December 1992

Robust Signal Restoration and Local Estimation of Image Structure

Visa Koivunen

University of Pennsylvania

Follow this and additional works at: http://repository.upenn.edu/cis_reports

Recommended Citation

http://repository.upenn.edu/cis_reports/468


This paper is posted at ScholarlyCommons. http://repository.upenn.edu/cis_reports/468
For more information, please contact repository@pobox.upenn.edu.
Robust Signal Restoration and Local Estimation of Image Structure

Abstract
A class of nonlinear regression filters based on robust theory is introduced. The goal of the filtering is to restore the shape and preserve the details of the original noise-free signal, while effectively attenuating both impulsive and nonimpulsive noise. The proposed filters are based on robust Least Trimmed Squares estimation, where very deviating samples do not contribute to the final output. Furthermore, if there is more than one statistical population present in the processing window the filter is very likely to select adaptively the samples that represent the majority and uses them for computing the output. We apply the regression filters on geometric signal shapes which can be found, for example, in range images. The proposed methods are also useful for extracting the trend of the signal without losing important amplitude information. We show experimental results on restoration of the original signal shape using real and synthetic data and both impulsive and nonimpulsive noise. In addition, we apply the robust approach for describing local image structure. We use the method for estimating spatial properties of the image in a local neighborhood. Such properties can be used for example, as a uniformity predicate in the segmentation phase of an image understanding task. The emphasis is on producing reliable results even if the assumptions on noise, data and model are not completely valid. The experimental results provide information about the validity of those assumptions. Image description results are shown using synthetic and real data, various signal shapes and impulsive and nonimpulsive noise.

Comments
Robust Signal Restoration and Local Estimation of Image Structure

MS-CIS-92-92
GRASP LAB 339

Visa Koivunen

University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

December 1992
Robust Signal Restoration and Local Estimation of Image Structure

Visa Koivunen
General Robotics and Active Sensory Perception (GRASP) Laboratory
University of Pennsylvania
300C 3401 Walnut Street
Philadelphia, PA 19104-6228

Abstract

A class of nonlinear regression filters based on robust theory is introduced. The goal of the filtering is to restore the shape and preserve the details of the original noise-free signal, while effectively attenuating both impulsive and nonimpulsive noise. The proposed filters are based on robust Least Trimmed Squares estimation, where very deviating samples do not contribute to the final output. Furthermore, if there is more than one statistical population present in the processing window, the filter selects adaptively the samples representing the majority for computing the output. We apply the regression filters on the geometric signal shapes which can be found, for example, in range images. Moreover, the proposed methods are also useful for extracting the trend of the signal without losing important amplitude information. We present experimental results demonstrating the restoration of the original signal shape using real and synthetic data and both impulsive and nonimpulsive noise.

In addition, we apply a robust approach for describing local image structure. We use the method for estimating spatial properties of the image in a local neighborhood. Such properties can be used, for example, as a uniformity predicate in the segmentation phase of an image understanding task. The emphasis is on producing reliable results even if the assumptions on the noise, data and model are not completely valid. The experimental results provide information about the validity of those assumptions. Image description results are shown using synthetic and real data, various signal shapes and impulsive and nonimpulsive noise.
1 Introduction

The goal of many signal processing tasks is to recover the original noise-free signal from noisy samples and to extract the structure of the signal. Typically filtering and estimation methods assume that the noise is stationary, zero mean Gaussian distributed noise. Real sensor data, however, often do not satisfy these classical assumptions. For example, laser range data include several different noise distributions [5], and very deviating observations due to steep surface slopes, specular reflection, or occlusion may occur as well. To be able to describe the structure of the underlying signal we must have some understanding about the signal, i.e., we must assume a parametric model or a set of models. Samples which deviate a lot from the majority of data assumed to represent the true signal have a large influence on linear filtering and on least squares estimation by pulling the fit towards them.

Linear FIR filters used for noise attenuation tend to perform poorly in the presence of very deviating or bad samples. Furthermore, they have a tendency to smear discontinuities which are important features in several signal processing tasks. Some nonlinear filters, on the other hand, can attenuate noise and simultaneously preserve details such as sharp edges. Median filtering is widely used for such tasks, for example in speech and image processing applications. The impulse response of the median is zero which is a very desirable feature when attenuating impulsive noise. Unfortunately, it does not effectively suppress nonimpulsive noise components and it distorts some signal shapes. In this paper we will consider a robust regression filter which can attenuate both impulsive and nonimpulsive noise while preserving the shape and important details of the signal.

Local window operators [2, 15] are widely used for estimating local image structure in several image processing and segmentation tasks [5]. The underlying surface is represented as linear combinations of polynomials and the coefficients are computed so that they minimize errors in the least squares sense [15]. The classical estimation methods assume that the noise is Gaussian distributed, and that all the data belong to one statistical population that can be represented using one model and one parameter set. However, there may occur outliers because of a very tailed noise distribution or there may be discontinuities in the data set. The robust approach we apply here produces reliable results in the presence of outliers and discontinuities.

The organization of this paper is as follows. In section 2 we describe briefly the background of some methods from robust theory used for noise attenuation and image structure estimation. In section 3 we propose regression filters based on robust theory for attenuating various type of noise while preserving the shape and important details of the original signal. Our approach is based on Least Trimmed Squares estimation [30]. We apply variable order regression filtering for restoration of geometric signal
shapes using up to second-order model. A special case of zero-order filter is also considered. We show experimental results demonstrating the capability of preserving the shape and discontinuities of the signal, and the performance under both impulsive and nonimpulsive noise. In section 4 we apply robust methods for estimating local image structure which is an important part of several segmentation methods used in image understanding. We use both simulated and real sensor data. The real sensor measurements are range data, where each sample measures the distance from the object surface to the sensor plane.

2 Robust Estimation

We provide a brief overview of some widely used robust estimation approaches. The concept of robustness means insensitivity to small departures from idealized assumptions for which the estimator is optimized. Robustness is usually used in context of distributional robustness, i.e., the actual noise distribution deviates from the nominal distribution. The nominal noise distribution is in most cases i.i.d. Gaussian with possibly unknown scale. The deviations, however, may also be due to model class selection errors, or there may be more than one statistical population present in the data set, and hence it is not possible to describe it with only one set of parameters. Robust methods can be considered to be approximately parametric, i.e. a parametric model is used but some deviations from the strict model is allowed.

The breakdown point of the estimator is describes formally the smallest percentage of outlying points which causes incorrect estimates. Least squares estimation has a breakdown point of 0 \% [30]. For high-breakpoint estimators the breakpoint is close to 50\%, and does not decrease so rapidly if the number of parameters to be estimated increases [30]. We feel that a high breakpoint is less important from the viewpoint of sensor noise because if almost 50\% of the measurements are bad because of a sensor, it is probably time to calibrate or replace it. The high breakpoint protects us from very influential observations from other data populations while computing the estimates, for example where discontinuities occur. In addition to a high breakpoint, we want to produce good estimates, described by the term efficiency. There is a trade-off between being a highly robust and a highly efficient estimator [25, 30].

Most of the robust statistical estimators can be classified into three categories: M-estimates, L-estimates and R-estimates. R-estimates are not considered here. M-estimates are generalized form of maximum likelihood estimators and minimize a function

\[ \sum \rho(r_i) \]

of the residuals \( r_i \), which are the difference between estimated and actual data. \( \rho \) is a
symmetric function with a unique minimum at zero \([30]\). Differentiating this expression with respect to regression coefficients gives the function \(\Psi(r_i)\). A lower weight is given to very deviant observations in the estimation procedure. The weights \(w_i\) are computed by \([30]\)

\[
w_i = \frac{\Psi(r_i)}{r_i}
\]

using the residuals of each point to determine the influence of each residual to the fit. Estimators using weighting functions which reject completely observations farther than certain distance are called redescending. Among the most widely employed functions for weighting are Huber's, Andrew's, Hampel's and Tukey's \(\Psi\)-function. The shape of each weighting curve is depicted in Figure 1. The breakpoint of M-estimators has been shown to be \(\varepsilon = 1/(p + 1)\), where \(p\) is the number of parameters to be estimated \([22]\).

L-estimators are linear combinations of order statistics. They are of the form:

\[
T_n(x_1, \ldots, x_n) = \sum_{i=1}^{n} a_i x_{i:n},
\]

where \(x_{1:n}, \ldots, x_{n:n}\) are the ordered samples and the \(a_i\)'s are coefficients. One of the most widely used L-estimators for location estimation is \(\alpha\)-trimmed mean, where \(\alpha n\) samples from the both ends of the ordered set of samples do not contribute to the estimate.

Least Median of Squares estimation (LMS or LMedS) is based on idea by Hampel and was later proposed by Rousseeuw \([29]\). The estimator is defined as follows:

\[
\text{Minimize median } r_i^2
\]
The estimated parameters should give the smallest value for the median of squared residuals for the whole set of samples. Note that no sum or weighted sum of residuals is minimized. The estimator is very robust but it has a slow convergence rate. The breakdown point of the estimator for \( n \) samples and \( p \) parameters is \( \varepsilon = ([n/2] - p + 2)/n \), if \( p > 1 \) [30].

The Least Trimmed Squares (LTS) estimation principle was introduced by Rousseeuw to overcome the efficiency problems of the LMedS estimation technique [30]. We chose this particular method because of the good convergence rate, smoother objective function and more stable algorithm than the LMedS method [31]. In a neighborhood with \( n \) data points it minimizes the sum

\[
\sum_{m=1}^{h} (r^2)_m
\]

where \((r^2)_1 \leq ... \leq (r^2)_h \leq ... \leq (r^2)_n\) are the ordered squared residuals and \( h \) is the number of residuals used in summation. The residuals are first squared and then ordered. The LTS method achieves the maximal breakdown point \( \varepsilon = ((n-p)/2 + 1)/n \) for \( h = [n/2] + [(p + 1)/2] \), where \( p \) is the number of parameters to be estimated.

To make the residuals equivariant with respect to scale, one has to standardize them by means of some estimation of standard deviation \( \sigma \). The median absolute deviation (MAD) scale estimator:

\[
\sigma = C \text{ median } (|r_i - \text{median } r_i|)
\]

is often used to create a scale invariant version of the estimator. The constant \( C = 1.4826 \) is for consistent estimation when Gaussian noise present. A specific scale estimator for the LTS estimator is [30]:

\[
\sigma = C_2 \sqrt{\frac{1}{n} \sum_{m=1}^{h} (r^2)_m}
\]

where \( C_2 \) is a correction factor. Standardized residuals \((r_i/\sigma)\) are very useful for outlier detection and for evaluating the validity of the assumptions.

3 Restoration Filtering

3.1 Related Work

Linear FIR filters tend to smooth out discontinuities, and perform poorly in the presence of impulsive noise or bad samples. In this section we will consider some nonlinear filters based on robust theory, and propose two filtering methods for signal processing purposes based on Least Trimmed Squares (LTS) estimation.
Median filtering is widely used for noise attenuation. It was proposed for signal processing by Tukey [33]. It attenuates impulsive noise components very effectively while preserving sharp step edges [10]. However, the median does not suppress non-impulsive noise very effectively. Some enhancements for median filtering have been proposed. Lee and Tantaratana [21] proposed a modified median (MMF) filter to overcome edge jittering problems caused by impulsive noise (especially one impulse). They use a hypothesis test to detect edges. Edge detection, however, faces the same noise attenuation problem [32, 8].

To be able to deal also with nonimpulsive noise components, methods that combine nonlinear and linear filtering approaches have been proposed [26, 16, 20]. \( \alpha \)-trimmed mean (\( \alpha \)-TM) filtering based on robust L-estimation is employed in [3]. Bovik et al. introduced an order statistic filter (OSF) [7] which is based on L-estimation as well. The filter output is a linear combination of ordered samples. Heinonen and Neuvo [16] proposed a FIR-median hybrid (FMH) filters that combine linear filtering with median filtering. The window for the filtering is subdivided into an odd number of subwindows. The linear filtering is performed in subwindows and the final output of the FMH filter is the median of subwindow outputs.

Lee and Kassam presented M-filters based on M-estimation [20]. They used the median as a reference signal for M-filters, and very deviating samples are downweighted so that they contribute less to the filter output. M-filters are in general not able to preserve sharp edges. If one is using redescending estimators, \textit{a priori} information about the height of the edges is needed to preserve them. Kashyap and Eom applied M-estimators to image restoration [17].

Lee and Kassam [20] proposed also a Modified Trimmed Mean (MTM) filter which chooses an interval \([x_{med} - q, x_{med} + q]\) for averaging, where \(x_{med}\) is the sample median and \(q\) is a preselected constant. The value of \(q\) should be approximately \(H - 2\sigma\), where \(H\) is the assumed minimum height of an edge and \(\sigma\) is the noise standard deviation. The filter preserves sharp edges if \(q \leq H\). They also proposed a double window (DW MTM) modification for MTM filters, using the smaller window for determining the median and the interval where the averaging is done. The mean of the samples that are within the interval in the larger window is the filter output. Gandhi and Kassam [12] investigated properties of combination filters (C-filters) that use rank-order based weighting of temporal order data within a window to produce the output. They also introduced a class of generalized C-filters (GC-filters). They allow the coefficients to be designed to optimally weight the observations that are not trimmed.
3.2 Least Trimmed Squares Filtering

The Least Trimmed Squares (LTS) algorithm described in [30] is related to the projection pursuit method which tries many low-dimensional sample clusters of higher dimensional data set to find the most informative cluster [14]. The output value and the sum of $h$ squared residuals from equation (5) for each subsample set are computed. The output parameters from the set with the smallest sum of squared residuals are chosen. The algorithm is computationally very intensive, hence they used random sampling to reduce it. The probability $P$ that at least one subsample contains good only points is $P = 1 - (1 - \left(1 - \varepsilon \right)^p)^m$, where $\varepsilon$ is the fraction of outliers, $p$ is the number of parameters to be estimated and $m$ is the number of sets randomly sampled. The number of sets needed remains high, especially for higher-order models, because the deviating samples are typically clustered, and the fraction of outlyingness is frequently close to 50\% due to discontinuities.

To make the method computationally more feasible for signal processing applications, we consider the fact that the sampling interval is typically constant. Hence, the dimensionality of the data set is not high, and the outliers are more likely to occur in actual sensor output (response variable), than as deviations from the tessellation of the samples (explanatory variables). Instead of using a subsampling method as used in [30], we chose to employ robust first estimates for the minimization to find an informative cluster of samples, and compute the estimates iteratively. The precautions we take are the same used successfully in Princeton Monte Carlo Study [1] for redescending estimators: the starting value of the iterations is the highly reliable median, and at least half of the observations are not trimmed nor severely downweighted. The computational cost of the filtering is lower, and the algorithm can easily be used for higher-order models as well, and is relatively easy to implement on hardware.

Let the input signal in a 1-D case be a sequence of noisy samples \{\ldots, x_{i-M}, \ldots, x_i, \ldots, x_{i+M}, \ldots\}. The underlying signal in each odd size processing window of $n$ samples is assumed to be piece-wise polynomial defined as follows:

$$f(x) = \sum_{k \leq K} a_k x^k,$$  

(8)

where $a_k$'s are the coefficients, $x$ is the index of the sample in the processing window, and $K$ is the maximum order of the polynomial. In 2-D case the input image is of type

$$z = f(x, y) = \sum_{k+l \leq K} a_{k,l} x^k y^l,$$  

(9)

where $a_{k,l}$'s are the coefficients, and $x$ and $y$ stand for row and column indices in the $U$-by-$V$ neighborhood of $n$ samples. The assumed noise distribution is zero mean Gaussian but we allow departures from it as long as the overall outlyingness is below.
the breakpoint, i.e. the majority of the data can be described using the assumed model. The deviations from the assumptions do not have to be symmetric either, which is the case, for example, when there are discontinuities present in the processing window.

The regression filter (RLTS) we propose is based on Least Trimmed Squares error measure from (5). The squared residuals from the reference value are first ordered, and then only the samples yielding smallest sum of $h$ squared residuals are used for computing the filter output. We use the median of the processing window as the reference value for computing the first set of residuals. The output from the previous iteration is used as a reference value for the next iteration. The application of reference signals, however, includes an assumption on the types of errors that will occur in the data. A block diagram of the filter using only a zero-order model is depicted in Figure 2.

![Block Diagram](image)

Figure 2: A block diagram of the zero-order Least Trimmed Squares filtering.

If the filtering is iterated recursively, the output from the previous iteration is used as the reference value for the next iteration. The process is continued until the filter output converges, or the maximum number of iterations is reached. The output is defined to convergence if the difference in quality factors between two successive iterations is smaller than a given threshold value. The iteration should be done in floating point mode to avoid quantization errors introduced after each iteration. The quality factor $E_l$ is defined as follows:

$$E_l = \sqrt{\frac{\sum_{m=1}^{h} (r^2)_m}{(h - p)}},$$

where $p$ is the number of parameters to be estimated.

We used zero-, first- and second-order ($K = 0, 1, 2$) functional models in the filter. The estimation is a forward selection process which begins with a simple regression
model and in subsequent steps a higher order polynomial is used. The parameters from the functional model giving the best quality factor from (10) are used to compute the filter output. A block diagram of the model selection is depicted in Figure 3.

![Block diagram of model selection process](image)

Figure 3: A block diagram of the model selection process.

An odd sized processing window of n samples is used for filtering. The size of the window determines the fineness of the details the filter can preserve. There is a trade-off in determining the window size. For a large window size, the filtering attenuates noise better as long as there is only one population present in the region of interest. A smaller processing window, on the other hand, can preserve finer details, for example waveforms, and is computationally cheaper. Any least squares technique can be applied for computing the output from the subset of h samples. The selection of h is a design factor. To be able to preserve edges, we select a value \( h = \lceil n/2 \rceil + \lceil (p + 1)/2 \rceil \) which yields the maximum breakdown point for each order. Note that the value of h yielding the maximum breakpoint varies depending on the order of the model. If we have a priori information about the location of the edges, we can use less trimming where no discontinuities occur to make the smoothing more efficient. All the samples where the residual is, for example, less than \( 2.0 \times \sigma \), where \( \sigma \) is a robust scale estimate, could be used to compute the output. Here, we assume that no such information about edges is available.

Regression filtering is a very useful tool for filtering complicated signals if one wants to restore the shape of the signal very accurately as, for example, in the case of filtering geometric shapes measured by a laser range finder [18]. Moreover, it can be used for estimating the trend of the signal without losing important amplitude information. The proposed regression approach is useful because it provides more signal understanding: rather than reject some samples as outliers, it detects a model
failure and applies an appropriate functional model to filter the signal. The proposed filtering is based on robust theory and it combines both nonlinear (ordering) and linear operations. Therefore, it attenuates both impulsive and nonimpulsive noise in images and preserves discontinuities. Furthermore, it is an efficient estimator for the subset of \( h \) samples if the inlier noise is Gaussian distributed.

3.3 Experimental Results

We apply two different regression filters based on Least Trimmed Squares error norm. The first one (RLTS) uses functional models up to second order \((K = 0, 1, 2)\), and the other one (RLTS-0) is an interesting special case of zero-order model \((K = 0)\). We compare the results to those obtained using median filtering. The set of experiments presented is designed to find out how well the proposed filtering methods preserve the shape of the signal while attenuating different types of noise. In addition to distributional robustness we address the validity of the assumption on the functional model of the filters and data smoothness. We demonstrate the properties of the filters using synthetic 1-D and 2-D signals, and show further experiments using real data. Zero mean Gaussian distributed noise and random bit error were added to the synthetic data. A maximum of 3 iterations is used in the experiments. Typically the filtering converges in 2 – 3 iterations.

3.3.1 Restoring the signal geometry

First, we use a perfect noise-free signal as test data to find out how much each filter distorts the original shape of the signal. The first test signal consists of piecewise-continuous zero-, first- and second-order data, and there occurs discontinuities as well. A 12 bit quantization is used. The median, LTS and RLTS filters are applied to the signal using a 5 point window. The original signal and the obtained output signals are depicted in Figure 4. All the filters preserve step discontinuities very effectively. The monotone first-order segment is also well-preserved. The median and RLTS-0 filters do not preserve roof edges and they flatten the top of the second-order segment as well. LTS filter behaves very much like the median because the median is used as a reference signal for computing the residuals. RLTS filter is able to choose the appropriate model for the filtering and preserves the shape of the signal very well, even in the case of roof edge and of the second-order data.

The second test data is similar to the first one except Gaussian distributed noise with \( \mu = 0 \) and \( \sigma = 5.0 \) and random bit error with probability \( P = 0.015 \) is added to noise-free signal. The random bit error produces very impulsive noise, i.e. samples that can be considered outliers. The noisy test data and the corresponding filter outputs
Figure 4: Filtering results for noise free signal: (a) The original noise-free signal, (b) the median filtered signal, (c) the RLTS-0 filtered signal, and (d) the RLTS filtered signal, respectively.

are depicted in Figure 5. The result indicates that all the filters are able to detect and reject outlying samples. The median and the RLTS-0 filter work well on flat segments but these filters are not able to restore the shape of the roof edge and parabolic segment. The output shows that the RLTS filter, on the other hand, is able to restore the shape of the original signal almost perfectly.

A synthetic 2-D image is employed to find out how 2-D versions of the filters restore geometric shapes in the image. The image consists of zero-, first- and second-order surface patches, with discontinuities between several regions. Each filter is run using a 7-by-7 processing window. Gaussian noise with $\sigma = 5.0$ and random bit error noise with probability $P = 0.01$ are added to the synthetic data set with 8-bit quantization. Figure 6 shows the differences between the original noise-free image and the outputs of different filters applied on the noisy image. All the filters suppress this noise very effectively. Moreover, they preserve step discontinuities very well. Each of these filters distorts corners because the corner point does not belong to the population representing the majority of the samples in the neighborhood. An example of the distortion of corners is shown in Figure 7, where perfect piecewise-constant image is filtered using
Figure 5: Filtering results for noisy signal: (a) The noisy signal, (b) the median filtered signal, (c) the RLTS-0 filtered signal, and (d) the RLTS filtered signal, respectively.

2-D versions of each filter. The distortion is due to a model failure, and a distinct corner model should be used to be able to tell that there is a corner instead of outlying data. LTS and median filtering assume a piecewise zero-order (constant) signal model, which is often a valid assumption when processing intensity images. However, if the signal model is more complicated (e.g., in the case of range images), the filters will distort some features analogously to the 1-D case. The median filtering and RLTS-0 filtering make the roof edges flat, and the top of the spherical surface is distorted as well. RLTS filter is able to preserve the shape due to the appropriate signal model.
Figure 6: Filtering results for noisy image: (a) the 3-D surface plot of the original, and (b) the noisy data, (c) the noisy image, and (d) the differences between the original and the RLTS-0, (e) the median, and (f) the RLTS filtered noisy image, respectively.
Figure 7: Distortion of corners caused by filtering: (a) original signal, (b) the RLTS-0 filtered signal, (c) the median filtered signal, and (d) the RLTS filtered signal, respectively.
3.3.2 Noise attenuation

The 2-D versions of the filters are applied on test images consisting of heavily contaminated data. 8-bit quantization is used, and Gaussian noise with three different $\sigma$-values and random bit error with probability $P = 0.01$ is added to the “wedding cake” image with step discontinuities. The amplitude of the background is 100, and the height of the step edges is $H = 50$. Figure 8 shows the noisy “wedding cake” ($\sigma = 10$), and the obtained filter outputs using a 5-by-5 processing window.

![Filtering results for noisy step edge signal: (a) the noisy signal, (b) the median filtered signal, (c) the RLTS-0 filtered signal, and (d) the RLTS filtered signal, respectively.](image)

The output shows that all the filters suppress impulsive noise very effectively. Furthermore, all the filters preserve step edges very well. All the filters distort (cut) the corners because the processing window is centered in the corner and the majority population is the one surrounding the corner. RMS errors of each filtering method at different noise levels for it image are shown in Table 1. The RMS errors indicate that RLTS-0 filter attenuates noise slightly better than the median, although the difference is not very significant. The zero-order model in RLTS-0 filter is appropriate here, because the signal consists of piecewise-constant data. The value of $h$ is such that it gives
Table 1: Different filtering methods and the RMS errors obtained for the "wedding cake" image.

<table>
<thead>
<tr>
<th>Noise/Filtering method</th>
<th>$\sigma = 5$</th>
<th>$\sigma = 10$</th>
<th>$\sigma = 32$</th>
</tr>
</thead>
<tbody>
<tr>
<td>median</td>
<td>2.52</td>
<td>4.03</td>
<td>9.70</td>
</tr>
<tr>
<td>LTS</td>
<td>2.40</td>
<td>3.79</td>
<td>10.72</td>
</tr>
<tr>
<td>RLTS</td>
<td>2.96</td>
<td>4.20</td>
<td>11.41</td>
</tr>
</tbody>
</table>

the highest breakpoint to be able to preserve the discontinuities. The RLTS filter suffers slightly more from very severe nonimpulsive noise because the second-order model employed may give high quality estimates for some noise patterns.

The real facemask data from NRCC [27] Range Image library is filtered using the median, RLTS-0 and RLTS filters in the 5-by-5 neighborhood. The differences between original noisy data and filter outputs are depicted in Figure 9 where dark areas mean large differences. The results indicate that the largest deviations occur on the boundary of the object. Moreover, the deviations are clustered on certain regions which indicates that the differences may be distortions caused by filtering. The median and RLTS-0 filter distort some important details from the facemask, for example by the nose, the eyebrows and the lips. The RLTS filter distorts those details significantly less. The RLTS-0 filter suffers from its zero-order model because it flattens the second-order surfaces.

We apply the proposed methods for restoration of standard gray-scale image as well. Contaminated version of the picture is made by adding Gaussian noise with zero mean and $\sigma = 5.0$. Moreover, the pixels are contaminated by impulsive noise which is produced by random bit error with probability $P=0.01$. Figure 10 shows the filtering outputs by using a 5-by-5 processing window. All the filters attenuate both impulsive and nonimpulsive noise. In qualitative comparison the output from both RLTS-0 and RLTS appears to be less blurred than the output of the median filter. The differences between the original image and the outputs of the filters applied on the contaminated image are shown in Figure 11 to illustrate the restoration capability and the distortion caused by the filters. The original signal is most distorted by RLTS-0 filter because the assumption on piecewise-constant signal is not valid. Moreover, it enhances step edges. The RLTS method distorts the signal the least. Table 2 shows the RMS errors between the outputs of different filters applied to the contaminated image and the original image with no added noise. The result indicates that RLTS filter has the smallest RMS errors except in the 3-by-3-neighborhood. It is apparent
Figure 9: The differences between the sensor data and the filter output where dark areas indicate large differences: a) the original data, b) the differences obtained using median filtering, c) RLTS-0 filtering and d) RLTS filtering, respectively.

that the assumption on piecewise-constant signal is not valid in the Lena image.
Figure 10: The filtering results for Lena image: a) the contaminated image, and b) the output of median, c) RLTS-0 and d) RLTS filter, respectively.

Figure 11: From left: the signal distortion using median, RLTS-0 and RLTS filter, respectively. Darker areas indicate more severe distortion.
Table 2: Different filtering methods and the obtained RMS errors for Lena image using different size processing windows.

<table>
<thead>
<tr>
<th>Neighborhood size/Filtering method</th>
<th>3-by-3</th>
<th>5-by-5</th>
<th>7-by-7</th>
</tr>
</thead>
<tbody>
<tr>
<td>median</td>
<td>5.73</td>
<td>8.19</td>
<td>10.37</td>
</tr>
<tr>
<td>LTS</td>
<td>6.62</td>
<td>8.47</td>
<td>11.13</td>
</tr>
<tr>
<td>RLTS</td>
<td>5.90</td>
<td>7.70</td>
<td>10.09</td>
</tr>
</tbody>
</table>
4 Application to Image Structure Estimation

4.1 Introduction

Computer vision applications typically want to subdivide digital pictures into regions or contours that have a certain uniformity. Local window operators [2, 15] are widely used for estimating spatial properties of digitized surfaces. These properties can be used as uniformity predicates in a segmentation task. The window operators are usually computed so that they minimize error in the least squares sense. However, the idealized assumptions on which the least squares estimation is based are not always valid in practice. There may be more than one statistical population present on the support area for the fit, or the sample may contain erroneous data that have a large influence on estimated surface coefficients.

An odd size rectangular $U$-by-$V$ local window of image data points is typically used for the surface parameter estimation. Methods using planar surface patch primitives require a large number of surface patches to adequately describe curved surfaces. Hence, second- or even higher-order models are often used. The underlying surface is assumed to take the parametric form of a polynomial in each neighborhood:

$$f(a, x, y) = \sum_{k+l \leq K} a_{k,l} x^k y^l,$$

where $a_{k,l}$'s are the coefficients, and $x$ and $y$ stand for row and column coordinates in the $U$-by-$V$ neighborhood. $K$ is the maximum order of the assumed surface model. In a typical application a more structured and descriptive characterization of the surface is computed based on the coefficients obtained from the fit [5, 19]. Furthermore, one would like to produce quantitative information about the quality of the processing.

The quality factor used for evaluating the goodness of the fit procedures in an odd sized $U$-by-$V$ processing window is defined as follows [24]:

$$E_1 = \sqrt{\frac{\sum_{x=-N}^{N} \sum_{y=-M}^{M} w_{x,y} r_{x,y}^2} {\sum_{x=-N}^{N} \sum_{y=-M}^{M} w_{x,y} - p}},$$

where $N = (U - 1)/2$, $M = (V - 1)/2$, $x$ and $y$ are the row and column indices of the samples in the processing window, $p$ is the number of parameters to be estimated, $r_{x,y}$ are the residuals and $w_{x,y}$ are the weights. In least squares estimation all the weights are set to $w_{x,y} = 1$, in M-estimation the weights are computed by using (2), and in the LTS estimation the weights for $h$ samples are set to 1, and for all the trimmed samples to zero. The convergence of the fit is defined good if the difference of the quality factors from two successive iterations is smaller than a given threshold value. In this section we apply Least Trimmed Squares method for estimating the surface coefficients from (11).
The method provides a powerful tool for recovering spatial properties of the surface reliably. We also emphasize production of quantitative data to be able to analyze the quality of the data description and validate the assumptions used in the fit procedure.

4.2 Related work

Förstner was the first to apply robust estimation in computer vision [11]. Kahyap and Eom applied M-estimation to image restoration [17]. Besl et al. applied the Iterative Reweighting Least Squares (IRLS) M-estimation technique for filtering impulse noise from the image [6]. Meer used the LMedS method to estimate polynomial surfaces [24, 25]. Koivunen employed LTS estimation method for surface description purposes [19]. Darrell et al. proposed a cooperative framework to be able to deal with occluding regions [9]. They use an array of parallel estimators instead of only one local estimator.

The Robust Window Operator by Besl [6] uses IRLS technique to estimate \( a_{k,l} \) from (11). The variable order method used selects the parameter set from the order yielding the best fit quality factor [4]. The median of the neighborhood is used to compute the first set of residuals. The algorithm employs first Huber’s minimax estimator followed by Hampel’s redescending estimator. A refined set of parameters is obtained by solving

\[
a = (A^T W A)^{-1} A^T W z
\]

where \( a = a_{k,l} \) is the vector of new parameters, the columns of \( A \) include the basis functions, the diagonal of the matrix \( W \) includes the weights and \( z \) is a vector of data points.

Least median of squares method was used for estimating zero- and first-order polynomial surfaces by Meer et al. [24]. To decrease the amount of computation random sampling is employed in order to choose a reduced set of \( p \)-tuples to be used for estimation, where \( p \) is the number of parameters. If higher-order models are used and the if the fraction of outlying data is high, the amount of \( p \)-tuples and computation required increases drastically ([30], p. 198). The contamination is often close to 50% because of discontinuities. Roth and Levine applied the same method as well [28]. They detect first jump and roof edges from the image by thresholding according to depth differences and differences in surface normal, respectively. Connected set of pixels that do not include edges are used as input to the LMedS fit. The fit quality is determined by comparing the least median of squares error to a threshold value which is set to \( 2.5 \times V \), where \( V \) is the variance of the noise. The fit is run using first order surfaces and if it is not successful, a second order model is used instead [28]. The method can effectively reject impulsive noise but it is not efficient under additive zero-mean Gaussian inlier noise.
4.3 Method Based on Least Trimmed Squares Estimation

In the Least Trimmed Squares surface estimation we assume that the underlying surface is piecewise-continuous and can be modeled using up to second-order patches. We employ variable order models [4] in the fit procedure to be able to select an appropriate model in each neighborhood. The quality of the fit is computed for the fit result of each order, and the parameter set giving the best quality is selected to be the final parameter set. Additional information can be stored for data analysis purposes, for example for detecting outliers and for evaluating the validity of the assumptions used.

The first estimate should be a typical representative of the true signal in the neighborhood to guarantee fast convergence. The median provides a very robust first estimate compared to the mean or to the parameters computed by least squares estimation (LSQ) [19]. The refined parameter vector can be easily computed by using standard LSQ method for the obtained subset of \( h \) pixels, for example by:

\[
a = (A^T A)^{-1} A^T z,
\]

where \( a, A \) and \( z \) are as in equation (13). The refinement is iterated until good convergence or the maximum number of iterations is reached [19]. The block diagram of the algorithm is depicted in Figure 12. We selected the value of \( h \) for each order to yield the maximum breakdown point. This is important because there may occur discontinuities in the neighborhood, and the high breakpoint estimator is able to exclude the samples from the minority populations from the fit. The breakpoint does not decrease so rapidly for LTS estimation compared to M-estimation as the number of parameters.

![Block diagram of surface parameter estimation for describing image structure based on Least Trimmed Squares estimation.](image)

Figure 12: A block diagram of surface parameter estimation for describing image structure based on Least Trimmed Squares estimation.
increases. LTS estimation rejects completely some samples from the fit, and hence its asymptotic variance ([30], p. 180) is large. Therefore, a one-step improvement is recommended in ([30], p. 191) to incorporate more samples to the fit, for example by using weighted least squares method. In addition to the ability to reject deviating samples, the LTS estimation is also efficient under Gaussian inlier noise.

4.4 Experimental Results

Our set of experiments is chosen to study how the classical least squares estimator and the robust estimator deal with departures from widely used assumptions on noise, data and functional model used in the fit. Impulsive and nonimpulsive noise are used to demonstrate distributional robustness. We also emphasize the diagnosis of the results using the fit quality factors and standardized residuals to detect the departures from the assumptions.

To visualize the performance of both methods in local surface differential property estimation, we classified all the pixels into different type of surfaces based on the those properties. This simple classification scheme is based on sign map of second fundamental form

\[ \Pi = Ldu^2 + 2Mdudv + Ndv^2 \]  

(15)

coefficients \( L, M \) and \( N \) from differential geometry [23]. This scheme labels surfaces into elliptic, parabolic, hyperbolic and planar patches. The coefficients are computed in each neighborhood, and the type of the surface is determined based on the sign of the discriminant \( LN - M^2 \). A threshold value for zero is set. The classification scheme is displayed in Table 3. The coefficients can also be used to compute, e.g., surface

<table>
<thead>
<tr>
<th>Discriminant</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>( LN - M^2 &gt; 0 )</td>
<td>Elliptic</td>
</tr>
<tr>
<td>( LN - M^2 &lt; 0 )</td>
<td>Hyperbolic</td>
</tr>
<tr>
<td>( LN - M^2 = 0, ) ( L^2 + M^2 + N^2 \neq 0 )</td>
<td>Parabolic</td>
</tr>
<tr>
<td>( L = M = N = 0 )</td>
<td>Planar</td>
</tr>
</tbody>
</table>

Table 3: Classification Scheme based on Second Fundamental Form Coefficients.

normals, and there exists methods that produce a richer description of the underlying surface [5, 19]. Figure 13 depicts range data from NRCC range image library [27] used for test images. Figure 14 illustrates the classification results using 7-by-7 neighborhood
for the both images. Bright regions stand for hyperbolic surfaces, medium bright for elliptic surfaces and medium dark for parabolic surfaces and dark patches for planar surfaces. The threshold value for zero was set to 0.5 for the image on the left and to 1.0 for the facemask image. The result indicates that the least squares method produces regions that tend to spread over their actual boundaries. The classification obtained using robust method covers the actual regions more accurately. Many of the errors that occur near surface boundaries can be avoided because the samples used for coefficient estimation are more likely to be from one population. Estimating differential properties of surfaces over discontinuities is an ill-posed problem [32]. The surfaces are usually made differentiable by Gaussian filtering which distorts the geometry of the original signal, i.e., the details we are interested in. By using high-breakpoint robust estimators, the differential properties are less likely estimated over discontinuities. On smooth surfaces and if no outliers occur, the classical method performs slightly better than the robust method.

The analysis of the fit quality and standardized residuals provides powerful tools for diagnostics. To detect outlying points the standardized residual \( r_{i,j}/\sigma \) of each sample from the obtained surface is plotted. The MAD scale estimate was used as an estimate for standard deviation \( \sigma \) in both methods. The samples where the standardized residual is greater than threshold value, for example \( 2.5 \times \sigma \), can be labeled as outliers. The absolute values of the normalized distances for the noisy image used in Figure 5 are shown in Figure 15 where dark intensity values indicate very deviating samples. A 7-by-7 neighborhood was used. In the data set, the impulsive noise peaks appear as outliers. Outliers occur also in the corners because of inappropriate model in the fitting procedure. The robust method detects more outliers from the image. To be able to find outliers using least squares estimator the distance measure have to be robust, therefore it should not be based on parameters like the mean and the standard deviation that are not robust themselves.
Figure 14: Classification results using second fundamental form coefficients: Bright regions are hyperbolic surfaces, medium bright elliptic surfaces, medium dark parabolic surfaces and dark areas are planar regions. On the left the results are obtained using least squares and on the right using robust method.

We also consider the influence of deviations from the assumption on the smoothness (differentiability) of the data. Local window operators that estimate surface differential properties assume that the underlying surface is continuous. This assumption is not always valid for real data because the neighborhood used for estimation may have samples from different populations. In order to demonstrate the performance when there is more than one population present in the neighborhood, the test images include zero-, first- and and second-order surface patches, and discontinuities in between. The quality factors $E_j$ from both fit procedures were plotted for to show where the largest errors occur. The quality factors for each neighborhood are depicted in Figure 16 where dark areas indicate poorer quality. The results indicate that for least squares estimation the largest errors occur near surface discontinuities, and LTS procedure is able to avoid them because it uses only the population that represents the majority of the points in the neighborhood.

The validity of the functional model was examined in 5-by-5 neighborhood by using variable-order model and only second-order model in the LTS fit. The fit quality factors for the noise-free data set used in Figure 15 are plotted in Figure 17 where dark
Figure 15: The normalized distances $r_{i,j}/\sigma$ are plotted to detect outlying points in the intensity image: (from left) the original image, standardized residuals using least squares method and robust method, respectively. The darker the value the higher the normalized distance.

Figure 16: Fit quality $E_l$ using :a) LSQ and b) LTS method. Dark areas indicate poorer quality.

pixels mean poorer fit quality. The results show that the variable-order model method provides better quality fits because it is able to select an appropriate model depending on the order of the surface. Some samples are less likely classified to be outliers because of model failure because a functional model that provides a better understanding of the underlying digital surface is used. If the image has substantial nonimpulsive noise the second-order model is selected more frequently. The models in the variable-order method are not appropriate for estimating corners in images, hence the corners are distorted in the final output.
Figure 17: Diagnosing model validity using a) variable order and b) only second order model in the fit procedure. Dark areas indicate poorer fit quality.
5 Conclusion

In this paper we propose two restoration filters based on robust theory. The filtering methods are based on the Least Trimmed Squares error norm. The first one (RLTS) is a regression filter which employs up to second-order functional models to compute the output while the second one (RLTS-0) is an interesting special case of regression filter using only zero-order (constant) model. We are interested in restoring the shape of the original noise-free signal with as little distortion as possible while effectively attenuating different types of noise. We demonstrate the performance of the filters if the noise is not Gaussian, if there occurs deviations from the assumptions on functional model of the underlying data, or if there is more than one statistical population present in the processing window. Experimental results are shown using both impulsive and nonimpulsive noise and piecewise zero-, first- and second-order data sets. Both synthetic and real data and 1-D and 2-D signals are used as test data. RLTS-0 and median filtering distort the shape of roof edges and second-order data, while RLTS filter is able to preserve more complicated signal shapes because an appropriate model is used in the filtering. It is useful for filtering geometric signal shapes, for example in range images, where roof edges and higher-order surfaces occur frequently. Moreover, robust filters are very likely to compute the filter output using the samples that represent the majority in the processing window, which is important to be able to preserve discontinuities. The proposed filters can suppress both impulsive and nonimpulsive noise effectively. The results indicate that the RLTS-0 filter attenuates noise slightly better than the median if the signal is piecewise-constant. In real images that assumption is typically not valid, and hence the RLTS filter has the best performance in restoring the original signal shape. The employed variable-order functional model provides more signal understanding. The filter output is selected from the order giving the best quality. The RLTS is a good choice for filtering tasks that require attenuation of various types of noise and simultaneously accurate restoration of the original signal shape. The computational complexity of the filters based on Least Trimmed Squares estimation is higher than that of median filtering because of the additional ordering of the residuals.

Least Trimmed Squares method is applied for estimating differential properties of digital surfaces. The proposed method can be used in segmentation phase of an image understanding task to extract homogeneous surfaces from the image. We compare the performance to constant coefficient window operators. We show experimental results emphasizing the ability to produce quantitative information about the quality of the processing that can be used for diagnosing the results, for example, for outlier detection. The experimental results show that the robust approach produces reliable fit results also near region boundaries without any a priori information about discontinuities. The surface coefficients are estimated using samples from the population representing
the majority in the neighborhood. This is important, because very large errors occur when the coefficients are estimated over discontinuities. A variable-order method is employed to use an appropriate functional model in the estimation process. We show results of simple surface classification based on the second fundamental form coefficients from differential geometry. The classification results were used to find geometrically homogeneous surface patches from the image.

We would like to thank Prof. Max Mintz and Prof. Saleem A. Kassam, University of Pennsylvania, for their valuable comments. The Academy of Finland, the University of Oulu, and the Foundation Suomen Kulttuurirahasto are gratefully acknowledged for financial support. The facilities were partially provided by Navy Grant N00014-92-J-1647, AFOSR Grant 88-0296; Army/DAAL 03-89-C-0031PRI; NSF Grants CISE/CDA 88-22719, IRI 89-06770, and ASC 91 0813; and Du Pont Corporation.

References


