A New Workspace For Principal Axes And Scaling Estimation

Stelios Krinidis and Michail Krinidis
Information Management Department, Technological Institute of Kavala
Ag. Loukas, 65404 Kavala, GREECE
stelios.krinidis@mycosmos.gr,mkrinidi@gmail.com

Abstract

A novel algorithm for 2D object orientation and scaling factor estimation, is proposed in this paper. The proposed method is accurate, effective, computationally efficient and fully-automated. The object orientation is calculated by using object principal axes estimation. The main contribution of the proposed approach is the utilization of a 2D empirical mode like decomposition (EMD-like) as a new workspace for principal axes and scaling determination. The EMD algorithm can decompose any nonlinear and non-stationary data into a number of intrinsic mode functions (IMFs). When the object is decomposed by empirical mode like decomposition (EMD-like), the IMFs of the object, provide a workspace with very good properties for calculating the principal axes. The method was evaluated on synthetic and real images. The experimental results demonstrate the effectiveness and the accuracy of the method, both in orientation and scaling estimations.

keywords: *Principal axes estimation, orientation estimation, scaling estimation, empirical mode decomposition, ensemble empirical mode decomposition, intrinsic mode.*

1 Introduction

The last two decades, the wide use of web and the easy access to cameras lead to an explosion of online and offline usage of image collections. Whether it has to do with personal albums, posted images to the web or with large image databases, always digital images should be displayed in the correct orientation and scale. Unfortunately, this is not a simple task and is often performed manually. Furthermore, many image and video processing algorithms assume a priori knowledge of the image or of the depicted objects orientation and scale.

The object orientation and scaling factor estimation problem are central problems in several research areas, such as robotics, computer graphics, image and video processing, pattern recognition and computer vision. Depending on the kind of the problem, object orientation and scaling estimation can be used in different ways [19], i.e., in robotics are commonly used in hand-eye coordination systems [6], in computer graphics aim at combining computer-generated objects with photographic scenes and in computer vision are extensively used to many approaches as a preprocessing step [34].

In the computer vision literature, several approaches have been mentioned for estimating the orientation of a target object [29, 34]. Most of them work for arbitrary 3D target point

configurations [2], some have been extended to use points and lines [1] and some work also for coplanar points [19]. Recent success has also been reported for online structure and motion estimation [20], where many interest points are extracted; frame-to-frame correspondence is rather easy and no a priori reference to a scene coordinate system is required. In many research approaches, characteristic features are extracted from the image *visual content* which include e.g. shape, texture, or color properties defined in the imagery domain [5, 11, 21, 22, 23, 25, 27, 28].

Most of the systems (computer vision, pattern recognition, content based image retrieval, etc.) which exploit shape, tries to utilize information and features of the image that are invariant to some geometrical transformations such as rotation and scaling. The objects under examination and the target images may not have the same orientation and scaling, a fact that raises many difficulties in various research areas. On the other hand, the utilization of invariant image features tends to throw out some other, often useful features. Thereby, the determination of the principal axes of an object and the estimation of differences in orientation and scaling between the objects under consideration, seems to be a main problem in computer vision and pattern recognition applications and research.

In [16], the authors determine the orientation and scaling of an object by exploiting the frequency-based features, derived by the free vibrations of an initial circular chain code (physics-based modeling), which parameterizes the contour of the object under consideration. The object shape is also exploited in [10] in order to examine the role of local and global orientation in visual search. It was demonstrated that search for the global orientation of configurations of contours was very efficient, provided that the stimuli had an unambiguous global orientation. It was also found that it is more reliable, when a method is based on global shape orientation, rather than on the orientation of local contours. Consistent with this, the authors described the global orientation as axial, which suggests that the assignment of an orientation to a shape is accomplished by deriving the principal axes on the basis of contour information.

The axes of symmetry of the shape could be detected using any of the well known techniques reported in the literature. In [8] the edge points of the shape was exploited in order to find the orientation of an object shape. The internal edges were used in addition to the external boundary edges to increase the orientation detection capabilities of the introduced algorithm. First, the edge map of the image was extracted by applying Canny edge detector [4]. Second, the center of the object is detected by calculating the average of the vertical and horizontal coordinates of the points of the edge map. Third, the total perpendicular absolute distances from the edge map points to the line that passes through the center point with specified angle were calculated. These calculations were repeated with different angles to find the angles of the minimum peaks of the calculated distances. Finally, if the shape has more than one minimum peak an averaging method was adopted to get the dominant direction angle of the shape or the shape orientation.

There is a large number of methods determining the principal axes (object pose) in the literature [7, 14, 18]. The majority of them, which are constrained to exploit only contour information, are based on the well-known principal component analysis (PCA) method [14]. PCA is a standard decorrelation technique which aims at deriving an orthogonal projection basis that directly leads to dimensionality reduction, and possibly to feature selection. In the 2D case, the dimensionality reduction, i.e., the orthogonal axes that are derived, can be concerned as the principal axes of the shape (points) under examination. A great variety of PCA algorithms lies in the literature offering akin results, dimension reduction and in

2D case the principal axes of the object, such as Kernel PCA [33], independent component analysis (ICA) [17] and principal coordinates analysis (PCoA or PCO) [12, 15].

This paper deals with the afore mentioned problem. It presents a novel workspace that drives existing principal axes determination methods to achieve more robust and accurate results. The contour of the object under consideration is decomposed using an Empirical Mode Like Decomposition (EMD-like) algorithm based on the original EMD method [13, 31], which produces the contour IMFs. The object shape IMFs assign a new workspace with very good properties for calculating its principal axes and scaling factor. Moreover, the proposed method seems to be robust against to missing contour data (noise, cropping, etc), which are frequently appeared due to the segmentation methods that are utilized.

The remainder of the paper is organized as follows. The 2D *Empirical Mode Like Decomposition* (EMD-like) with its ensemble mode (EEMD-like), as well as the estimation of the principal axes and the scaling factor of the object under examination, are introduced in Section 2. Experimental results are presented in Section 3 and conclusions are drawn in Section 4.

2 The New Workspace and the Algorithm for Object Principal Axes and Scaling Estimation

In this Section, a new workspace for object orientation and scaling determination derived from the exploitation of the 2D empirical mode like decomposition (EMD-like) and its intrinsic mode functions (IMFs), will be introduced. More details regarding the original decomposition process, its properties and all the adopted assumptions are presented in [13, 31]. Furthermore, the algorithm for object principal axes and scaling estimation based on the introduced workspace will be presented in the end of this Section.

The basic idea embodied in the EMD analysis is the decomposition of any complicated data set into a finite and often small number of intrinsic mode functions, which have different frequencies, one superimposed on the other. The main characteristic of the EMD, in contrast to almost all previous decomposition approaches, is that EMD works directly in temporal space, rather than in the frequency space. The EMD method, as Huang *et al.* pointed out [13], is direct intuitive and adaptive with an *a-posteriori* defined basis based on and derived from the data and therefore, highly efficient. Since the decomposition of the input signal is based on the local characteristic time scale of the data, the EMD is applicable to nonlinear and non-stationary process.

The IMFs obtained by the decomposition method, constitutes an adaptive basis, which satisfies the majority of properties for a decomposition method, i.e., the convergence, completeness, orthogonality and uniqueness. Moreover, EMD algorithm copes with stationarity (or the lack of it) by ignoring the concept and embracing non-stationarity as a practical reality [13].

The possibly non-linear signal, which may exhibit varying amplitude and local frequency modulation, is linearly decomposed by EMD into a finite number of (zero mean) frequency and amplitude modulated signals. The remainder signal, called as a residual function, is a monotonic trend or is simply a constant.

2.1 The 1D Original Empirical Mode Decomposition (1D EMD)

Let us briefly review the original 1D empirical mode decomposition (EMD). In the EMD algorithm, the data $x(t)$ is decomposed in terms of IMFs c_i, as follows:

$$x(t) = \sum_{i=1}^{N} c_i + r_N, \tag{1}$$

where r_N is the residue of data $x(t)$, after N number of extracted IMFs. IMFs are simple oscillatory functions with varying amplitude and frequency, and hence have the following basic properties:

- Throughout the whole length of a single IMF, the number of extrema and the number of zero-crossings must either be equal or differ at most by one (although these numbers could differ significantly for the original data set).
- At any data location, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero.

In practice, the EMD is implemented through a "sifting process" that uses only local extrema. From any data r_{i-1}, the procedure is as follows:

1. Identify all the local extrema (the combination of both maxima and minima), connect all these local maxima (minima) with a cubic spline as the upper (lower) envelope, and calculate the local mean m_i of the two envelopes.
2. Obtain the first component $h = r_{i-1} - m_i$ by taking the difference between the data and the local mean of the two envelopes.
3. Treat h as the data and repeat steps 1 and 2 as many times as required until the envelopes are symmetric with respect to zero mean under certain criteria.

The final h is designated as c_i. The procedure can be repeatedly applied to all subsequent r_i, and the result is

$$\begin{array}{lcl} x(t) - c_1 & = & r_1 \\ r_1 - c_2 & = & r_2 \\ & \cdots & \\ r_{N-1} - c_N & = & r_N. \end{array} \tag{2}$$

The decomposition process finally stops when the residue, r_N, becomes a monotonic function or a function with only one extremum from which no more IMF can be extracted. By summing up equation (2), one can derive the basic decomposition equation (1). This means that a signal $x(t)$ is decomposed to N IMFs (c_i) and a residual r_N signal.

The very first step of the sifting process is depicted in Figure 1. Figure 1(a) depicts the original input data, while Figures 1(b) and 1(c) show the extrema (maxima and minima) of the data with their corresponding (upper and lower) envelopes. Figure 1(d) depicts the average of the two (upper and lower) envelopes, and Figure 1(e) illustrates the residue signal, that is the difference between the original data and the mean envelope. This procedure is repeated, as mentioned above, and all the IMFs are extracted from the original input signal. An example of the EMD algorithm and the extracted IMFs for the input data shown in Figure 1(a), is presented in Figure 2.

Based on this simple description of EMD, Flandrin *et al.* [9] and Wu and Huang [30] have shown that, when the data consists of white noise, the EMD behaves as a dyadic

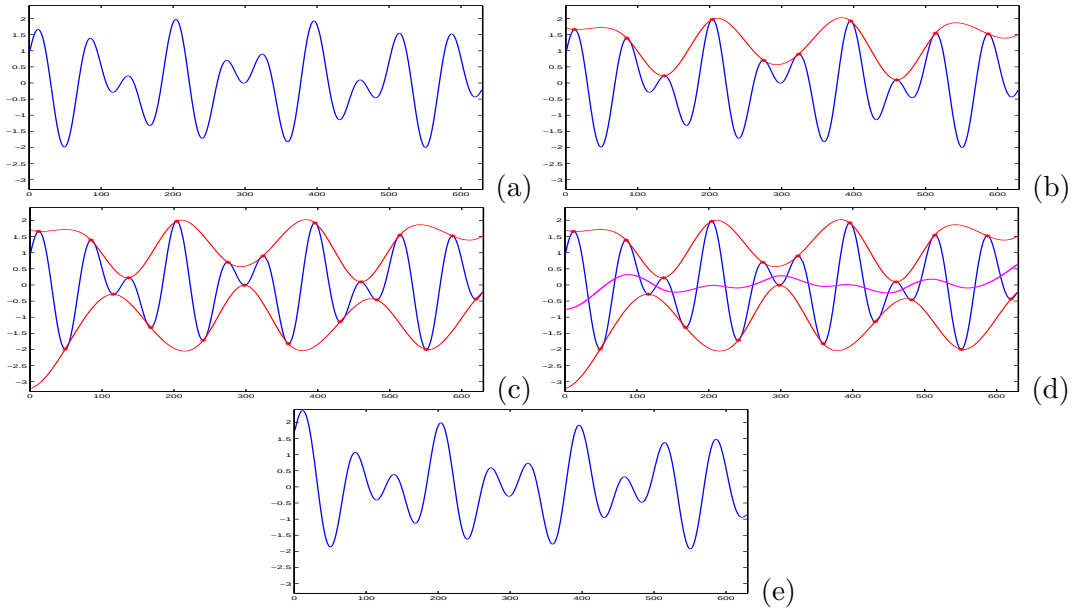

Figure 1. *The very first step of the sifting process.* ***(a)*** *is the input data,* ***(b)*** *identifies local maxima and plots the upper envelope,* ***(c)*** *identifies local minima and plots the lower envelope,* ***(d)*** *plots the the mean of the upper and lower envelope, and* ***(e)*** *the residue, the difference between the input data and the mean of the envelopes.*

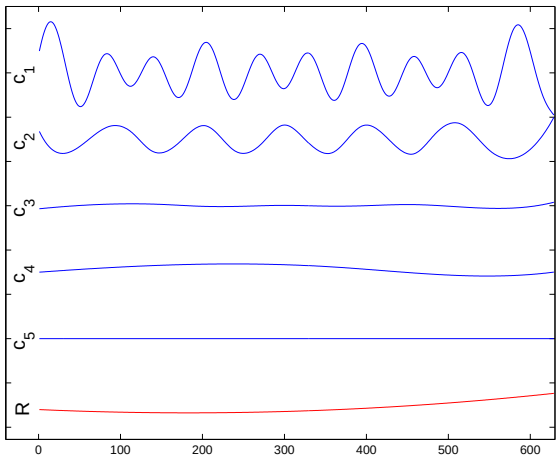

Figure 2. *The intrinsic mode functions (IMFs) of the input data displayed in Figure 1(a).*

filter bank: the Fourier spectra of various IMFs collapse to a single shape along the axis of logarithm of the period or the frequency. Then the total number (N+1) of IMFs of a data set is close to $log_2 N'$, with N' being the number of total data points. On the other hand, when the data is not pure noise, some scales could be missing, and as a consequence, the total number of the IMFs might be fewer than $log_2 N'$. Additionally, the intermittency of signals in certain scale would also cause mode mixing.

One of the major drawbacks of EMD is mode mixing. Mode mixing is defined as a single IMF either consisting of signals with widely disparate scales or consisting of a signal with

a similar scale residing in different IMF components. Mode mixing is a consequence of signal intermittency. The intermittency could not only cause serious aliasing in the time-frequency distribution but could also make the individual IMF lose its physical meaning [13]. Another side effect of mode mixing is the lack of physical uniqueness. Supposing that two observations of the same oscillation are made simultaneously, one contains a low level of random noise and the other does not. The EMD decompositions for the corresponding two records are significantly different [32].

However, since the cause of the problem is due to mode mixing, one expects that the decomposition would be reliable if the mode mixing problem is alleviated or eliminated. To achieve the latter goal, i.e., to overcome the scale mixing problem, a new noise-assisted data analysis method was proposed, named as the ensemble EMD (EEMD) [32]. The EEMD defines the true IMF components as the mean of an ensemble of trials, each one consisting of the signal with white noise of finite amplitude.

The ensemble EMD (EEMD) algorithm could be summarized as follows:

1. add a white noise series $w(t)$ to the original input data $x_i(t) = x(t) + w_i(t)$,
2. decompose the data with added white noise into a number of IMFs $c_{jk}(t)$,
3. repeat steps 1 and 2 but with different white noise series each time, and
4. obtain the (ensemble) means of corresponding IMFs $c_j(t) = \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} c_{jk}(t)$ of the decomposition as the final result.

The critical concepts advanced in EEMD are based on the following observations:

- A collection of white noise cancels each other out in a time-space ensemble mean. Therefore, only the true components of the input data can survive and persist in the final ensemble mean.
- Finite, not infinitesimal, amplitude white noise is necessary in order to force the ensemble to exhaust all possible solutions.
- The physically meaningful result of the EMD is not derived from the data without noise, but it is designated to be the ensemble mean of a large number of EMD trials of the input data with the added noise.

The mode mixing is largely eliminated using EEMD, and the consistency of the decompositions of slightly different pairs of data is greatly improved. Indeed, EEMD represents a major improvement over the original EMD. Furthermore, since the level of the added noise is not of critical importance and of finite amplitude, EEMD can be used without any significant intervention. Thus, it provides a truly adaptive data analysis method. The EMD, with the ensemble approach (EEMD), has become a more mature tool for nonlinear and non-stationary time series (and another one dimensional data) analysis.

2.2 The New Workspace - The 2D Empirical Mode Like Decomposition (2D EMD-like)

In the proposed 2D EMD-like algorithm, the input data is a number of marked points on the object contour, which do not represent interior object regions. The coordinates of the object under examination are stacked in the vector:

$$\mathbf{v} = [\mathbf{v_1}, \mathbf{v_2}, \ldots, \mathbf{v_M}], \tag{3}$$

where $\mathbf{v_i} = [x_i, y_i]^T$ denotes the i-th object contour point, and M is the number of contour points. In 2D EMD-like algorithm, the data $\mathbf{v}$ is decomposed in terms of IMFs $\mathbf{c_i}$, as follows:

$$\mathbf{v} = \sum_{i=1}^{N} \mathbf{c_i} + \mathbf{r_N}, \tag{4}$$

where $\mathbf{r_N}$ is the residue of data $\mathbf{v}$, after N number of IMFs are extracted. In practice, the 2D EMD-like is implemented through a "sifting process" that uses only local extrema. For any data r_{i-1}, the procedure is as follows:

1. Identify all the local extrema (the combination of both maxima and minima), connect all these local maxima (minima) with lines as the upper (lower) envelope, and calculate the local mean m_i of the two envelopes.
2. Obtain the first component $h = r_{i-1} - m_i$ by taking the difference between the data and the local mean of the two envelopes.
3. Treat h as the data and repeat steps 1 and 2 as many times as required.

The final h is designated as $\mathbf{c_i}$. The procedure can be repeatedly applied to all subsequent $\mathbf{r_i}$, and the result is:

$$\begin{aligned} \mathbf{v} - \mathbf{c_1} &= \mathbf{r_1} \\ \mathbf{r_1} - \mathbf{c_2} &= \mathbf{r_2} \\ &\cdots \\ \mathbf{r_{N-1}} - \mathbf{c_N} &= \mathbf{r_N}. \end{aligned} \tag{5}$$

The decomposition process finally stops when the residue, r_N, becomes a monotonic-like function. As experimentally has been proven, this happens, when the total number of extracted IMFs is close to $log_2 M$, with M being the number of total object contour points. By summing up equation (5), one can derive the basic decomposition equation (4). This means that a signal $\mathbf{v}$ is decomposed to N IMFs ($\mathbf{c_i}$) and a residual $\mathbf{r_N}$ signal.

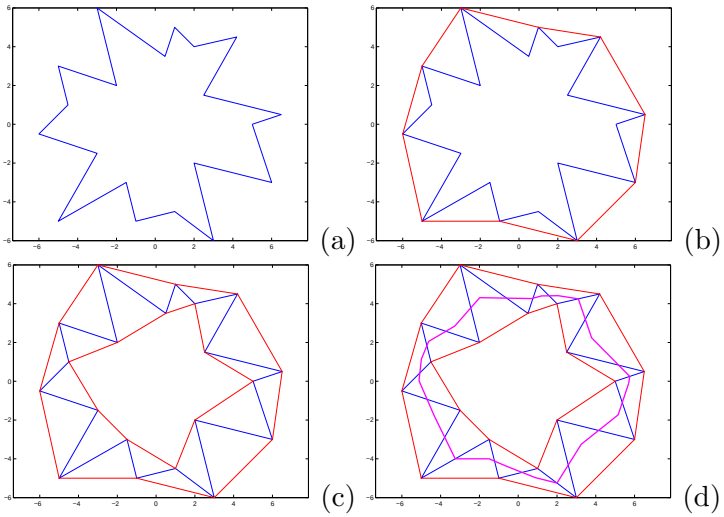

Figure 3. *The very first step of the sifting process.* ***(a)*** *is the input contour,* ***(b)*** *identifies local maxima and plots the upper envelope,* ***(c)*** *identifies local minima and plots the lower envelope,* ***(d)*** *plots of the average between the upper and lower envelope.*

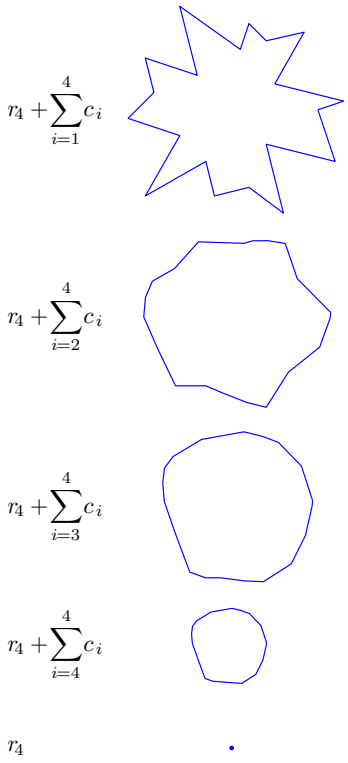

Figure 4. *The intrinsic mode functions (IMFs) of the input contour displayed in Figure 3(a).*

The very first step of the sifting process is depicted in Figure 3. Figure 3(a) depicts the original input data, while Figures 3(b) and 3(c) show the extrema (maxima and minima) of the data with their corresponding (upper and lower) envelopes. Figure 3(d) depicts the average of the two (upper and lower) envelopes. This procedure is repeated, as mentioned above, and all the IMFs are extracted from the original input object contour. An example of the 2D EMD-like algorithm and the extracted IMFs for the input data shown in Figure 3(a), is presented in Figure 4.

Similar to the original 1D EMD, the total number $(N+1)$ of IMFs of the data set in the 2D EMD-like algorithm, is very close to $log_2 M$, with M being the number of total object contour points. In practice the number M of object contour points is calculated after the removal of all the collinear points.

Furthermore, the EMD drawback of mode mixing also existes in the previously described 2D EMD-like case. Thus, we define an *Ensemble EMD-like* (EEMD-like) algorithm, which defines the true IMF components as the mean (ensemble) of an ensemble of trials, each one consisting of the signal with white noise of finite amplitude.

The ensemble EMD-like (EEMD-like) algorithm could be summarized as follows:

1. add a white noise series $\mathbf{w}$ to the original input data $\mathbf{v_i} = \mathbf{v} + \mathbf{w_i}$,
2. decompose the data with added white noise into a number of IMFs $\mathbf{c_{jk}}$,
3. repeat steps 1 and 2 but with different white noise series each time, and

4. obtain the (ensemble) means of corresponding IMFs $\mathbf{c_j} = \lim_{L\to\infty} \frac{1}{L} \sum_{k=1}^{L} \mathbf{c_{jk}}$ of the decomposition procedure as the final result.

2.3 Object Principal Axes and Scaling Estimation

The algorithm for object principal axes and scaling estimation based on the introduced workspace (2D EEMD-like), will be consequently introduced.

One can easily see from the aforementioned examples that the first and foremost IMFs $(c_1, \ldots, c_{N_A-1})$ mainly carries the object contour "noise", missing the data or outliers (contour noise, cropping, etc.), while the latest IMFs $(c_{N_B+1}, \ldots, c_N)$ and the residue r_N mostly describe the trend of the object contour. On the other hand, intermediate IMFs $(c_{N_A}, \ldots, c_{N_B})$ describe the initial object contour with simple and uniform curves. This is the main reason that the proposed method is focused on c_{N_A} to c_{N_B} IMFs, where N_A and N_B define the range of IMFs under consideration. Let us define the summation c_{N_A,N_B} of these IMFs as follows:

$$c_{N_A,N_B} = \sum_{i=N_A}^{N_B} c_i. \tag{6}$$

Once the new workspace has been defined for the principal axes and scaling estimation, and the outlier IMFs have been removed, the proposed method applied a well known algorithm [14, 16] to define the principal axes and the scaling factor of each produced summation $c_{i,N}$:

$$\{\theta_i, s_i\} = \text{estim}\left\{c_{i,N} + r_N\right\}, \tag{7}$$

where θ_i and s_i are the estimated principal axis and the scaling factor respectively, derived from the *estim* algorithm, and $i = N_A, \ldots, N$. The final principal axis θ and the scaling factor s is extracted as follows:

$$\begin{aligned} \theta &= \frac{1}{N - N_A + 1} \sum_{k=N_A}^{N} \theta_k, \\ s &= \frac{1}{N - N_A + 1} \sum_{k=N_A}^{N} s_k. \end{aligned} \tag{8}$$

Figure 5 illustrates an example of estimating the principal axes of an image (Figure 5i) and the same image rotated by 35^o and scaled with factor 1.5 (Figure 5i). The axes that are depicted in both images are the principal axes, as estimated by the proposed workspace exploiting the method proposed in [16]. In Figure 5a, the summations $c_{1,8}$ of the derived IMFs of images 5i and their extracted principal axes are respectively depicted, while Figures 5b show the summations $c_{2,8}$ of the derived IMfs of images 5i and their principal axes respectively. In the same way, Figures 5c - 5h depict the summations $c_{i,8}$ of the derived IMfs of images 5i and their extracted principal axes respectively. One can easily notice that the estimated principal axes in the produced summations (Figures 5a - 5h) have small deviations. This fact occurs because the first and foremost IMFs summations carry a lot of contour outliers (contour noise, cropping, etc), while the latest form the main object shape without important details. Thus, the intermediate IMFs summations, $c_{3,8} - c_{5,8}$ in our example, carry the main shape information free of outliers. Based on those IMFs

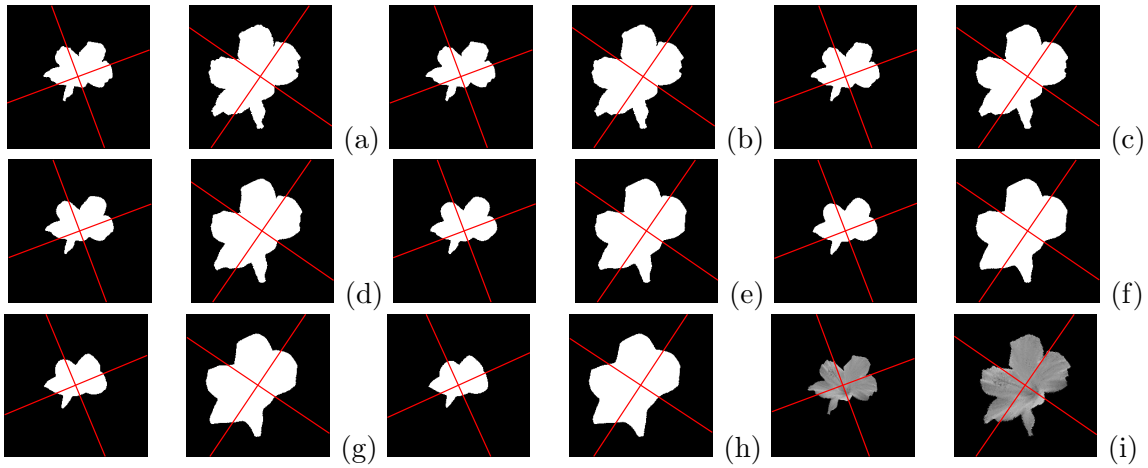

Figure 5. ***(a) - (h)*** *The summations $c_{1,8}$ of the IMFs of the images under consideration and their corresponding principal axes as they produced by the proposed algorithm.* ***(i)*** *The final image set illustrate the images under consideration (original image and its copy scaled with factor 1.5 and rotated 35^o) and their final principal axes.*

summations, the proposed algorithm extracts the principal axes exploiting equations (7) and (8) of the objects under consideration (Figure 5i).

The proposed object principal axes and scaling estimation algorithm could be summarized as follows:

1. compute the 2D EEMD-like algorithm on the object contour under consideration,
2. remove the IMFs of the outliers (first and foremost),
3. estimate the principal axes and scaling factor of each remaining IMF separately by exploiting algorithm [14, 16],
4. obtain the (ensemble) means of corresponding estimations as the final result.

3 Experimental Results

To evaluate the proposed workspace, the presented algorithm was applied to a number of natural images (over 500 images) collected from the Internet and also applied to the shape database used by Sebastian *et al.* [24] (e.g. Figures 6 and 7). They were manually, artificially scaled and rotated, i.e. they were randomly transformed using scalings s varying from 0.25 to 4.0 and rotations θ varying from -90^o to +90^o, in order to measure and confirm the principal axes and scaling factor estimation accuracy and efficiency of the proposed algorithm. In other words, the accuracy of the derived transformation parameters (θ and s) by the principal axes estimation have been measured. Moreover, the proposed algorithm has been used in combination of the well-known *principal component analysis* (PCA) [14], as well as of a physics-based method [16]. Furthermore, the results of the proposed algorithm are compared with results obtained by the pure aforementioned methods (without the use of the introduced workspace) applied to the same transformations with the proposed method. Since, the principal axes of an object is, in most cases, an objective subject, the proposed method has been examined in estimating the principal axes of an object and the

transformed one, and the accuracy of transformation retrieval among the objects has been computed. Also, the proposed algorithm is evaluated by an objective criterion, the *Mean Alignment Error (MAE)* [7]:

$$MAE(A,B) = \frac{1}{M}\sum_{i=1}^{M}\sqrt{\left(x_i^A - x_i^B\right)^2 + \left(y_i^A - y_i^B\right)^2}, \tag{9}$$

where $A = \left\{\left(x_i^A, y_i^A\right)\right\}_{i=1..M}$ and $B = \left\{\left(x_i^B, y_i^B\right)\right\}_{i=1..M}$ are the *M-point* shape instances. The shape instance A is said to be aligned to instance B if the $MEA(A,B)$ is minimum. Furthermore, a segmentation procedure has been applied to the input images, since the proper function of the proposed algorithm demands the existence of an obvious foreground and background in the testing images. The segmentation was performed either manually or automatically [3, 26].

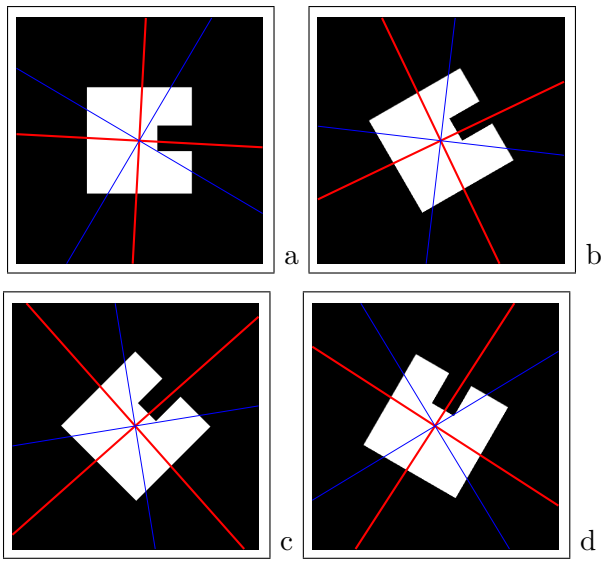

Figure 6. *Orientation and scaling recovery of an artificial image.* ***(a)*** *The initial image. The initial image after* ***(b)*** $+30^o$ *rotation,* ***(c)*** $+45^o$ *rotation, and* ***(d)*** $+60^o$ *rotation. The thickly line represents the principal axes estimated by the proposed method exploiting a physics-based method [16], while the hairline depicts the principal axes computed by the PCA algorithm.*

As aforementioned, the proposed method was applied to the orientation and scaling recovery of an artificially transformed image (Figure 6). Thus, the original image [Figure 6a] was transformed using rotations θ varying from -90^o to $+90^o$ with no scaling. Transformations in each case follow a uniform distribution, which means that all the integer rotation angles between -90^o and $+90^o$ were tested in order not to privilege any angle [Figs. 6b - 6d]. Table 1 presents statistics on the rotation and scaling recovery errors $\Delta\theta$ and Δs. The errors $\Delta\theta$ and Δs were computed between the acquired and the initial images. The median, average, maximum and standard deviation values were calculated for all the errors $\Delta\theta$ and Δs computed for each rotation angle between -90^o and $+90^o$. As it can be seen, median and mean scaling and rotation errors are much less than 0.001 and 0.10^o respectively, while the corresponding results of pure PCA and physics-based algorithms are

Table 1. An artificial image (Figure 6) was artificially rotated using various rotation parameters. The image was uniformly rotated using angles varying from -90^o to +90^o with no scaling. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.09^o	1.03^o	0.07^o	0.13^o
	Δs	0.0001	0.0009	0.0001	0.0001
	MAE	0.21	0.35	0.20	0.06
Proposed Workspace exploiting PCA	$\Delta\theta$	0.72^o	5.75^o	1.36^o	1.89^o
	Δs	0.0001	0.0018	0.0002	0.0002
	MAE	1.60	3.65	0.91	0.84
Physics-based method [16]	$\Delta\theta$	0.49^o	1.39^o	0.52^o	0.37^o
	Δs	0.0049	0.0068	0.0058	0.0011
	MAE	0.65	1.54	0.70	0.32
PCA	$\Delta\theta$	5.68^o	8.81^o	5.69^o	4.80^o
	Δs	0.0048	0.0072	0.0051	0.0010
	MAE	5.24	10.14	4.59	2.20

a little worst. Also maximum errors are less than 0.001 and 1.50^o respectively, showing the robustness of the proposed workspace, a fact that is also enhanced by the MAE, which illustrates the average pixel restoration error. Hence, exploiting the proposed workspace for the principal axes determination of an object, one can achieve better results than using a pure method. In Figure 6 is also depicted the principal axes of the object estimated by the proposed technique exploiting a physics-based method [16] (thickly line) and the pure PCA (hairline).

Furthermore, the proposed method was tested using the same image (object), but in this set of experiments, the rotation angle was set equal to 0^o and the scaling factor varied between 0.25 to 4. Table 2 presents statistics of the rotation and scaling recovery errors, similar to Table 1. As it can be seen, median and mean scaling and rotation errors are less than 0.001 and 0.10^o respectively. Also maximum errors are less than 0.01 and 0.30^o respectively, showing the effectiveness of the proposed technique, which is also established by MAE results. Both the afore mentioned experiments show that no particular orientation or scaling factor has been privileged by the proposed workspace. On the other hand, pure PCA and physics-based algorithms in both of the experiments achieved worst performance than when the proposed workspace was utilized.

Moreover, the proposed workspace was applied on another set of experiments with an artificially transformed image (Figure 6), where the rotation parameter took uniformly random angles varying from -90^o to $+90^o$ and the scaling parameter also took uniformly random values varying from 0.25 to 4. Table 3 presents statistics of the rotation and scaling recovery errors from this set of experiments. As it can be seen, median and mean scaling and rotation errors are much less than 0.01 and 0.50^o respectively. Also maximum errors are less than 0.01 and 1.50^o respectively, proving that by exploiting the proposed workspace, an algorithm can recover very accurately the initial configuration (orientation and scaling)

Table 2. An artificial image (Figure 6) was artificially scaled using various scaling parameters. The image was uniformly scaled using values varying from 0.25 to 4.0 with no rotation. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.03^o	0.30^o	0.08^o	0.09^o
	Δs	0.0007	0.0018	0.0008	0.0007
	MAE	0.27	0.50	0.29	0.10
Proposed Workspace exploiting PCA	$\Delta\theta$	0.64^o	6.63^o	1.88^o	2.23^o
	Δs	0.0016	0.0020	0.0010	0.0009
	MAE	3.40	6.05	3.63	1.16
Physics-based method [16]	$\Delta\theta$	0.22^o	0.52^o	0.21^o	0.11^o
	Δs	0.0088	0.0210	0.0106	0.0058
	MAE	0.33	0.61	0.34	0.13
PCA	$\Delta\theta$	5.56^o	8.93^o	6.76^o	5.96^o
	Δs	0.0022	0.0107	0.0132	0.0067
	MAE	4.05	10.02	5.24	3.76

Table 3. An artificial image (Figure 6) was artificially rotated and scaled using various rotation and scaling parameters. The image was uniformly randomly rotated and scaled using angles varying from -90^o to $+90^o$ and scaling varying from 0.25 to 4. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.13^o	1.02^o	0.18^o	0.19^o
	Δs	0.0012	0.0027	0.0012	0.0007
	MAE	0.43	1.23	0.57	0.20
Proposed Workspace exploiting PCA	$\Delta\theta$	0.52^o	18.67^o	4.95^o	6.12^o
	Δs	0.0016	0.0033	0.0014	0.0008
	MAE	1.14	15.23	3.98	3.66
Physics-based method [16]	$\Delta\theta$	0.18^o	1.19^o	0.23^o	0.19^o
	Δs	0.0166	0.0201	0.0180	0.0058
	MAE	0.57	1.83	0.66	0.21
PCA	$\Delta\theta$	0.69^o	20.67^o	5.25^o	6.41^o
	Δs	0.0156	0.0287	0.0171	0.0061
	MAE	1.46	20.26	4.54	3.69

of an object.

Moreover, the proposed workspace was applied on a variety of other artificial and natural images, a sample of which is shown in Figure 7, using akin transformations as in the previous experiments. Table 4 presents statistics of the rotation and scaling recovery errors from

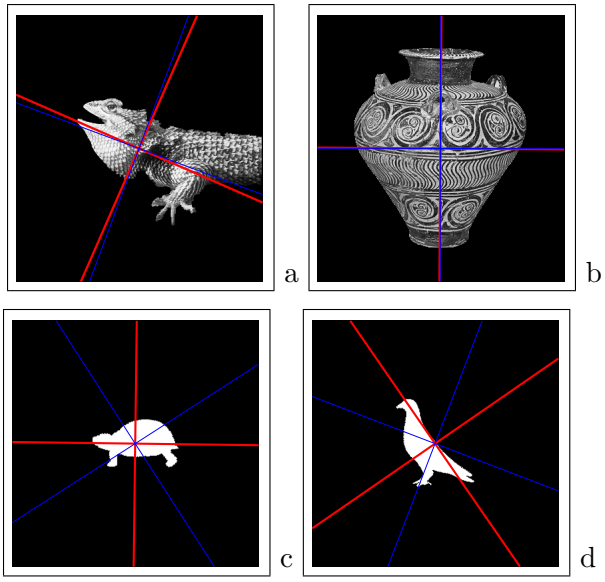

Figure 7. *Examples of artificial and natural images applied on the proposed algorithm exploiting physics-based method [16].*

Table 4. Synthetic and natural images were artificially rotated and scaled using various rotation and scaling parameters. The images were uniformly randomly rotated and scaled using angles varying from -90^o to $+90^o$ and scaling varying from 0.25 to 4. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.13^o	7.99^o	0.33^o	1.25^o
	Δs	0.0007	0.0377	0.0112	0.0558
	MAE	0.49	1.81	0.59	0.32
Proposed Workspace exploiting PCA	$\Delta\theta$	0.27^o	20.90^o	5.15^o	7.19^o
	Δs	0.0008	0.0455	0.0117	0.0581
	MAE	0.98	9.62	2.18	1.67
Physics-based method [16]	$\Delta\theta$	0.39^o	8.48^o	0.55^o	0.47^o
	Δs	0.0232	0.0834	0.0248	0.0153
	MAE	0.82	6.53	0.83	0.23
PCA	$\Delta\theta$	1.09^o	48.24^o	20.38^o	28.97^o
	Δs	0.0306	0.1331	0.0407	0.0282
	MAE	1.08	20.26	5.61	5.79

those images similar to Table 1. As it can be seen, median and mean scaling and rotation errors are about 0.01 and 0.35^o respectively. Also maximum errors are about 0.05 and 8.00^o respectively. The errors from the results of pure PCA and physics-based algorithms are worse than exploiting the proposed workspace in all kind of experiments indicating its efficiency and its reliableness.

Also, Figures 6 and 7 illustrate the principal axes of the objects under consideration

estimated by the proposed method using the physics-based algorithm [16] (thickly line) and the pure PCA algorithm (hairline). In some cases, the estimated principal axes of the objects under consideration, acquired by algorithms which once used the proposed workspace and then the pure version of the algorithm, were almost the same (Figure 7a and 7b). However, in most of the cases, the axes differ a lot (Figure 6, 7c and 7d), and the principal axes computed by algorithms which exploit the introduced workspace seem to be more reasonable (visually) than the ones derived by the original versions of the algorithms (PCA, etc).

Table 5. Synthetic and natural images were artificially rotated and scaled using various rotation and scaling parameters. The images were uniformly randomly rotated and scaled using angles varying from -90^o to $+90^o$ and scaling varying from 0.25 to 4. Gaussian noise (0,1) was also added to the objects contour. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.09^o	6.30^o	0.22^o	0.46^o
	Δs	0.0010	0.0250	0.0014	0.0061
	MAE	0.50	1.44	0.55	0.27
Proposed Workspace exploiting PCA	$\Delta\theta$	2.14^o	13.04^o	5.20^o	5.19^o
	Δs	0.0010	0.0250	0.0016	0.0062
	MAE	3.00	8.63	3.37	1.67
Physics-based method [16]	$\Delta\theta$	0.36^o	2.82^o	0.46^o	0.38^o
	Δs	0.0233	0.0678	0.0244	0.0129
	MAE	0.82	6.53	0.83	0.23
PCA	$\Delta\theta$	9.68^o	27.17^o	10.25^o	22.41^o
	Δs	0.0567	0.1261	0.0574	0.0305
	MAE	13.08	24.52	12.21	5.67

A frequently encountered task in many applications is that the adopted image preprocessing steps (rotations, cropping, interpolation methods, segmentation, etc.) insert a "noise"-like variation in the image and as a consequence on the contour of the image. Those contour missing data and outliers make difficult and inaccurate the computation of the principal axes and the scaling factor of the objects under examination. In the next set of experiments, those outliers and missing data were simulated by adding gaussian noise (0,1) to the object contour (Figures 8a and 8b) or by cropping it (10% of the contour points) as shown in Figures 8c and 8d. The error statistics of those experiments are presented in Tables 5 and 6 respectively. It is noticeable that using the proposed workspace, in both cases, the median and mean scaling and rotation errors are much less than 0.01 and 0.30^o respectively. Also, maximum errors are about 0.03 and 6.50^o respectively. On the other hand, pure PCA and physics-based algorithms cannot provide better accuracy than the proposed workspace, which is proved to be robust to contour loss data and outliers.

Finally, in order not to privilege the proposed workspace against pure PCA, a new set of experiments was performed, where the object contour was pre-smoothed (only in PCA algorithm) by a median filter of size varying from 3 to 9. The error statistics of the above

Table 6. Synthetic and natural images were artificially rotated and scaled using various rotation and scaling parameters. The images were uniformly randomly rotated and scaled using angles varying from -90^o to $+90^o$ and scaling varying from 0.25 to 4. The contour of the object under examination was cropped by 10%. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting [16]	$\Delta\theta$	0.05^o	7.11^o	0.19^o	0.47^o
	Δs	0.0009	0.0297	0.0020	0.0010
	MAE	0.51	2.44	0.56	0.26
Proposed Workspace exploiting PCA	$\Delta\theta$	0.09^o	38.28^o	5.42^o	12.07^o
	Δs	0.0009	0.0353	0.0021	0.0058
	MAE	3.05	8.65	3.41	1.66
Physics-based method [16]	$\Delta\theta$	0.08^o	17.21^o	3.69^o	2.15^o
	Δs	0.0012	0.0048	0.0062	0.0059
	MAE	0.27	5.27	2.45	1.46
PCA	$\Delta\theta$	1.18^o	48.98^o	7.32^o	5.28^o
	Δs	0.0406	0.0899	0.0412	0.0216
	MAE	2.48	9.73	4.12	4.57

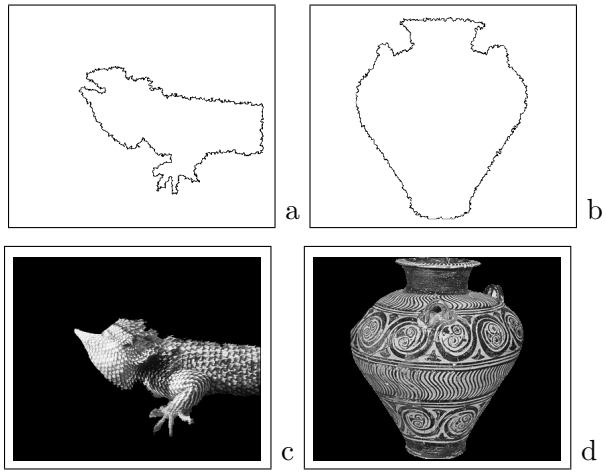

Figure 8. *Examples of images (**a**) and (**b**) with gaussian noise (0,1) applied on the object contour, and images (**c**) and (**d**) with* 10% *of their contour points cropped.*

mentioned experiment are presented in Table 7. It is remarkable, that whereas the presmoothing of the contour enforces pure PCA to provide akin results to those utilizing the proposed workspace, the improvement was not large enough to reach the accurate results of the proposed workspace, which was proved more efficient and reliable than pure PCA method.

Table 7. Synthetic and natural images were artificially rotated and scaled using various rotation and scaling parameters. The images were uniformly randomly rotated and scaled using angles varying from -90^o to $+90^o$ and scaling varying from 0.25 to 4. The contour of the object under examination (only in PCA algorithm) was pre-smoothed by a median filter of size varying from 3 to 9. Different statistics on the derived errors for the rigid transformation parameters are presented. Rotation errors are expressed in degrees and MAE in pixels.

		median	maximum	mean	s. dev
Proposed Workspace exploiting PCA	$\Delta\theta$	0.06^o	3.40^o	0.07^o	0.20^o
	Δs	0.0007	0.0059	0.0011	0.0010
	MAE	0.50	1.45	0.56	0.27
PCA	$\Delta\theta$	0.93^o	46.39^o	18.11^o	23.67^o
	Δs	0.0565	0.1254	0.0564	0.0306
	MAE	0.98	10.53	4.86	4.71

4 Conclusion

A novel and robust workspace for the estimation of 2D objects principal axes (orientation) and scaling factor, was presented. The contour of the object under consideration was decomposed by an EEMD-like algorithm, which produces the IMFs of the object. The obtained IMFs provide a very good workspace for well-known algorithms that determine the principal axes and scaling factor. The proposed workspace was experimentally proven to produce very low orientation and scaling errors comparing to pure methods for principal axes and scaling estimation (PCA, etc). The estimated principal axes and scaling factor were accurate and reasonable.

Furthermore, no particular orientation and scaling value was privileged by the proposed technique. Also, the usage of the intermediate IMFs of the object contours makes the technique robust to missing data or outliers (contour noise, cropping, etc.) on the object contour.

The good quality of the object orientation and scaling determination, as well as, its outlier resistance makes the workspace a promising tool for the computer vision and pattern recognition systems.

References

[1] A. Ansar and K. Daniilidis. Linear pose estimation from points or lines. In *Proceedings of European Conference on Computer Vision*, volume 4, pages 282 – 296, May 2002.

[2] H. Araujo, R. Carceroni, and C. Brown. A fully projective formulation to improve the accuracy of lowes pose-estimation algorithm. *Vision and Image Understanding*, 71(2):227–238, 1998.

[3] A. G. Bors and I. Pitas. Optical flow estimation and moving object segmentation based on median radial basis function network. *IEEE Transactions on Image Processing*, 7(5):693–702, May 1998.

[4] J. Canny. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6):679–698, 1986.

[5] V. Castelli and L. D. Bergman. *Image Databases: Search and Retrieval of Digital Imagery*. John Wiley & Sons, Inc, New York, 2002.

[6] G. N. DeSouza and A. C. Kak. Vision for mobile robot navigation: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):237–267, Feb. 2002.

[7] N. Duta, A. Jain, and M. Dubuisson-Jolly. Automatic construction of 2D shape models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(5):433–446, May 2001.

[8] E. El-Sayed, R.F. Abdel-Kader, and R.M. Ramadan. Orientation of multiple principal axes shapes using efficient averaging method. In *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2010*, pages 377 – 381, 2010.

[9] P. Flandrin, G. Rilling, and P. Goncalves. Empirical mode decomposition as a filter bank. *IEEE Signal Processing Letters*, 11(2):112–114, February 2004.

[10] A. Found and H. Müller. Local and global orientation in visual search. In *Perception & Psychophysics*, volume 59, pages 941–963, 1997.

[11] B. Furht, S. W. Smoliar, and H. J. Zhang. *Video and Image Processing in Multimedia Systems*. Kluwer Academic Publishers, 1995.

[12] J. C. Gower. Some distance properties of latent root and vector methods used in multivariate analysis. *Biometrika*, 53:325–338, 1966.

[13] N. Huang, Z. Shen, S. Long, M. Wu, E. Shih, Q. Zheng, C. Tung, and H. Liu. The empirical mode decomposition method and the Hilbert spectrum for non-stationary time series analysis. *Proceedings of the Royal Society of London*, A454:903–995, 1998.

[14] I. T. Jolliffe. *Principal Component Analysis*. 2nd Edition Springer-Verlag, 2002.

[15] S. Kaveti, E. Khwang, and H. Wang. Localization and pose estimation of free-form shapes under Euclidean transformation. In *Proceedings of IEEE International Conference on Industrial Electronics, Control, and Instrumentation, (IECON'96)*, volume 2, pages 1070–1075, August 1996.

[16] S. Krinidis and V. Chatzis. Principal axes estimation using the vibration modes of physics-based deformable models. *IEEE Transactions on Image Processing*, 17(6):1007–1019, June 2008.

[17] T. W. Lee. *Independent Component Analysis - Theory and Applications*. Kluwer Academic Publishers, 1998.

[18] C. Liu and H. Wechsler. A shape and texture based enhanced fisher classifier for face recognition. *IEEE Transactions on Image Processing*, 10(4):598–608, April 2001.

[19] C.-P. Lu, G. D. Hager, and E. Mjolsness. Fast and globally convergent pose estimation from video images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(6):1, June 2000.

[20] D. Nister, O. Naroditsky, and J. Bergen. Visual odometry. In *Proceedings of Computer Vision and Pattern Recognition*, volume 1, pages 652–659, 2004.

[21] Y. Rubner and C. Tomasi. *Perceptual Metrics for Image Database Navigation*. Kluwer Academic Publishers, 2001.

[22] Y. Rui, T. S. Huang, and S. F. Chang. Image retrieval: Current techniques, promising

directions and open issues. *Journal of Visual Communication and Image Representation*, 10:39–62, 1999.

[23] M. H. Safar and C. Shahabi. *Shape Analysis and Retrieval of Multimedia Objects*. Kluwer Academic Publishers, 2003.

[24] T. B. Sebastian, P. N. Klein, and B. B. Kimia. Recognition of shapes by editing shock graphs. In *Proceedings of IEEE International Conference on Computer Vision (ICCV'01)*, pages 755–762, 2001.

[25] A. M. W. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12):1349–1380, Dec 2000.

[26] L. Staib and J. Duncan. Boundary finding with parametrically deformable models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(11):1061–1075, November 1992.

[27] H. Tamura and N. Yokoya. Image database systems: A survey. *Pattern Recognition*, 17(1):29–43, 1984.

[28] R. C. Veltkamp, H. Burkhardt, and H. P. Kriegel. *State-of-the-Art in Content-Based Image and Video Retrieval*. Kluwer Academic Publishers, 2001.

[29] B. Wrobel. *Minimum Solutions for Orientation, Calibration and Orientation of Cameras in Computer Vision*. Springer-Verlag, chapter 2, 2001.

[30] Z. Wu and N. Huang. A study of the characteristics of white noise using the empirical mode decomposition method. *Proceedings of the Royal Society of London*, A460:1597–1611, 2004.

[31] Z. Wu and N. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1):1–41, 2009.

[32] Z. Wu and N. Huang. Ensemble empirical mode decomposition: A noise-assisted data analysis method. *Advances in Adaptive Data Analysis*, 1(1):1–41, 2009.

[33] L. W. Zhao, S. W. Luo, and L. Z. Liao. 3D object recognition and pose estimation using kernel PCA. In *Proceedings of International Conference on Machine Learning & Cybernetics*, pages 3258–3262, 2004.

[34] J. Zunic and M. Stojmenovic. Boundary based shape orientation. *Pattern Recognition, Elsevier*, 41:1785–1798, 2007.

Authors

Krinidis Stelios was born in Kavala, Greece, in 1978. He received the B.Sc. in informatics in 1999 and the Ph.D. degree in informatics in 2004, both from the Aristotle University of Thessaloniki, Thessaloniki, Greece. From 1999 to 2004, he was a researcher and teaching assistant in the Department of Informatics, University of Thessaloniki. From 2005 to 2012, he is a visitor lecturer in the Department of Information Management, Technological Institute of Kavala where he is currently a senior researcher. His current research interests include computational intelligence, pattern recognition, digital signal and 2D and 3D image processing and analysis, and computer vision.

Krinidis Michail was born in Kavala, Greece, in 1981. He received the B.Sc. in informatics in 2002 and the Ph.D. degree in informatics in 2009, both from the Aristotle University of Thessaloniki, Thessaloniki, Greece. From 2002 to 2009, he was a researcher and teaching assistant in the Department of Informatics, University of Thessaloniki. From 2009 to 2012, he is a visitor lecturer in the Department of Information Management, Technological Institute of Kavala where he is currently a senior researcher. His current research interests include face detection and object tracking in video sequences, image clustering and segmentation, face pose estimation and image compression.