A SSA-Based New Framework Allowing for Smoothing and Automatic Change-Points Detection in the Fuzzy Closed Contours of 2D Fuzzy Objects

Vasile Georgescu

Department of Mathematical Economics, University of Craiova, 13 A.I.Cuza str., 200585 Craiova, Romania
vasile.georgescu@feaa.ucv.ro

Abstract. The aim of this paper is to propose a new framework, based on Singular-Spectrum Analysis, allowing for smoothing and automatic change-point detection in the fuzzy closed contours of 2D fuzzy objects. The representation of fuzzy objects is first addressed, by distinguishing between fuzzy regions and fuzzy closed curves. Fuzzy shape signatures are derived in special cases, from which fuzzy time series can be subsequently sampled. Geodesic and Euclidean fuzzy paths and distances between two points in a fuzzy region are next contrasted. Finally, a novel approach to decomposing and reconstructing a fuzzy shape and to automatic change-point detection is proposed, based on a generalization of Singular-Spectrum Analysis so as to deal with complex-valued trajectory matrices. The coordinates themselves, represented as complex numbers are used as a shape signature. This approach is suitable for non-convex and non-star-shaped fuzzy contours.

Keywords: 2D fuzzy regions and fuzzy closed curves, Fuzzy geodesic paths and distances, Singular-spectrum analysis, Complex-valued trajectory matrices encoding fuzzy closed contours, SSA-based change-point detection.

1 Representations of Fuzzy Objects: Fuzzy Regions vs. Fuzzy Closed Curves

Shapes and textures are extremely important features in human as well as machine vision systems. Shape analysis is concerned with two main classes of methods: boundary-based (when only the shape boundary points are used for the description) and region-based (when the whole interior of a shape is considered for description).

In fuzzy shape analysis, however, boundary points are neither strictly delimited, nor independent from texture information, but have assigned to them a fuzzy membership value according to the extent of their belongingness to the object; there is a progressive transition of the membership values from the support outline to the core outline.

Continuous fuzzy shapes can be described as fuzzy geometric objects. A continuous fuzzy geometric object S in $\Re^n$ is defined as a set of pairs $\{(x, \mu_S(x)) \mid x \in \Re^n\}$ where

V. Torra et al. (Eds.): MDAI 2011, LNAI 6820, pp. 174–185, 2011.

$\mu_S : \Re^n \rightarrow [0, 1]$ is the membership function of S in $\Re^n$. An alternative representation of fuzzy geometric objects is given by a set of $\alpha -$ cuts (also called $\alpha -$ supports). For any value $\alpha \in [0, 1]$, the $\alpha - support$ of S, denoted by $S_\alpha = Supp_\alpha(S)$, is the hard subset $\{x \mid x \in \Re^n$ and $\mu_S(x) \geq \alpha \}$ of $\Re^n$. The 0-support will often be referred to as *support* and be denoted by $Supp\ (S)$, while the 1-support will be referred to as *core*. A fuzzy subset with a bounded support is called *bounded*. *S* is said to be *convex* if, for every three collinear points *x*, *y*, and *z* in $\Re^n$ such that *y* lies between *x* and *z*, $\mu_S(y) \geq \min[\mu_S(x), \mu_S(z)]$. A fuzzy subset is called *smooth* if its membership function is differentiable at every location $x \in \Re^n$.

We have to take into account two distinct classes of fuzzy geometric objects: fuzzy regions and fuzzy closed curves. The major difference between them is the shape of the fuzzy boundary.

The membership function of a fuzzy region is non-increasing away from the interior of the object. This means that the $\alpha -$ supports of a fuzzy region are nested, i.e., for membership values $1 = \alpha_1 > \ldots > \alpha_{n+1} = 0$, one has $S_{\alpha_1} \subseteq \ldots \subseteq S_{\alpha_{n+1}}$.

By contrary, the membership function of a fuzzy closed curve has values greater than zero only on the fuzzy boundary and is typically LR-shaped (i.e., it is first increasing from the interior to the modal point of the frontier and then is decreasing to the exterior).

Let's start with the first case. Figure 1 shows a star-shaped fuzzy region. The centroid of the fuzzy shape will be denoted by $C_S = (x_c, y_c)$, where

$$x_c = \frac{\iint x \cdot \mu(x, y)\, dxdy}{\iint \mu(x, y)\, dxdy}; \qquad y_c = \frac{\iint y \cdot \mu(x, y)\, dxdy}{\iint \mu(x, y)\, dxdy} \tag{1}$$

The straight path from the centroid along a radial direction defined with respect to a given angle θ can be parameterized with respect to a parameter $t \in [0,1]$ as follows:

$$\pi^\theta(t) = \{(x^\theta(t), y^\theta(t)) \mid x^\theta(t) = x_c + \rho^\theta(t) \cdot \cos\theta,\ y^\theta(t) = y_c + \rho^\theta(t) \cdot \sin\theta\} \tag{2}$$

where

$$\rho^\theta(t) = \left\| x^\theta(t) - x_c,\ y^\theta(t) - y_c \right\| = \begin{cases} 2t\rho_1^\theta & t \in [0, 1/2] \\ (3-2t)\rho_1^\theta + (2t-1)\rho_0^\theta & t \in (1/2, 1] \end{cases} \tag{3}$$

Given that a fuzzy region is non-increasing away from the interior of the object, its fuzzy boundary along the path π^θ is delimited by two points: $(x_c + \rho_1^\theta \cos\theta,\ y_c + \rho_1^\theta \sin\theta)$ from the interior and $(x_c + \rho_0^\theta \cos\theta,\ y_c + \rho_0^\theta \sin\theta)$ from the exterior, where:

$$\begin{aligned} \rho_1^\theta &= \max\left(\rho^\theta = \left\| x^\theta - x_c,\ y^\theta - y_c \right\| \mid (x^\theta, y^\theta) \in \pi^\theta,\ \mu(x^\theta, y^\theta) = 1\right) \\ \rho_0^\theta &= \min\left(\rho^\theta = \left\| x^\theta - x_c,\ y^\theta - y_c \right\| \mid (x^\theta, y^\theta) \in \pi^\theta,\ \mu(x^\theta, y^\theta) = 0\right) \end{aligned} \tag{4}$$

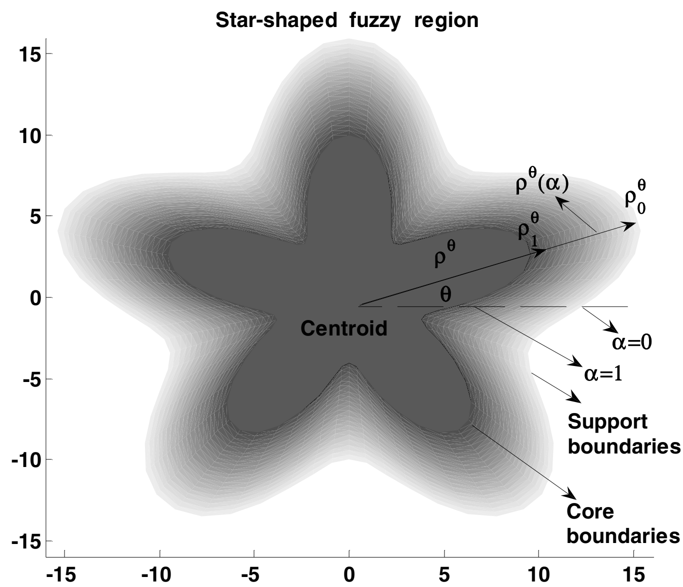


Fig. 1. A star-shaped fuzzy region; the straight path from the centroid along a radial direction

The simplest case is when the membership function along the path π^{θ} is piece-wise linear:

$$\mu\left(x^{\theta}(t),\, y^{\theta}(t)\right)=\mu\left(\rho^{\theta}(t)\right)=\alpha(t)=\begin{cases}1 & t\in[0,1/2]\\ 2-2t & t\in(1/2,1]\end{cases} \tag{5}$$

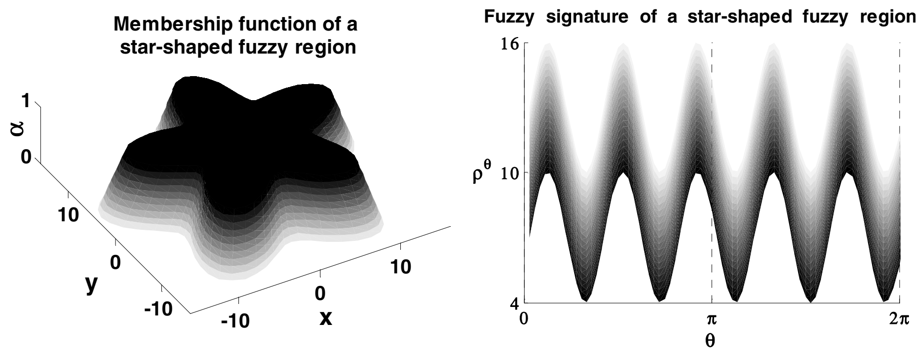


Fig. 2. 3D representation of the membership function of a star-shaped fuzzy region (left); the fuzzy signature of a star-shaped fuzzy region, based on the fuzzy radial distance (right)

Since $\alpha \to \rho_{\alpha}^{\theta}$ is an inverse of the membership function along the path $\pi^{\theta}(t)$, $t\in[1/2,1]$, we can use α to parameterize the Euclidean distance across the $\alpha-$support boundary points:

$$\rho_\alpha^\theta = \rho_1^\theta \cdot \alpha + (1-\alpha) \cdot \rho_0^\theta\,, \quad \alpha \in [0,1], \quad \rho_\alpha^\theta \in [\rho_1^\theta, \rho_0^\theta] \tag{6}$$

Thus, ρ^θ is defined as a fuzzy distance, with the membership function given by:

$$\mu(\rho^\theta) = \begin{cases} 1 & \rho^\theta \in [0, \rho_1^\theta] \\ \dfrac{\rho - \rho_0^\theta}{\rho_1^\theta - \rho_0^\theta} & \rho^\theta \in [\rho_1^\theta, \rho_0^\theta] \end{cases} \tag{7}$$

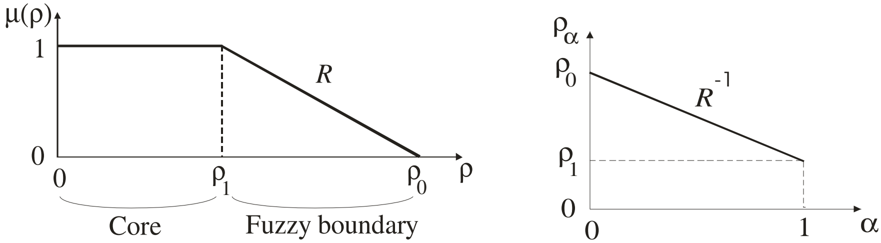


Fig. 3. The membership function of a star-shaped fuzzy region along a radial direction from the centroid

Alternatively, the linear membership function $\mu(\rho^\theta)$ can be replaced by a nonlinear function along with an appropriate parameterization of the path π^θ.

A fuzzy shape signature of a continuous star-shaped fuzzy region can be defined by ρ_α^θ as a fuzzy function of the radial angle θ, with $tg\theta$ being the slope of the straight path between the centroid and the fuzzy boundary (Figure 2, right)

In contrast with a fuzzy region, the membership function of a fuzzy closed curve is typically LR-shaped.

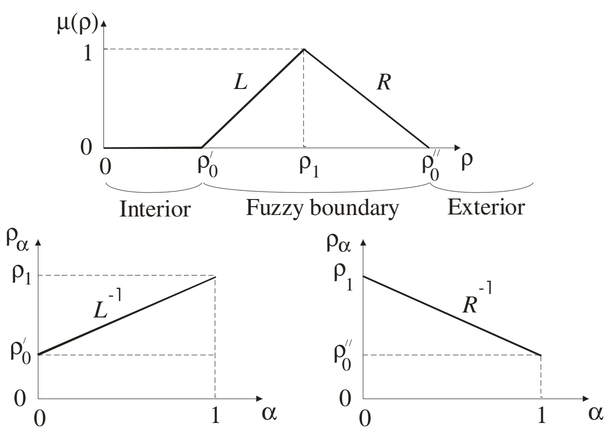


Fig. 4. The membership function of a star-shaped fuzzy closed curve along a radial direction from the centroid

A star-shaped fuzzy closed curve is depicted in Figure 5.

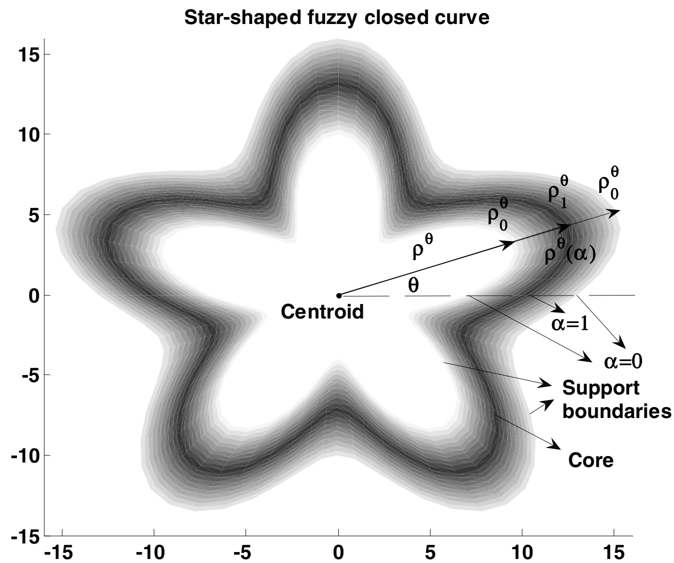


Fig. 5. A star-shaped fuzzy closed curve

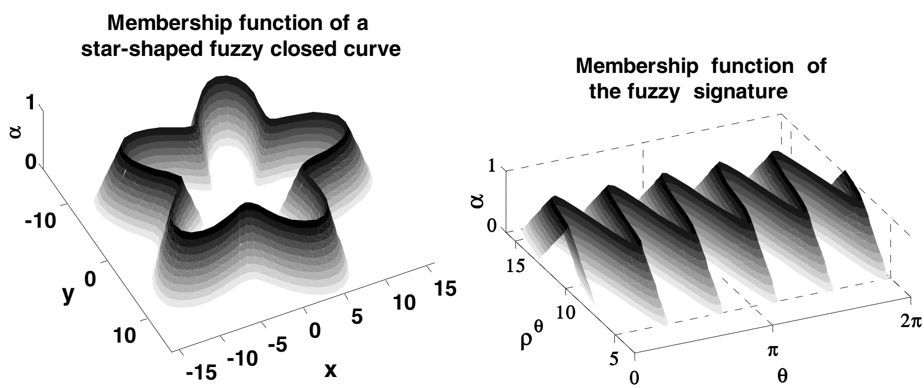


Fig. 6. 3D representation of the membership function of a star-shaped fuzzy closed curve (left); 3D representation of the membership function of the fuzzy signature (right)

The fuzzy shape signature of a star-shaped fuzzy closed curve has the membership function depicted in Figure 6 (right) and can be defined only in terms of the Euclidean notions of path and distance from the centroid to the fuzzy boundary (rather then in a geodesic sense), since the region containing the centroid does not belong to the fuzzy object itself, but to its complement (see the next section).

For a sequence of indices $\tau = \{0, 1, \ldots, n\}$ and the corresponding discrete sequence of angles $\theta_0 = 0 < \ldots < \theta_\tau < \ldots < \theta_n = 2\pi$, a fuzzy-valued time series $\{\rho_\tau\} = \{\rho(\theta_\tau)\}$, $\tau = 0, 1, \ldots, n$, can be sampled from the continuous fuzzy signature (Figure 7).

Moreover, for each α_i in a discrete sequence $\alpha_0 = 0 < \alpha_1 < \ldots < \alpha_{p-1} < 1$, a pair of real-valued time series $\left(\underline{\rho}_\tau(\alpha_i),\ \overline{\rho}_\tau(\alpha_i)\right) = \left(L_\tau^{-1}(\alpha_i), R_\tau^{-1}(\alpha_i)\right)$ can be sampled, as well as a real-valued time series from the modal values in the core ($\alpha_p = 1$).

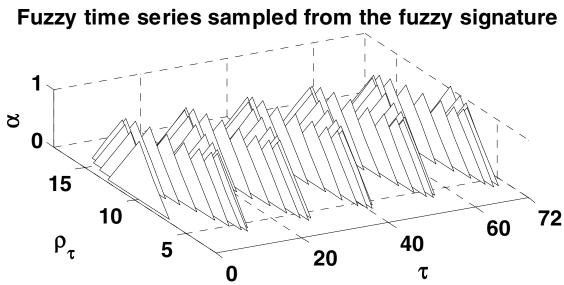


Fig. 7. Fuzzy-valued time series sampled from the continuous fuzzy signature (fuzzy radius-vector function)

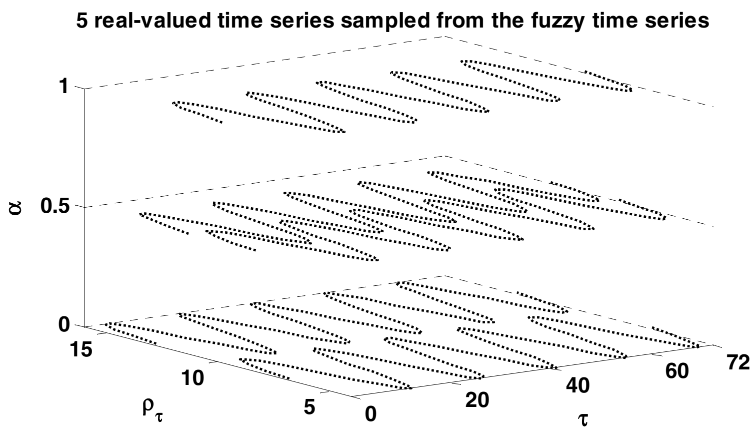


Fig. 8. Five real-valued time series sampled from the fuzzy time series, for $\alpha = 0, 0.5, 1$

2 Geodesic vs. Euclidian Fuzzy Paths and Distances

The notion of a geodesic path between two points in a fuzzy subset has been introduced with respect to different purposes and formal settings, in both the continuous and digital fuzzy geometry. For instance, this notion is the main ingredient in the definition of the fuzzy distance transform (FDT), proposed in [8] by Saha et al. (2002).

A *path* π in $\Re^n$ from a point $x \in \Re^n$ to another point (not necessarily distinct) $y \in \Re^n$ is a continuous function π : [0, 1]→ $\Re^n$ such that $\pi(0)=x$ and $\pi(1)= y$. The *length* of a path π in S, denoted by $\Pi_S(\pi)$, is the value of the following integration

$$\Pi_S(\pi) = \int_0^1 \mu_S(\pi(t)) \left| \frac{d\pi(t)}{dt} \right| dt \tag{8}$$

i.e., $\Pi_S(\pi)$ is the integral of Euclidian distances weighted by the membership values (in S) along π.

For the path defined by the equations (2), (3) and (5), we have

$$\frac{d\pi(t)}{dt} = \begin{cases} \left(-2\rho_1^\theta \sin\theta, 2\rho_1^\theta \cos\theta\right) & t \in [0, 1/2] \\ \left((2\rho_1^\theta - 2\rho_0^\theta)\sin\theta, (-2\rho_1^\theta + 2\rho_0^\theta)\cos\theta\right) & t \in (1/2, 1] \end{cases} \tag{9}$$

$$\left| \frac{d\pi(t)}{dt} \right| = \begin{cases} 2\rho_1^\theta & t \in [0, 1/2] \\ 2\left|\rho_0^\theta - \rho_1^\theta\right| & t \in (1/2, 1] \end{cases} \quad \Rightarrow \quad \Pi_S(\pi) = \rho_1^\theta + \frac{1}{2}\left|\rho_0^\theta - \rho_1^\theta\right| \tag{10}$$

which means that the length of the path across the fuzzy boundary is contracted by a factor of 1/2 with respect to the length of the equivalent Euclidean path.

When a path passes through a low density (low membership) region, its length increases slowly and the portion of the path in the complement of the support of S contributes no length. This approach is useful to measure regional object depth, object thickness distribution, etc.

Let $\zeta_S(x, y)$ denote a subset of positive real numbers defined as

$$\zeta_S(x, y) = \{\Pi_S(\pi) \mid \pi \in P(x, y)\} \tag{11}$$

i.e., $\zeta_S(x, y)$ is the set of all possible path lengths in S between x and y. The *fuzzy distance* from $x \in \Re^n$ to $y \in \Re^n$ in S, denoted as $\omega_S(x, y)$, is the infimum of $\zeta_S(x, y)$; i.e.,

$$\omega_S(x, y) = \inf \zeta_S(x, y) \tag{12}$$

Actually, the fuzzy distance ω_S is a geodesic distance, which means that the shortest paths (when they exist) in a fuzzy subset S between two points $x, y \in \Re^n$ are not necessarily a straight line segment even when S is convex.

Concepts such as connectivity and geodesic distance between pixels or voxels in a 2D or 3D digital image have been proved to play a key role in the fuzzy digital geometry, since it has been introduced in [6] by Rosenfeld (1984).

There are two main approaches in measuring distances when considering fuzzy spatial objects: the first one basically compares only the membership functions representing the concerned fuzzy object, while the other one combines spatial distance between objects and membership functions, thus taking into account both spatial information and information related to the imprecision attached to the image object.

Distances between two points in a fuzzy set are typically addressed in order to find the best path in the geodesic sense in a spatial fuzzy set.

Distances from a point to a set are used when computing distance from a point to a complement of a fuzzy set, i.e., performing distance transform.

The distances between sets are used in shape matching.

A geodesic distance between points in a fuzzy set was introduced in [1] by Bloch (2000), being defined conditionally to a reference set *X*. It naturally incorporates some concepts involved in its crisp equivalent, such as Euclidian distance, path lengths and connectivity. Thus, a geodesic distance $d_X(x, y)$ from *x* to *y* is the length of a shortest path from *x* to *y*, completely included in *X*. Let μ be a fuzzy set on the space S. The definition of the geodesic distance relies on the degree of connectivity in μ between two points *x* and *y* of *S*, as defined by Rosenfeld (1984),

$$c_\mu(x, y) = \max_{L(x,y)} \left[\min_{t \in L(x,y)} \mu(t) \right] \tag{13}$$

where $L(x, y) = t_1, \ldots, t_n$ denotes a path from $x = t_1$ to $y = t_n$, consisting of a sequence of points in *S* according to the discrete connectivity defined on *S*. Let $L^*(x, y)$ denote the shortest path between *x* and *y* on which c_μ is reached; this path is not necessarily unique and can be interpreted as a geodesic path descending as little as possible in terms of membership degrees. Let $l(L^*(x, y))$ denote its length (the number of points along the path). Then the geodesic distance in μ between *x* and *y* is defined as

$$d_\mu(x, y) = \frac{l\left(L^*(x, y)\right)}{c_\mu(x, y)} \tag{14}$$

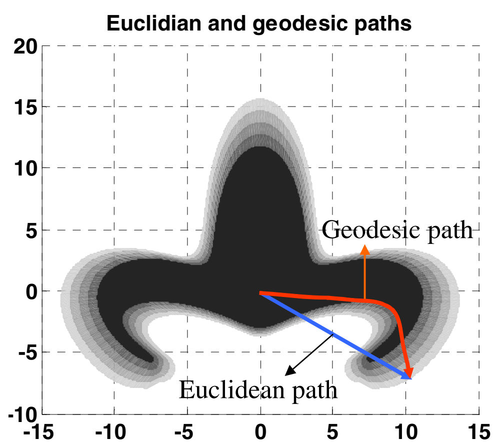


Fig. 9. Euclidean vs. geodesic paths in a non-convex, non-star-shaped fuzzy region

If $c_\mu(x, y) = 0$, then $d_\mu(x, y) = \infty$, which corresponds to the result obtained for the classical geodesic distance in the case where *x* and *y* belong to different connected components.

Figure 9 contrasts the Euclidean and geodesic paths between points in a non-convex, non-star-shaped fuzzy region.

3 A Novel Approach to Decomposing and Reconstructing a Fuzzy Shape and to Automatic Change Point-Detection, Based on SSA

In case of fuzzy shapes that are not star-shaped, abstracting a fuzzy signature based on a radial fuzzy distance is inappropriate. A fuzzy region may has the centroid belonging to its interior (or not!), while a fuzzy closed contour typically has the centroid belonging to its complement. For a connected fuzzy region with the centroid belonging to its interior, a fuzzy geodesic path (not necessarily unique) usually exists. However, geodesic paths cannot be used for the decomposition and the reconstruction of a fuzzy shape because they use weighted distances and thus may severely distort the reconstructed shape, when comparing to the original coordinates. On the other hand, a geodesic path is meaningless in case of fuzzy closed curves, when the centroid typically belongs to the complement of the fuzzy object. A Euclidean radial path is also inappropriate in case of non-star-shaped fuzzy shapes since the path may intersect the fuzzy boundary many times (see Figure 9).

The proposed novel strategy is to use the coordinates themselves as a shape signature, but represented as complex-valued numbers in the complex plane. A complex-valued fuzzy time series can then be sampled from such a continuous complex-valued fuzzy signature, allowing the machinery of time series analysis to be subsequently used. The method is general, that is, not constrained to convex, or star-shaped fuzzy regions.

Both denoising and change-point detection can be carried out by a powerful method called Singular-Spectrum Analysis (SSA). SSA is a nonparametric method for time series structure recognition and identification ([4]). The SSA algorithm has two basic stages: decomposition and reconstruction. Basically, it first builds the trajectory matrix associated to a time series, whose columns are formed by a sequence o lagged vectors extracted from the time series with a sliding window. Afterward, the *singular value decomposition* is applied to the trajectory matrix. For the reconstruction of the de-noised part of the time series the most dominant singular values (and the corresponding singular vectors) are considered, whereas the remaining singular values are used to compute the residuals (associated with noise).

Let start by assuming that the fuzzy object is represented in terms of $\alpha-$supports.

There are two distinct cases:

1. the case of a fuzzy region, where the membership function is non-increasing away from the interior of the object: assuming a counterclockwise rotation along the fuzzy boundary, one complex-valued $\alpha-$level time series can be sampled from the continuous complex-valued $\alpha-$support contour, which can be written as $z_t^\alpha = x_t^\alpha + i\, y_t^\alpha \in C$, for $t = 1, \ldots, N$. Here α is chosen from a finite non-increasing sequence $\alpha_1 = 1 < \alpha_2 < \ldots < \alpha_n = 0$.

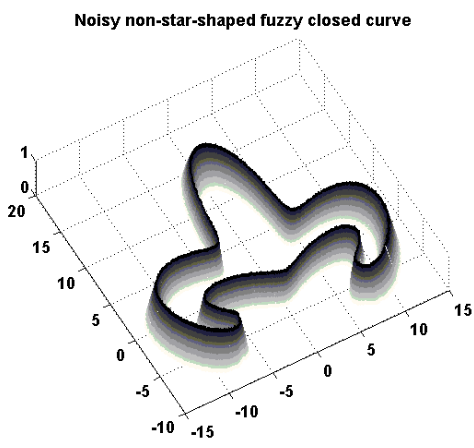
2. the case of a fuzzy closed curve, where the membership function is typically LR-shaped: assuming a counterclockwise rotation along the fuzzy boundary, two complex-valued $\alpha-$level time series can be sampled from the continuous complex-valued $\alpha-$support contour, a left-side one and a right-side one, i.e., $(z_t^\alpha)^L = (x_t^\alpha)^L + i(y_t^\alpha)^L \in C$, $(z_t^\alpha)^R = (x_t^\alpha)^R + i(y_t^\alpha)^R \in C$, for $t = 1, \ldots, N$. Eventually, if the core reduces to a singleton, then $(z_t^1)^L = (z_t^1)^R$.

Now, a complex-valued trajectory matrix can be defined for each complex-valued time series in turn:

$$X^\alpha = \begin{pmatrix} z_1^\alpha & z_2^\alpha & \cdots & z_K^\alpha \\ z_2^\alpha & z_3^\alpha & \cdots & z_{K+1}^\alpha \\ \vdots & \vdots & \vdots & \vdots \\ z_M^\alpha & z_{M+1}^\alpha & \cdots & z_N^\alpha \end{pmatrix}, \quad z_t^\alpha \in C \tag{15}$$

The proposed generalization of SSA is founded on the ACM Algorithm 358 for Singular Value Decomposition of a complex matrix. The decomposition theorem (Businger and Golub, [2]) can be stated as follows: each and every $M \times K$ complex-valued matrix X can be reduced to diagonal form by unitary transformations U and V, $X = U\, diag[\sigma_1, \ldots, \sigma_K] V^H$, where $\sigma_1 \geq \ldots \geq \sigma_K \geq 0$ are real-valued scalars, called the singular values of X. Here U is an $M \times K$ column orthogonal matrix, V an $K \times K$ unitary matrix and V^H is a Hermitian transpose of V. The columns of U and V are called the left and right singular vectors of X, respectively.

As it is shown in Figures 10 and 11, the proposed generalization of SSA-based algorithm for complex-valued trajectory matrices can be successfully applied to decomposing and reconstructing (de-noising) non star-shaped fuzzy contours.



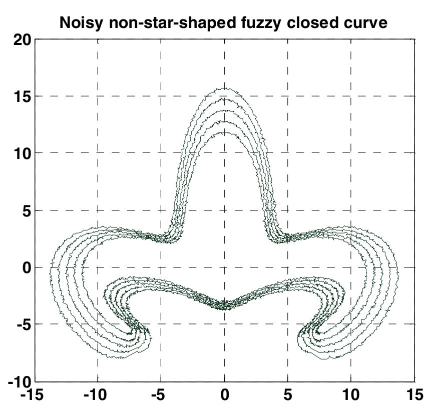


Fig. 10. The membership function of a noisy, non-star-shaped fuzzy closed curve (left); several α-cuts (right)

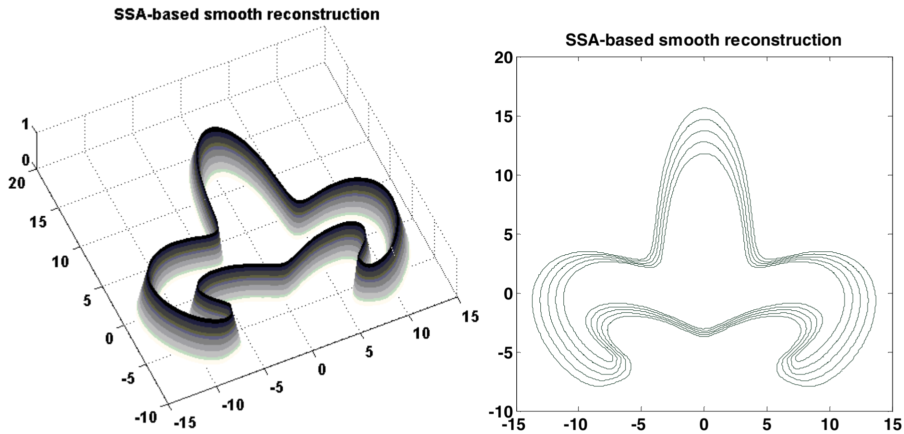


Fig. 11. The SSA-based smooth reconstruction of a non-star-shaped fuzzy closed curve (left); several α-cuts (right)

SSA is also at the core of a powerful change-point detection algorithm ([5]). The detection statistic is defined with respect to the squared distance to the subspace spanned by the eigenvectors of the lag-covariance matrix, computed in a sequence of moving time intervals. The novelty is that all calculations are adapted to be made in complex spaces (for instance, the distance between complex numbers is considered).

The detected change-points and the distance-based test statistics are shown in Figure 12. The location of change points is in the local minima of test statistic function.

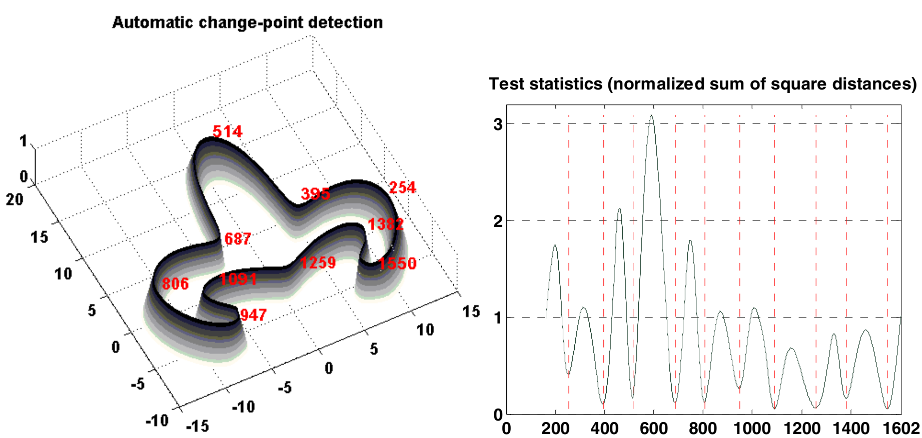


Fig. 12. Automatic selection of landmark positions through change-point detection (left); detecting change points from Distance detection statistic (right)

4 Conclusion

This paper extends my approach in [3] to a fuzzy context. SSA is generalized to deal with complex-valued trajectory matrices, encoding the coordinates themselves as complex numbers, in order to carry out the decomposition and reconstruction of a fuzzy closed contour, as well as change-point detection. This approach is suitable for any kind of fuzzy contours, including non-convex and non-star-shaped ones.

References

1. Bloch, I.: Geodesic balls in a fuzzy geodesic mathematical morphology. Pattern Recognition 33, 897–905 (2000)
2. Businger, P.A., Golub, G.H.: Algorithm 358: Singular value decomposition of a complex matrix. Comm. ACM 12, 564–565 (1969)
3. Georgescu, V.: A Novel and Effective Approach to Shape Analysis: Nonparametric Representation, De-noising and Change-Point Detection, Based on Singular-Spectrum Analysis. In: Torra, V., et al. (eds.) MDAI 2011. LNCS(LNAI), vol. 6820, pp. 162–173. Springer, Heidelberg (2011)
4. Goljadina, N., Nekrutkin, V., Zhigljavky, A.: Analysis of Time Series Structure: SSA and related techniques. Chapman and Holl, London (2001)
5. Moskvina, V.: Applications of the singular-spectrum analysis for change point detection in time series. Ph.D. thesis, School of Mathematics, Cardiff University, Cardiff (2001)
6. Rosenfeld, A.: The fuzzy geometry of image subsets. Pattern Recognition Letters 2, 311–317 (1984)
7. Rosenfeld, A.: Fuzzy geometry: An updated overview. Information Sciences 110, 127–133 (1998)
8. Saha, P.K., Wehrli, F.W., Gomberg, B.R.: Fuzzy distance transform: Theory, algorithms, and applications. Computer Vision and Image Understanding 86, 171–190 (2002)