IMAGE ESTIMATORS BASED ON MARKED BINS

Amparo Baíllo and Antonio Cuevas*

Abstract

The problem of approximating an "image" *S* in *R*^d from a random sample of points is considered. If *S* is included in a grid of square bins, a plausible estimator of *S* is defined as the union of the "marked" bins (those containing a sample point). We obtain convergence rates for this estimator and study its performance in the approximation of the border of *S*. The estimation of "digitalized" images is also addressed by using a Vapnik-Chervonenkis approach. The practical aspects of implementation are discussed in some detail, including some technical improvements on the estimator, whose performance is checked through simulated as well as real data examples.

Keywords: Set estimation; histogram; rates of convergence; image analysis.

*Baíllo, Statistics Department, University Carlos III de Madrid, C/ Madrid 126, 28903 Getafe (Madrid), e-mail: abaillo@est-econ.uc3m.es; Cuevas, Mathematics Department, University Autónoma de Madrid, 28049 Madrid, e-mail: antonio.cuevas@uam.es. Research partially supported by Spanish grant MTM2004-00098.

Image estimators based on marked bins[*]

Amparo Baíllo

Universidad Carlos III de Madrid

28903 Getafe (Madrid), Spain

Antonio Cuevas

Universidad Autónoma de Madrid

28049 Madrid, Spain

Abstract

The problem of approximating an "image" $S \subset \mathbb{R}^d$ from a random sample of points is considered. If S is included in a grid of square bins, a plausible estimator of S is defined as the union of the "marked" bins (those containing a sample point). We obtain convergence rates for this estimator and study its performance in the approximation of the border of S. The estimation of "digitalized" images is also addressed by using a Vapnik-Chervonenkis approach. The practical aspects of implementation are discussed in some detail, including some technical improvements on the estimator, whose performance is checked through simulated as well as real data examples.

Keywords: Set estimation; histogram; rates of convergence; image analysis.
AMS Classification 2000: 62G05 (62G07, 62H35)

1 Introduction

Given a dot pattern $X_1, \ldots, X_n$ on the real plane $\mathbb{R}^2$ or, more generally, in $\mathbb{R}^d$, one interesting problem is to extract the "essential" external shape or border of the pattern. When the pattern is a sample that has been drawn from a compactly supported probability distribution F with support $S \subseteq \mathbb{R}^d$, this external shape approximates that of the original set S. Thus, if S were a black image defined over a white background, the border of the sample would provide the boundary for a binary image estimator. This easy-to-state problem has application in several areas of research apart from image analysis such as astronomy, quality control, biomedical imaging, wildlife radiotracking, forestry or geophysical statistics.

The estimation of the support S has been extensively studied in the literature (see, e.g., Cuevas and Rodríguez-Casal 2003 for a review). When S is convex the natural estimator is the convex

---

hull of the sample (see Rényi and Sulanke 1963, 1964; Schneider 1988; Korostelev and Tsybakov 1993,...). When this restrictive assumption fails (as is most likely in image analysis) there is no unique, totally satisfactory and computationally convenient way of recovering S, and least of all as the dimension d increases. Devroye and Wise (1980) proposed a simple estimator of S: a union of balls centered at the sample points (see also Baíllo, Cuevas and Justel 2000). Although this estimator has obvious computational advantages and does not impose any geometrical restriction on S, it produces a "rough" approximation to the original image. Hall, Park and Turlach (2002) and Rodríguez-Casal (2003) have considered a generalization of the convexity assumption: the so-called r-convexity, under which a ball of radius $r > 0$ is allowed to roll along the exterior of S.

In the present work we will mainly focus on recovering the binary image S through the analysis of a sample of points randomly drawn from S. In more precise terms, we want to approximate S by an estimator S_n based on this sample. The performance of this estimator will be evaluated by checking that the "distance in measure" between S_n and S (i.e., the measure of the non-common part in $S_n \cup S$) is small. We are also interested in estimating the "shape" of S, as given by its boundary ∂S. We will use a computationally efficient set-estimation method studied by Ray Chaudhuri, Chaudhuri and Parui (1997) and Ray Chaudhuri *et al.* (1999): a histogram type estimator of S whose bins are marked as 1 (black) or 0 (white) depending on whether they contain sample points or not (see also a related approach in Klemelä 2004). The main theoretical results are included in Section 2, which is devoted to study the asymptotic properties of this estimator (called the H-shape). Section 3 focuses on the more general problem of identifying homogeneous regions in a colour or gray-scale image from a sample of (colour-marked) random points $(X_i, Z(X_i))$ extracted from the image. In Section 4 we construct a more refined piecewise constant image estimator, the ASH-shape, a gray-scale image approximating the original black-and-white one, and it is applied to a set of real data. Finally Section 5 displays the performance of the H- and ASH-shape over some simulated samples.

2   Estimation of a Binary Image. The H-Shape

Let $X_1, \ldots, X_n$ be a random sample of i.i.d. observations from a distribution F defined on $\mathbb{R}^d$, which is absolutely continuous with respect to the Lebesgue measure Leb and whose support is a compact set S. Assume without loss of generality that $S \subseteq [0, 1]^d$, the unit cube in $\mathbb{R}^d$. Let f denote the density of F. Consider a sequence of cubic partitions $\mathcal{P}_n = \{A_{nj}, j = 1, 2, \ldots\}$ of $[0, 1]^d$, $n = 1, 2, \ldots$ Let each element A_{nj} of $\mathcal{P}_n$ be a bin given by $\prod_{i=1}^d [k_i h, (k_i + 1)h)$, where the k_i's are integers and $h = h(n)$ is the bin width. The point $(k_1 h, \ldots, k_d h)$ is the origin of the bin A_{nj} and the bin origin given by $k_1 = \ldots = k_d = 0$ will be considered the grid origin. In this case the grid origin coincides with the coordinate origin, but later on (in Section 4) we will consider the possibility of shifting the partitions $\mathcal{P}_n$.

Ray Chaudhuri *et al.* (1999) define the following estimator of the set S

$$S_n = \bigcup \{A_{nj} \in \mathcal{P}_n : \ X_i \in A_{nj} \text{ for some } i = 1, \ldots, n\}. \tag{1}$$

As the closure of S_n coincides with the support of the histogram density estimator defined on the partition $\mathcal{P}_n$, we will call S_n the *H-shape*. We denote by B_n the union of the elements A_{nj} of the partition $\mathcal{P}_n$ intersecting ∂S and by I_n the union of the A_{nj} contained in the interior of S.

As a measure of discrepancy between the set S and its estimator S_n, consider the following pseudodistance between them

$$d_\mu(S_n, S) := \mu(S_n \Delta S) = \mu(S_n \cap S^c) + \mu(S_n^c \cap S),$$

where μ is a positive σ-finite measure on $\mathbb{R}^d$ (typically $\mu = \text{Leb}$), whose restriction to S is absolutely continuous with respect to F.

It can be observed that S_n and S can be quite close in d_μ and still the shapes of both sets be quite different. This is why we are also concerned with the accurate estimation of the shape of S, measured in terms of the "visual distance" between the boundaries $d_H(\partial S_n, \partial S)$, where d_H denotes the well-known Hausdorff metric between compact sets, defined by

$$d_H(S, T) = \inf\{\epsilon > 0 : S \subset B(T, \epsilon), \ T \subset B(S, \epsilon)\},$$

where $B(T, \epsilon) := \bigcup_{x \in T} B(x, \epsilon)$, for any set T.

2.1 Shape restrictions

Throughout this work we will use some shape conditions on the set S. For ease of understanding we list them below before stating any result:

(B1) The number of bins in $\mathcal{P}_n$ intersecting the boundary of S, $\sharp\{A_{nj} \subset B_n\}$, is $O(h^{-(d-1)})$.

This can be seen as a "thinness" condition imposed on the boundary of S. In particular, it is satisfied if $\text{Leb}(B(\partial S, \epsilon)) = O(\epsilon)$ as $\epsilon \to 0$. To see this observe that, for n sufficiently large,

$$h^d \sharp\{A_{nj} \subset B_n\} = \text{Leb}(B_n) \leq \text{Leb}(B(\partial S, \sqrt{d}h)) = O(h).$$

(B2) The bounded set S is *partly expandable* if there exist constants $r > 0$ and $C(S) \geq 1$ such that

$$d_H(\partial S, \partial B(S, \epsilon)) \leq C(S)\epsilon, \qquad 0 \leq \epsilon < r.$$

(B3) The bounded set $S \subset \mathbb{R}^d$ is *standard* (with respect to a measure μ) if for every $\lambda > 0$ there exists $\delta \in (0, 1)$ such that

$$\mu(B(x, \epsilon) \cap S) \geq \delta \text{Leb}(B(x, \epsilon)), \quad \text{for all } x \in S, \ 0 < \epsilon \leq \lambda.$$

3

Conditions (B2) and (B3) have already appeared in the context of set estimation (see Baíllo, Cuevas and Justel 2000, Cuevas and Rodríguez-Casal 2004). Basically, they have been used to exclude the presence in S of too sharp outward and inward peaks respectively. For example, if a ball of radius $r > 0$ rolls inside and outside S, then the set S satisfies (B1), (B2) and (B3) (see, e.g., Walther 1997 and Cuevas and Rodríguez-Casal 2004). We will also see that expandability is related to the following well-known shape restriction that has been extensively used, for instance, in Partial Differential Equations (see, e.g., Gilbarg and Trudinger 1983).

(B4) S satisfies a (uniform) outer *cone condition* if, for some constants $L \geq R > 0$ and all $x \in \partial S$, there exists $y \in \partial B(x, L)$ such that $\Psi_x \cap S = \{x\}$, where

$$\Psi_x := \{(1 - \delta)x + \delta z, \delta \in [0, 1], z \in B(y, R) \cap \partial B(x, L)\}.$$

S is said to fulfil a (uniform) interior cone condition if S^c satisfies the exterior cone condition.

Under (B4) the set S is a sort of "watchtower" where every point x in ∂S can "observe" a cone Ψ_x (outside S) of fixed width and height. If S is convex then the exterior cone condition is satisfied for any $L = R > 0$. If S is star-shaped and its kernel contains a ball of positive radius, then (B4) also holds. It is easy to see that if S satisfies the (uniform) interior cone condition then it fulfills the standardness condition (B3), with respect to the Lebesgue measure.

Note also that the cone property (B4) is a sufficient condition for the expandability assumption (B2). To see this we just have to check that, if a set S satisfies condition (B4), then there exists some $r > 0$ such that

$$d_H(\partial S, \partial B(S, \epsilon)) = \max(\max_{x \in \partial B(S, \epsilon)} d(x, \partial S), \max_{x \in \partial S} d(x, \partial B(S, \epsilon)) \leq C(S)\epsilon, \quad 0 < \epsilon < r.$$

Indeed, it is clear that $d(x, \partial S) \leq \epsilon$ for all $x \in \partial B(S, \epsilon)$. On the other hand, if $x \in \partial S$ and $0 < \epsilon < R$, using notation in condition (B4), there exists a point p along the segment joining x and y and at distance $a \leq 2\epsilon\sqrt{1 + (L/R)^2}$ from x, such that $B(p, \epsilon) \subset \text{int}(\Psi_x) \subset S^c$ and thus $p \notin B(S, \epsilon)$. This implies that $d(x, \partial B(S, \epsilon)) \leq a$ for all $x \in \partial S$.

2.2   Rates of convergence: the general case

For dimension $d = 2$, let λ_n denote the Lebesgue measure of the minimum rectangle, with sides parallel to the coordinate axes, containing all the sample points X_i. Ray Chaudhuri *et al.* (1999) prove that, if F is the uniform distribution on S, $\mu = \text{Leb}$, and $h = n^{-\delta}\sqrt{\lambda_n}$, where $0 < \delta < 1/2$, then $d_F(S_n, S) = O_P(n^{-\delta})$ as $n \to \infty$. In Theorem 1 we obtain rates of convergence of $d_\mu(S_n, S)$ to 0. Concretely, in (b) the above mentioned result of Ray Chaudhuri *et al.* (1999) is generalized to any dimension d and distributions F other than the uniform. We will see that the convergence rates are faster for those measures F decaying faster (larger γ) to zero as they approach the boundary

of S. This result is reminiscent of the minus-sampling solution to the well-known problem of edge-effects in spatial statistics (see Stoyan, Kendall and Mecke 1995, pg. 133). Observe that the aim in both contexts is to reduce the bias term (in our case $\mu(S_n \cap S^c)$).

Part (c) in Theorem 1 focuses on estimation of "digital" images, defined as a union of square pixels with sides parallel to the coordinate axes. The structure of the H-shape estimator looks especially well-suited for approximating this kind of images. In fact, we will see that in this case, stronger approximation results can be obtained. The mathematical tools in Theorem 1 (c) are different from those in (a) and (b) and rely on the use of Vapnik-Chervonenkis (VC) theory for empirical processes. Let us briefly recall some key concepts in this theory to be used below (see Devroye, Györfi and Lugosi (1996), ch. 12 and 13 for a review of the VC theory).

Given a class $\mathcal{A}$ of sets in $\mathbb{R}^d$ and a finite set $\{x_1, \ldots, x_n\} \subset \mathbb{R}^d$, we define the *shatter coefficient associated with* $\mathcal{X}_n = \{x_1, \ldots, x_n\}$ (denoted by $s(\mathcal{A}, \mathcal{X}_n)$) as the maximum number of different subsets of $\mathcal{X}_n$ which can be obtained by intersecting $\mathcal{X}_n$ with all the members of the class $\mathcal{A}$. The *shatter coefficient of the class* $\mathcal{A}$, denoted by $s(\mathcal{A}, n)$, is the maximum of $s(\mathcal{A}, \mathcal{X}_n)$ when $\mathcal{X}_n$ ranges over all the possible subsets of $\mathbb{R}^d$ with cardinality n. Of course, $s(\mathcal{A}, n) \leq 2^n$. The *Vapnik-Chervonenkis dimension* of $\mathcal{A}$, $V_{\mathcal{A}}$, is the largest (possibly infinite) value of n such that $s(\mathcal{A}, n) = 2^n$. The value of $V_{\mathcal{A}}$ quantifies in some sense the complexity of the class $\mathcal{A}$. The central core of Vapnik-Chervonenkis theory is an exponential bound (useful when $V_{\mathcal{A}}$ is finite), which will be the basic tool for the proof of Theorem 1 (c).

Theorem 1:

a) *If $h \to 0$ and $nh^d \to \infty$ as $n \to \infty$, then $d_\mu(S_n, S) \to 0$ a.s.*

b) *Assume S satisfies condition (B1). Assume also that f is bounded and there exists $\gamma \geq 0$ such that*

   (P_f) *There exists a constant $m > 0$ such that $F(B(x, \epsilon)) \geq m\epsilon^{\gamma+d}$, for every ball $B(x, \epsilon) \subset int(S)$ with $\epsilon > 0$.*

   and

   (P_μ) *There exist constants $M > 0$ and $\epsilon_0 > 0$ such that, for all $0 < \epsilon < \epsilon_0$, if $x \in \partial S$ then $\mu(B(x, \epsilon)) \leq M\epsilon^{\gamma+d}$.*

   If $h = n^{-\delta}$ for some $0 < \delta < 1/(d + \gamma)$, then $d_\mu(S_n, S) = O_P(h^{\gamma+1})$.

c) *Assume that F is uniform on S and that $S \in \mathcal{A}$, where $\mathcal{A}$ denotes a collection of sets which are unions of bins of the type $\prod_{i=1}^d [k_i/K, (k_i+1)/K)$ and the k_i's are integers ranging from 0 to $K - 1$. Take $\mu =$ Leb. Let*

$$h_n = 1/(c_n K), \tag{2}$$

5

with $\{c_n\}$ a nondecreasing sequence of positive integers. If $c_n = O(n^\delta)$ for some $0 \le \delta < 1/d$, then $d_{\mathrm{Leb}}(S_n, S) = O_P(\epsilon_n)$ for any $\epsilon_n \to 0$ such that $\epsilon_n^{-1} = o(n^{(\delta d - 1)/2})$.

Proof: a) Recall that

$$d_\mu(S_n, S) = \mu(S_n^c \cap S) + \mu(S_n \cap S^c). \tag{3}$$

With probability one $\mu(S_n \cap S^c) \le \mu(B(S, \sqrt{d}h) \cap S^c)$, which converges to zero as $n \to \infty$ by the Lebesgue dominated convergence theorem. On the other hand, $F(S_n^c \cap S) \le J_n$, where $J_n := \int |f_n - f|$ is the L^1 error of the histogram estimator, $f_n(x) := (nh_n^d)^{-1} \sum_{i=1}^n \mathbb{I}\{X_i \in A_{nj}\}$, and $\mathbb{I}_A$ or $\mathbb{I}(A)$ denotes the indicator function on the set A. Under the assumptions in (a) the error J_n converges to zero a.s. (see, e.g., Devroye and Györfi 1985). Then, as the restriction of μ to S is absolutely continuous with respect to F, we have that $\mu(S_n^c \cap S)$ converges to zero a.s. too.

b) The first term in the right-hand side of (3) may be split in the following way

$$\mu(S_n^c \cap S) = \mu(S_n^c \cap I_n) + \mu(S_n^c \cap S \cap B_n). \tag{4}$$

On the one hand, for the first term on the right-hand side of (4) we have

$$
\begin{aligned}
E[\mu(S_n^c \cap I_n)] \quad &= E\left(\sum_{A_{nj} \subset I_n} \mu(A_{nj})\mathbb{I}\{X_i \notin A_{nj}, i = 1, \ldots, n\} \right) \\
&\le C_1 h^d \sum_{A_{nj} \subset I_n} P\{X_i \notin A_{nj}, i = 1, \ldots, n\} \le C_2 \exp(-C_3 nh^{d+\gamma}),
\end{aligned} \tag{5}
$$

for some constants $C_i > 0$. The first inequality above follows from the fact that f is bounded and the restriction of μ to S is absolutely continuous with respect to F. In the second one we have used $\#\{A_n \subset I_n\} = O(h^{-d})$ together with (P_f) and the inequality $(1 - x)^n \le \exp(-nx)$ for $x \in (0, 1)$.

On the other hand, using condition (P_μ) we can bound $E[\mu(S_n^c \cap S \cap B_n)] + E[\mu(S_n \cap S^c)]$ by $C_4 h^{\gamma+d}\sharp\{A_{nj} \subset B_n\}$, which is $O(h^{\gamma+1})$ under (B1). Thus the distance $d_\mu(S_n, S)$ is also $O(h^{\gamma+1})$ since the exponential bound given in (5) for the other term goes to zero faster than any power of h.

c) Due to the choice $h_n = 1/(c_n K)$, we have $S_n \subseteq S$ and thus

$$
\begin{aligned}
\mathrm{Leb}(S_n \Delta S) \quad &= \mathrm{Leb}(S_n^c \cap S) = \mathrm{Leb}(S)\left[1 - \frac{\mathrm{Leb}(S_n \cap S)}{\mathrm{Leb}(S)} \right] \\
&= \mathrm{Leb}(S)\left[F_n(S_n) - F(S_n) \right] \le |F_n(S_n) - F(S_n)| \\
&\le \sup_{A \in \mathcal{A}_n} |F_n(A) - F(A)|,
\end{aligned} \tag{6}
$$

where $\mathcal{A}_n$ denotes the collection of all possible unions of bins $A_{nj} \in \mathcal{P}_n$. Notice that the shatter coefficient $s(\mathcal{A}_n, n)$ of this class is bounded by $\min(2^n, 2^{(Kc_n)^d})$, which is equal to $2^{(Kc_n)^d}$ for n large. Using VC inequality (see e.g., Devroye, Györfi and Lugosi, p. 197-198) on (6) yields

$$P\{\mathrm{Leb}(S_n \Delta S) > \epsilon\} \le 8 \cdot 2^{K^d} \exp[-c_n^d(-\log 2 + nc_n^{-d}\epsilon^2/8)].$$

This, together with the assumptions made on c_n and h_n gives the desired result.

$\square$

By imposing further restrictions on the class $\mathcal{A}$ of possible images the VC approach can be used to obtain faster rates. For example, convexity is a quite natural condition. However, as every set in $\mathcal{A}$ is a union of square pixels, the only convex sets in $\mathcal{A}$ are rectangles and this additional assumption would be too restrictive. We could instead consider the following weaker condition.

(B5) A set $S \subset \mathbb{R}^d$ is *marginally convex* if its intersection with every line parallel to any coordinate axis is at most one interval.

Let now $\mathcal{A}_n$ be the class of all possible marginally-convex unions of bins $A_{nj} \in \mathcal{P}_n$ and S_n be the smallest set in $\mathcal{A}_n$ containing all the sample points. If we choose h as in (2), then inequality (6) still holds. It can be shown that the VC dimension of $\mathcal{A}_n$ is $V_{\mathcal{A}_n} = 2(c_n K)^{d-1}$. The n-th shatter coefficient of this class can be bounded by $s(\mathcal{A}_n, n) \leq (n+1)^{V_{\mathcal{A}_n}} \leq (n+1)^{2(c_n K)^{d-1}}$ (see Devroye, Györfi and Lugosi (1996, p. 218). Using VC inequality now yields

$$P\{\mathrm{Leb}(S_n \Delta S) > \epsilon\} \quad \leq 8 \, \exp(2K^{d-1} c_n^{d-1} \log(n+1) - n\epsilon^2/8),$$

which implies that, for c_n chosen as in Theorem 1(c), we have $d_\mu(S_n, S) = O_P(\epsilon_n)$ for any $\epsilon_n \to 0$ such that $\epsilon_n^{-1} = o(n^{(\delta(d-1)-1)/2} \log n)$, a slightly faster rate than that obtained in the theorem. It is interesting, thus, to note that, under appropriate shape conditions, the reduction in the complexity of the class of possible estimators redounds to an increase in the convergence rate.

As a consequence of Theorem 1 (b) we can also consider the case of automatic (data driven) choice of the bin width h_n. Generalizing the idea in Ray Chaudhuri *et al.* (1999), we may take

$$h_n = \frac{\lambda_n^{1/d}}{n^\delta}, \tag{7}$$

where $0 < \delta < 1/d$ and λ_n is the Lebesgue measure of the smallest rectangle with sides parallel to the coordinates axes and containing all the sample points $X_1, \ldots, X_n$. This choice of the bin width provides the same convergence rates as those obtained in Theorem 1 (b) for the case $\gamma = 0$.

Corollary: *Let S satisfy condition (B1) and contain a ball of positive radius. Let also f be bounded away from zero on S. If the bin width is chosen as in (7) and $\mu = \mathrm{Leb}$, then $d_\mu(S_n, S) = O_P(n^{-\delta})$.*

Proof: The steps are the same as in Theorem 1 (b), that is, use expressions (3) and (4) and bound the resulting terms. This bounding is simplified by the fact that $\mu = \mathrm{Leb}$ (so P_μ is fulfilled and f does not have to be bounded), f is bounded away from zero on S (then P_f is also fulfilled) and h

is enclosed between two non-random sequences of order $n^{-\delta}$. To check this last statement remark that, as $S \subseteq [0,1]^d$, with probability one, λ_n is upper bounded by 1, so $h_n \leq n^{-\delta}$ a.s. On the other hand, there exists a ball B_r of radius $r > 0$ in the interior of set S. Thus, with probability one for n large, the smallest rectangle containing the whole sample contains also B_r and then $\lambda_n \geq \mathrm{Leb}(B_r)$ a.s.

$\square$

2.3   Shape estimation

The next result establishes rates of convergence to 0 of $d_H(\partial S, \partial S_n)$, the Hausdorff distance between the boundaries of S_n and S. Choosing an adequate bin width h the border of the H-shape will approximate the unknown border of S, which is the natural aim in image analysis.

Theorem 2: *Let S satisfy conditions (B2) and (B3) with $\mu = F$. Take*

$$h_n = C \left(\frac{\log n}{n} \right)^{1/d} \qquad \text{for some} \quad C > \left(\frac{2}{\delta w_d} \right)^{1/d},$$

where w_d denotes the Lebesgue measure of the unit ball in $\mathbb{R}^d$. Then, with probability 1,

$$d_H(\partial S, \partial S_n) \leq \sqrt{d} \max\{1 + d^{-1/2}, C(S)\} \, h_n \quad \text{eventually.}$$

Proof: Denote by $S_n(h)$ the H-shape with bin width h, and define the following estimator of S

$$\hat{S}_n(\epsilon) := \bigcup_{i=1}^{n} B(X_i, \epsilon), \qquad \epsilon > 0 \tag{8}$$

(see, e.g., Devroye and Wise 1980). We want to bound

$$d_H(\partial S_n(h_n), \partial S) = \max\{ \max_{x \in \partial S_n(h_n)} d(x, \partial S), \max_{x \in \partial S} d(x, \partial S_n(h_n)) \}.$$

If $x \in \partial S_n(h_n)$ then there exists an X_i such that $d(x, X_i) \leq \sqrt{d} h_n$. Thus, for every $x \in \partial S_n(h_n)$, we have $d(x, \partial S) \leq d(x, X_i) + d(X_i, \partial S) \leq \sqrt{d} h_n + d_H(\aleph_n, \partial S)$, where $\aleph_n = \{X_1, \ldots, X_n\}$. Then, by Theorem 3 in Cuevas and Rodríguez-Casal (2004), with probability 1,

$$\max_{x \in \partial S_n(h_n)} d(x, \partial S) \leq (\sqrt{d} + 1) h_n \quad \text{eventually.}$$

On the other hand, if $x \in \partial S \cap S_n^c(h_n)$, let X_i be the sample point nearest to x. As $X_i \in S_n(h_n)$ there must be a point $y \in \partial S_n(h_n)$ in the segment joining x with X_i. Thus, for every $x \in \partial S \cap S_n^c(h_n)$, we have that $d(x, \partial S_n(h_n)) \leq d(x, X_i) \leq d_H(\aleph_n, \partial S) \leq h_n$ eventually almost surely.

Finally, if $x \in \partial S \cap S_n(h_n)$ then $x \in \partial S \cap \hat{S}_n(\sqrt{d}h_n)$. By Theorem 4 in Cuevas and Rodríguez-Casal (2004), with probability 1,

$$\max_{x \in \partial S \cap S_n(h_n)} d(x, \partial S_n(h_n)) \leq \max_{x \in \partial S \cap \hat{S}(\sqrt{d}h_n)} d(x, \partial \hat{S}_n(\sqrt{d}h_n)) \leq C(S)\sqrt{d}h_n \quad \text{eventually.}$$

$\square$

Observe that the rates of convergence obtained in Theorem 2 for the H-shape are the same as those obtained by Cuevas and Rodríguez-Casal (2004) for the set estimator in (8). Both estimators have a non-smooth border. As we will see (Sections 4 and 5) in the case of the H-shape a smoother version can be obtained by averaging several shifted versions of the original estimator.

3   Estimation of a Gray-Scale Image: Generalizing the H-Shape

We can model data sampled from a colour or gray-scale image (defined on the unit cube) in the following way:

$$(X_1, Z(X_1)), (X_2, Z(X_2)), \dots, (X_n, Z(X_n)),$$

where the $X_i \in [0,1]^d$ are random variables representing the data location (usually $d = 2$) and $Z(X_i)$ represent the colour, the gray intensity or, in general, the mark associated to observation X_i. In the case of colour images, $Z(X_i)$ is a three-dimensional vector in $[0,1]^3$ specifying the amount of red, blue and green that determine the colour, while for gray-scale images $Z(X_i)$ only takes values in $[0,1]$. Since this fact greatly simplifies notation, the images considered from now on are just gray-scale, but we note that the proposed image estimator can be generalized straightaway to the coloured case.

We will assume that there exists a fixed image $z(x)$ (a measurable function from $[0,1]^d$ to $[0,1]$) which we want to estimate from its values, $z(X_1), \dots, z(X_n)$, on a sample of independent points $X_1, \dots, X_n$ taken uniformly into $[0,1]^d$. In other situations, though, it might be interesting to model the observed mark as $z(X_i) + \epsilon_i$, where $z(\cdot)$ is a fixed function and ϵ_i is some noise. Then the mark $Z(X_i)|X_i$ is a random variable in $[0,1]$, as in the general setting of marked point processes.

In order to estimate the image z, define a cubic partition $\mathcal{P}_n$, $n = 1, 2, \dots$, of the unit cube $[0,1]^d$ as in the previous section

$$\mathcal{P}_n = \{A_{nj}, j = 1, 2, \dots\} \quad \text{where} \quad A_{nj} = \prod_{i=1}^{d} [k_i h, (k_i + 1)h), \quad k_i \in \mathbb{Z}$$

Assume that

$$h_n \to 0 \text{ as } n \to \infty \quad \text{and} \quad \sum_{n=1}^{\infty} \exp(-nh^d) < \infty. \tag{9}$$

9

By Borel-Cantelli lemma, with probability one, for n sufficiently large at least one sample point X_i falls into every $A_{nj} \in \mathcal{P}_n$ whose intersection with the unit cube has positive Lebesgue measure. Then we can define a piecewise constant estimator $\hat{z}$ of z as follows. For every $x \in \text{int}(A_{nj})$,

$$\hat{z}(x) = \hat{z}_{nj} := G_{nj}^{-1}(1/2),$$

where G_{nj} denotes the empirical distribution of $\{z(X_i) : X_i \in A_{nj}\}$ and, for any distribution G and any $0 < p < 1$, $G^{-1}(p) := \min\{x : G(x) \geq p\}$ denotes the p-quantile of G. For $x \in \partial A_{nj}$ define $\hat{z}(x)$ in such a way that $\hat{z}$ is upper semicontinuous (USC), that is, if $x \in \partial A_{nj} \cap \partial A_{nj'}$ with $j \neq j'$ then $\hat{z}(x) = \max(\hat{z}_{nj}, \hat{z}_{nj'})$.

We would like to know if $\hat{z}$ is a consistent estimator of z. We will restrict ourselves to the simplest case, a *polychotomous* image, that is, a finite union of constant-coloured pieces. Thus z is a piecewise constant USC function

$$z(x) = \sum_{k=1}^{M} z^{(k)} \mathbb{I}_{S^{(k)}}(x), \tag{10}$$

where the positive integer M denotes the maximum number of different colours, $z^{(1)}, \ldots, z^{(M)} \in [0, 1]$, appearing in the image and $S^{(k)}$ denotes the set $\{x \in [0, 1]^d : z(x) = z^{(k)}\}$, $k = 1, \ldots, M$. From now on it is assumed that $\text{Leb}(S^{(k)}) > 0$ and that $\text{cl}(\text{int}(S^{(k)})) = \text{cl}(S^{(k)})$, for all $k = 1, \ldots, M$, where $\text{cl}(A)$ denotes the closure of a set A. We can express $\hat{z}$ in a way analogous to (10):

$$\hat{z}(x) = \sum_{k=1}^{m} z^{(k)} \mathbb{I}_{\hat{S}^{(k)}}(x), \tag{11}$$

where $\hat{S}^{(k)} := \{x \in [0, 1]^d : \hat{z}(x) = z^{(k)}\}$ and $m(\leq M)$ is the number of different colours observed in the sample $X_1, \ldots, X_n$. It is easy to see that, with probability one, there exists an n from which on $m = M$. Therefore we will assume, from now on, that a sample point has fallen into the interior of every A_{nj} intersecting the unit cube, that all M colours have been observed (so that m in (11) is equal to M) and that $[0, 1]^d = \bigcup_{k=1}^{M} S^{(k)}$.

As a measure of the discrepancy between the estimated image $\hat{z}$ and the real one z we can use the Hausdorff distance between the boundaries of their hypographs

$$d_{H^*}(\hat{z}, z) := d_H(\partial H_{\hat{z}}, \partial H_z),$$

where $H_z = \{(x, y) \in [0, 1]^{d+1} : x \in [0, 1]^d, 0 \leq y \leq z(x)\}$ denotes the hypograph of function z. This distance is properly defined since z and $\hat{z}$ are USC. The reason for using d_{H^*} (instead of, say, the supremum metric) is that d_{H^*} is a "visual" distance which characterizes well the "physical" proximity between two images and allows this proximity to be checked in any direction apart from the vertical one (see, e.g., Cuevas and Fraiman 1998).

10

The following result analyses the rates of convergence of $d_{H^*}(\hat{z}, z)$. In particular if $h = C(\log n/n)^{1/d}$ for any $C > 1$ (in order to fulfill condition (9)) then $d_{H^*}(\hat{z}, z) = O(\log n/n)$ a.s., which means that the order of convergence appearing in Theorem 2 for binary images is still preserved for the coloured ones.

Theorem 3:   *If $S^{(k)}$ satisfies an interior cone condition, for all $k = 1, \ldots, M$, and h satisfies (9) then $d_{H^*}(\hat{z}, z) \leq Ch$ eventually a.s., where C is a constant which depends only on d and on the cone condition constants.*

Proof: Observe that

$$d_{H^*}(\hat{z}, z) \leq \max_{k=1,\ldots,M} d_H(\text{cl}(\hat{S}^{(k)}), \text{cl}(S^{(k)}))$$

and

$$d_H(\text{cl}(\hat{S}^{(k)}), \text{cl}(S^{(k)})) = \max\{ \sup_{x \in \hat{S}^{(k)}} d(x, \text{cl}(S^{(k)})), \sup_{x \in S^{(k)}} d(x, \text{cl}(\hat{S}^{(k)}))\}.$$

If $x \in \hat{S}^{(k)}$ then there exists $A_{nj} \in \mathcal{P}_n$ such that $x \in \text{cl}(A_{nj})$ and there is at least one sample point $X_i \in A_{nj}$ with $z(X_i) = z^{(k)}$, that is, $X_i \in A_{nj} \cap S^{(k)}$. Thus, for all $x \in \hat{S}^{(k)}$, we have that $d(x, \text{cl}(S^{(k)})) \leq \sqrt{d}\,h$ a.s.

On the other hand, if $x \in S^{(k)} \cap [\hat{S}^{(k)}]^c$ then $x \in S^{(k)} \cap \hat{S}^{(k')}$ for some $k' \neq k$. This implies that there exists $A_{nj} \in \mathcal{P}_n$ such that $x \in \text{cl}(A_{nj})$ and $A_{nj} \cap S^{(k')} \neq \emptyset$ and thus $d(x, \partial S^{(k)} \cap \partial S^{(k')}) \leq \sqrt{d}h$. Consequently for every $x_0 \in S^{(k)} \cap [\hat{S}^{(k)}]^c$ we have that

$$d(x_0, \text{cl}(\hat{S}^{(k)})) \leq \sqrt{d}h + \max_{x \in \partial S^{(k)}} d(x, \text{cl}(\hat{S}^{(k)})).$$

The final step in the proof will be to show that, as $S^{(k)}$ satisfies the interior cone condition, then if n is sufficiently large we have that $\max_{x \in \partial S^{(k)}} d(x, \text{cl}(\hat{S}^{(k)})) \leq Ch$.

The reasoning is done for $d = 2$ (see Figure 1) but it is generalizable to any dimension. Using notation in condition (B4), take the point z_1 on the segment $\overline{xy}$ joining x and y, such that the intersection of Ψ_x with the straight line $l(z_1)$ passing through z_1 and orthogonal to $\overline{xy}$ has length $2\sqrt{2}h$. Similarly take the point z_2 such that the length is $4\sqrt{2}h$. Observe that, due to the uniform cone condition, these points z_i exist for n suficiently large and not dependent on x. Also the Euclidean distance from x to z_i, $\|x - z_i\|$, is $C_i h$ where $C_i > 0$, $i = 1, 2$, are constants not dependent on x either. Then the ring section given by $\Psi_x \cap [B(x, C_2 h) \backslash B(x, C_1 h)]$ necessarily contains an $A_{nj'} \in \mathcal{P}_n$ such that $A_{nj'} \subseteq S^{(k)}$ and thus, with probability 1, $A_{nj'} \subseteq \hat{S}^{(k)}$. As $x \in [\hat{S}^{(k)}]^c$ this implies that $d(x, \partial \hat{S}^{(k)}) \leq C_2 h$ and, as the n from which on this holds or the C_i do not depend on x, in general we have that $\max_{x \in S^{(k)} \cap [\hat{S}^{(k)}]^c} d(x, \partial \hat{S}^{(k)}) \leq C_2 h$ eventually a.s.
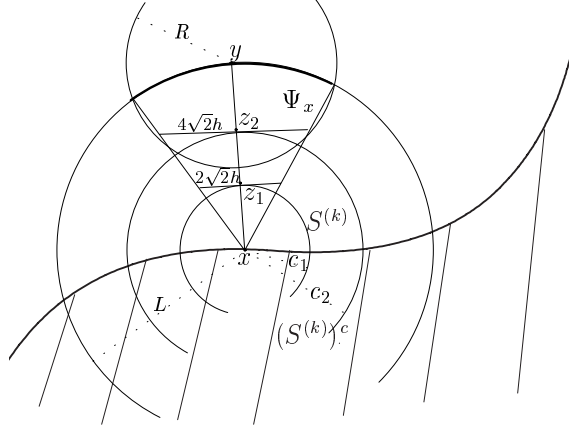
$\square$

Figure 1: $S^{(k)}$ satisfies an interior cone condition

4 Averaged Shifted H-Shapes. A Real Data Example

In spite of its computational advantages, the H-shape has a rough boundary which hides some of the features suggested by the original data at first sight. For instance, in Figure 2a we can see several locations of drumlins in Northern Ireland (Upton and Fingleton 1985). Drumlins are elongated wave-shaped hills formed by glacial action and whose axis is parallel to the movement of the ice. In Geology it is interesting to determine the outline of drumlin fields or clusters, since they indicate continental ice sheet glaciation. The H-shape of these data depicted in Figure 2c is not very informative regarding the outline or orientation of drumlin fields. It has very sharp edges and also displays a large number of "holes", though some of these are usually swept away (Figure 2d) by the simple morphological operations of "opening" and "closing" the estimator with a structuring element (see, e.g., Serra 1982).

Following the lines described in Scott (1992, Ch. 5) for the *averaged shifted histograms*, we may define a gray-scale image $\hat{z}$ to estimate the original binary image z determined by the indicator function on S, $z(x) = \mathbb{I}_S(x)$ for all $x \in [0,1]^d$. This estimator $\hat{z}$, which we will call *ASH-shape*, is the result of averaging several of the H-shapes with the same bin width h, but differing grid origins. The ASH-shape is intended to retain the computational efficiency of the H-shape, but at the same time produce a "smoother" image, in the sense that, the further a bin is from the sample, the lighter gray it is assigned.

Consider a collection of H-shapes, as defined in (1), each with bin width h, but with grid origins differing only in coordinate shifts that are multiples of $\delta = h/m$, where m is a positive integer. These origins are given by $t_{\mathbf{i}} = \delta\mathbf{i} = (i_1\delta, i_2\delta, \ldots, i_d\delta)$, where $\mathbf{i} = (i_1, i_2, \ldots, i_d)$ and the i_j's are integers between $1 - m$ and 0. Observe that the total number of shifted H-shapes is m^d. We will
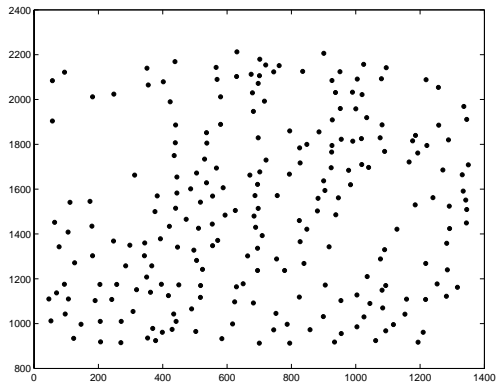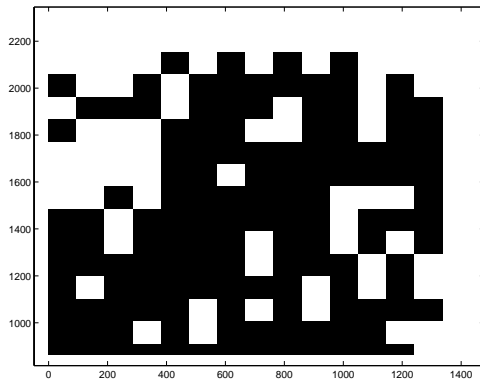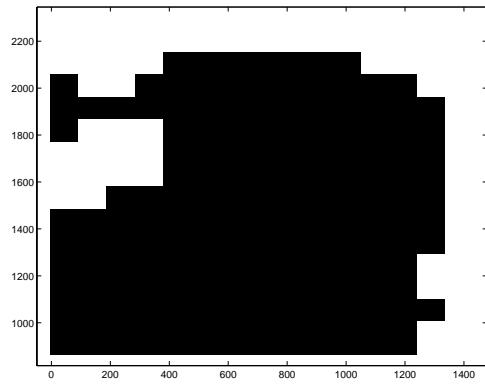
Figure 2: (a) Drumlin locations in Northern Ireland; (b) Map of the area where drumlins were tracked; (c) H-shape of data in (a); (d) Closing of H-shape in (c).

denote by $\hat{S}_{\mathbf{i}}$ and by $\mathcal{P}_{\mathbf{i}}$ respectively the H-shape and the partition of $[0,1]^d$ with grid origin in $t_{\mathbf{i}}$:

$$\mathcal{P}_{\mathbf{i}} = \{A = t_{\mathbf{i}} + \prod_{j=1}^{d} [k_j h, (k_j + 1)h), A \cap [0,1]^d \neq \emptyset, k \in \mathbb{Z}\}, \quad \mathbf{i} \in \{1 - m, \ldots, 0\}^d$$

$$\hat{S}_{\mathbf{i}} = \bigcup \{A \in \mathcal{P}_{\mathbf{i}} : \ X_j \in A \text{ for some } j = 1, \ldots, n\}, \quad \mathbf{i} \in \{1 - m, \ldots, 0\}^d.$$

Let $\hat{z}_{\mathbf{i}}$ be the indicator function of $\hat{S}_{\mathbf{i}}$ in $[0,1]^d$. Then the ASH-shape can be expressed in the following way

$$\hat{z}(x) = \frac{1}{m^d} \sum_{\mathbf{i}} \hat{z}_{\mathbf{i}}(x) = \frac{1}{m^d} \sum_{i_1 = 1-m}^{0} \cdots \sum_{i_d = 1-m}^{0} \mathbb{I}_{S_{\mathbf{i}}}(x), \quad x \in [0,1]^d,$$

and its support is given by

$$\hat{S} = \mathrm{cl} \left(\bigcup_{i_1 = 1-m}^{0} \cdots \bigcup_{i_d = 1-m}^{0} \hat{S}_{\mathbf{i}} \right).$$

In particular, in the two-dimensional case the ASH-shape $\hat{z}$ is piecewise constant over the bins

$$B_{\mathbf{i}} = [i_1 \delta, (i_1 + 1)\delta) \times [i_2 \delta, (i_2 + 1)\delta), \quad \mathbf{i} = (i_1, i_2) \in \mathbb{Z}^2.$$

For example, the height of the ASH-shape in bin $B_{\mathbf{0}}$ is the average of the indicators of the m^2 shifted H-shapes, each of bin width $h = m\delta$, over bin $B_{\mathbf{0}}$:

$$\hat{z}(x) \ = \ \frac{1}{m^2} \sum_{i_1 = 1-m}^{0} \sum_{i_2 = 1-m}^{0} \mathbb{I}_{t_{\mathbf{i}} + [0,h)^2}(X_k, \text{ for some } k), \quad x \in B_0.$$

Figure 3 displays the resulting ASH estimators and their closings for the drumlin data. Observe the diagonal orientation of the drumlin field which agrees with the fact that the glacial ice sheet moved into Ireland from the Northeast end into the Southwest direction.

5   Implementation of the Set Estimators. Simulations

In order to check the performance of the H- and ASH-shape we have implemented both in digital domain. In our simulations we have considered the foreground (the class of object or black pixels) of a certain digital image as the original set S. For example, in Figure 4a the unit square has $121 \times 137 = 16577$ pixels and the foreground has a butterfly shape with 6812 object pixels. We have drawn uniformly distributed samples of size $n = 100$, $n = 300$ and $n = 1000$ from the foreground (see Figures 4b, 4c and 4d).

The choice of the bin width h is crucial. Ray Chaudhuri, Chaudhuri and Parui (1997) proposed the following iterative method to select h. Let the grid origin be fixed throughout the whole procedure. The initial value $h_{n,1}$ is given by (7) and the subsequent bin widths are defined by

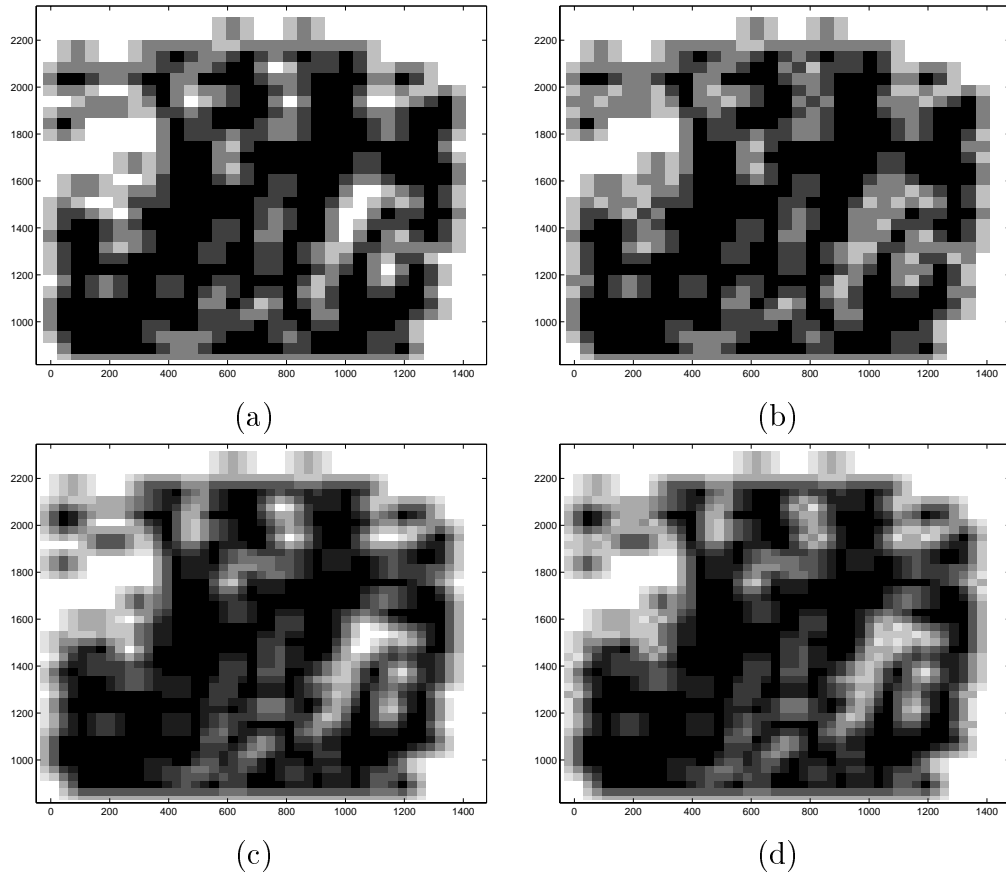$$h_{n,i} = \sqrt{\mathrm{Leb}(S_{n,i-1})/n} \quad \text{if } i > 1.$$

Figure 3: Drumlin data. (a) ASH estimator for $M = 2$; (b) Closing of ASH estimator for $M = 2$; (c)ASH estimator for $M = 3$; (d) Closing of ASH estimator for $M = 3$.
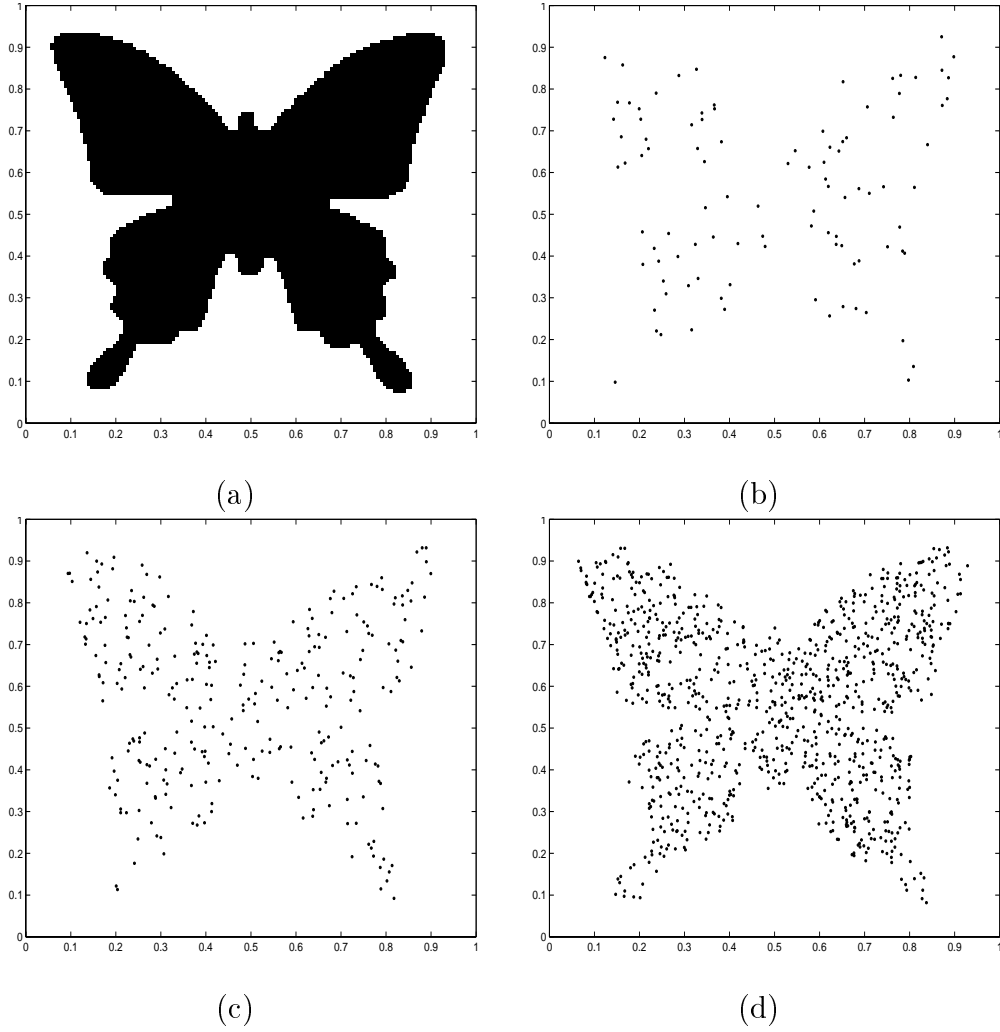
Figure 4: (a) Original binary digital image. Samples of size (b) $n = 100$, (c) $n = 300$ and (d) $n = 1000$.

where $S_{n,i}$ denotes the H-shape determined by bin width $h_{n,i}$. For a fixed $\epsilon > 0$ let the ϵ-measure of dispersion be defined by $h_{n,i_0} = \max\{h_{n,i} : |h_{n,i-1} - h_{n,i+1}| < \epsilon h_{n,i}\}$. Ray Chaudhuri, Chaudhuri and Parui (1997) used the bin width given by the ϵ-measure of dispersion for $\epsilon = 0.4$. However, our simulations have shown that, in general, $\epsilon = 0.4$ produces a small h_{n,i_0}. Actually the ϵ-measure of dispersion is not sufficiently sensitive to variation of ϵ, which means that finding an "appropriate", data-driven ϵ might not be an easy task. This is the reason why we have gone one step further and chosen, from the grid of decreasing bin widths $\{h_{n,i}\}_{i=1}^{i_0}$, the following one. For a fixed $p \in (0,1)$ compute

$$
i^* := \begin{cases} \max\left\{i \in \{1, \ldots, i_0\} : \sharp\{j : X_j \in S_{n,i}^{(j)}\} \geq p \cdot n\right\} & \text{if } \sharp\{j : X_j \in S_{n,i}^{(j)}\} > 0 \\ 1 & \text{if } \sharp\{j : X_j \in S_{n,i}^{(j)}\} = 0 \end{cases}
$$

where $S_{n,i}^{(j)}$ denotes the H-shape with bin width $h_{n,i}$ and constructed from the sample $X_1, \ldots, X_{j-1}$, $X_{j+1}, \ldots, X_n$. Thus h_{n,i^*} approximates the smallest bin width such that the corresponding H-shape intersects at least an area p of S. In the simulations we have taken $\epsilon = 0.4$ and $p = 0.7$.

In Figures 5, 7 and 6 the resulting ASH-shape for the butterfly samples is displayed for $m = 1, 2$ and 3. As the images point out, the choice $m = 2$ or $m = 3$ already offers a remarkable improvement over the original H-shape. The foreground of these estimators was also morphologically operated by a "binary closing" (as indicated in Ray Chaudhuri *et al.* 1999), where the structuring element was 2×2 and had all entries equal to 1. This means that we merged into the estimator those bins which were not a part of a 2×2 lattice of pixels completely contained in the background (the class of white pixels) of the image. The mark $\hat{z}$ associated to this new bin A was obviously 1 in the case of $M = 1$ and the mean of the non-zero marks in the eight pixels surrounding A in the case of $M > 1$.

6   Concluding Remarks

In this work we aim at highlighting some of the advantages associated to the use of a simple set estimator such as the H-shape. For instance, the fact that it can incorporate intuitive and (digitally reasonable) shape restrictions make it attractive from the theoretical point of view, since this controls the VC dimension of the class of possible estimators and thus opens the door to using empirical process theory (Section 2). Under very general geometrical conditions we have also obtained (nonparametric) rates of convergence of the boundary of the H-shape to that of the original image (Section 2). The rough border of this histogram-type estimator may thus serve as a first orientative (but asymptotically efficient) approximation to extracting a more elaborate and smoother description of the dot pattern shape.
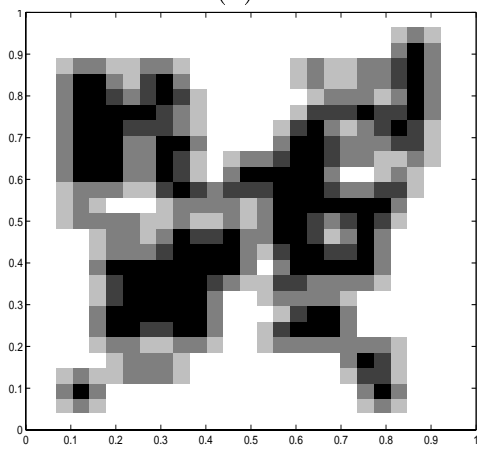
From a more practical viewpoint we have seen (Sections 4 and 5) that the lack of smoothness of the H-shape can be overcome, to a large extent, by a softer gray-scale image, the ASH-shape. This estimator, which averages several shifted H-shapes, clearly profits from (and retains) the
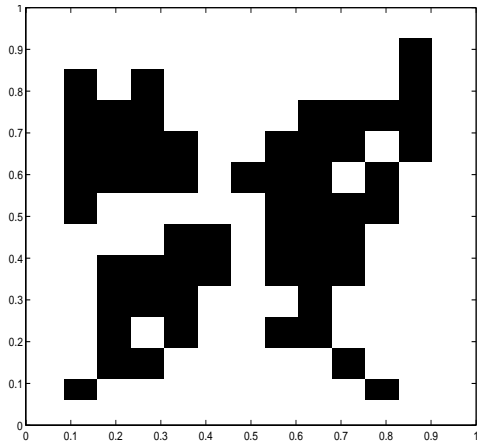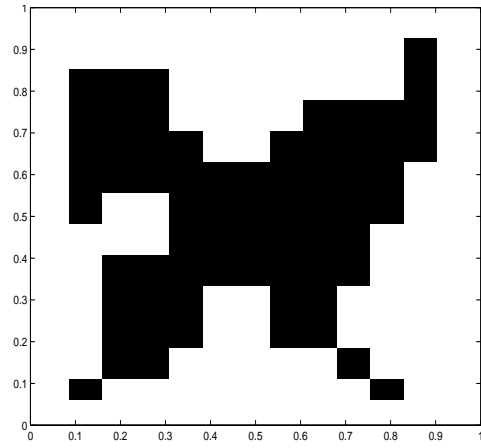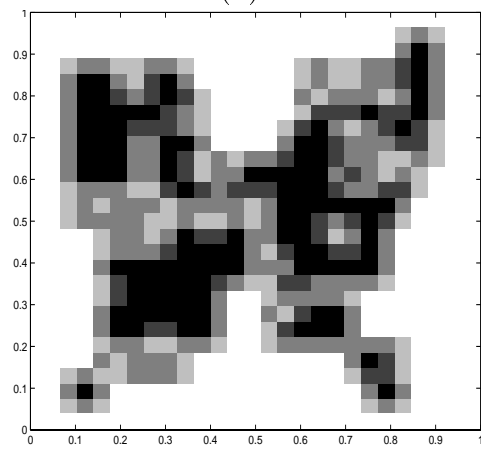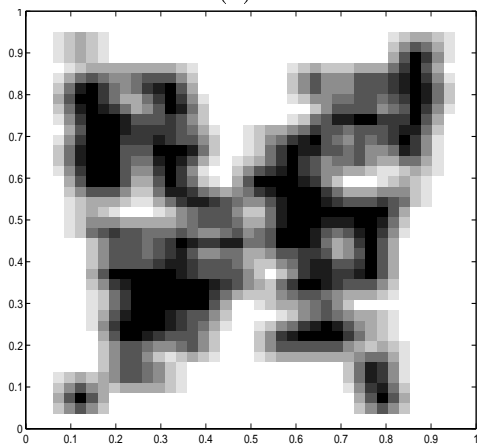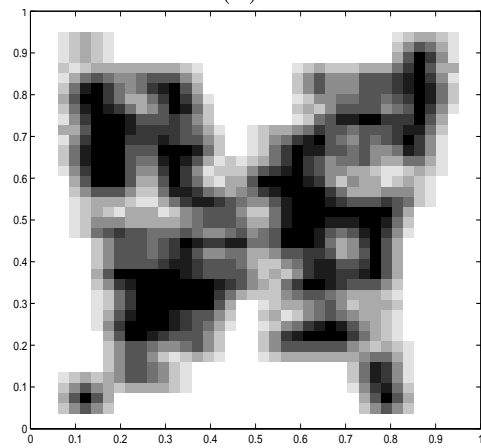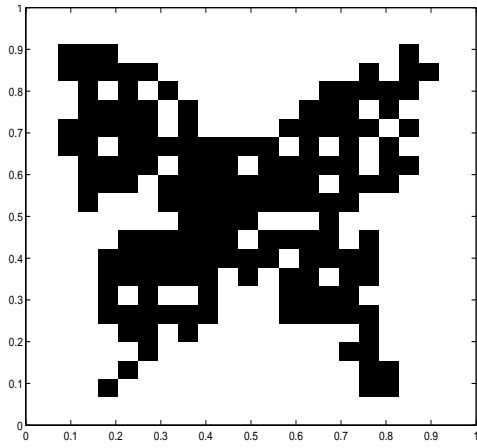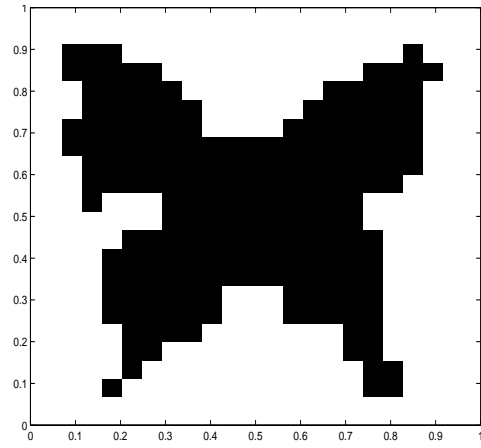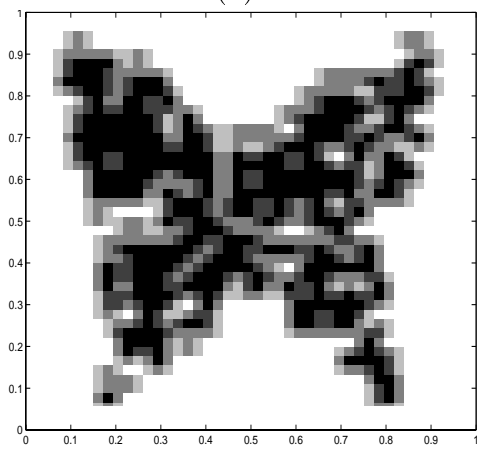
Figure 5: ASH estimator for $n = 100$ and (a) $M = 1$, (c) $M = 2$, (e) $M = 3$. Closing of ASH estimator for $n = 100$ and (b) $M = 1$, (d) $M = 2$, (f) $M = 3$.
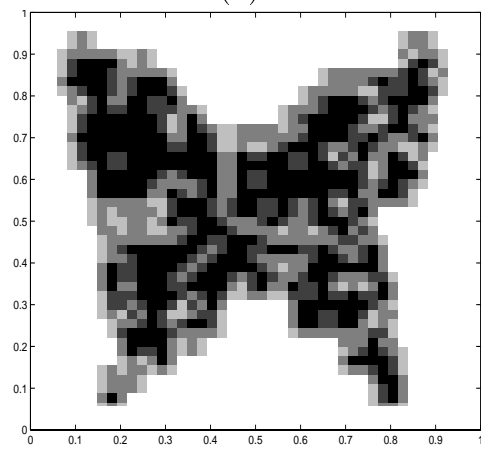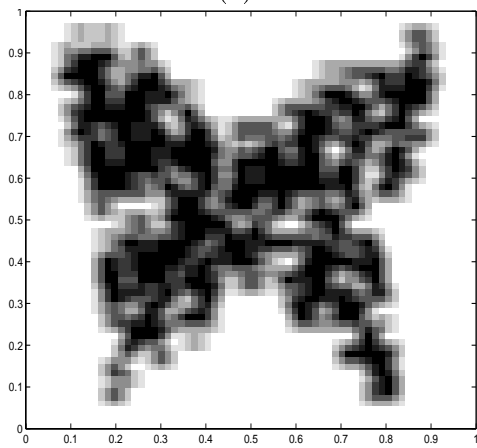
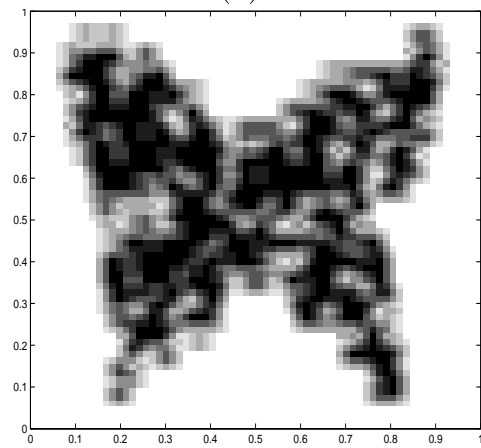Figure 6: ASH estimator for $n = 300$ and (a) $M = 1$, (c) $M = 2$, (e) $M = 3$. Closing of ASH estimator for $n = 300$ and (b) $M = 1$, (d) $M = 2$, (f) $M = 3$.
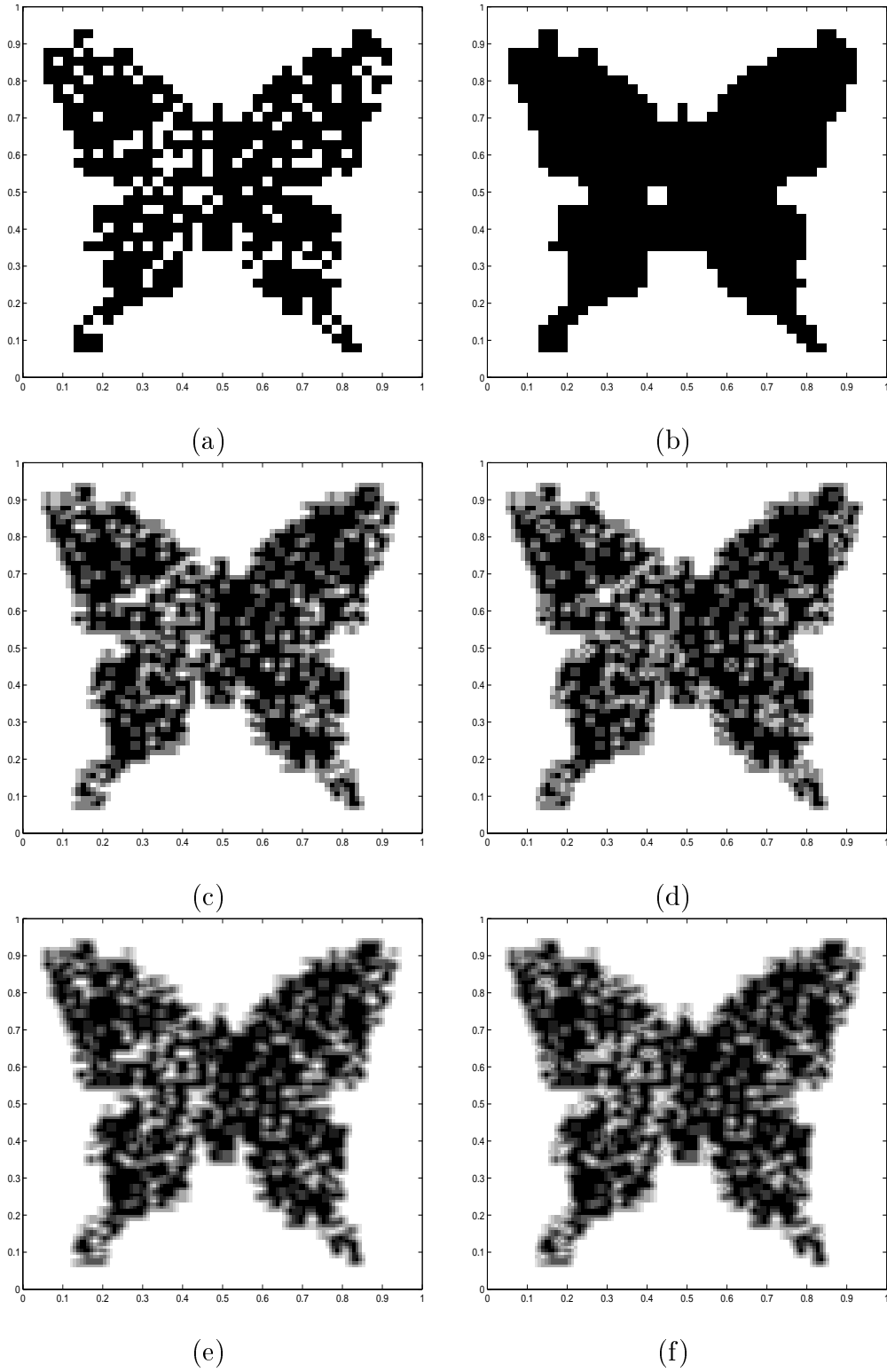
Figure 7: ASH estimator for $n = 1000$ and (a) $M = 1$, (c) $M = 2$, (e) $M = 3$. Closing of ASH estimator for $n = 1000$ and (b) $M = 1$, (d) $M = 2$, (f) $M = 3$.

computational simplicity of the H-shape. One further advantage of the H-shape as an image estimator is that it can be adapted to approximate colour images (Section 3). The resulting marked-bin estimator preserves both the computational efficiency of the H-shape and the convergence rates to the boundaries separating different colours.

References

Baíllo, A., Cuevas, A. and Justel, A. (2000). Set estimation and nonparametric detection. *Canad. J. Statist.*, 28, 765-782.

Braker, H., Hsing, T. and Bingham, N. H. (1998). On the Hausdorff distance between a convex set and an interior random convex hull. *Adv. Appl. Prob.*, 30, 295-316.

Cuevas, A. and Fraiman, R. (1998). On visual distances in density estimation: the Hausdorff choice. *Statist. Prob. Letters*, 40, 333-341.

Cuevas, A. and Rodríguez-Casal, A. (2003). Set estimation: an overview and some recent developments. In *Recent Advances and Trends in Nonparametric Statistics*, M. G. Akritas and D. N. Politis eds., Elsevier (North Holland), pp. 251-264.

Cuevas, A. and Rodríguez-Casal, A. (2004), On boundary estimation. *Adv. Appl. Prob.*, 36, 340-354.

Devroye, L. and Györfi, L. (1985). *Nonparametric Density Estimation: The L_1-View*. John Wiley, New York.

Devroye, L., Györfi, L. and Lugosi, G. (1996). *A Probabilistic Theory of Pattern Recognition*. Springer, New York.

Devroye, L. and Wise, G. L. (1980). Detection of abnormal behavior via nonparametric estimation of the support. *SIAM J. Appl. Math.*, 38, 480-488.

Gilbarg, D. and Trudinger, N. S. (1983). *Elliptic Partial Differential Equations of Second Order*. Springer-Verlag.

Hall, P., Park, B. U. and Turlach, B. A. (2002), Rolling-ball method for estimating the boundary of the support of a point-process intensity. *Ann. Inst. Henri Poincaré Prob. Statist.*, 38, 6, 959-971.

Klemelä, J. (2004). Complexity penalized support estimation. *J. Multivariate Anal.*, 88, 274-297.

Korostelev, A. P. and Tsybakov, A. B. (1993). *Minimax Theory of Image Reconstruction*. Springer-Verlag, New York.

Ray Chaudhuri, A., Basu, A., Bhandari, S. K. and Chaudhuri, B. B. (1999). An efficient approach to consistent set estimation. *Sankhyā Ser. B*, 61, 3, 496-513.

Ray Chaudhuri, A., Chaudhuri, B. B. and Parui, S. K. (1997). A novel approach to computation of the shape of a dot pattern and extraction of its perceptual border. *Computer Vision and Image Understanding*, 68, 3, 257-275.

Rényi, A. and Sulanke, R. (1963). Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrsch. Verw. Gebiete*, 2, 75-84.

Rényi, A. and Sulanke, R. (1964). Über die konvexe Hülle von n zufällig gewählten Punkten (II). *Z. Wahrsch. Verw. Gebiete*, 3, 138-147.

Rodríguez-Casal, A. (2003). Set estimation under convexity type assumptions. Manuscript.

Schneider, R. (1988). Random approximation of convex sets. *J. Microscopy*, 151, 211-227.

Scott, D. W. (1992). *Multivariate Density Estimation. Theory, Practice, and Visualization.* Wiley, New York.

Serra, J. (1982). *Image Analysis and Mathematical Morphology.* Academic Press, New York.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995). *Stochastic Geometry and its Applications.* Wiley.

Upton, G. and Fingleton, B. (1985). *Spatial Data Analysis by Example.* Wiley, New York.

Walther, G. (1997). Granulometric smoothing. *Ann. Statist.*, 25, 6, 2273-2299.