

INDEXING OF SATELLITE IMAGES OF DIFFERENT SPATIAL RESOLUTIONS USING MULTI FEATURES

Bin Luo, Shujing Jiang, Liangpei Zhang, Xin Huang

Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University
Wuhan, China
robinlb2002@gmail.com

ABSTRACT:

In this paper, we investigate the issue of the retrievals of remote sensing images with different spatial resolutions by using texture features. Since the texture features extracted on images with different resolutions are not directly comparable, we use the resolution or scale invariance of the texture features to make them comparable. Image retrieval experiments are carried out on the database composed by SPOT images at two different resolutions.

1 INTRODUCTION

Space agencies have collected databases with huge amount of remote sensing images over the last decades. The indexing of the images from such databases is a key issue for the space agencies. Since 1990s, the Content Based Image Retrieval (CBIR) approaches have been proposed for the natural image or multimedia databases (see (Lew et al., 2006) for a review). The CBIR systems, such as the KES (Colapicchioni and iee, 2004), KIM (Datcu et al., 2003), for remote sensing images have also been developed.

One particularity of the remote sensing image databases, when compared to the natural images databases, is that they are constituted by images with different but known spatial resolutions. Classical low-level features, such as Gabor features, GLCM features, wavelet features, etc., extracted from images with different resolutions are not directly comparable. Thus the images with different resolutions can't be jointly indexed. Fortunately, many feature extraction methods are either scale invariant or resolution invariant. For example, the Scale Invariant Feature Transformation (SIFT) has been proposed (Lowe, n.d.). The scale invariance of the Gabor filter bank has been investigated in (Kamraining et al., 2006). In (Luo et al., 2008), the authors have proposed the resolution invariance for the Gaussian derivative wavelets in order to compare the Gaussian wavelet features obtained from the images with different resolutions. The invariances of these low-level features make it possible to compare the features extracted from images with different resolutions, which can hence be jointly indexed.

The main contribution of this paper is to propose the schemes for comparing the Gaussian wavelet features, the Gabor features, the GLCM features and the shape features by using the resolution/scale invariances for the joint retrieval of remote sensing images at different resolutions. The well-known SIFT features have not been investigated in this study due to the fact that the SIFT features are proposed for describing local salient points over scales, rather than the whole scene containing in the image.

The paper is organized as follows. In Section 2, we briefly introduce the features used for indexing the remote sensing images. In addition, in the same section, we present how to compare the features extracted from the images at different resolutions for the indexing. In Section 3, we present the data sets used for experiments, as well as the parameters used for extracting the features.

In Section 4, the retrieval results of remote sensing images at different resolutions are shown. In Section 5, we conclude.

2 EXTRACTION OF FEATURES OBTAINED AT DIFFERENT RESOLUTIONS

The radiometric features and the texture features are used for indexing of remote sensing images at different resolutions. In this section, we briefly introduce the methods for extracting the features. Since the low level features extracted on the images at different resolutions are not always comparable, in addition, the approaches which allow to compare the features extracted from images at different resolutions are also proposed for each type of features in this section.

2.1 Radiometric features

The mean values and the standard deviations of the images are computed as radiometric features for indexing. More concretely, the radiometric features for an image is defined by:

$$\Theta_R = \{\mu, S\}, \quad (1)$$

where $\mu = \frac{\sum_{x,y} I(x,y)}{N}$ ($I(x,y)$ is the gray value of the pixel x on the image and N is the number of pixels in an image), $S = \sqrt{\frac{\sum_{x,y} (I(x,y) - \mu)^2}{N}}$.

If the histogram of the remote sensing images on the same scene but with different spatial resolutions are similar, their radiometric features can be directly compared without any additional steps.

2.2 Texture features

Three classical methods are used for extracting texture features in this paper: the continuous Gaussian wavelets and the Gray Level Co-occurrence Matrix (GLCM).

2.2.1 Gaussian wavelet features The Gaussian scale-space representation of an image I is defined as:

$$L_t = I \star k_t, \quad (2)$$

where $k_t = \frac{1}{\sqrt{2\pi}t} e^{-\frac{x^2+y^2}{2t^2}}$ and t is the scale parameter. The features of the image I_r (of resolution r) at scale t are computed

on the Gaussian scale-space representation by:

$$\theta_{W;t;r} = \{M_{1;t;r}^x, M_{1;t;r}^y, M_{2;t;r}^x, M_{2;t;r}^y\}, \quad (3)$$

where $M_{1;t}^x = \frac{\sum_{x,y} |\partial_x L_t|}{N_r}$, $M_{1;t}^y = \frac{\sum_{x,y} |\partial_y L_t|}{N_r}$, $M_{2;t}^x = \frac{\sum_{x,y} |\partial_x L_t|^2}{N_{r^2}}$ and $M_{2;t}^y = \frac{\sum_{x,y} |\partial_y L_t|^2}{N_{r^2}}$. ∂_x and ∂_y represent the discrete first order derivatives on the horizontal and vertical direction. Usually the features are extracted on a serie of scales $T = \{t_1, t_2, t_3, \dots\}$.

It has been shown in (Luo et al., 2008) that the image acquisition process can be modelled by a Gaussian convolution followed by a sampling. An image I_r of resolution r is obtained by:

$$I_r = \Pi_r \{f \star k_r p\}, \quad (4)$$

where Π_r is the sampling at resolution r , f is the continuous function representing the scene, $k_r p$ is a Gaussian function with standard deviation rp representing the MTF of the instrument, p is the characteristic parameter of the MTF. The larger is p , the smoother is the image. The Gaussian scale-space is scale invariant.

According to (Luo et al., 2008), by using the causality of the Gaussian convolution and Equations (2) and (4), we can obtain the resolution invariance of the Gaussian scape-space features. More concretely, the features extracted on the images of the same scene with different spatial resolutions can be the same. More concretely, for two images I_{r_1} (with resolution r_1) and I_{r_2} (with resolution r_2) on the same scene, the features $\theta_{W;t;r_1}$ and $\theta_{W;t';r_2}$ extracted respectively on I_{r_1} and I_{r_2} are equal, if

$$r_1 \sqrt{t^2 + p^2} = r_2 \sqrt{t'^2 + p^2}. \quad (5)$$

It is shown in (Luo et al., 2008) that the resolution invariance of the Gaussian waveletfeatures yields better indexing results than using only the scale invariance of the Gaussian scale-space.

Therefore, for the images of two different resolutions r_1 and r_2 , we extract the features by Gaussian waveletat respectively two series of scales $T = \{t_1, t_2, t_3, \dots\}$ and $T' = \{t'_1, t'_2, t'_3, \dots\}$ and keep $r_1 \sqrt{t_i^2 + p^2} = r_2 \sqrt{t'_i^2 + p^2}$, the extracted features at different resolutions can thus be compared for indexing.

2.2.2 Gray Level Cooccurrence Matrix (GLCM) features The Gray Level Co-occurrence Matrix computed on an image I_r (with resolution r) is defined as:

$$P(i, j; \delta_x, \delta_y; r) = \# \{(x, y) | I_r(x, y) = i, I_r(x + \delta_x, y + \delta_y) = j\}, \quad (6)$$

where $\#A$ represents the number of elements contained in the set A . Usually, for an image I_r , four co-occurrence matrices are computed at four oriatations:

$$P_{0;d;r} = P(i, j; d, 0; r) \quad (7)$$

$$P_{90;d;r} = P(i, j; 0, d; r) \quad (8)$$

$$P_{45;d;r} = P(i, j; d, d; r) \quad (9)$$

$$P_{135;d;r} = P(i, j; -d, -d; r). \quad (10)$$

where d is the distance parameter. In this paper, 6 features (the Angular Second Moment, the Contrast, the Variance, the Inverse Different Moment, the Shadow of Clustering and the Prominence of Clustering which are proposed in (Haralick et al., 1973)) are respectively extracted from each matrix. The definitions of these

features are shown in the Appendix. We note $H_{l;\theta;d;r}$ ($l = 1, 2, \dots, 14$, $\theta = 0, 90, 45, 135$) the l th feature extracted on the GLCM $P_{\theta;d;r}$.

The GLCM feature set computed on an image I_r with resolution r is defined as:

$$\Theta_{H;d;r} = \{H_{l;\theta;d;r}\}, \quad (11)$$

where $l = 1, 2, \dots, 6$, $\theta = 0, 90, 45, 135$. There are thus $4 \times 6 = 24$ GLCM features extracted for each image I_r .

For the indexing of images at different resolutions, we extract the GCLM features from images with resolution r_1 by using distance d_1 and the GCLM features from images with resolution r_2 by using distance d_2 . Though equality between the GLCM features extracted on two images I_{r_1} and I_{r_2} with two different resolutions (but on the same scene) can not be rigorously established, we can do a rough approximation between $\Theta_{H;d_1;r_1}$ and $\Theta_{H;d_2;r_2}$ by setting

$$d_1 \times r_1 = d_2 \times r_2. \quad (12)$$

3 DATA SETS AND PARAMETERS

3.1 Data sets

The SPOT5 image taken on the Nanjing, China is used. The spatial resolution of the panchromatic image is $2.5m$. Since the SPOT5 provides also multispectral data with $10m$ resolution on the same region, we synthetize a panchromatic image with $10m$ resolution by using the multispectral product in order to assure the terrain types and the weather conditions are similar for the classification for avoiding at most the influences of other factors rather than the difference of the spatial resolutions.

The images are then cutted into small image patches. For the images with the resolution of $2.5m$, the size of each patch is 512×512 pixels; While for the reoslution of $10m$, the size of each patch is 128×128 pixels. Among all the image patches of each resolution, 3 classes of terrains are chosen: the Building class (which contains 194 images representing the urban areas), the Vegetation class (which contains 194 images representing the forest areas) and the Farm class (which contains 202 images representing the rural and agricultural areas). In total, 670 image patches with $2.5m$ resolution and the same number of image patches with $10m$ resolution are used for the classification.

Some examples of the image patches at different resolutions are shown in Figure 1. It can be seen that even though the scene on the ground is the same, the visual appearances (mainly the contrast) of the images at different resolutions are slightly different (especially the images at the first column).

3.2 Parameters

Radiometric, texture and shape features are extracted from the image patches for the experiments.

For extracting the Gaussian waveletfeatures, the parameter p (see Equation (4)) is set to be 0.4, which is experimentally found to be the most appropriate for the SPOT5 instrument. For extracting features from the images with $10m$ resolution, the scale parameters used are $T = \{1.5; 2; 2.5; 3; 3.5; 4\}$. The scale parameters used for the images with $2.5m$ resolution are computed by using Equation (5).

For extracting GLCM features, the GLCM at the four directions $\{0, 45, 90, 135\}$ are computed. For the image with $10m$ resolution, at each direction, the GLCM features of the distances 1, 2 and 3 are extracted. According to Equation (12), for the image with $2.5m$ resolution, at each direction, the GLCM features of the distances 4, 8 and 12 are extracted.

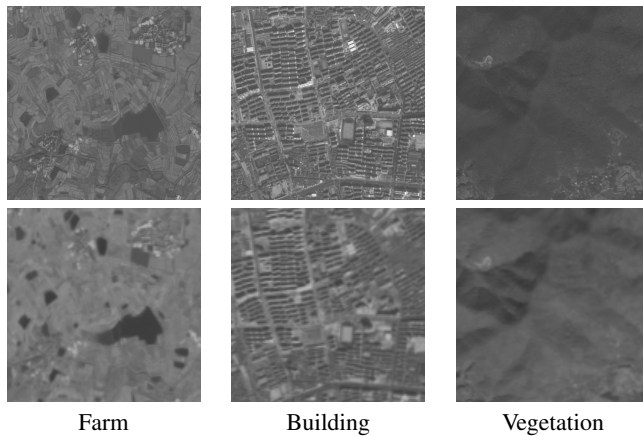


Figure 1: Examples of the SPOT5 image patches taken on Nanjing used for experiments ©SPOTIMAGE. Top row: image patches of the the classes with 2.5m resolution (512×512 pixels). Bottom row: image patches of the the classes with 10m resolution (128×128 pixels).

4 EXPERIMENTS AND RESULTS

We combine the the radiometric features (Θ_R), the GLCM features (Θ_H) and the Gaussian wavelet features (Θ_W) for the retrieval of the SPOT images at different resolutions.

In order to validate this combination of features, we have evaluated the classification performance of different combinations of the Gabor features, the GLCM features, the shape features and the radiometric features (see (Luo et al., n.d.) for details of these features). The results are shown in Figure 2, where the images of 10m resolution (Figure 2(a)) and of 2.5 resolution (Figure 2(b)) are randomly selected as training samples. The percentage of the training samples vary from 10% to 60%. The test sets include all the images of 10m and 2.5m resolutions. It can be seen that the combination of the radiometric features (Θ_R), the GLCM features (Θ_H) and the Gaussian wavelet features (Θ_W) can give the best results.

For each retrieval, a key image patch (with 2.5m or 10m resolution) is selected from from the database for the request. Its radiometric, GLCM and Gaussian scale-space features are then compared with the features of the other image patches in the database. The Euclidean distance between the features are computed as similarity measurement. The most similar image patches are selected as the retrieval result. For each retrieval, 47 image patches are shown. In Figures 3 to 5, the results of three retrieval experiments are shown. The three key image patches belong to the three different classes: Building, Farm and Vegetation.

Several remarks can be drawn from the retrieval results:

- Although the visual appearances (the contrasts) of the images with 2.5m resolution and the images with 10m resolution are different, the retrieval results are quite good. For the Farm and the Building classes, all the retrieved images belong to the same class of the key image. While for the Building and the Vegetation class, there is only one retrieved image belongs to a different class to the key image. This indicates that though the radiometric features are usually pertinent, the other two feature sets (the GLCM and the Gaussian wavelet features) are also important. Moreover, the comparison schemes for the features extracted from images at different resolutions proposed in Section 2 are accurate enough for the joint retrieval.

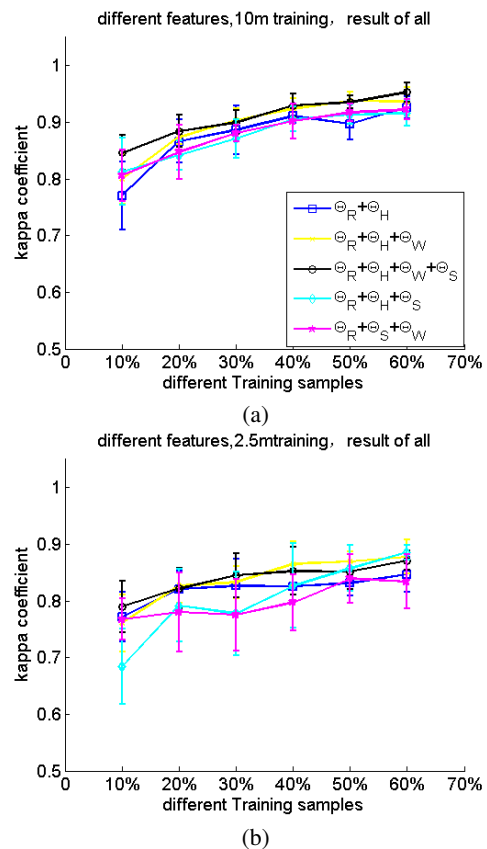


Figure 2: Classification results of different feature sets (a) when the images of 10m resolution is selected as training samples; and (b) when the images of 10m resolution is selected as training samples.

- For the images of the Farm and the Vegetation classes, the images on the same scenes of the key images but at different resolutions are the first ones retrieved by using the feature sets $(\Theta_R + \Theta_H + \Theta_W)$. This indicates that their features are the most similar to that of the key images among all the images in the database, though their spatial resolutions are very different. While for the Building class, the image on the same scene of the key image has also been retrieved by using the feature sets $(\Theta_R + \Theta_H + \Theta_W)$.

5 CONCLUSIONS

In this paper, we have proposed the method for indexing the remote sensing images at different resolutions. The radiometric features, the texture features (including the Gaussian wavelet features, the Gabor features and the GCLM features) and the shape features have been used in this paper. For the Gaussian wavelet features, we have proposed to use the resolution invariance in order to compare the features extracted from images at different resolutions. While for the GLCM features, the distance parameters are tuned according to the resolutions of the images. According to the image retrieval results of remote sensing images at different resolutions, the combination of the radiometric features, the GLCM features and the Gaussian wavelet features is very efficient though the difference of the spatial resolutions is important.

REFERENCES

- Colapicchioni, A. and IEEE, 2004. Kes: Knowledge enabled services for better EO information use. In: Igarss 2004: IEEE International Geoscience and Remote Sensing Symposium Proceedings, pp. 176–179.
- Datcu, M., Daschiel, H., Pelizzari, A., Quartulli, M., Galoppo, A., Colapicchioni, A., Pastori, M., Seidel, K., Marchetti, P. G. and D'Elia, S., 2003. Information mining in remote sensing image archives: System concepts. *IEEE Transactions on Geoscience and Remote Sensing* 41(12), pp. 2923–2936.
- Haralick, R. M., Shanmuga, K. and Dinstein, I., 1973. Textural features for image classification. *IEEE Transactions on Systems Man and Cybernetics* SMC3(6), pp. 610–621.
- Kamarainen, J. K., Kyrki, V. and Kalviainen, H., 2006. Invariance properties of Gabor filter-based features - overview and applications. *IEEE Transactions on Image Processing* 15(5), pp. 1088–1099.
- Lew, M. S., Sebe, N., Djeraba, C. and Jain, R., 2006. Content-based multimedia information retrieval: State of the art and challenges. *ACM Transactions on Multimedia Computing Communications and Applications* 2(1), pp. 1–19.
- Lowe, D., n.d. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, pp. 91–110.
- Luo, B., Aujol, J. F., Gousseau, Y. and Ladjal, S., 2008. Indexing of satellite images with different resolutions by wavelet features. *IEEE Transactions on Image Processing* 17(8), pp. 1465–1472.
- Luo, B., Jiang, S. and Zhang, L., n.d. Performance comparisons of the multi-scale low-level features for the indexing of remote sensing images with different resolutions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, submitted.

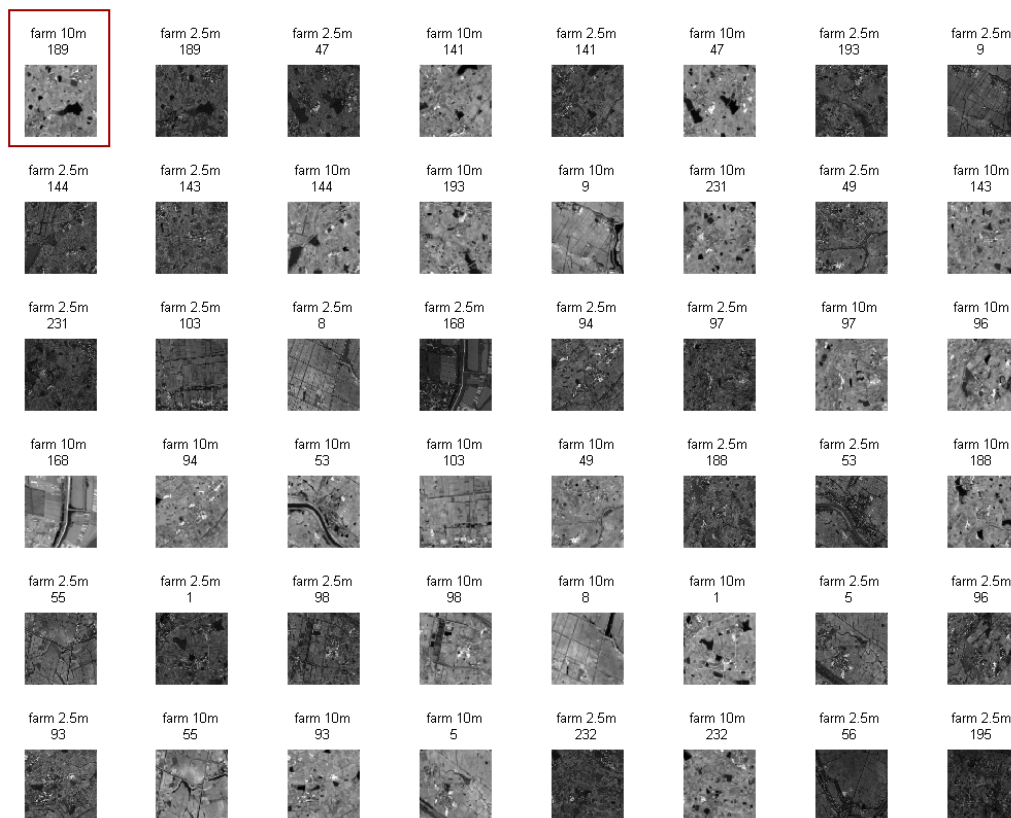


Figure 3: Retrieval results for an image patch of the Farm class. The key image patch, which is on the top left, belongs to the Farm class with 10m resolution. On the top of each image patch, the class which the image patch belongs to and its spatial resolution are shown on the first row. And the number of the image patch in the database is shown on the second row.

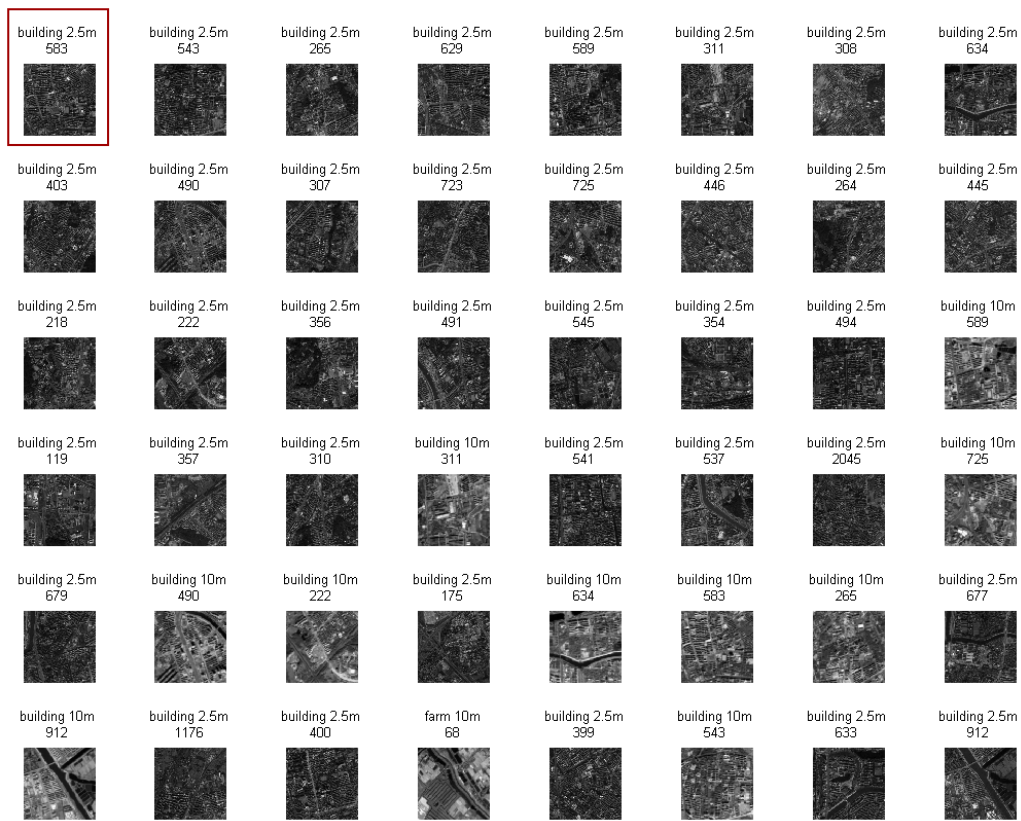


Figure 4: Retrieval results for an image patch of the Building class.

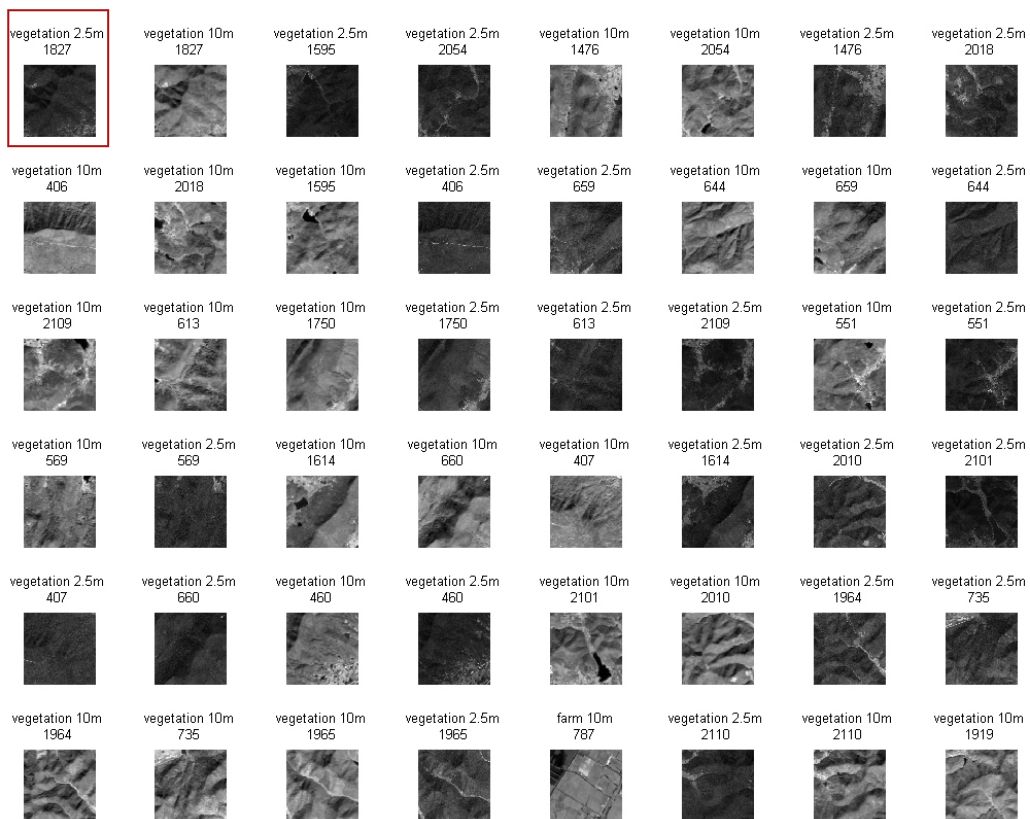


Figure 5: Retrieval results for an image patch of the Vegetation class.