Machine Learning 32, 207–224 (1998)

Localization vs. Identification of Semi-Algebraic Sets*

SHAI BEN-DAVID shai@cs.technion.ac.il
MICHAEL LINDENBAUM mic@cs.technion.ac.il
Department of Computer Science, Technion, Haifa 32000, Israel

Editors: Thomas Dietterich and David Haussier

Abstract. How difficult is it to find the position of a known object using random samples? We study this question, which is central to Computer Vision and Robotics, in a formal way. We compare the information complexity of two types of tasks: the task of *identification* of an unknown object from labeled examples input, and the task of *localization* in which the identity of the target is known and its location in some background scene has to be determined.

We carry out the comparison of these tasks using two measuring rods for the complexity of classes of sets; The Vapnik-Chervonenkis dimension and the ϵ-entropy of relevant classes. The VC-dimension analysis yields bounds on the sample complexity of performing these tasks in the PAC-learning scenario whereas the ϵ-entropy parameter reflects the complexity of the relevant learning tasks when the examples are generated by the uniform distribution (over the background scene). Our analysis provides a mathematical ground to the intuition that *localization* is indeed much easier than *identification*.

Our upper-bounds on the hardness of *localization* are established by applying a new, algebraic-geometry based, general tool for the calculation of the VC-dimension of classes of algebraically defined objects. This technique was independently discovered by Goldberg and Jerrum. We believe that our techniques will prove useful for further VC-dimension estimation problems.

Keywords: learning theory, PAC, Vapnik-Chervonenkis dimension, localization, identification, recognition, computer vision

1. Introduction

Object recognition, a fundamental task of computer vision, usually deals with the following situation: one observes a scene, extracts some measurements out of it, and uses them to judge whether certain objects are present in the scene, and what are their positions. In the basic form of this task, called *localization*, the identity of the object is known, and one tries to guess its position correctly. A different and more general task is *identification*, where the only advance information about the object is its membership in some known library of objects. The *identification* task is to discover both the shape and the position of the target object.

The aim of this work is to provide some rigorous mathematical analysis of the information-complexity of these tasks. Our analysis yields a justification to the clear intuition that *localization* is an easier task than *identification*.

Let a *class of images* be a class of objects that are transformed instances of one particular object. Such a class depends on the original object and on the type of transformations

*An extended abstract of this paper appeared in the Proceedings of the Sixth Annual ACM Conference on Computational Learning Theory (Ben-David & Lindenbaum, 1993).

allowed. One can view *localization* as the task of *identification* from a library that is a class of images.[1]

We wish to quantify the 'complexity' of classes of images, for different objects and groups of transformations, and to compare it to the 'complexity' of some natural library classes to which the objects belong. The measures of complexity of a class that we shall investigate are two: The Vapnik-Chervonenkis dimension and the metric (ϵ-) entropy. These measures are relevant to the learning difficulty of classes in the distribution-free PAC model and with respect to the fixed uniform distribution (respectively).

If no limitations are imposed on the shape of the object, classes of images may be arbitrarily difficult to learn (see Claim 4 in Section 3 for an example). We shall therefore limit our discussion to objects that are semi-algebraic sets in $\mathbb{R}^n$, i.e., can be defined by boolean combinations of polynomial inequalities.

As for the families of allowed transformations, we shall consider affine transformations of $\mathbb{R}^n$ as well as some subgroups such as isometries (or Euclidean transformations), which correspond to repositioning of rigid bodies, and Similarity transformations, which allow also uniform scale change. These groups of transformations are commonly used to model image acquisition distortions that arise when the input information is derived from a (two-dimensional) picture taken by a camera whose position relative to the object is unknown.

In Section 2 we define our objects of research—the Semi-Algebraic sets. We then develop some tools for proving lower bounds on the VC-dimension of classes of such sets. Section 3 investigates the VC-dimension of classes of transformed images of a semi-algebraic object. We introduce a technique for bounding the VC-dimension of classes parameterized by algebraically defined sets of real numbers. We believe that this technique, relying on a classical result of Milnor, is a powerful tool that will be applicable far beyond the issues discussed in this paper. Goldberg and Jerrum have, independently, discovered and analyzed a very similar technique (Goldberg, 1992; Goldberg & Jerrum, 1995).

Viewed from the distribution-free PAC learnability angle, these results imply, on one hand, upper bounds on the number of examples needed for the *localization* of an object of some semi-algebraic degree, and on the other hand, much higher lower bounds on the number of examples needed for the *identification* of such an object from among all objects of the same degree.

Section 4 carries these results over to the setting of learnability with respect to the (fixed) uniform distribution. This is done by analyzing the ϵ-entropy of the relevant classes, and showing that, for wide classes of semi-algebraic objects, the entropies, under the metric induced by the uniform distributions, approach their maximum possible values (over all probability distributions).

Finally, in Section 5, we discuss the relevance our results to object recognition.

2. The VC-dimension of semi-algebraic classes

We wish to show that *localization* is, in some sense, an easy task. This statement may fail when the object one wishes to localize is very 'wild', an example of such a case is given later (see Claim 4). We shall therefore focus on well behaved geometrical objects—Semi-Algebraic subsets of $\mathbb{R}^n$.

Definition 1 (Semi-algebraic and polynomial sets).

- A *semi-algebraic* (*open*) *set* of degree (k, m) in $\mathbb{R}^n$ is a set that can be represented as a boolean combination of k sets of the form $\{\bar{x} \in \mathbb{R}^n : f_j(\bar{x}) > 0\}$ where the functions f_j are real polynomials of maximal degree m.
- The *semi-algebraic class* of degree (k, m) over $\mathbb{R}^n$ is the collection of all (k, m)-semi-algebraic subsets of $\mathbb{R}^n$, namely,

$$SA^n_{(k,m)} \stackrel{\text{def}}{=} \{A \subseteq \mathbb{R}^n : A \textit{ is a semi-algebraic} \text{ open set of degree } (k, m)\}.$$

- A *polynomial set* is a semi-algebraic open set of degree $(1, m)$, for some finite m.

From the Computer Vision point of view, even polynomial objects of modest degrees (e.g., 4) seem to enable the description of complicated objects, thereby providing sufficient representation power (Keren, Cooper & Subrahmonia, 1992; Terzopoulos et al., 1987). The class we consider here is even richer: besides polynomial objects it also contains combinations of them which include, e.g., polygonal objects (which, for k being the number of polygon sides, are semi-algebraic sets of degree $(k, 1)$).

The VC-dimension of a concept class is defined as follows:

Definition 2 (Vapnik-Chervonenkis dimension). Let X be some set and $\mathcal{K}$ a collection of its subsets (a concept class).

- We say that $\mathcal{K}$ *shatters* a set of points $A \subseteq X$, if, for every $B \subseteq A$, there exists some $C \in \mathcal{K}$ such that $C \cap A = B$.
- The *Vapnik-Chervonenkis dimension* (in short, VC-dimension) of $\mathcal{K}$ is the maximum size of a set shattered by $\mathcal{K}$. (If $\mathcal{K}$ shatters sets of unbounded size, we say that $VCdim(\mathcal{K})$ is ∞.)

A finite concept class is always associated with a finite VC-dimension:

Example 1. Let $\mathcal{K}$ be a concept class containing a finite number, N, of concepts. A finite point set of n points contains 2^n different subsets, but only N subsets of any point set A may be represented as $c \cap A$ (for any $c \in \mathcal{K}$). Therefore, no set larger than $\lfloor \log N \rfloor$ may be shattered, and $VCdim(\mathcal{K}) \leq \lfloor \log N \rfloor$.

Infinite concept classes, like the semi-algebraic classes specified above, may have finite or infinite VC-dimension. The following theorem of Dudley is a key tool for their analysis.

Theorem 1 (Dudley, 1984). *For a real-valued function f on some domain X, let $pos(f)$ denote $\{x \in X : f(x) > 0\}$. If $\mathcal{H}$ is a real vector space of functions from X to $\mathbb{R}$ then the VC-dimension of $\{pos(f) : f \in \mathcal{H}\}$ equals the linear (vector space) dimension of $\mathcal{H}$. Furthermore, for h being any real-valued function on X, the VC-dimension of $\{pos(f + h) : f \in \mathcal{H}\}$ is also equal to this linear dimension.*

Let $CC(n,m) = \binom{n+m}{m}$ denote the number of subsets of size m of a set of size $m+n$. The following claim is a straightforward application of Dudley's theorem:

Claim 1. *The VC-dimension of $SA^n_{(1,m)}$—the class of all degree m polynomial subsets of $\mathbb{R}^n$ is $CC(n,m)$.*

(This claim also follows from results of Cover (1965).) We can now readily compute the VC-dimension of the classes of regions in $\mathbb{R}^2$ that are bounded by graphs of polynomial functions.

Corollary 1. *Let $B^2_m \stackrel{\text{def}}{=} \{\hat{p} : p$ is a polynomial of degree at most $m\}$, where $\hat{p} \stackrel{\text{def}}{=} \{(x,y) \in \mathbb{R}^2 : y < p(x)\}$. Then, for every $m \in \mathbb{N}$, the VC-dimension of the class B^2_m is $m+1$.*

The proof is immediate by noting that, for every polynomial p, $\hat{p}$ equals $\{\mathrm{pos}(f) : f(x,y) = p(x) - y\}$ where the degree of p is at most m and y plays the role of the fixed function, h, from Dudley's theorem. Another natural family of classes of semi-algebraic sets in $\mathbb{R}^2$ is the classes of convex k-edge polygons. Such a class is a subclass of $SA^2_{(k,1)}$.

Claim 2. *The VC-dimension of the class of convex k-edge polygons is at least $2k$.*

This can be easily verified by considering a set of $2k$ points equally spaced on the boundary of a circle and noting that, for every subset of it, there is a convex k-gon that contains this subset and no other point of the set.

2.1. The VC-dimension of classes generated by classes of known dimension

We shall now present some tools for the evaluation of the VC-dimension of classes of subsets of $\mathbb{R}^n$ from the dimensions of their generating classes.

Definition 3. Let C be a class of subsets of $\mathbb{R}^n$,

- C, is *shift invariant* if, for every $c \in C$ and for every $\bar{t} \in \mathbb{R}^n$, $c + \bar{t} \stackrel{\text{def}}{=} \{\bar{x} + \bar{t} : \bar{x} \in c\}$ is also in C.
- C is *scale invariant* if, for every subset $c \in C$ and for every $\alpha \in \mathbb{R}$, $\alpha c \stackrel{\text{def}}{=} \{\alpha x \mid x \in c\}$, is also in C.
- The *element-wise union* of two classes, $C_1 \underline{\cup} C_2$, is $\{a \cup b \mid a \in C_1, b \in C_2\}$.
- The *element-wise intersection* of two classes, $C_1 \underline{\cap} C_2$, is $\{a \cap b \mid a \in C_1, b \in C_2\}$.

Lemma 1. *Let C_1, C_2 be classes of bounded subsets in $\mathbb{R}^n$ whose VC-dimension is finite. Then*

1. *If C_1 and C_2 are shift invariant, or scale invariant, then so are $C_1 \underline{\cup} C_2$ and $C_1 \underline{\cap} C_2$.*
2. *If C is scale invariant, then, for every $r > 0$, bounding C to the ball around the origin $B(O,r) = \{\bar{x} \in \mathbb{R}^n : \|\bar{x}\| \leq r\}$, does not affect its VC-dimension, i.e., $VCdim(C) = VCdim(C \underline{\cap} \{B(O,r)\})$.*
3. *If C_1 and C_2 are shift invariant then $VCdim(C_1 \underline{\cup} C_2) \geq VCdim(C_1) + VCdim(C_2)$.*

Proof: We leave the proof of the first two claims to the reader. For the third claim, consider two sets of points S_1, S_2, of sizes $VCdim(C_1)$ and $VCdim(C_2)$ respectively, each shattered by the corresponding class. Let SC_1, SC_2 be subclasses of C_1, C_2, that contain the minimal number of subsets needed to shatter these sets. The shift invariance of C_1, C_2 and the boundedness of their members, imply that the S_1, S_2 sets can be chosen such that the intersection between any member of SC_1 and any member of SC_2 is null. Every union of an element of SC_1 with an element of SC_2 is a member of $SC_1 \underline{\cup} SC_2$. It follows that the class $SC_1 \underline{\cup} SC_2$, which is a subclass of $C_1 \underline{\cup} C_2$, shatters $S_1 \cup S_2$, implying that its VC-dimension is at least $VCdim(C_1) + VCdim(C_2)$. $\square$

To see how tight this bound is, we compare it to the upper bound derived in (Dudley, 1984). There, Dudley considers the element-wise union (or intersection) of two classes, C_1 and C_2. He applies Sauer lemma (Sauer, 1972) to get:

$$VCdim(C_1 \underline{\cup} C_2) \leq \sup\left\{ r \in \mathbb{N}; \sum_{i=0}^{VCdim(C_1)} \binom{r}{i} \sum_{j=0}^{VCdim(C_2)} \binom{r}{j} > 2^r \right\}.$$

Now, letting $r^* = VCdim(C_1 \underline{\cup} C_2)$, Dudley's inequality becomes:

$$2^{r^*} < \sum_{i=0}^{VCdim(C_1)} \binom{r^*}{i} \sum_{j=0}^{VCdim(C_2)} \binom{r^*}{j} < (r^*)^{VCdim(C_1)} (r^*)^{VCdim(C_2)}$$

implying

$$\frac{VCdim(C_1 \underline{\cup} C_2)}{\log(VCdim(C_1 \underline{\cup} C_2))} \leq VCdim(C_1) + VCdim(C_2).$$

2.2. *Lower-bounding the VC-dimension of semi-algebraic classes*

We use Lemma 1 to obtain a general lower bound on the VC-dimension of semi-algebraic sets.

Claim 3. *The VC-dimension of the class $SA^n_{(k,m)}$ of semi-algebraic subsets of $\mathbb{R}^n$ of degree (k, m), is at least $\frac{k}{2}\binom{m+n}{m}$.*

Proof: Consider the class $A = B \underline{\cap} C$, where B is the class of polynomial objects of degree m, and C the class of balls (of finite radii). By Claim 1 there exists a set of size $CC(n, m)$ that is shattered by the class B. Since there exists some ball in C that contains it, the VC-dimension of the new class A is at least $CC(n, m)$. Note that A contains only bounded sets, and is shift and scale invariant. Applying Lemma 1 we conclude that the VC-dimension of the class $\underline{\bigcup}_{i=1\ldots k/2} A$ is at least $\frac{k}{2} CC(n, m)$. The claim now follows by noting that this latter class is a class of semi-algebraic sets of degree (k, m). $\square$

In fact, as we prove later in Claim 8, there is an upper bound on the VC-dimension of classes of semi-algebraic objects which is not very different from this lower bound—it is larger by only a logarithmic factor.

3. The VC-dimension of classes of images

In this section we wish to bound from above the VC-dimension of classes of images, i.e., collections of images of a given object under a family of transformations. As the Claim 4 below demonstrates, such a bound depends on the object generating the class. We shall consider objects that are semi-algebraic sets of a given degree. We will show that, for wide classes of transformations, the dimension of the corresponding classes of images of any (k, m)-semi-algebraic set in $\mathbb{R}^n$ is substantially below the dimension of $SA^n_{(k,m)}$—the class of all sets of the same degree.

The families of transformations we consider include the families of translations, rotations, scale changes, combinations of them, and the group of all affine transformations.

Definition 4.

- An object, V, is a subset of $\mathbb{R}^n$.
- A transformation, t, is a mapping, $t(\bar{x})$ from $\mathbb{R}^n$ to $\mathbb{R}^n$. A class of transformations is denoted T.
- For any $t \in T$, let V_t denote the t-transformed image of V, i.e., $V_t = \{t(\bar{x}) : \bar{x} \in V\}$.
- The class of T-images of V, denoted by $C_T(V)$, is $\{V_t : t \in T\}$.
- A translation is a transformation of $\mathbb{R}^n$ defined by an equation of the form $t(\bar{x}) = \bar{x} + \bar{y}$, for some vector, $\bar{y} \in \mathbb{R}^n$. The class of all translations is denoted $\mathcal{T}$.

First, we show that even the most simple transformation class: one-dimensional translation, may yield a class of images associated with infinite VC-dimension.

Claim 4. *There exists a one-dimensional object*, V, *such that VCdim*$(C_{\mathcal{T}}(V))$ *is infinite.*

Proof: Consider the set $S = \{1/2, 1/3, 1/4, \ldots\}$. Let $\{A_n\}_{n\in\mathbb{N}}$ enumerate all its finite subsets. (A possible enumeration of these subsets is, for example, $A_1 = \emptyset$, $A_2 = \{1/2\}$, $A_3 = \{1/3\}$, $A_4 = \{1/2, 1/3\}$, $A_5 = \{1/4\}$, $A_6 = \{1/2, 1/4\}, \ldots$.) Now define the object as $V^* = \bigcup_{n\in\mathbb{N}}(A_n + n)$. We claim that the class of its translations, $C_{\mathcal{T}}(V^*) = \{V^* + t : t \in \mathbb{R}\}$, has an infinite VC-dimension. To see this, just note that $C_{\mathcal{T}}(V)$ shatters any finite subset of S. □

The above claim makes it clear that simple transformations do not necessarily elicit simple image classes. In the rest of the section we shall show that the complexity of classes of images depends on the complexity of both the object being transformed and the family of the transformations.

3.1. A mapping between points and transformation subclasses

The main idea behind the following analysis is to translate the question of the combinatorial complexity of a class $C_T(V)$ to some geometric question in the space of transformations T. Once this translation is established, the mighty tools of algebraic geometry can be called in to solve our combinatorial problem. We shall now describe a general framework for such translations.

Given a family of transformations of $\mathbb{R}^n$, T, the first step we take is to note that any object, $V \subseteq \mathbb{R}^n$, induces a mapping of points of $\mathbb{R}^n$ to subsets of T. Namely, every point $\bar{x} \in \mathbb{R}^n$ may be mapped into the subset of transformations $K_{\bar{x}}^V = \{t \in T : \bar{x} \in V_t\}$. Note that this mapping is dual to the mapping from members of T to subsets of $\mathbb{R}^n$ defined by mapping t to the set V_t.

Consider now a set of points $S = \{x_1, x_2, \ldots, x_N\}$ in $\mathbb{R}^n$. Fixing an object $V \subseteq \mathbb{R}^n$, every subset $A \subset S$ corresponds to the subset

$$W(A, S) \stackrel{\text{def}}{=} \{t \in T : V_t \cap S = A\}.$$

of the family, T, of transformations. The following claim follows immediately from the definitions:

Claim 5. *For any $A \subseteq S$ and $t \in T$,*
- *For $t \in W(S, A)$ and $x \in S$, $t \in K_x^V$ iff $x \in A$.*
- *For any object $V \subseteq \mathbb{R}^n$ and a family T of transformations, the class of images, $C_T(V)$, shatters a set $S \subseteq \mathbb{R}^n$ iff, all of the members of $\{W(A, S) : A \subseteq S\}$ are nonempty.*[2]

We have therefore reduced the calculation of the VC-dimension of classes of images to counting the number of nonempty $W(A, S)$ sets of transformations. This reduction is the basis for the subsequent derivations of this section.

3.2. Parameterizing the transformations

The next step we take is to represent T parametrically, with parameters forming some parameter space $\mathbb{R}^p$. For example, if T is the family of translations of $\mathbb{R}^n$, it can be naturally parameterized by assigning the parameter $\bar{y} \in \mathbb{R}^n$ to the translation (of $\mathbb{R}^n$) $t(\bar{x}) = \bar{x} + \bar{y}$.

As mentioned in Section 1, the transformations of most practical interest are Translations, Euclidean, Similarity and Affine transformations, a family that includes all the former families. A linear affine transformation on $\mathbb{R}^n$ is defined by a pair, (A, b), where A is an $n \times n$ matrix and $b \in \mathbb{R}^n$. Such a transformation $H = (A, b)$ acts on $\mathbb{R}^n$ by $H(x) = Ax + b$. We shall restrict our attention to nonsingular transformations, i.e., transformations that are one-to-one or, equivalently, their defining matrix A is regular. For these transformations the inverse transformation, denoted $H' = (A', b')$, always exists, and its parameters, the components of $(A', b') \in \mathbb{R}^{n^2+n}$ will be used to represent the transformation $H(x) = Ax + b$. Allowing a slight abuse of notation, we shall identify sets of transformations with the sets of their corresponding parameters. For example, $K_{\bar{x}}^{V}$ will also denote the set of parameters that correspond to all transformations t for which V_t includes the point $\bar{x}$.

The families of Euclidean transformations and of similarity transformations of $\mathbb{R}^n$ can always be represented in $\mathbb{R}^{n^2+n}$, but can also be represented in lower dimensional parameter spaces. This will enable us to obtain, in some cases, better bounds on the VC-dimension of classes of images of an object under such transformations.

The following lemma demonstrates a useful property of Affine transformations, namely, that simple objects, V, give rise to simple subsets $K_{\bar{x}}^{V}$ of the parameter space.

Lemma 2. *If V is a semi-algebraic subset of $\mathbb{R}^n$ of degree (k, m) then, for every $\bar{x} \in \mathbb{R}^n$, the set of transformation parameters $K_{\bar{x}}^{V}$ is a semi-algebraic set of degree (k, m) (in the parameter space of affine transformations $\mathbb{R}^{n^2+n}$).*

Proof: A sufficient and necessary condition for a point $\bar{x}$ to be inside the transformed object is that the result of applying the inverse transformation on it, $\bar{y} = A'\bar{x} + b'$ will be in the original (nontransformed) semi-algebraic set V. The expression for $\bar{y}$ is a linear function of the parameters, and inserting it into the polynomials $\{f_j\}$ that specify V, induces polynomial sets of equal degree in the parameter space $\{\bar{t} \mid f_j(\bar{t}, \bar{x}) > 0\}$. The union and intersection operations on the polynomial sets of $\mathbb{R}^n$ transform into similar operations on the polynomial sets of the parameter space, and therefore, the set $K_{\bar{x}}^{V}$ is also a semi-algebraic set of degree (k, m). □

Let $C_{\mathcal{A}}(V)$ denote the class of all objects obtained by transforming a semi-algebraic object V via affine transformations. Our next step is to set upper bounds on the VC-dimension of classes of the form $C_{\mathcal{A}}(V)$.[3] For this proof we shall employ (a small modification of) the classical result of Milnor (1964), regarding the number of connected components of polynomial sets.

Lemma 3 (A modification of Milnor (1964)). *Suppose $X \subset \mathbb{R}^n$ is specified by the polynomial inequalities $f_1(x) > 0, \ldots, f_l(x) > 0, f_{l+1}(x) \geq 0, \ldots, f_p(x) \geq 0$ with total degree $d = deg\, f_1 + \cdots + deg\, f_p$. Then $\Psi(X)$, the number of connected components of the set X, satisfies*

$$\Psi(X) \leq \frac{1}{2}(2+d)^n$$

Proof: The original theorem of Milnor provides this relation when all inequalities that specify X are weak (that is, all of them are of the type $f_j(x) \geq 0$). Consider now the sequence of sets $\{X_q\}_{q\in\mathbb{N}}$ specified by the weak inequalities $f_1(x) - 1/q \geq 0, \ldots, f_l(x) - 1/q \geq 0, f_{l+1}(x) \geq 0, \ldots, f_p(x) \geq 0$. Clearly, $X_1 \subset X_2 \subset \cdots \subset X$ and $X = \bigcup_{q\in\mathbb{N}} X_q$. Each of the sets X_q does satisfy Milnor's condition; therefore, the number of connected components cannot increase unboundedly with q.

The next claim, a purely point-set-topological statement, suffices to conclude the proof of the lemma. □

Claim 6. *Let $X_1 \subseteq X_2 \subseteq \cdots \subseteq X_q \subseteq \cdots$ be a sequence of sets in some topological space and let X denote their union, $\bigcup_{q\in\mathbb{N}} X_q$. Let $\sharp A$ denote the number of connected components of a set A, then*

$$\sharp X \leq \sup_{q\in\mathbb{N}} \sharp X_q$$

Proof (of the claim): First, note that if two of points, x, y, are in the same connected component in some X_q, then they share the same connected component in X as well. Assume $\sup_{q\in\mathbb{N}} \sharp X_q$ is some finite number, d (otherwise there is nothing to prove). Let $x_1, x_2, \ldots, x_{d+1}$ be points in X, and let $\hat{q}$ be such that they are all members of $X_{\hat{q}}$. As $\sharp X_{\hat{q}} \leq d$, there are some $i \neq j$ such that x_i and x_j share the same connected component of $X_{\hat{q}}$. It follows that x_i and x_j are in one connected component of X. □

3.3. *The VC-dimension of affine transformed objects*

Theorem 2. *For every semi-algebraic set V of degree (k, m) in $\mathbb{R}^n$,*

$$VCdim(C_{\mathcal{A}}(V)) = O(n^2 \log nkm)$$

Proof: Let $S = \{x_1, \ldots, x_N\}$ be a subset of $\mathbb{R}^n$ that is shattered by the class of images $C_{\mathcal{A}}(V)$. Let us focus on the parameter space of transformations $\mathbb{R}^{n^2+n}$. The union of boundaries $B_S = \bigcup_{i=1}^{i=N} \partial K_{x_i}^V$ of the semi-algebraic open parameter sets $\{K_{x_i}^V\}$ divides the parameter space into connected components. We shall show that the number of such components bounds the number of nonempty sets of the form $W(A, S)$, and therefore, the exponent of the size of the shattered set S. We shall then apply Lemma 3 to bound the number of the connected components.

Claim 7. *For any pair A, A' of distinct subsets of S, any connected component of $\mathbb{R}^{n^2+n} \backslash B_S$ that has a nonempty intersection with $W(A, S)$ necessarily has an empty intersection with $W(A', S)$.*

Proof (of the claim): Let x_{i_0} be a point in $A \backslash A'$. By Claim 5, any $t \in W(A, S)$ is a member of $K_{x_{i_0}}^V$ while for any $t \in W(A', S)$, $t \notin K_{x_{i_0}}^V$. Consider the sets $\mathrm{int}(K_{x_{i_0}}^V)$ and $\mathrm{int}(\mathbb{R}^{n^2+n} \backslash K_{x_{i_0}}^V)$. They are disjoint open sets and they cover $(\mathbb{R}^{n^2+n} \backslash B_S)$. It follows that any subset of $(\mathbb{R}^{n^2+n} \backslash B_S)$ that meets both $W(A, S)$ and $W(A', S)$ is partitioned by the sets $\mathrm{int}(K_{x_{i_0}}^V)$ and $\mathrm{int}(\mathbb{R}^{n^2+n} \backslash K_{x_{i_0}}^V)$ into nonempty disjoint open subsets, and is therefore not connected. □

To calculate the number of connected components of $(\mathbb{R}^{n^2+n} \backslash B_S)$, recall that each of the parameter sets $K_{x_i}^V$ is specified by k-many polynomial sets of the form $\{\bar{t} \mid f_j(\bar{t}, x_i) > 0\}$ and note that at least one of the functions $f_{ij}(\bar{t}) = f_j(\bar{t}, x_i)$ vanishes on each point of the boundary of V_{x_i}.

Consider now the product function $G(\bar{t}) = \prod_{i,j} f_{ij}(t)$. Any connected component of $\mathbb{R}^{n^2+n} \backslash B_S$ is a union of one or more connected components of $\{\bar{t} : G(\bar{t}) > 0\}$ or of

$\{\bar{t} : G(\bar{t}) < 0\}$. $G(\bar{t})$ is a (kmN)-degree polynomial in $n^2 + n$ real variables (where N stands for the cardinality of the shattered set, S), and, by Lemma 3, the total number of connected components of its positive set $\{P \mid G(\bar{t}) > 0\}$ and its negative set (which is the pos-set of $-G$) is not higher than $2 \cdot \frac{1}{2}(2 + kmN)^{n^2+n}$. Consequently, our assumption that the set S is shattered by $C_{\mathcal{A}}(V)$ implies that $2^N \leq (2 + kmN)^{n^2+n}$. The theorem now follows by taking the logarithm of the last inequality and noting that it fails whenever $N \geq 2(n^2 + n)\log((n^2 + n)(2 + km)$, for sufficiently large values of k and m. ($N \leq a\log(bN)$ implies that, for $a \geq 2$ and $b \geq 1$, $N \leq 2a\log(ab)$, as otherwise, $a \geq \frac{N}{\log(bN)} > \frac{2a\log(ab)}{\log 2 + \log ab + \log\log(ab)} \geq \frac{2a\log(ab)}{2\log(ab)} \geq a$, which is a contradiction. In fact, for sufficiently large b, the condition $N \leq a\log(bN)$ implies that $N \leq (1+\gamma)a\log(ab)$ for any positive γ. This enhanced bound is not needed for this proof but will be useful below, for the asymptotic bounds.) □

In the Computer Vision and Robotics context, one is usually interested only in two- or three-dimensional object spaces. For these cases, it is easy to see that subgroups of affine transformations, such as Translation, Euclidean and Similarity transformations can be represented in a parameter space smaller than $\mathbb{R}^{n^2+n}$, implying that the corresponding classes of images $C_{\mathcal{T}}(V)$, $C_{\mathcal{E}}(V)$, $C_{\mathcal{S}}(V)$, have lower VC-dimensions.

In the two-dimensional case, for example, the translation transformation may be represented only by the two components of the (inverse) translation vector b'. To represent the Euclidean transformation we need also to specify the rotation matrix A'. To keep the transformation linear in the parameters, which is essential for the derivation, we use a four-dimensional space $\bar{t} = \{t_1 = a'_{11}(= a'_{22}), t_2 = a'_{21}(= -a'_{12}), t_3 = b'_1, t_4 = b'_2\}$ with a constraint $t_1^2 + t_2^2 = 1$ kept in mind. For Similarity transformations, which also allow uniform scale change, five-dimensional parameter space is used, with four parameters identical to the Euclidean transformation parameters, and the fifth induced by the weaker constraint $t_1^2 + t_2^2 = t_5^2$. The most general affine transformation, is represented, as described above by the $(n^2 + 2 = 6)$-dimensional parameter space.

Now, let $B_{\mathcal{T}}(V)$, $B_{\mathcal{E}}(V)$, $B_{\mathcal{S}}(V)$, and $B_{\mathcal{A}}(V)$ be upper bounds on the VC-dimension of the corresponding classes of images when the complexity of the object V, represented by the product km, increases to infinity. Applying the parameterizations introduced above, with the method presented in the proof of Theorem 2 will yield the required asymptotic bounds. The additional degree-2 polynomial constraint on the parameters, which occurs in some of the cases, is easily incorporated by expressing it as two weak inequality constraints.

Corollary 2.

$$B_{\mathcal{T}}(V) = 2\log(km)$$
$$B_{\mathcal{E}}(V) = 4\log(km)$$
$$B_{\mathcal{S}}(V) = 5\log(km)$$
$$B_{\mathcal{A}}(V) = 6\log(km)$$

3.4. An upper bound on the VC-dimension of semi-algebraic sets

Using the same technique, it is possible to prove the following upper bound on the class of semi-algebraic subsets.

Claim 8. *The VC-dimension of the class of semi-algebraic subsets of* $\mathbb{R}^n$ *of degree* (k, m), *is at most* $2k\binom{m+n}{m}\log(k(k+1)\binom{m+n}{m})$.

Proof (sketch): Assuming that a set of N points is shattered, we look at the induced partition of a parameter space. The dimension of this parameter space, representing the coefficients of the k polynomials of degree m is $k\binom{m+n}{m}$. This space is partitioned here by kN linear polynomials, implying that the inequality $2^N \leq (2+kN)^{k\binom{m+n}{m}}$, must be satisfied, which, in turn, implies the claim. □

Note that this upper bound differs only by a logarithmic factor from the lower bound derived above.

4. Learnability of geometric objects under the uniform distribution

The Vapnik-Chervonenkis dimension of a class can be viewed as a measure of the (information-theoretic) hardness of its distribution-free learnability. That is, its learnability in a setting where the underlying distribution—the distribution according to which examples are provided and relative to which the accuracy of hypotheses is defined—is unknown to the student and, furthermore, his performance is analyzed in a worst-case setting. Consequently, the Vapnik-Chervonenkis dimension of a class may be readily used to provide upper bounds on the difficulty of learning the class, in less demanding models of learnability, e.g., when the underlying distribution is known to the student, or is chosen from a limited family of candidate distributions.

The relevance of the VC-dimension to hardness (i.e., lower bound) results is not as clear as its applicability to upper bounds. ‘Real-life’ settings are usually much more restricted than what the distribution-free model may reflect. It may very well be the case that a class, whose distribution-free learnability is hard, is easily learnable once the underlying distribution is chosen from among a family of ‘realistic’ or ‘relevant’ distributions.

In this section we wish to show that, in the realm of classes of geometric objects in a Euclidean space, this is not the case. That is, the lower bounds on the difficulty of learnability of such classes, as provided by the Vapnik-Chervonenkis dimension, hold even in the restricted model of one fixed underlying distribution—the uniform probability measure.

4.1. Learning in metric spaces and covering numbers

The difficulty of learning a concept class under a fixed distribution is best analyzed in the context of metric spaces. Let us begin our discussion by introducing some basic concepts from the theory of metric spaces.

Definition 5.

- A metric space (E, d) is a set E associated with a distance function $d(\cdot, \cdot)$ between its members, satisfying, for every $x, y, z \in E$, three conditions:
  1. $d(x, x) = 0$,
  2. $d(x, y) = d(y, x) \geq 0$,
  3. $d(x, z) \leq d(x, y) + d(y, z)$.

  (We use a weaker definition of a metric space, which is often referred to also as "pseudo-metric".)
- Let (E, d) be a metric space, let A be a subset of E and $\epsilon > 0$.
  1. $B \subseteq E$ is an ϵ-*cover* for A if for every $a \in A$ there exists some $b \in B$ such that $d(a, b) < \epsilon$.
  2. $\mathcal{N}_d(\epsilon, A)$ is the minimal cardinality of an ϵ-cover for A. (If there is no such finite cover then it is defined to be ∞.) $\mathcal{N}_d(\epsilon, A)$ is sometimes referred to as the ϵ-*covering number* of A.
  3. $A \subseteq E$ is ϵ-*separated* if, for any distinct $a, b \in A$, $d(a, b) > \epsilon$.
  4. $\mathcal{M}_d(\epsilon, A)$ is the maximal size of an ϵ-separated subset of A. $\mathcal{M}_d(\epsilon, A)$ is sometimes referred to as the ϵ-*capacity* or the ϵ-*entropy* of A.

The ϵ-covering numbers and ϵ-capacities are closely related. The following inequalities can be verified (see, e.g., (Kolmogorov & Tichomirov, 1961)):

Claim 9. *For every metric space* (E, d), $A \subseteq E$ *and* $\epsilon > 0$

$$\mathcal{M}_d(2\epsilon, A) \leq \mathcal{N}_d(\epsilon, A) \leq \mathcal{M}_d(\epsilon, A).$$

Given a probability space $(X, \mathcal{O}, P)$, a natural pseudo-metric, d_P, is induced over the set of measurable sets: For every $a, b \in \mathcal{O}$, $d_P(a, b) = P(a \Delta b)$.[4] We shall use $\mathcal{N}_P(\epsilon, A)$ to denote $\mathcal{N}_{d_P}(\epsilon, A)$ (and similarly for $\mathcal{M}$). Note that a and b are subsets in the space X but are points in the induced metric space. Similarly, a set of X subsets (e.g., a concept class) is a subset of the induced metric space.

Benedek and Itai (1988) investigate learnability with respect to a fixed distributions. The results of Section 4 there imply the following bounds:

Theorem 3 (Benedek & Itai (1988)). *Let* $l_C^P(\epsilon, \delta)$ *denote the number of random examples needed for* (ϵ, δ)*—learning of a class* C *with respect to a probability distribution* P*. For any probability space* $(X, \mathcal{O}, P)$*, any concept class* $C \subseteq \mathcal{O}$ *and any positive* ϵ *and* δ,

$$l_C^P(\epsilon, \delta) \geq \log(1 - \delta) + \log \mathcal{M}_P(2\epsilon, C),$$
$$l_C^P(\epsilon, \delta) \leq \frac{54}{\rho}(\ln 1/\rho + \ln \mathcal{N}_P(\rho/2, C)),$$

where $\rho = \min(\epsilon, \delta)$.

The above result reduces the assessment of the information-complexity of the learnability of a class, under a given distribution P, to finding its covering numbers relative to the (pseudo-) metric d_P. A fundamental result of Dudley now brings us back to the VC-dimension.

Theorem 4 (Dudley (1984)). *For any measurable space $(X, \mathcal{O})$ and any class $B \subseteq \mathcal{O}$,*

$$\inf\{w : \textit{for any finite subset } A, \|\{b \cap A : b \in B\}\| = O(\|A\|^w)\}$$
$$= \inf\left\{w : \sup_P\{\mathcal{N}_P(\epsilon, B)\} = O(\epsilon^{-w})\right\},$$

where $\sup_P$ *denotes the supremum over all probability distributions over* $(X, \mathcal{O})$.

The connection of Dudley's theorem to the VC-dimension goes through Sauer's lemma (Sauer, 1972). The lemma implies that, for every class B of sets, $\|\{b \cap A : b \in B\}\| = O(\|A\|^{VCdim(B)})$. Furthermore, there exist classes for which $VCdim(B)$ is the minimal exponent satisfying this equation (for example, this is the case when X is infinite and B is the class of all subsets of X of some fixed finite cardinality d). Consequently, the theorem implies that, for a class B having dimension d, $(1/\epsilon)^d$ is an upper bound on its ϵ-covering numbers relative to any distribution, and that there exist classes and distributions that give rise to covering numbers that are arbitrarily close to this function of ϵ.

Wishing to establish lower bounds on the difficulty of learning under some fixed distribution, we shall have to show that for the classes we care about, the covering numbers relative to this distribution approach the upper bound stated by the theorem.

4.2. The covering numbers may approach the VC-dimension limit even for a uniform distribution

One should note that if a class B has a finite VC-dimension d, then, there always exists a probability distribution, P, such that for $\epsilon = 1/d$, $\mathcal{N}_P(\epsilon, B) = 2^d$ (which is of the same order of magnitude as Dudley's upper bound). To establish this claim just pick a set of size d that is shattered by B and let P be the uniform distribution over this set.

The interesting questions are to show that such 'maximum capacity' behavior can be attained for arbitrarily small ϵ's and, maybe more important, to demonstrate such behavior relative to natural distributions.

The first probability measure that comes to one's mind, when considering a bounded region of a Euclidean space, is the uniform distribution over that region. Clearly, many classes in such a space may have much lower capacities than the bounds derived from their VC-dimension. For example, classes of finite (and co-finite) sets may have arbitrarily large VC-dimension, yet their ϵ-capacity (with respect to the pseudo-metric induced by the uniform distribution) is just 1, for any $\epsilon > 0$. We wish to set forth the thesis that, as long as the classes under consideration are natural classes of geometric objects, the bounds derived from the combinatorial considerations are indeed matched by the ϵ-capacities under the uniform distribution. Given the vagueness of the notion of a 'natural class of geometric objects' we settle for demonstrating the above claim through various examples of such classes.

Let $C(n, d)$ denote $\sum_{i=0}^{d}\binom{n}{i}$. Sauer's lemma (Sauer, 1972) states that, for any class B having VC-dimension d, $C(n, d)$ is an upper bound on the cardinality of $\Pi_B(A)(=\{b\cap A : b \in B\})$ for sets A of cardinality n, and is the minimal such upper bound. Note that $d = \inf\{w : C(n, d) = O(n^w)\}$. It follows that Dudley's bound is established once one shows that, for any class B having dimension d and for any $\epsilon > 0$, $\mathcal{M}(\epsilon, B) \geq C(1/\epsilon, d)$.

To prove his theorem, Dudley picks, for every n, a set A of cardinality n for which this bound is attained. He then constructs a probability distribution P that concentrates on these sets and gives each member of such a set equal probability weight (of about $1/|A|$)). Under such a distribution, $\Pi_B(A)$ is ϵ-separated and, therefore, $\mathcal{M}_P(1/|A|, B) = C(|A|, d)$, meeting the upper bound of the theorem.

When one wishes to apply this idea to the uniform distribution, the sets A must be chosen more carefully. Having no control over the distribution, our tool for giving the needed weights to members of A is to make sure that, with every point x in such a set, there is an attached neighborhood U_x such that the members of B that participate in defining $\Pi_B(A)$ do not divide these neighborhoods (i.e., there exists $B' \subseteq B$ such that $\Pi_{B'}(A) = \Pi_B(A)$ and, for every $b \in B'$ and $x \in A$, either $U_x \subseteq b$ or $U_x \cap b = \emptyset$). The probabilities of these neighborhood sets, under the uniform distribution, will now play the role that $P(x)$—the probability weight of a singleton $\{x\}$—plays in Dudley's construction of P.

Let us demonstrate the theme discussed above by applying it to a couple of example classes of subsets of the unit square. Let US denote $[0, 1] \times [0, 1]$ and let U denote the uniform distribution over it.

Claim 10. *For every* $0 < \epsilon < 1$, $\mathcal{M}_U(\epsilon, SA^2_{(k,2)}) \geq C(\pi/(4\epsilon), k)$.

Proof: For any $s, t \in \mathbb{R}$ and $\rho > 0$, let $B((s, t), \rho)$ denote the circle $\{(x, y) : (x - s)^2 + (y-t)^2 < \rho^2\}$. For every $0 < \epsilon' = \frac{4}{\pi\epsilon} < 1$, let $\mathcal{B}_{\epsilon'}$ denote $\{B((i\sqrt{\epsilon'} - \frac{\sqrt{\epsilon'}}{2}, j\sqrt{\epsilon'} - \frac{\sqrt{\epsilon'}}{2}), \sqrt{\epsilon'}/2) : i, j \in \{1/\epsilon', 2/\epsilon', \ldots, 1/\sqrt{\epsilon'}\}\}$. $\mathcal{B}_{\epsilon'}$ is a set of cardinality $\frac{1}{\epsilon'} = \frac{\pi}{4\epsilon}$ of disjoint circles, each having area $\frac{\pi}{4}\epsilon' = \epsilon$. (See figure 1 (left).) Every union of $\leq k$ of such circles is a member of $SA^2_{(k,4)}$, every pair of distinct such unions is at least ϵ apart (in the metric d_U), and there are $C(\pi/(4\epsilon), k)$ such distinct unions. □

Repeating the proof of the previous claim, replacing the sets of circles $\mathcal{B}_\epsilon$ by the sets of rectangles $\mathcal{R}_\epsilon \stackrel{\text{def}}{=} \{\{(x, y) : i\sqrt{\epsilon} < x < (i + 1)\sqrt{\epsilon}\,;\ j\sqrt{\epsilon} < y < (j + 1)\sqrt{\epsilon}\} : i, j = 0, 1, 2, \ldots, \frac{1}{\sqrt{\epsilon}} - 1\}$, one readily gets:

Claim 11. *For every* $0 < \epsilon < 1$, $\mathcal{M}_U(\epsilon, SA^2_{(4k,1)}) \geq C(1/\epsilon, k)$.

Using a somewhat more complicated construction this can be improved to:

Claim 12. *For every* $0 < \epsilon < 1$, $\mathcal{M}_U(\epsilon, SA^2_{(k,1)}) \geq C(1/2\epsilon, k)$.

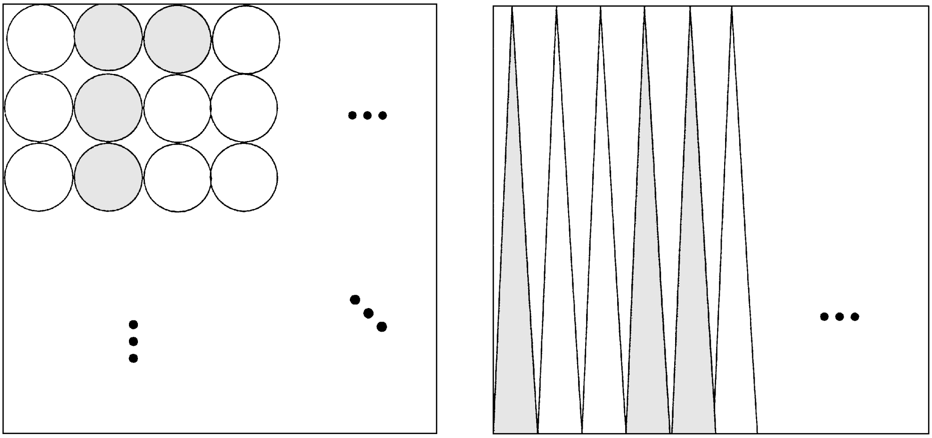

Figure 1. Constructing the metric space required for proving claims 10 and 12. In the left figure, a concept is a union of k circles (say the 4 darkened ones) and corresponds to a point in the metric space. Similarly, in the right figure a concept is a union of k triangles.

Proof: For any $k \in \mathbb{N}$ and $0 < \epsilon < 1$, define a set of functions

$$\Theta_k^\epsilon = \{\theta : \{2i\epsilon : 0 \leq i \leq 1/2\epsilon\} \to \{0, 1\} : \theta \text{ assumes the value 1 at most } k \text{ times}\}.$$

For any such function θ, let p_θ be the polygonal object obtained by successively connecting, by linear segments, the points $(2i\epsilon, \theta(2i\epsilon))$. Define a class

$$\mathcal{P}_{(\epsilon,k)} \stackrel{\text{def}}{=} \left\{\hat{p}_\theta : \theta \in \Theta_k^\epsilon\right\}.$$

Recall that $\hat{p}_\theta$ denotes $\{(x, y) : y < p_\theta(x)\}$, i.e., any member of $\mathcal{P}_{(\epsilon,k)}$ is determined by a polygonal object p_θ that is obtained by assigning values 0 or 1 to reals that are multiples of $1/2\epsilon$ and connecting the resulting points of the plan by linear segments. (See figure 1 (right).) Furthermore, all the p_θ defining $\mathcal{P}_{(\epsilon,k)}$ have at most k many 1's. Note that, for every ϵ and k there exist $C(1/2\epsilon, k)$—many such functions θ and each class $\mathcal{P}_{(\epsilon,k)}$ is ϵ-separated. □

So far, we have established lower bounds on the covering numbers of classes defined by combinations of linear segments or quadratic functions. Our next result pushes the ideas employed above to obtain similar bounds for classes whose members are sets defined by single polynomials. The idea is to replace the polygonal objects p_θ by interpolating polynomials.

Claim 13. *For every integer m and every $1/m < \epsilon < 1$, $\mathcal{M}_U(\epsilon, B_m^2) \geq C(1/3\epsilon, m)$. (Note that, as $B_m^2 \subset SA_{(1,m)}^2$, this bound applies also to $SA_{(1,m)}^2$.)*

Proof (sketch): Given any m and ϵ, we repeat the construction of the classes $\mathcal{P}_{(\epsilon,m)}$ but replace the polygonal objects p_θ by their interpolations via polynomials of degree m. □

4.3. Localization and identification under the uniform distribution

The significance of these results to our discussion stems from their comparison to the upper bounds of Section 3 on the VC-dimensions of corresponding classes of images. Consider a student who is trying to learn by viewing labeled examples drawn independently according to the uniform distribution on the unit square. We compare the information complexity (i.e., the number of labeled examples needed) of two types of tasks: the task of *identification*, in which all the student knows are the parameters (k, m) of the semi-algebraic class to which the target object belongs, and the task of *localization* in which he knows that the target is a transformed image of some given object V.

Combining the results of this section with the lower bound of Theorem 3, we get the following lower bounds on the information complexity of the task of *identification*: For $C = SA^2_{(k,2)}$ or $SA^2_{(k,1)}$ and, when $\epsilon > 1/k$, also for $C = SA^2_{(1,k)}$,

$$l^U_C(\epsilon, \delta) \geq k \log(1/\epsilon) + \log(1 - \delta).$$

On the other hand, for the task of *localization* of any object $V \in SA^2_{(k,m)}$ under the class of all affine transformations, the basic distribution-free upper bound of Blumer et al. (1989), combined with our results of Section 3 yields:

$$l^U_C(\epsilon, \delta) \leq \max\left\{\log(km)\frac{48}{\epsilon}\log\frac{13}{\epsilon}, \frac{4}{\epsilon}\log\frac{2}{\delta}\right\}.$$

5. The implications on the shape recognition problem

Consider a student who is trying to learn by viewing labeled examples drawn independently according to some distribution on the unit square. We compare the information complexity (i.e., the number of labeled examples needed) of two types of tasks: the task of *identification*, in which all the student knows are the parameters (k, m) of the semi-algebraic class to which the target object belongs, and the task of *recognition* in which he knows that the target is a transformed image of some given object V.

Combining the results of Section 2 with the lower bound of Blumer et al. (1989), implies that for some ‘unfortunate’ sampling distribution the student will need a large, number of samples, proportional to km^2, to identify a semi-algebraic object of degree (k, m) up to a prediction error of ϵ.

The results of the last section strengthen this conclusion and imply that this difficulty is not due to some ‘peculiar’ distribution but exists also when the sampling distribution is uniform. Our results, combined with Theorem 3, imply that the number of samples needed for *identification* is at least proportional to k. We conjecture here that this bound is not tight and that *identification* of semi-algebraic sets under the uniform distribution is as hard as *identification* under arbitrary distribution.

On the other hand, for the task of *localization* of any object $V \in SA^2_{(k,m)}$ under the classes of Affine, Euclidean and Similarity transformations, the basic distribution-free upper bound of Blumer et al. (1989), combined with our results of Section 3 implies that, asymptotically, the number of samples needed is not more than $\alpha \log(km)$ where α is 2, 4, 5 or 6, for Translation, Euclidean, Similarity and Affine transformation, respectively. This is, of course, much smaller than the number of samples required for *identification*.

Note, however, that the actual VC-dimension of such classes may be even lower than these logarithmic bounds. This is the case for, e.g., the class of translations of convex polygonal objects: Even though the number of sides may be arbitrarily large, the VC-dimension is only three (Pach & Woeginger, 1990). We conjecture here that the true VC-dimension of more general classes of images will be indeed logarithmic in the complexity of the object.

The ability to make the good hypotheses that achieve a small prediction error may be tied to the ability to localize the object within certain precision, by constructing the following mapping between the *localization* imprecision and the maximal prediction error associated with it: For any (known or bounded) sampling distribution, and any metric used to measure the *localization* error, one may consider all pairs of object instances whose *localization* error is above a certain threshold d_0, and find a lower bound on the associated distribution-weighted symmetric differences. A successful learning procedure that results in prediction error smaller than this bound implies that the error in the location of the object is smaller than d_0. Such a procedure was suggested and used in (Lindenbaum, 1995) to set bounds on the probability of achieving several recognition tasks for specific objects.

Acknowledgments

We thank Ilan Newman for insightful discussions concerning this work. The comments of the anonymous referees contributed a lot to the paper and we thank them for that.

Notes

1. In the field of computer vision, the term ‘recognition’, or ‘model-based recognition’ refers to *identification* from a library that is generated by transforming a finite number of base objects.
2. Dudley (1984) defines a collection of sets $\{A_1, A_2, \ldots, A_N\}$ to be *independent,* if for every function θ : $\{1, \ldots, N\} \to \{-1, 1\}$, the intersection, $\bigcap_{i=1}^{N} A_i^{\theta(i)}$, is nonempty. With this notation the above claim says that the class of images, $C_T(V)$, shatters a set $S \subseteq \mathbb{R}^n$ iff $\{K_{x_i}^V \mid x_i \in S\}$ is an independent set.
3. Such bounds can be deduced also from the results obtained by the independent work of Goldberg and Jerrum (1995), which was first published together with our work (Ben-David & Lindenbaum, 1993) and was discovered even before (Goldberg, 1992). Their work considers concept classes specified by logic formula or algorithm and characterizes a wide set of concept classes, associated with polynomial VC-dimension. Their technique is similar to ours except that they do not use the parameter space directly as we do and use more recent, and sometimes tighter, bounds derived by Waren, instead of the bounds by Milnor we use.
4. $a \,\Delta\, b$ denotes the symmetric difference between the two subsets, a and b, and is the set $(a \cap \bar{b}) \cup (b \cap \bar{a})$.

References

Ben-David, S., & Lindenbaum, M. (1993). Localization vs. identification of semi-algebraic sets. *Proc. of COLT93* (pp. 327–336). ACM Press.

Benedek, G.M., & Itai, A. (1988). Learnability by fixed distributions. *Proc. of COLT88* (pp. 80–90). *Theoretical Computer Science* (to appear).

Blumer, A., Ehrenfeucht, A., Haussler, D., & Warmuth, M.K. (1989). Learnability and the Vapnik-Chervonenkis dimension. *JACM*, *36*(4), 929–965.

Cover, T.M. (1965). Geometrical and statistical properties of systems of linear equations with applications in pattern recognition. *IEEE Trans. Electron. Comput.*, 326–334.

Dudley, R.M. (1984). A course on empirical processes. *Lecture Notes in Mathematics*, *1097*, 2–142.

Goldberg, P.W. (1992). PAC-learning geometrical figures, Ph.D. thesis, Department of Computer Science, University of Edinburgh, UK.

Goldberg, P., & Jerrum, M. (1995). Bounding the Vapnik-Chervonenkis dimension of concept classes parameterized by real numbers. *Machine Learning*, *18*, 131–148.

Keren, D., Cooper, D., & Subrahmonia, J. (1992). Describing complicated objects by implicit polynomials. *IEEE Transaction on Patt. Anal. Mach. Intell.* (accepted for publication).

Kolmogorov, A.N., & Tichomirov, V.M. (1961). ϵ-entropy and ϵ-capacity of sets in functional spaces. *Amer. Amth. Soc. Translations (Ser. 2)*, *17*, 277–364.

Lindenbaum, M. (1995). Bounds on shape recognition performance. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *PAMI-17*(7), 666–680.

Milnor, J. (1964). On the Betti numbers of real varieties. *Proc. Amer. Math. Soc.*, *15*, 275–280.

Pach, J., & Woeginger, G. (1990). Some new bounds on ϵ-nets. *Proc. 6th Int. Symp. on Comp. Geometry* (pp. 10–15).

Sauer, N. (1972). On the density of family of sets. *Journal of Combinatorial Theory (Series A)*, *13*, 145–147.

Terzopoulos, D., Platt, J., Barr A., & Fleischer, K. (1987). Elastically deformable models. *ACM Computer Graphics*, *21*, 205–214.

Received January 24, 1996
Accepted December 30, 1997
Final Manuscript May 1, 1998