Supplementary material for manuscript:

Object detection with few training data: detection of subsiding troughs in SAR interferograms

Tab. 1. Test images

Image name	Resolution	Description
Q	11 215 x 3 984	Original name of binary file: *Phase_ifg_srd_IW2_VV_28Jan2016_04Mar2016.img* Relatively good quality (distinctive troughs, areas of low-frequency and high-amplitude noise separated from areas of troughs)
Q_{zoom1}	2 102 x 1 187	Sub-image of Q with area of occurrence of troughs
Q_{zoom2}	1 324 x 623	Sub-image of Q_{zoom1} with particular density of troughs and without low-frequency and high-amplitude noise
P	2 501 x 2 001	Example of bad quality image: areas of troughs are mixed with high-amplitude noise. Areas of noise are irregular and sometimes their shape is similar to the shape of troughs.
Ellipses	845 x 557	Artificial image with ellipses, circles and rectangular shapes, used for preliminary tests of the method based on convolution of image with circular wavelets.

1. Detection of troughs' centres without learning samples

The method is not based on automatic learning, therefore there is no need for a training set. We convolve the input image with a circular kernel derived from the Gabor impulse response. 2D Gabor kernel used for extraction of local frequency features along *x* axis (see Fig. 1a) is given by:

$$g(x,y)=\frac{1}{2\pi\sigma_x\sigma_y}\exp\left[-\frac{1}{2}\left(\frac{x^2}{\sigma_x{}^2}+\frac{y^2}{\sigma_y{}^2}\right)+2\pi jf_{centr}x\right], \tag{1}$$

and its 1D intersection along *x* axis is:

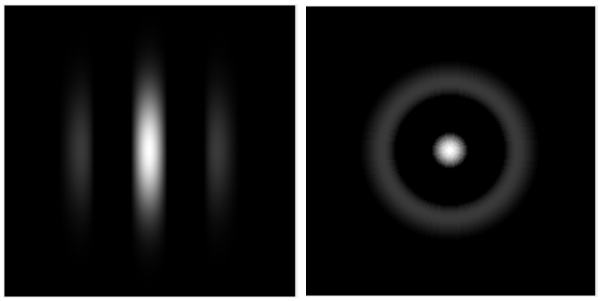
$$g(x)=\frac{1}{2\pi\sigma}\exp\left[-\frac{x^2}{2\sigma^2}+2\pi jf_{centr}x\right], \tag{2}$$

where f is the central frequency of the filter and σ is the standard deviation. Based on (2) we propose a circular kernel (see Fig. 1b):

$$g(r)=\frac{1}{2\pi\sigma}\exp\left[-\frac{r^2}{2\sigma^2}+2\pi jf_{centr}r\right], \tag{3}$$

where r is the distance from the centre (x_0, y_0) of the mask:

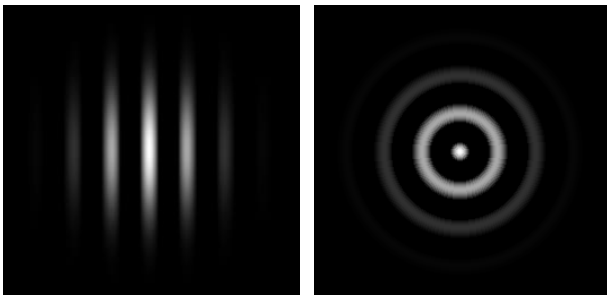
$$r=\sqrt{(x-x_0)^2+(y-y_0)^2}\,. \tag{4}$$



a b

Fig. 1 Real part of a Gabor impulse response (a) and the corresponding circular sine wave for default Spatial Frequency Bandwidth.

The idea is that convolution of the kernel (3) with the input image should yield magnitude peaks in centres of circular shapes.



a b

Fig. 2 Real part of a Gabor impulse response (a) and the corresponding circular sine wave for Spatial Frequency Bandwidth set to 0.5.

Centres of troughs are detected at points where the module of filter output exceed the threshold.

1.1 Tests on artificial image

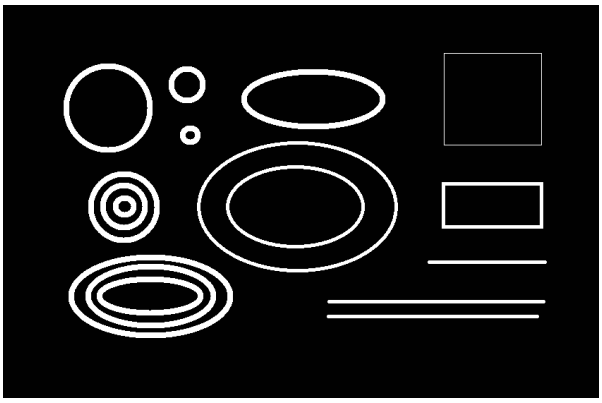
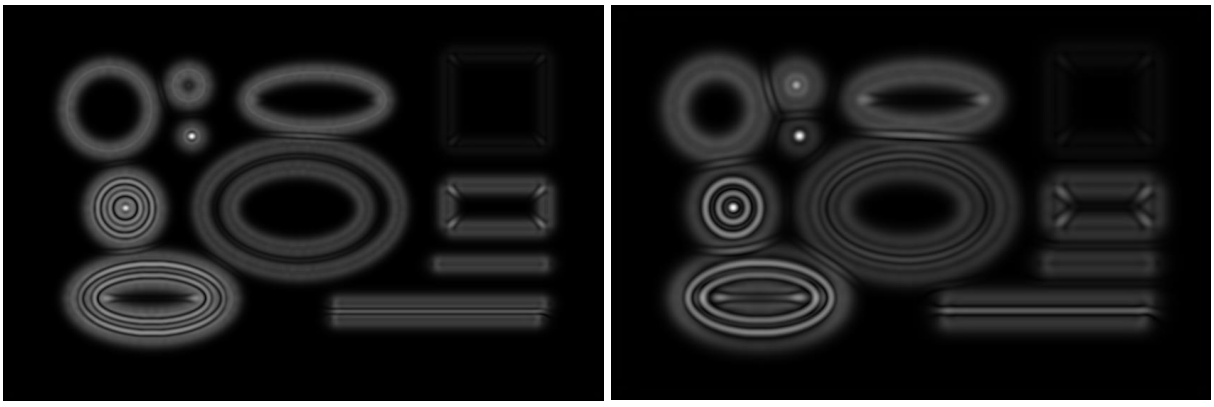
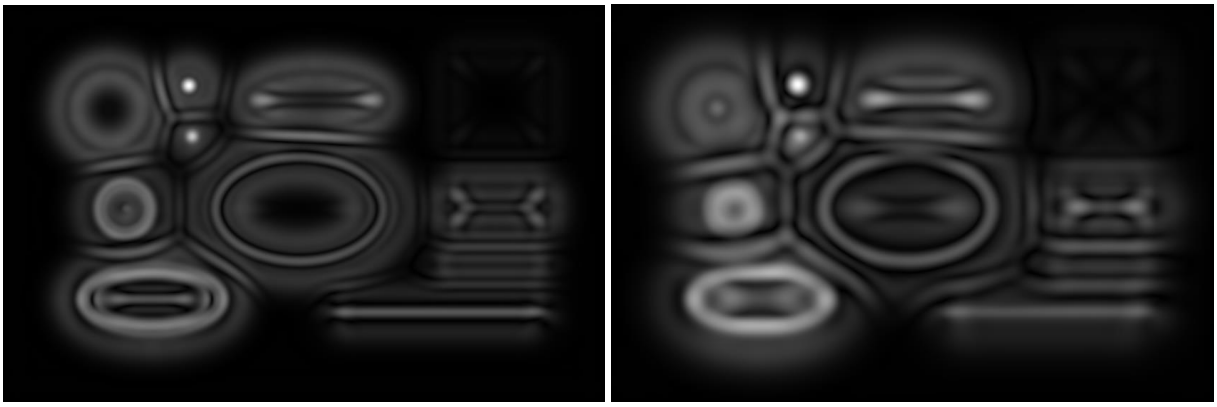


Fig. 3 Artificial test image.

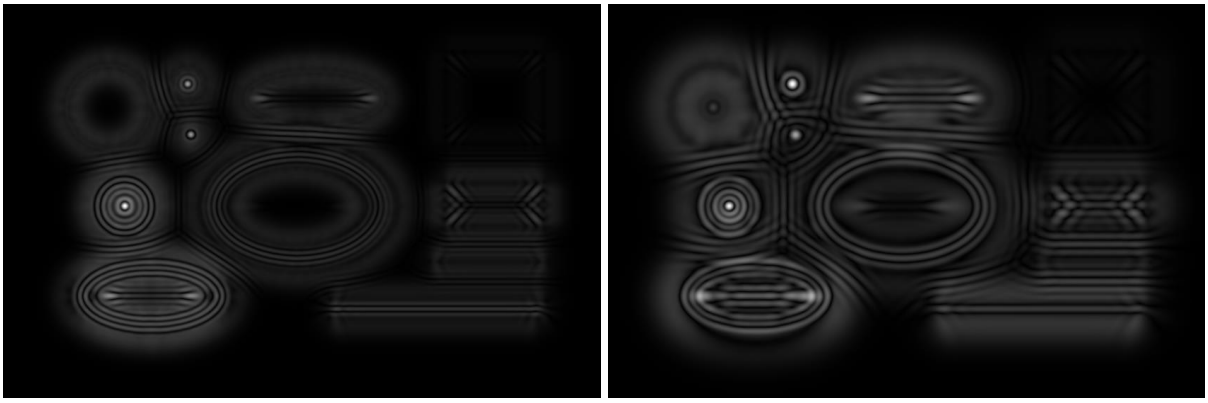


a b

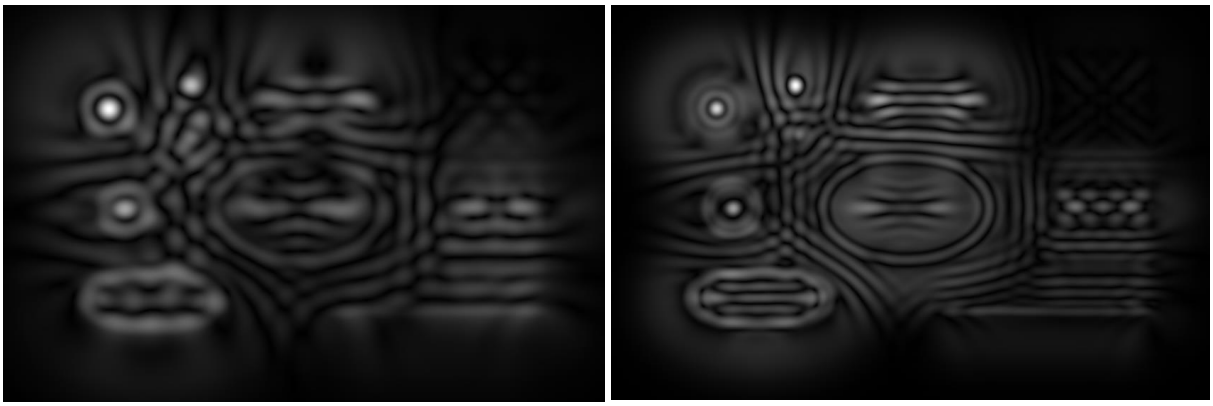


c d

Fig. 4 Magnitude of response for wavelengths: 15, 22.8, 34.2 and 51.2 pixels.



a b



c d

Fig. 5 Magnitude of response for wavelengths: 15, 22.8, 34.2 and 51.2 pixels for Spatial Frequency Bandwidth set to 0.5.

1.2 Tests on real data: image Q_{zoom2}

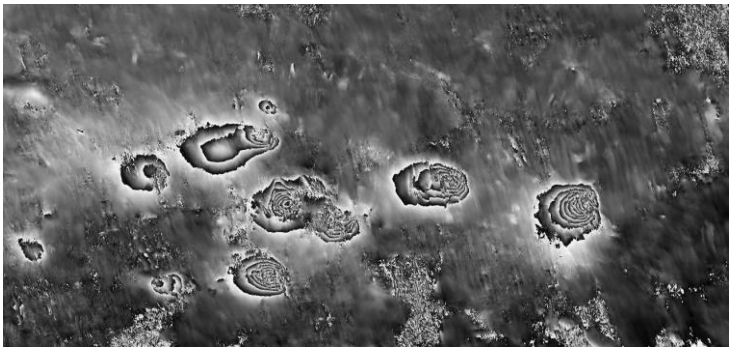
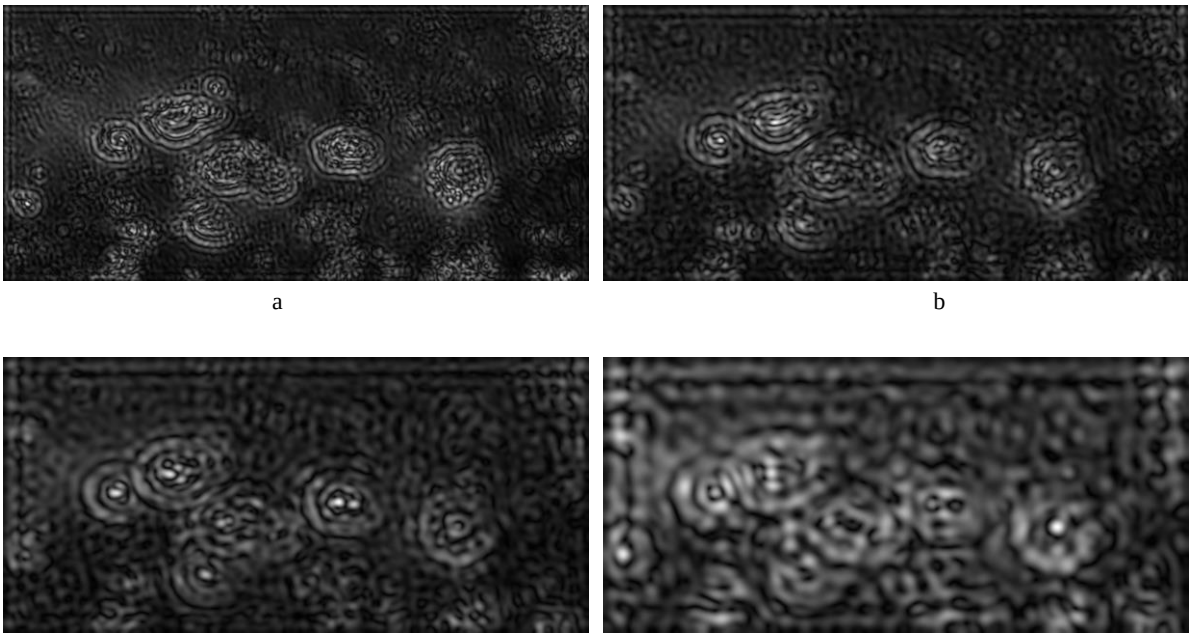


Fig. 6 Image Q_{zoom2}



a b

c d

Fig. 7 Examples of magnitude of response for wavelengths: 15, 22.8, 34.2 and 51.2 pixels for Spatial Frequency Bandwidth set to 0.5.

In Fig. 8 we present the result of trough detection in image Q_{zoom2}. Green points are spots where for any frequency the magnitude response is above the threshold. Wavelength vector was:
T=[10.00 11.49 13.20 15.16 17.41 20.00 22.97 26.39 30.31 34.82 40.00 45.95 52.78 60.63].
Parameters of Gabor filter bank used to generation of convolution filter were: SpatialAspectRatio=1, SpatialFrequencyBandwidth=0.5.
Threshold: [1 0.85 1 0.95 0.85 0.85 0.85 0.85 0.85 0.65 0.85 0.85 0.85 0.85];

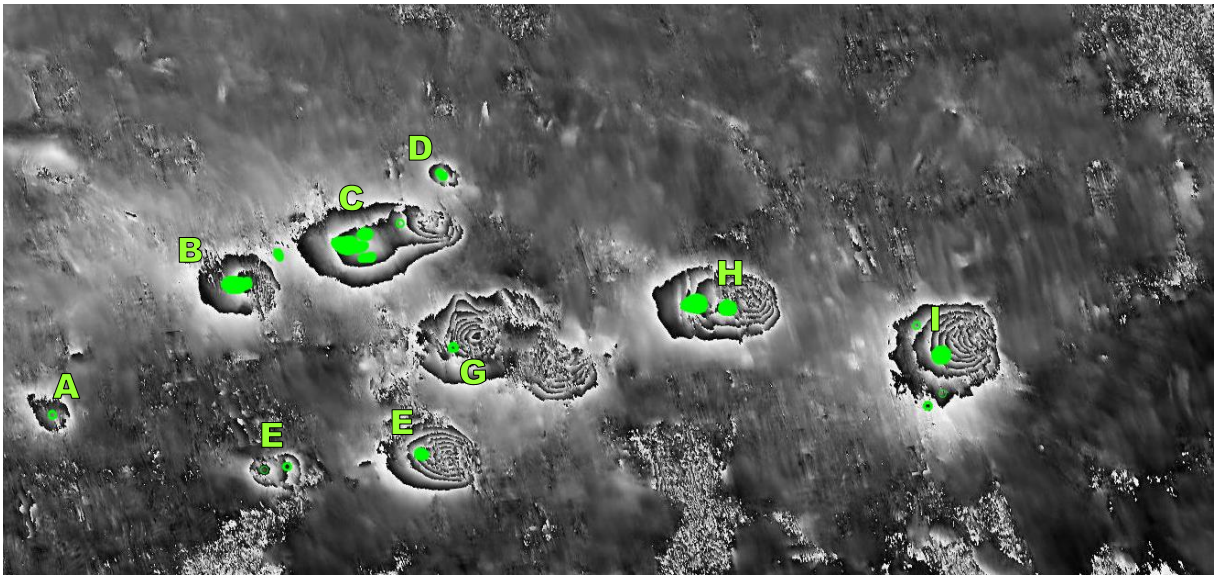


Fig. 8 The result of detection.

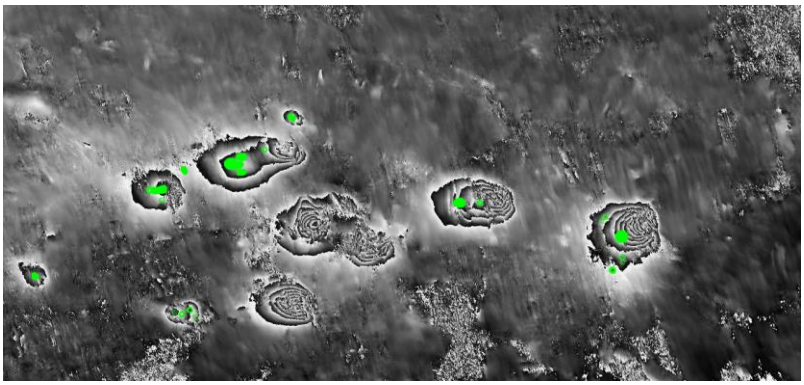
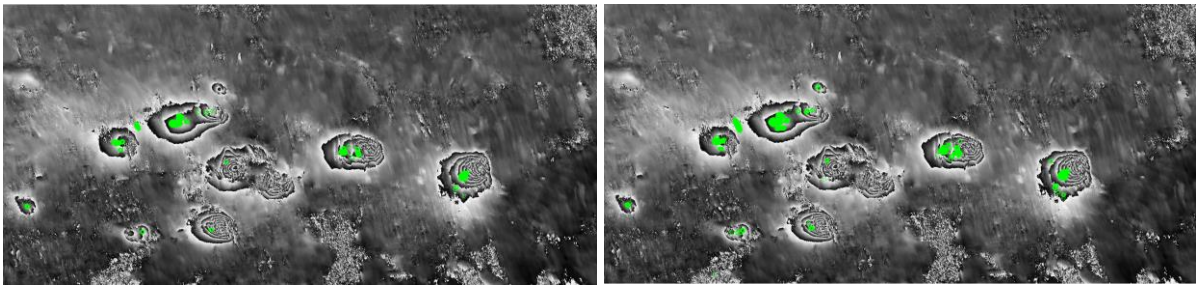


Fig. 9. The result when threshold was set to 0.85 maximum magnitude of response for all wavelengths.*

1.2.1 Tuning Spatial Frequency Bandwidth

The default value of Spatial Frequency Bandwidth equal 1 is definitely not appropriate for this method. Below we present results for Spatial Frequency Bandwidth = 0.3. Size of our circular kernel varies from 127x127 for wavelength T=10 to 767x767 for T=60.63. Real part of a circular sine wave is presented in Fig. 10.


Fig. 10. Real part of a circular sine wave for Spatial Frequency Bandwidth = 0.3



a b

Fig. 11. a) The result when threshold was set to 0.85 maximum magnitude of response for all wavelengths for SpatialFrequencyBandwidth=0.3 b) we were able to decrease the common threshold to 0.8 without new false detections*

1.3 Test on image Q_{zoom1}

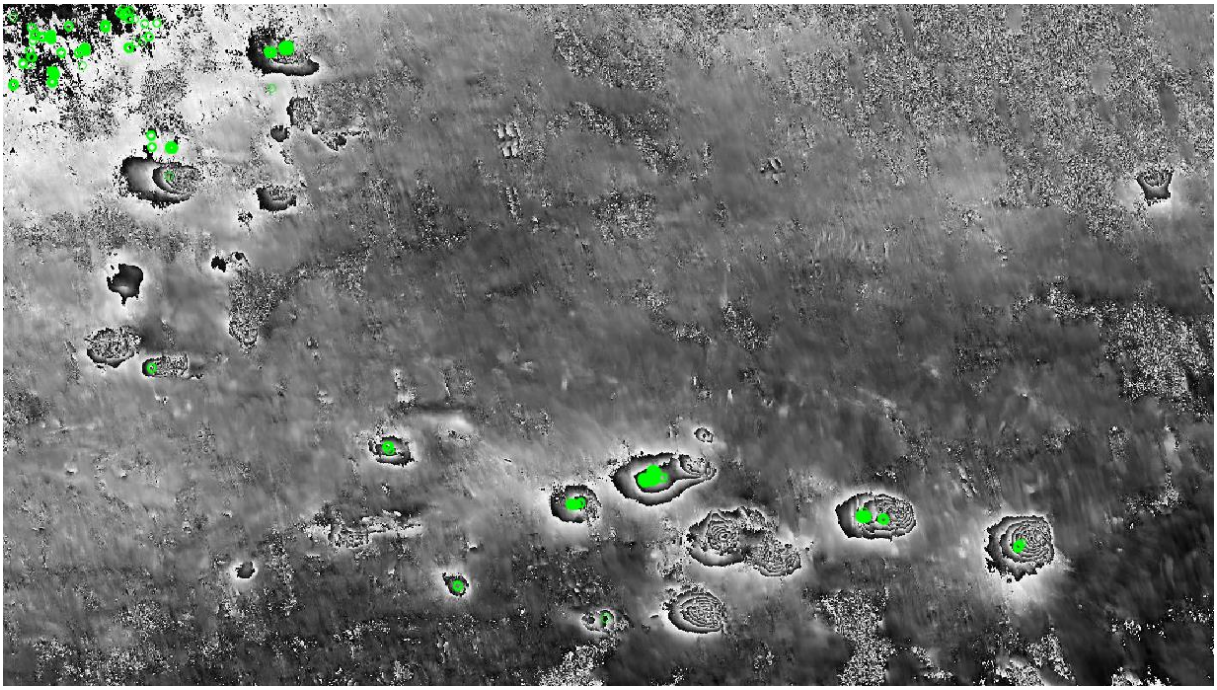


Fig. 12 The result when threshold was set to 0.85 maximum magnitude of response for all wavelengths for SpatialFrequencyBandwidth=0.3*

In order to get rid of false detections in the top left corner we excluded all wavelengths except 40 and 45.95. With such parameters we received only one false detection, caused by a nearby trough. However, using only two frequencies does not allow detection of most troughs (Fig. 13). For this image there is no set of parameters that could ensure correct detection without false alarms.

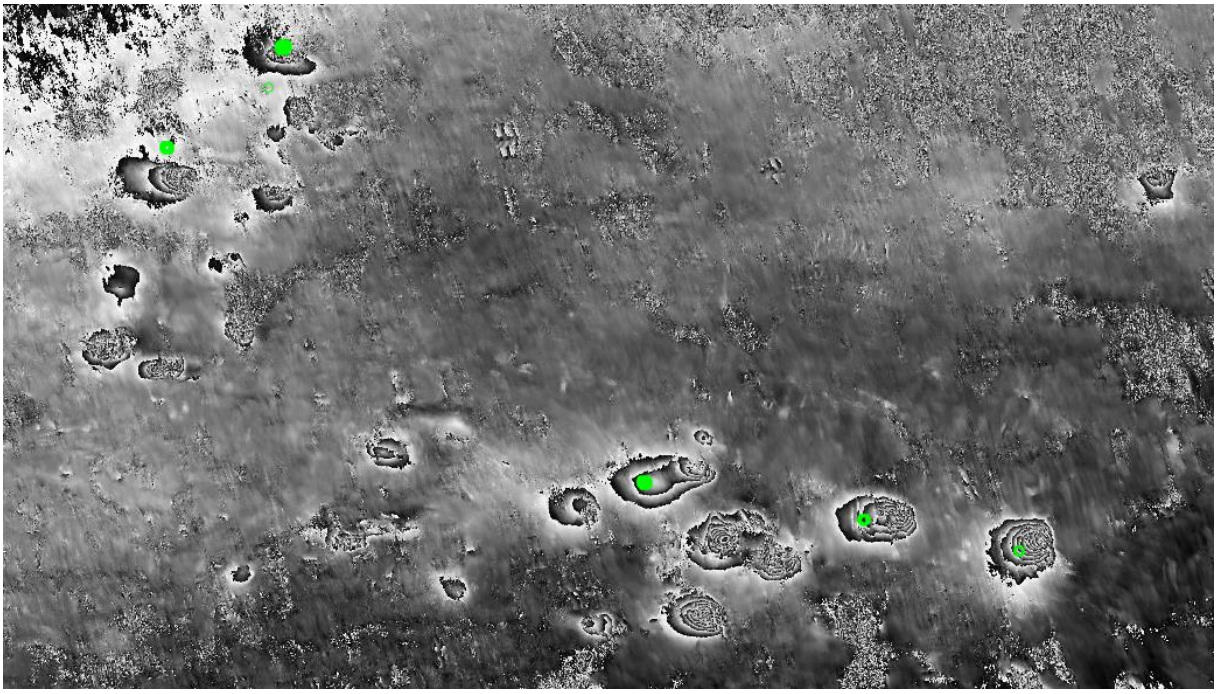
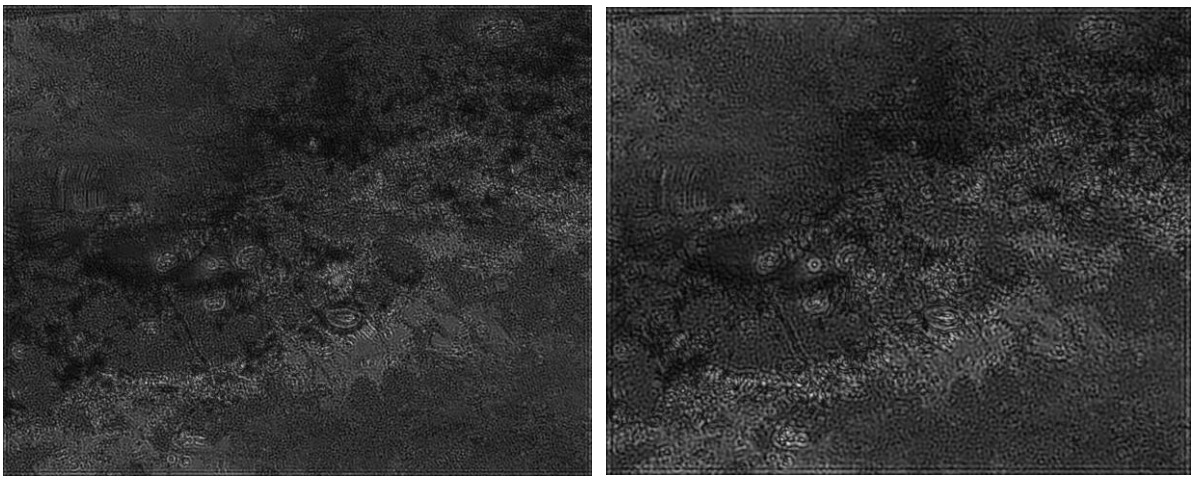


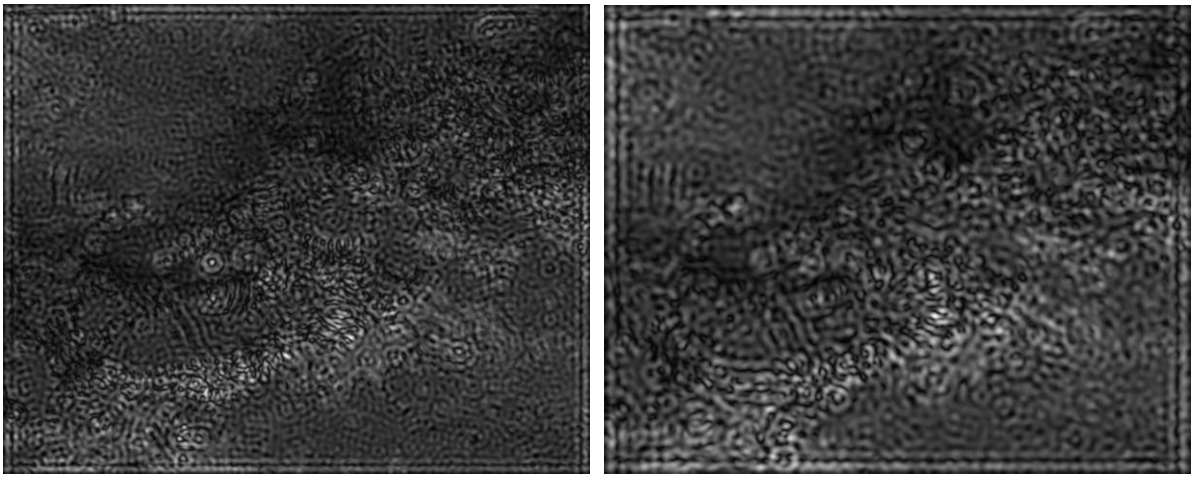
Fig. 13 Thresholds for wavelengths: 40 and 45.95 are equal 0.85, all other are set to 1.

1.4 Difficult case: image P

In this chapter we test the method on an image with a large amount of noise with diversified spectral features, where some troughs are difficult to distinguish from the background.



a b



c d

Fig. 14 Magnitude of response for wavelengths: 15, 22.8, 34.2 and 51.2 pixels for Spatial Frequency Bandwidth set to 0.5.

In Fig. 15 we present the results of detection for wavelength vector:
T=[10.00 11.49 13.20 15.16 17.41 20.00 22.97 26.39 30.31 34.82 40.00 45.95 52.78 60.63] threshold 0.8 for all wavelengths and Spatial Frequency Bandwidth set to 0.3.

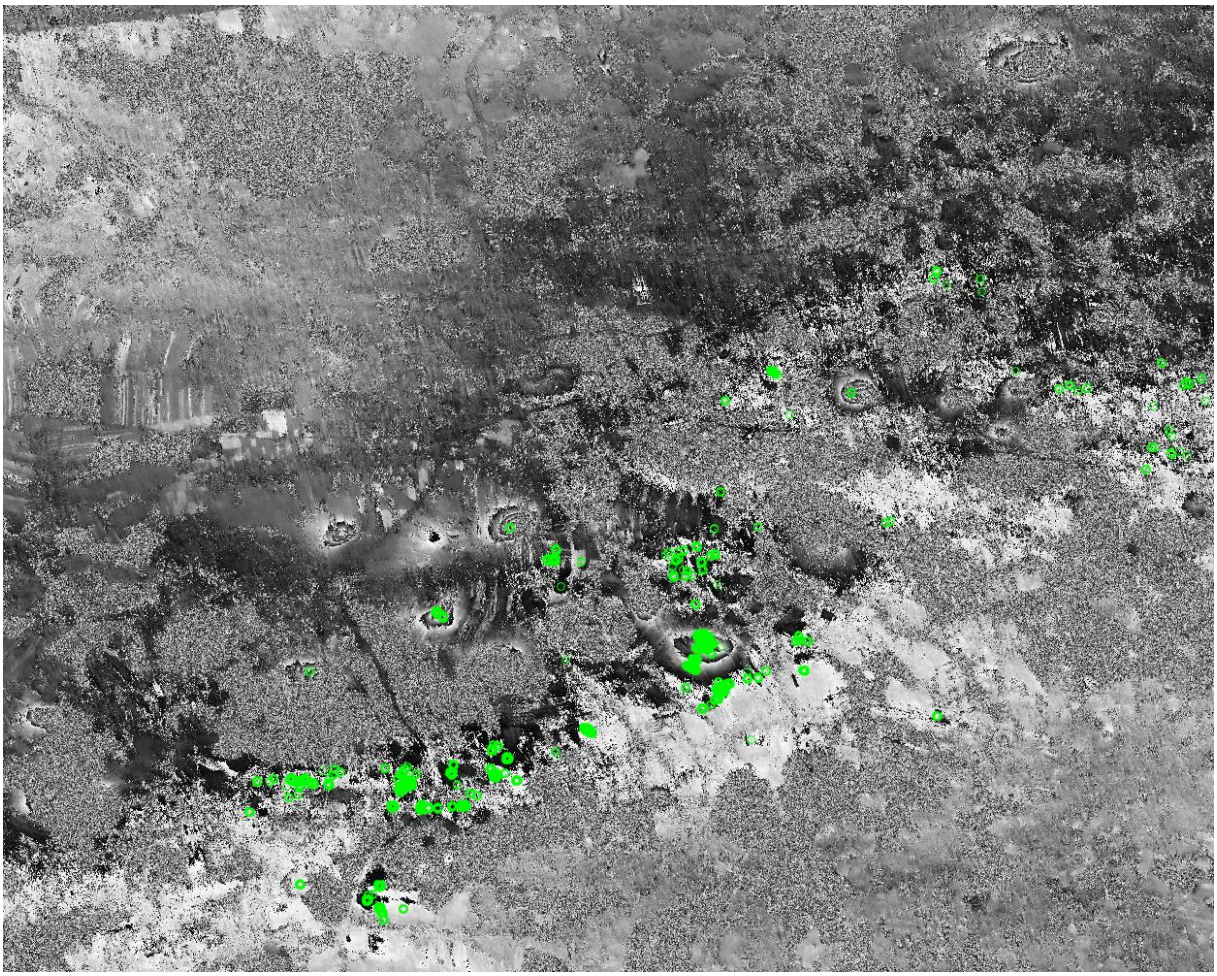


Fig. 15 The result of detection threshold 0.8 for all wavelengths and Spatial Frequency Bandwidth set to 0.3

Only a part of troughs were found and there are many false alarm. We conclude that applicability of this method depends on the quality of the input image.

2. Gabor features calculated directly from the image in Cartesian coordinates

2.1 Parameters of the Gabor filter bank

We used Gabor filter bank with wavelengths: $T_i = T_0 k^i$, where T_0 is the shortest wavelength and k=1.5. In input images the distance between consecutive extrema of troughs varies between 4 and 45 pixels, so we set the shortest wavelength 3 pixels (to ensure a certain margin) and the vector of wavelengths is:

```
T = [3 4.5 6.8 10 15 22.8 34.2 51.2].
```

We used four orientations:

```
FI = [0 45 90 135],
```

so there are 32 filters in the bank.

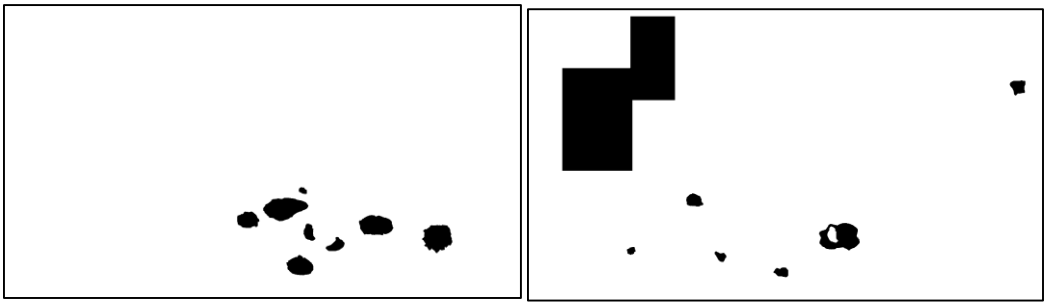
After calculating of Gabor outputs, each pixel of the input image is described with a set of 32 features, each of them is the magnitude of output of the corresponding Gabor filter. The we used Support Vector Machines (SVM) to classify pixels into two categories: areas of troughs and other areas.

2.2 Tests on image Q

The image used in the first series of tests is presented in Fig. 16.

Fig. 16. Original image Q_{zoom1} used in this test.

For training of SVM we indicated manually areas of troughs – see mask in Fig. 17a. The black area was used as a positive example. In Fig. 17b we show mask for areas that were excluded from training and they were neither used as positive nor negative examples. The rest of image was used as negative examples.



a b

Fig. 17. Mask for positive examples (a) and for part of the image excluded from training (b). Part of the image not covered by any of these two masks was used for training as negative examples.

In Fig. 18 we present the results of troughs identification.

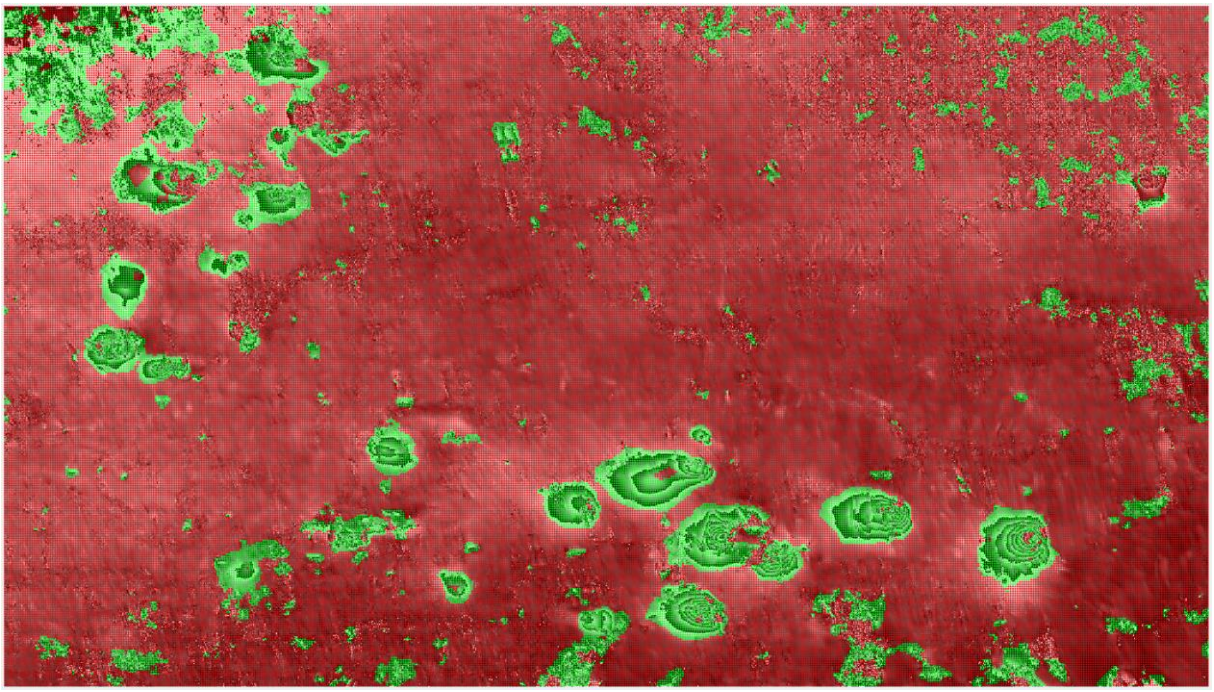


Fig. 18. Results of linear SVM classifier, learned on grid 3x3 for positive and 20x20 for negative samples (5230 positive and 5726 negative samples).

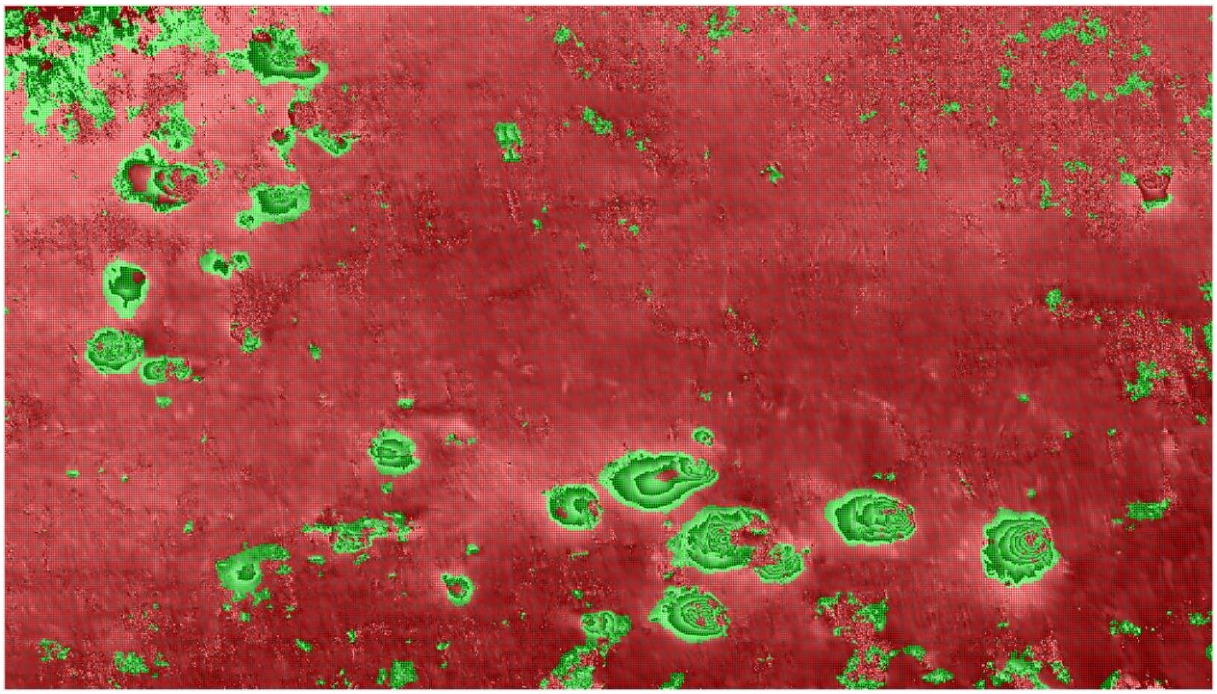


Fig. 19. Results of linear SVM classifier, learned on grid 2x2 for positive and 10x10 for negative samples (11 760 positive and 22 606 negative samples).

2.3 Difficult case: image P

2.3.1 Test of the classifier learned from Q

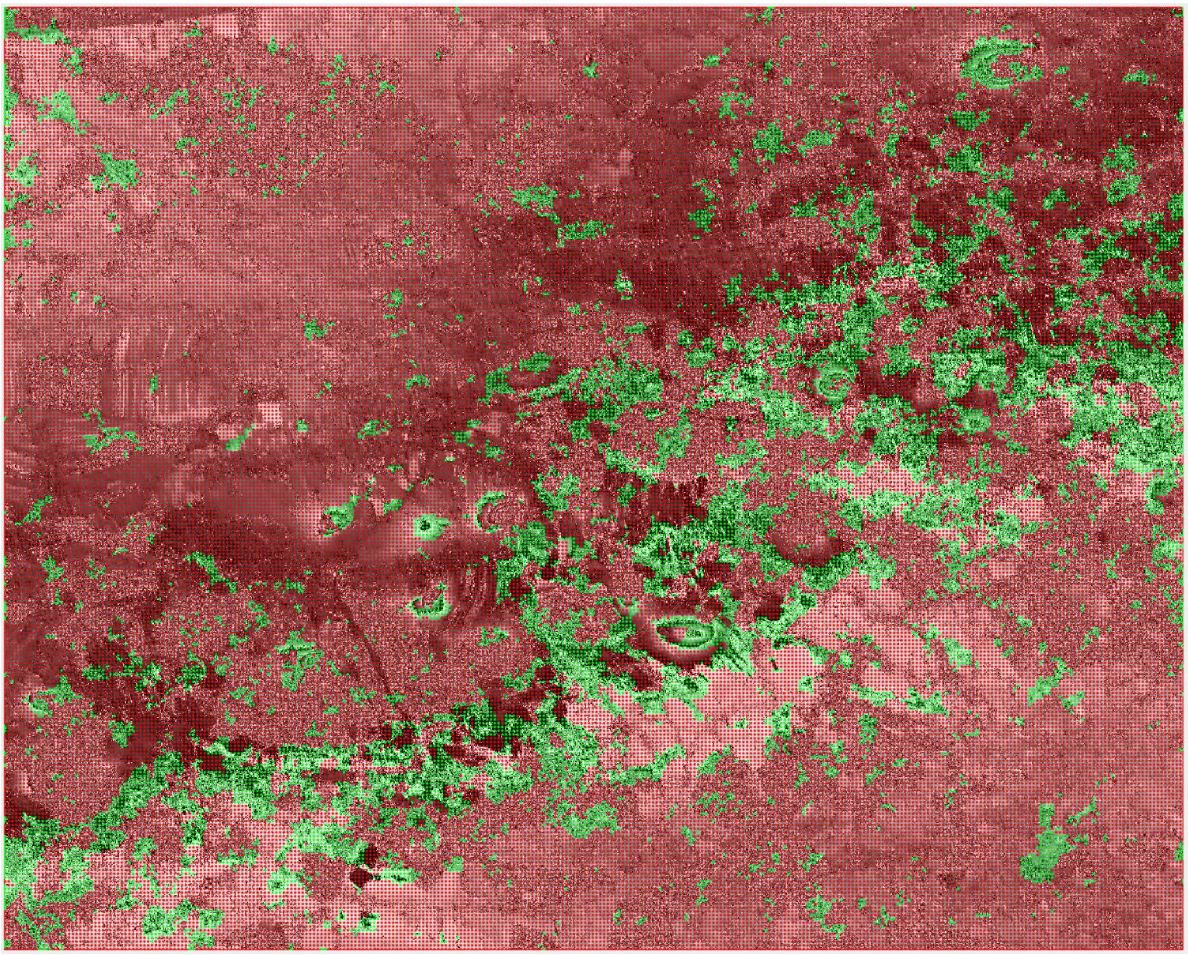


Fig. 20. Results for image P, while the classifier was trained on image Q_{zoom1}. Linear SVM classifier was used, learned on grid 2x2 for positive samples and on grid 10x10 for negative samples (total 11 760 positive and 22 606 negative samples).

2.3.2 Tests of the classifier learned from P

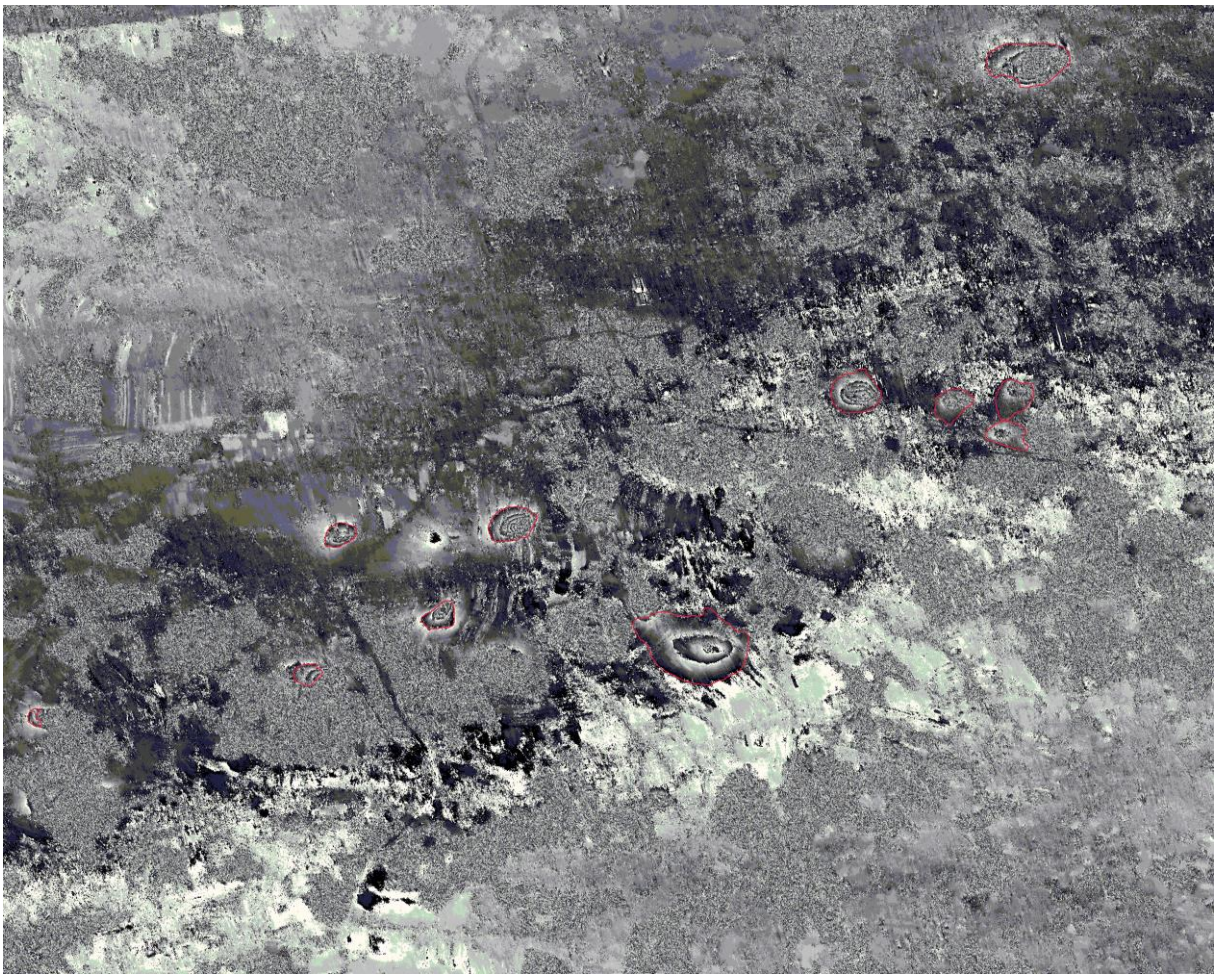


Fig. 21. Positive samples marked for learning the SVM classifier.

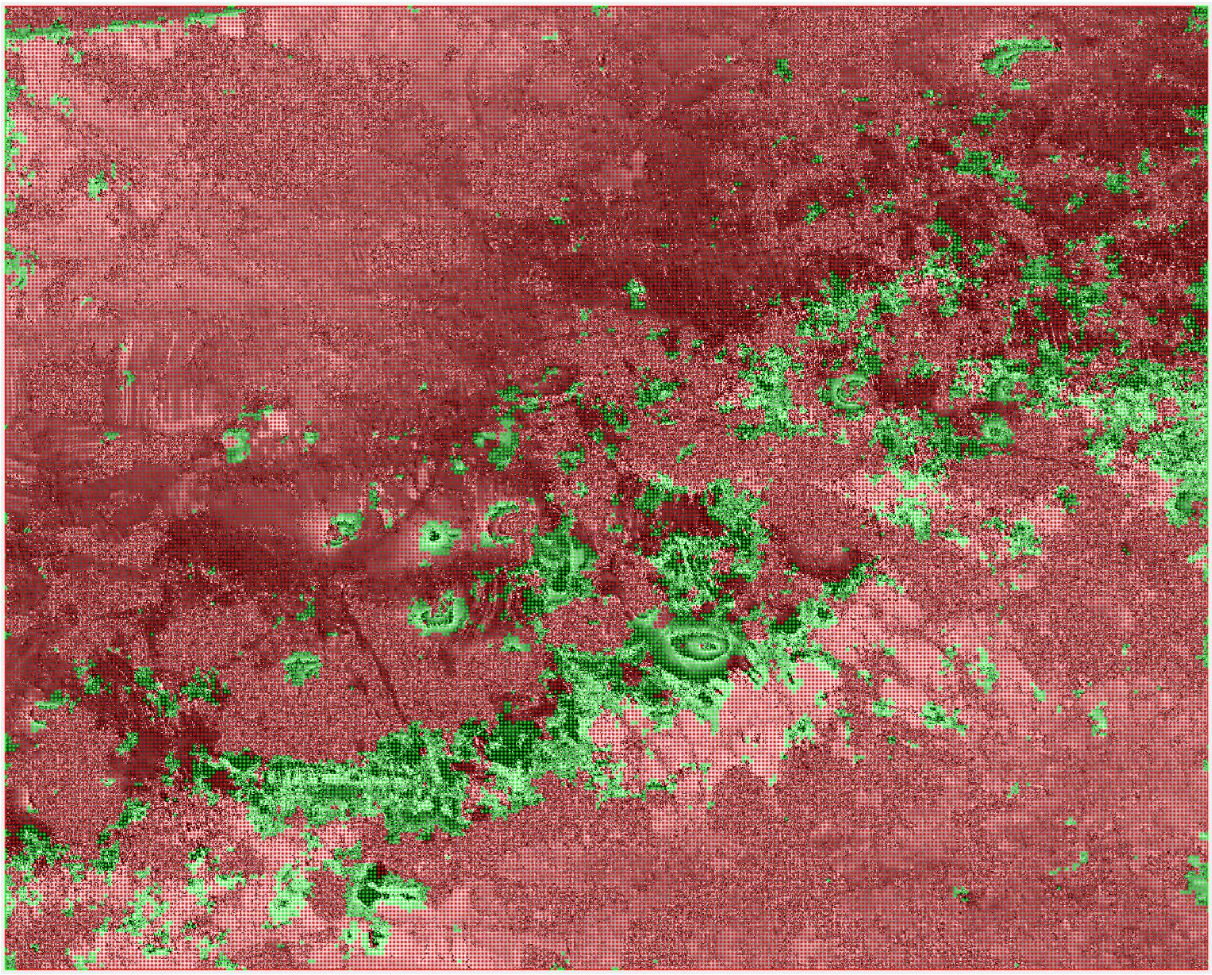


Fig. 22. Results for image P, while the classifier was also trained on image P. Linear SVM classifier was used, learned on grid 3x3 for positive samples and on grid 20x20 for negative samples (total 8 099 positive and 12 535 negative samples).

For wavelet vector: T = [6.8 10 15 22.8 34.2 51.2] the result was practically identical as presented in Fig. 22, so the highest frequencies in our filter bank that correspond to wavelengths 3 and 4.5 do not influence the discriminative hyperplane parameters.

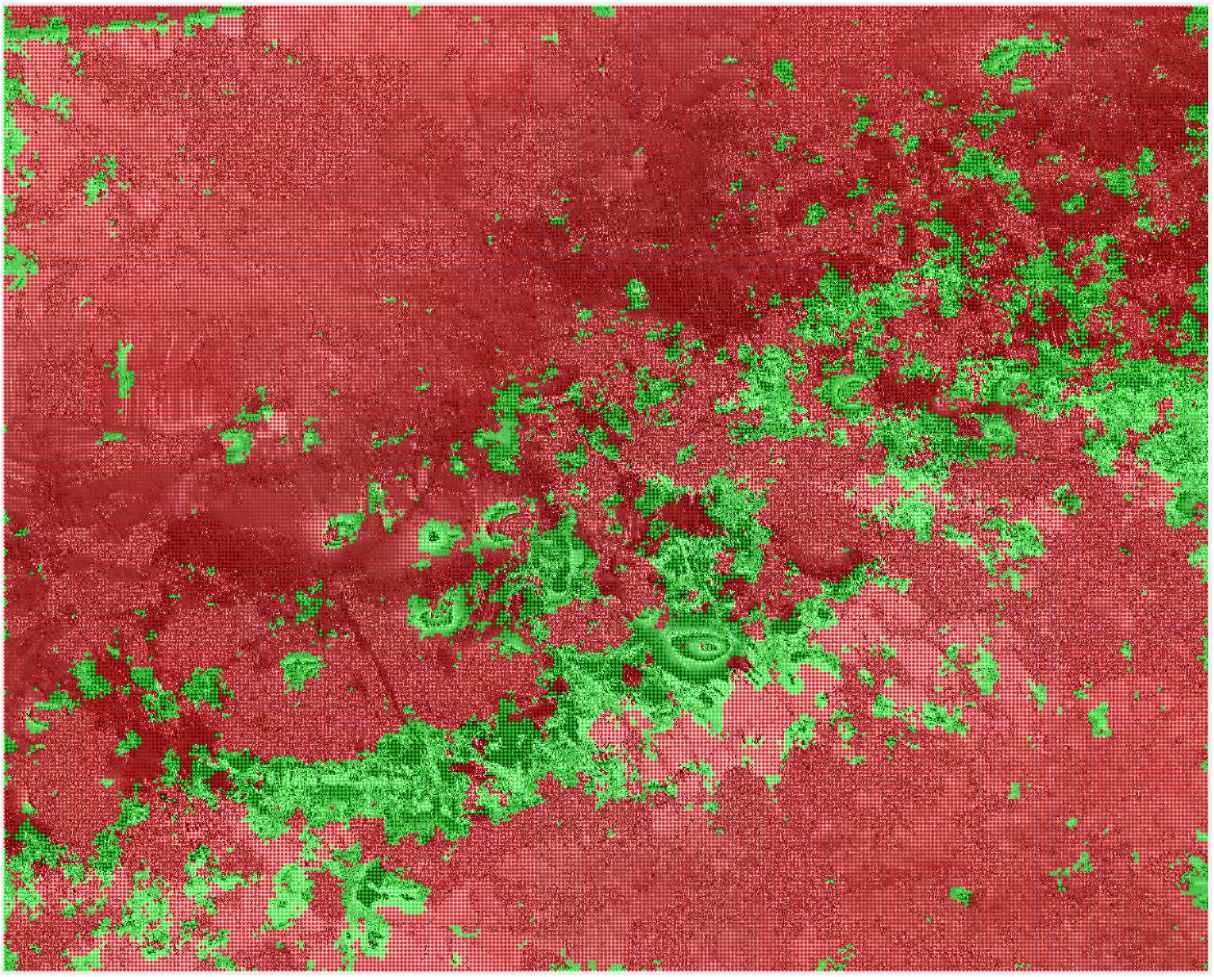


Fig. 23. Results for image P, while the classifier was also trained on image P. Linear SVM classifier was used, learned on grid 2x2 for positive samples and on grid 15x15 for negative samples (total 18 243 positive and 22 056 negative samples).

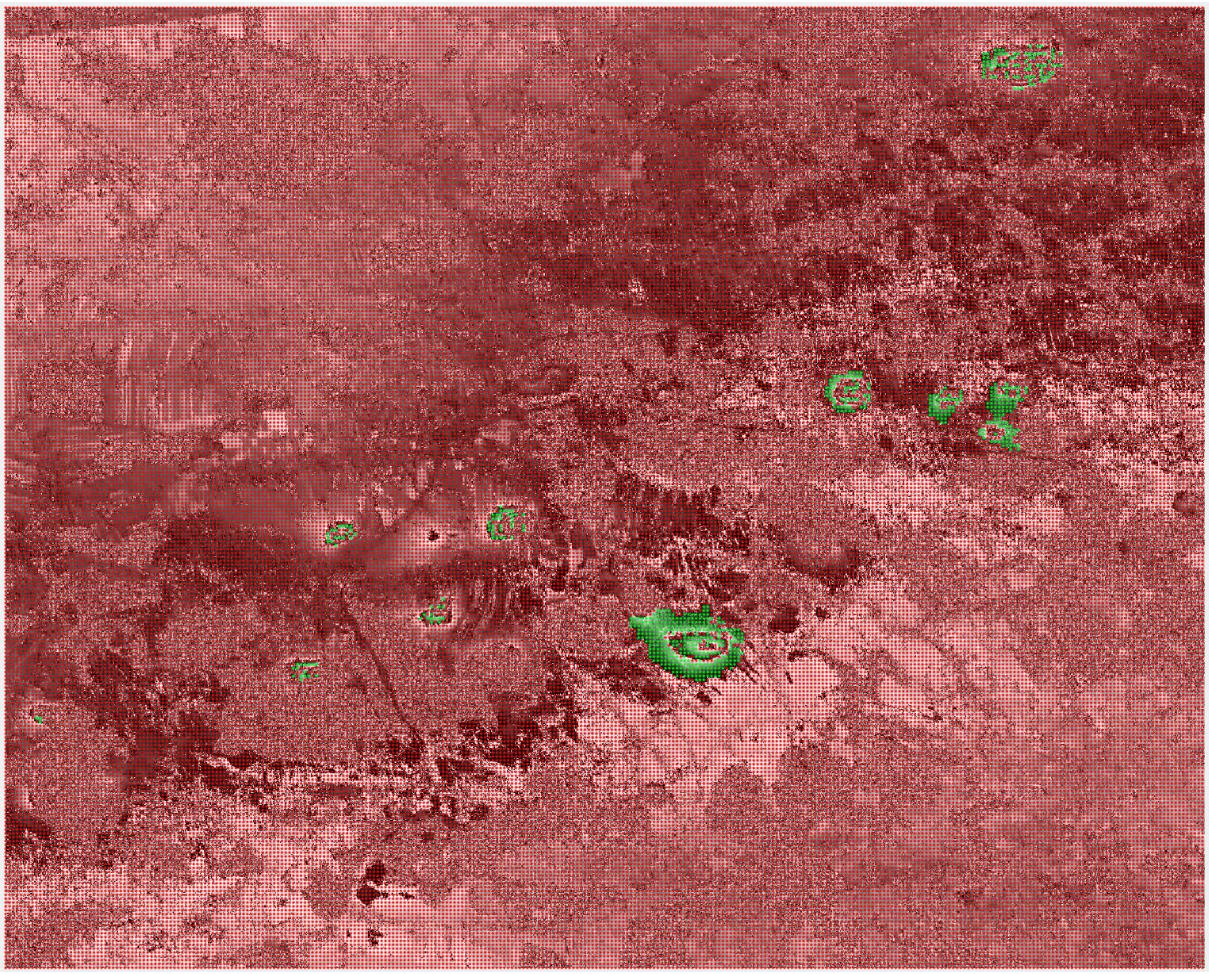
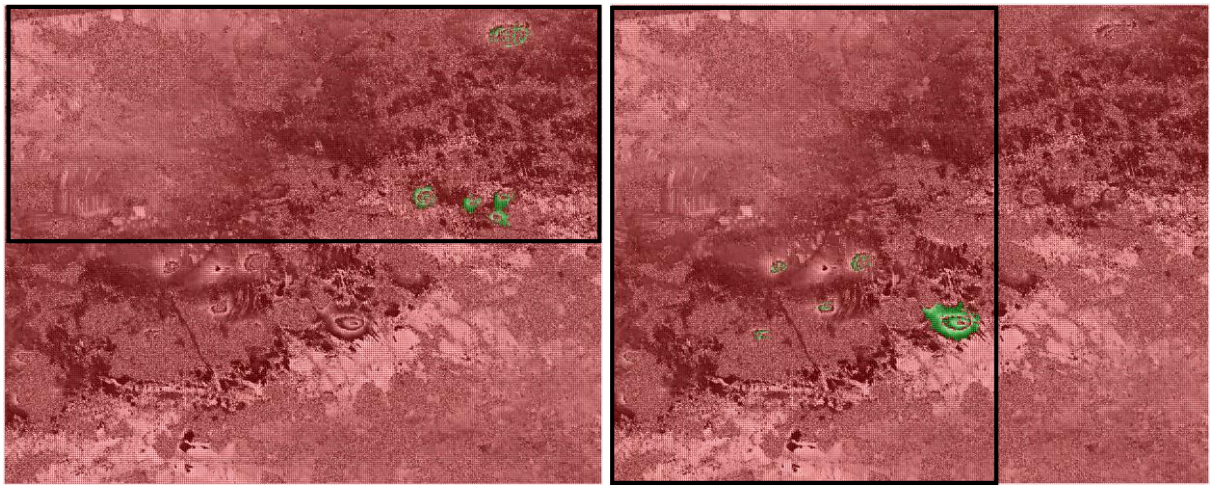


Fig. 24. Results for image P, while the classifier was also trained on image P. SVM classifier with RBF kernel was used, learned on grid 3x3 for positive samples and on grid 20x20 for negative samples (total 8 099 positive and 12 535 negative samples).



a b

Fig. 25. Results for image P, when the classifier was trained on a part of image P in the black frame. SVM classifier with RBF kernel was used, learned on grid 3x3 for positive samples and on grid 20x20 for negative samples

Fig. 26 presents the results of classification with SVM with RBF kernel after modification of kernel scale parameter from 1 to 100. This should enlarge a part of the decision space classified as trough areas and it could improve classification if the reason for errors was to high rejection rate. However, the result presented in the image confirms that in case of this noisy data, the classifier is unable to discriminate troughs based on the set of features that we used. The classifier was trained on a part of image P in the black frame.

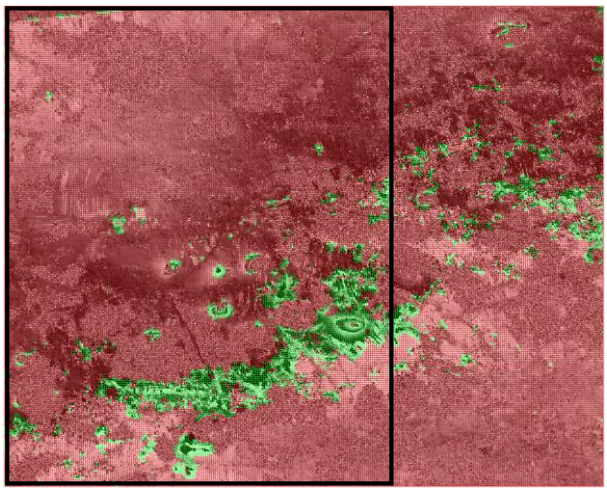


Fig. 26. Classification with RBF after changing KernelScale parameter from 1 to 100. The classifier was trained on a part of image P in the black frame

2.3.3 Learning from P and tests on Q

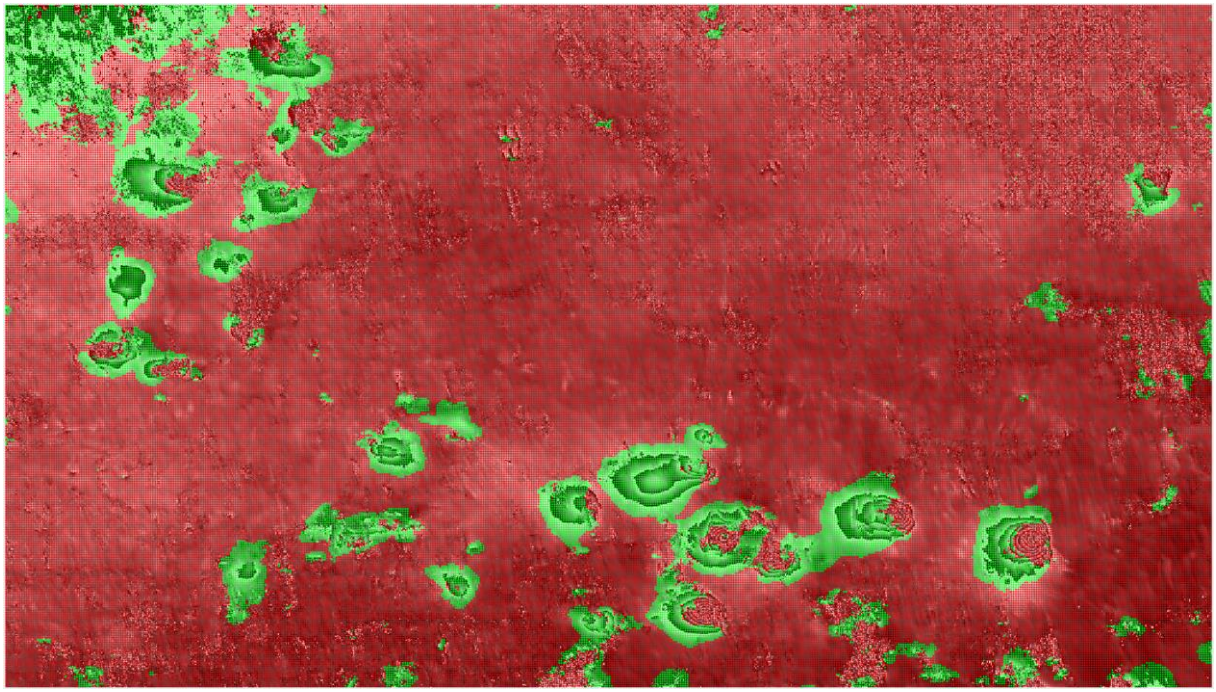


Fig. 27. Tests on image Q of SVM classifier learned on image P (8099 positive samples and 12535 negative samples, linear SVM classifier).

The result is obviously worse than when learning data were taken from Q (Fig. 18, Fig. 19) but it can be regarded as almost satisfactory and after some processing it could be used in next stages. We can conclude that the main reason for incorrect classification of image P is that in some areas the background has similar spectral properties to trough areas.

2.3.4 Dependence of Gabor filter parameters

We checked if increasing the number of frequencies in the Gabor filter bank can improve its discriminative properties. We increased the number of frequencies to 5 per octave, like recommended for SIFT method. The smallest and the biggest wavelengths were set to 4 and 50 pixels, based on measurements of troughs' geometry from images. Wavelengths are therefore calculated as: $T_i = T_0 k^i$, where $k = \exp(\ln 2 / S)$, where S=5 is the number of scales per octave:

T = [4. 4.59 5.28 6.06 6.96 8.00 9.19 10.56 12.13 13.93 16.00 18.38 21.11 24.25 27.86 32.00 36.76 42.2200 48.50]

The number of orientations remains 4, so there are 76 filters in the bank. Other parameters of learning and classification were identical like in experiment presented in Fig. 22. The result is presented in Fig. 28.

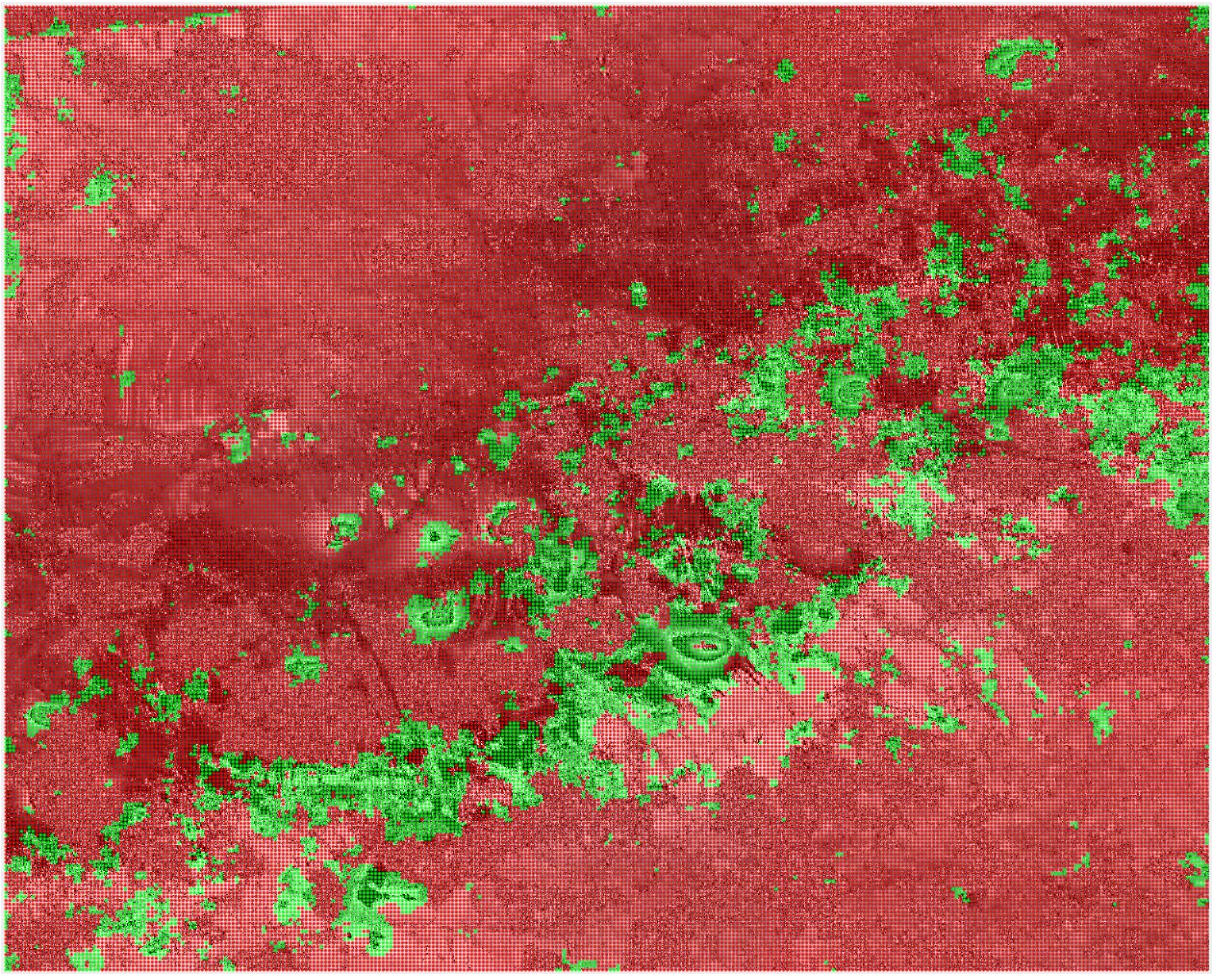


Fig. 28. For 19 frequencies of Gabor filters (76 features)

The result is very similar to what we got from filter bank with 8 frequencies (and 32 features) so we can conclude that increasing the number of frequencies is not relevant for discriminative properties of the feature set.

2.3.5 Feature extraction from gradient image

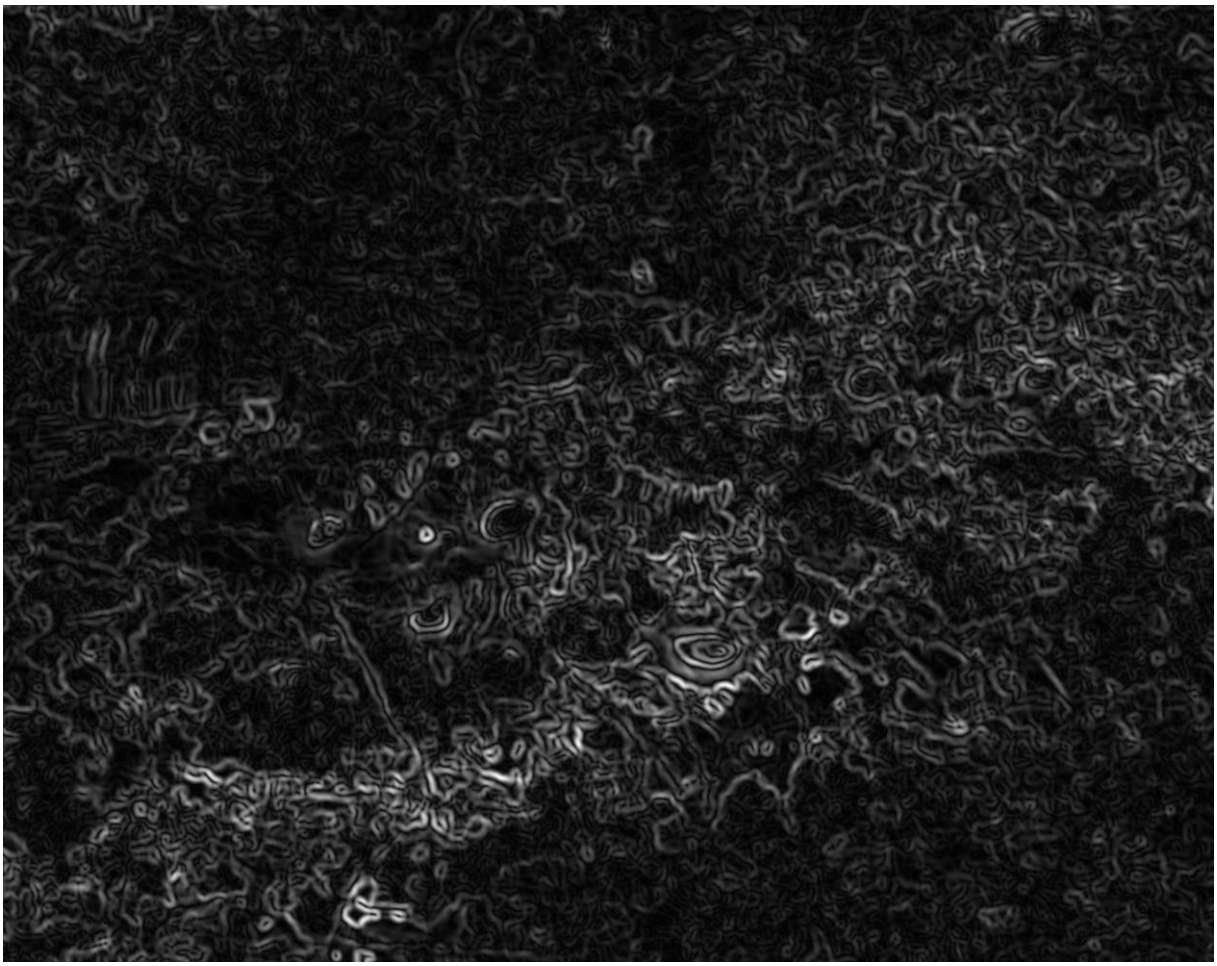


Fig. 29 Gradient of image P blurred with Gauss (30,5).

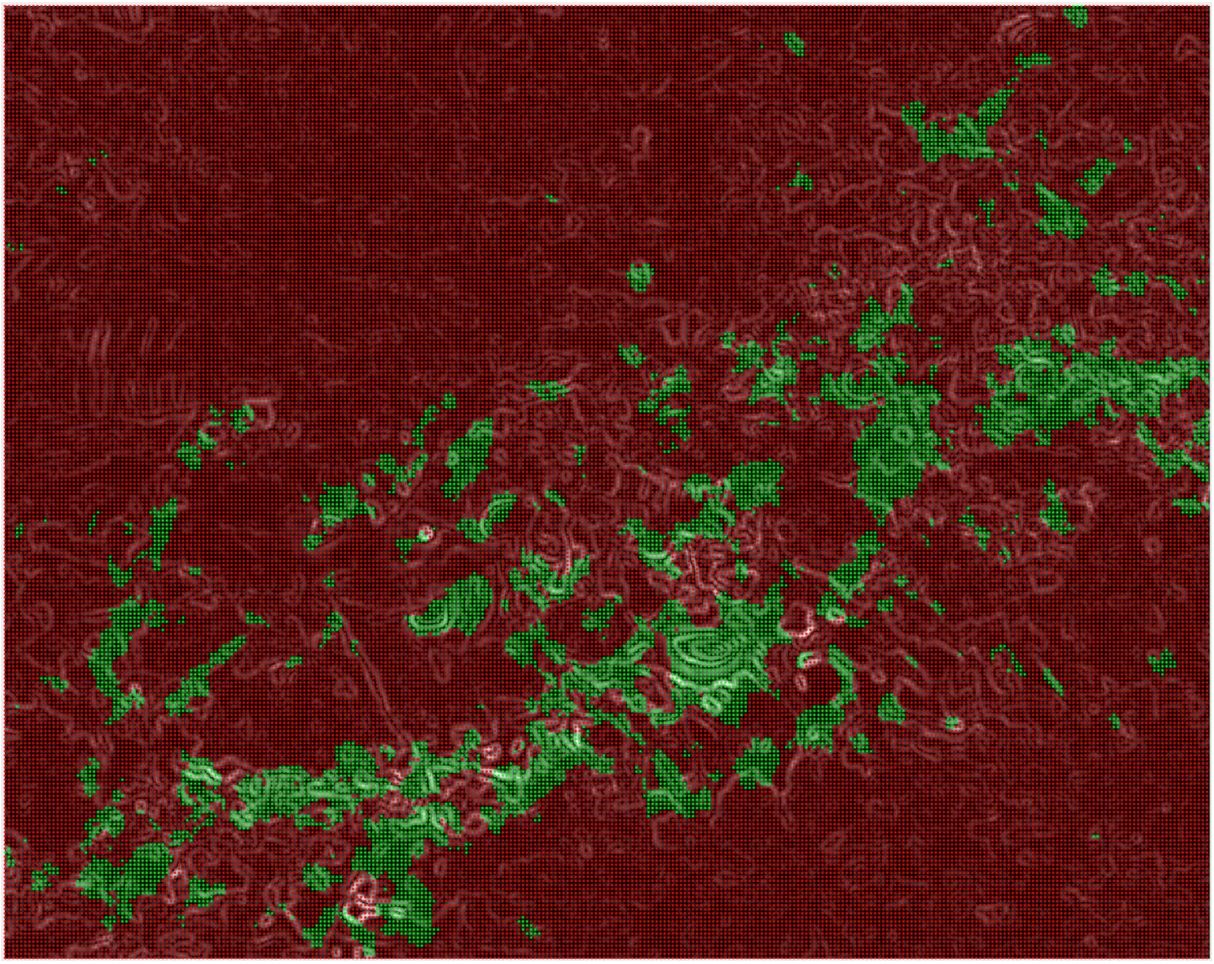


Fig. 30 The result of classification of gradient image of P (gradient image was also used as a training set).

2.4 Time of learning and classification

Time of Gabor filter bank calculating for image resolution 1187x2102 pixels, for 32 filters (8 frequencies, 4 orientations) was 27 seconds. Time of SVM learning and classification is shown in Tab. 2 and Tab. 3.

Tab. 2. Time of learning

Image resolution	**Positive grid**	**Negative grid**	**Positive samples**	**Negative samples**	**Total samples**	**SVM type**	**Time of learning**
Q: 1187x2102	3	20	5230	5726	10 956	linear	25.2 s
Q: 1187x2102	2	10	11 760	22 606	34 366	linear	294.6 s
P: 2001x2501	2	15	18 243	22 056	40 299	linear	721.8 s
P: 2001x2501	3	20	8 099	12 535	20 634	linear	222.1 s
P: 2001x2501	3	20	8 099	12 535	20 634	rbf	61.8 s

Tab. 3. Time of SVM classification (linear)

Image resolution	Grid	Total samples	Time of data preparation [s]	Time of classification	Time of plot generation
1187x2102	20x20	6360	0.07	0.1 s	0
1187x2102	10x10	25 109	0.44	0.4 s	0
1187x2102	5x5	100 198	2.9	1.75 s	0.007
1187x2102	2x2	624 294	25.8	11.5	0.02
1187x2102	1x1	2 495 074	16.9	45.6	0.08

3. Verification of troughs' centres using local conversion to polar coordinates and wavelets

3.1 General concept

This method is designed to verify whether a specified point (x_c, y_c) lays in the centre of a trough. In the first step, image 200x200 pixels centred at (x_c, y_c) is retrieved from the input image. Then this sub-image is converted to polar coordinates (r, θ) – see Fig. 31. The idea is to compare frequency spectrum in horizontal and in vertical directions in the converted image. Classification is made based on the average magnitude of frequency response in vertical direction divided by the average magnitude of frequency response in horizontal direction, at several (between 3 and 5) selected frequencies.

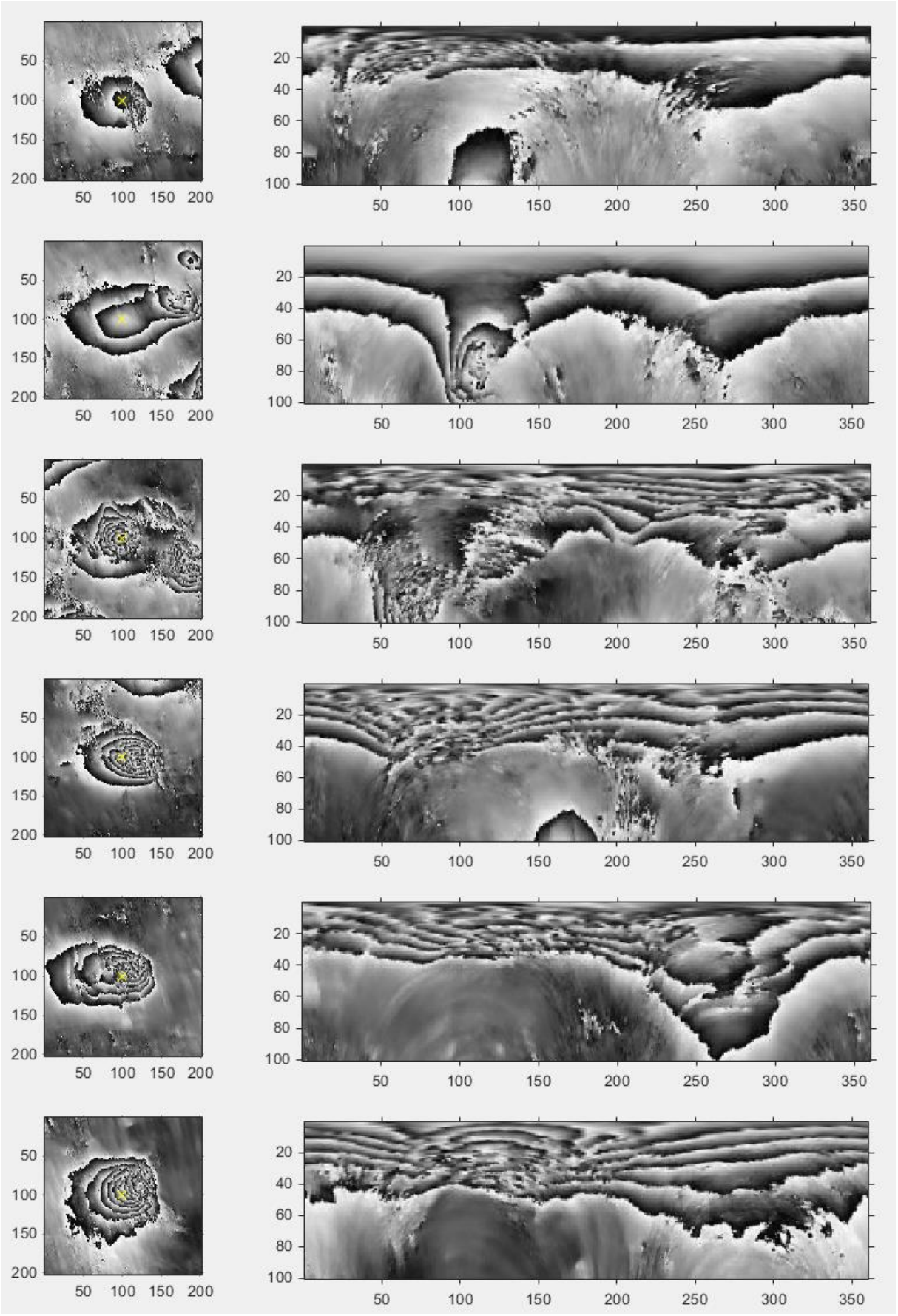


Fig. 31. 120x120 images centred at points marked manually and corresponding image in polar coordinates. In left images angle increases clockwise, starting from pointing down.

Because of image quantization and inhomogeneous distribution of pixels during transformation from Cartesian to polar coordinate system, the output image has always more distinctive spectrum in vertical direction than in horizontal, see Fig. 32. Therefore ratio of vertical to horizontal magnitude above 1 is typical and it does not necessarily mean that a trough is likely to occur in the input image. It is necessary to produce some positive and negative examples and train our classifier.

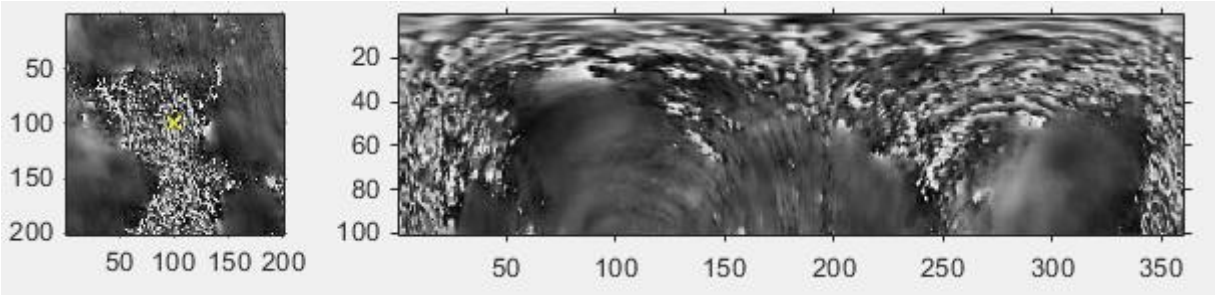


Fig. 32. Example of 100x100 image that does not contain ellipse. Due to Cartesian to polar conversion, horizontal frequencies are stronger than vertical. Gabor ratios for wavelengths [2 4 8 16] are: [6.76, 4.17, 2.5, 1.36].

3.2 Training set

For tests we used 12 positive samples extracted from image Q_{zoom1}, see Fig. 33. Each image is centred at the centre of the trough. The set of 1521 negative samples was collected from image Q and samples were taken every 50 pixels. Twelve examples of negative samples are presented in Fig. 34.

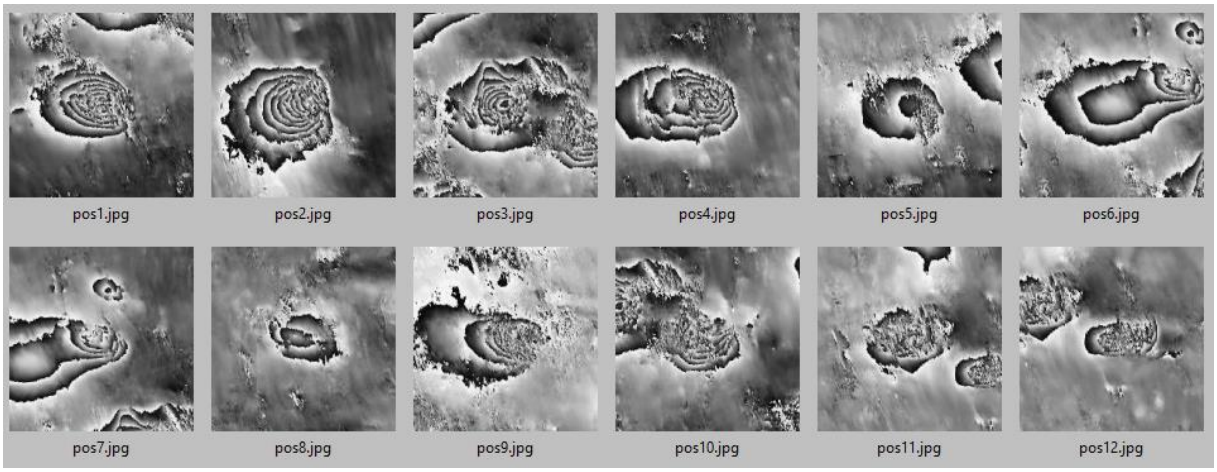


Fig. 33 Positive samples extracted from image Q_{zoom1}.

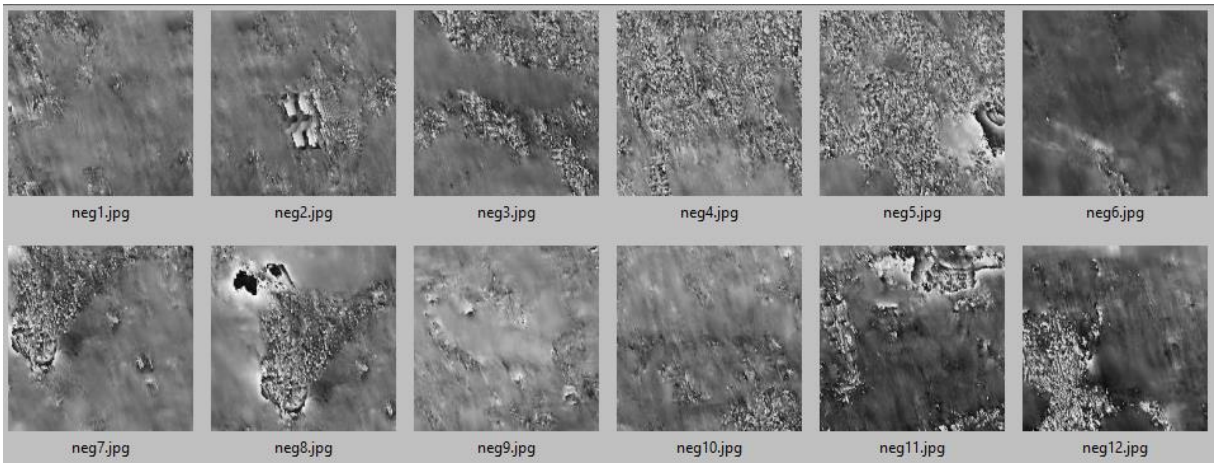


Fig. 34. Examples of negative samples extracted from image Q_{zoom1}. Total number of negative samples was 1521.

3.3 Results

In Tab. 4 we present the set of features for positive and 12 negative samples of the training set. Features are means of the magnitude of Gabor filter response in vertical direction divided by the mean of magnitude of the response in horizontal direction for a specified wavelength. For averaging we took radius *r* from 2 to 35, because this range in the most meaningful and for all troughs it corresponds to area within the trough. The distribution of features is graphically presented in Fig. 35 and we can see that in this case positive and negative samples form two linearly separable clusters.

Tab. 4 Values of features for positive samples and 12 selected negative samples. T is a wavelength of Gabor wavelet.

Sample	**T=3**	**T=6**	**T=12**
pos1	5.4680	4.3786	2.9994
pos2	6.7159	4.5722	2.4028
pos3	9.4640	7.0741	4.7459
pos4	5.8296	5.2174	6.0120
pos5	4.9588	3.8830	5.1056
pos6	3.4251	4.5189	5.9332
pos7	8.4044	7.3313	4.7165
pos8	8.2522	7.2082	2.4152
pos9	7.6227	6.6028	3.3909
pos10	6.4191	4.6322	2.6291
pos11	6.7129	6.0481	3.0151
pos12	7.7110	7.3675	2.5192
neg1	2.1525	1.5158	1.4146
neg2	6.1506	3.6626	4.5064
neg3	3.2660	2.7580	2.2709
neg4	5.2012	2.6156	1.4172
neg5	5.3131	3.1813	1.8501
neg6	1.5640	2.4201	1.4809
neg7	3.0539	2.5002	1.9123
neg8	4.5064	2.6790	1.0734
neg9	1.5826	1.5434	1.4621
neg10	3.5592	3.3148	2.0284
neg11	4.4119	2.6043	2.2896
neg12	3.2549	2.4387	1.9127

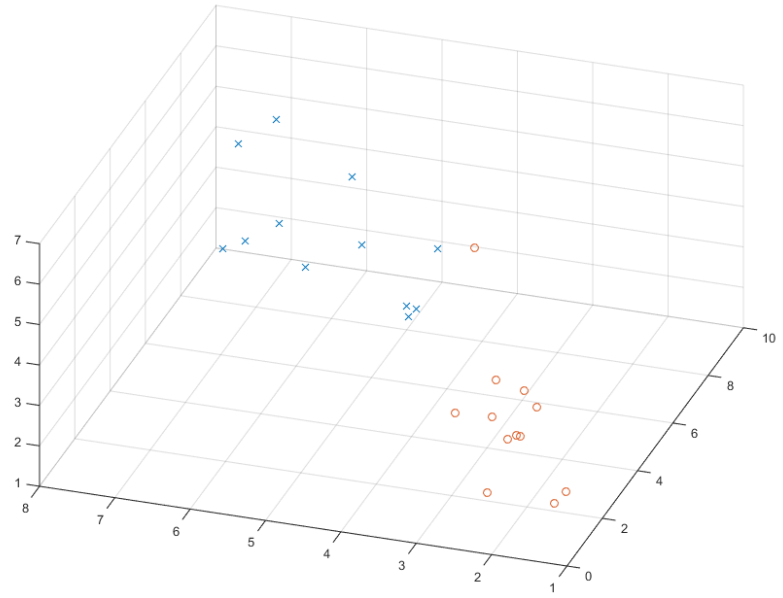


Fig. 35. Positive (x) and 12 negative (o) samples in the space of Gabor descriptors

In Fig. 36 we present positive samples and the whole set of 1521 negative samples in feature space. When rotating this image, we observed that some positive samples are located within the area of negative samples. This means that there is no classifier that separates perfectly these two areas.

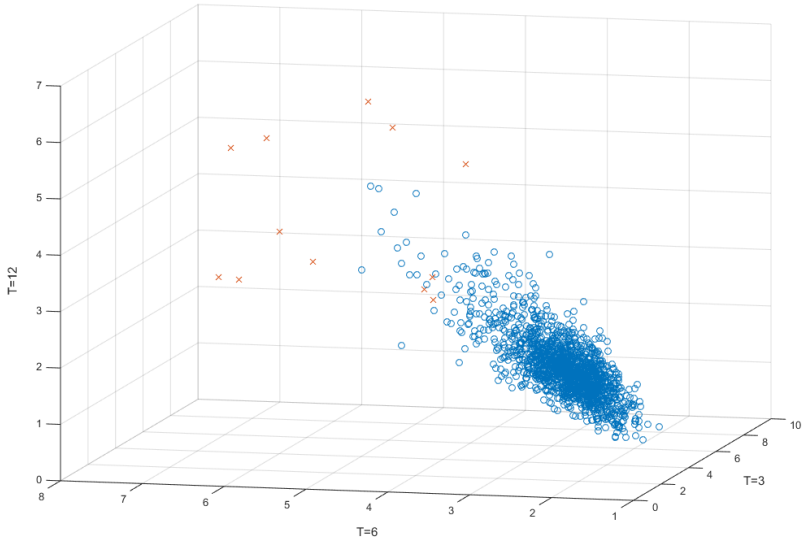


Fig. 36. Positive (x) and all 1521 negative (o) samples in the space of Gabor descriptors

Conclusions: This method could be useful, although some false alarms should be expected. Increasing the set of features (more filters, different ranges of r) may improve the performance and it could be used for verification in the future.

3.4 Local polar coordinates for gradient image

We have checked how the proposed method works on gradient image. Original image, presented in Fig. 37 is blurred using a Gaussian filter (Fig. 38) to make it less susceptible to noise and then magnitude of its gradient is calculated (Fig. 39).

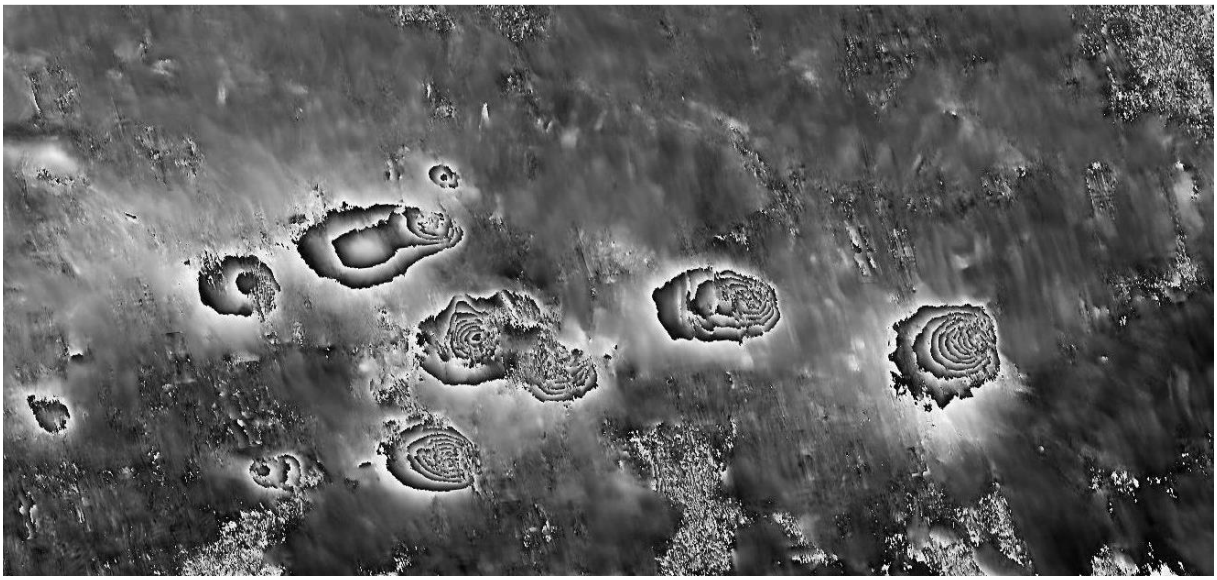


Fig. 37. Original image

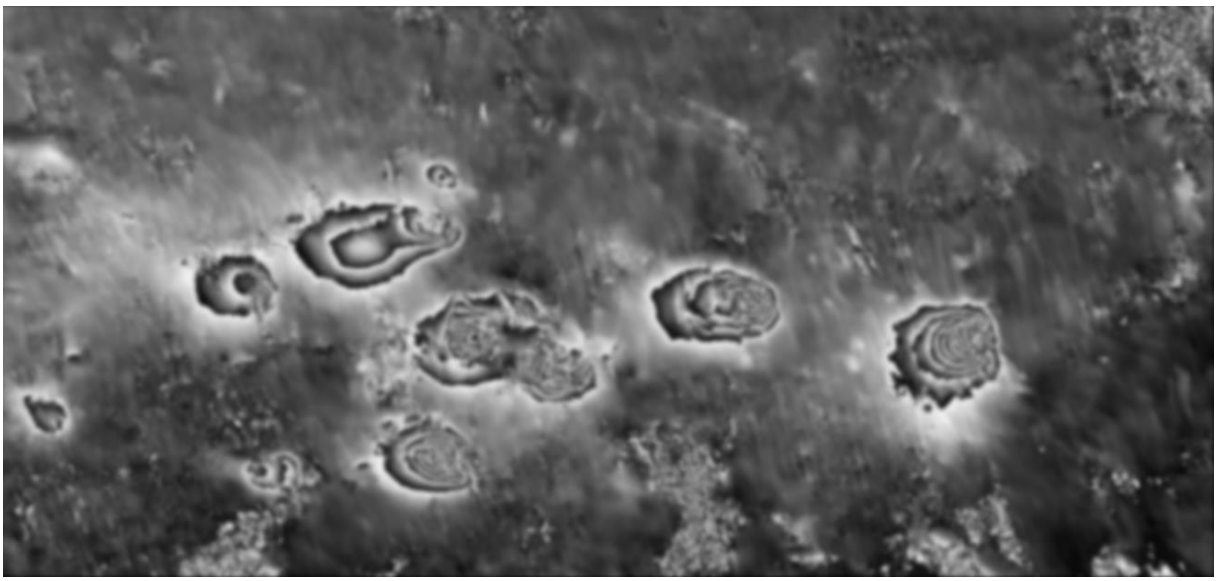


Fig. 38. Image after filtration (Gaussian 15, 2.5)

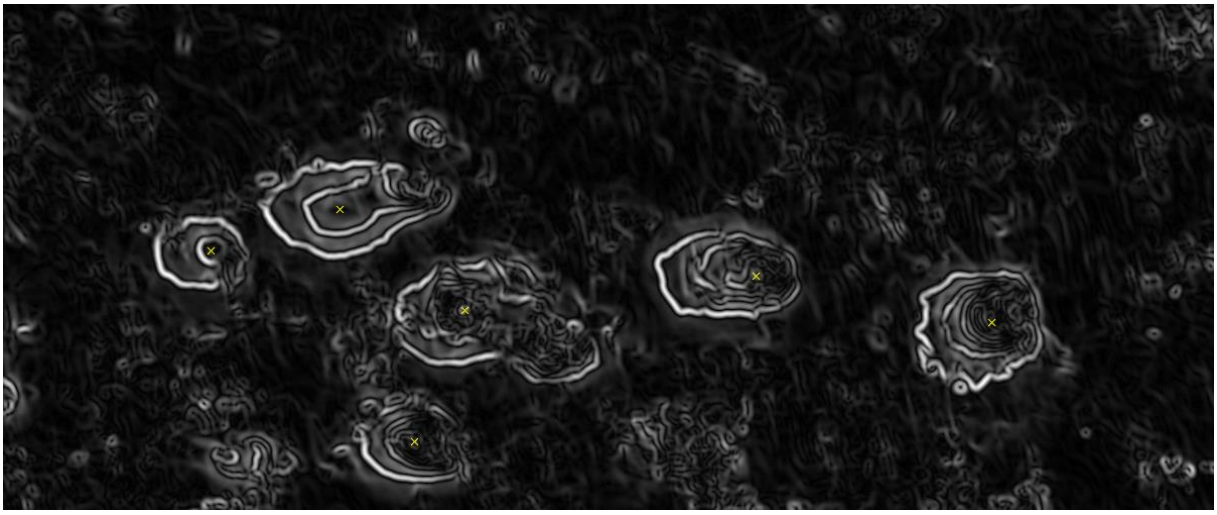


Fig. 39. Image of gradient magnitude

In Fig. 40 we present positive samples and corresponding images in polar coordinates. After some experiments with feature extraction for different sets of parameters we found out that there is no improvement so we resigned from further tests.

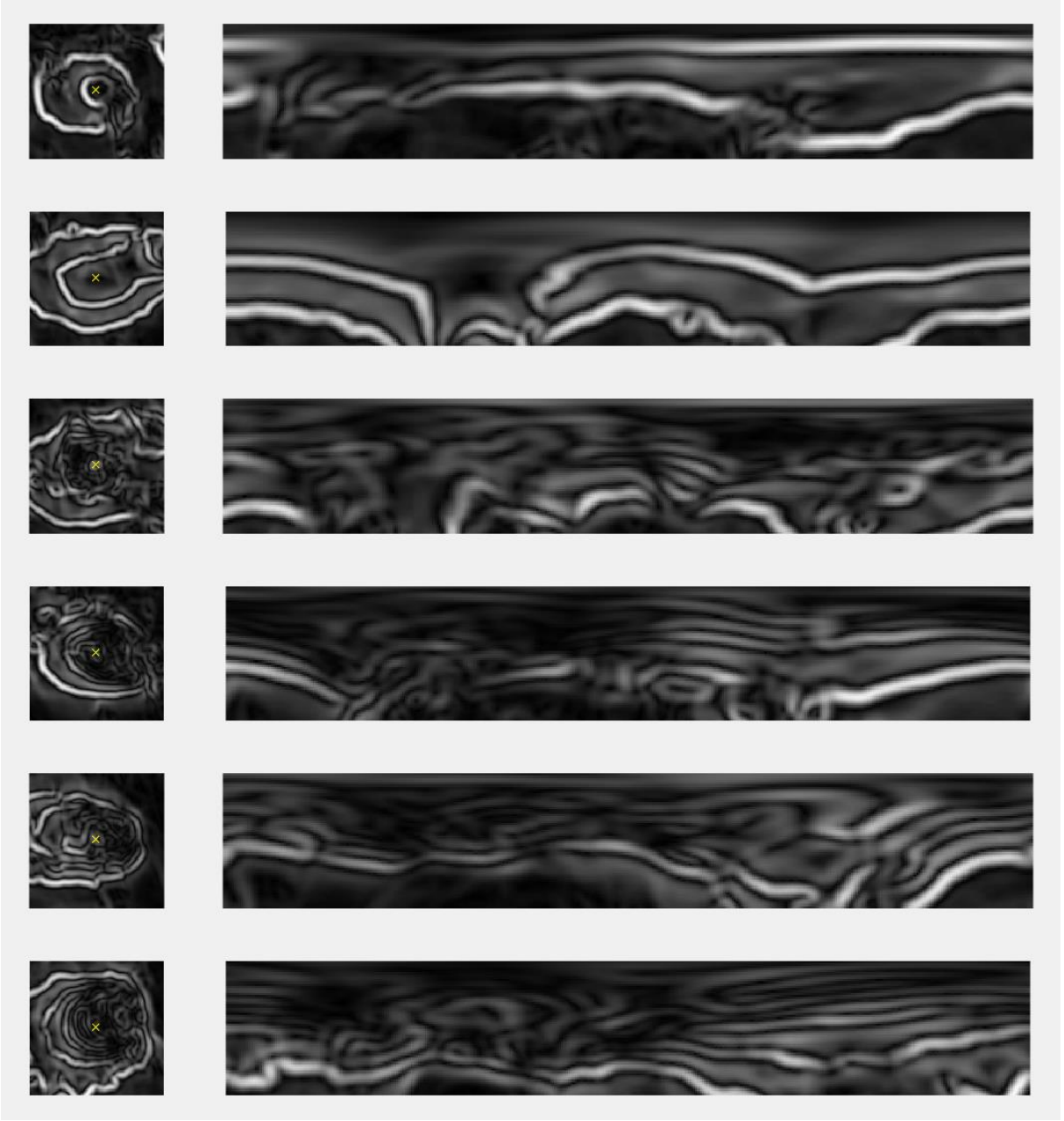


Fig. 40. 120x120 images centred at points marked manually and corresponding image of gradient magnitude in polar coordinates for r=3(top) to r=60 (bottom) and angle from 0 (left) to 360 (right). Angle increases clockwise, starting from pointing down.