

Polar shapelets

Richard Massey^{1,2★} and Alexandre Refregier^{3★}

¹*Institute of Astronomy, Madingley Road, Cambridge CB3 0HA*

²*California Institute of Technology, 1200 E. California Blvd, Pasadena, CA 91125, USA*

³*Service d’Astrophysique, CEA/Saclay, 91191 Gif-sur-Yvette, France*

Accepted 2005 July 13. Received 2005 July 6; in original form 2004 August 25

ABSTRACT

The shapelets method for image analysis is based upon the decomposition of localized objects into a series of orthogonal components with convenient mathematical properties. We extend the ‘Cartesian shapelet’ formalism from earlier work, and construct ‘polar shapelet’ basis functions that separate an image into components with explicit rotational symmetries. These frequently provide a more compact parametrization, and can be interpreted in an intuitive way. Image manipulation in shapelet space is simplified by the concise expressions for linear coordinate transformations, and shape measures (including object photometry, astrometry and galaxy morphology estimators) take a naturally elegant form. Particular attention is paid to the analysis of astronomical survey images, and we test shapelet techniques upon real data from the *Hubble Space Telescope*. We present a practical method to automatically optimize the quality of an arbitrary shapelet decomposition in the presence of observational noise, pixelization and a point spread function. A central component of this procedure is the adaptive choice of the scale size and the truncation order of the shapelet expansion. A complete software package to perform shapelet image analysis is made available on the World Wide Web.

Key words: methods: analytical – methods: data analysis – techniques: image processing – galaxies: fundamental parameters.

1 INTRODUCTION

In the shapelets formalism (Refregier 2003; hereafter Shapelets I), individual objects in an image are decomposed into weighted sums of orthogonal basis functions. The particular set of basis functions has been chosen to be mathematically convenient for image manipulation and analysis. In astronomical images, it also provides a compact representation for the shapes of galaxies of all morphological types. Refregier & Bacon (2003; hereafter Shapelets II) showed how these properties could be used to measure the slight distortions in galaxy shapes due to weak gravitational lensing. The elegant behaviour of shapelets under a Fourier transform also enabled Chang & Refregier (2002) to reconstruct images from interferometric observations. Massey et al. (2004) used shapelets to simulate realistic astronomical images containing galaxies with complex morphologies. A classification scheme for galaxy morphologies using principal-component analysis of the shapelet basis set was discussed in that paper and applied to the Sloan Digital Sky Survey by Kelly & McKay (2004). A method similar to shapelets has also been independently suggested by Bernstein & Jarvis (2002; hereafter BJ02).

In this paper, we expand upon the earlier work of Shapelets I, Shapelets II and BJ02, developing the formalism of ‘polar

shapelets’, in which an image is decomposed into components with explicit rotational symmetries. Whilst the original Cartesian shapelets remain useful in certain situations, the polar shapelets, which are separable in r and θ , frequently provide a more elegant and intuitive form. We find estimators of the flux, position and size of an object, that form naturally from groups of its polar shapelet coefficients. We calculate the behaviour of a polar shapelet model during linear coordinate transformations. We also improve the basic shapelet decomposition by incorporating treatments of pixelization, observational noise and point-spread functions, and optimizing the overall quality of image reconstruction while maximizing data compression. To test our method upon real data, we extract isolated galaxies from the *Hubble Deep Fields* (hereafter HDFs; Williams et al. 1996, 1998). These deep, high-resolution images from the *Hubble Space Telescope* (*HST*) provide typical examples of the irregular shapes of distant galaxies.

A complete IDL software package to perform the image decomposition and shape analyses presented in this paper can be downloaded from <http://www.astro.caltech.edu/~rjm/shapelets/>.

This paper is laid out as follows. In Section 2, we introduce the Cartesian and polar shapelet basis functions, and their relation to each other. In Section 3, we investigate the qualitative effects of varying the shapelet scale size β and set quantitative goals for the optimization of this choice. In Section 4, we develop practical techniques to cope with the effects of pixelization, seeing and noise in

*E-mail: rjm@astro.caltech.edu (RM); refregier@cea.fr (AR)

real data. We then demonstrate various applications of shapelets: in Section 5, we illustrate the manipulation of images in terms of their changing polar shapelet coefficients under coordinate transforms. In Section 6, we construct basic shape estimators for a shapelet model, including flux, centroid and size measures. In Section 7, we develop more advanced shape measures that can be used to quantitatively distinguish galaxies of various morphological types. We conclude in Section 8.

2 SHAPELETS FORMALISM

2.1 Cartesian basis functions

The shapelet image decomposition method was introduced in Shapelets I, and a related method has been independently suggested by BJ02. The idea is to express the surface brightness of an object $f(x, y)$ as a linear sum of orthogonal two-dimensional (2D) functions,

$$f(\mathbf{x}) = \sum_{n_1=0}^{\infty} \sum_{n_2=0}^{\infty} f_{n_1, n_2} \phi_{n_1, n_2}(\mathbf{x}; \beta), \quad (1)$$

where the f_{n_1, n_2} are the ‘shapelet coefficients’ to be determined. The dimensionful shapelet basis functions ϕ_{n_1, n_2} are

$$\phi_{n_1, n_2}(\mathbf{x}; \beta) \equiv \frac{H_{n_1}(x/\beta) H_{n_2}(y/\beta) e^{-|x|^2/2\beta^2}}{\beta 2^n \sqrt{\pi n_1! n_2!}}, \quad (2)$$

where $H_n(x)$ is a Hermite polynomial of the order of n , and β is the shapelet scale size. They are shown in Fig. 1. These Gauss–Hermite polynomials form a complete and orthonormal basis set; this ensures that the shapelet coefficients for any image can be simply and uniquely determined by evaluating the ‘overlap integral’

$$f_{n_1, n_2} = \iint_{\mathbb{R}} f(\mathbf{x}) \phi_{n_1, n_2}(\mathbf{x}; \beta) d^2x. \quad (3)$$

In practice, a shapelet expansion (equation 1) must be truncated at a finite order $n_1 + n_2 \leq n_{\max}$. The array of shapelet coefficients is sparse for typical galaxy morphologies, which therefore can be accurately modelled using only a few coefficients. As shown in Shapelets I, data compression ratios as high as 60:1 can be achieved for well resolved *HST* images. Note, however, that our choice of Gauss–Hermite basis functions was not governed by the physics of galaxy morphology and evolution but by the mathematics of image manipulation. As we shall see throughout this paper, a shapelet parametrization is mathematically convenient for many tasks common in astronomy and other sciences.

2.2 Polar shapelet basis functions

Polar shapelets were introduced in Shapelets I as an orthogonal transformation of the Cartesian basis states, and were independently proposed by BJ02. They have all the useful properties of Cartesian shapelets, and a similar Gaussian weighting function with a given scale size β . However, the polar shapelet basis functions are instead separable in r and θ . This renders many operations more intuitive, and makes polar shapelet coefficients easy to interpret in terms of their explicit rotational symmetries.

The polar shapelet basis functions $\chi_{n, m}(r, \theta; \beta)$ are also parametrized by two integers, n and m , and a smooth function $f(r, \theta)$ in polar coordinates may be decomposed into

$$f(r, \theta) = \sum_{n=0}^{\infty} \sum_{m=-n}^n f_{n, m} \chi_{n, m}(r, \theta; \beta). \quad (4)$$

The polar shapelet coefficients $f_{n, m}$ are again given by the ‘overlap integral’

$$f_{n, m} = \iint_{\mathbb{R}} f(r, \theta) \chi_{n, m}(r, \theta; \beta) r dr d\theta. \quad (5)$$

BJ02 showed that the ‘polar Hermite polynomials’ $H_{n_1, n_2}(x)$, which were described in Shapelets I, are related to associated Laguerre polynomials

$$L_p^q(x) \equiv \frac{x^{-q} e^x}{p!} \frac{d^p}{dx^p} (x^{p+q} e^{-x}) \quad (6)$$

for $n_r > n_l$ by

$$H_{n_l, n_r}(x) \equiv (-1)^{n_l} (n_l!) x^{n_r - n_l} L_{n_l}^{n_r - n_l}(x^2), \quad (7)$$

where n_l and n_r are any non-negative integers. In this paper we shall instead prefer the simpler n, m notation, where $n = n_r + n_l$ and $m = n_r - n_l$. In this scheme, n can be any non-negative integer, and m can be any integer between $-n$ and n in steps of 2. We truncate the series at $n \leq n_{\max}$. Although the only allowed states are those with n and m both even or both odd, this condition will not be written explicitly alongside every summation for the sake of brevity.

As plotted in Fig. 2, the dimensionful polar shapelet basis functions are therefore

$$\chi_{n, m}(r, \theta; \beta) = \frac{(-1)^{(n-|m|)/2}}{\beta^{|m|+1}} \left\{ \frac{[(n-|m|)/2]!}{\pi [(n+|m|)/2]!} \right\}^{1/2} \times r^{|m|} L_{(n-|m|)/2}^{|m|} \left(\frac{r^2}{\beta^2} \right) e^{-r^2/2\beta^2} e^{-im\theta}. \quad (8)$$

These are different from the Laguerre expansion used by BJ02 in two ways. Those are the complex conjugate of our basis functions: i.e. their m is equivalent to our $-m$. The Laguerre expansion in BJ02

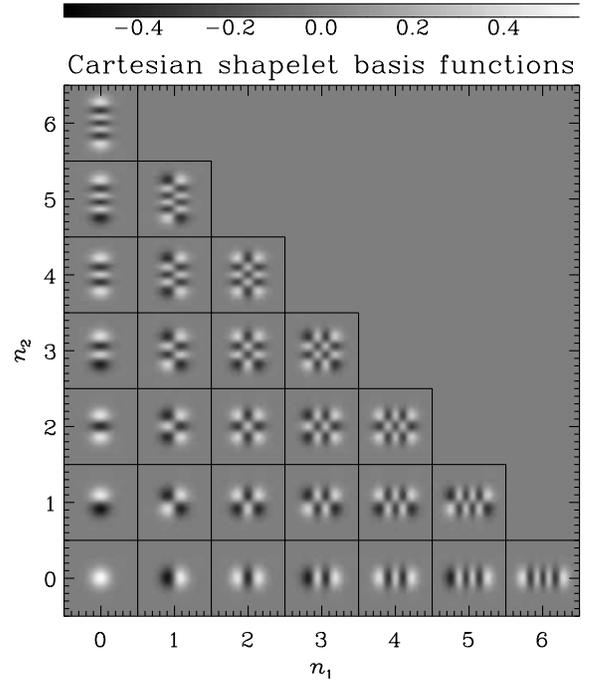


Figure 1. Cartesian shapelet basis functions, parametrized by two integers n_1 and n_2 , and here truncated at $n_{\max} = 6$. An image can be decomposed into a weighted sum of these functions. This basis is particularly convenient for many aspects of image analysis and manipulation commonly used in astronomy and other sciences.

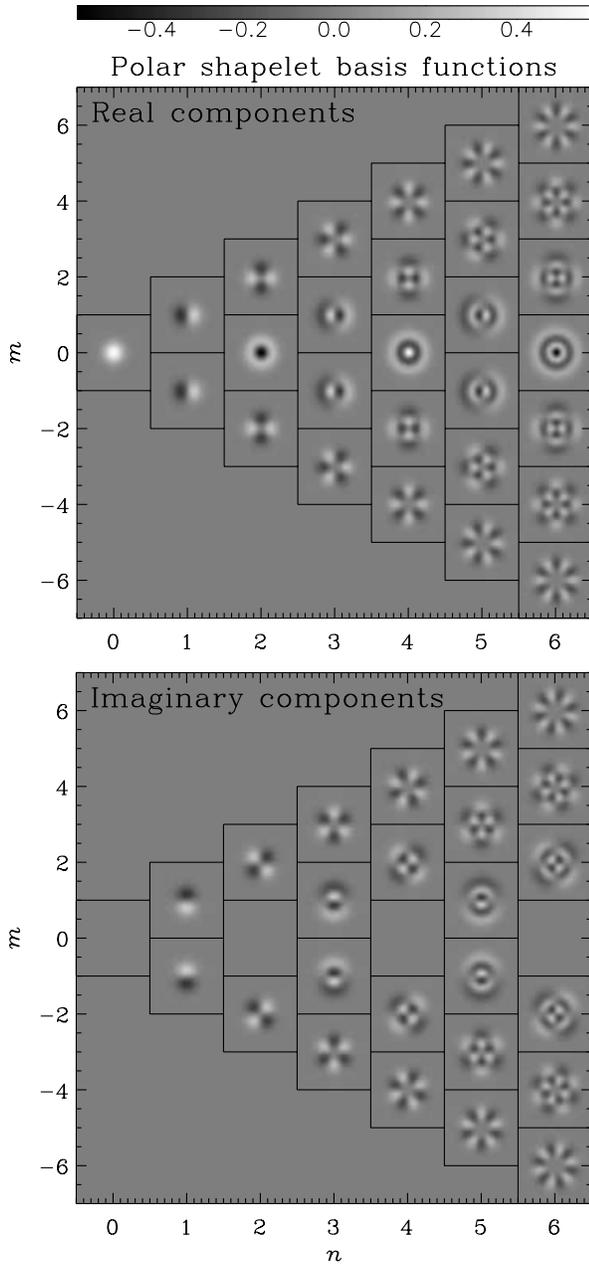


Figure 2. Polar shapelet basis functions. The real components of the complex functions are shown in the top panel, and the imaginary components in the bottom. The basis functions with $m = 0$ are wholly real. In a shapelet decomposition, all of the basis functions are weighted by a complex number, the magnitude of which determines the strength of a component and the phase of which sets its orientation.

is also normalized by one less factor of β . This dimensionality ensures that, as in the case of Cartesian shapelets, the polar shapelet basis functions are orthonormal

$$\iint_{\mathbb{R}} \chi_{n,m}^*(r, \theta; \beta) \chi_{n',m'}(r, \theta; \beta) r dr d\theta = \delta_{n,n'} \delta_{m,m'} \quad (9)$$

and complete (see, e.g., Wünsche 1998)

$$\sum_{n=0}^{\infty} \sum_{m=-n}^n \chi_{n,m}(r, \theta; \beta) \chi_{n,m}(r', \theta'; \beta) = \delta(r - r') \delta(\theta - \theta'), \quad (10)$$

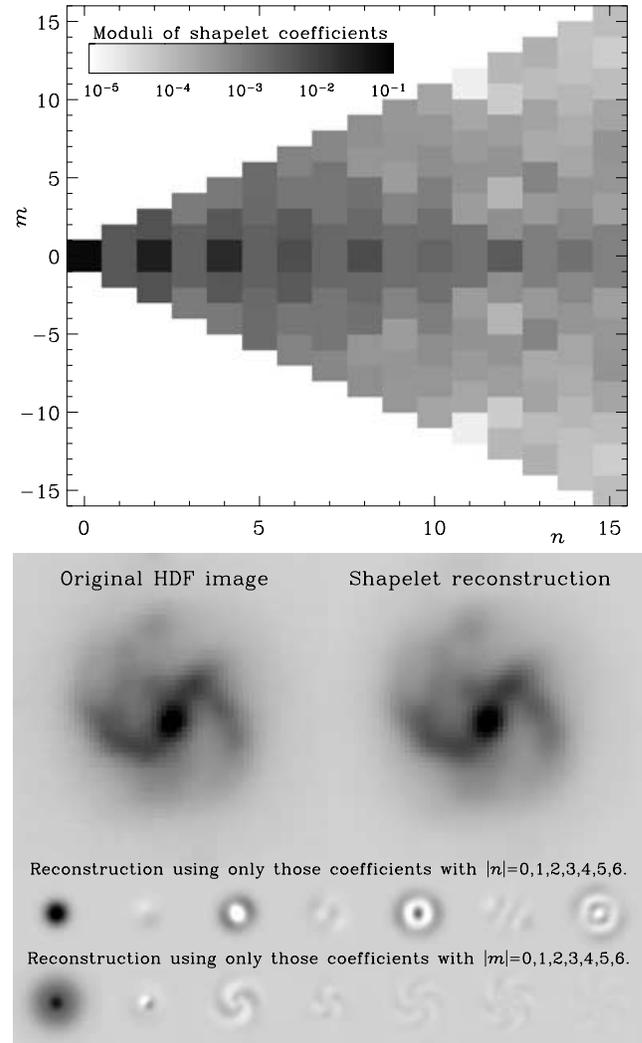


Figure 3. Example polar shapelet decomposition of a HDF galaxy. Top panel: the moduli of the polar shapelet coefficients, with a logarithmic colour scale. Bottom panel: the original galaxy image using a linear colour scale and its shapelet reconstruction using $n_{\max} = 20$. Additional reconstructions are shown using only particular sets of coefficients, to highlight the contribution of components containing different symmetries.

where δ is the Kronecker delta and the asterisk denotes complex conjugation. Only those basis functions with $m = 0$ contain net flux.

$$\iint_{\mathbb{R}} \chi_{n,m}(r, \theta; \beta) r dr d\theta = 2\sqrt{\pi}\beta \delta_{m0}. \quad (11)$$

Fig. 3 demonstrates the polar shapelet decomposition of a galaxy found in the HDF. The original image (middle left-hand panel) agrees well with the reconstruction using $n_{\max} = 20$. The top panel shows the modulus of the polar shapelet coefficients as a function of the n and m indices. The dominant coefficients have small values of both indices, demonstrating the compactness of a polar shapelet representation, and further improved prospects for data compression. The bottom panel shows the reconstruction of the galaxy using only coefficients with given values of n or $|m|$, thus highlighting the contributions of terms with specific rotational symmetries. The off-central bulge is captured in the $|m| = 1$ coefficients and the main spiral arms in the $|m| = 2$ coefficients. The spiral arms can also be

seen as the rotation of the n -only reconstructions with increasing radius. The fainter spiral arms appear as an interplay of the $|m| = 4$ and 5 coefficients.

2.3 Conversion between Cartesian and polar shapelets

Cartesian shapelets are real functions, but polar shapelet basis functions $\chi_{n,m}$ and coefficients $f_{n,m}$ are both complex. However, their symmetries

$$\chi_{n,-m}(r, \theta; \beta) = \chi_{n,m}^*(r, \theta; \beta) = \chi_{n,m}(r, -\theta; \beta), \quad (12)$$

simplify matters somewhat if we are concerned only with the representation of real functions $f(\mathbf{x})$, such as the surface brightness of an image. Equations (9) and (12) imply that $f(\mathbf{x})$ is real if and only if

$$f_{n,-m} = f_{n,m}^*. \quad (13)$$

Coefficients with $m = 0$ are thus wholly real. All polar shapelet coefficients are paired with their complex conjugate on the other side of the line $m = 0$. Therefore, even though the polar shapelet coefficients $f_{n,m}$ are generally complex, the number of independent parameters in the shapelet decomposition of a real function is conserved from the Cartesian case.

A set of Cartesian shapelet coefficients f_{n_1, n_2} with $n_1 + n_2 \leq n_{\max}$ can be transformed, into polar shapelet representation with $n \leq n_{\max}$, using

$$f_{n,m} = 2^{-n/2} i^m \left\{ \frac{n_1! n_2!}{[(n+m)/2]! [(n-m)/2]!} \right\}^{1/2} \delta_{n_1+n_2, n} \\ \times \sum_{n'_1=0}^{n_1} \sum_{n'_2=0}^{n_2} i^{m'} \binom{(n+m)/2}{n'_1} \binom{(n-m)/2}{n'_2} \delta_{n'_1+n'_2, n_1} f_{n_1, n_2}. \quad (14)$$

The particular choices of truncation scheme for Cartesian and polar shapelets now make sense as a way to keep this mapping one-to-one.

3 CHOICE OF SHAPELET SCALE SIZE

A shapelet decomposition requires values for the scale size β and for the centre of the basis functions \mathbf{x}_c to be specified in advance. Choosing the centre is relatively easy: there are many methods well known in the astronomical literature to accurately determine astrometry from the flux-weighted moments of objects. However, the selection of β is a less well-posed problem. In this section, we shall first use some properties of polar shapelets to describe the effect that

the choice of the scale size has upon a shapelet decomposition. We shall then set quantitative criteria for the selection of β in arbitrary galaxy images that we can implement in a practical algorithm.

Note that the selection of β will be linked to the selection of n_{\max} . As shown in Shapelets I Section 2.4, these two parameters determine the maximum extent θ_{\max} and finest resolution θ_{\min} that can be successfully captured by a shapelet model. If $n_{\max} \rightarrow \infty$, any object can be represented using any scale size β . However, if the shapelet expansion is truncated at finite n_{\max} , the shape information needs to be more efficiently contained within fewer coefficients. It is clearly desirable in this situation to select a scale size β that compresses information, and lets us store the smallest possible number of coefficients.

3.1 Radial profiles

Our discussion can be simplified by initially considering only the radial profile of an object, thus reducing the task to a one-dimensional problem. Let us consider an object with surface brightness $f(\mathbf{x})$. The radial profile of the object $\bar{f}(r)$ is its brightness averaged in concentric rings about its centre, i.e.

$$\bar{f}(r) = \frac{1}{2\pi} \int_0^{2\pi} f(r, \theta) d\theta. \quad (15)$$

With the object decomposed into polar shapelets as in equation (5), it is easy to show that this is given by

$$\bar{f}(r) = \sum_n^{\text{even}} f_{n0} \chi_{n0}(r; \beta). \quad (16)$$

This simple expression results from the fact that only the $m = 0$ basis functions are invariant under rotations. These are given by

$$\chi_{n0}(r; \beta) = \frac{(-1)^{n/2}}{\beta \sqrt{\pi}} L_{n/2}^0(r^2/\beta^2) e^{-r^2/2\beta^2}. \quad (17)$$

The first few rotationally invariant basis functions are written explicitly in Table 1.

As a concrete example, we consider the decomposition of galaxy images from the Hubble Deep Fields (Williams et al. 1996, 1998). The mean radial profile of spiral galaxies is typically exponential, $f(r) \propto e^{-r/r_0}$, with some characteristic radius scale r_0 . Fig. 4 shows the shapelet reconstruction of an exponential radial profile using various values of β , with $n_{\max} = 20$ and the integral in equation (5) calculated numerically.

As can be seen in the top panel of Fig. 4, the quality of the reconstruction depends on the choice of β . For small values

Table 1. The first few rotationally invariant polar Shapelet basis functions.

$\chi_{0,0}(r) = \frac{1}{\beta \sqrt{\pi}} e^{-r^2/2\beta^2}$
$\chi_{2,0}(r) = \frac{-1}{\beta \sqrt{\pi}} \left(1 - \frac{r^2}{\beta^2} \right) e^{-r^2/2\beta^2}$
$\chi_{4,0}(r) = \frac{1}{\beta \sqrt{\pi}} \left(1 - 2 \frac{r^2}{\beta^2} + \frac{1}{2} \frac{r^4}{\beta^4} \right) e^{-r^2/2\beta^2}$
$\chi_{6,0}(r) = \frac{-1}{\beta \sqrt{\pi}} \left(1 - 3 \frac{r^2}{\beta^2} + \frac{3}{2} \frac{r^4}{\beta^4} - \frac{1}{6} \frac{r^6}{\beta^6} \right) e^{-r^2/2\beta^2}$
$\chi_{8,0}(r) = \frac{1}{\beta \sqrt{\pi}} \left(1 - 4 \frac{r^2}{\beta^2} + \frac{6}{2} \frac{r^4}{\beta^4} - \frac{4}{6} \frac{r^6}{\beta^6} + \frac{1}{24} \frac{r^8}{\beta^8} \right) e^{-r^2/2\beta^2}$
$\chi_{10,0}(r) = \frac{-1}{\beta \sqrt{\pi}} \left(1 - 5 \frac{r^2}{\beta^2} + \frac{10}{2} \frac{r^4}{\beta^4} - \frac{10}{6} \frac{r^6}{\beta^6} + \frac{5}{24} \frac{r^8}{\beta^8} + \frac{1}{120} \frac{r^{10}}{\beta^{10}} \right) e^{-r^2/2\beta^2}$

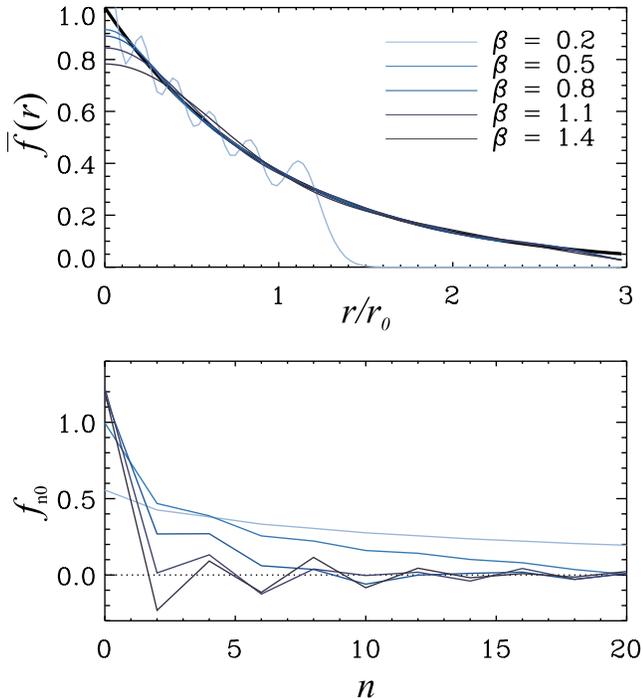


Figure 4. Decomposition of an exponential profile into radial polar shapelets. Top panel: the thick dark line shows the input exponential profile. The reconstructed profile is shown for different values of the shapelet scale β with $n_{\max} = 20$. Bottom panel: the corresponding shapelet coefficient profile f_{n0} versus shapelet order n .

($\beta \lesssim 0.4r_0$) the reconstruction is oscillatory and cuts off the profile at large radii ($r \gtrsim 1.5r_0$). For large values ($\beta \gtrsim 0.8r_0$), the reconstruction fails to reproduce the cusp at small radii ($r \lesssim 0.4r_0$) and exceeds the true profile at $r \simeq 0.6r_0$. However, for intermediate values ($0.5r_0 < \beta < 1.1r_0$), the reconstruction is good throughout the range $0.1r_0 \lesssim r \lesssim 2.8r_0$. This range can of course be expanded by including more shapelet coefficients of higher order. As $n_{\max} \rightarrow \infty$, the input model can be recovered with arbitrary precision using *any* value of β .

The corresponding behaviour in shapelet space is apparent in the bottom panel of Fig. 4. The $f_{n,0}$ coefficients can be thought as the profile of the galaxy in shapelet space or the ‘shapelet profile’. For low values of β the shapelet profile is very flat, showing that the power is distributed almost evenly throughout all orders. For $\beta = 0.5r_0$, the coefficients $a_{n,m}$ are seen numerically to be $\propto (n+1)^{-2}$. This will be an important result for the convergence of shapelet estimators formed from series of shapelet coefficients in Section 6. Convergence is fastest at $\beta \approx 0.8r_0$, with $a_{n,m} \propto (n+1)^{-5/2}$. For higher values of β , the signs of $a_{n,m}$ begin to alternate and the convergence once again falls below $\propto (n+1)^{-2}$ at $\beta \approx 1.1r_0$.

Fig. 5 demonstrates the importance of a proper choice of the parameters β and n_{\max} for the practical decomposition of a spiral galaxy in the HDF. Its spiral arms complicate measurement, but its radial profile is approximately an exponential with a scalelength of $r_0 \approx 3$ arcsec (12 pixel). The left-hand column shows the growth in complexity of a shapelet model using increasing n_{\max} . Note, in particular, the rotation of the core ellipticity as n_{\max} is increased from 2 to 8 and higher-order moments are included to resolve the spiral arms. In this column, β is allowed to vary in order to minimize the least-squares difference between the model and the HDF image, shown at the bottom. The middle column shows shapelet

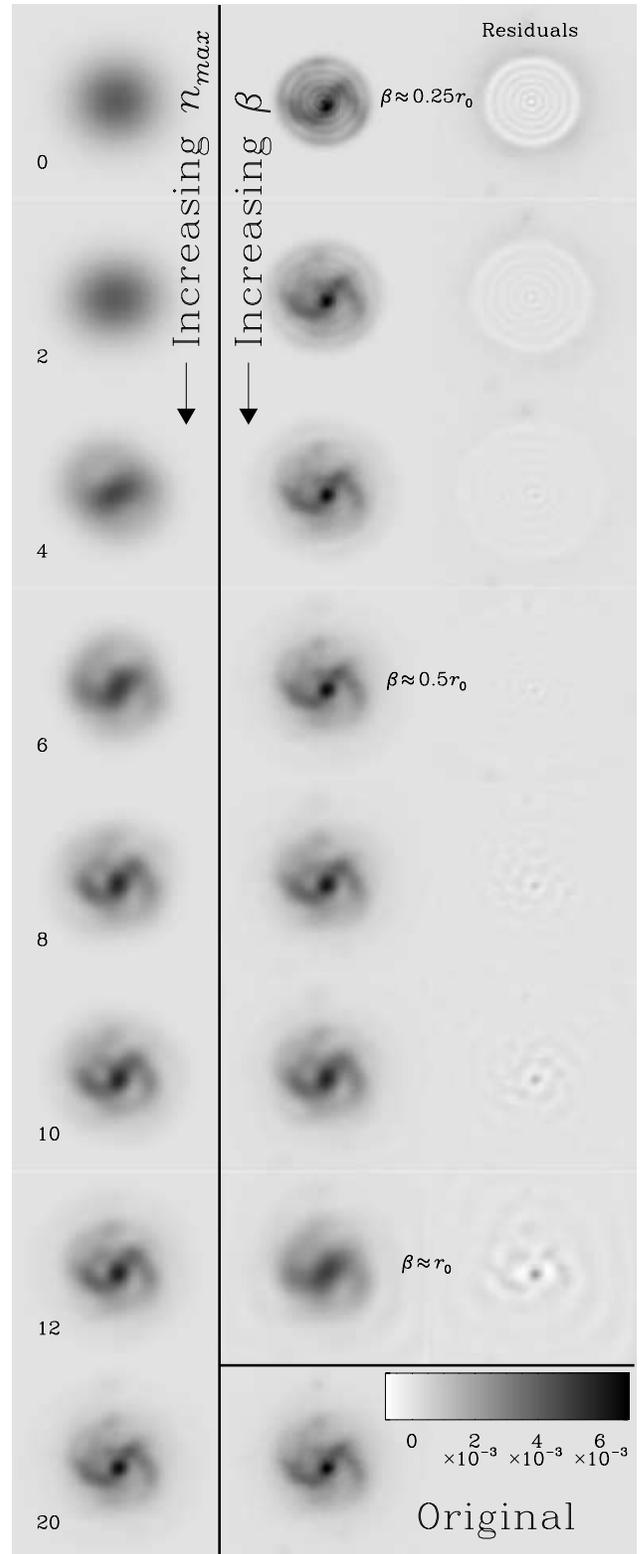


Figure 5. Shapelet decomposition of a real spiral galaxy in the HDF. The best-fitting de Vaucouleurs profile has $r_0 \simeq 12$ pixels. Left-hand column: the shapelet model shows growing complexity with increasing n_{\max} . For each of these fits, β is varied to minimize the least-squares difference between the data and the model. Right-hand columns: the shapelet decomposition has a preferred scale size. The residual between the original (in the bottom right-hand panel) and these models with fixed $n_{\max} = 20$ and varying β , is smallest with $\beta \simeq 0.5r_0$.

decompositions at fixed $n_{\max} = 20$, with varying β . The residuals are plotted in the right-hand column. As in Fig. 4, we find that the best overall image reconstruction uses $0.5r_0 \lesssim \beta \lesssim 0.7r_0$. This is perhaps at the low end of the range suggested by Fig. 4 because of the extra high-frequency detail contained in the spiral arms.

By experimentation we have found a fairly wide range of β values that produce a faithful shapelet reconstruction. The information is then concentrated into the few lowest shapelet states, with fast convergence to the final model, and truncation is possible at a computationally acceptable value of n_{\max} . We shall now consider ways to formalize this process, and hone our choice of x_c , β and n_{\max} using quantitative criteria.

3.2 Existing optimization methods

Methods in the literature that face similar choices suggest several distinct philosophies for the quantitative selection of parameters equivalent to x_c , β and n_{\max} . The suggestions, outlined below, differ both in the goals set for an ideal decomposition and the method used to achieve it.

(i) Shapelets I suggested a geometrical argument using $\theta_{\min}, \theta_{\max}$: the minimum [point spread function (PSF) or pixel] and maximum (entire image) sizes on which information exists. This could be iterated using functional rules on x_c and r_f^2 as defined by shapelet coefficients. However, the coefficients change as a function of n_{\max} , and it is not clear what the rules should be.

(ii) Van der Marel & Franx (1993) fit one-dimensional (1D) Gauss–Hermite polynomials to spectral lines. They arbitrarily fix $n_{\max} = 6$, probably finding this sufficient because their spectra have relatively high signal-to-noise (S/N) ratios and their lines have a nearly Gaussian profile. Parameters equivalent to x_c and β are obtained from the best-fitting Gaussian. This also determines f_0 and in 1D is equivalent to constraining $f_1 = f_2 = 0$, i.e. the derivatives of the Gaussian with respect to x_c and β . The number of variables is reduced and the problem rendered tractable. Unfortunately, this does not help us in 2D because while both $a_{1\pm 1}$ can be forced to zero by varying x_c , no unique recipe can be found for setting the three $n = 2$ states using only one value β .

(iii) Van der Marel et al. (1994) relax the constraint on f_1 . This is an improvement as f_1 is only the first term of an expression for the centroid, expanded using *all* odd f_n in equation (51). Without the higher-order corrections, the true object centroid is moved slightly from the origin: amongst other things rendering rotations and shear operations more complicated. Instead, they set x_c from the theoretical rest wavelength of a line. Unfortunately, astrometric calibration clearly cannot be performed with such accuracy. Nor has the $n = 2$ problem been solved.

(iv) Kaiser, Squires & Broadhurst (1995) combine fitting with a stand-alone object detection algorithm, HFINDPEAKS. Translated into shapelet language, their approach is roughly equivalent to placing x_c at data peaks and then finding a width β such that the signal-to-noise ratio ν in f_0 is maximized.

(v) Bernstein & Jarvis (2002) propose a similar approach. They prescribe β by requiring $f_{20} = 0$, while moving x_c to ensure $f_{1\pm 1} = 0$. Higher coefficients are then determined afterwards by linear decomposition. To first order, this β constraint is equivalent to that for HFINDPEAKS. This β is generally larger than values chosen by our χ^2 method below, and it can be several times larger for a high signal-to-noise ratio object containing lots of substructure such as the galaxy in Fig. 3. This method may indeed prescribe the optimal decomposition for weak lensing as the shear signal in the quadrupole

moments becomes concentrated in one number; however, a predisposition towards particular states often leads to poor overall image reconstruction and PSF deconvolution, so it is not necessarily ideal for all applications.

(vi) Kelly & McKay (2004) were able to set a fixed *physical* scale of $\beta = 2$ kpc for galaxies in the Sloan Digital Sky Survey, where photometric redshifts were available. However, galaxies have a broad distribution of physical sizes, and it may in fact become more difficult to interpret a shapelet model derived using this method.

(vii) Marshall (in preparation) describes a fully Bayesian approach to applying the shapelet transform in the context of image reconstruction. Here, x_c , β and n_{\max} are varied in order to maximize the evidence (the probability of observing the data, marginalized over all shapelet coefficients). At high S/N ratio, this method gives a value of β which approaches the same as that from our χ^2 method below, but otherwise tends to prefer a fractionally larger β , conservatively eliminating some ‘noise’ in favour of a smoother image reconstruction. However, this is computationally slow, a serious issue when analysing large images.

3.3 Optimization of image reconstruction

We shall adopt a choice of β and n_{\max} that is suitable for many applications, including overall image reconstruction and PSF deconvolution. Different models will be quantitatively compared via the overall reconstruction residual

$$\chi_r^2 \equiv \frac{[f_{\text{obs}}(\mathbf{x}) - f_{\text{rec}}(\mathbf{x}; \beta)]^T V^{-1} [f_{\text{obs}}(\mathbf{x}) - f_{\text{rec}}(\mathbf{x}; \beta)]}{n_{\text{pixels}} - n_{\text{coeffs}}}, \quad (18)$$

where $f_{\text{obs}}(\mathbf{x})$ is the observed image and $f_{\text{rec}}(\mathbf{x}; \beta)$ is the reconstructed image from the shapelet model. V is the covariance matrix between pixel values, i.e. its diagonal elements are the noise variance in each pixel. We will need to know this a priori, or estimate it from blank areas of the image. n_{pixels} is the number of pixels in the observed image and n_{coeffs} is the number of shapelet coefficients used in the model. The residual itself has variance (Lupton 1993)

$$\sigma^2(\chi_r^2) = \frac{2}{n_{\text{pixels}} - n_{\text{coeffs}}}. \quad (19)$$

An example of typical χ_r^2 isocontours on an n_{\max} versus β plane is shown in Fig. 6 for an elliptical galaxy from the HDF. The horizontal trough is present for all galaxies (and many other isolated objects). This demonstrates that there is indeed an optimum β for the reconstruction of this image. As one might expect, it is roughly independent of n_{\max} , but decreases very slightly as more coefficients are added. By increasing $n_{\max} \rightarrow \infty$, the reconstruction can be improved to arbitrary precision. However, stopping at the $\chi_r^2 = 1$ contour produces a model where the residual is consistent with the noise. Additional coefficients would just model the background noise and should be excluded.

The form of these typical contours thus suggests a unique location in parameter space. We will choose β and n_{\max} so that the model lies at the intersection of the trough and the $\chi_r^2 = 1$ contour, i.e. at the leftmost point on the contour. To achieve this, we set quantitative goals of

$$\frac{\partial \chi_r^2}{\partial \beta} = 0, \quad (20)$$

$$x_c = 0 \quad (21)$$

and $\chi_r^2 = 1$ or flattens out

$$\frac{\partial \chi_r^2}{\partial n_{\max}} < \sigma(\chi_r^2) \approx \sqrt{\frac{2}{n_{\text{pixels}}}}. \quad (22)$$

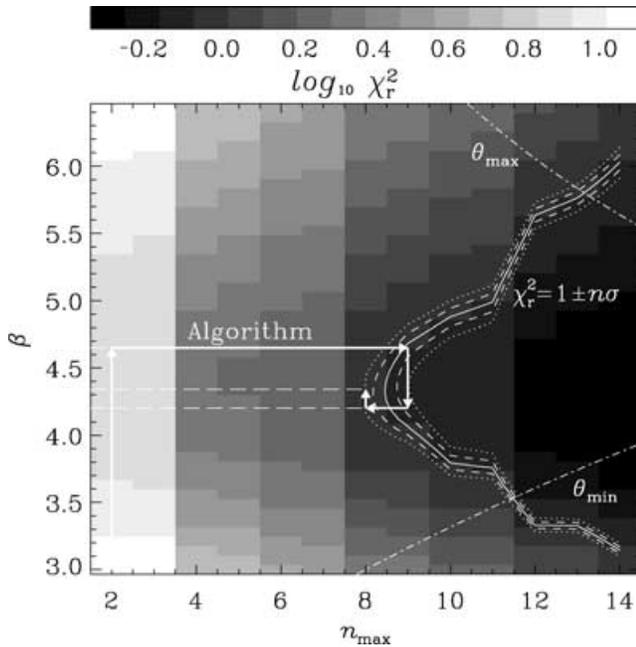


Figure 6. χ_r^2 isocontours on an n_{\max} versus β plane for an elliptical galaxy in the HDF. The roughly horizontal trough is typical, with a well-pronounced excursion of the χ_r^2 contours to lower n_{\max} for well-chosen values of β . The challenge is to locate the leftmost section of the $\chi_r^2 = 1$ contour in an automated and efficient way. The arrows show individual steps (each containing several substeps) taken by our optimization algorithm described in Section 4. Also shown are geometrical θ_{\min} , θ_{\max} constraints and the target $\chi_r^2 = 1$ contour.

The first constraint ensures that the scale size is well suited to efficiently model the image. The second ensures that the shapelet centre matches the object centroid. The third guarantees that sufficient coefficients (n_{\max}) are included to model an object, but with truncation that ‘smooths over’ observational noise. A flatness constraint is also included (in the right-hand side of equation 22). This is particularly important for galaxies with a near neighbour or for very faint objects that have noisy and fragmented χ_r^2 contours. In these cases, including additional shapelet coefficients may not significantly improve a fit, so the series is truncated early.

We apply extra geometrical constraints to the minimum θ_{\min} and maximum θ_{\max} scales of the decomposition, to prevent the model from containing features smaller than the pixel scale or extending off the edge of an image, where it would be unconstrained.

3.4 Automatic optimization algorithm

Satisfying the three conditions (20)–(22) would ensure that a shapelet decomposition uses the optimum values of n_{\max} , β and \mathbf{x}_c . It is easy to determine the values of these parameters once the entire n_{\max} versus β plane has been examined, as in Fig. 6. However, this is a slow process, so we need a practical algorithm to more efficiently explore this parameter space, and to iterate rapidly towards our targets. The numerical implementation of this iteration will inevitably be non-trivial, because it combines both minimization and root finding, in a space with one axis discrete. Here we describe a code that we have developed to repeatedly decompose an object into shapelets, test the residual, and improve the decomposition parameters. Its stepwise approach is shown in Fig. 6, and the full code can be downloaded from the World Wide Web.

Objects are first detected in an image using SETRACTOR (Bertin & Arnouts 1996), a friends-of-friends peak-finding algorithm. After experimenting on various data sets, we have found the results of SETRACTOR highly sensitive to input settings. To avoid reliance upon these settings, we use SETRACTOR as sparingly as possible. We set low detection thresholds in order to obtain a complete catalogue, and filter out false detections later. We use the measurement of the FWHM of each object to make an initial guess at β , and also to set the size of the fixed, circular ‘postage stamp’ region that is extracted around each object. We aim for a postage stamp large enough to contain the entire object, but small enough to isolate it from its neighbours and to make the routine computationally efficient. We then use the SETRACTOR segmentation map to identify pixels in the postage stamp but well away from any object. These are used to estimate the background noise level, or to locally renormalize the pixel weight map. Within reasonable limits, the process is stable with respect to such parameters and we shall not be too concerned as to the exact SETRACTOR settings.

Using constant $n_{\max} = 2$ for speed, β is varied in order to minimize χ_r^2 and satisfies the criterion in equation (20), via a 1D version of the *Numerical Recipes* AMOEBA routine: crawling vertically in Fig. 6. During each step of this iteration, the centroid is simultaneously shifted to re-zero the series in equation (51) in the shapelet coefficients and thus satisfy the criterion in equation (21). As the calculation of the centroid is independent of β for isolated objects (see Section 7), this part of the iteration is both stable and fast. Fig. 6 also shows the additional geometrical constraints of $\theta_{\min} > 0.2$ pixel and θ_{\max} not falling off the edge of the postage stamp. These act as hard boundaries to the region of parameter space that the amoeba is allowed to explore.

Once the optimum β has been found, n_{\max} is increased until the criterion in equation (22) is satisfied: crawling horizontally in Fig. 6. The increases are performed in steps of two, because even n states frequently improve the fit more than odd n states (primarily due to the addition of a new $f_{n,0}$ circular state). The value of n_{\max} is fine-tuned to the exact best value at the end. If two values of n_{\max} both allow a decomposition with $\chi_r^2 = 1 \pm 1\sigma$, the lower value is taken.

If the object warranted more coefficients than the initial guess of $n_{\max} = 2$, β and \mathbf{x}_c are again readjusted at the new n_{\max} , using our 1D AMOEBA routine. Another n_{\max} search then starts back at $n_{\max} = 2$ and increases again in steps of two. The algorithm terminates when either the horizontal or vertical search returns to the same value as it started. All three conditions in equations (20)–(22) have then been met. Computation time for each object increases $\propto n_{\max}^4$. On a single 2-GHz processor, our algorithm takes ~ 45 min to process all of the 3596 objects detected in the HDF North.

A selection of reasonably bright HDF galaxies is shown with their shapelet models in Fig. 7. The right-hand column shows the reconstruction residuals, which are consistent with noise even for irregular galaxy morphologies. A comparison of their shapelet-based shape estimators to traditional SETRACTOR measurements is shown in Fig. 8.

4 DECOMPOSITION OF REAL DATA

4.1 Least-squares fitting

Unlike the continuous, analytic formalism presented in Section 2, real images are complicated by pixelization, PSF convolution and noise. In order to incorporate these effects, we shall first adopt a somewhat different approach to shapelet decomposition than the overlap integrals (3) and (5). We shall instead fit shapelet coefficients

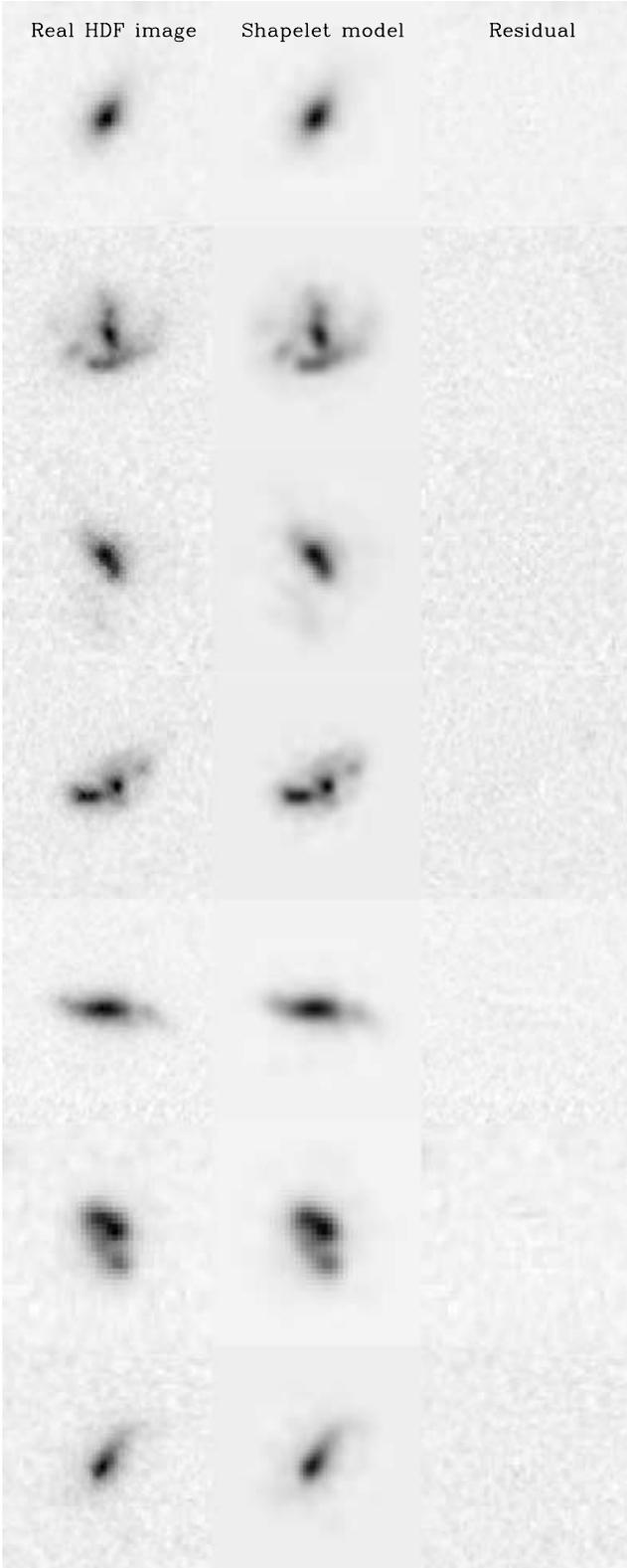


Figure 7. Shapelet models of a selection of HDF galaxies, with their shapelet scale size β and maximum order n_{\max} determined automatically. In all cases, the image residuals are entirely consistent with noise. Our code to perform this task, by minimizing the least-squares difference between the model and input images, is described in the text.

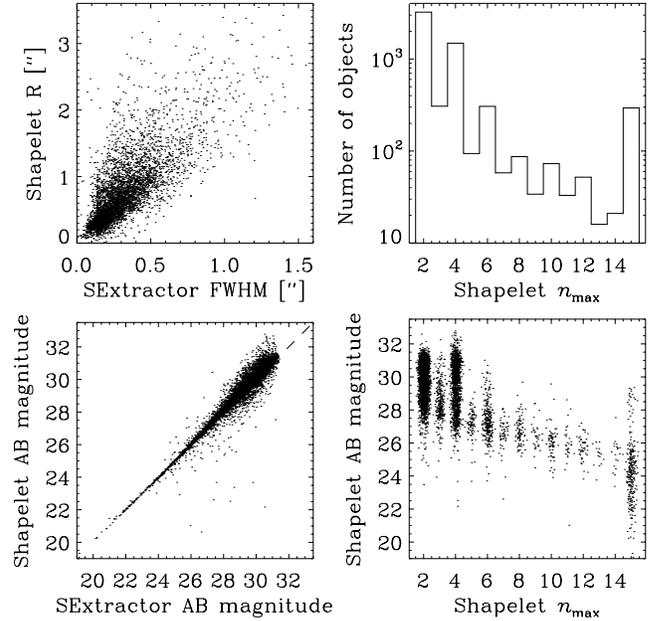


Figure 8. The successful recovery of object statistics from the shapelet parameters of HDF galaxies. For comparison to the `SEXTRACTOR` measurements, shapelet size measurements are shown without PSF deconvolution. In the right-hand panels, galaxies requiring $n_{\max} \geq 15$ coefficients have been forced into the final bin, and in the bottom-right panel, points have been randomly offset a small amount for clarity.

to the data using a least-squares method. As the model $f_{\text{rec}}(\mathbf{x})$ in equation (18) is linear in the shapelet coefficients, we can solve for the minimum χ_r^2 solution (18) exactly. We obtain (see Lupton 1993; Chang & Refregier 2002)

$$\mathbf{f}_{n,m} = (M^T V^{-1} M)^{-1} M^T V^{-1} \mathbf{f}_{x,y}, \quad (23)$$

where $\mathbf{f}_{n,m}$ is a vector of the derived shapelet coefficients, $\mathbf{f}_{x,y}$ is the surface brightness in each pixel arranged as a data vector, V is the covariance matrix between pixel values and M is a matrix of each shapelet basis function evaluated in each pixel. A fit achieving $\chi_r^2 = 1$ has successfully modelled all significant spatial variation in the image and removed observational noise.

If the noise per pixel is known, 1σ confidence limits can be derived on all of the assigned coefficients using this fitting method (Lupton 1993). If a complete pixel noise map is available (e.g. from multiple exposures stacked using `DRIZZLE` software – Fruchter & Hook 2002), it can be used to down-weight noisy pixels where cosmic rays or hot/cold pixels were present in some of the exposures. Although the code available on the World Wide Web simply uses a diagonal matrix for V that contains only the noise level in each pixel, the method is, in general, able to use the full covariance matrix that contains the amount of covariance between different pixels. In real data, the flux in adjacent pixels is indeed slightly correlated because of convolution with the PSF and also because of additional aliasing effects introduced by `DRIZZLE`. If this effect is important, the pixel-to-pixel covariances could be estimated from empty regions of an image and included in the calculation. In particular, this may have a small improvement on statistics measured from very small objects (cf. Massey et al. 2004).

A constant background level can also be removed using this method, by adding an undetermined constant to the set of basis functions. Poor flat-fielding or local background gradients near a bright object can also be fitted and removed by adding a plane with a

variable slope. Although these functions are not strictly orthogonal, the procedure works well in practice as long as there are sufficient pixels around the fitted object that contain only background noise.

4.2 PSF deconvolution

All real images are inevitably seen after convolution with a point spread function. In astronomy, this is typically caused by atmospheric turbulence or ‘seeing’ (for ground-based observatories), aperture diffraction at the primary mirror and imperfect telescope tracking or optics. The combination of such effects can be measured from the size and shape of stars observed in an image (because these distant objects would be point-like in the absence of a PSF), and can be fitted with a shapelet model in the same way as the galaxies. Shapelets I presented the matrix operation for convolving an image with a Gaussian PSF in shapelet space. Shapelets II extended this derivation to a general PSF and demonstrated PSF *deconvolution* via matrix inversion. However, the inversion of the PSF matrix is potentially slow and may be numerically unstable. Our least-squares fitting method will allow us to elegantly sidestep this process by convolving the basis functions with the PSF model in advance, then fitting this new basis set to the data. The returned shapelet model, reconstructed using the unconvolved basis functions, will be automatically deconvolved from the PSF.

The formalism for convolution in shapelet space is presented in Shapelets I Section 4 and involves three separate scale sizes for three separate objects: α for the unconvolved model, β for the PSF and γ for the convolved model (there are also corresponding values of n_{\max}^{α} , n_{\max}^{β} and n_{\max}^{γ}). We assume that β is known. We can optimize α as in Section 3.3. However, the choice of γ is a matter entirely internal to the fitting procedure. Just as before, if $n_{\max}^{\gamma} \rightarrow \infty$, any γ will work (but this time without increasing the number of external free parameters in the model). In practice, however, it is still necessary to truncate this series somewhere. Note that $\gamma^2 = \alpha^2 + \beta^2$ was incorrectly suggested as a ‘natural choice’ for this parameter in Shapelets I. Another choice would be $\gamma = \alpha$, which, with $n_{\max}^{\gamma} = n_{\max}^{\alpha}$, makes the convolution matrix $P_{n,m}$ symmetric and thus simplifies its calculation.

The optimum values for γ and n_{\max}^{γ} are, in fact, obtained from an argument concerning the information present in shapelet coefficients. A shapelet model contains information only between minimum and maximum scales

$$\theta_{\min} = \frac{\beta}{\sqrt{n_{\max} + 1}} \quad \text{and} \quad \theta_{\max} = \beta \sqrt{n_{\max} + 1}. \quad (24)$$

During convolution, θ_{\min}^{α} and θ_{\min}^{β} add in quadrature to produce θ_{\min}^{γ} ; similarly for θ_{\max}^{γ} . γ and n_{\max}^{γ} should therefore be chosen to precisely capture the information contained on all of these new scales. Writing $(n_{\max}^{\alpha} + 1)$ as N_{α} , etc. for brevity, we find

$$\gamma = \sqrt{[4] \frac{(\alpha^2 N_{\alpha} + \beta^2 N_{\beta})(\alpha^2 N_{\beta} + \beta^2 N_{\alpha})}{N_{\alpha} N_{\beta}}} \quad (25)$$

and

$$n_{\max}^{\gamma} = \sqrt{\frac{\alpha^2 N_{\alpha} + \beta^2 N_{\beta}}{\alpha^2 N_{\beta} + \beta^2 N_{\alpha}}} N_{\alpha} N_{\beta} - 1. \quad (26)$$

The PSFs of cameras on board the *Hubble Space Telescope* are well known and stable. Fig. 9 shows an oversampled TINYTIM (Krist & Hook 1997) model of the *Wide-Field Planetary Camera 2* (WFPC2) PSF, raytraced through an engineering model, plus charge diffusion to simulate photon capture within the CCD cameras. This is easy to model with shapelets, except for the fact that its

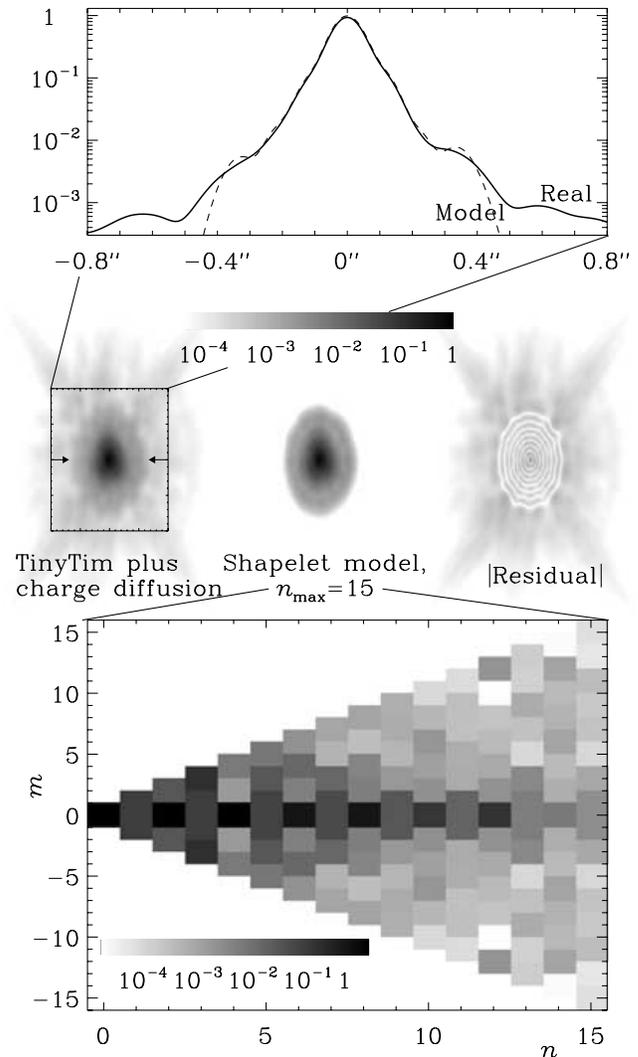


Figure 9. Shapelet model of the TINYTIM (Krist 1995) WFPC2 PSF plus charge diffusion. Top panel: a horizontal slice through the centre of the PSF. Bottom panel: the moduli of its polar shapelet coefficients to $n_{\max} = 15$. Note that the amplitude scales are all logarithmic: the core is actually modelled very successfully out to the second diffraction ring. For speed we do not bother capturing the wings.

steep cusp and extended wings are intrinsically ill-matched to the Gaussian around which shapelets are constructed. The PSF is shown in the figure beside a shapelet decomposition up to $n_{\max}^{\text{PSF}} = 15$. This is sufficient to accurately capture the core and the first two diffraction rings, which are already more than two orders of magnitude below the maximum, but does not extend to the four faint diffraction spikes or far into the low-level wings (note that the colour scales are logarithmic). In principle, this could be further extended at a cost to processor time by using more shapelet coefficients.

Fig. 10 demonstrates successful PSF deconvolution. A galaxy from the HDF is convolved with the WFPC2 PSF (in real space). This is treated as the observed image, and deconvolved from the PSF using a shapelet fit. The resulting reconstruction is in good agreement with the original galaxy image, as can be seen from its small residual. Note that the optimum scale size β for the model is slightly lower when PSF deconvolution is performed. This reflects the need to capture finer details.

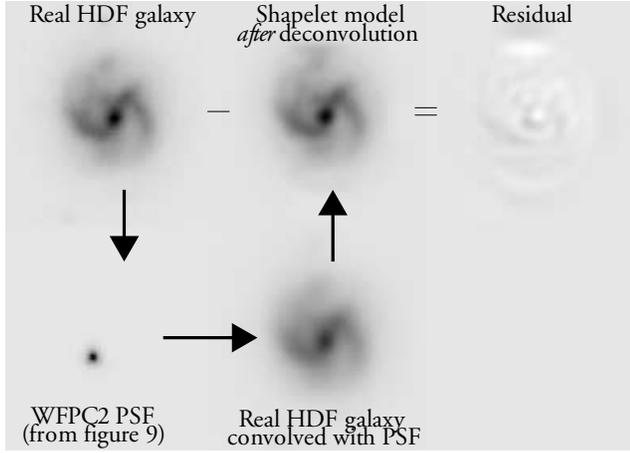


Figure 10. Demonstration of deconvolution from an observational PSF. Top left-hand panel: a real HDF galaxy. Bottom left-hand panel: the WFPC2 PSF model, from Fig. 9 but displayed here with a linear colour scale. Bottom middle panel: the galaxy image convolved with the PSF. This convolution has been performed in real space, to disassociate the operation from anything involving shapelets. Top middle panel: a shapelet reconstruction and deconvolution of the galaxy to $n_{\max} = 20$, obtained from a fit to the convolved image, assuming knowledge of the PSF. Top right panel: the difference between the true galaxy image and the shapelet model after deconvolution. This small residual demonstrates the success of PSF deconvolution using shapelets.

4.3 Pixelization

Real image data are typically stored in discrete pixels. To link this to the analytic shapelet formalism, one must either smooth the data or pixelate the shapelet basis functions. Smoothing the data requires an arbitrary interpolation scheme to be defined, and resampling the data on to smaller pixels can be very slow. A better approach is to leave the data alone and to discretize the smooth shapelet basis functions. This reduces the integrals in equations (3) and (5) to sums over pixel values, which are fast to compute. However, they are no longer analytically exact. We therefore need to define a discretization scheme that keeps the basis functions as orthogonal as possible, and the integrals as accurate as possible.

As pointed out by Berry, Hobson & Withington (2004), one cannot simply adopt the value of basis functions at the centre of each pixel. Basis functions that contain oscillations on scales smaller than the pixel size are sampled in an essentially random manner. Their discrete versions are then neither representative of the analytic function nor orthogonal. Degeneracies are introduced between shapelet coefficients during the decomposition that destabilize the inversion of coefficient matrices in the reconstructed model, and bias quantities such as the flux of an object. Fortunately, this is rarely a problem in practical cases, because we can choose n_{\max} and β in advance to isolate only those basis functions that contain oscillations on scales larger than the pixel (or seeing) size. Under these conditions, Berry et al. (2004) show that the shapelet basis functions are indeed orthogonal.

We suggest an even safer alternative here. The Cartesian basis functions are separable in x and y , and may be analytically integrated within rectangular pixels. This is exactly the same process undergone by photons arriving at a CCD, where the smooth function of a real scene becomes binned into digital squares. Once we have convolved the basis functions with the PSF, and integrated them within pixels, they can be suitably matched to the data.

To integrate the 2D Cartesian basis functions, first consider the 1D basis functions from Shapelets I,

$$\phi_n(x) \equiv [2^n \pi^{1/2} n! \beta]^{-1/2} H_n \left(\frac{x}{\beta} \right) e^{-x^2/2\beta^2}. \quad (27)$$

Integrating by parts and using two well-known identities (see, e.g., Boas, chapter 12)

$$H_n(x) = 2x H_{n-1}(x) - 2(n-1)H_{n-2}(x) \quad (28)$$

and

$$\frac{dH_{n-1}(x)}{dx} = 2(n-1)H_{n-2}(x), \quad (29)$$

one can obtain the recurrence relation

$$I_n \equiv \int_a^b \phi_n(x) dx \quad (30)$$

$$= -\beta \sqrt{\frac{2}{n}} [\phi_{n-1}(x)]_a^b + \sqrt{\frac{n-1}{n}} I_{n-2}. \quad (31)$$

Finally, note that

$$I_0 = \sqrt{\frac{\beta \pi^{1/2}}{2}} [\text{erf}(x)]_a^b \quad \text{and} \quad (32)$$

$$I_1 = -\sqrt{2\beta} [\phi_0(x)]_a^b. \quad (33)$$

This supplies all the necessary integrals. As the 2D Cartesian basis functions are separable in x and y , it is easy to extend this derivation to integrate within square CCD pixels:

$$I_{n_1, n_2} = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \phi_{n_1}(x) \phi_{n_2}(y) dx dy = I_{n_1} \times I_{n_2}, \quad (34)$$

where, if there is no ‘dead zone’ around the edge of a pixel, $(b_1 - a_1) \times (b_2 - a_2)$ is the angular size of a pixel. A missing pixel border, due for instance to electronics which is unresponsive to light, can be included by altering the limits on the integral.

We can either use this result to obtain a model in Cartesian shapelet space, which can later be converted to a polar shapelet representation using equation (14), or we can integrate the polar shapelet basis functions within pixels using the same equations. This integration is a particularly important advance for small galaxies or for shapelet basis functions at high n , that can contain oscillations smaller than a single pixel.

The symmetries of polar shapelets can also be used to integrate models within circular apertures using equations (46)–(49).

5 COORDINATE TRANSFORMATION

Image manipulation via linear transformations is simple in shapelet space. As in Shapelets I, let us consider an infinitesimal coordinate transformation $\mathbf{x} \rightarrow (1 + \Psi)\mathbf{x} + \epsilon$, where $\epsilon = \{\epsilon_1, \epsilon_2\}$ is a displacement and Ψ is a 2×2 matrix parametrized as

$$\Psi = \begin{pmatrix} \kappa + \gamma_1 & \gamma_2 - \rho \\ \gamma_2 + \rho & \kappa - \gamma_1 \end{pmatrix}. \quad (35)$$

The parameters ρ , κ , ϵ and γ_i correspond to infinitesimal rotations, dilations, translations and shears.

An image transforms as $f(\mathbf{x}) \rightarrow f'(\mathbf{x}) \simeq f(\mathbf{x} - \Psi\mathbf{x} - \epsilon)$, which can be written as

$$f' \simeq (1 + \rho \hat{R} + \kappa \hat{K} + \gamma_i \hat{S}_i + \epsilon_i \hat{T}_i) f, \quad (36)$$

where \hat{R} , \hat{K} , \hat{S}_i and \hat{T}_i are the operators generating rotation, convergence, shears and translations, respectively. We adopt a notation from weak gravitational lensing, where a ‘convergence’ κ corresponds to a change in the radius of an object by a factor $(1 - \kappa)^{-1}$. These transformations can be viewed as a mapping of $f_{n,m}$ coefficients in shapelet space. For example, a finite rotation is

$$\hat{R} : f_{n,m} \rightarrow f'_{n,m} = f_{n,m} e^{im\rho}, \quad (37)$$

so a rotation through 180° can be written as

$$\hat{R}_{180^\circ} : f_{n,m} \rightarrow f'_{n,m} = (-1)^m f_{n,m}. \quad (38)$$

An (infinitesimal) dilation can be performed in polar shapelet space by mapping the shapelet coefficients as

$$\begin{aligned} \hat{K} : f_{n,m} \rightarrow f'_{n,m} &= (1 + \kappa) f_{n,m} \\ &+ \frac{\kappa}{2} \sqrt{(n-m)(n+m)} f_{n-2,m} \\ &- \frac{\kappa}{2} \sqrt{(n-m+2)(n+m+2)} f_{n+2,m}. \end{aligned} \quad (39)$$

The shapelet model may require more coefficients after this transformation. Note that this dilation operation increases both the flux and the image area by a factor of $1 + 2\kappa$, thus conserving surface brightness. To instead perform a dilation that conserves the total flux, divide the right-hand side of equation (39) by this factor. To first order, this is

$$\begin{aligned} \hat{K} : f_{n,m} \rightarrow f'_{n,m} &= (1 - \kappa) f_{n,m} \\ &+ \frac{\kappa}{2} \sqrt{(n-m)(n+m)} f_{n-2,m} \\ &- \frac{\kappa}{2} \sqrt{(n-m+2)(n+m+2)} f_{n+2,m}. \end{aligned} \quad (40)$$

In Section 6, we shall ensure that shape estimators for a shapelet model are independent of the scalefactor chosen for the decomposition by ensuring that the estimators are unchanged under this mapping.

Rather than these first-order approximations, finite dilations can be performed to all orders using the rescaling matrix in the appendix of Shapelets I. This is identical to the convolution matrix, but the image is convolved with a δ -function.

Shears and translations can be performed using

$$\begin{aligned} \hat{S} : f_{n,m} \rightarrow f'_{n,m} &= f_{n,m} \\ &+ \frac{\gamma_1 + i\gamma_2}{4} [\sqrt{(n+m)(n+m-2)} f_{n-2,m-2} \\ &- \sqrt{(n-m+2)(n-m+4)} f_{n+2,m-2}] \\ &+ \frac{\gamma_1 - i\gamma_2}{4} [\sqrt{(n-m)(n-m-2)} f_{n-2,m+2} \\ &- \sqrt{(n+m+2)(n+m+4)} f_{n+2,m+2}] \end{aligned} \quad (41)$$

and

$$\begin{aligned} \hat{T} : f_{n,m} \rightarrow f'_{n,m} &= f_{n,m} \\ &+ \frac{\epsilon_1 + i\epsilon_2}{2\sqrt{2}} [\sqrt{(n+m)} f_{n-1,m-1} \\ &- \sqrt{(n-m+2)} f_{n+1,m-1}] \\ &+ \frac{\epsilon_1 - i\epsilon_2}{2\sqrt{2}} [\sqrt{(n-m)} f_{n-1,m+1} \\ &- \sqrt{(n+m+2)} f_{n+1,m+1}] \end{aligned} \quad (42)$$

with the translation specified in units of β .

Other image manipulations can also be represented as mappings of shapelet coefficients. Changes of flux by a factor B are trivially implemented by the mapping

$$\hat{B} : f_{n,m} \rightarrow f'_{n,m} = B \times f_{n,m}. \quad (43)$$

It is also possible to circularize an object with the mapping (see Section 3.1)

$$\hat{C} : f_{n,m} \rightarrow f'_{n,m} = f_{n,m} \delta_{m0}, \quad (44)$$

or to flip the parity of an object by reflection in the x -axis using

$$\hat{P} : f_{n,m} \rightarrow f'_{n,m} = f_{n,m}^*. \quad (45)$$

Combining this \hat{P} with the rotation operator allows reflections to be performed in any axis.

The actions of these operators are demonstrated upon a real galaxy image in Fig. 11.

6 OBJECT SHAPE MEASUREMENT

The above symmetries of the polar shapelet basis functions can be used to identify combinations of shapelet coefficients that measure the flux (photometry), centroid position (astrometry) and the size of an object. Similar weighted combinations of Cartesian shapelet coefficients were found in Shapelets I, but we find the interpretation of polar shapelets more intuitive, and the expressions below are usually simpler than their Cartesian equivalents. For example, the rotationally invariant part of an object is isolated into its $m = 0$ coefficients. The linear offset of an object from the origin is described by its $m = \pm 1$ coefficients and the ellipticity of an object by its $m = \pm 2$ coefficients. In the latter cases, the magnitude of the coefficients indicate an amplitude and the phases a direction.

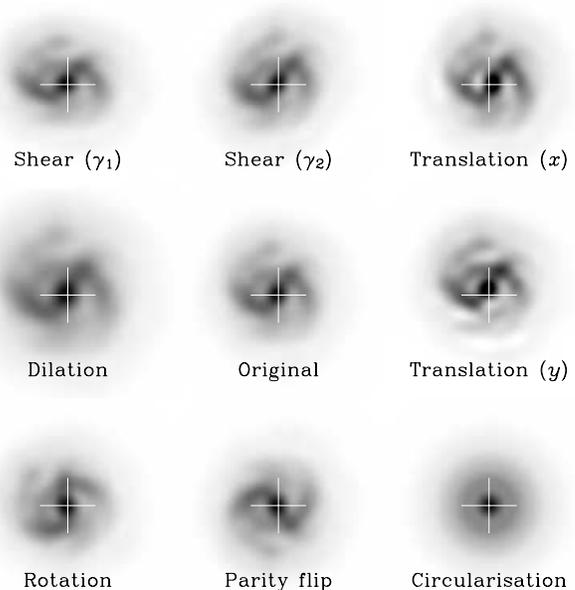


Figure 11. Some simple operations applied to a real galaxy image, by using the polar shapelet ladder operators or coefficient mappings as described in the text. The central image is the original galaxy. Starting at the bottom left and proceeding clockwise, the other images show rotation by 40° , dilation of $\kappa = 0.15$, shears of $\gamma = 10$ per cent, translations, circularization and reflection in the x -axis.

6.1 Photometry

Practical measurements of the flux of an object usually introduce a Gaussian or top-hat weight function in order to limit contamination from surrounding noise and nearby objects. The flux of a shapelet model inside a circular aperture can be calculated using only the coefficients with $m = 0$. All other coefficients correspond to basis functions with positive and negative regions that cancel out under integration around θ . From equations (16) and (17) for the radial profile of an object, we find that

$$\int_0^{2\pi} \int_0^R f(r) r dr d\theta = (4\pi)^{1/2} \beta \sum_n^{\text{even}} f_{n,0} I_n, \quad (46)$$

where

$$I_n = \frac{(-1)^{n/2}}{\beta^2} \int_0^R L_{n/2}^0 \left(\frac{r^2}{\beta^2} \right) e^{-r^2/2\beta^2} r dr. \quad (47)$$

Using relation (A9) to integrate by parts, we can find a recursion relation

$$I_n = (-1)^{n/2} \left[1 - L_{n/2}^0 \left(\frac{R^2}{\beta^2} \right) e^{-R^2/2\beta^2} - 2 \sum_{i=0}^{(n-2)/2} (-1)^i I_{2i} \right], \quad (48)$$

and a closed form

$$I_n = 1 - e^{-R^2/2\beta^2} \left[2 \sum_{i=0}^{n/2} (-1)^i L_i^0 \left(\frac{R^2}{\beta^2} \right) - (-1)^{n/2} L_{n/2}^0 \left(\frac{R^2}{\beta^2} \right) \right]. \quad (49)$$

However, the imposition of an integration boundary is unnecessary with shapelets because the model is analytic and noise-free. In the limit of $R \rightarrow \infty$, we obtain a simple expression for the total flux in a shapelet model

$$F \equiv \iint_{\mathbb{R}} f(\mathbf{x}) d^2x = (4\pi)^{1/2} \beta \sum_n^{\text{even}} f_{n,0}, \quad (50)$$

a result that can also be recovered by transforming the sum over Cartesian shapelet coefficients from Shapelets I into polar shapelet space via equation (14). Cartesian shapelet models can also be integrated within square apertures using equations (30)–(34).

This extrapolation to large radii does rely upon the faithful representation of an object by a shapelet expansion, and the removal of its noise via series truncation. Such truncation restricts the completeness of the basis functions, and a weight function (constructed from a combination of the allowed basis functions) akin to a ‘prior probability’ is subtly implicit inside our fitting procedure. However, a fitting method such as ours can beat a direct, pixel-by-pixel measurement. Our fit is able to include flux from the extended wings of an object, by integrating it over a large area, even when the signal lies beneath the noise level in any individual pixel. The wings of galaxies in Fig. 7 are indeed well captured by the shapelet models.

6.2 Astrometry

It can similarly be shown that the unweighted centroid (x_c, y_c) is

$$\begin{aligned} x_c + iy_c &\equiv \frac{\iint_{\mathbb{R}} (x + iy) f(\mathbf{x}) d^2x}{\iint_{\mathbb{R}} f(\mathbf{x}) d^2x} \\ &= \frac{(8\pi)^{1/2} \beta^2}{F} \sum_n^{\text{odd}} (n+1)^{1/2} f_{n,1}, \end{aligned} \quad (51)$$

where the summation is over only odd values of n , because only these have the $m = \pm 1$ coefficients that possess the desired rotational symmetries.

6.3 Size

Measures for the size and ellipticity of an object can be derived from the unweighted quadrupole moments of the object,

$$F_{ij} \equiv \iint_{\mathbb{R}} x_i x_j f(\mathbf{x}) d^2x. \quad (52)$$

The rms radius R of an object is given by

$$R^2 \equiv \frac{\iint_{\mathbb{R}} |\mathbf{x}|^2 f(\mathbf{x}) d^2x}{\iint_{\mathbb{R}} f(\mathbf{x}) d^2x} \quad (53)$$

$$= \frac{F_{11} + F_{22}}{F} = \frac{(16\pi)^{1/2} \beta^3}{F} \sum_n^{\text{even}} (n+1) f_{n,0}. \quad (54)$$

Integrals (46)–(49) can also be used to calculate Petrosian radii that enclose a specified fraction of the total flux within a circular aperture.

6.4 Ellipticity

The unweighted ellipticity of an object can also be calculated from its quadrupole moments.

$$\varepsilon \equiv \frac{F_{11} - F_{22} + 2iF_{12}}{F_{11} + F_{22}} = \frac{(16\pi)^{1/2} \beta^3}{FR^2} \sum_n^{\text{even}} [n(n+2)]^{1/2} f_{n,2}, \quad (55)$$

where the complex ellipticity notation of Blandford et al. (1991), with $\varepsilon = |e| \cos 2\theta + i |e| \sin 2\theta$, arises here naturally.

7 GALAXY MORPHOLOGY CLASSIFICATION

The shapelet decomposition of an object captures its entire structure, and useful information is frequently found in coefficients of higher order than those considered above. In particular, galaxy morphologies are well known to provide an indication of their physical properties, local environment and formation history. The classical ‘Hubble sequence’ of morphological types has been recently improved by several shape estimators that attempt to classify galaxies in a more quantitative manner, which correlates directly with the physical properties of interest (e.g. Simard 1998; Bershadsky, Jangren & Conselice 2000; van den Bergh 2002).

It is possible to manufacture such morphology diagnostics from weighted combinations of shapelet coefficients. Introducing shapelets to this field allows a measurement to take advantage of our robust treatment of noise, pixelization and PSF deconvolution. The shapelet expressions for existing shape measures are frequently elegant; and the natural symmetries in shapelets also suggest new diagnostics.

7.1 General scale-invariant quantities

One approach is to consider general shape estimators Q , formed from a linear combination of shapelet coefficients, as an extension of the previous section

$$Q = \beta^s \sum_{n,m} w_{n,m} f_{n,m}, \quad (56)$$

where $w_{n,m}$ are arbitrary weights and the exponent s sets the dimension of the estimator. These are also linear in the surface brightness of a galaxy. We initially restrict ourselves to using those combinations which are independent of β to at least first order. This ensures that the choice of the scalefactor does not affect the final result, and is also equivalent to invariance under object dilations (40). We can then impose further constraints that the estimator must be independent of or linearly dependent upon the various other operations described in Section 5. Setting $\partial Q/\partial\beta = 0$ and using the result that

$$\frac{\partial f_{n,m}}{\partial\beta} = \frac{1}{2\beta} [\sqrt{(n+m+2)(n-m+2)} f_{n+2,m} - \sqrt{(n+m)(n-m)} f_{n-2,m}] \quad (57)$$

it is easy to show that we require

$$w_{n,m} = \frac{2s}{\sqrt{(n+m)(n-m)}} w_{n-2,m} + \sqrt{\frac{(n+m-2)(n-m-2)}{(n+m)(n-m)}} w_{n-4,m}. \quad (58)$$

Note that all quantities so formed mix coefficients with only one value of $|m|$. This can be chosen to give Q the desired properties under rotation. Any term on the right-hand side should be ignored if it refers to non-existent states with negative n . The normalization of the first term in each series, $w_{m,m}$, is arbitrary: this can be set to ensure independence to changes of object flux.

Setting $(s, m) = (1, 0), (2, 1), (3, 0)$ and $(3, 2)$ recovers the flux F , centroid x_c , rms square radius R^2 and ellipticity ε , up to the normalization factor of F^{-1} for the latter three quantities. This proves that these are indeed the *only* β -invariant linear quantities with such dimensionality and rotational symmetries. Furthermore, as equations (50), (51), (54) and (55) describe for *unweighted* moments, they must in fact be independent of β to all orders.

All of these basic shape estimators converge for any galaxy with a shapelet spectrum steeper than n^{-2} . This includes both spiral galaxies with an exponential profile, and elliptical galaxies with a ‘de Vaucouleurs’ profile, as long as n_{\max} is kept sufficiently low to prevent the high- n coefficients from modelling background noise at large radii. The flux and centroid estimators converge most rapidly, so are least sensitive to the choice of n_{\max} . The error on these series due to truncation can be calculated using any of a range of methods for generic Taylor series in, for example, Boas (1983).

7.2 Concentration

We can extend this sequence by raising s further. For example, setting $s = 5$ and $m = 0$ gives the 2D unweighted kurtosis of the image, producing an estimate of the concentration of the object. Unfortunately, such a high value of s yields a series of shapelet coefficients that does not converge for galaxies with a de Vaucouleurs or exponential radial profile.

We have also noted that a ratio of the two existing shapelet scale sizes, R and β , also works rather well as a concentration index (although this estimator is not independent of β). Further work will need to be performed to calibrate this estimator to the physical properties of galaxies.

An alternative approach is to mimic pre-existing, and pre-calibrated, morphology diagnostics. Bershady et al. (2000) define a concentration index

$$C \equiv 5 \log \left(\frac{r_{80}}{r_{20}} \right), \quad (59)$$

where r_{80} and r_{20} are the radii of circular apertures containing 80 and 20 per cent of the total flux of the object. This correlates well with a Hubble-type galaxy (Bershady et al. 2000) and also its mass (Conselice, Gallagher & Wyse 2002). Integrals (46)–(49) can be evaluated for various values of R , to find r_{80} and r_{20} , and thus calculate this quantity for a shapelet model.

7.3 Asymmetry

Conselice, Bershady & Jangren (2000) define an asymmetry index

$$A \equiv \frac{\sum_{\text{pixels}} |f(x, y) - f^{180^\circ}(x, y)|}{\sum_{\text{pixels}} f(x, y)}, \quad (60)$$

where the superscript denotes an image rotated through 180° . A term dealing with the background noise and sky level has been omitted here, as these are automatically dealt with during the shapelet decomposition process in Section 4. The asymmetry correlates with the star formation rate of a galaxy (Conselice et al. 2000), and high asymmetry values often indicate recent galaxy interactions or mergers.

In a shapelet expansion, all of the information concerning galaxy asymmetry is contained in coefficients with odd m (and n). Using the orthonormality condition (9) and rotating via equation (38), we find the simple form

$$A = \frac{\sqrt{2}\beta}{\pi F} \sum_{n,m}^{\text{odd}} |f_{n,m}|. \quad (61)$$

Estimators of asymmetry under rotations of 120° or 90° can also be formed from sums of shapelet coefficients with m not divisible by 3 or 4, respectively.

7.4 Chirality

A quadratic combination of shapelet coefficients can be used to describe the ‘chirality’ or ‘handedness’ of an object. One dimensionless estimator $\chi_{|m|}$ can be formed for each value of $|m|$, to trace the relative rotation of those coefficients, with increasing n . This is roughly equivalent to tracing the rotation of the isophotes of a galaxy with increasing radius. For example, the galaxy shown in Fig. 3 has two prominent spiral arms that unwind in a clockwise sense, so it has a large, positive value of χ_2 .

We require that the chirality estimators should be invariant under global rotation of the object, invariant under changes of flux, invariant to first order under changes of β and to flip sign when the object is mirror-imaged. These conditions uniquely specify

$$\chi_{|m|} = \frac{\beta^2}{F^2} \sum_{n=m}^{\infty} \sum_{n'=n+2}^{\infty} w_{n,n'} f_{n,m}^* f_{n',m}, \quad (62)$$

where $w_{m,m+2} = 1$ and

$$\sqrt{(n'+m)(n'-m)} w_{n,n'} = 4w_{n,n'-2} + \sqrt{(n+m)(n-m)}, \quad (63)$$

thus mixing all coefficients with the same value of m .

This estimator has yet to be calibrated against physical quantities. However, this approach ought to be able to automatically distinguish between elliptical galaxies and spiral galaxies in way that mimics a visual classification, and could also be adapted as a function of n_{\max} to find bars in the cores of spiral galaxies.

8 CONCLUSIONS

We have extended the formalism of shapelets for image analysis from basis functions separable in Cartesian to polar coordinates. Cartesian shapelets are convenient for the initial object decomposition. In particular, we have shown that they can analytically be integrated inside a square boundary, thus facilitating the pixelization of the smooth basis functions. On the other hand, polar shapelets decompose an object into components with explicit rotational symmetries, and often have a more direct physical interpretation. In addition, they yield more compact representations of typical galaxy images, as terms with low orders of rotational symmetry tend to dominate.

We have quantitatively investigated the effects of the choice of the shapelet scale size parameter, β . For most objects in astronomical images, one scale size is clearly optimal for high-quality image reconstruction, data compression and the fast convergence of shapelet estimators. We have developed a practical algorithm to find this value of β for arbitrary objects in real images, plus optimum values for the shapelet centre \mathbf{x}_c and truncation order n_{\max} . This algorithm can also take into account observational effects including noise, pixelization and PSF deconvolution.

We have then described a number of applications of polar shapelets. Shapelet models can be rotated, enlarged and sheared by simple analytic operations. As the shapelet basis functions are invariant under Fourier transform, analytic convolutions and deconvolutions (e.g. from a PSF) are also easy to perform. Linear combinations of the polar shapelet coefficients of an object produce elegant expressions for its flux (photometry), position (astrometry) and size. We also showed how other combinations of shapelet coefficients can be used to distinguish between morphological types of galaxies. A complete IDL software package to perform the image decomposition and shapelet image analysis is publicly available on the World Wide Web at <http://www.astro.caltech.edu/~rjm/shapelets/>.

ACKNOWLEDGMENTS

The authors thank David Bacon, Gary Bernstein, Sarah Bridle, Tzu-Ching Chang, Mark Coffey, Chris Conselice, Phil Marshall and Jean-Luc Starck for invaluable insights and constructive conversations. Astute suggestions from the referee helped shape this paper into a more logical progression.

REFERENCES

Bernstein G., Jarvis M., 2002, AJ, 123, 583 (BJ02)
 Berry R., Hobson M., Withington S., 2004, MNRAS, 34, 199
 Bershady M., Jangren A., Conselice C., 2000, AJ, 119, 2645
 Bertin E., Arnouts S., 1996, A&AS, 117, 393
 Blandford R., Saust A., Brainerd T., Villumsen J., 1991, MNRAS, 251, 600
 Boas M., 1983, *Mathematical Methods in the Physical Sciences*, Wiley, Chichester
 Conselice C., Bershady M., Jangren A., 2000, ApJ, 529, 886
 Chang T., Refregier A., 2002, ApJ, 570, 447
 Conselice C., Gallagher J., Wyse R., 2002, AJ, 123, 2246

Fruchter A., Hook R., 2002, PASP, 114, 144
 Gerhard O., 1993, MNRAS, 265, 213
 Goldberg D., Bacon D., 2004, ApJ, 619, 741
 Goldberg D., Natarajan P., 2002, ApJ, 564, 65
 Kaiser N., Squires G., Broadhurst T., 1995, ApJ, 449, 460
 Kelly B., McKay T., 2004, AJ, 127, 3089
 Krist J., 2003, *Instrument Sci. Rep. ACS 2003-06*. STScI, Baltimore
 Krist J., Hook R., 1997, *The Tiny Tim User's Guide*. STScI, Baltimore
 Lupton R., 1993, *Statistics in Theory and Practice*. Princeton Univ. Press, Princeton, NJ
 Massey R., Refregier A., Conselice C., Bacon D., 2004, MNRAS, 348, 214
 Refregier A., 2003, MNRAS, 338, 35 (Shapelets I)
 Refregier A., Bacon D., 2003, MNRAS, 338, 48 (Shapelets II)
 Simard L., 1998, in *ADASS VII, ASP, Conf. Ser., Vol., 145*. Astron. Soc. Pac., San Francisco, p. 108
 van den Bergh S., 2002, AJ, 124, 786
 Van der Marel R., Franx M., 1993, ApJ, 407, 525
 Van der Marel R. et al., 1994, MNRAS, 268, 521
 Williams R. et al., 1996, AJ, 112, 1335
 Williams R. et al., 1998, A&AS, 193, 7501
 Wünsche A., 1998, J. Phys. A, 31, 8267

APPENDIX: LAGUERRE POLYNOMIALS

Different conventions have been used to define the Laguerre polynomials, especially before the 1960s. The $p!$ term is omitted from equation (6) in many older books, and caution must be observed with the resulting relations. Several useful recursion relations can be derived to simplify their calculation (e.g. Boas 1983, chapter 12), which we gather here, using our convention, for future reference:

$$L_0^q(x) = 1 \quad (\text{A1})$$

$$L_1^q(x) = 1 - x + q \quad (\text{A2})$$

$$L_p^q(x) = \left(2 + \frac{q-1-x}{p}\right) L_{p-1}^q(x) - \left(1 + \frac{q-1}{p}\right) L_{p-2}^q(x) \quad (\text{A3})$$

$$= L_p^{q+1}(x) - L_{p-1}^{q+1}(x) \quad (\text{A4})$$

$$= \sum_{i=0}^p L_i^{q-1} \quad (\text{A5})$$

$$\frac{dL_p^q(x)}{dx} = x^{-1} [pL_p^q(x) - (p+q)L_{p-1}^q(x)] \quad (\text{A6})$$

$$= -L_{p-1}^{q+1}(x) \quad (\text{A7})$$

$$\frac{dL_p^0(x)}{dx} = \frac{dL_{p-1}^0(x)}{dx} - L_{p-1}^0(x) \quad (\text{A8})$$

$$= -\sum_{i=0}^{p-1} L_i^0. \quad (\text{A9})$$

This paper has been typeset from a $\text{\TeX}/\text{\LaTeX}$ file prepared by the author.