Statistical Characterisation of Stack Filters

Olli Yli-Harja[1]

Abstract **- In this paper the properties of the joint distribution function of the outputs of stack filters with common arguments are examined. The special characteristics of these functions are discussed. Different approaches for characterising them are considered. Empirical tests are performed and reported. Different ways of extracting useful correlation information are compared.**

INTRODUCTION

Stack filters, including median filters have been applied to a variety of applications, c.f. [3], since median filtering was introduced by Tukey in 1974 [6]. These filters have some useful properties, not shared by strictly linear systems, e.g. robustness in the presence of heavy tailed noise. They are also well suited for image processing, where their non-linear effects are useful. The standard median filter, for example, removes impulsive noise and preserves sharp edges. Many median type filters, e.g. weighted median and order statistic filters, can be thought of as special cases of stack filters [3] and thus expressed as combinations of MIN- and MAX-operations. Threshold decomposition and the stacking property [2] provide a useful link between stack filters and positive Boolean functions. Statistical properties of median and stack filters have been studied in [1], [2] and [3].

In this paper we concentrate on examining the joint distribution of two stack filters. A formula for computing the cumulative joint distribution function has been derived [4], assuming independent input distributions. The output distribution function can be expressed as piecewise defined multinomial of the input distribution, if iid. input is assumed. For two stack filters this means that the output distribution is defined by two different multinomials, and that there is discontinuity of derivative on the diagonal. This discontinuity has an interesting effect when an expression for the corresponding joint density is derived - the density function will have to include the Dirac delta function. An infinitely small area on the plane, a mere line, possesses a finite portion of the probability mass.

In examining median filtered sequences the joint distribution of two subsequent samples, (for example), behaves in the above described way - the same input sample can be chosen to output several times. We will investigate the properties of the autocorrelation function of such sequences, and discuss the alternative ways of defining it. One such alternative is that described by Maragos [5]. We will also propose a new way of defining a useful measure of correlation, a way which is especially suited to signal processing problems involving median type filters. We will define this correlation measure so that it will estimate the amount of dependence caused by median filtering, independent of the input distribution.

We will report the results of simulations comparing these methods and the traditional cross correlation.

JOINT DENSITY FUNCTIONS OF STACK FILTERS

The cumulative joint distribution function of the output of two stack filters can be expressed as piecewise defined multinomial of the input distribution, if iid. input is assumed [4]. This means that the output distribution is defined by two different multinomials. The two functions are equal on the diagonal, this can be seen by setting $s=t$ in (from [4])

[1]O. Yli-Harja is with the Department of Computer Science, University of Helsinki, P.O.Box 26, 00014 University of Helsinki, Finland.

$$\Psi(s,t) = \begin{cases} \sum\limits_{\substack{a \in f^{-1}(0) \\ b \in g^{-1}(0)}} \prod\limits_{i=1}^{N} \Phi_i(s)^{\bar{a}_i \bar{b}_i} 0^{\bar{a}_i b_i} \\ \cdot (\Phi_i(t) - \Phi_i(s))^{a_i \bar{b}_i} (1 - \Phi_i(t))^{a_i b_i} \\ \text{if } s \le t \\ \sum\limits_{\substack{a \in f^{-1}(0) \\ b \in g^{-1}(0)}} \prod\limits_{i=1}^{N} \Phi_i(t)^{\bar{a}_i \bar{b}_i} 0^{\bar{a}_i b_i} \\ \cdot (\Phi_i(s) - \Phi_i(t))^{a_i \bar{b}_i} (1 - \Phi_i(s))^{a_i b_i} \\ \text{if } s > t \end{cases} \tag{1}$$

The derivatives of the two multinomials, however, are not equal on the diagonal.

This discontinuity has an interesting effect when an expression for the corresponding joint density is derived - the density function will have to include the Dirac delta function. An infinitely small area on the plane, a line, possesses a finite portion of the probability mass (Figure 1).

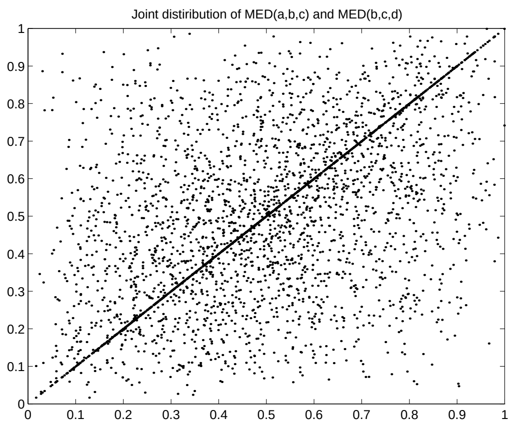

Figure 1. Joint distribution of samples one step apart in a 4000 point sequence, formed by filtering a (0,1)-uniformly distributed input signal with the 3-point median filter. 1332 points are lying on the diagonal line.

Let us now proceed to derive the joint density function of two stack filters. For simplicity, let input arguments be (0,1)-uniformly distributed iid. random variables. We can now restrict our attention to a unit square in (s,t)-plane. The results can be generalised for any continuous distribution. Let

$$\Psi(s,t) = \begin{cases} \Psi_1(s,t), & \text{if } s \le t \\ \Psi_2(s,t), & \text{if } s > t \end{cases} \tag{2}$$

and

$$\frac{\partial}{\partial s}\Psi(s,t) = \begin{cases} \frac{\partial}{\partial s}\Psi_1(s,t), & \text{if } s \le t \\ \frac{\partial}{\partial s}\Psi_2(s,t), & \text{if } s > t \end{cases}, \tag{3}$$

which can be expressed with the aid of Heaviside step function

$$H(x) = \begin{cases} 0, x < 0 \\ 1, x \ge 0 \end{cases} \tag{4}$$

as

$$\begin{aligned} &\Psi^s(s,t) = \Psi_1^s(s,t) \\ &+ H(s-t)(\Psi_2^s(s,t) - \Psi_1^s(s,t)), \end{aligned} \tag{5}$$

where

$$f^s(s,t) = \frac{\partial}{\partial s} f(s,t). \tag{6}$$

Note that,

$$\frac{\partial}{\partial x} H(x) = D(x), \tag{7}$$

where D denotes the Dirac delta function with the properties

$$D(x) = \begin{cases} \infty, & x = 0 \\ 0, & x \ne 0 \end{cases} \tag{8}$$

and

$$\int_{-\infty}^{\infty} D(x)dx = 1. \tag{9}$$

Differentiating (5) with respect to t we get the joint density function,

$$\begin{aligned} &\Psi^{st}(s,t) = \Psi_1^{st}(s,t) \\ &+ H(s-t)(\Psi_2^{st}(s,t) - \Psi_1^{st}(s,t)) \\ &- D(s-t)(\Psi_2^{s}(s,t) - \Psi_1^{s}(s,t)). \end{aligned} \tag{10}$$

The density function (10) can be used, for example, to calculate the probability P of the event that two different stack filters, sharing common arguments, produce equal outputs.

This can be done either by integrating over the density function, disregarding the diagonal, that is

$$1-P=\int_0^1\int_0^1(\Psi_1^{st}(s,t) + H(s-t)(\Psi_2^{st}(s,t)-\Psi_1^{st}(s,t)))dtds \qquad (11)$$

or by setting r=t=s and thus integrating over the diagonal part,

$$P=\int_0^1 -(\Psi_2^{s}(r,r)-\Psi_1^{s}(r,r))dr \qquad (12)$$

Example 1. Let the (0,1)-uniformly distributed iid. input signal be filtered with the three point median filter, and let us investigate the joint distribution two consecutive samples in the output sequence. By [4], their joint cumulative distribution function will be

$$\Psi(s,t)=\begin{cases} s^2+4s^2t-2s^2t^2-2s^3, s\le t \\ t^2+4t^2s-2t^2s^2-2t^3, s>t. \end{cases} \qquad (13)$$

Differentiating with respect to s, and applying (5) we get

$$\frac{\partial}{\partial s}\Psi(s,t)=(2s+8st-4st^2-6s^2) + H(s-t)(-8st-2s+6s^2+4t^2), \qquad (14)$$

and differentiating with respect to t we get the density function

$$\Psi^{st}(s,t)=(8s-8st) + H(s-t)(-8s+8t) - D(s-t)(-8st-2s+6s^2+4t^2). \qquad (15)$$

Applying (11) we get 1-P=2/3, and applying (12) P=1/3.

The finite probability mass on the diagonal line can be interpreted as follows: the two stack filters have a certain probability of producing exactly equal outputs - even for a continuous input distribution. This would not happen in the case of two different linear filters. The fact that stack filters always produce one of the input samples to the output, accounts for this phenomenon.

CORRELATION MEASURES

We will consider three correlation measures, and apply them to template matching, namely linear cross correlation (L2-correlation), morphological correlation (L1-correlation) [5], and a measure based on exact equality of samples (EQ-correlation).

Let $f(n)$ be an arbitrary signal and $g(n)$ a pattern to searched from f. To find the best match, an error criterion such as *mean squared error* (MSE) can be minimised

$$E_2(k)=\sum_{n\in W}(f(n+k)-g(n))^2. \qquad (16)$$

Since $(a\text{-}b)^2=a^2+b^2-2ab$, minimising (16) equals maximising

$$\gamma_{fg}(k)=\sum_{n\in W}f(n+k)g(n), \qquad (17)$$

yielding the classical (sum of products) linear correlation.

Using the mean absolute error (MAE) criterion

$$E_1(k)=\sum_{n\in W}\left|f(n+k)-g(n)\right|, \qquad (18)$$

and noticing that $|a-b|=a+b-2\min(a,b)$, we can define morphological correlation [5]

$$\mu_{fg}(k)=\sum_{n\in W}\min(f(n+k),g(n)). \qquad (19)$$

Correlation based on measuring the number of exactly equal samples, the so called EQ-correlation can be defined by

$$\sigma_{fg}(k)=\sum_{n\in W}(f(n+k)=g(n)), \qquad (20)$$

where the result of sample-wise comparison is taken to be real 0 or 1.

It has been shown in [5] that morphological correlation yields sharper peaks than linear correlation. EQ-correlation, on the other hand, is severely handicapped in many situations by the requirement of exact matching. Still, it may be useful in situations where sample amplitudes are restricted to a small number of discrete values (image processing). Another useful application may be examination of sequences filtered by stack filters. EQ-correlation is also totally independent of the shape of the continuous input distribution.

Let us now consider statistical characterisation of the density function (10), using the above described correlation measures. Let

$$\Psi^{st}(s,t) = P(s,t) + H(s-t)Q(s,t) + D(s-t)R(s,t), \tag{21}$$

where P,Q and R are multinomials. Linear cross correlation can now be calculated as

$$E(st) = \int_0^1\int_0^1 st\Psi^{st}(s,t)dsdt \tag{22}$$

and, denoting

$$\min(s,t) = s + H(s-t)(t-s), \tag{23}$$

morphological correlation as

$$E(\min(s,t)) = \int_0^1\int_0^1 \min(s,t)\Psi^{st}(s,t)dsdt \tag{24}$$

EQ-correlation is defined by (12), measuring the probability mass lying on the diagonal.

We can now continue with example 1, and calculate the autocorrelation function of the 3-point median filter analytically from the density function (15), recalling the output distribution functions from [4]. The results are illustrated in table 1.

Correlation of samples k steps apart

	$k=0$	1	2	3
$\mathrm{L2-corr.}$	$3/10$	$5/18$	$83/315$	$1/4$
$\mathrm{L1-corr.}$	$1/2$	$13/30$	$2/5$	$13/35$
$\mathrm{EQ-corr.}$	1	$1/3$	$2/15$	0

Table 1. Correlation of (0,1)-uniformly distributed iid. sequence, filtered with 3-point median filter.

EXPERIMENTS

The results presented in table 1. Are verified empirically. Autocorrelation functions are calculated by filtering a 2000-point sequence of iid. (0,1)-uniformly distributed input signal with the 3-point median filter. Then a portion of the output signal, (samples 800-1199), are matched against the filtered signal, applying L2-, L1- and EQ-correlation criteria. *a*-figures show the whole autocorrelation function, and *b*-figures a zoomed-in version.

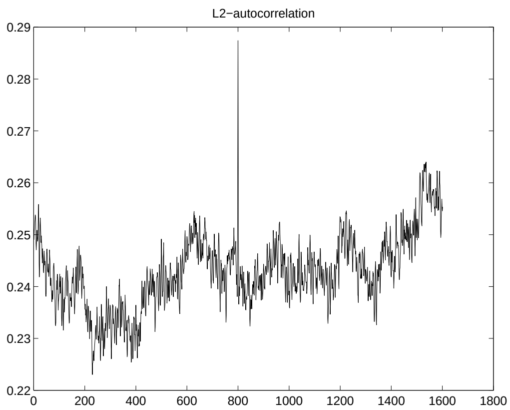

Figure 2a.

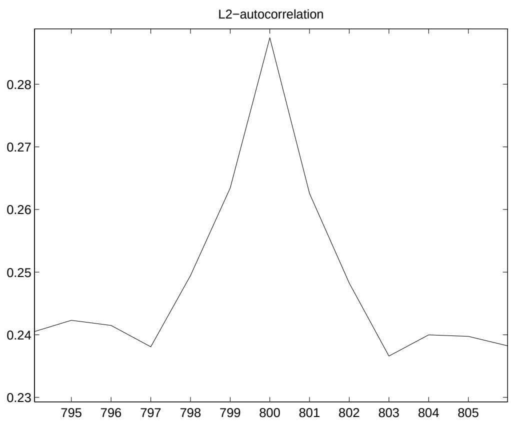

Figure 2b.

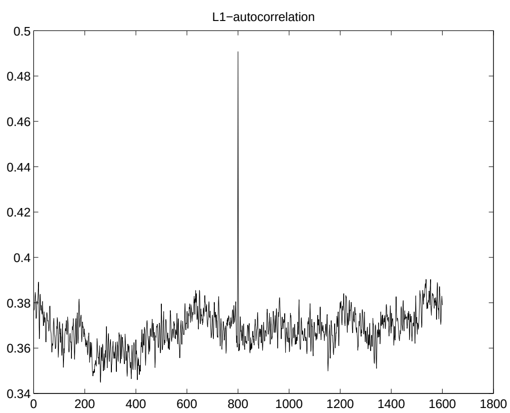

Figure 3a.

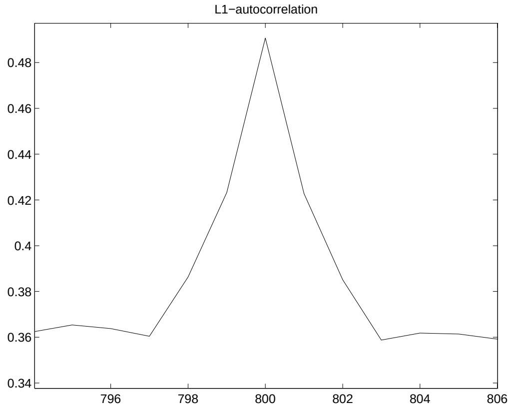

Figure 3b.

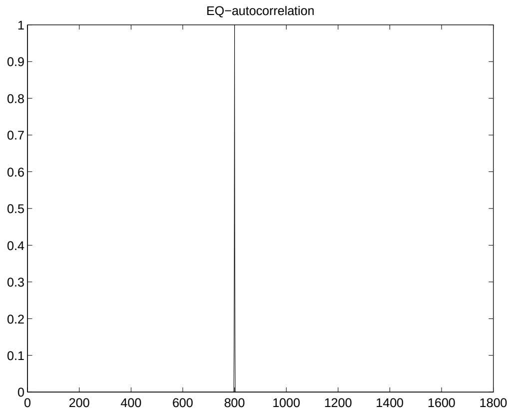

Figure 4a.

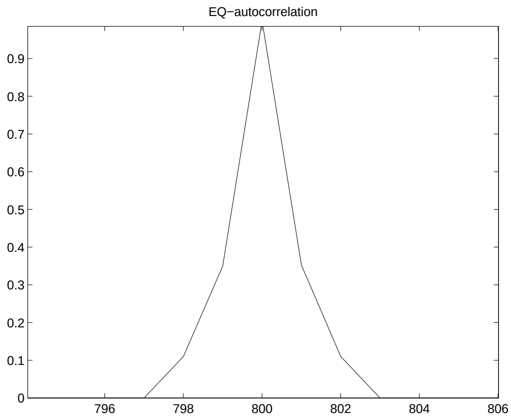

Figure 4b.

REFERENCES

[1] B. I. Justusson, "Median filtering: Statistical properties", in *Topics in Applied Physics, Two Dimensional Digital Signal Processing II*, T. S. Huang, Ed., Berlin: Springer, 1981.
[2] P. D. Wendt, E. J. Coyle and N. C. Gallagher Jr., "Stack filters", *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. ASSP-34, no. 4, pp. 898-911, Aug. 1986.
[3] O. Yli-Harja, J. Astola and Y. Neuvo, "Analysis of the properties of median and weighted median filters using threshold logic and stack filter representation", *IEEE Transactions on Signal Processing*, vol. SP-39, no. 2, pp. 395-410, Feb. 1991.
[4] O. Yli-Harja, "Formula for the Joint Distribution of Stack Filters", IEEE Signal Processing Letters, vol. 1, no. 9, pp. 129-130, Sept. 1994.
[5] P. Maragos, "Morphological Correlation and Mean Absolute Error Criteria", IEEE CH2673-2, pp. 1568-71, 1989.
[6] J. W. Tukey, “Nonlinear (Nonsuperposable) Methods for Smoothing Data”, Cong. Rec. EASCON’74, p. 673, 1974.