Restoration of Hyperspectral Push-Broom Scanner Data

Rasmus Larsen, Allan Aasbjerg Nielsen & Knut Conradsen
Department of Mathematical Modelling, Technical University of Denmark

ABSTRACT: Several effects combine to distort the multispectral data that are obtained from push-broom scanners. We develop an algorithm for restoration of such data, illustrated on images from the ROSIS scanner. In push-broom scanners variation between elements in the detector array results in a strong striping along flight lines. A non-systematic striping is also present along flight lines. Furthermore, line drop-outs occur, and finally, various types of electronic noise of salt-and-pepper type are also present. We describe techniques for the correction for all these types of effects. Line drop-outs are located automatically using line means statistics, and if present new pixel values are interpolated from the neighbouring lines. Striping along and across flight lines is corrected for by adjusting the line and column means, respectively. This restoration is carried out in stationary parts of the image, for instance over water. Following these initial corrections we use minimum/maximum autocorrelation factor (MAF) analysis in order to separate the spatially coherent signal components from the noise components. The MAF transformation is a linear transformation into new orthogonal variables that are ordered by decreasing autocorrelation. In this way noise channels (with low autocorrelation) can be identified and cleaned or eliminated. Also, the MAF transformation enables us to isolate electronic or aircraft engine induced noise components that have a special spatial structure. Subsequent inverse transformation back into the original spectral space results in noise corrected variables. The noise components will now have been removed from the entire original data set by working on a smaller set of noise contaminated transformed variables only. The application of the above techniques results in a dramatic increase in visual image quality.

1 INTRODUCTION

In this paper we will consider the restoration of multispectral data from pushbroom scanners, illustrated on images from the ROSIS scanner. In pushbroom scanners variation across the elements of the detector array often results in strong striping along flight lines. Striping across flight tracks may also be present due to other types of variation in the instrument. Furthermore, line drop-outs occur, and finally various electronic and aircraft engine noise components are also present.

2 DATA

ROSIS (Reflective Optics System Imaging Spectrometer) is a compact airborne imaging spectrometer [6]. It is a pushbroom scanner with 500 detector elements across track. The scanner has two operational modes, imaging mode and spectral mode. In the spectral mode 81 channels are available in the range 440 to 852 nm. The sampling interval of the first 12 channels (440 to 572 nm) is 12 nm. For the remaining 69 channels (580 to 852 nm) the sampling interval is 4 nm. In the imaging mode only 29 of the 81 channels are available. However, we may obtain full spatial resolution across track in the imaging mode, i.e. 500 pixels. In the spectral mode the resolution across track is reduced by a factor 3.

In this paper we will illustrate the restoration techniques on a 1024×167 of a flight track recorded in spectral mode near Gyllingnæs i Jutland, Denmark.

3 METHODOLOGIES

3.1 Correction for line drop-outs

In Figure 1(a) the first channel is shown. The pixel counts are scaled linearly between mean ± 3 standard deviations. A series of black lines can be seen. From the line means shown in Figure 1(b) it is seen that the black lines do not correspond to zero counts. They do, however, differ signicantly in level from the other lines. We will use the deviation from the mean of the line means in a moving window as criterion for detec-

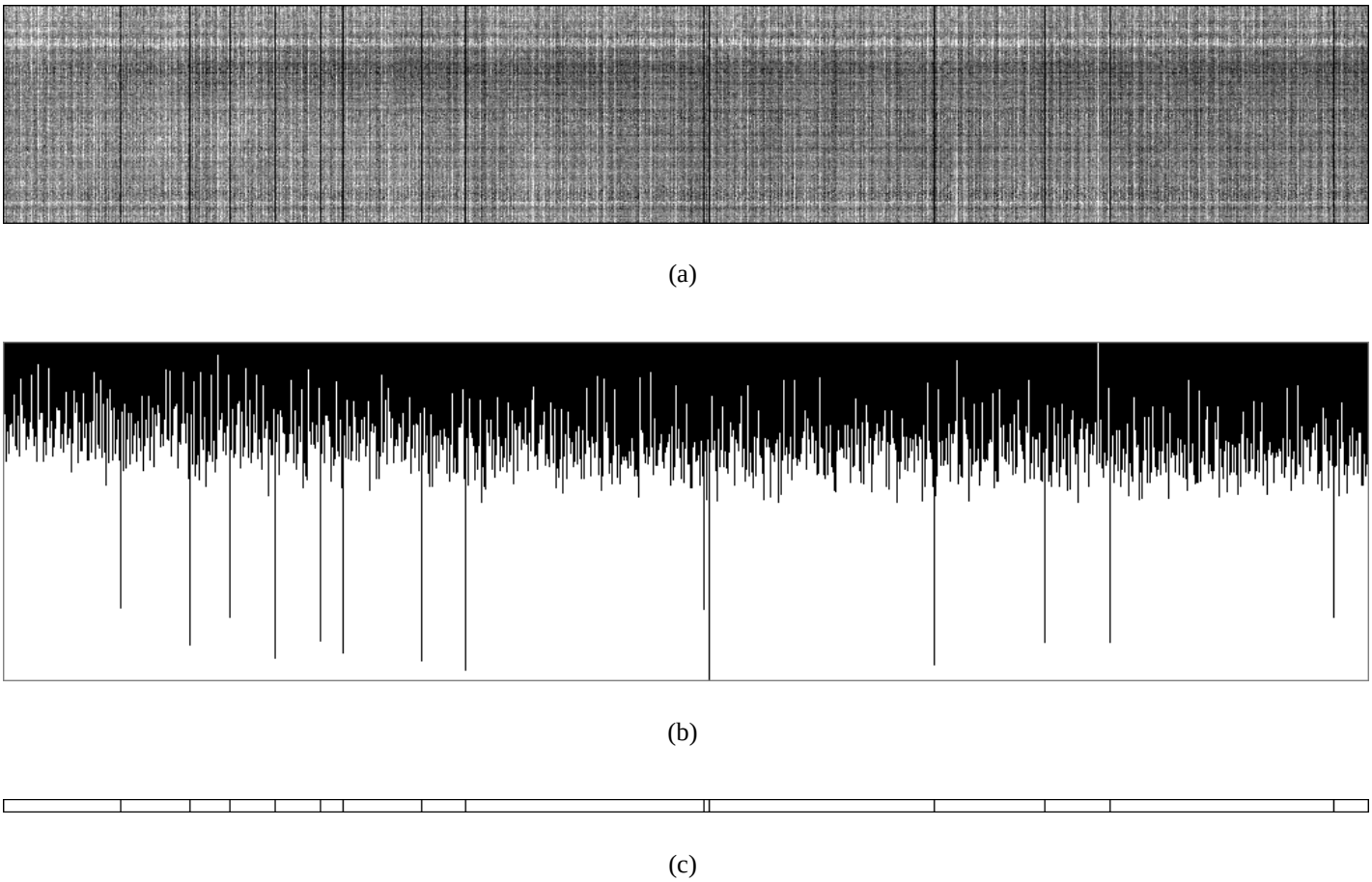


Figure 1: The defect lines in the original channels are identified by the deviation from the local mean level as defined by a robust statistic. In (a) the original channel 1 is shown, in (b) we see the line means, and on (c) we see the identified drop-out lines.

ting the drop-outs. The local means and the standard deviation are estimated using robust statistics [2]. First the local mean and standard deviation are estimated using all line means. After this all lines that deviate more from the local mean than 2 standard deviations are excluded from the estimations. New estimates of local mean and standard deviation are then obtained. Now all lines whose mean deviate more than 5 standard deviations from the robust estimate of the local mean using the yardstick given by the robust estimate of the standard deviation are defined as being drop-outs. The result is shown in Figure 1(c).

When the missing lines have been identified they are substituted by the mean value of the line above and the line below.

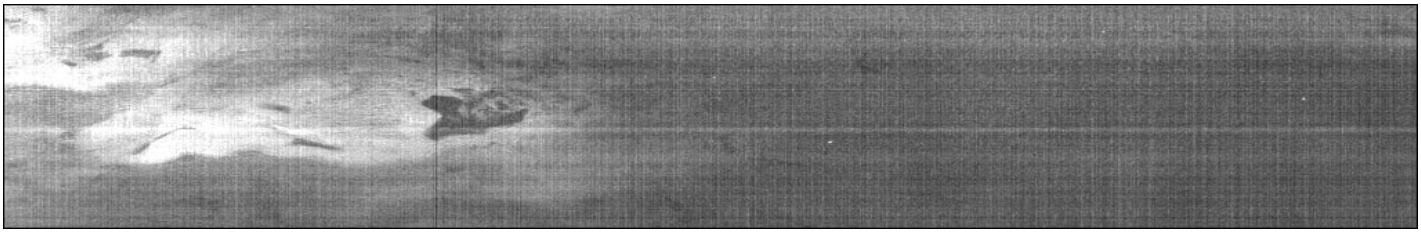
3.2 Correction for horizontal and vertical striping

We assume independence between the horizontal and the vertical striping. Furthermore, we assume that the striping is caused by additive row and column dependent terms. Because we only have the data it self to perform the analysis from, we will estimate the striping terms for each channel over an area with homogeneous signal. This area may be extracted in a number of ways. One way is to follow the robust statistics approach from the previous section. In our case we choose to manually define an homogeneous area that spans all columns and all rows. Under this area we estimate the additive striping terms. In Figure 2(a) the drop-out corrected channel 7 is shown, and in Figure 2(b) we see this channel first corrected for the column striping and in Figure 2(c) also for the line striping.

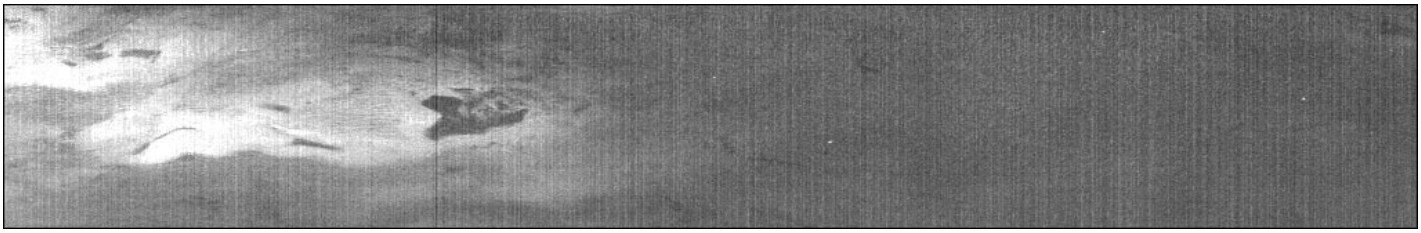
3.3 Correction for other noise components

The last procedure we will apply for the restoration is based on utilizing the redundancy that is present in hyperspectral datasets. We will extract those linear combinations that hold the signal variation, throw away the remaining subspace, and transform back to the original feature space.

As opposed to the celebrated principal components (PC) transformation the minimum/maximum autocorrelation factors (MAF) transformation allows for the spatial nature of image data. The application of this transformation requires knowledge of or an estimate of the dispersion matrix and the dispersion matrix of



(a)



(b)

(c)

Figure 2: Vertical and horizontal striping is assumed independent and removed by adjusting the level of each line and column. In (a) channel 7 corrected for drop-outs is shown, in (b) we see the same channel corrected for vertical striping, and in (c) corrected for horizontal striping as well.

the difference between the original and the spatially shifted image.

The MAF transformation minimizes the autocorrelation rather than maximizing the variance (PC). In reverse order the MAFs maximize the autocorrelation represented by each component. MAF one is the linear combination of the original bands that contains minimum autocorrelation between neighboring pixels. A higher order MAF is the linear combination of the original bands that contains minimum autocorrelation subject to the constraint that it is orthogonal to lower order MAFs. The MAF procedure thus constitutes a (conceptually) more satisfactory way of orthogonalizing image data than PC analysis. The MAF transformation is equivalent to a transformation of the data to a coordinate system in which the covariance matrix of the difference between the spatially shifted image data and the original image data is the identity matrix followed by a principal components transformation. An important property of the MAF procedure is its invariance to linear transformations, a property not shared by ordinary PC analysis. This means that it doesn't matter whether the data have been scaled e.g. to unit variance before the analysis is performed.

The MAF procedure was suggested in [5]. PCs, MAFs and other orthogonal transformations are described in [3, 1, 4].

We consider the random variable $\boldsymbol{Z}(\boldsymbol{x}) = [Z_1(\boldsymbol{x}), \ldots, Z_m(\boldsymbol{x})]^T$ and we assume that

$$\begin{aligned} \mathrm{E}\{\boldsymbol{Z}(\boldsymbol{x})\} &= \boldsymbol{0} && (1)\\ \mathrm{D}\{\boldsymbol{Z}(\boldsymbol{x})\} &= \boldsymbol{\Sigma}. && (2) \end{aligned}$$

We denote a spatial shift by $\boldsymbol{\Delta}^T = (\Delta_1, \Delta_2)$. We are interested in the correlations between projections of the variables and the shifted variables. Introducing

$$\begin{aligned} \boldsymbol{\Sigma}_\Delta &= \mathrm{D}\{\boldsymbol{Z}(\boldsymbol{x}) - \boldsymbol{Z}(\boldsymbol{x}+\boldsymbol{\Delta})\} \\ &= \mathrm{E}\{[\boldsymbol{Z}(\boldsymbol{x}) - \boldsymbol{Z}(\boldsymbol{x}+\boldsymbol{\Delta})] \\ &\qquad [\boldsymbol{Z}(\boldsymbol{x}) - \boldsymbol{Z}(\boldsymbol{x}+\boldsymbol{\Delta})]^T\}, \end{aligned} \tag{3}$$

which considered as a function of $\boldsymbol{\Delta}$ is a multivariate variogram, we have

$$\text{Corr}\{\boldsymbol{a}^T\boldsymbol{Z}(\boldsymbol{x}), \boldsymbol{a}^T\boldsymbol{Z}(\boldsymbol{x}+\boldsymbol{\Delta})\} = 1 - \frac{1}{2}\frac{\boldsymbol{a}^T\boldsymbol{\Sigma}_\Delta\boldsymbol{a}}{\boldsymbol{a}^T\boldsymbol{\Sigma}\boldsymbol{a}}. \quad (4)$$

If we want to minimize that correlation we must maximize the Rayleigh coefficient

$$R(\boldsymbol{a}) = \frac{\boldsymbol{a}^T\boldsymbol{\Sigma}_\Delta\boldsymbol{a}}{\boldsymbol{a}^T\boldsymbol{\Sigma}\boldsymbol{a}}. \quad (5)$$

Let $\lambda_1 \geq \cdots \geq \lambda_m$ be the eigenvalues and $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m$ corresponding conjugate eigenvectors of $\boldsymbol{\Sigma}_\Delta$ with respect to $\boldsymbol{\Sigma}$. Then

$$\boldsymbol{Y}_i(\boldsymbol{x}) = \boldsymbol{a}_i^T\boldsymbol{Z}_i(\boldsymbol{x}) \quad (6)$$

is the i'th minimum/maximum autocorrelation factor or shortly the i'th MAF. The reverse numbering of MAFs so that the signal MAF is referred to as the first MAF or MAF1 is often used, also here.

The first five of the odd numbered MAFs are shown in Figure 3. It is evident that the content of spatially autocorrelated signal is decreasing with component number. It should be noted at this point that in this case the MAF analysis performs satisfactorily only after the preprocessing steps described in Sections 3.1 and 3.2. Some of the components that contain signal with high autocorrelation are composed of natural surface variation, others like MAF7 in Figure 3(d) contain noise. This particular noise component is present in a number of the original bands. The MAF analysis isolates it very neatly from the rest of the signal (and noise). The origin of this component is unknown to the authors. Also, in MAFs 3 and 5 it is evident that the detector elements in the right hand side of the detector array have induced some sort of vertical blurring that is not present in the left hand side.

In order to use the MAF analysis for restoration we exclude all noise components before transforming from the MAF space to the original space. By excluding all high number MAFs (from number 16 and up) and also MAF7 in Figure 3(d) and a similar one in MAF8 prior to transforming back to the original space all the (noise) variation in these components is removed.

Examples of the restored channels are shown in Figure 4

4 CONCLUSION

In this paper we have used a series of techniques for restoring data from the pushbroom ROSIS scanner. The data were disturbed by a number of defects. First, line drop-outs were identified by measuring the deviation of the line mean from the local mean of line means. This was done using robust statistics. Then, the data were corrected for horizontal and vertical striping by adjusting line and column means, respectively. This approach is sensitive to correlation between the signal and the striping. However, by using areas of signal homogeneity as described by the operator the adjustment terms may be estimated. Finally the minimum/maximum autocorrelation factor algorithm was used to extract the relevant information in a lower dimensionality space. It should be noted here that this subspace potentially also contains noise that expresses high autocorrelation. In the sample case two components corresponding to some low (spatial) frequency variation in the system were isolated. These special noise components as well as those corresponding to noise of the salt-and-pepper type were removed before transforming back to the original space, thus resulting in the final restoration of the data.

5 ACKNOWLEDGEMENTS

The image data were put at our disposal by M.Sc. Peter Viskum Jørgensen of Computer Resources International A/S.

REFERENCES

[1] Knut Conradsen, Bjarne Kjær Nielsen, and Tage Thyrsted. A comparison of min/max autocorrelation factor analysis and ordinary factor analysis. In *Proceedings from Symposium in Applied Statistics*, pages 47–56, Lyngby, January 1985.

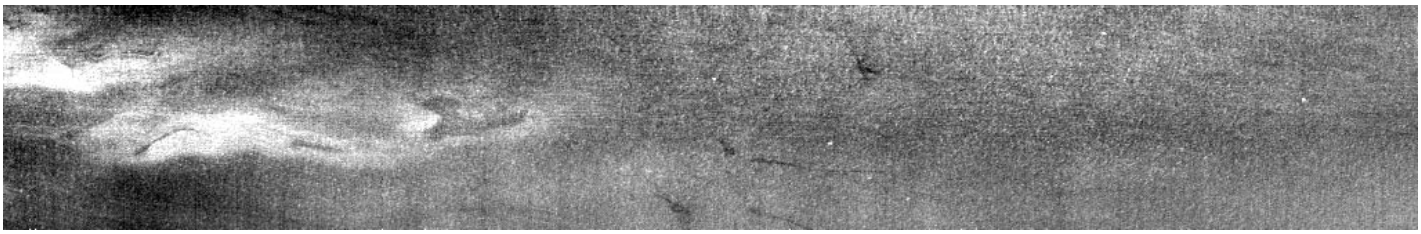
[2] Susan J. Devlin, R. Gnanadesikan, and J. R. Kettenring. Robust estimation of dispersion matrices and principal components. *Journal of the American Statistical Association*, 76:354–362, 1981.

[3] Bjarne Kjær Ersbøll. *Transformations and classifications of remotely sensed data*. PhD thesis, Institute of Mathematical Statistics and Operations Research, Technical University of Denmark, Lyngby, 1989. 297 pp.

[4] Allan Aasbjerg Nielsen. *Analysis of Regularly and Irregularly Sampled Spatial, Multivariate, and Multi-Temporal Data*. PhD thesis, Department of Mathematical Modelling, Technical University of Denmark, Lyngby, 1994. xxiv + 213 pp.

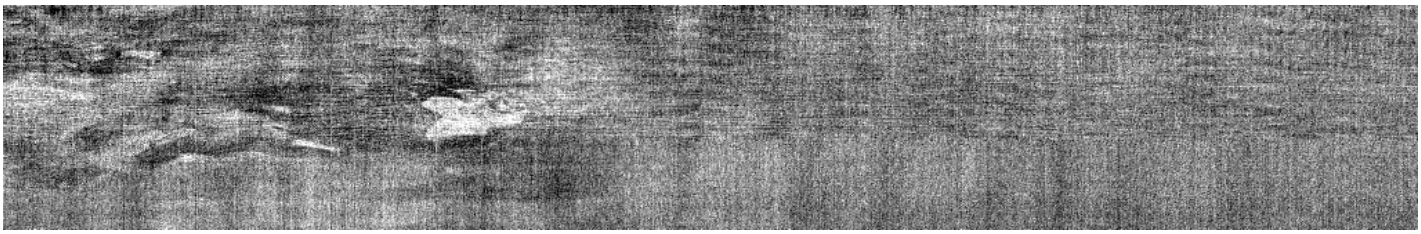
[5] Paul Switzer and Andrew A. Green. Min/max autocorrelation factors for multivariate spatial statistics. Technical Report 6, Stanford University, 1984. 10 pp.

[6] H. van der Piepen and R. Doerffer. The imaging spectrometer ROSIS specifications. Technical report, Deutsche Luft- und Raumfart, Institute of Optoelectronics, Oberpfaffenhofen, 1994. 5 pp.

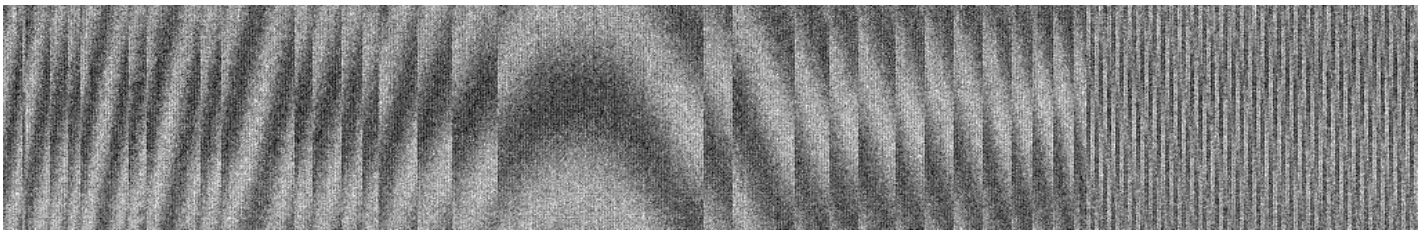


(a) MAF1

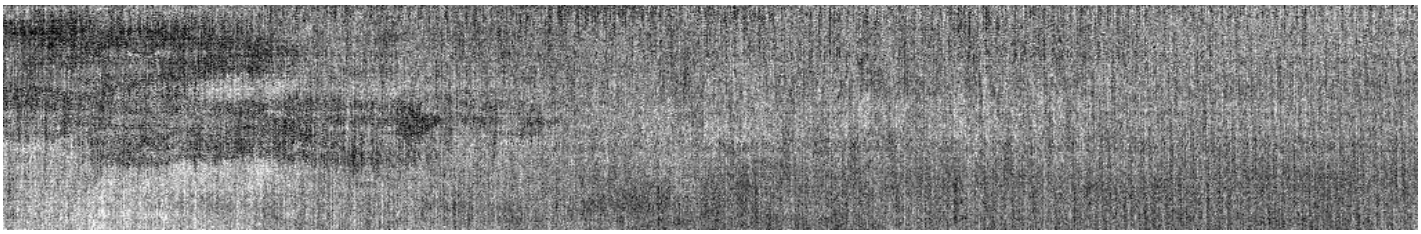
(b) MAF3



(c) MAF5

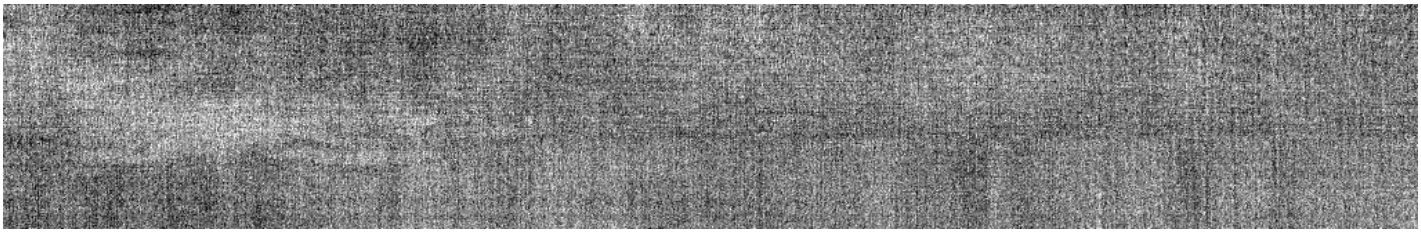


(d) MAF7



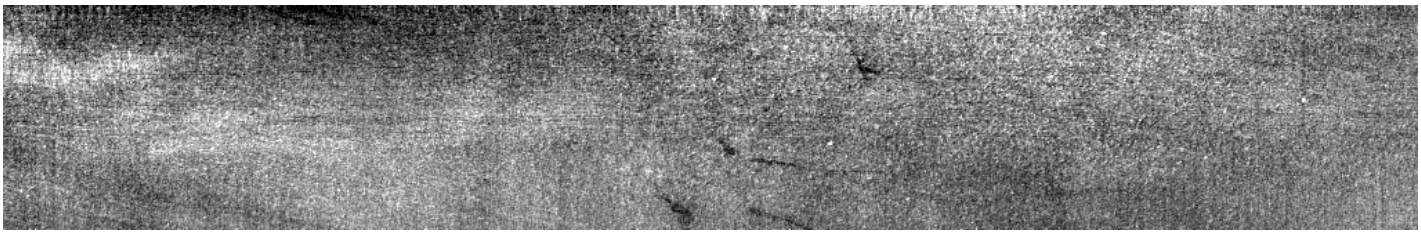
(e) MAF9

Figure 3: Here the first five odd-numbered MAFs are shown. We see that the content of spatially autocorrelated signal is decreasing.



(a) CH1

(b) CH7



(c) CH47

Figure 4: Three restored channels.