

A Classification Method Based on Two Separating Hyper Surfaces

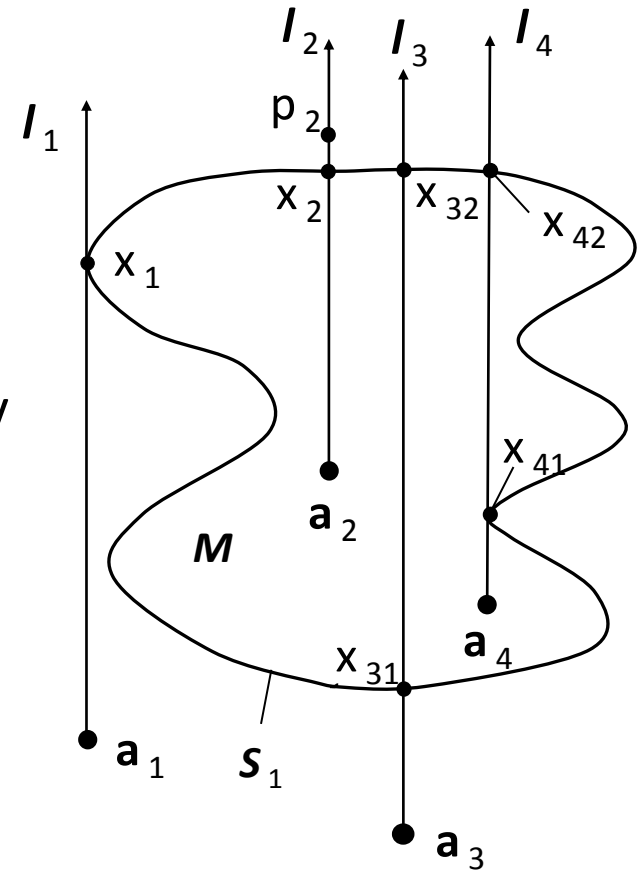
Valeri Ilchev

Svetozar Ilchev

- Hyper-Surface Classification (HSC)
- Definition of HSC_2
- Our Approach for determining the HSC_2
- Classification Rule
- Experimental Results & Comparison to the Classic Approach
- Conclusions

Hyper-Surface Classification

- There are multiple classes (sets of points):
 - ⊙ Defined in multi-dimensional feature space
 - ⊙ Separated by hyper-surfaces.
- The problem is:
 - ⊙ Find a second hyper-surface classification (HSC_2) for each class
 - ⊙ Define a rule that determines if an arbitrary point is located inside or outside of a class.
- Classic definition of the rule
 - ⊙ Set a ray starts at the point to be classified.
 - ⊙ If the ray crosses the HSC an odd number of times the point belongs to the set; otherwise it is out of the set.
 - ⊙ **PROBLEM:** A ray tangent to the HSC leads to a false classification.



Definition of HSC_2

- Create a second HSC (HSC_2), which is close enough to original HSC (HSC_1) so that $HSC_1 \subset HSC_2$ (Inr).

- Mathematical definition of HSC_2 :

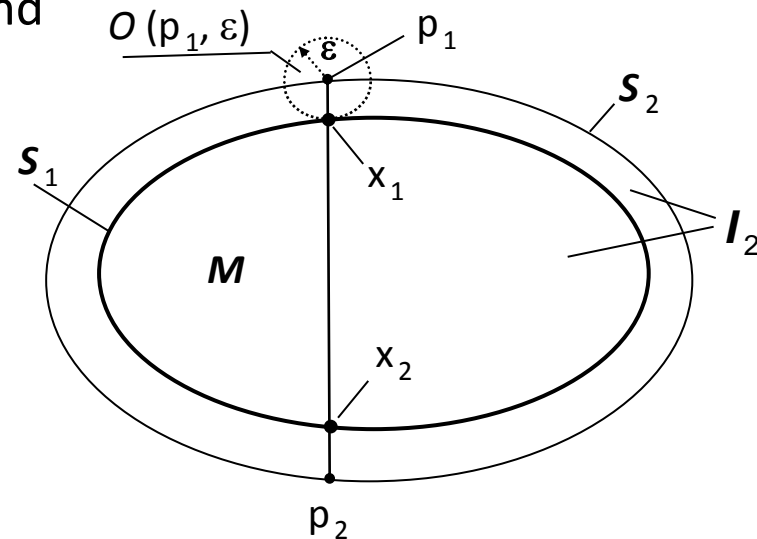
- ⊙ Jordan-Brouwer Theorem: Let S be a closed set of $n-1$ dimensions in the space \mathbf{R}^n . If it is homeomorphic to a sphere in \mathbf{R}^{n-1} , then its complement $\mathbf{R}^n \setminus S$ consists of two connected components, which have S as their common boundary. One of them is an inner component with regard to S , the other one is an outer component.

- ⊙ Definition: The hyper-surface S_2 is an outer envelope of the set M (delimited by the hyper-surface S_1) if $S_1 \cap S_2 = \emptyset$ and every closed ball $O(p, \varepsilon)$ with a center at the point p and a radius $\varepsilon = const$ (a sufficiently small number):

$O(p, \varepsilon) = \{x : |x - p| \leq \varepsilon; \varepsilon > 0\}$ is such that for $\forall p \in S_2$, the following conditions hold:

$O(p, \varepsilon) \cap S_1 \neq \emptyset \wedge O(p, \varepsilon) \subset K_M$ and $O(p, \varepsilon) \cap Int(M) = \emptyset$, where

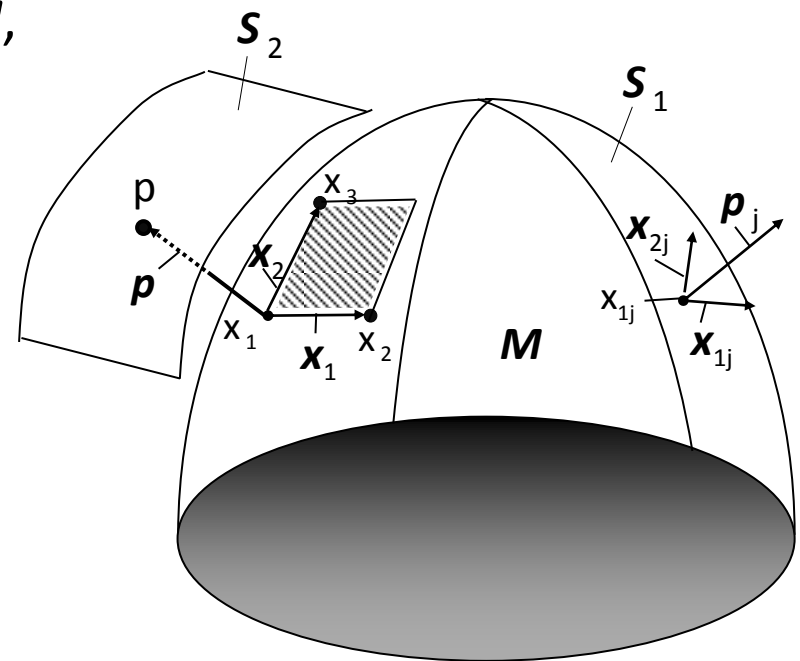
$K_M = \mathbf{R}^n \setminus Int(M)$ and $M = S_1 \cup Int(M)$.



Our Approach for determining the HSC_2

■ Practical calculation of HSC_2 :

- ⊙ Choose n points x_1, x_2, \dots, x_n on the hyper-surface S_1 , which are sufficiently close to one another and define the vectors $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_{n-1}$, situated on S_1 , where all vectors have a common initial point x_1 :
 $\mathbf{x}_1 = x_2 - x_1, \mathbf{x}_2 = x_3 - x_1, \dots, \mathbf{x}_{n-1} = x_n - x_1$.
- ⊙ Compute the vector product $\mathbf{p} = \mathbf{x}_1 \times \mathbf{x}_2 \times \dots \times \mathbf{x}_{n-1}$.
- ⊙ Assuming the point p is outside of M , the vector product, normalized to a sufficiently small number ε , will represent the image of every point $x_j \in S_1$, on the outer envelope S_2 .
- ⊙ We can regard \mathbf{p} as the radius of the ball $O(p, \varepsilon)$, centered at the point p . Then we have: $O(p, \varepsilon) \cap S_1 = x_1$.



Classification Rule

■ Definition of the *Classification Theorem*:

⊙ Let us define a ray I_a starting at the point a such that: $I_a \cap S_1 \neq \emptyset$ and $I_a \cap S_2 \neq \emptyset$, which defines two sets of points: $X_I = \{x_i : x_i \in I_a \cap S_1\}$ and $P_I = \{p_i : p_i \in I_a \cap S_2\}$.

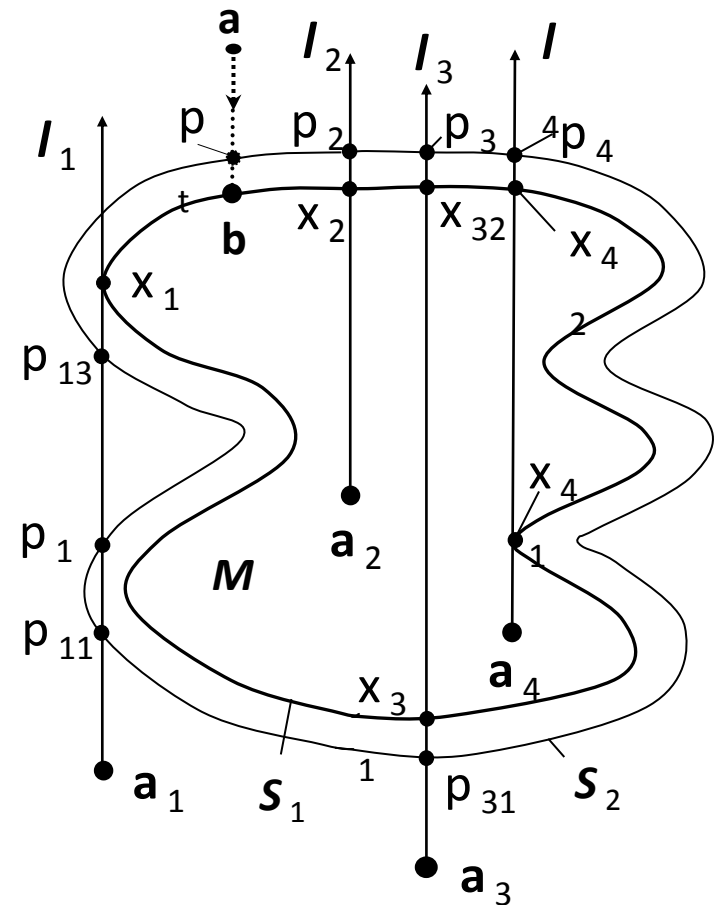
⊙ If the point closest to the point a of all points belonging to the set $X_I \cup P_I$ is a point of the set X_I , then $a \in M$;

⊙ if the point closest to the point a of all points belonging to the set $X_I \cup P_I$ is a point of the set P_I , then $a \notin M$.

■ For the the ray I_4 :

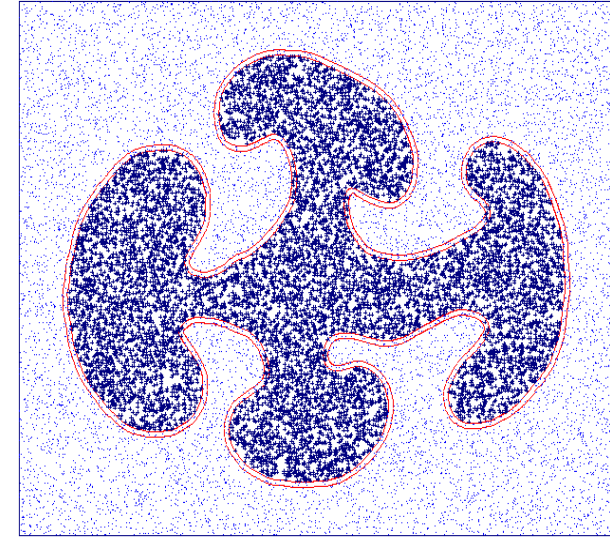
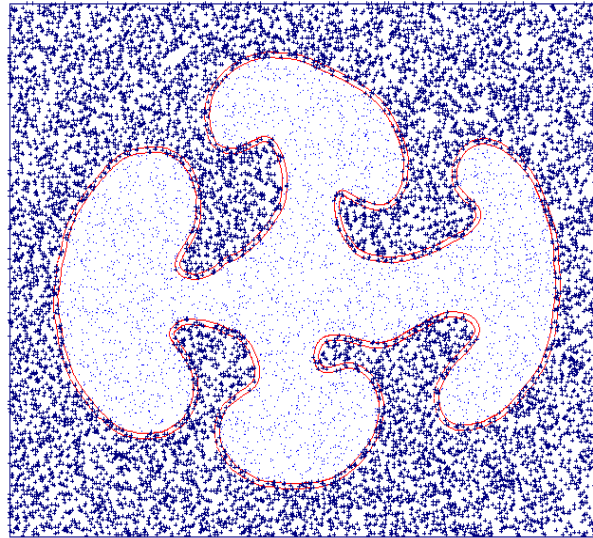
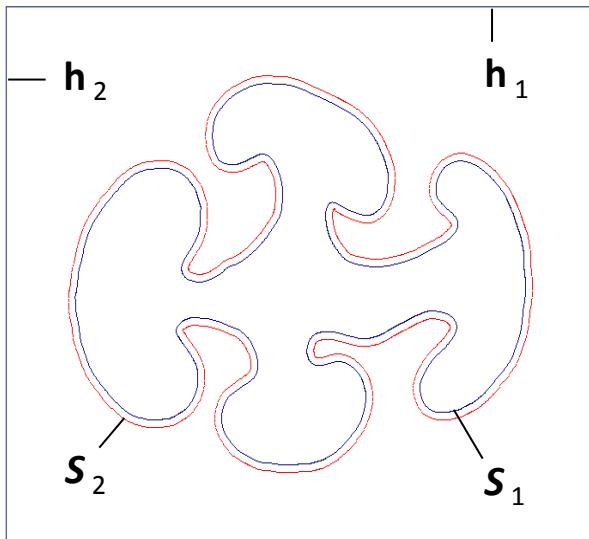
$$d(a_4, x_{41}) < d(a_4, x_{42}) < d(a_4, p_4)$$

leading to a correct classification compared to the classic rule.



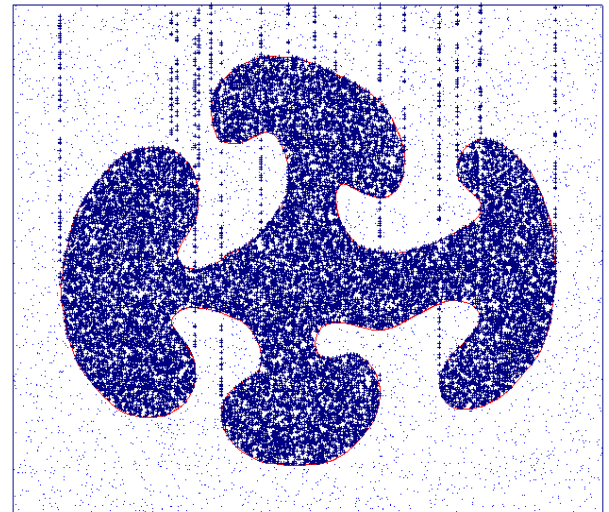
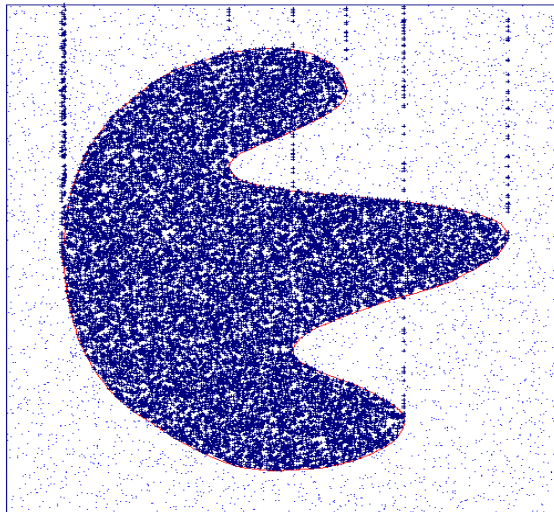
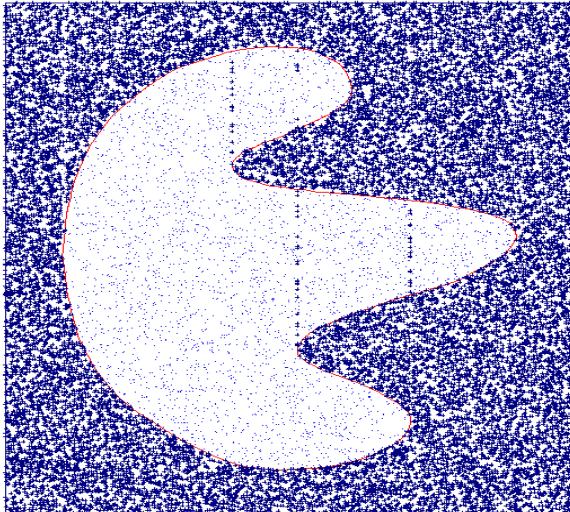
Experimental Results

- S_1 is the contour of the figure; S_2 is created by means of the vector product.
- The rectangular borders are filled with 10^4 individual points obtained by the generation of random numbers on the axes h_1 and h_2 .
- The set of points filling the rectangle is divided into one external subset and one internal subset by means of the classification rule proposed in our paper.
- The direction of the rays is not important. In our examples, the rays are parallel to h_2 and directed upwards – in the direction of the increasing y -values.
- All internal and external points have been classified correctly, which demonstrates in practice the reliability of our classification method.



Comparison to the Classic Approach

- A single classification closed curve is used and the number of intersections (even or odd) of the ray with this curve is used as a classification criterion.
- As before, the initial point of each ray is the point which is currently under examination.
- The dark vertical lines represent points that are classified incorrectly according to the classic rule – the points corresponding to the rays that are tangent to the classification closed curve
- The results of the proposed classification rule (previous slide) are reliable for the whole set of points, which is not true for the results of the classic rule (this slide), especially when the figures have complex outlines (more tangent rays).



- The proposed method and its innovative classification rule solve successfully the problem of the separation of the points inside a given class from the points outside the class – regardless of the complexity of the form of the classification closed curve \mathcal{S}_1 .
- If we assume that the classes are separable (they do not intersect), the test examples of our experiment show one of the most complex classification cases – the classes have a common classification closed curve \mathcal{S}_1 (in the general case – a HSC), which separates them into one internal and one external subset. In practice, the more common case is the case of partially non-strictly separable sets, i.e. their HSC partially coincide without intersecting the other class (if the classes are separable).
- In summary, the method proposed in the paper is reliable and applicable to a wide variety of complex practical classification problems.

Thank You for Your Attention!