# Nonparametric Density Estimation using Wavelet Transformation and Scale-space zero-crossing reconstruction

*Wu Ying* *, *Li Bin* ** ,*Ping Fan Yan* *
* Pattern Recognition & Artificial Control Laboratory Automation Dept. , Tsinghua Univ.
E_Mail address : ys4wy@info.au.tsinghua.edu.cn
** CIMS Center Room 609 , Automation Dept. ,Tsinghua Univ.
100084 Beijing , P.R.China

**Abstract:** Parzen window method requires relatively larger sample set , and the result of estimation is subject to the selection of window width , so the Parzen window method cannot get good estimation to complex distribution that needs multi-resolution . A novel approach ,which is based upon wavelet transformation is presented . On the viewpoint of wavelet transformation , the result of Parzen window method is only the smoothing approximation of p.d.f. . Scale space filter technology is used to get rid of the noise produced by smaller sample set . Six simulations show out that this method can successfully solve the dilemma of estimating of p.d.f. by small sample set and complex distribution .

**Key word** : density estimation , the Parzen windows , wavelet transformation , zero-crossing reconstruction

## I. Introduction

Density estimation is an important aspect of signal processing . In statistical signal processing , probability density function (p.d.f.) play an key role in the Bayes estimation which is the base of statistical pattern recognition. However, where can we get the probability density function? In many cases, we assume that data are produced from Gaussian distribution which is universal in our life .But this assumption is not always tenable . So we have to find the p.d.f. from the data .

Parzen window method , one of the traditional approaches to p.d.f . estimation, requires relatively larger sample set , and the result of estimation is subject to the selection of window width , so the Parzen window method cannot get good estimation to complex distribution that needs multi-resolution .

Wavelet transformation is a multiresolution analysis which provides an elegant frame in which both smooth signal and suddenly changing signal can be described in a integrated form .Wavelet decomposes signal into smoothing part and detail parts in many scales .Unlike Fourier transformation and windowed Fourier transformation , wavelet does not need to select the width of window .The feature of the given signal can be described under specific resolution.

The motivation of our work is to use the multiresolution idea of wavelet analysis .Because the Parzen window method need to select window width ,the introduction of multiresolution will be overcome the difficulties of acquiring the knowledge to determine a appropriate width.

Section two discusses the Parzen window method to density estimation. Section three presents the motivation of our work and describes the viewpoint of wavelet transformation to the Parzen window method .Section four explains the approach to density estimation using wavelet transformation .Section five gives some simulation result to illustrate the efficiency of this method .Some conclusions are made in section six .

## II . Parzen windows Approach to Density estimation

The Parzen window[1] approach to estimating densities can be introduced by assuming that the region

This work is supported by the Chinese program,the National Key Project for Fundamental Research in China,Grant NSL 92097.

0-7803-2912-0

$R_n$ is a d_dimensioned hypercube . If $h_n$ is the length of an edge of that hypercube , then its volume is given by $V_n=h_n^d$ . If we define the following window function :

$$\Phi(u)=\begin{cases}1, |u_j|\le 1/2, j=1,\ldots,d.\\ 0, otherwise\end{cases} \quad (1)$$

then , $k_n$ , the number of samples falling in the hypercube , is given by

$$k_n=\sum_{i=1}^{n}\Phi\left(\frac{x-x_i}{h_n}\right) \quad (2)$$

and we obtain the estimation :

$$p_n(x)=\frac{1}{n}\sum_{k=1}^{n}\frac{1}{V_n}\Phi\left(\frac{x-x_i}{h_n}\right) \quad (3)$$

Generally , we allow a more general class of window functions other than the hypercube window defined by Eq(1) , if we require that $\Phi(u) > 0$ and $\int\Phi(u)\,du=1$ .

$h_n$ affects the effect of estimation . If $h_n$ is very large , $p_n(x)$ is very smooth , "out_of_focus" estimation for p(x) . On the other hand , if $h_n(x)$ is very small , $p_n(x)$ is an erratic , "noisy" estimation of p(x) . With limited number samples , the best we can do is to seek an acceptable compromise .

With enough samples , we are essentially insured of convergence to an arbitrarily complicated unknown density . But the number of samples needed may be too large to get . This limitation severely restricts the practicability of such nonparametric procedures .

## III . The Viewpoint of Wavelet to Density estimation

### 1 . $p_n(x)$ is the smoothing approximation of p(x) at certain resolution .

From Eq(3) , when $n\to\infty$ ,

$$p_n(x)=\int\frac{1}{V_n}p(t)\Phi\left(\frac{t-x}{h_n}\right)dt \quad (4)$$

If we regard $\Phi(u)$ as the scale function of wavelet transformation , where $\Phi(u)$ satisfies :

$\int\Phi(t)\,dt=1$ and $\Phi(t)\ge 0$ ,

$p_n(x)$ is the smoothing approximation of p(x) at scale $h_n$ .

### 2 . Constructing the wavelet transformation of p(x)

Because p(x,hn) is the smoothing approximation , we can find a space $W_{h_n}$ , which satisfies $V_{h_n}+W_{h_n}=V_{h_n/2}$ and $V_{h_n}\perp W_{h_n}$ , where $V_{h_n}=span\{\frac{t-x}{h_n})\}$ . We use dyadic wavelet . $h_n=2^n$ , that is , $V_n+W_n=V_{n-1}$ and $V_n\perp W_n$ .

Because $\{\Phi(\frac{t-x}{h_n})\}$ is orthonomal bases of $V_{h_n}$ .

We can find $\{\Psi(\frac{t-x}{h_n})dt\}$ , which is wavelet bases .

From Eq(4) we can easily construct

$$WT_{x,h_n}p(x)=\frac{1}{V_n}\int p(x)\Psi\left(\frac{t-x}{h_n}\right)dt \quad (5)$$

where $\Psi(u)$ is wavelet function , which is $\odot\,\Phi(u)$ .

In this case , we get

$$WT_{x,h_n}\Phi(x)=\frac{1}{N}\sum_{i=1}^{n}\frac{1}{V_n}\Psi\left(\frac{x-x_i}{h_n}\right) ,$$

the discrete form .

## IV . Wavelet approach to density estimation

In our method , we get the discrete form of wavelet transformation of p(x) by using wavelet window , which is nearly the same as the Parzen window method with only the different window form . For example , if we use normal Gaussian window , then we use Marr function as wavelet window which is orthogonal to the Gaussian function . For convenience , we use dyadic wavelet , and we use 4 to 5 levels to construct wavelet decomposition . At the largest resolution , we use

Parzen window as the smoothing approximation . After that , we construct filters h(n) and g(n) to construct p(x) .

$$h(n)=\frac{1}{\sqrt{2}}\int\Phi(\frac{t}{2})\Phi^{*}(t-n)dt=<\Phi_{10},\Phi_{0n}> \quad (6)$$

$$g(n)=\frac{1}{\sqrt{2}}\int\Psi(\frac{t}{2})\Phi^{*}(t-n)dt=<\Psi_{10},\Phi_{0n}> \quad (7)$$

Because of the relatively small sample set , the estimation of p(x) is not very smooth . When we observe this in our simulation experiments , we find a way to eliminate the noise . It is assumed that the amplitude of the noise decrease as resolution level increase . So we use tracing zero-crossing approach which is described in Mallat's paper[2] to get rid of the noise .Following is our algorithm:

(1)search the maximum point of each wavelet decomposition, then use the maximum point set to represent the wavelet decompositions.

(2)using tracing technic[2] to eliminate the noise in the sample data.

(3)use iterate project method in Mallat's paper[2] to get the refined wavelet decomposition.

(4)reconstruct the distribution p(x) from the refined wavelet decomposition.

Six simulations show that even with 100 or less samples , this method can achieve relatively much better results than Parzen window method .

## V . Simulation result

In our experiment, we use several typical probability density functions. We used Gauss distributed function, uniform distributed function (including very smooth distribution and very erratic one), and the miscellany of the former two. And we use only about 100 samples to test the efficiency of our approach to small sample set .

The result of miscellany of Gauss distribution and uniform distribution is presented below.

As shown in Fig 1., the probability distribution is a miscellany of a gauss function and a uniform distribution .The probability density function is:

$$f(x)=\begin{cases}\frac{1}{2}\cdot\frac{1}{\sqrt{2\pi}\cdot 20}\exp(-\frac{(x-150)^2}{2\times 20^2})\ldots\ldots\ldots 0\le x\le 300\\ \frac{1}{2}\cdot\frac{1}{150}\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 350\le x\le 500\\ 0\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots otherwise\end{cases}$$

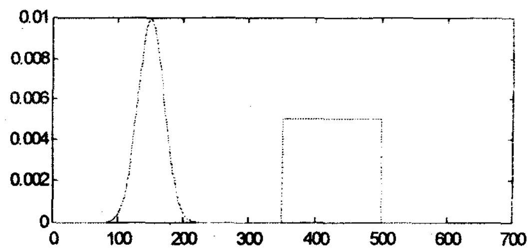


Fig 1. The probability density function

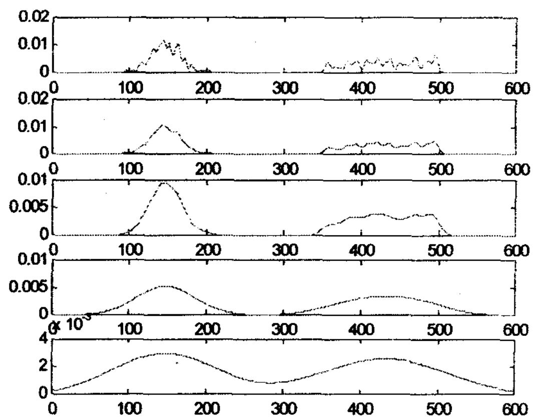

Fig 2.The result of Parzen method

When we use Parzen method, the results are shown in Fig 2. The first wave is a result of a Parzen approach with window width h=32. Others are results of Parzen approach with h doubling the upper one. We can see from the graphs that the efficiency of Parzen method depends on it's window width ,that is ,the resolution. Then the window width is too narrow, the estimation is erratic. On the contrary, the estimation may be ignorant of the rapid change of probability distribution if the window size is too large . Although the problem of erratic estimation can be solved with large number of sample set, the requirement is unrealistic in general condition.

Fig 3. refer to the wavelet approach without scale space filter.The upper three subgraphs are the dyadic wavelet decomposition of sample set. The fourth is the smooth approximation of the sample set. The last one is the result of wavelet reconstruction which is fairly well without consideration of the window width. In the wavelet estimation, the information in different frequencies of sample set is employed to reconstruct the probability distribution, so the result is more reliable and accurate than the Parzen method.

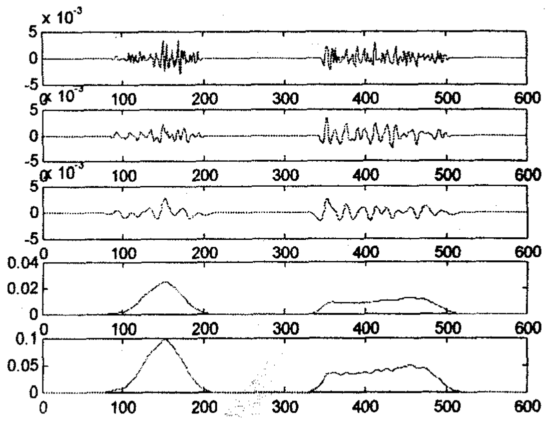

Fig 3.Wavelet approach

Considering the random noise of sample data, we further applied the maximum value wavelet reconstruction to eliminate the affect of the noise, which is shown in the Fig 4. Because the randomness in sample data, there are some random noise in wavelet decomposition, we can get more precise estimation using scale space filter technic to eliminate these noise.

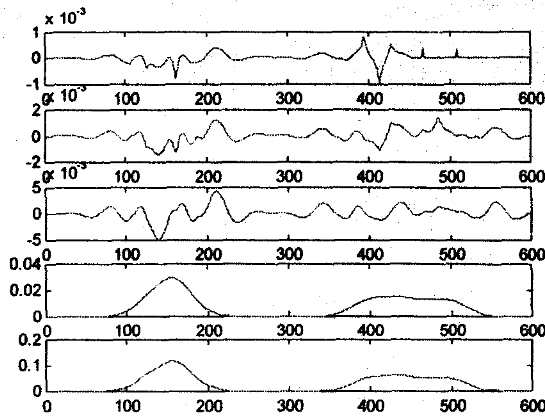

Fig 4. Wavelet approach using scale space filter technic to estimate p.d.f..

# VI . Conclusion

Wavelet approach is a multiresolution method , so the estimation will be affected little by $h_n$ . Compared with the Parzen window approach , our method use the information in the samples better than Parzen window method , by using not only smooth approximation but also detail information . The simulation shows that , for the same sample number , this method is much better than Parzen's approach . It does not need to find a proper $h_n$ painstakingly, and can largely reduce the sample number needed .

If window length is properly selected , the Parzen window method can give good estimation to some p.d.f. distributions . But when a distribution is complex and is composed of high frequency part , Parzen window can't get good result . However , our method can achieve fairly well estimation to complex p.d.f. distribution .

## REFERENCE

1. P. A. Devijer and J. Kittler , *Pattern Recognition : A statistical Approach* , 1982 .

2. S. Mallat , *Zero-crossings of A Wavelet Transform* , IEEE trans. on information Theory, Vol 37, No. 4 , July 1991.