Robust Pixel Unmixing

Paul L. Rosin

***Abstract*—Pixel unmixing is commonly performed by employing a least squared (LS) error criterion, making it sensitive to outliers. As an alternative, the least median of squares (LMedS) method is proposed. Not only is it extremely robust, but it is efficient and straightforward both to implement and use.**

***Index Terms*—Image classification, least mean square methods, robustness.**

I. Introduction

A COMMON problem in remote sensing is that the limited spatial resolution of the scanner inevitably leads to "mixed" pixels at class boundaries. That is, individual pixels cover more than one ground cover type. This leads to the spectral response at a pixel being a mixture of the underlying pure classes—the so-called endmembers. If standard single class per pixel classification is performed, then in the best case, accuracy is compromised since a fraction of the pixel is incorrectly classified. In the worst case, mixing may produce a confusing spectral combination such that the pixel's classification is totally incorrect. A better solution which reduces these two sources of error is to model the spectral mixture and classify at each pixel the proportions of the endmember classes.

The simplest, and most widely used approach is the linear mixture model [1], [6], which models the mixture as a linear combination of the endmembers

$$r_b = \sum_{i=1}^{N} f_i e_{bi}$$

where

- r_b measured reflectance in band b;
- f_i fraction of the ith endmember;
- e_{bi} known reflectance of that endmember in that band;
- N number of endmembers.

In addition, the two constraints $\sum f_i = 1$ and $f_i \geq 0$ should be satisfied. Assuming that the endmembers are not linearly dependent, the mixing fractions can be recovered from the data. For a given number of spectral bands B if $B = N - 1$, then an exact solution can be found for the mixture model for each pixel, while if $B > N - 1$, then a least squares (LS) fitting procedure can be applied to obtain the best fit. Alternatively, the measurements from a set of pixels can be pooled together and the whole set classified. Combining data in this way potentially enables more accurate classification. However, the danger lies in outliers, that is, pixels with atypical values, being contained in the set.

More precisely, the sensitivity of an estimator to outliers can be quantified by its *breakdown point*. This specifies the fraction of data that can be contaminated (i.e., become outliers) without affecting the result of the estimator. The standard LS method has a zero breakdown point since even a single data value can have an enormous effect on the result. This has led to the field of *robust statistics* [9], in which more reliable methods have been developed. The following sections describe one such approach—*the least median of squares (LMedS) estimator*—and show its application to pixel unmixing.

Manuscript received January 19, 2001; revised April 25, 2001.

The author is with the Department of Computer Science, Cardiff University, Cardiff CF24 3XF, U.K. (e-mail: Paul.Rosin@cs.cf.ac.uk).

Publisher Item Identifier S 0196-2892(01)07620-3.

II. LMedS Pose Estimation

The LMedS [9] estimator is simple in essence. Instead of minimizing the usual squared residuals, it minimizes the *median* of the squared residuals. At the expense of accuracy the median operation gains the LMedS substantial robustness as the high errors incurred by outliers are ignored. In fact, it has a breakdown point of 0.5, so that, as long as the outliers make up less than half the dataset, the LMedS should return a reasonable estimate.

In one dimension, its implementation is simple and efficient. The data samples $\{x_i\}_{i=1\ldots n}$ are presorted, and the estimate $\hat{x}$ satisfying

$$\min_{\hat{x}} \operatorname*{med}_i (\hat{x} - x_i)^2$$

is calculated as

$$\hat{x} = \frac{x_{t+n/2} + x_t}{2}$$

such that

$$t = \arg\min_i x_{i+n/2} - x_i.$$

In other words, the window containing half the data elements is found that minimizes the range in the values, and the mean of the range is returned. It should be noted that the LMedS estimate is not guaranteed to be unique (e.g., consider its application to a uniform distribution), but in practice this is rarely a problem.

For higher dimensions, a different procedure is required. In this paper, we use the minimal subset approach [8] to approximate the optimal LMedS solution. A minimal subset is the minimum number of data values sufficient to determine a unique solution. Combinations of values are used to generate different minimal subsets, each of which provides a possible estimate. For each minimal subset, the median squared error over the complete dataset is computed. The final result is the estimate which produces the minimum median squared error.

An important practical issue is that larger minimal subsets permit more combinations of data values, and therefore give rise to more possible subsets. For instance, fitting ellipses to two-dimensional (2-D) data points requires minimal subsets of size six [8]. This leads to $O(n^6)$ possible subsets, but it is usually too computationally expensive to analyze them all. The solution is to randomly select a smaller number of subsets while maintaining a high probability that at least one of them will be

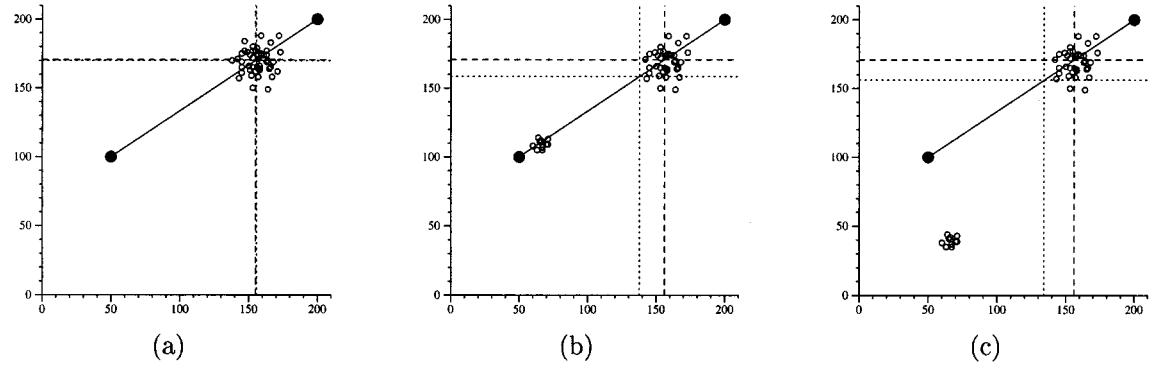

Fig. 1. Plot of synthetic two-band reflectance data. The mixed values obtained from the LS estimate of the mixing proportions is shown by dotted-line cross hairs, while the LMedS estimate is shown by the dashed-line cross hairs.

uncorrupted by outliers. Rousseeuw [9] showed that if there is a proportion of outliers ϵ in a subset of size S then the number of subsets N that must be chosen to include with $100 \times C\%$ confidence one composed entirely of inliers (which is necessary for obtaining a correct estimate) is

$$N = \frac{\ln(1-C)}{\ln(C-(1-\epsilon)^S)}. \tag{1}$$

Even for high levels of confidence and high dimensions, this results in a small workable number (e.g., only 59 for 95% confidence, 50% outliers, and $S = 100$), and ensures that the overall complexity is linear.

As noted above, the *statistical efficiency* of the LMedS is inferior to the LS estimator. Moreover, the minimal subset method is still less accurate. The estimate is only calculated from a few values which even if they are inliers are still prone to low levels of noise, and so it is unlikely to provide a optimal minimum median squared error. To combine the benefits of both the LMedS and LS approaches, a polishing stage can be applied. The estimate produced by the LMedS is refined by a LS fit applied to the inliers only. The unpolished estimate provides residuals e_i for each pixel. Outliers are then identified by testing if

$$|e_i| > 3 \times 1.4826 \times \text{mad}$$

where the median absolute deviation (mad) robustly estimates the standard deviation and is calculated as

$$\text{mad} = \underset{i}{\text{med}}\,|e_i|$$

and 1.4826 is a normalizing factor to make the mad equivalent to the standard deviation for Gaussian distributed data.

In summary, the algorithm is as follows:

1) Randomly sample pixel tuples.
2) Determine the mixing proportions from each tuple.
3) Select the mixing proportions that minimize the median error over the complete dataset.
4) Identify outliers.
5) Polish the estimate, i.e., redetermine the mixing proportions from the inliers.

Many approaches to mixture modeling are possible, e.g., support vector machines [3], neural networks [4], and the EM algorithm [10] have all been applied to remote sensing data. These techniques could be used to provide local estimates that are fed into the LMedS method as described above. In most of the experiments in this paper, the most straightforward approach has been applied, namely, linear mixing with the endmember fractions estimated by simple LS if the problem is overdetermined. The fractions are not constrained to be positive; this requires more sophisticated techniques such as convex quadratic programming.[1]

Fig. 2. Blue band of the Glandorf data with superimposed boundaries.

The LMedS estimator is demonstrated on a simple two-band synthetic example containing two endmembers at (50, 100) and (200, 200) mixed with proportions 0.3 and 0.7, respectively. Fifty samples with added Gaussian noise were generated and are shown in Fig. 1(a) where the axes represent the two bands and the small circles show the reflectances of the samples. The endmembers are plotted as larger filled circles. The reflectances obtained by applying the estimated mixing proportions are shown by the cross hairs (dotted and dashed lines for LS and LMedS, respectively). For such uncorrupted data, both the LS and LMedS estimates are naturally good, giving very similar estimates. However, when some outliers are added, as in Fig. 1(b), the LS estimate is pulled toward them, degrading the fit to the inliers. Fig. 1(c) shows that the more distance the

[1]In practice, a simpler approach is often carried out: unconstrained fitting is applied, any negative fractions are set to zero, and the remaining fractions renormalized. However, this step was not performed in the experiments described in this paper.

TABLE I
MIXING PROPORTIONS ESTIMATED BY THE LS AND LMedS ESTIMATORS ON MEDIUM SIZED DATASETS (150–200 PIXELS)

ground truth			LS estimate				LMedS estimate			
house	road	fields	house	road	fields	error	house	road	fields	error
0.0	0.9	0.1	-0.049	0.878	0.171	0.142	-0.028	0.825	0.203	0.207
0.1	0.0	0.9	0.167	0.081	0.752	0.296	0.145	0.012	0.843	0.113
0.2	0.0	0.8	0.195	0.072	0.733	0.145	0.207	0.025	0.768	0.065
0.3	0.0	0.7	0.274	0.174	0.552	0.348	0.298	0.095	0.608	0.189
0.4	0.0	0.6	0.358	0.130	0.511	0.261	0.352	0.134	0.514	0.267
0.5	0.0	0.5	0.595	0.031	0.375	0.250	0.405	0.038	0.557	0.190
0.6	0.0	0.4	0.645	0.090	0.265	0.270	0.631	-0.017	0.386	0.063
0.7	0.0	0.3	0.667	0.020	0.314	0.067	0.691	-0.019	0.328	0.056
0.8	0.0	0.2	0.763	0.001	0.237	0.075	0.795	-0.038	0.243	0.085
0.9	0.0	0.1	1.031	0.041	-0.072	0.345	0.888	-0.017	0.129	0.058
			average absolute error:			0.220	average absolute error:			0.129

outliers lie from the inliers, the greater their effect. In contrast, the LMedS estimate is essentially unaffected by the relatively small group of outliers.

III. RELATED WORK

In previous work on pixel unmixing, there has been little emphasis on ensuring robust performance. An exception is the work by Bosdogianni *et al.* [2], who use the Hough Transform. The approach is somewhat similar to the minimal subset one, as tuples of pixels generate estimates. These are cast as votes and accumulated in a parameter space. Peaks in parameter space then correspond to the final estimates.

As with LMedS, the Hough Transform is robust since outliers do not greatly detract from the main peaks. It also has the advantage that, if the outliers are coherent, then they can be detected as subsidiary peaks. However, there are several disadvantages. The first is that the volume of the parameter space scales to the power of the number of parameters, which becomes prohibitive for more than a few dimensions. Second, the parameter space needs to be quantized. If the wrong bin size is selected this can affect the likelihood of detecting the peak and the accuracy of the estimate. Third, peak detection is problematic, and often requires some filtering (e.g., smoothing) of the parameter space as well as thresholding to eliminate insignificant peaks. In contrast, the LMedS approach is nonparametric, and its performance does not depend on tuning to ensure that the various stages work effectively.

There are also many other possible approaches found in the robust statistics literature. However, not all have such high breakdown points. For instance, the breakdown point of some, such as the Theil–Sen Estimator [11], decreases as a function of the dimensionality of the estimation problem. A problem with others is that various parameters need to be set. For instance, Huber's M-estimators [5] use a redescending kernel to reduce the effect of outliers, and the size and shape of the kernel need to be specified. Thus, the LMedS remains competitive in terms of its ease of use as well as its effectiveness.

IV. EXPERIMENTAL RESULTS

Like Bosdogianni *et al.* [2], we demonstrate the approach on a three-class and three-band example. Fig. 2 shows the blue band from the Glandorf dataset provided by the ISPRS Working Group III/3. Overlaid are the three classes (streets, houses, and fields). Although these are the endmembers, in practice, they can be seen to be rather varied. The image was then averaged over 30×30 blocks to generate mixed pixels. The ground truth for the mixing fraction values was taken as the proportions of the three classes in their generating blocks. From these were extracted ten sets that contained similar proportions of classes, each containing between 150 and 200 pixels,[2] and another ten sets containing between 20 and 30 pixels.

Since for three classes only two bands are required to estimate the mixing proportions, the exact solution is quite tractable

$$f_1 = \frac{r_2y_1 - r_1y_2 - r_2z_1 + y_2z_1 + r_1z_2 - y_1z_2}{x_2y_1 - x_1y_2 - x_2z_1 + y_2z_1 + x_1z_2 - y_1z_2}$$

$$f_2 = \frac{r_2x_1 - r_1x_2 - r_2z_1 + x_2z_1 + r_1z_2 - x_1z_2}{-x_2y_1 + x_1y_2 + x_2z_1 - y_2z_1 - x_1z_2 + y_1z_2}$$

$$f_3 = 1 - f_1 - f_2$$

where for ease of notation we have written the spectral responses for the pure endmembers classes in band i as $\{x_i, y_i, z_i\}$; the pixel intensities in the two bands are denoted r_1 and r_2. However, in practice all three bands per pixel are used to provide some additional accuracy. In other words, each pixel is considered as a slightly expanded minimal subset. This means that there is no combinatorial explosion of the minimal subsets (the complexity remains linear), and it is possible to use every pixel to generate an estimate.

The results of the LS and LMedS estimates of the mixing proportions are given in Table I. In most cases, the LMedS method performed better, and the overall accuracy of its estimates is clearly superior. To gain a better understanding of the estimation process, the red- and green band values are plotted for the data points used to generate the results in rows 5 and 8 of Table I, and are shown in Fig. 3(a) and (b). All pixels are plotted as small circles while detected inliers are also filled with a cross. The fields and houses endmembers are represented by the large square and circle, respectively. We note that there is a wide spread in values, leading to the relatively low accuracies in the estimates.[3] Never-

[2]The larger sets contain only two mixed classes as combinations of three classes did not occur frequently with the same mixing proportions.

[3]Better results could be obtained by splitting classes into subcategories to ensure that they are more homogeneous. However, the data and models have been left in their raw state to show the effectiveness of the technique even under such difficult circumstances.

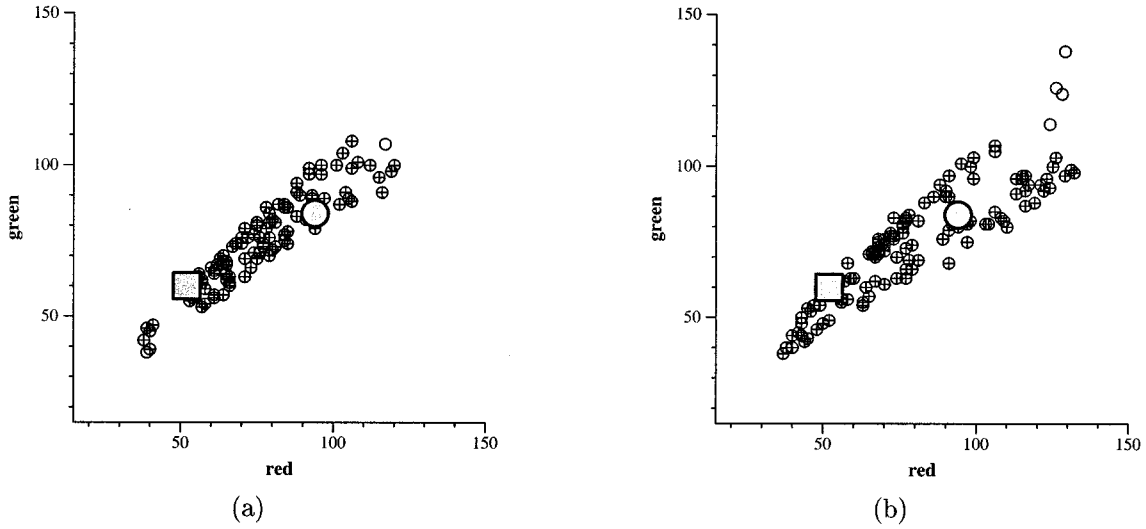

Fig. 3. Pixels from the medium sized datasets analyzed in row 5 and 8 of Table I. Inliers are crossed. Endmembers are shown as a large circle (houses) and a rectangle (fields). (a) Row 5 endmember fractions: 40% house, 60% fields. (b) Row 8 endmember fractions: 70% house, 30% fields.

TABLE II
MIXING PROPORTIONS ESTIMATED BY THE LMedS ESTIMATOR USING ONLY FOUR MINIMAL SUBSETS

ground truth			without polishing				with polishing			
house	road	fields	house	road	fields	error	house	road	fields	error
0.0	0.9	0.1	-0.560	1.107	0.452	1.119	-0.028	0.825	0.203	0.207
0.1	0.0	0.9	0.408	-0.105	0.697	0.615	0.164	0.031	0.806	0.189
0.2	0.0	0.8	0.086	0.019	0.895	0.228	0.207	0.025	0.768	0.065
0.3	0.0	0.7	-0.170	0.343	0.827	0.941	0.298	0.095	0.608	0.189
0.4	0.0	0.6	-0.428	0.486	0.942	1.655	0.352	0.128	0.519	0.257
0.5	0.0	0.5	0.159	0.078	0.763	0.682	0.423	0.034	0.543	0.155
0.6	0.0	0.4	0.403	-0.014	0.611	0.421	0.631	-0.017	0.386	0.063
0.7	0.0	0.3	0.021	0.223	0.757	1.358	0.691	-0.019	0.328	0.056
0.8	0.0	0.2	-0.199	0.269	0.930	1.998	0.785	-0.038	0.253	0.106
0.9	0.0	0.1	-0.283	0.426	0.857	2.367	0.888	-0.017	0.129	0.058
			average absolute error:			1.138	average absolute error:			0.135

theless, the values are fairly well clustered in Fig. 3(a) which is why very few points have been marked as outliers. This explains why the LS and LMedS gave similar results for this dataset. Fig. 3(b) shows a more distinct group of outliers that are correctly flagged, leading to an improvement in the LMedS estimate.

Equation (1) specifies that only four pixels are required to obtain with 95% confidence a robust fit. The results of testing this are given in Table II. Although the minimal subsets provide noisy estimates, they are still able in most cases to identify outliers accurately enough that the polishing stage (applied to all inliers) is able to recover estimates comparable with the estimates obtained when all pixels are used as minimal subsets.

Working in the opposite direction, since larger subsets provide the LMedS algorithm with more accurate estimates (as long as they are not contaminated by outliers) this could be used to improve the unpolished estimate, leading to more accurate detection of outliers, and thereby a more accurate overall polished fit. As an example, from (1) we see that if ten tuples are used, then 58 subsets are enough to probably choose one without outliers. As an experiment, the collection of single pixel estimates used in Table I were augmented with 58 randomly selected subsets of ten pixels before the LMedS solution was determined. In this case, it was found to improve the estimator's performance. However, there is no guarantee that this procedure will not lead to a drop in performance in some instances. For instance, Fig. 4

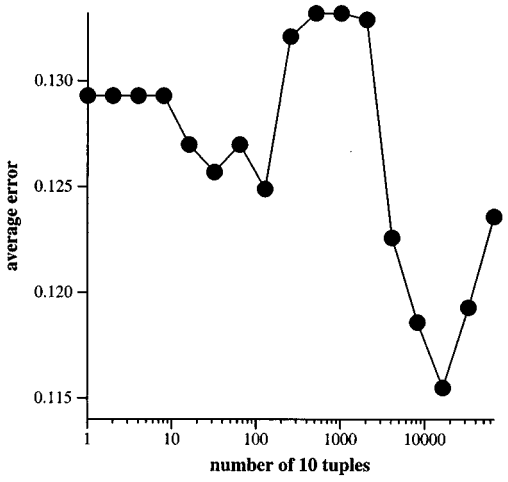

Fig. 4. Average error rate for the ten test datasets when additional ten tuples are used by the LMedS estimator in addition to the individual pixel minimal subsets.

TABLE III
MIXING PROPORTIONS ESTIMATED BY THE LS AND LMedS ESTIMATORS ON SMALL DATASETS (20–30 PIXELS)

ground truth			LS estimate				LMedS estimate			
house	road	fields	house	road	fields	error	house	road	fields	error
0.02	0.15	0.83	-0.018	0.589	0.430	0.877	-0.110	0.244	0.866	0.261
0.20	0.50	0.30	0.234	0.564	0.202	0.195	0.234	0.564	0.202	0.195
0.10	0.50	0.40	0.237	0.556	0.207	0.387	0.237	0.556	0.207	0.387
0.05	0.10	0.85	-0.017	0.526	0.490	0.853	-0.120	0.311	0.808	0.423
0.35	0.50	0.15	0.203	0.668	0.129	0.337	0.203	0.668	0.129	0.337
0.30	0.50	0.20	0.192	0.669	0.139	0.338	0.192	0.669	0.139	0.338
0.25	0.50	0.25	0.288	0.719	-0.007	0.513	0.322	0.744	-0.066	0.632
0.40	0.50	0.10	0.291	0.803	-0.095	0.606	0.388	0.648	-0.037	0.297
0.15	0.50	0.35	0.290	0.672	0.038	0.624	0.324	0.357	0.319	0.348
0.45	0.50	0.05	0.283	0.893	-0.177	0.787	0.352	0.716	-0.068	0.431
			average absolute error:			0.552	average absolute error:			0.365

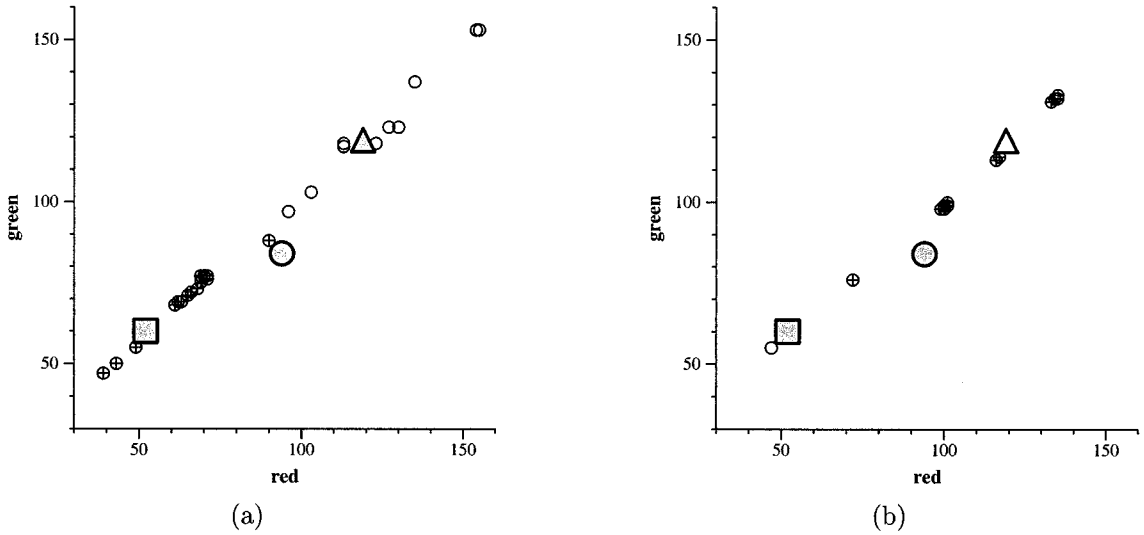

Fig. 5. Pixels from the small datasets analyzed in rows 1 and 7 of Table III. Inliers are crossed. Endmembers are shown as a large circle (houses), a triangle (streets), and a rectangle (fields). (a) Row 1 endmember fractions: 2% house, 15% road, 83% fields. (b) Row 7 endmember fractions: 25% house, 50% road, 25% fields.

TABLE IV
MIXING PROPORTIONS ESTIMATED BY THE LS AND LMedS ESTIMATORS WITH ADDED CONSTRAINTS ON MEDIUM SIZED DATASETS (150–200 PIXELS)

ground truth			LS estimate				LMedS estimate			
house	road	fields	house	road	fields	error	house	road	fields	error
0.0	0.9	0.1	0.000	0.854	0.146	0.093	0.000	0.811	0.189	0.179
0.1	0.0	0.9	0.167	0.081	0.752	0.296	0.185	0.049	0.766	0.267
0.2	0.0	0.8	0.195	0.072	0.733	0.145	0.212	0.063	0.725	0.150
0.3	0.0	0.7	0.274	0.174	0.552	0.348	0.303	0.100	0.597	0.206
0.4	0.0	0.6	0.358	0.130	0.511	0.261	0.373	0.122	0.504	0.244
0.5	0.0	0.5	0.595	0.031	0.375	0.250	0.616	0.013	0.371	0.259
0.6	0.0	0.4	0.645	0.090	0.265	0.270	0.692	0.002	0.305	0.189
0.7	0.0	0.3	0.667	0.020	0.314	0.067	0.665	0.000	0.335	0.070
0.8	0.0	0.2	0.763	0.001	0.237	0.075	0.741	0.000	0.259	0.117
0.9	0.0	0.1	0.895	0.105	0.000	0.210	0.967	0.000	0.033	0.134
			average absolute error:			0.201	average absolute error:			0.181

shows the effect of adding a variable number of extra ten tuples. In most but not all instances, the error rate is unaffected or reduced.

Table III lists the results of processing the small datasets. As expected both the LS and LMedS methods are less accurate than for the larger datasets, but the LMedS estimates are generally better. Again, two rows (1 and 7) are visualized to aid understanding of the process. Fig. 5(a) shows all three endmembers and it can be seen that they are almost linear, making unmixing less reliable. Because of the sparsity of the data it is more difficult to reliably determine the outliers. This is especially difficult in these examples since the data values are widely spread out. A cluster about the fields endmember is correctly detected in Fig. 5(a), leading to the top end of the spread of values being rejected, and substantially improving the final estimate. Again, Fig. 5(b) shows that the worst performance of the LMedS oc-

curs when there is no distinct cluster so that outliers cannot be identified and rejected.

Finally, results are shown in Table IV where the estimation of the endmembers is constrained to ensure that they are nonnegative and sum to one.[4] As before, the LMedS method gives an improved overall estimate although the difference is less marked than appeared in the results of the unconstrained fitting presented in Table I.

V. Conclusion

This paper has proposed the LMedS estimator for pixel unmixing. Its advantages are that it is **computationally efficient** (i.e., linear in the number of points), **nonparametric** (i.e., easy to use without requiring user intervention and tuning), and **robust**, (i.e., it can operate effectively as long as 50% of the data remains uncorrupted). Experiments show its effectiveness compared to standard nonrobust estimation techniques.

References

[1] J. B. Adams and M. O. Smith, "Spectral mixing modeling: A new analysis of rock and soil types at the Viking Lander I site," *J. Geophys. Res.*, vol. 91, pp. 8098–8112, 1986.

[2] P. Bosdogianni, M. Petrou, and J. V. Kittler, "Mixed pixel classification with robust statistics," *IEEE Trans. Geosci. Remote Sensing*, vol. 35, pp. 551–559, May 1997.

[3] M. Brown, S. R. Gunn, and H.G. Lewis, "Support vector machines for optimal classification and spectral unmixing," *Ecol. Modeling*, vol. 120, pp. 167–179, 1999.

[4] G. M. Foody, "Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data," *Int. J. Remote Sens.*, vol. 17, pp. 1317–1340, 1996.

[4]For convenience, the constrained fitting was performed using simulated annealing [7]. The function optimized was the mean squared error with the addition of a highly weighted penalty term for each constraint.

[5] P. J. Huber, *Robust Statistics*. New York: Wiley, 1981.

[6] S. E. Marsh, P. Switzer, W. S. Kovalik, and R. J. P. Lyon, "Resolving the percentage of component terrains within single resolution elements," *Photogramm. Eng. Remote Sens.*, vol. 46, pp. 1079–1086, 1980.

[7] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vettering, *Numerical Recipes in C*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[8] G. Roth and M. D. Levine, "Extracting geometric primitives," *Comput. Vis. Graph. Image Process., Image Understanding*, vol. 58, no. 1, pp. 1–22, 1993.

[9] P. Rousseeuw and A. Leroy, *Robust Regression and Outlier Detection*. New York: Wiley, 1987.

[10] S. Tadjudin and D. A. Landgrebe, "Robust parameter estimation for mixture model," *IEEE Trans. Geosci. Remote Sensing*, vol. 38, pp. 439–445, Jan. 2000.

[11] H. Theil, "A rank-invariant method of linear and polynomial regression analysis," *Nederlandse Akademie Wetenchappen A.*, vol. 53, pp. 386–392, 1950.

Paul L. Rosin received the B.Sc. degree in computer science and microprocessor systems from Strathclyde University, Glasgow, U.K., in 1984, and the Ph.D. degree in information engineering from City University, London, U.K., in 1988.

Currently, he is a Senior Lecturer with the Department of Computer Science, Cardiff University, Cardiff, U.K. He was also a Lecturer with the Department of Information Systems and Computing, Brunel University, London, U.K., a Research Scientist with the Institute for Remote Sensing Applications, Joint Research Center, Ispra, Italy, and a Lecturer at Curtin University of Technology, Perth, Australia. His research interests include the representation, segmentation, and grouping of curves, knowledge-based vision systems, early image representations, machine vision approaches to remote sensing, and the analysis of shape in art and architecture.

Dr. Rosin is Secretary of the British Machine Vision Association.