Estimating statistics in arbitrary regions of interest

Timor Kadir[1,2] and Michael Brady[2]
[1]Siemens Molecular Imaging Advanced Applications
23-38 Hythe Bridge Street, Oxford, OX1 2EP, UK
[2]Robotics Research Group, University of Oxford, Oxford, OX1 3PJ, UK
{timork,jmb}@robots.ox.ac.uk

Abstract

We address the problem of estimating statistics in regions of interest (ROIs) containing both whole and partial pixels. Such ROIs arise frequently in vision problems such as segmentation and registration. For example, even where the control points of an ROI, say the vertices of a polygon, are forcibly aligned with the pixel grid, the connecting edges will rarely do so. In medical image analysis, for instance, this can be a cause of significant error. More generally, any cost function that includes statistics estimated from the image will often exhibit irregularities due to such partial pixels.

Our proposed solution addresses this problem by correctly accounting for the partial pixel area. Moreover, the method has no arbitrary parameters such as bin widths or kernel sizes. It implicitly addresses the issue of independence and gives rise to continuous density estimates whose quality is, in principle at least, independent of the number of pixels in the ROI. We present results to compare our proposed method with conventional techniques such as weighted histograms and Parzen windowing.

1 Introduction

Probabilistic and statistical methods are of fundamental importance in image (and signal) analysis, but particularly in medical image analysis. For example, image segmentation based on variants to region competition [8] requires iterative estimation of the probability density function (PDF) inside and outside a putative region that is growing or shrinking during the segmentation process. The segmentation of MRI images of the brain into regions corresponding to grey matter (GM), white matter (WM), cerebrospinal fluid (CSF) etc, such as the hidden Markov random field (HMRF) approach developed by [7], require either that pdfs for those tissue classes be known in advance or, more usually, that they be estimated directly from the (3D) image. The simultaneous segmentation and registration (alignment) of images [6] based on HMRFs also requires the estimation of PDFs from images. Examples such as these, and probabilistic modelling techniques such as those motivated through Bayesian analysis, rely upon the accurate and reliable estimation of the PDF of some appropriate signal (class).

PDF estimators may be classified using three categories: parametric, non-parametric and semi-parametric. Parametric techniques are suitable where a particular form of func-

BMVC 2005 doi:10.5244/C.19.64

tion can be assumed for application specific reasons. For example, Rician and Rayleigh functions are often used in ultrasound signal processing. Of the non-parametric techniques, the simplest and most widely used method is the histogram. Its limitations, principally the requirement to pre-specify the number of bins, the arbitrary bin boundaries and the block-like nature of the resulting PDF estimate have lead to the development of a number of alternative methods. For example, Parzen windowing avoids arbitrary bin assignments and leads to smoother PDFs. However, a suitable kernel shape and size must still be chosen. Finally, semi-parametric techniques such as Gaussian Mixture Models, in which the superposition of a number of parametric densities are used to approximate the underlying density, offer a useful compromise between these two approaches.

Recently, [2] noted that common to all such methods is the assumption that the samples from which the PDF is estimated are Independent and Identically Distributed (IID), despite the fact that image samples are evidently *not* independent, a fact which is universally acknowledged, and ignored. However, images (and signals), from which we wish to estimate PDFs are both Band-Limited and (at least) Critically Sampled (BL-CS). For such signals, the samples represent the band-limited continuous signal, evaluated and quantised at some (arbitrary) spatial or temporal points. This differs fundamentally from the IID assumption in several important ways: the samples are *not* independent; the samples are ordered; and knowledge of the samples, their order and the sampling pre-filter are enough to uniquely specify the band-limited continuous signal exactly (upto quantisation[4]).

Starting from this observation, [2] have developed a general, non-parametric technique for the accurate and stable estimation of PDFs. The method is briefly sketched in the next Section. Their method has several useful properties. First, the domain resolution of the resulting estimate is continuous, hence is independent of the number of sample points and signal quantisation. That is, the number of points at which the PDF can be usefully evaluated is infinite, regardless of the number of samples from which it is calculated. Usefully means that new information can be obtained by evaluating the PDF at finer intervals, in contrast to, for example, Parzen windows, for which evaluation at points finer than the quantisation reveals information only about the smoothing kernel.

In this paper, we extend the method introduced in [2] to demonstrate its applicability for handling partial pixels arising from non-pixel aligned regions of interest (ROIs). Such cases are quite common and give rise to the so-called partial volume effect (PVE), which is particularly important in medical image analysis. Consider, as an example, MRI, which is justifiably the leading method for imaging soft tissue, such as those in the brain. Different imaging (pulse) sequences can give exquisite contrast between grey and white matter, or between CSF and grey/white matter. However, in clinical practice, the resolution of current MRI systems is approximately $1mm^3$, and this is large compared to the size of brain structures that need to be viewed and, more importantly, measured. Increasingly, medical image analysis is quantitative: we seek for example to measure the areas and/or volumes of structures such as temporal lobes which have complex shapes, and we seek to measure changes that are characteristic of disease progression (eg degenerative brain disease) or changes wrought by a drug or radiotherapy. Many brain structures have portions that are thinner than the MRI resolution and are, at first sight, blurred out. Indeed, it has been shown that if one does not analyse the PVE, then the error in estimates of the volumes of small or elongated, articulated structures can be as much as 40%. Fortunately, MRI has the convenient property that the signal resulting from a voxel that is composed of two or more tissues is the weighted sum of the signal that would result from pure tissue

voxels, the weight being the relative amount of the various tissues. Several algorithms have been developed for overcoming the PVE, for example [7] but all of them depend on assumptions about the PDFs of pure tissues (typically that they are Gaussian). Note that, in most cases, PVE analysis follows segmentation though it is evidently the case that PVE may affect considerably estimation of the PDFs used in segmentation. In this paper, we investigate whether the PDF estimation technique introduced by [2] can be extended to accommodate the partial pixels. More precisely, we aim to demonstrate that the PDF estimation is stable and accurate as a result of subpixel (subvoxel) displacements in the image and ROI. This is an important ingredient in developing well-conditioned algorithms that can segment and measure complex structures in images.

2 Estimating a PDF

In this section, we briefly recall the method introduced by [2]. The observation that images are band-limited and critically sampled leads directly to the Whittaker-Shannon sampling theory, which states that three pieces of information are necessary to specify the original band-limited signal at any point: the samples, their order and the pre-filter characteristics. Conventional PDF estimation methods use only the first of these. Since the positions of the sample points (pixels) in standard sampling, are arbitrary with respect to the image, small changes in their position can give rise to different values; pixel values are unstable. One way in which this instability can be overcome is to generate many more of them, for example by upsampling the signal and using these to build a histogram. The only factor which controls the goodness of the histogram are the number of samples we choose to take and the accuracy of the interpolator. Moreover, we are free to choose an arbitrarily small bin size given sufficient samples. Upsampling requires the use of interpolation. [2] notes that many common interpolation schemes proceed by fitting piecewise functions to the signal samples and re-sampling these at the required points. In this method, however, the aim is to estimate the signal PDF, we can avoid the re-sampling step and calculate the PDF of each piecewise section directly in closed-form. Such an approach is in general more accurate and more efficient to implement than the over-sampling method. In fact, its accuracy is dependant only on the accuracy of the piecewise representation, not the number of samples nor the number of bins.

To calculate the PDF (or CDF) of the piecewise function, the signal is considered a function of a uniform random variable representing its domain. From standard probability theory, a function of a random variable creates another random variable whose distribution may be determined by the Transformation formula or, alternatively, the distribution method [3]. The algorithm to calculate the PDF/CDF consists of three main steps:

1. Calculate the polynomial coefficients for the signal samples;
2. Calculate the PDF/CDF for each piecewise section;
3. Populate the appropriate bins for each piecewise section.

The final step is necessary only if an explicit numerical representation of the PDF is required. Step 1 is standard and will not be discussed further. The details of steps 2 and 3 for 2D Bilinear interpolation are recalled below. In this section (but not the next), we adopt the convention that the piecewise spans start at zero and are of unit length. Consequently,

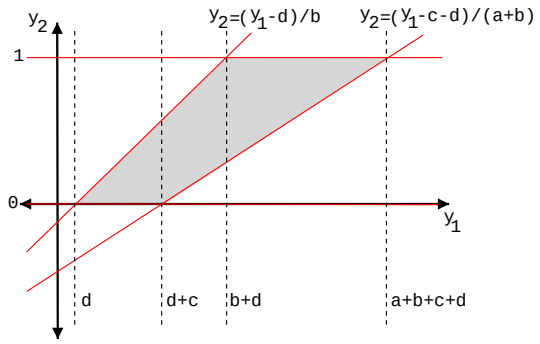


Figure 1: The integration ranges shown graphically for a particular configuration of the bilinear spline. Shown here is the case for $\{a,b,c,d\} > 0$ and $b > c$.

the PDF of the domain variable is unity over its range: $f_x(x) = 1,\ 0 \le x \le 1$. Note that bilinear interpolation will, in general, only approximate the true band-limited signal and hence the true PDF. Better approximations can be obtained if an accurate representation of the sampling pre-filter is available through prior knowledge of the sensor or alternatively by empirical estimation [5].

For a 1D signal, each pair of adjacent samples uniquely defines a straight line, which we denote $y(x) = ax + b$. Then the PDF is given by:

$$\begin{aligned} f_y(y) &= \frac{1}{|a|} f_x\left(\frac{y-b}{a}\right) \\ &= \frac{1}{|a|} \qquad b \le y \le a+b \end{aligned} \tag{1}$$

This has a straightforward and intuitive implementation: the PDF is simply the superposition of piecewise constant sections of magnitude $\frac{1}{|a|}$ between domain values b and $a+b$. The derivation for the two dimensional case requires the introduction of a dummy variable which must be integrated out in the final step, denoted x_2 in the following:

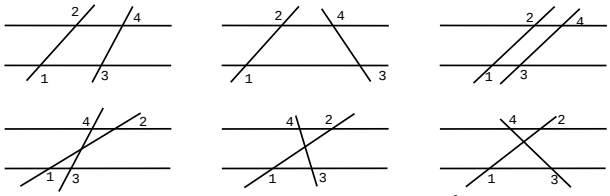
$$\begin{aligned} y_1(x_1,x_2) &= ax_1x_2 + bx_1 + cx_2 + d \qquad y_2(x_1,x_2) = x_1 \\ x_2(y_1,y_2) &= \frac{y_1 - by_2 - d}{ay_2 + c} \qquad x_1(y_1,y_2) = y_2 \end{aligned}$$

The derivative used in the univariate case becomes a Jacobian in the multi-variate case:

$$\begin{aligned} |J| &= \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} \\ &= \begin{vmatrix} 0 & 1 \\ \frac{-1}{(ay_2+c)} & \frac{-b}{(ay_2+c)} + \frac{(-by_2-y_1+d)a}{(ay_2+c)^2} \end{vmatrix} \\ &= \frac{1}{ay_2+c} \end{aligned} \tag{2}$$

The joint PDF between y_1 and y_2 is given by:

$$f_{y_1,y_2} = f_{x_1,x_2}(y_2, \frac{y_1 - by_2 - d}{ay_2 + c})|J| = \frac{1}{ay_2+c} \qquad \begin{matrix} 0 \le\ y_2\ \le 1 \\ by_2 + d \le\ y_1\ \le y_2(a+b) + c + d \end{matrix} \tag{3}$$



$$
\begin{array}{rcrcrcl}
1 & : & y_2 = 0 & \cap & y_2 & = & \frac{y_1-d}{b} \\
2 & : & y_2 = 1 & \cap & y_2 & = & \frac{y_1-d}{b} \\
3 & : & y_2 = 0 & \cap & y_2 & = & \frac{y_1-c-d}{a+b} \\
4 & : & y_2 = 1 & \cap & y_2 & = & \frac{y_1-c-d}{a+b}
\end{array}
$$

Figure 2: The 6 basic configurations of Equation 4 which determine the integration ranges.

The inequalities in Equation 3 define the range over which the dummy variable, y_2, should be integrated out. Graphically, the integration must be carried over the range of y_2 defined by the lines:

$$y_2 = 0, \quad y_2 = 1, \quad y_2 = \frac{y_1 - d}{b}, \quad y_2 = \frac{y_1 - c - d}{a+b}. \tag{4}$$

For example, Figure 1 shows the required integration graphically a particular case where the integration proceeds over three ranges:

$$
\begin{array}{rcrcl}
\int_0^{\frac{y_1-d}{b}} f_{y_1,y_2}.dy_2 & : & d \leq & y_1 & < d+c \\
\int_{\frac{y_1-c-d}{a+b}}^{\frac{y_1-d}{b}} f_{y_1,y_2}.dy_2 & : & d+c \leq & y_1 & < b+d \\
\int_{\frac{y_1-c-d}{a+b}}^{1} f_{y_1,y_2}.dy_2 & : & d+b \leq & y_1 & \leq a+b+c+d
\end{array}
$$

The final result is given by:

$$
\begin{array}{rcrcl}
\frac{1}{a}\ln\left(\frac{ay_1-d+cb}{cb}\right) & : & d \leq & y_1 & < d+c \\
\frac{1}{a}\ln\left(\frac{a+b}{b}\right) & : & d+c \leq & y_1 & < b+d \\
\frac{1}{a}\ln\left(\frac{(a+c)(a+b)}{ay_1-d+cb}\right) & : & d+b \leq & y_1 & \leq a+b+c+d
\end{array} \tag{5}
$$

Note, that the specific integrals are determined by the values of the coefficients, or more precisely, the intersections of the lines defined by Equation 4. This complicates the implementation since there are 24 cases to consider (permutations of 4 intersections). However, for computational convenience these may be grouped into 6 basic arrangements as shown Figure 2 where the numbering scheme refers to intersections of the various lines. For example, orderings $\{2\ 1\ 4\ 3\}$, $\{3\ 4\ 1\ 2\}$ and $\{4\ 3\ 2\ 1\}$ all result in configurations similar to that of $\{1\ 2\ 3\ 4\}$. [2] shows that this method generates remarkably accurate estimates of PDFs from images. See Figure 5 for examples.

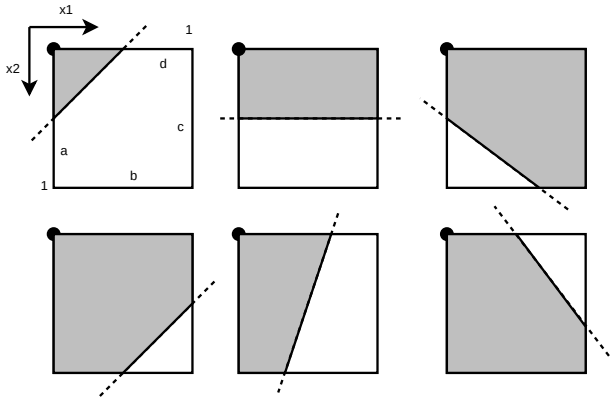


Figure 3: Polygonal ROIs give rise to partial pixels defined by lines (aside from the vertices). Six cases can arise as shown here: a-d, a-b, a-c, b-c, b-d and c-d. Where the origin of the pixel is outside the ROI the defined region is complement of that shown.

3 Partial pixels

In this section, we extend the above analysis to include partial pixels - pixels that are only partially covered by a pre-defined region of interest. In this paper, we restrict our analysis to polygonal ROIs therefore each partial pixel is split by a single line, except for the vertices which must be handled separately. The modification is, in principle at least, quite straightforward and requires simply that the ranges for x_1 and x_2 are modified appropriately. However, in practice there are a number of cases that must be handled explicitly, as illustrated in Figure 3. In all of these cases the valid range can be defined by $0 \leq x_1 \leq 1$, $0 \leq x_2 \leq m * x_1 + k$, where m is the gradient of the line and k is the y-intercept. Where the origin is outside of the ROI the $\leq$ should be replaced with $\geq$. Alternatively, the partial patch PDF could be substracted from the full patch PDF. For case b-d, a positive gradient switches the polarity of the comparison.

As an illustrative example we re-examine the case shown in Figure 1 with partial pixel case a-d with the origin inside the ROI. Up to Equation 3 the derivation is unchanged, but now the ranges are defined by:

$$\begin{aligned} f_{y_1,y_2} &= f_{x_1,x_2}(y_2, \frac{y_1 - by_2 - d}{ay_2 + c})|J| \\ &= \frac{1}{ay_2 + c} \quad 0 \leq y_2 \leq 1; 0 \leq \frac{y_1 - by_2 - d}{ay_2 + c} \leq my_2 + k \\ &= \frac{1}{ay_2 + c} \quad \begin{array}{rl} 0 \leq & y_2 \leq 1 \\ by_2 + d \leq & y_1 \leq amy_2{}^2 + (cm + ak + b)y_2 + ck + d \end{array} \end{aligned} \tag{6}$$

The effect of the partial pixel is to change the ranges of the final integration which must now be performed between a line and a quadratic (as a function of y_2) as shown in Figure 3. In this particular case, four ranges are required:

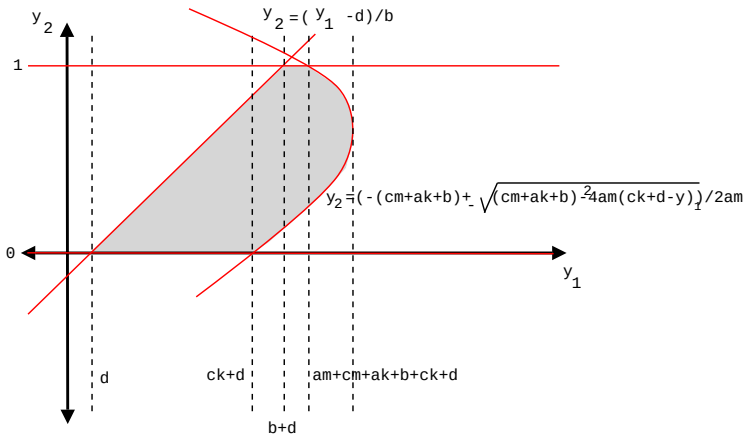


Figure 4: Partial pixels change ranges over which the final integration must be evaluated.

$$
\begin{aligned}
&\int_{0}^{\frac{y_1-d}{b}} && d \le y_1 < ck+d \\
&\int_{\frac{-B-\sqrt{B^2-4A(C-y_1)}}{2A}}^{\frac{y_1-d}{b}} && ck+d \le y_1 < b+d \\
&\int_{\frac{-B-\sqrt{B^2-4A(C-y_1)}}{2A}}^{1} && b+d \le y_1 < a(m+k)+c(m+k)+b+d \\
&\int_{\frac{-B-\sqrt{B^2-4A(C-y_1)}}{2A}}^{\frac{-B-\sqrt{B^2-4A(C-y_1)}}{2A}} && a(m+k)+c(m+k)+b+d < y_1 < y_1^{max}
\end{aligned}
\tag{7}
$$

where $A = am$, $B = cm + ak + b$, $C = ck + d$ and $y_1^{max} = y1(-B/2A)$.

Other cases can be handled in a similar fashion. Bilinear patches at the vertices of the ROI can be handled by splitting the patch into two parts, one for each edge, and setting the range of x_1 appropriately. For example, if a vertex is placed at $x_1 = 0.6$ then the ranges may be set $0 \le x_1 < 0.6$ and $0.6 \le x_1 < 1$. In all of the experiments presented in this paper, for simplicity of implementation the majority of the partial pixel cases are handled numerically.

4 Experiments

In this section, we present two sets of experiments to examine the performance of six PDF estimators: histogramming, weighted histogramming, Parzen windowing, weighted Parzen windowing, bilinear PDF and partial pixel bilinear PDF. The Parzen window methods are implemented as Gaussian filtered histograms where the sigma of the kernel is set at twice the bin size. The weighted methods adjust the counts to the histograms according to the area of the partial pixels. In all cases the aim is to estimate a 256 bin histogram,

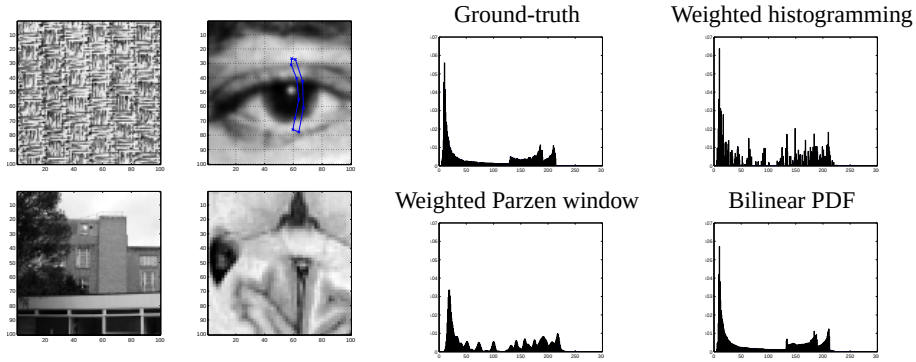


Figure 5: Left: The images used for the experiments. Right: Histograms for the image and ROI on the top right for the ground-truth,weighted histogramming, weighted Parzen windowing and bilinear PDF estimators.

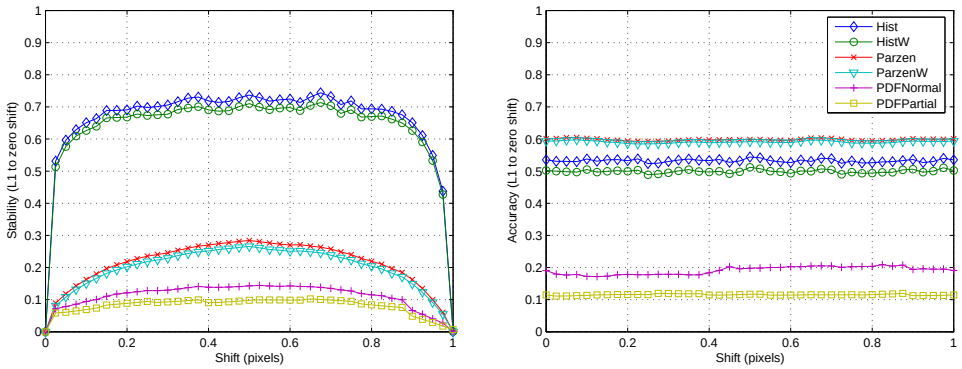
however, it should be noted that the proposed technique is not restricted to the quantisation of the image. In fact, if required the estimate can be left in continuous form and not discretised into a histogram. Mean average results are reported over 5 ROIs, 5 locations, 4 images and 41 sub-pixel shifts; a total of 5125 PDF estimations.

We measure performance in two ways: stability and accuracy. Ideally, PDF estimators should be invariant to simultaneous sub-pixel translations of the image and ROI. That is, the PDF of a ROI in an image and the shifted ROI of an equivalently shifted image should be identical[1]. In practice however, this is not the case. We measure PDF stability by calculating the $\mathscr{L}_1$ distance (sum of absolute differences of the histograms) between the zero shifted histogram for each method and that resulting from the translated image and ROI. The choice of $\mathscr{L}_1$ is somewhat arbitrary; there are many alternatives available such as Kullback-Leibler or χ^2. Our motivation is that it has a simple interpretation and its value is bounded between 0, for an identical histogram, and 2, for its complement. The sub-pixel shifts are generated by 40x upsampling the image using a FIR interpolator with a Gaussian filter ($\sigma = 0.7 pixels$). This method can generate sub-pixel shifts in steps of $\frac{1}{40}$ of a pixel.

Figure 6(a) shows the stability results for the image and ROI shown in Figure 5. All the plots start and end with zero $\mathscr{L}_1$ value corresponding to zero and one pixel shift respectively. Over all sub-pixel shifts, the proposed method is consistently the best performer (legend PDFPartial), while the worst is the standard histogram, with Parzen windowing in between. Weighting the histogram counts improves the performance for both the standard histogram and the Parzen windowing. The improvements in the bilinear spline PDF estimator brought about by considering the partial pixel effects is evident; compare the PDFNormal and PDFPartial plots.

Stability, however, is not the only factor to consider — an estimate can be stable but inaccurate. Consider increasing the sigma of the Gaussian kernel used for the Parzen window. In the limit, the resulting estimate will be very stable but would not reflect the true density. In the Machine Learning community this is often referred to as the bias-variance trade-off. To quantify this effect, we measure the accuracy of the histogram to an estimate

[1] More generally estimators should be invariant to similarity transformations of the image and ROI.



(a) (b)

Figure 6: (a) The mean stability and (b) accuracy as a function of sub-pixel translation of 6 PDF estimators.

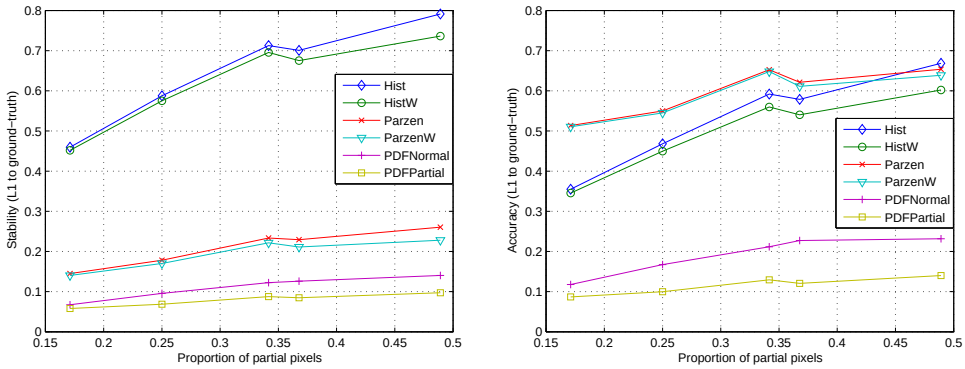
of the ground-truth. To obtain this estimate, the ROI is transformed to the resolution of the upsampled image and a histogram built. We are assuming that Gaussian filtering is a good approximation to the camera point spread function [1, 5] and 40x oversampling is sufficient to obtain a good estimate. Figure 5 shows the ground-truth and as well example histograms from the various methods.

Figure 6(b) show the corresponding accuracy results for the same image and ROI. The bias effect is clear. Both Parzen window estimators perform the worst but the proposed method is still outperform all other including the normal spline method. The accuracy margin is somewhat larger than is evident in the stability graph. It is also worth noting that the accuracy error for the proposed method is largely stable over all shifts indicating the difference arises from the interpolation and the partial pixels are handled well.

The final experiment examines the performance of the various estimators as a function of the fraction of partial pixels. The protocol is identical to the previous experiments except that it is repeated for ROIs with different fractions of partial pixels and the mean stability and accuracy recorded. The results are shown in Figure 7. For ROIs with a low proportion of partial pixels, the performance of the proposed technique and that of [2] appear to be converging one would expect. An increase in the proportion of partial pixels results in a decrease in performance for all methods, however, the slope is much less for the proposed method compared to the others.

5 Conclusions

We have presented a method that can estimate accurate and stable PDF estimates in an ROI containing partial pixels. The method has no parameters that need to be set and can generate continuous PDF estimates. In this paper we have concentrated on polygonal ROIs. This is not a limitation of the method and further work remains to explore more general parameterisations or special cases such as circles. Also, we have estimated only 1st order, or marginal, statistics. Similar derivations are possible to joint or conditional distributions or even moments such as mean, variance and kurtosis. Different interpo-



(a) (b)

Figure 7: (a) The mean stability and (b) mean accuracy as a function of the proportion of partial pixels in the ROI.

lation strategies are also worth investigating since improving this should result in more accurate estimates or simpler implementations.

6 Acknowledgements

The work was partly carried out while the Timor Kadir was at the University of Oxford under EC project CogViSys.

References

[1] K. Astrom and A. Heyden. Stochastic modelling of image acquisition, interpolation and scale-space smoothing. In *Proc. Int. Conf. on Pattern Recognition*, 1996.

[2] Timor Kadir and Michael Brady. Non-parametric estimation of probability distributions from sampled signals. Technical Report 2283/05, University of Oxford, 2005.

[3] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes.* McGraw-Hill, 2002.

[4] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: principles, algorithms and applications.* Macmillan, 1992.

[5] B. Triggs. Empirical filter estimation for subpixel interpolation and matching. In *Proc. Int. Conf. on Computer Vision*, pages 550–557, 2001.

[6] Chen Xiaohua, Daniel Rueckert, and Michael Brady. Simultaneous segmentation and registration for medical image. In *Proc. of the 7th Int. Conf. on Medical Image Computing and Computer Assisted Intervention (MICCAI04)*, pages pp. 663–670, 2004.

[7] Y. Zhang, Michael Brady, and S. Smith. Segmentation of brain MR images through a hidden Markov Random Field model and the expectation maximization algorithm. *IEEE Trans. on Medical Images*, 20(1):45–57, 2001.

[8] S.C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 18(9):884–900, 1996.