

Non-parametric Estimation of Probability Distributions from Sampled Signals

Timor Kadir and Michael Brady

Robotics Research Laboratory, Department of Engineering Science,
University of Oxford,
Parks Road, Oxford OX1 3PJ, U.K.

July 4, 2005

Abstract

This paper is concerned with the non-parametric estimation of probability distributions from band-limited and at least critically sampled signals such as images. Conventional approaches employing histograms or Parzen windows often perform unsatisfactorily since they ignore two important properties of such signals, namely that they are ordered and contain sufficient information to reconstruct the original continuous band-limited signal exactly.

Based on these observations we propose a method to estimate the density and distribution of such signals. The technique has a number of interesting and useful properties. First, that domain resolution of the estimate is independent of the number of sample points. Second, the estimate is continuous and consequently no arbitrary bin widths or smoothing kernel parameters have to be set. Third, if a suitable analytic form of the band-limited signal or equivalently the sampling pre-filter is available then the resulting density is exact (aside from signal quantisation).

1 Introduction

The accurate estimation of Probability Density Functions (PDF) or Cumulative Distribution Functions (CDF) from signals is an important step in many signal processing, learning and communication algorithms. For example, the success of probabilistic modelling techniques such as those motivated through Bayesian analysis, rely on accurate and reliable estimation of the PDF of some signal or signal class.

Whilst PDF estimation techniques and their resulting models differ greatly in their generality and efficacy, most make the following assumption: the PDF is to be estimated from samples that are Independent and Identically Distributed (IID) instances from some continuous underlying PDF. Consequently, such techniques tend to make use of only the samples themselves to approximate the PDF with the corollary that the number of samples tends to determine the quality of the final estimate.

In this paper, we consider the estimation of PDFs from Band-Limited and (at least) Critically Sampled (BL-CS) signals which constitute a wide class of signals of practical interest, such as images or speech. For such signals, the samples represent the band-limited continuous signal evaluated and quantised at some (arbitrary) spatial or temporal points. This differs from the IID assumption in several important ways. Firstly, the samples are not independent, secondly the samples are ordered and finally, that knowledge of the samples, their order and the sampling pre-filter are enough to uniquely specify the band-limited continuous signal exactly, aside from quantisation effects [6].

We propose a general non-parametric technique for accurate and stable PDF estimation and describe specific implementations for one and two dimensional signals. The proposed method has a number of useful properties. Firstly, the domain resolution of the resulting estimate is independent of the number of sample points and signal quantisation; indeed, it is continuous. 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. By usefully, we mean that new information can

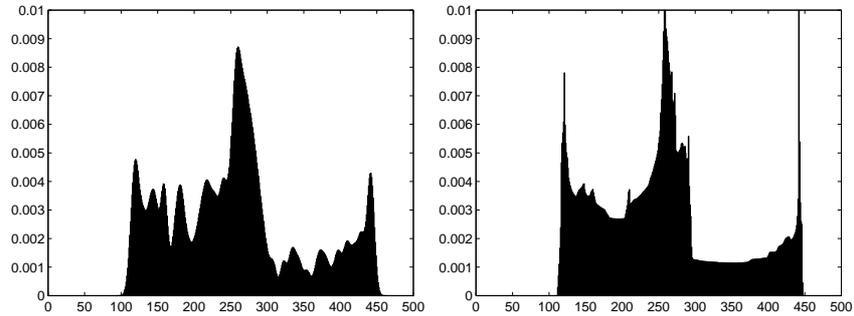


Figure 1: 512 bin histograms estimated using Parzen windowing (left) and the proposed approach (right). Both methods can produce histograms at finer binning than the quantisation of the image, but for Parzen windowing this only reveals information about the smoothing kernel.

be obtained by evaluating the PDF at finer intervals in contrast to, say Parzen windows where evaluation at points finer than the quantisation reveals only information about the smoothing kernel; see Figure 1.

Secondly, no arbitrary bin widths or smoothing kernel parameters have to be set; this is an open issue with many conventional non-parametric techniques such as histogramming or Parzen windowing. Thirdly, the samples are not assumed to be independent. Finally, if certain criteria are met, then resulting density is the exact PDF of the band-limited signal (upto quantisation).

The paper is structured as follows. After briefly discussing related work in Section 2, we motivate our approach by examining the histogramming process for a synthetic image. In Section 3 we define the measure of PDF quality used in the experiments. Sections 4 and 5 detail our approach and finally in Section 6 we present our results.

2 Related work

PDF estimators generally fall into one of three categories, parametric, non-parametric and semi-parametric. Parametric techniques are suitable where a particular form of function can be

assumed due to some application specific reasons. For example, Rician and Rayleigh functions are often used in ultrasound signal processing applications.

Of the non-parametric techniques probably the simplest and most widely used method is the histogram. Its limitations, principally the requirement to define 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. Parzen windowing avoids arbitrary bin assignments and leads to smoother PDFs, however, a suitable kernel shape and size must be chosen. It has been noted that conventionally this choice has been somewhat arbitrary and largely driven by aesthetics [8], although some work has been done on defining systematic methods for selecting kernel sizes; see [1, 10, 7]. Other, more exotic methods such as Wavelet density estimators [2] have been also been proposed.

Semi-parametric techniques such as Gaussian Mixture Models offer a useful compromise between these two approaches whereby the superposition of a number of parametric densities are used to approximate the underlying density.

Common to all such methods is the assumption that the samples from which the PDF is estimated are IID. Clearly for the BL-CS class of signals under consideration here, they are not independent. This is widely acknowledged, though frequently ignored. While it is not inaccurate to consider the samples arising from some unknown distribution, doing so tends to ignore their generative process; ordered but arbitrary point evaluations of the underlying band-limited signal. By arbitrary, it is meant that they are unrelated to the signal but instead determined by the sensor.

This is a critical point and is illustrated in the following example which examines the process of building a histogram from an image. The argument generalises to all methods that treat BL-CS signals in this manner.

Figure 2(b) shows the histogram of the synthetic image shown in (a). In this simple case the method seems to have worked well with the two modes in the histogram capturing the two principle populations in the image.

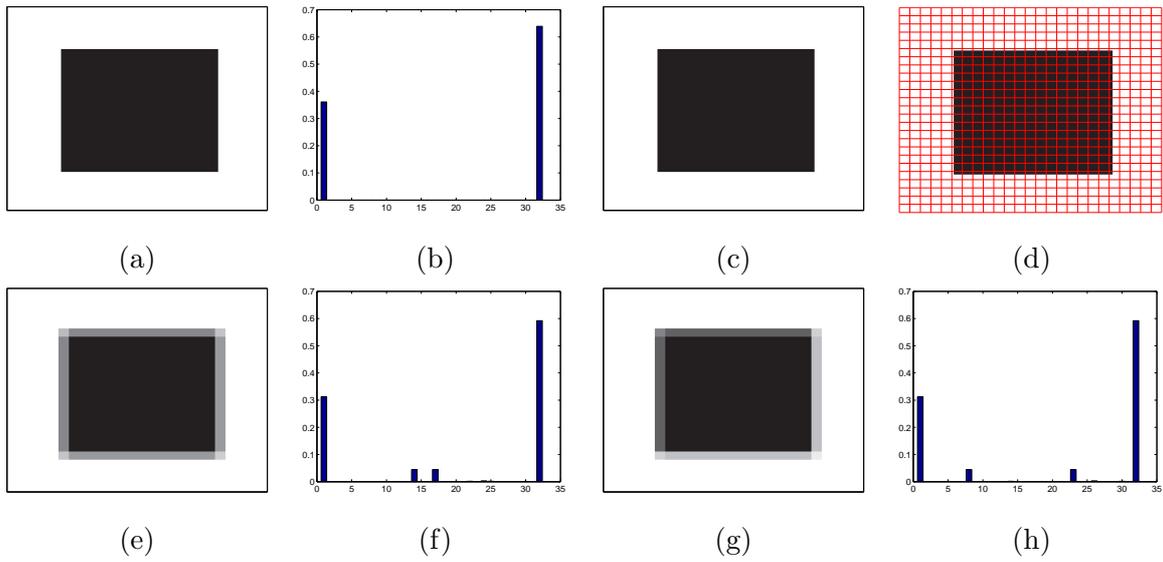


Figure 2: Estimating the PDF of a synthetic image (a) using a histogram (b). Including the sensor in the process: the real world scene, (c), is imaged by a finite resolution camera, (d), with the resulting image and histogram shown (e) and (f) respectively. Images (g,h) repeat the steps with a sub-pixel translation of the camera position.

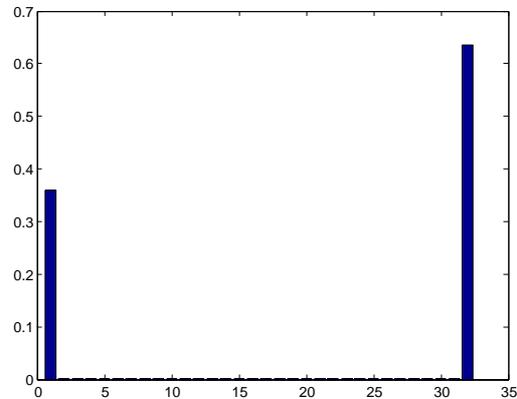


Figure 3: The histogram obtained by upsampling the image 50 times using bilinear interpolation.

However, in practice this is not what happens. Typically, the image we wish to process has been acquired by some finite resolution sensor, say a CCD camera. Figures 2(c-f) illustrates a simple model of this process. Here, the CCD is modelled as a grid of square cells where each cell measures the average intensity over its area. The resulting histogram is shown in (f). It is evident that the acquisition process has produced two ‘spurious’ peaks in between the main two modes present in the real scene. These are a result of the ‘partial volume’ pixels found at the border of the black square. Worse still, these peaks move; Figures 2 (g-h) show the process repeated but this time with a (simulated) sub-pixel translation of the camera position.

Note that this rather unpleasant behaviour exists in the absence of noise or quantisation errors. Furthermore, in this trivial example the only source of the ‘spurious’ peaks are at the border of the black square; in general all pixels contain measurements from a mixture of scene elements. Another limitation is that the number of pixels determines the number of counts in the histogram and hence the reliability of the PDF estimate.

3 PDF estimate quality

Evidently, histogramming is not invariant to translations and by extension rotations. Such invariance is desirable for many applications and has been used as a means of evaluating performance of low-level operators e.g. [3, 4]. We adopt this approach for assessing PDF estimate quality. The advantage of such an approach is that it very general.

We define quality in two ways: stability and accuracy. Stability is measured as the variation of the PDF under the geometric Euclidean group of transformations, namely translation and rotation. In one dimension, this reduces to a shift. We choose \mathcal{L}_1 as the distance metric. It is acknowledged this is somewhat arbitrary as there are many alternative ways in which to compare PDFs, e.g. Integrated squared error, Kullback-Leibler. Our choice is primarily motivated by the fact that \mathcal{L}_1 is bounded between 0 and 2 irrespective of the number of bins and each bin difference is added uniformly. However, as is shown in the experiments in Section 6 such a simple measure does not necessarily reflect the accuracy of the PDF shape.

In cases where the ground-truth PDF is known or can be estimated as is done in Section 6, we examine the accuracy of the PDF estimate. This is done by measuring the distance between the estimate and the ground-truth; again we choose to use \mathcal{L}_1 as the distance metric.

4 Order matters

It is clear that the cause of the unstable modes in the histograms in Figure 2 are the pixels on the boundary of the black square. More generally, in images of real scenes most pixels would typically exhibit such behaviour since each typically represents a mixture of scene elements.

The assumption that a statistical sampling process is responsible for the samples leads to the conclusion that little can be done about their inherent uncertainty. The best that can be done is to consider the uncertainty as a source of noise or make further assumptions on the form or smoothness of the source PDF.

However, by assuming a signal sampling process it is apparent that the problem can be overcome. The Whittaker-Shannon sampling theory 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.

Therefore, one straightforward way by which to overcome the instability in the samples is to generate more by upsampling the signal and using these to building a histogram. Note, the only factor which controls the goodness of the histogram are the number of samples we choose to take. Moreover, we are free to choose an arbitrarily small bin size given sufficient samples.

Figure 3 shows the 32 bin histogram resulting from a 50 times upsampled (using bilinear interpolation) version of the synthetic image shown in Figure 2(a). Once again the two peaks corresponding to the two main populations present, but all the bins in between these have non-zero values. Since bilinear interpolation was used here, these correspond to the interpolated rectangular patches at the border of the black and white regions. Moreover, these are much more stable under translation than with histogramming alone.

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 [9].

For such a scheme to be practical the degree of oversampling must be computationally feasible. To investigate this issue, we examine the accuracy and stability of histograms generated from samples from one period of the function, $x(t) = \sin(t + \phi)$, in the following manner. Equally spaced samples are selected from a sine function offset by a random phase shift and a histogram built. This process is repeated 400 times with different phase offsets and the mean \mathcal{L}_1 distance between the histograms is recorded. The results are shown in Figure 4 for different numbers of samples and histogram bins. The mean accuracy of the histograms is also recorded. This is the mean \mathcal{L}_1 distance to the ground-truth PDF calculated analytically and the results shown in Figure 5.

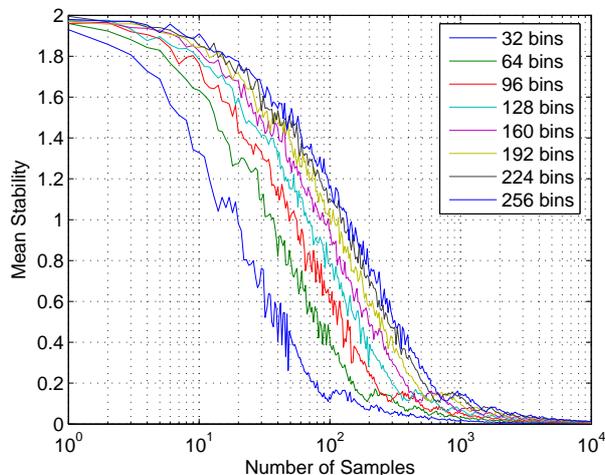


Figure 4: The stability of histograms of one sine period, as a function of sample number for different numbers of bins.

As one would expect, an increase in sample number improves the accuracy and stability of the PDF estimate. Also, fewer bins require fewer samples to obtain a given stability value; this is a well-known rule of thumb. However, of importance here is the relatively high number of samples required to achieve even a modest stability performance. For example, a 32 bin histogram requires at least 150 samples for a 0.1 stability value, equivalent to an average 5% change in the histogram and 754 samples for mean stability of 0.01, equivalent to a 1% change in the histogram. Oversampling signals to such a degree and building histograms is somewhat impractical, especially for higher dimensional signals such as images or 3D volumes. However, the results serve to demonstrate that assuming the Nyquist conditions are met, the quality and resolution of the PDF estimate is independent of the original number of samples (ignoring the effects of quantisation).

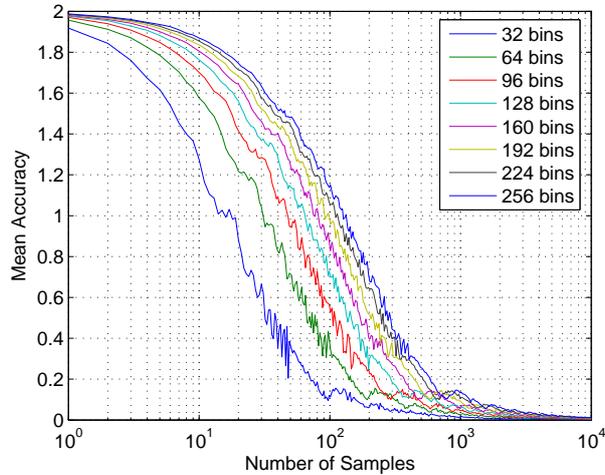


Figure 5: The accuracy of histograms of one sine period, as a function of sample number for different numbers of bins.

5 PDFs from piecewise functions

Many interpolation schemes proceed by fitting piecewise functions to the signal samples and re-sampling these at the required points. However, since our 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 alternatiely the distribution method. A detailed presentation of this background may be found in [5].

To use the Transformation formula the function must be solvable and its derivative cal-

culated. For the CDF, the resulting PDF must be integrated. In 2D a dummy variable is introduced and the function solved. A Jacobian must be obtained and an extra integration performed to remove the dummy variable. In all cases, the function must be monotonic in the range under consideration and any exceptions to this handled explicitly.

The question remains as to the form of piecewise function: we choose splines and manipulate them in their polynomial form. The splines can be fitted to the samples using interpolation or approximation procedure. We assume the samples are essentially noise free and choose the former. In this case, the splines pass through the sample points exactly. Polynomial splines, at least at low order and in one or two dimensions, fulfil the above requirements.

Clearly, there are many choices for the interpolating function; for example a sinc-based interpolation scheme would probably more accurately approximate the true signal. Our choice is primarily motivated by mathematical tractability and computational speed. Splines also have an interesting interpretation when variations in order are considered. At one end a zero order spline corresponds to a piecewise constant interpolation which is exactly the same as the conventional histogramming procedure. At the other extreme an infinite order spline corresponds to a Gaussian interpolation kernel which makes an interesting link with Scale-space ideas. In between, the order of the spline controls the approximation of the local function in a truncated Taylor expansion manner.

The algorithm to calculate the PDF/CDF and histogram 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 in any further detail. The following sections detail steps 2 and 3 for three cases: 1D linear, 1D quadratic and 2D Bilinear. We adopt the convention

that the piecewise spans start at zero and are of unit length. Consequently, the PDF of the domain variable is unity over its range: $f_x(x) = 1, 0 \leq x \leq 1$.

5.1 1D - Linear

In this case, each piecewise section is represented as a polynomial of the form, $y(x) = ax + b$, i.e. straight lines. The PDF is given by:

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

This has a straightforward and intuitive implementation. The PDF is simply the super-position of piecewise constant sections of magnitude $\frac{1}{|a|}$ between domain values b and $a+b$. This corresponds to adding all values between consecutive sample points in equal proportion. The CDF is given by:

$$F_y(y) = \frac{y}{|a|} \quad b \leq y \leq a+b \quad (2)$$

5.2 1D - Quadratic

Each span is represented by a polynomial of the form, $y(x) = ax^2 + bx + c$. The derivation of the PDF is slightly complicated by the fact that quadratics are in general non-monotonic. Such cases can be handled by either by detecting points at which the curve becomes monotonic and modifying the PDF calculation appropriately or by resampling the spline spans such that each section is strictly monotonic. The latter approach could provide a fast implementation of the PDF estimation stage at the expense of complexity of the spline fitting step. However, our interest here is primarily the design of a general PDF estimation algorithm, so we choose the former approach.

$$x(y) = \frac{-b \pm \sqrt{b^2 - 4a(c-y)}}{2a} \quad (3)$$

Due to the non-monotonicity of quadratic, the inverse quadratic function will in general be multi-valued, as indicated by the two roots in Equation 3. However, within the spline span it

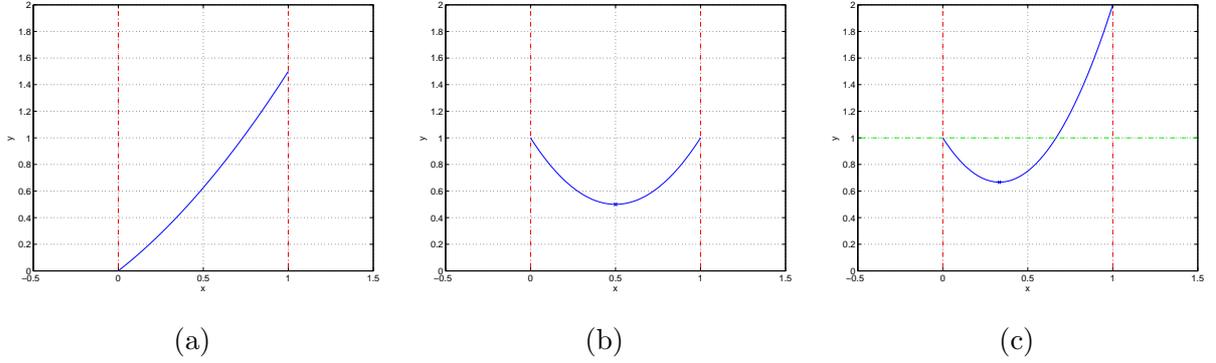


Figure 6: As a function of y , a 1D quadratic can exhibit (a) single values, (b) multiple values or (c) a combination of both across the span of the spline.

may exhibit single values, multiple values or a combination of both, as illustrated in Figure 6. Fortunately, since quadratics are symmetric about the extrema point, multiple valued sections can be accounted for by considering only one root and multiplying the PDF by two in that part. For each monotonic section of the spline the PDF can be calculated as follows:

$$\begin{aligned}
 f_y(y) &= \frac{1}{|2ax + b|} f_x(x) \\
 &= \frac{1}{|\sqrt{b^2 - 4a(c - y)}|} \quad c \leq y \leq a + b + c
 \end{aligned} \tag{4}$$

The CDF is given by:

$$F_y(y) = \frac{|\sqrt{b^2 - 4a(c - y)}|}{2a} \quad c \leq y \leq a + b + c \tag{5}$$

5.3 2D - Bilinear

The derivation for the two dimensional case requires the introduction of a dummy function and variable which must be integrated out in the final step, denoted x_2 in the following:

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

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

$$\begin{aligned}
|J| &= \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} \\
&= \begin{vmatrix} 0 & 1 \\ \frac{-1}{(ay_2+c)} & \frac{-b}{(ay_2+c)} + \frac{(-by_2-y_1+d)a}{(ay_2+c)^2} \end{vmatrix} \\
&= \frac{1}{ay_2+c}
\end{aligned} \tag{6}$$

The joint PDF between y_1 and y_2 is given by:

$$\begin{aligned}
f_{y_1, y_2} &= f_{x_1, x_2}(y_2, \frac{y_1 - by_2 - d}{ay_2 + c}) |J| \\
&= \frac{1}{ay_2 + c} \quad \begin{matrix} 0 \leq y_2 \leq 1 \\ by_2 + d \leq y_1 \leq y_2(a + b) + c + d \end{matrix}
\end{aligned} \tag{7}$$

The inequalities in Equation 7 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:

$$y_2 = 0, \quad y_2 = 1, \quad y_2 = \frac{y_1 - d}{b}, \quad y_2 = \frac{y_1 - c - d}{a + b}. \tag{8}$$

For example, Figure 7 shows the required integration graphically a particular case where the integration proceeds over three ranges:

$$\begin{aligned}
\int_0^{\frac{y_1-d}{b}} f_{y_1, y_2} \cdot dy_2 &: \quad d \leq y_1 < d + c \\
\int_{\frac{y_1-c-d}{a+b}}^{\frac{y_1-d}{b}} f_{y_1, y_2} \cdot dy_2 &: \quad d + c \leq y_1 < b + d \\
\int_{\frac{y_1-c-d}{a+b}}^1 f_{y_1, y_2} \cdot dy_2 &: \quad d + b \leq y_1 \leq a + b + c + d
\end{aligned} \tag{9}$$

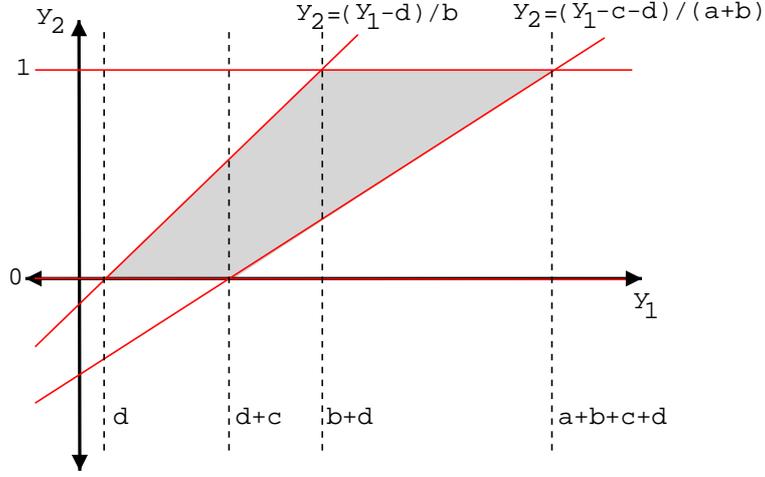


Figure 7: The integration ranges shown graphically for a particular configuration of the bilinear spline. Shown here is the case for $\{a, b, c, d\} > 0$ and $b > c$.

The final result is given by:

$$\begin{aligned}
 \frac{1}{a} \ln \left(\frac{ay_1 - d + cb}{cb} \right) &: d \leq y_1 < d + c \\
 \frac{1}{a} \ln \left(\frac{a+b}{b} \right) &: d + c \leq y_1 < b + d \\
 \frac{1}{a} \ln \left(\frac{(a+c)(a+b)}{ay_1 - d + cb} \right) &: d + b \leq y_1 \leq a + b + c + d
 \end{aligned} \tag{10}$$

Note, that the specific integrals are determined by the values of the coefficients, or more precisely, the intersections of the lines defined by Equation 8. 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 8 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 $\{1\ 2\ 3\ 4\}$.

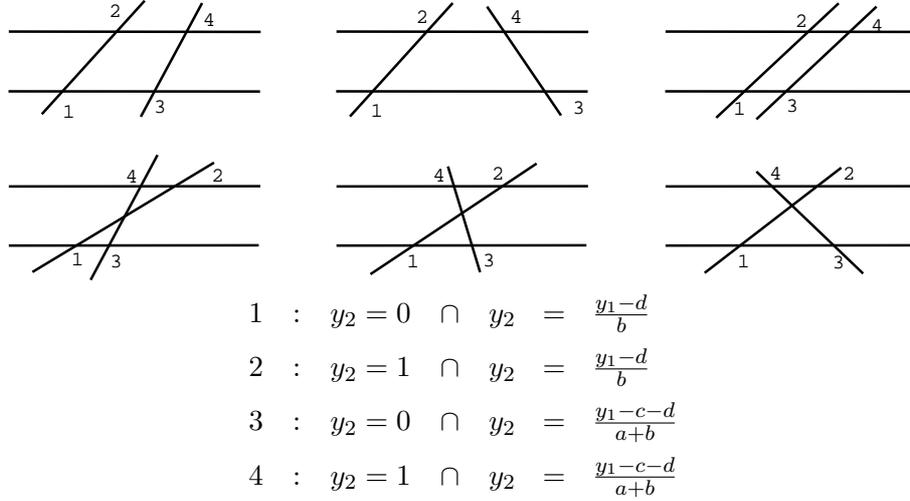


Figure 8: The 6 basic configurations of Equation 8 which determine the integration ranges.

5.4 Joint PDFs

Estimation of joint PDFs can proceed in manner similar to that of the 2D case described in the previous section. However, instead of the dummy variable and function we use the piecewise polynomials of the second signal and do not perform the final integration step. For clarity of presentation we illustrate the joint case using the 1D linear case from which the bilinear case follows.

$$\begin{aligned}
 y_1(x_1, x_2) &= ax_1 + b & y_2(x_1, x_2) &= cx_2 + d \\
 x_1(y_1, y_2) &= \frac{y_1 - b}{a} & x_2(y_1, y_2) &= \frac{y_2 - d}{c}
 \end{aligned}$$

The Jacobian in this case is:

$$|J| = \begin{vmatrix} \frac{1}{a} & 0 \\ 0 & \frac{1}{c} \end{vmatrix} = \frac{1}{ac} \quad (11)$$

The joint PDF between y_1 and y_2 is given by:

$$f_{y_1, y_2} = f_{x_1, x_2} \left(\frac{y_1 - b}{a}, \frac{y_2 - d}{c} \right) |J|$$

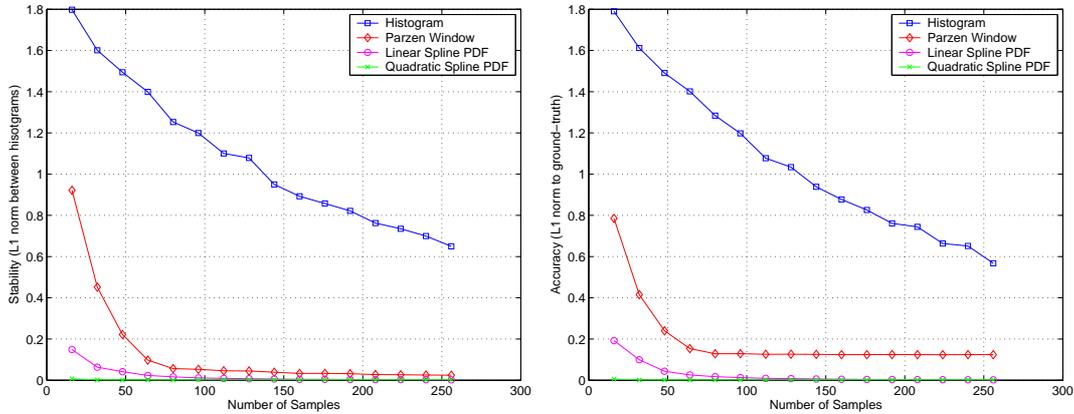


Figure 9: Comparing the quality of conventional histograms, Parzen windows, linear and quadratic spline PDF estimators at estimating the sine PDF as a function of sample number. The top graph shows the stability and the bottom graph the accuracy.

$$= \frac{1}{ac} \quad b \leq y_1 \leq a + b; d \leq y_2 \leq c + d \quad (12)$$

In some cases the variable of one function is a transformation of another. For example, in the maximisation of Mutual Information (MI) [11] algorithm one signal, typically an image, is some assumed to be some unknown transformation of the other. Conventionally, an interpolation step is necessary to calculate the joint PDF for arbitrary transformations which is computationally slow and introduces errors. Utilising this property avoids this.

6 Experiments

The first experiment compares the PDF estimate stability and accuracy for four techniques: conventional histogram, Parzen window (Gaussian $\sigma = 4$), linear spline and quadratic spline. The objective is to estimate a 256 bin histogram from one period of the sine function, $127.5 \sin(x + \phi) + 127.5$, using a pre-determined number of samples. The ground-truth histogram can be calculated from $F_y(y) = \frac{\arcsin(y)}{\pi} + \frac{255}{2}$. Figure 9 shows the results. Clearly, the spline methods

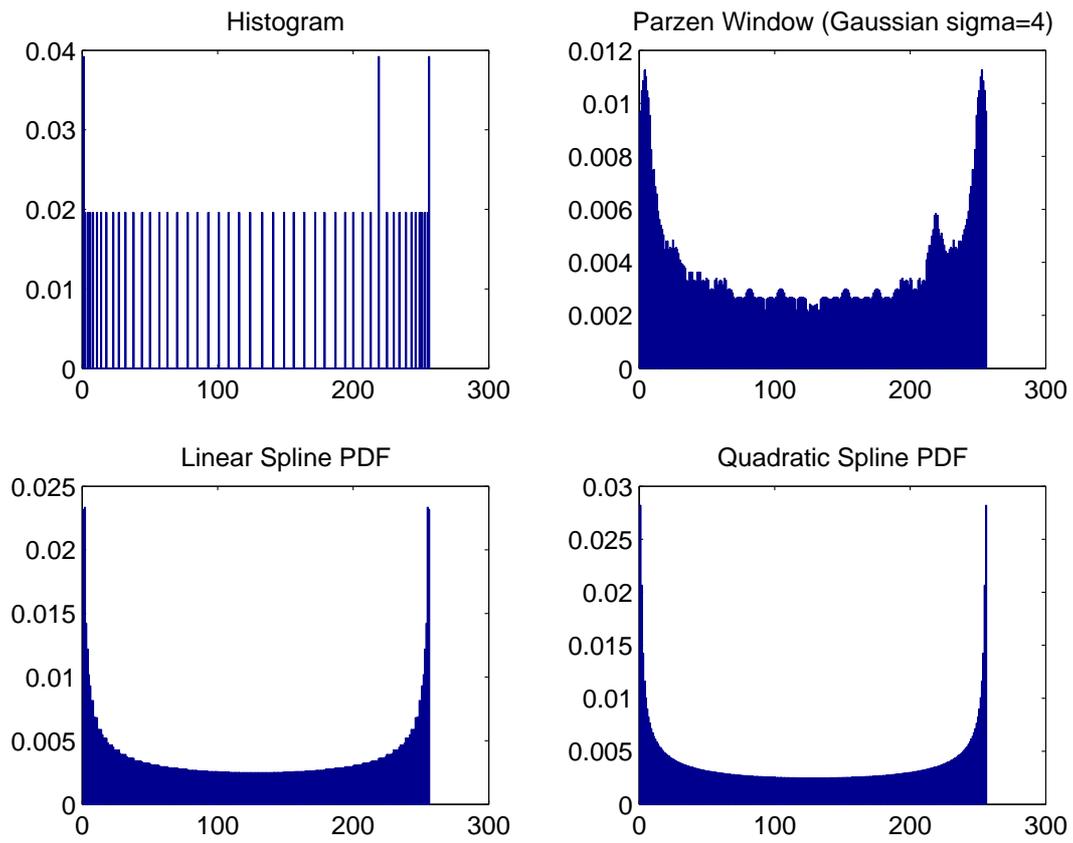


Figure 10: Examples of the PDF estimates of the sine function for four methods at 50 samples.

outperform the conventional methods by some significant margin, however the Parzen window method performs surprisingly well considering it is just a Gaussian smoothed histogram. The reason is that in this particular case smoothing across bins is a good thing to do since in a sine function adjacent locations vary smoothly. In general this is not true. Moreover, whilst the stability of the Parzen window estimator converges to zero with increasing sample number, the accuracy does not. This is because the Gaussian smoothing causes a distortion in the PDF which cannot be removed despite an increase in the number of samples. This confirms the rule of thumb that the width of the Parzen kernel should be inversely related to the number of samples available. It should also be said that the sine function also favours the quadratic spline method since a quadratic can approximate a sine quite well with only a few samples. Figure 10 shows examples of the PDFs generated by each method with 50 samples. It is interesting to note the sharp peaks at 0 and 255. These correspond to singularities in the PDF caused by the zero gradients in the sine function at these points. In fact, the PDF of any smooth non-monotonic function will exhibit such singularities. This is somewhat surprising given that it is often assumed that smooth functions give rise to smooth PDFs - an assumption upon which many kernel PDF estimators are based.

Next we examine the performance of the proposed technique for real images as follows. The test image is divided into non-overlapping blocks of a fixed size and the PDF estimated in each. However, since the image statistics are unlikely to be stationary we cannot use the stability as in the previous section. Instead, we calculate the PDF accuracy to the estimated ground-truth PDF. This is found by histogramming an interpolated version of the image as discussed in Section 4. To avoid biasing the test towards the proposed method, polynomial spline interpolators are not used; we use a FIR interpolator - the image is upsampled with zeros then filtered with a Gaussian ($\sigma = 0.7$ pixels). The results are shown in Figure 11 for patches of size 3x3, 5x5, 10x10, 15x15 and 20x20 and for 32 and 256 bin histograms. The bilinear PDF estimator outperforms the other methods in all cases. At the largest patch size, 400 samples are available yet the performance of conventional methods is still somewhat unsatisfactory; see Figure 12.

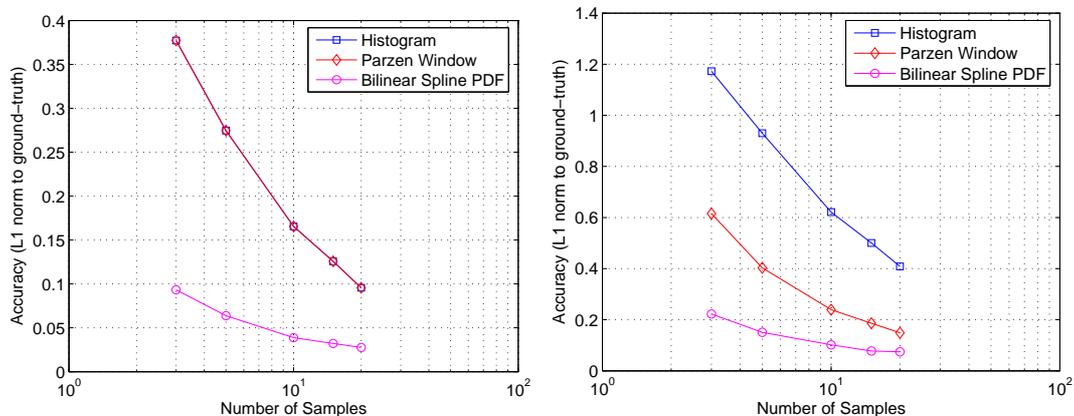


Figure 11: Accuracy of 32 (top) and 256 bin (bottom) histograms for the histogramming, Parzen windowing and bilinear spline PDF estimators for real images

7 Discussion and Future Work

What are the practical consequences of this work? One of the key benefits is that it enables the accurate estimation of PDFs from very small image patches which is useful for methods such as [3] which are at the mercy of the underlying PDF estimator. Moreover, no binning or smoothing parameters have to be set. The method can be applied to filtered, e.g. high-pass images.

If the spline function exactly replicates the true interpolation of the sampling pre-filter then the resulting PDF is exact. In general however, this is unlikely to be the case. One approach could be to upsample the signal using an accurate interpolator and then use a bilinear or quadratic PDF estimator at this higher sampling rate. Non-uniform spatial sampling can be used to reduce the overhead associated with higher sampling rates.

References

- [1] D. Comaniciu, V. Ramesh, and P. Meer. The variable bandwidth mean shift and data-driven scale selection. In *Proc. Int. Conf. on Computer Vision*, pages 438–445, 2001.

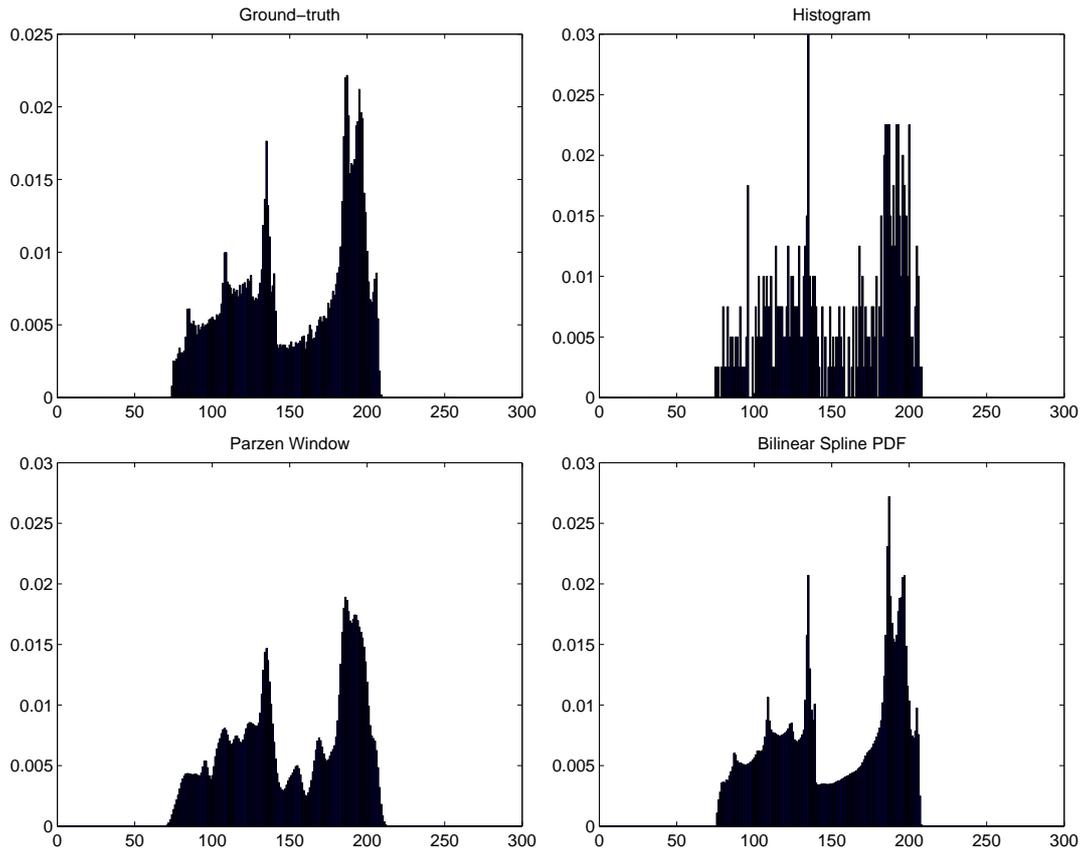


Figure 12: Example PDFs from the real image experiments. \mathcal{L}_1 distances to the ground-truth are 0.35, 0.14 and 0.09 for the histogramming, Parzen and bilinear spline methods.

- [2] D.L. Donoho, I.M. Johnstone, G. Kerkyacharian, and D. Picard. Density estimation by wavelet thresholding. *The Annals of Statistics*, 24(2):508–539, 1996.
- [3] T. Kadir, A. Zisserman, and M. Brady. An affine invariant salient region detector. In *Proc. European Conf. on Computer Vision*, 2004.
- [4] K. Mikolajczyk and C. Schmid. An affine invariant interest point detector. In *Proc. European Conf. on Computer Vision*, 2002.
- [5] A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.
- [6] J. G. Proakis and D. G. Manolakis. *Digital Signal Processing: principles, algorithms and applications*. Macmillan, 1992.
- [7] J. S. Simonoff. *Smoothing methods in Statistics*. Springer-Verlag, 1996.
- [8] G.R. Terrell. The maximal smoothing principle in density estimation. *Journal of the American Statistical Association*, 85:470–477, 1990.
- [9] B. Triggs. Empirical filter estimation for subpixel interpolation and matching. In *Proc. Int. Conf. on Computer Vision*, pages 550–557, 2001.
- [10] B.A. Turlach. Bandwidth selection in kernel density estimation: A review. Discussion Paper 9317, Institut de Statistique, Universite Catholique de Louvain, 1993.
- [11] P. Viola and William M. Wells III. Alignment by maximization of mutual information. In *Proc. Int. Conf. on Computer Vision*, pages 16–23, 1995.