Rough set approach to approximate reasoning via calculus of standards

Adam Szmigielski

Polish-Japanese Institute of Information Technology
Koszykowa 86 02008 Warsaw, Poland
Institute of Computing Science, Warsaw University of Technology
Nowowiejska 15/19 Warsaw, Poland
`email: szmigielski@ia.pw.edu.pl`

Abstract. In this paper we present methods of calculus of standards based on rough mathematical morphology. Many methods of mathematical morphology are originally based on Minkowski operators in affine spaces. In the paper [8] Polkowski proposed a scheme of mathematical morphology based on ideas of rough sets theory. Our goal is to develop this idea in approximate reasoning by means of calculus of standards. We propose a measure of information loss in this reasoning and present simple numerical experiments to illustrate this.

keywords: *rough set theory, rough mathematical morphology, calculus of standards, approximate reasoning*

1 Introduction

1.1 Mathematical morphology

Mathematical morphology is concerned with problems of filtering of complex objects by means of some geometrical operations. We begin with a brief introduction to mathematical morphology [4] [10] and for purposes of simplicity and applicability we restrict ourselves to the euclidean case in which objects are either subsets of a euclidean space E^n of n-dimensions or subsets of an affinely closed subspace $V \subseteq E^n$(in our experiments here we assume that V is the digital space Z^n). Morphological operations are generated by two binary operations called, respectively, the *Minkowski sum* and the *Minkowski difference.*

Definition 1 *The Minkowski sum, denoted $\oplus$, is defined for any pair $A, B \subseteq V$ by the formula*

$$A \oplus B = \{x + y : x \in A, y \in B\} \tag{1}$$

where $+$ denotes the addition in the space E^n. Similarly, the Minkowski difference is denoted $\ominus$ and defined via the formula

$$A \ominus B = \{x \in V : \{x\} \oplus B \subseteq A\}. \tag{2}$$

We fix a set St of *standard objects* (structuring objects) for an object $B \in St$ and an object $X \subseteq V$. We define the *dilation* $d_B(X)$ of a given object X by the standard object B as the object $d_B(X) = X \oplus B$ and the *erosion* $e_B(X)$ of X by B is defined via the identity $e_B(X) = X \ominus B$. Dilations and erosions generate new morphological operations the *opening* and the *closing.*

Definition 2 *Opening $o_B(X)$ of X by B is the composition of erosion and dilation*

$$o_B(X) = d_B(e_B(X)) \tag{3}$$

and closing $c_B(X)$ of X by B is the composition of dilation and erosion

$$c_B(X) = e_B(d_B(X)). \tag{4}$$

It follows immediately from the definitions above that the local characterization of opening is the following

$$o_B(X) = \{x \in V : \exists y.(x \in \{y\} \oplus B \subseteq X)\} \tag{5}$$

and similarly for closing

$$c_B(X) = \{x \in V : \forall y.(x \in \{y\} \oplus B \Rightarrow X \cap (\{y\} \oplus B) \neq \emptyset)\}. \tag{6}$$

From (5) and (6) it is easily inferred that the operations of morphological opening, respectively, closing, resemble topological operations of interior, respectively, closure [3] with the difference that the former are defined with respect to a fixed object B while the latter are defined with respect to the whole family St satisfying the requirements for an open basis for topology on V.

1.2 Mathematical morphology based on rough sets theory

Basic notions of rough sets theory The *theory of rough sets* [7] is also concerned with finding a useful set of classifying features for a given class of objects. Objects under consideration form a set U and the description of objects is provided by means of a set A of attributes (features) where each attribute $a \in A$ is represented as a mapping $a : U \longrightarrow V_a$ of the set U into the value set of a. Objects in the set U are perceived through the attribute set A; in consequence, objects having identical descriptions relative to A are identified into a common class and treated as one generalized object. Formally, this is expressed by means of the relation of indiscernibility $IND(A) = \{(x, y) \in U \times U : a(x) = a(y)\}$ for each $a \in A$. Equivalence classes of this relation form elementary granules of our knowledge: all objects in the equivalence class $[x]_{IND(A)}$ are pairwise indistinguishable.

Definition 3 *For a given subset $X \subseteq U$, the lower approximation of set X defined via the formula*

$$A_{-}X = \{x \in U : [x]_{IND(A)} \subseteq X\}, \tag{7}$$

consists of objects which are elements of X with certainty. The upper approximation of set X is defined via the formula

$$A^{+}X = \{x \in U : [x]_{IND(A)} \cap X \neq \emptyset\}, \tag{8}$$

and it consists of objects which cannot be excluded from X with certainty.

The set X is approximately defined by the pair $(A_{-}X, A^{+}X)$ of its approximations. A set X is *rough* in case $A_{-}X \neq A^{+}X$. The difference $A^{+}X - A_{-}X$ is called the *boundary region of* X. It may be observed that rough set theoretic approximations and morfological approximations are formally related to each other via formulae (5) and (6). Let us observe that approximations $A_{-}X, A^{+}X$ may be expressed in topological language. Denoting by $\pi_{IND(A)}$ the partition topology, obtained by taking the set of equivalence classes of $IND(A)$ as its open base, we have:

$$A_{-}X = int_{\pi_{IND(A)}}X \tag{9}$$

and

$$A^{+}X = cl_{\pi_{IND(A)}}X \tag{10}$$

where *int*, respectively, *cl* are, respectively, the *interior* the *closure operators* [11] , [12]. Let us observe, that in case of the indiscernibility relation $IND(A)$ on the set U of objects, rough sets may be characterized topologically as those sets X for which $int_{\pi_{IND(A)}}X \neq cl_{\pi_{IND(A)}}X$.

Rough morphology For our purpose we will assume that $IND(a_{m+1}) \subseteq IND(a_m)$ for each m i.e. the attribute a_{m+1} induces a finer classification than the attribute a_m. In this context a natural topology appears: we denote by π_A the topology obtained by taking the union $\cup\{[x]_{IND(a_m)} : x \in U,\ m = 1, 2, ...\}$ of equivalence classes of all relations $IND(a_m)$ as an open base. To give a practical example consider the Euclidean n-space E^n along with a compact set $Z \subseteq E^n$.

Example: Let a sequence $(a_m)_m$ of attributes be given, where the attribute a_m partitions the space E^n into the cubes $\prod_{1 \leq i \leq n}(j_i + k_i \cdot 2^{-m}, j_i + (k_i + 1) \cdot 2^{-m}]$ where $j_1, j_2, ..\ ., j_n$ are integers , $k_i = 0, 1, ...\ 2^m - 1$ is an integer for each $i \leq n$ [8].

In this context the rough sets metrics can be defined (cf. [8]). We recall these definitions of metrics D_o, D and $D*$ [8] and we begin with the function ρ_n on U, defined for each n, such that:

- $\rho_n(x, y) = 1$ in case $[x]_{IND(a_n)} \neq [y]_{IND(a_n)}$
- $\rho_n(x, y) = 0$ otherwise.

The function

$$\rho(x, y) = \sum_{n=1}^{\infty} \rho_n(x, y) \cdot 10^{-n} \tag{11}$$

is then a pseudo - metric compatible with the topology π_A. In practical case we have finitely many attributes, say n, and then

$$\rho(x, y) = \sum_{i=1}^{n} \rho_i(x, y) \cdot 10^{-i}. \tag{12}$$

We denote by the symbol $\rho(x, C)$ the ρ *-distance* of x from $C \subseteq U$

$$\rho(x, C) = \inf\{\rho(x, y) : y \in C\}. \tag{13}$$

The metric D_o on the space of π_A- closed sets is defined in the standard way

$$D_o(C,D)=\max\{\max\{\rho(x,C):x\in D\},\max\{\rho(x,D):x\in C\}\} \tag{14}$$

and the metric D is defined as follows:

$$D((Q,T),(Q',T'))=\max\{D_o(Q,Q'),D_o(T,T')\}, \tag{15}$$

where pair (Q,T) represents a set $X\subseteq U$ in the following way:

$$Q=A^+X=cl_{\pi_{IND(A)}}X \tag{16}$$

and

$$T=U-A_-X=U-int_{\pi_{IND(A)}}X. \tag{17}$$

Similarly, the metric D^* is defined as follows [8]

$$D^*((Q,T),(Q',T'))=\max\{D_o(Q,Q'),D_o(T,T'),D_o(Q\cap T,Q'\cap T')\}. \tag{18}$$

We are interested also in sets which may be rough for some initial number of attributes $a_1,a_2,...a_k$ but become exact with respect to a_{k+1}. We will represent such sets as sequences of the form $(Q_m,T_m)_m$, where Q_m and T_m are defined as Q and T above except that we replace $IND(A)$ with $Ind(a_m)$ (see Polkowski [8] for details). We call these sets *almost - rough sets.*

2 Calculus of standards based on rough sets

2.1 Description of Standards

By a *standard St* we understand a distinguish element (often "virtual" not accessible in reality) in universe U of objects. A standard usually describes well-known objects, which is useful in some practical sense. Many problems can be reduced to find a set of standards (i.e. useful objects) and then, in decision process, to classify objects to standards. The classification process itself can be difficult and may take a lot of effort. It would be desired to reduce this effort. This goal may be obtained by the reduction of standard description. To do this, we may consider the sequences of standard approximations. We recall that all important applications of a complete context are based either directly or implicitly (via some generalizations) on the renowned *Banach contraction principle*:

Theorem 1 *For given a complete metric space* (X,ρ) *and a contracting mapping* $f:X\longrightarrow X$ *with a contraction constant* $c\in(0,1)$ *(i.e.* $\rho(f(x),f(y))\le c\cdot\rho(x,y))$*) there exists a unique fixed point* x_f *(i.e.* $f(x_f)=x_f$*) of the mapping* f*. Moreover,* x_f *can be found as the limit of any sequence* $(a_n)_n$ *where*

1. $a_o\in X$ *is arbitrary and* $a_{n+1}=f(a_n)$ *for* $n=0,1,2...$ *. Even the member of the sequence* $(a_n)_n$ *giving a desired approximation to the fixed point can be found: we have*
2. $\rho(a_n,x_f)\le c^n\cdot(1-c)^{-1}\cdot\rho(a_o,f(a_o))$.

Proof can be found in [1].

In [8] the approximation problem of set F depending of granularity of set F' was investigated. The equality $a^+_mF^{'}=a^+_mF$ implies that the Hausdorff distance between $F^{'}$ and F is less than $2^{-m+\frac{1}{2}}$ hence for large enough values of m both sets $F^{'}$ and F are satisfactorily close also with respect to the metric D' - introduced for the case in example in section 1.2. The metric D' on almost rough sets is constructed as follows:
For each n we define the metric ρ_n as above. Then we define the metric D_{o_n} induced by ρ_n via the formula (14) and we induce the metric D'_n on pairs of $\pi_{IND(a_n)}$ closed sets by means of the formula:

$$D^{'}_n((Q_n,T_n),(Q^{'}_n,T^{'}_n))=\max\{D_{o_n}(Q_n,Q^{'}_n),D_{o_n}(T_n,T^{'}_n)\}. \tag{19}$$

Finally, the metric D' is defined as follows:

$$D^{'}((Q_n,T_n)_n,(Q^{'}_n,T^{'}_n)_n)=\sum_{i=1}^{\infty}D^{'}_n((Q_n,T_n),(Q^{'}_n,T'_n))\cdot 10^{-n}. \tag{20}$$

We have

$$If\quad \lim_n D'(F',F)=0\quad then\quad \lim_n D(F',F)=0. \tag{21}$$

Assume now that $f : C \longrightarrow C$ is a contracting mapping on a compact set $Z \subseteq E^n$ with a contraction coefficient $c \in (0,1)$ (*e.g.* one selects compact sets $C_o^1, .., C_o^k$ and lets $F_o = \cup_i C_o^i, F_1 = \cup_i f(C_o^i)$ *etc.* Let $K = D(F_0, F_1)$. In order to have $D(F', F) < \varepsilon$ for a chosen positive threshold ε, it is sufficient to check that $a_m^+ F' = a_m^+ F$ with $m = \lceil \frac{1}{2} - \log_2 \varepsilon \rceil$ and

$$n \geq n_o = \frac{log_2[2^{-m+\frac{1}{2}} \cdot K^{-1} \cdot (1-c)]}{\log_2 c} \tag{22}$$

This compression causes some loss of information. It would be useful to measure this lost information. We would like to propose the measure of information lost as follow.

Definition 4 *For the given set F let $F_m^{'}$ denotes approximation of the set F with respect to the granularity m. The loss of information of this approximation is the number I_m determined via the formula*

$$I_m = -\log_2 \frac{|F \cap F_m^{'}|}{|F|} \tag{23}$$

where $|F|$ denotes the cardinality of F.

Let us observe, that the value of I_m is equal to zero when $F = F_m^{'}$ for some granularity m and is positive in other cases.

Let us give an example.

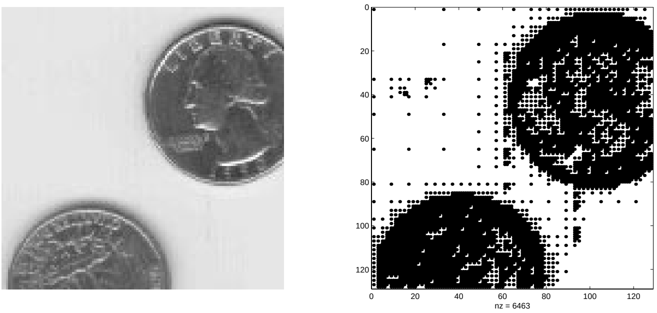

Fig. 1. Image and its black and white decomposition

We can generate the sequences of binary images (black and white pixels) which are presented in Fig. 3.

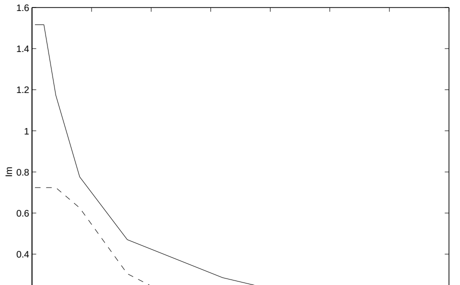

Fig. 2. Loss of information: dashed line - lower approximation, solid line -upper approximation

Fig. 2 shows the information loss in the approximation process. The values of I_m are plotted against granularity. The image consists of 128 x 128 pixels and we start with the granularity, which all image is taken as either one white or one black pixel and at each step we divide a given granularity pixel into 4 new pixels. It could be observed (in Fig 2), that even at the 4^{th} level of granularity, we have quite good description of the standard (in the sense they are clearly distinguished one from another).

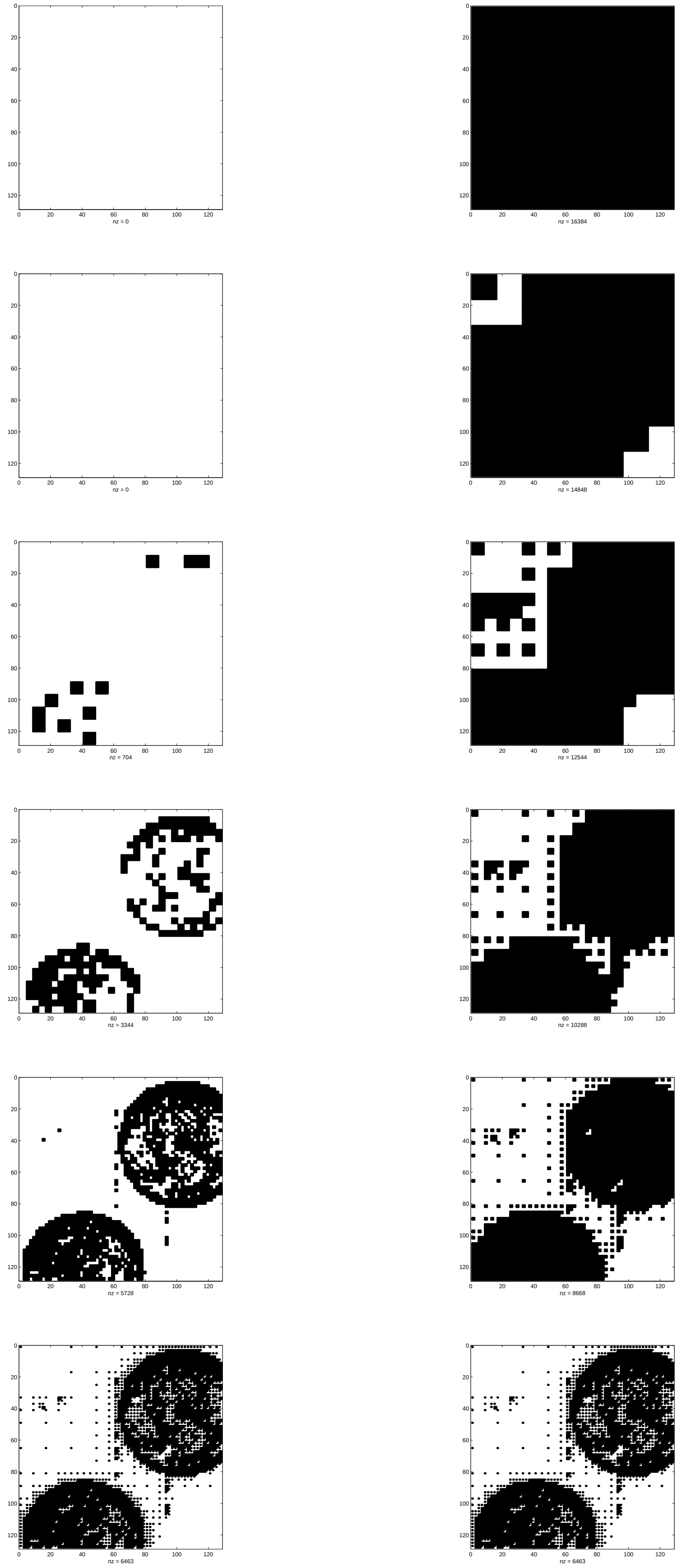

Fig. 3. Morphology - the sequences of lower and upper approximation in granularity

2.2 Calculus of standards

In real world applications standards are often given with some tolerance. Classifying the objects, we compare them to the standards. The results of comparing objects to standards are distances between them. It is quite clear, that these distances are usually greater then zero. As shown in section 2.1, description of the standards depends on granularity of information. Algorithms of classical mathematical morphology are often NP-complete, so it takes a lot of time to compare an object to the standards [5]. Complexity of those comparing algorithms often makes not possible to work in real time tasks. One of the ways of solving this problem is changing the granularity level of comparison. We would like to show experimental results of comparing objects to standards depending on the information granularity. Similarity to the standard is calculated via the formula

$$\mu(F, St) = \frac{|F \cap St|}{|St|} \tag{24}$$

where μ is the *rough membership function* [6]

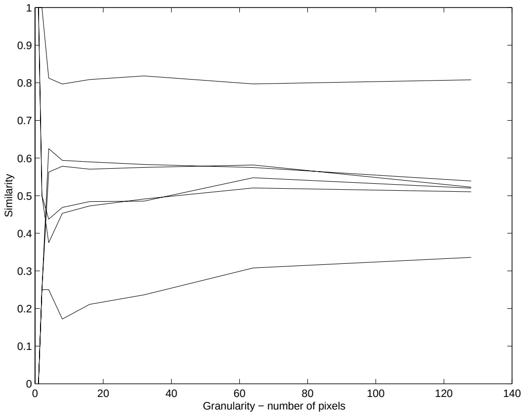

Fig. 4. Similarities to the standards

In [2] authors research mobile robots' ability to learn in the context of sonar granularity. They obtain interesting results, which demonstrate, that mobile robots can best learn with very low sonar granularity. This result is similar and compatible with our results in this paper. Our simulations indicate that high resolution cameras may be not necessary in the task of robot vision. We expect to carry out experiments with a mobile robot control to support our analysis.

3 Conclusions

In this scheme, the underlying space of points (e.g. pixels) is partitioned into disjoint cells (classes) by means of some primitive attributes (features) and morphological operations are performed on classes, which allows for compression of data (by reducing complexity). Calculus of standard methods may be useful in approximate representation of information.

Acknowledgement

The author appreciates help from the Polish - Japanese Institute of Information Technology during the preparation of this paper.

References

1. S. Banach, 1922, Sur les opérations dans les ensembles abstraits et leur application aux équations intégrales, Fundamenta Mathematicae 3 pp. 133 - 181.
2. M.Dorigo, M. Colombetti, 1998, Robot shaping an experiment in behavior engineering, A Bradford Book The MIT Press Cambridge, Massachusetts, London, England.
3. K. Kuratowski, 1966, Topology, vols.I, II, Academic Press, New York.
4. G. Matheron, 1975, Random Sets and Integral Geometry, Wiley, New York.
5. M. Nieniewski, 1998, Mathematical morphology in image processing, PLJ, Warsaw (in Polish).
6. Z. Pawlak, A.Skowron, 1994, Rough membership function, R.R. Yager, M. Fedrizzi, J. Kacprzyk (eds.) Advances in Dempster Shafer Theory, John Wiley, New York, pp.251-271
7. Z. Pawlak, 1991, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer, Dordrecht.
8. L.Polkowski, 1993, Mathematical morphology of rough sets, Bull. Polish Acad. Ser. Sci.Math. 41(3) (1993), pp. 241 - 273.
9. L. Polkowski, A. Skowron (eds.), 1998, Rough Sets in Knowledge Discovery and Data Mining. Methodology and Applications, Springer (Physica Verlag).
10. J. Serra, 1988, Image Analysis and Mathematical Morphology: Theoretical Advances, Academic Press, New York.
11. A. Skowron, 1988, On topology in information systems, Bull. Polish Acad. Ser. Sci. Math. 36 pp. 477 - 479.
12. A. Wiweger, 1988, On topological rough sets, Bull. Polish Acad. Ser . Sci. Math. 37, pp. 89 - 93.