# PRINCIPAL MOMENTS FOR EFFICIENT REPRESENTATION OF 2D SHAPE

*João F. P. Crespo,   Gustavo A. S. Lopes,   Pedro M. Q. Aguiar*

Institute for Systems and Robotics / IST, Lisboa, Portugal
aguiar@isr.ist.utl.pt

## ABSTRACT

The analytic signature is a recently proposed 2D shape representation scheme. It is tailored to the representation of shapes described by arbitrary sets of unlabeled points, or landmarks, because its most distinctive feature is the maximal invariance to a permutation of those points. The shape similarity of two point clouds can then be obtained from a direct comparison of their representations. However, since the analytic signature is a continuous function, performing the comparison of their densely sampled versions may result excessively time-consuming, *e.g.*, when dealing with large databases, even of simple shapes. In this paper we address the problem of efficiently storing and comparing such powerful representations. We start by showing that their frequency spectrum is related to particular complex moments of the shape. From this relation, we derive the bandwidth of the representation in terms of the shape complexity. Using this result, we show that the analytic signature can be described by a small set of complex moments. We call this compact description the *Principal Moments* (PMs) of a shape and show how to efficiently compare shapes using PMs. Our experiments illustrate that the gain in efficiency comes at no cost in performance.

***Index Terms*—** Object recognition, Image shape analysis, Moment methods, Frequency domain analysis, Signal sampling

## 1. INTRODUCTION

Representing shape is a challenging task. In fact, unlike local characteristics like color, which can be uniquely determined by a small set of parameters, or texture, which has been successfully captured by using statistical descriptors, the visual information conveyed by shape, of more global nature and easily perceived by humans, is hard to represent in an appropriate way. This paper deals with two-dimensional (2D) shape representation.

When the 2D shapes to describe are simply connected regions, researchers have used contour-based descriptions, *e.g.*, [1]. Naturally, for more general shapes, usually consisting in arbitrary sets of points, or landmarks, these descriptors are not adequate. If the points describing the shape are labeled, *i.e.*, if the correspondences between the landmarks of two shapes to compare are known, the problem reduces to the impact of geometrical transformations and disturbances, elegantly addressed through the statistical theory of shape [2]. However, in many practical scenarios, the shape points are obtained from an automatic process, *e.g.*, edge or corner detection, thus come without labels or natural ordering.

Estimating the above mentioned correspondences leads to a combinatorial problem, which requires prohibitively time-consuming

algorithms, even for shapes described by a moderate number of landmarks. To circumvent this problem, the analytic signature (ANSIG) of a 2D shape was recently introduced. This signature exhibits not only invariance to the point labels but also fully discrimination, in the sense that different shapes lead to different signatures (being then maximally invariant to the permutation of shape points) [3, 4]. Under this framework, the hard task of comparing the shapes of two sets of unlabeled points reduces to the direct comparison of the corresponding ANSIGs. However, since these signatures are continuous (analytic functions), they are compared through their densely sampled versions, which results computationally demanding when dealing with large databases, even if the shapes are simple.

In this paper, we address the problem of combining the need to work with compact descriptions of 2D shape, for computational efficiency, with the invariance and discriminative power of maximally invariant representations. We start by studying the frequency spectrum of the restriction of the ANSIG to the unit-circle, which, from Cauchy's integral formula, univocally determines the shape signature. We obtain a closed-form expression for the coefficients of the corresponding Fourier series, in terms of particular complex moments of the shape. Then, we derive an upper bound for the magnitude of these coefficients, in terms of the shape complexity, *i.e.*, the number of landmarks. Using these results, we propose a compact representation, which we call the *Principal Moments* (PMs) of the shape, whose length depends on the shape complexity.

While moments of image patterns have been extensively used due to their invariance properties (since at least the early sixties [5]), their discriminative properties have not been studied even in more recent related work, *e.g.*, [6, 7, 8, 9]. In opposition, our PMs have the distinguishing feature of inheriting the discriminative characteristic of ANSIG (due to its maximal invariance). In fact, as it will become clear below, the usage of PMs can also be seen as the adoption of a guided choice for the sampling interval of the ANSIG, in terms of the shape complexity. Since, under the proposed framework, shapes of distinct complexities are described by vectors of different lengths, we also describe how to efficiently compare them.

## 2. THE ANALYTIC SIGNATURE OF A SHAPE

Consider a 2D shape described by a set of $P$ (unlabeled) points, or landmarks, $\{(x_i, y_i), 1 \leq i \leq P\}$. The shape is then given by the complex vector $\mathbf{z} = [z_1, z_2, \ldots, z_P]^T$, where $z_i = x_i + jy_i$. Naturally, since the points do not have labels, the same shape can be described by any vector obtained from $\mathbf{z}$ by re-ordering its entries.

To cope in a simple way with this myriad of descriptions for the same object, a permutation invariant representation was introduced in [3, 4]: the analytic signature (ANSIG). The ANSIG is an analytic function on the complex plane, obtained from $\mathbf{z}$ through

$$a(\boldsymbol{z}, \xi) = \frac{1}{P} \sum_{i=1}^{P} e^{z_i \xi} . \qquad (1)$$

The adequacy of representing the shape in **z** by its ANSIG hinges on the following facts, formalized and proved in [3] as the maximal invariance with respect to the permutation group: (i) any vector obtained from **z** by re-ordering its entries, leads to the same ANSIG (permutation invariance); and (ii) any vector with landmarks lying at different positions than the ones in **z**, leads to a different ANSIG (discrimination). This justifies that (1) is in fact as a *shape signature*.

The ANSIG copes with geometric transformations. In [3] it is shown that the simple pre-processing step of working with $\sqrt{P}(\mathbf{z} - \bar{\mathbf{z}})/\|\mathbf{z} - \bar{\mathbf{z}}\|$, rather than directly with **z**, extends the maximal invariance of the ANSIG to translation and scale transformations. It is also shown that the rotation of a shape "propagates" to its signature, being easily factored out. In [4], it is shown that, with the pre-processing step just referred, the ANSIG also deals well with shapes described by point sets of different cardinality, which is often the case in practice, when the shape points come from automatic processing, *e.g.*, edge detection, of images of different resolutions.

The properties above make the task of comparing arbitrary shapes very simple, requiring only the direct comparison of the corresponding ANSIGs. This was thoughtfully illustrated in [3, 4], with several shape-based image classification experiments. In these experiments, the analytic function (1) is described in the computer by its 512 samples uniformly taken on the unit-circle of the complex plane. Working with these high dimensional vectors may be adequate for tasks requiring the comparison of a small number of shapes but certainly not for applications that deal with large databases, *e.g.*, the internet. In the sequel, we show that a much more compact shape description can be used without performance decrease.

## 3. COMPLEX MOMENTS AND THE ANSIG SPECTRUM

A direct consequence of Cauchy's integral formula [10] is that any analytic function is fully specified by the values it takes on a closed contour on the complex plane. Thus, the ANSIG in (1) is equivalently described by its restriction to the unit-circle ($\xi = e^{j\theta}$),

$$h(\theta) = a(\mathbf{z}, e^{j\theta}) = \frac{1}{P} \sum_{i=1}^{P} \exp\left(z_i e^{j\theta}\right), \qquad (2)$$

where, for simplicity, we omitted the dependence of $h$ on **z**.

To approximate the (unit-circle restriction of the) ANSIG by a finite-dimensional computer representation, we study its frequency spectrum. Note that $h(\theta)$ in (2) can be seen as a real-argument complex-valued periodic function, with fundamental period $T = 2\pi$ and fundamental frequency $\omega_0 = 2\pi/T = 1$. Thus, it can be written in terms of its Fourier series:

$$h(\theta) = \sum_{k=-\infty}^{+\infty} H_k e^{jk\theta} \qquad (3)$$

where each coefficient $H_k$ is given by, see, *e.g.*, [11],

$$H_k = \frac{1}{2\pi} \int_0^{2\pi} h(\theta) e^{-jk\theta} \, d\theta. \qquad (4)$$

The analysis expression (4) is hard to carry out in the case of $h(\theta)$ given by (2). However, the coefficients of the Fourier series easily follow from the comparison of the synthesis expression (3) with the definition of $h(\theta)$ in (2), after some manipulations. In fact, expressing the exponential in (2) by its Maclaurin series (*i.e.*, its Taylor series at the origin), we get

$$h(\theta) = \frac{1}{P} \sum_{i=1}^{P} \sum_{k=0}^{\infty} \frac{(z_i e^{j\theta})^k}{k!} = \sum_{k=0}^{\infty} \left(\frac{1}{Pk!} \sum_{i=1}^{P} z_i^k\right) e^{jk\theta}. \quad (5)$$

Comparing (5) with (3) and since the coefficients of the Fourier series representation are unique [11], we conclude that:

$$H_k = \begin{cases} \frac{1}{Pk!} \sum_{i=1}^{P} z_i^k & \text{for } 0 \le k < +\infty, \\ 0 & \text{for } -\infty < k \le -1. \end{cases} \quad (6)$$

Thus, given a 2D shape described by a set of $P$ landmarks, expression (6) relates the values of the coefficients of the Fourier series of (the unit-circle restriction of) its ANSIG, $\{H_k, -\infty < k < +\infty\}$, to the positions of the landmarks in the plane, $\{z_i, 1 \le i \le P\}$.

Before proceeding, we relate the Fourier series coefficients in (6) with the so-called *complex moments* (CMs) introduced in [6]. The CM of order $(p, q)$ of an image $g(x, y)$ is defined by

$$C_{pq} = \int\int_{-\infty}^{+\infty} (x + jy)^p (x - jy)^q g(x, y) \, dx \, dy, \quad (7)$$

where $p \ge 0$ and $0 \le q \le p$ [6]. Considering an image composed by a set of $P$ mass points located at the shape landmarks $\{z_i, 1 \le i \le P\}$, the integral in (7) becomes a sum:

$$C_{pq} = \sum_{i=1}^{P} z_i^p \bar{z}_i^q, \qquad (8)$$

where, $\bar{z}_i$ denotes the complex conjugate of $z_i$. It is now clear that, for $k \ge 0$, the Fourier series coefficients in (6) are scaled versions of particular complex moments in (8): $H_k = C_{k0}/Pk!$. Finally, note also that, for any shape, we have $H_0 = \frac{1}{P} \sum_{i=1}^{P} 1 = 1$, and, assuming the shapes were pre-processed as described in the previous section, $H_1 = \frac{1}{P} \sum_{i=1}^{P} z_i = 0$.

We now derive an upper bound $b(k)$ for the magnitude of the coefficient $H_k$, for $k \ge 2$, in terms of the number of landmarks describing the shape. This is done through the following chain of equalities and inequalities:

$$|H_k| = \frac{1}{Pk!} \left| \sum_{i=1}^{P} z_i^k \right| \le \frac{1}{Pk!} \sum_{i=1}^{P} |z_i|^k \qquad (9)$$

$$= \frac{1}{Pk!} \sum_{i=1}^{P} \left(|z_i|^2\right)^{\frac{k}{2}} \le \frac{1}{Pk!} \left(\sum_{i=1}^{P} |z_i|^2\right)^{\frac{k}{2}} \quad (10)$$

$$= \frac{P^{\frac{k}{2}-1}}{k!} = b(k), \qquad (11)$$

where (9) uses the triangle inequality, (10) uses the fact that the function $(a + b)^n$ is convex for $a, b$ real positive and $n \ge 1$, and (11) is due to the fact that $\sum_{i=1}^{P} |z_i|^2 = P$, for pre-processed shapes.

## 4. PRINCIPAL MOMENTS FOR SHAPE DESCRIPTION

We now estimate the bandwidth of the ANSIG in terms of the number of landmarks. Usually, bandwidth is estimated by seeking the smallest $k$ such that the ratio $|H_k|/|H_0|$ is below a given threshold $p$. In our case, since $H_0 = 1$, we get $|H_k| < p$. Using the upper bound $b(k)$ as a proxy for $|H_k|$, we are not sure to obtain the smallest $k$, but we guarantee the satisfaction of the inequality, since $|H_k| \le b(k)$, see (9)-(11).

Finding the smallest $k$ such that $b(k) < p$ requires solving the limit case equation $P^{\frac{k}{2}-1}/k! = p$, for which there is not an analytic solution. We propose a simple method to solve for $k$ numerically.

Due to the fast grow of $k!$, we use logarithms to increase stability. Denoting the natural logarithm of $b(k)$ by $B(k)$, we have:

$$B(k) = \ln b(k) = \left(\frac{k}{2} - 1\right)\ln P - \ln k! \quad (12)$$

$$\simeq k\left(\frac{\ln P}{2} + 1\right) - k\ln k - \frac{1}{2}\ln k - \ln P - \frac{1}{2}\ln 2\pi, (13)$$

where (12) uses the definition of $b(k)$ in (11) and (13) uses the Stirling's approximation $\ln k! \simeq k\ln k - k + \frac{1}{2}\ln 2\pi k$, see, *e.g.*, [12].

To analyze the behavior of $B(k)$ given by (13), we relax $k$ to the reals and express the first two derivatives:

$$B'(k) = \frac{\ln P}{2} - \ln k - \frac{1}{2k}, \qquad B''(k) = -\frac{1}{k} + \frac{1}{2k^2}. \quad (14)$$

From these expressions, we see that $B(k)$ has an inflection at $k = 1/2$, where $B''(k) = 0$, and two extrema at $k = k_1 < 1/2$ and $k = k_2 > 1/2$, where $B'(k) = 0$. Also, $B(k)$ monotonically decreases for $k > k_2$, where $B'(k) < 0$, being $\lim_{k\to+\infty} B(k) = -\infty$. The plot in Fig. 1 illustrates the behavior of $B(k)$ for $P = 10$. To find $k$ such that $B(k)$ is below a given threshold $\ln p$, we thus propose the following strategy in two steps: first, solve $B'(k) = 0$ in the interval $k \in [1/2, +\infty)$, obtaining $k_2$. Then, solve $B(k) - \ln(p) = 0$ in the interval $k \in [k_2, +\infty)$, obtaining $k = k_B$, the desired upper bound for the bandwidth of the ANSIG. Note that the first step is necessary to specify the lower limit of the search region for the second one: without that limit, we could obtain a spurious solution $k_B < k_2$.
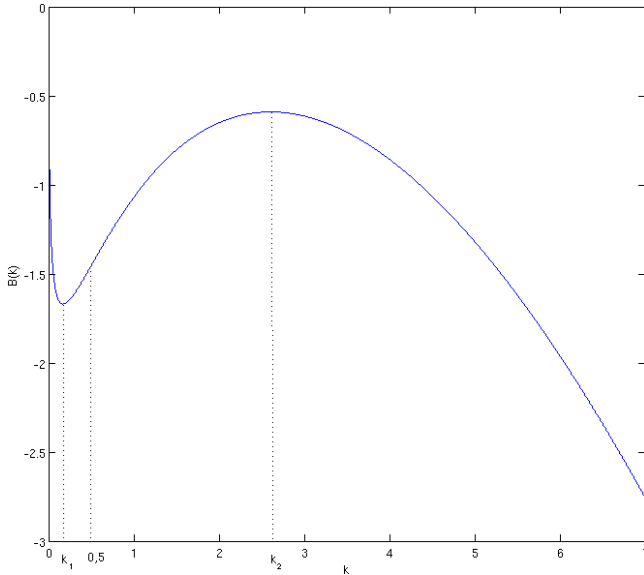


**Fig. 1**. Upper bound $B(k)$ for the (logarithm of the) magnitude of the spectrum of the analytic signature, for $P = 10$ points.

We thus conclude that the most of the energy (the parameter $p$ controls the amount) of the ANSIG of a shape described by $P$ landmarks is contained in a number $k_B$ of complex coefficients, which, naturally, depends on $P$. Fig. 2 plots the number of coefficients $k_B$, computed as described above, as a function of the number of landmarks $P$, for $p = 0.1$ ($-20$dB). Obviously, $k_B$ can be indistinctly interpreted as either the required number of Fourier series coefficients $\{H_k\}$, which we call the *Principal Moments* (PMs) of the shape (due to the relation between $H_k$ and the complex moments $C_{k0}$ pointed out in the previous section), or the required number of samples in the unit-circle. In fact, since the fundamental

frequency is $\omega_0 = 1$, the approximate bandwidth of the signal is $\omega_B = k_B\omega_0 = k_B$. Since the spectrum is zero for negative frequencies (6), it suffices to sample at a rate (number of points) of $N = \omega_s = \omega_B = k_B$ (the Nyquist sampling rate $\omega_s$ of twice the bandwidth is only required for two-sided spectra [11, 13]). Naturally, to recover the original continuous ANSIG from these samples, it suffices to use a (complex coefficient) filter with passband $\omega \in [0, \omega_s)$ (in opposition to the traditional low-pass filter with cutting frequency $\omega_s/2$). The plot in Fig. 2 also compares the required number $k_B$ of samples, or of PMs, with 512, the fixed number of samples used in [3, 4]: while for shapes described by an huge number of points (more than $\simeq 40000$), 512 samples may not be enough, for the majority of cases that may arise in practice (a few hundred of landmarks), the required number is much smaller (a few dozens).
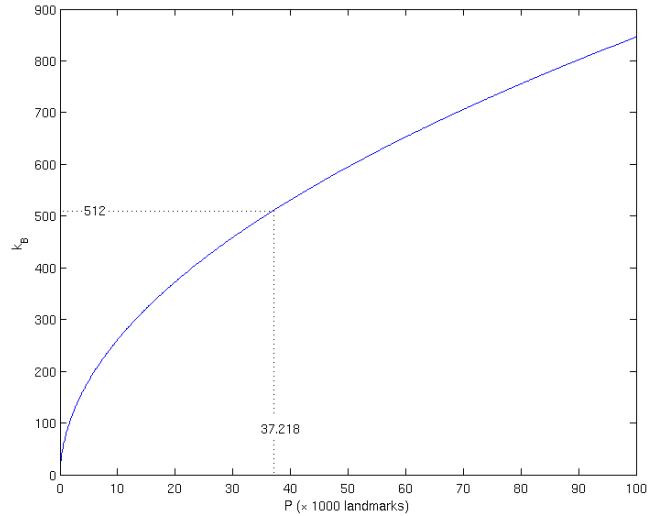


**Fig. 2**. The number of coefficients $k_B$ needed to represent a shape described by $P$ landmarks.

Having motivated the usage of only $k_B$ PMs as a way to compactly represent shape, in terms of its complexity, *i.e.*, its number of landmarks, we now discuss the comparison of descriptions of shapes of distinct complexity. Since the PMs are the (most relevant) coefficients of the ANSIG Fourier series, it suffices to pad with zeros the smaller vector, before performing the comparison in frequency domain. If, in opposition, the shapes are equivalently described by the sparse set of $k_B$ ANSIG samples, say $N_1$ samples for one of the shapes and $N_2$ samples for the other, it is necessary to use multirate signal processing techniques to convert both to a common sampling rate, see, *e.g.*, [13]. For example, perform upsampling of the ANSIGs by a factor of, respectively, $L_1 = \mathrm{lcm}(N_1, N_2)/N_1$ and $L_2 = \mathrm{lcm}(N_1, N_2)/N_2$, where lcm stands for the least common multiple, followed by interpolation using (complex coefficient) filters with passband, respectively, $\omega \in [0, 2\pi/L_1)$ and $\omega \in [0, 2\pi/L_2)$.

## 5. EXPERIMENTS

Since the ANSIG was extensively demonstrated in shape-based classification of real images [3, 4], we focus on showing that the computational saving that arises from using our PMs, a much more compact description, does not degrade performance. We start by illustrating this point by using a shape described by an arbitrary number (7 in this case) of unlabeled landmarks. Proceeding as described in the previous section, we obtain the required number of PMs for

this shape, $k_B = 6$. The plot of Fig. 3 shows the magnitude of the spectrum of the corresponding ANSIG. As easily perceived, the magnitude of the $6^{\text{th}}$ coefficient, $|H_5|$, is very small, indicating that the first 6 coefficients, *i.e.*, the PMs $\{H_k, 0 \leq k \leq 5\}$, containing the majority of the energy, adequatelly describe the shape. See also that $|H_0| = 1$ and $H_1 = 0$, as anticipated above. In Fig. 4, we compare the dense ANSIG with the one obtained by interpolating our very compact representation. The densely sampled ANSIG is represented by the solid lines, the $k_B = 6$ samples required by our study are dots, and the reconstruction obtained from this sparse set is represented by the dashed lines. We see that the lines of both plots are visually indistinguishable, showing that the PMs are adequate to represent the ANSIG and, consequently, the underlying shape.
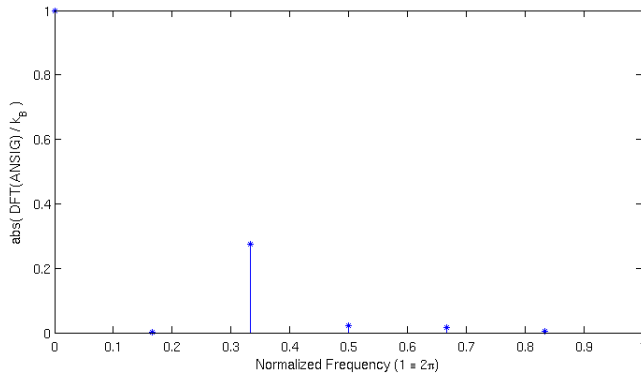


**Fig. 3**. Magnitude of the spectrum of the ANSIG of a shape described by a small number of landmarks.

The behavior just illustrated was observed in general. To confirm the results, we then compared PMs with the densely sampled ANSIG in terms of shape classification performance. In particular, we generated noisy versions of prototype shapes, like the ones used in [3, 4], and classified them by using 1-NN, *i.e.*, by just selecting the prototype that had most similar description. The number of PMs $k_B$ ranged from 15 to 22, thus our descriptions are much shorter than the vectors of 512 samples used in [3, 4]. We performed hundreds of tests for each shape, obtaining the same performance (100% correct classifications, except for shapes that are visually indistinguishable), for both the densely sampled ANSIG and the PMs.

## 6. CONCLUSION

We studied the frequency spectrum of the ANSIG for 2D shape representation. The major relevance of this study is that it enables the adoption of a compact description for 2D shape, in terms of what we call its *Principal Moments*, where the description length depends on the shape complexity. Our experiments show that this compact representation exhibits the same performance than the fixed rate densely sampling scheme originally proposed for the ANSIG.

## 7. REFERENCES

[1] I. Bartolini, P. Ciaccia, and M. Patella, "Warp: accurate retrieval of shapes using phase of fourier descriptors and time warping distance," *IEEE T-PAMI*, vol. 27, no. 1, 2005.

[2] D. Kendall, D. Barden, T. Carne, and H. Le, *Shape and Shape Theory*, John Wiley and Sons, 1999.

[3] J. Rodrigues, P. Aguiar, and J. Xavier, "ANSIG – an analytic signature for permutation-invariant two-dimensional shape representation," in *IEEE CVPR*, USA, 2008.
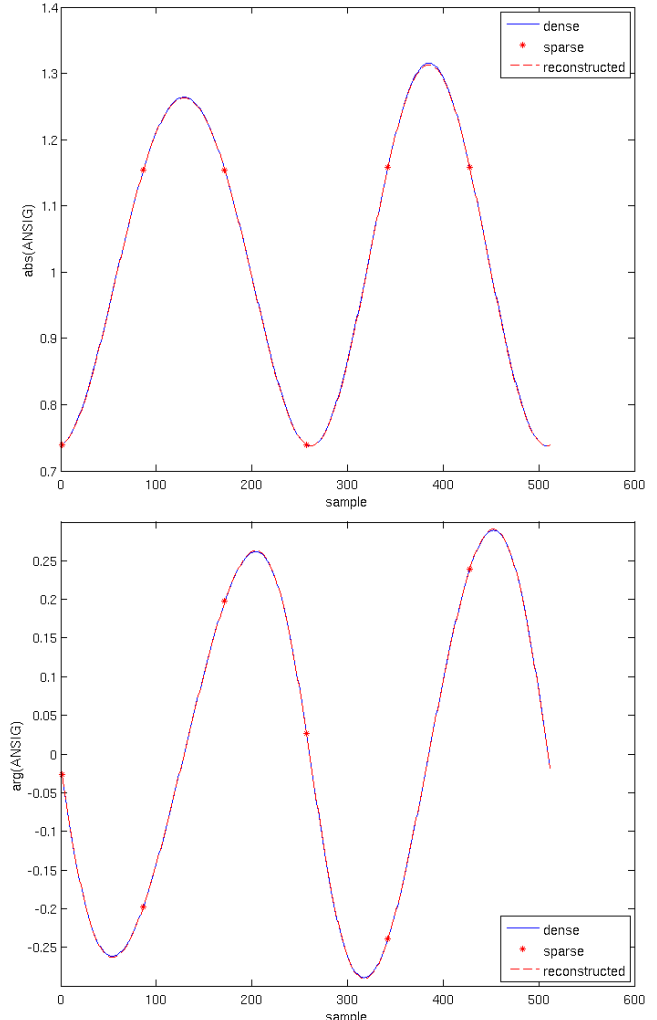
**Fig. 4**. Magnitude and phase of the ANSIG. Note how the very sparse representation we propose leads to plots that result visually indistinguishable from those obtained by dense sampling.

[4] J. Rodrigues, J. Xavier, and P. Aguiar, "Classification of unlabeled point sets using ANSIG," in *IEEE ICIP*, USA, 2008.

[5] M. Hu, "Visual pattern recognition by moment invariants," *IRE Trans. on Information Theory*, vol. 8, pp. 179–187, 1962.

[6] Y. Abu-Mostafa and D. Psaltis, "Image normalization by complex moments," *IEEE T-PAMI*, vol. 7, no. 1, 1985.

[7] A. Khotanzad and Y. Hong, "Invariant image recognition by Zernike moments," *IEEE T-PAMI*, vol. 12, 1990.

[8] D. Shen and H. Ip, "Symmetry detection by generalized complex moments," *IEEE T-PAMI*, vol. 21, no. 5, 1999.

[9] R. Mukundan, S. Ong, and P Lee, "Image analysis by Tchebichef moments," *IEEE T-IP*, vol. 10, 2001.

[10] L. Ahlfors, *Complex Analysis*, McGraw-Hill, USA, 1978.

[11] A. Oppenheim, A. Willsky, and S. Hamid, *Signals and Systems*, Prentice Hall, 1996.

[12] R. Paris and D. Kaminsky, *Asymptotics and the Mellin-Barnes Integrals*, Cambridge University Press, 2001.

[13] A. Oppenheim, R. Schafer, and J. Buck, *Discrete-Time Signal Processing*, Prentice Hall, 1999.