.

[28] T. Ojala, M. Pietikinen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognit.*, 29(1):51–59, 1996.

[29] T. Ojala, M. Pietikinen, and Menp T. Multiresolution gray-scale and rotation invariant texture classification with local binary pattern. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(7):971–987, 2002.

[30] M. Oujaoura, B. Minaoui, and M. Fakir. Image annotation by moments. In *Moments and Moment Invariants - Theory and Applications*, volume 1, pages 227–252, 2014.

[31] S. H. Peckinpaugh. An improved method for computing gray-level co-occurrence matrix based texture measures. *CVGIP: Graph. Models Image Process.*, 53(6):574–580, 1991.

[32] L. Putzu and C. Di Ruberto. Rotation invariant co-occurrence matrix features. In *19th International Conference ICIAP on Image Analysis and Processing*, volume 10484 of *Lecture Notes in Computer Science*, pages 391–401. Springer International Publishing, 2017.
[33] T.J. Rivlin. *An Introduction to the Approximation of Functions*. Dover Publications Inc., New York, 1981.

[34] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. https://arxiv.org/abs/1409.1556, 2014.

[35] P. Soille. Morphological texture analysis: An introduction. In K. Mecke and D. Stoyan, editors, *Morphology of Condensed Matter: Physics and Geometry of Spatially Complex Systems*, pages 215–237. Springer Berlin Heidelberg, 2002.

[36] J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics*. Springer-Verlag, New York, second edition, 1991.

[37] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–9, June 2015.

[38] A. Tahmasbi, F. Saki, and S. B. Shokouhi. Classification of benign and malignant masses based on Zernike moments. *Comput. Biol. Med.*, 41(8):726–735, 2011.

[39] M. R. Teague. Image analysis via the general theory of moments. *J. Opt. Soc. Am.*, 70(8):920–930, 1980.

[40] M. Tuceryan. Moment-based texture segmentation. *Pattern Recognit. Lett.*, 15(7):659–668, 1994.

[41] B. Vijayalakshmi and V. S. Bharathi. Classification of CT liver images using local binary pattern with Legendre moments. *Curr. Sci.*, 110(4):687–691, 2016.

[42] K. Wu, C. Garnier, J. L. Coatrieux, and H. Shu. A preliminary study of moment-based texture analysis for medical images. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 5581–5584, 2010.

[43] P. T. Yap, R. Paramesran, and S.-H. Ong. Image analysis by Krawtchouk moments. *IEEE Trans. Image Process.*, 12(11):1367–1377, 2003.