

ADAPTIVE RECONSTRUCTION METHOD OF MULTISPECTRAL IMAGES

Michal Haindl
Institute of Information Theory and Automation
Czech Academy of Sciences
Czech Republic
haindl@utia.cas.cz

Stanislava Šimberová
Astronomical Institute
Czech Academy of Sciences
Czech Republic
ssimbero@asu.cas.cz
Commision III, Working Group 2

KEY WORDS: Image, Reconstruction, Multispectral, Theory, Application

ABSTRACT

A new adaptive method is introduced to reconstruct missing or corrupted lines in multi-spectral image data. The reconstruction uses available information from the failed pixel surrounding due to spectral and spatial correlation of multi-spectral data. Missing lines are assumed to be modelled with a multi-dimensional regression model but this model cannot be identified, so a special approximation is introduced. The reconstruction is based on two mutually competing adaptive approximations of the regression model from which the locally optimal predictor is selected. A directional forgetting concept is introduced to support parameter adaptation.

1 INTRODUCTION

There are several ways of reconstructing corrupted or missing image data. The simplest method is to replace the missing detector scan line by the scan line of the detector immediately above or below it (we will refer to this method further as A). This scheme can cause [Bernstein, 1984] very observable distortions in the final image products, especially images of high contrast features. As a variant of mentioned method it has been suggested to linearly interpolate between the lines above and below the corrupted detector line - method B, or between six neighbouring pixels - method C. This does not solve the problem. Even interpolation with higher order curves, such as quadratic fit, is of no help see [Bernstein, 1984]. Three more sophisticated template-like methods were suggested in [Bernstein, 1984]: Template Replacement - D, Template Replacement with Error Adjustment - E and Quadratic Vertical Fit with Template Data - F. The Template Replacement method directly substitutes a corrupted detector line with a detector line from a similar (well correlated) band, after scaling its output intensity so that its range is similar to the other lines of the failed line spectral band. The coefficients of the quadratic are determined by a least squares fit to the actual data in a five - pixel vertical slice centered around each bad detector pixel. The value used for the bad (center) pixel in the slice is calculated as in a D algorithm. Test results in [Bernstein, 1984] show in low contrast regions slightly off colour stripe after applications of algorithms D and E. The problem of algorithm F is that it produces a lower contrast value than expected in light contrast areas.

These template-like methods cannot be used for reconstruction of multi-spectral pixels with several spectral components missing, while the A,B,C methods can be used also in these cases.

We have proposed the regression method [Haindl, 1992], which clearly outperforms the above - mentioned recon-

struction methods. The regression method was improved in [Haindl, 1996] to select a locally optimal predictor from two mutually competing symmetrical adaptive predictors for each pixel to be reconstructed - G. In this paper we present further improvements of our reconstruction method. The regression model is generalised to reconstruct a multi-spectral line with all spectral components missing - H. Finally a modification of the method based on directional forgetting idea - method I, which improves parameter estimation is presented.

Note that all the above mentioned methods, as well as our method, do not use any data from bad pixels, i.e. there is no difference between reconstruction of corrupted or missing data using these methods.

The present paper is organized as follows. In Section 2, a proposed method general concept under a Bayesian framework is introduced. Section 3 completes the algorithm with a locally optimal model selection rule design. Section 4 deals with a multi-spectral line reconstruction and Section 5 introduces the concept of directional forgetting. Section 6 discuss numerical realization problems while Section 7 contains an application to radio-spectrograph observations of the solar radio emissions (mono-spectral case) and remote sensing imagery data.

2 MONO-SPECTRAL LINE REGRESSION MODEL

Our method uses high spectral bands correlation and spatial correlation between neighbours of unusable pixels. We assume the mono-spectral line to be modelled as:

$$Y_t = \sum_{i \in I_t} a_i Y_{t-i} + E_t \quad (1)$$

with a multi-index $t = (m, n, d)$; Y_t is a reconstructed mono-spectral pixel value, m is the row number, n the column number, $d(d \geq 1)$ denotes the number of spectral bands and also the spectral band with line to be reconstructed

(the arrangement of spectral bands can be chosen at will), a_i are unknown model parameters, E_t is the white noise component. I_t is some neighbour index shift set excluding unknown data $(0, j, 0), (m^a - m^f, j, 0) \notin I_t$, where m^a is an approximation line (see (18)) and m^f is a reconstructed line, respectively.

Note that although the model reconstructs a mono-spectral corrupted line, the model can use information from all other spectral bands of an image ($d \geq 1$) as well. For mono-spectral images (e.g. radiospectrograph data) $d = 1$.

Let us denote another multi-index $t = (m, n, d)$ and choose a direction of movement on the image plane to track the bad line $t-1 = (m, n-1, d), t-2 = (m, n-2, d), \dots$. E_t is the white noise component with zero mean and constant but unknown dispersion Ω . We assume that the probability density of E_t has a normal distribution independent of previous data and is the same for every time t . Let us formally assume the knowledge of the bad data, then the task consists in finding the conditional prediction density $p(Y_t|Y^{(t-1)})$ given the known process history (2) and taking its conditional mean estimation \tilde{Y} for the reconstructed data.

$$Y^{(t-1)} = \{Y_{t-1}, Y_{t-2}, \dots, Y_1, Z_t, Z_{t-1}, \dots, Z_1\} \quad (2)$$

where Z is defined by (7). We have chosen the conditional mean estimator for data reconstruction, because of its optimal properties [Broemeling, 1985]:

$$\tilde{Y}_t = E[Y_t|Y^{(t-1)}] \quad (3)$$

Let us rewrite the regressive model (1) into a matrix form:

$$Y_t = P^T Z_t + E_t \quad (4)$$

where

$$P^T = [a_1, \dots, a_\beta] \quad (5)$$

is the $1 \times \beta$ unknown parameter vector.

$$\beta = \text{card} I_t \quad (6)$$

We denote the $\beta \times 1$ data vector

$$Z_t = [Y_{t-i} : \forall i \in I_t]^T \quad (7)$$

Data arrangement in (7) corresponds to the arrangement of parameters in (5).

Assuming normality of the white noise component E_t , conditional independence between pixels and an a priori probability density for the unknown model parameters chosen in the form (this normal form of a priori probability results in analytically manageable form of a posteriori probability density)

$$p(P, \Omega^{-1} | Y^{(0)}) = (2\pi)^{-\frac{\gamma(0)}{2}} |\Omega|^{-\frac{\gamma(0)}{2}} \exp \left\{ -\frac{1}{2} \text{tr} \left\{ \Omega^{-1} \begin{pmatrix} -I \\ P \end{pmatrix}^T V_0 \begin{pmatrix} -I \\ P \end{pmatrix} \right\} \right\}, \quad (8)$$

where V_0 is a positive definite $(\beta+1) \times (\beta+1)$ matrix and

$$\gamma(0) > \beta - 2, \quad (9)$$

we have shown [Haindl, 1992] that the conditional mean value is:

$$\tilde{Y}_t = \hat{P}_{t-1}^T Z_t \quad (10)$$

The following notation is used in (8) and (10) :

$$\hat{P}_{t-1} = V_{z(t-1)}^{-1} V_{zy(t-1)}, \quad (11)$$

$$V_{t-1} = \tilde{V}_{t-1} + V_0, \quad (12)$$

$$\tilde{V}_{t-1} = \begin{pmatrix} \tilde{V}_{y(t-1)} & \tilde{V}_{zy(t-1)}^T \\ \tilde{V}_{zy(t-1)} & \tilde{V}_{z(t-1)}^T \end{pmatrix}, \quad (13)$$

$$\tilde{V}_{y(t-1)} = \sum_{k=1}^{t-1} Y_k Y_k^T, \quad (14)$$

$$\tilde{V}_{zy(t-1)} = \sum_{k=1}^{t-1} Z_k Y_k^T, \quad (15)$$

$$\tilde{V}_{z(t-1)} = \sum_{k=1}^{t-1} Z_k Z_k^T. \quad (16)$$

It is easy to check [Haindl, 1992] also the validity of recursive (17).

$$\hat{P}_t = \hat{P}_{t-1} + (1 + Z_t^T V_{z(t-1)}^{-1} Z_t)^{-1} V_{z(t-1)}^{-1} Z_t (Y_t - \hat{P}_{t-1}^T Z_t)^T \quad (17)$$

To evaluate predictor (10) we need to compute the parameter estimator (11) or (17), but we do not know the past necessary data Y_t , because they are those to be reconstructed. On the other hand the data from Z_t in (10) are known: we can select a contextual support of the model in such a way to exclude unknown data. This problem is solved using the approximation based on spatial correlation between close lines

$$\tilde{Y}_t \doteq \tilde{P}_{t-1}^T Z_t, \quad (18)$$

where \tilde{P}_{t-1} is the corresponding parameter estimator (11), (17) for the nearest known line (including known contextual neighbours (7)) to our reconstructed one in the spectral band d . Note the different Z (7) in (18) and (15), (16), (17). This approximation assumes similar directional correlations on both lines, but not necessarily a mutual correlation of these lines themselves.

3 OPTIMAL MODEL SELECTION

Let us assume two regression models (4) M_1 and M_2 with the same number of unknown parameters ($\beta_1 = \beta_2 = \beta$) and mutually symmetrical neighbour index shift sets $I_{1,i}, I_{2,i}$ with the missing line being their symmetry axis. According to the Bayesian theory, the optimal decision rule for minimizing the average probability of decision error chooses the maximum a posteriori probability model, i.e. a model whose conditional probability given the past data is the highest one. The presented algorithm can be therefore completed as in (19):

$$\tilde{Y}_t \doteq \begin{cases} \tilde{P}_{1,t-1}^T Z_{1,t} & \text{if } p(M_1|Y^{(t-1)}) > p(M_2|Y^{(t-1)}) \\ \tilde{P}_{2,t-1}^T Z_{2,t} & \text{otherwise} \end{cases} \quad (19)$$

where $Z_{i,t}$ are data vectors corresponding to $I_{i,t}$. Following the Bayesian framework used in our paper and choosing uniform a priori model in the absence of contrary information, $p(M_i|Y^{(t-1)}) \sim p(Y^{(t-1)}|M_i)$, the simultaneous conditional probability density can be evaluated from

$$p(Y^{(t-1)}|M_i) = \int \int p(Y^{(t-1)}|P, \Omega^{-1}) p(P, \Omega^{-1}|M_i) dP d\Omega^{-1} \quad (20)$$

Under the already assumed conditional pixel independence, the analytical solution has the form

$$p(M_i|Y^{(t-1)}) = k |V_{i,z(t-1)}|^{-\frac{1}{2}} \lambda_{i,t-1}^{-\frac{\gamma(t-1)-\beta+2}{2}}, \quad (21)$$

where k is a common constant. To evaluate $p(M_i|Y^{(t-1)})$, we have to use a similar approximation (18) as for the predictor (10). All statistics related to a model M_1 (15), (16), (17), (21) are computed from data on one side of the reconstructed line while symmetrical statistics of the model M_2 are computed from the opposite side.

The solution of (21) uses the following notations:

$$\gamma(t-1) = \gamma(0) + t - 1, \quad (22)$$

$$\lambda_{t-1} = V_{y(t-1)} - V_{zy(t-1)}^T V_{z(t-1)}^{-1} V_{zy(t-1)}. \quad (23)$$

The determinant $|V_{z(t)}|$ as well as λ_t can be evaluated recursively see [Haindl, 1992]:

$$|V_{z(t)}| = |V_{z(t-1)}| (1 + Z_t^T V_{z(t-1)}^{-1} Z_t), \quad (24)$$

$$\lambda_t = \lambda_{t-1} + (Y_t - \hat{P}_{t-1}^T Z_t) (Y_t - \hat{P}_{t-1}^T Z_t)^T (1 + Z_t^T V_{z(t-1)}^{-1} Z_t)^{-1}. \quad (25)$$

In the case when some data necessary for the approximation are missing the corresponding model probability is set to zero

$$p(M_i|Y^{(t-1)}) = 0. \quad (26)$$

If the reconstructed line is located in a boundary image area then the reconstruction algorithm uses only one model (one of the model probabilities is permanently zero).

4 MULTI-SPECTRAL LINE REGRESSION MODEL

Let us assume that all spectral components of the multi-spectral line are missing. Such a multi-spectral line can be modelled using a multi-dimensional regression model:

$$Y_i = \sum_{i \in I_t} A_i Y_{i-1} + E_i \quad (27)$$

where $i = (m, n)$; Y_i is $\nu \times 1$ reconstructed multi-spectral pixel value, A_i are unknown $\nu \times \nu$ model parameters matrices, E_i is the white noise vector.

Parameter vector P^T (5) become now the $\nu \times \beta\nu$ matrix:

$$P^T = [A_1, \dots, A_{\beta}] \quad (28)$$

and

$$\beta^* = \nu \text{card} I_t = \nu\beta \quad (29)$$

Z_t (7) is the β^* data vector, V_0 become a positive definite $(\beta^* + \nu) \times (\beta^* + \nu)$ matrix and $\gamma(0) > \beta^* - 2$. Equations (7)-(26) remain unchanged and can be used for the multi-dimensional prediction and optimal multi-dimensional model selection (19),(21) as well.

If $A_i = \text{diag}[a_{1,i}, \dots, a_{\nu,i}] \forall i$ then the multi-dimensional model reconstruction is identical with separately applied single-dimensional model reconstruction G on every spectral line component.

5 DIRECTIONAL FORGETTING

The reconstruction model in sections 2 and 3 was developed under the assumption that model parameters are strictly location-invariant. This assumption is not realistic for most real image reconstruction problems. The Bayesian solution of the case of location-variant parameters is given by equation

$$p(P_{t+1}, \Omega_{t+1}^{-1}|Y^{(t)}) = \int \int p(P_{t+1}, \Omega_{t+1}^{-1}|P_t, \Omega_t^{-1}, Y^{(t)}) p(P_t, \Omega_t^{-1}|Y^{(t)}) dP_t d\Omega_t^{-1}. \quad (30)$$

Unfortunately the required conditional distribution for this recursion $p(P_{t+1}, \Omega_{t+1}^{-1}|P_t, \Omega_t^{-1}, Y^{(t)})$ is seldom known. Usual solution of this problem is the constant exponential forgetting. It increases the uncertainty of the old parameter estimate by a constant factor equal for all data. This results in modification of equations (17),(24),(25) see [Haindl, 1996] for details.

The exponential forgetting permanently loses old information even if there is lack of a new one (parameters remain unchanged in some directions). Some ideas to overcome this insufficiency were suggested in [Hägglund, 1983], [Kulhavý, 1993]. We propose another solution based on directional forgetting coefficients to control individually every data item $[Y_{t-i} : i \in I_t]^T$ forgetting depending on the corresponding directional derivation change, i.e.

$$\alpha_i = \frac{\min\{|\frac{\delta Y_t}{\delta i}|, |\frac{\delta Y_{t-1}}{\delta i}|\}}{\max\{|\frac{\delta Y_t}{\delta i}|, |\frac{\delta Y_{t-1}}{\delta i}|\}}. \quad (31)$$

If there is no change in the direction i then $\alpha_i = 1$ (no data forgetting), otherwise $\alpha_i < 1$ and the increase of old information uncertainty is proportional to the directional change of the derivation.

Let us denote the matrix of directional forgetting parameters

$$\alpha = \begin{pmatrix} \alpha_y & 0 \\ 0 & \alpha_z \end{pmatrix} = \begin{pmatrix} \alpha_1 & \dots & 0 \\ & \ddots & \\ 0 & \dots & \alpha_{\beta^* + \nu} \end{pmatrix}, \quad (32)$$

where

$$\alpha_y = \text{diag}[\alpha_1, \dots, \alpha_\nu]$$

and

$$\alpha_z = \text{diag}[\alpha_{\nu+1}, \dots, \alpha_{\nu+\beta^*}].$$

The model equations (17),(24),(25) become (33)-(35), respectively.

$$\hat{P}_t = \alpha_z^{-1} [\hat{P}_{t-1} \alpha_y + (1 + \tilde{Z}_t^T V_{z(t-1)}^{-1} \tilde{Z}_t)^{-1} V_{z(t-1)}^{-1} \tilde{Z}_t (Y_t - \alpha_y \hat{P}_{t-1}^T \tilde{Z}_t)^T] \quad (33)$$

$$|V_{z(t)}| = \left(\sum_{j=\nu}^{\nu+\beta^*} \alpha_j \right)^2 |V_{z(t-1)}| (1 + \tilde{Z}_t^T V_{z(t-1)}^{-1} \tilde{Z}_t) \quad (34)$$

$$\lambda_t = \alpha_y \lambda_{t-1} \alpha_y + (Y_t - \alpha_y \hat{P}_{t-1}^T \tilde{Z}_t) (Y_t - \alpha_y \hat{P}_{t-1}^T \tilde{Z}_t)^T (1 + \tilde{Z}_t^T V_{z(t-1)}^{-1} \tilde{Z}_t)^{-1} \quad (35)$$

$$\tilde{Z}_t = \alpha_z^{-1} Z_t \quad (36)$$

The filtering matrices α_y, α_z are diagonal so the increase of computational complexity for the algorithm I is very moderate.

6 NUMERICAL REALIZATION

The predictors in (19) can be evaluated using updating of matrices $V_{i,t}$ (12) and their following inversion. Another possibility is the direct updating of $\hat{P}_{i,t}$ (17), (33). To ensure the numerical stability of the solution, it is advantageous to calculate $\hat{P}_{i,t}$ (17), (33) using a square-root filter, which guarantees the positivity of matrix (12). The filter updates directly the Cholesky square root of matrices $V_{i,t}^{-1}$.

Alternatively it is possible to use the UDU filter (a factorization into two triangular and one diagonal matrices) for this purpose. Note that the same square-root filters can be used also for the updating of statistics of the directional forgetting algorithm version. They only difference is in input filter values. Initialization of recursive (17), (24) and (25) must keep the condition of positive definiteness of matrices $V_{i,0}$ (8). We implemented in our algorithm the uniform a priori start :

$$V_{i,0} = I . \quad (37)$$

This solution not only conforms with the initial lack of information at the start of algorithm, but also simplifies the calculation of the integral (20). Another possibility could be for example a local condition start, which ensures a quicker adaptation .

7 RESULTS

In this section we present simulation results of the proposed reconstruction method and compare them with methods briefly surveyed in the introductory section. The performance of the methods is compared on artificially created bad lines (removed from the unspoiled parts of the images so that the original data are known) using the criterion of mean absolute difference between original and replaced pixel values

$$MAD = \frac{1}{n\nu} \sum_{j=1}^{\nu} \sum_{i=1}^n |Y_{i(j)} - \tilde{Y}_{i(j)}| , \quad (38)$$

where $\nu = 1$ for the mono-spectral model.

method	monospectral		multispectral	
	radar MAD	SPOT MAD	TM MAD	SPOT MAD
A	701	2.8	24	1.25
B	412	1.3	20	0.7
C	394	1.8	19	0.9
D	-	-	23	2.6
E	-	-	21	1.1
F	-	-	19	1.5
G	135	1.23	16	0.7
I	132	1.1	14	0.5
I	130	0.72	10.1	0.34

Table 1: Single spectral band reconstruction.

Pixels corresponding to the $I_{i,t}$ are denoted * and the reconstructed pixel o , respectively.

The first example is the defective radiospectrograph image shown in Fig.1 from the Ondřejov Observatory 1000 - 2000 MHz radiospectrograph observation of the solar radio emission. The frequency band is divided into 256 channels (the frequency resolution of about 4 MHz) and the grey level range of pixels is 0-2800.

The optimal reconstruction models M_i for the radiospectrograph were found to be:

$$M_1 \begin{matrix} & * & & * & * & & & & * & * & * & & * \\ * & * & & o & * & * & & & M_2 & * & * & o & * & * \\ & * & & * & * & & & & & * & * & & * & \end{matrix}$$

The second tested image Fig.2 was the agricultural type of the Thematic Mapper seven spectral band sub-scene from North Moravia. The failed line was located in the TM1 band and the selected models are:

$$M_1 \begin{matrix} & & * & & * & & & & & * & & * & & & \\ * & * & o & * & & & & & M_2 & * & o & * & * & & \\ & & * & & * & & & & & * & & * & & & \end{matrix}$$

The last two examples are SPOT multi-spectral image Fig.3 (agricultural scene from Moravia, failed line located in the green visible band)

$$M_1 \begin{matrix} * & & & & & & & & & & & & & & * \\ & * & o & * & * & & & & M_2 & * & * & o & * & & \end{matrix}$$

and SPOT panchromatic image (agricultural scene from vicinity of La Rochelle, both models are * * o * * *).

Table 1 contains monospectral line reconstruction results. Method I (regression method with the directional forgetting) demonstrates improvement in comparison with the regression method using a constant exponential "forgetting factor" $\alpha = 0.99$. These results show the superiority of our method over the classical ones. The last table row demonstrates isolate pixels reconstruction. In this case there is no need for approximation, because the predictor is used with complete knowledge of all past data.

The radar example demonstrates properly found better estimation data side (approximation line for model parameters estimation) in the case of one superior side, on the

method	multispectral	
	TM MAD	SPOT MAD
A	53	2.75
B	49	1.5
C	48	1.9
G	37	0.9
H	33	0.7
H	26	0.4

Table 2: Multispectral line reconstruction.

remaining examples the optimal side is oscillating so single model method cannot reach performance of the double-model method G even if repeated on both corresponding sides.

Table 2 shows multi-spectral line reconstruction results. The regression model is again superior over the classical methods applied separately on every missing spectral component line. The multi-dimensional model has ν^2 times more parameters than the single-dimensional model and so it is more sensitive to overparametrization resulting in degraded model performance. Equation (20) is again used for the optimal model structure selection, but in this case $\beta_1^* \neq \beta_2^*$ so $k_1 \neq k_2$ in (21).

8 CONCLUSION

The results of our test are encouraging. The proposed methods were always the best ones in all our experiments. The advantage of the regression-type method increases with an increasing number of correlated spectral bands but even on monospectral images (the radiospectrograph and SPOT panchromatic examples) it is also the best one.

Applying the presented reconstruction method in the radiospectrograph image reconstruction problem, we obtained images without missing lines or pixels. These reconstructed radiospectrograms were successfully used for the evaluation of the observation of fast drift bursts during the solar activity. While the former practice was to discard such unusable radiospectrograms with the unfortunate consequence of disruption of observation series.

We have not seen any other scheme for correcting image defects (off colour stripe, colour gaps, horseshoe effect) applying the regression method. The advantage of the present method I over our previously published regression method G [Haindl, 1992] with constant exponential forgetting lies in more precise parameter estimation and consequently the model prediction - reconstruction quality improvement.

Our method can also be used if more than one of the monospectral line components is missing. In this case the reconstruction is done with the multi-dimensional model (H) with dimensionality ν equal to the number of missing spectral line components. Alternatively, the single-dimensional model can be applied repeatedly to all missing monospectral lines. Single-dimensional model results are worse than the multi-dimensional model ones but the model structure optimization is easier. It is also possible to combine the multi-dimensional model with the directional forgetting concept.

The algorithm can be used to remove scratches as well if it is applied sequentially on linear parts of a scratch.

Finally if the method is used for isolated image pixels reconstruction then the predictor and similarly the model probability expression do not need any data approximation and the regression method performs better than for line reconstruction and much better than any of the classical methods.

The proposed method is fully adaptive, numerically robust and still with moderate computation complexity so it can be used in an on-line image acquisition system.

REFERENCES

- [Bernstein, 1984] Bernstein, R., Lotspiech, J.B., Myers, H.J., Kolsky, H.G., Lees, R.D., Analysis and Processing of Landsat-4 Sensor Data Using Advanced Image Processing Techniques and Technologies. IEEE Trans on Geosci., GE-22(3), pp. 192-221.
- [Broemeling, 1985] Broemeling, L.D. Bayesian Analysis of Linear Models. Dekker, New York.
- [Haindl, 1992] Haindl, M., Šimberová, S., A Multispectral Image Line Reconstruction Method. In: Theory & Applications of Image Analysis. P. Johansen, S. Olsen Eds., World Scientific Publishing Co. Singapore.
- [Haindl, 1996] Haindl, M., Šimberová, S., A high - resolution radiospectrograph image reconstruction method. Astronomy and Astrophysics, Suppl. Ser., 115(1), pp. 189-193.
- [Hägglund, 1983] Hägglund, T. The problem of forgetting old data in recursive estimation. Proc. IFAC Workshop on Adaptive Systems in Control and Signal Processing, San Francisco.
- [Kulhavý, 1993] Kulhavý, R. Zarrop, M.B. On a general concept of forgetting. Int. J. of Control, 58(4), pp. 905-924.

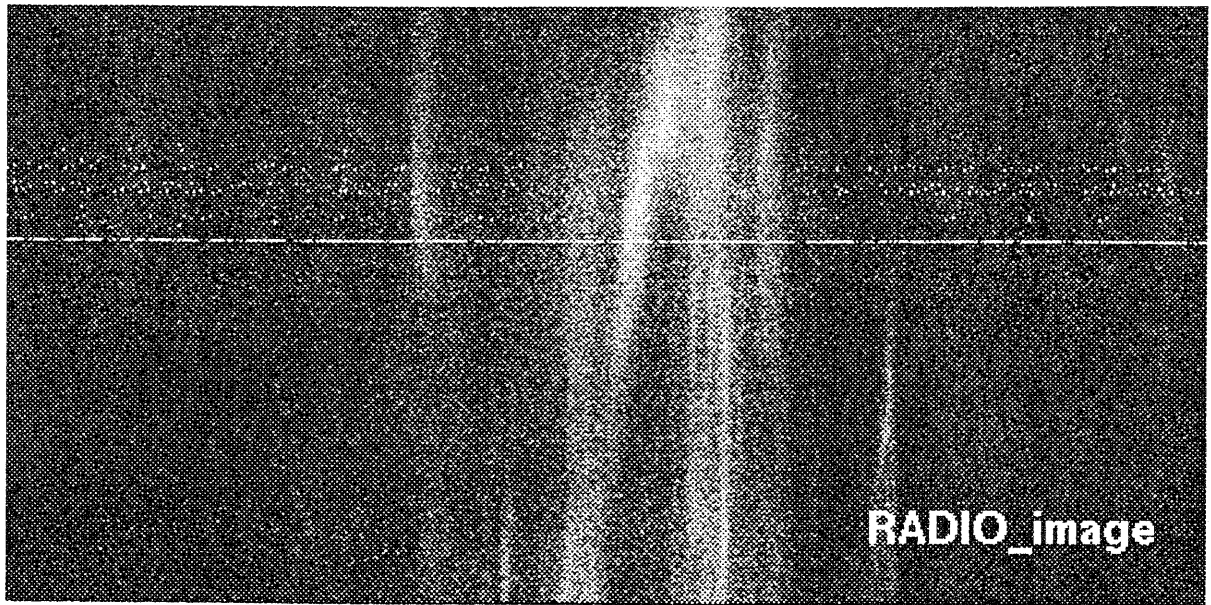


Fig.1 The original defective radiospectrograph image



Fig.2 The Thematic Mapper - subscene from North Moravia



Fig.3 The SPOT multispectral image