ISPRS Journal of Photogrammetry and Remote Sensing 149 (2019) 14–28

Contents lists available at ScienceDirect

ISPRS Journal of Photogrammetry and Remote Sensing

journal homepage: www.elsevier.com/locate/isprsjprs

3D gray level co-occurrence matrix and its application to identifying collapsed buildings

Luis Moya[a,*], Homa Zakeri[b], Fumio Yamazaki[c], Wen Liu[c], Erick Mas[a], Shunichi Koshimura[a]

[a] International Research Institute of Disaster Science, Tohoku University, Aoba 468-1-E301, Aramaki, Aoba-ku, Sendai 980-0845, Japan
[b] Housing System Department, NIKKEN SEKKEI, 1-4-27, Koraku, Bunkyo-ku, Tokyo 112-0004, Japan
[c] Department of Urban Environment System, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263-8522, Japan

ARTICLE INFO

Keywords:
Three-dimensional GLCM
Texture
Change detection

ABSTRACT

With the remarkable progress in access to remote sensing imagery data, nowadays research very often utilizes more than one image. We are often able to use multitemporal, hyperspectral, and/or full polarization of microwave radar images. In addition, it has become the general consensus that texture analysis plays an important role in remote sensing. It has been found in several publications that texture analysis was applied to each layer separately; however, this procedure requires a significant amount of computation and produces a massive volume of data. One alternative, and perhaps a better procedure, is to arrange the images into a multi-layered structure and perform texture analysis within some sort of three-dimensional domain. This manuscript extends the concepts of the gray level of co-occurrence matrix (GLCM) texture analysis applied for a single image to a multi-layered set of images, referred to in this paper as 3DGLCM. We then presented an interpretation of the 3DGLCM within the context of building damage identification. A set of 3DGLCM-based features were computed and evaluated as well. As a result, it was observed that some texture features have certain similarities with other methods proposed in previous studies, whereas other features have not been used before. Furthermore, this paper evaluates the performance of the Support Vector Machine (SVM) classifier in learning and detecting collapsed buildings using 3DGLCM-based features. Thus, the empirical evaluation focuses on the identification of collapsed buildings caused by the 2011 Tohoku earthquake and tsunami, where individual polarized TerraSAR-X intensity images are used to compute the texture features, and the collapsed buildings caused by the 2016 Kumamoto earthquake, where LIDAR-based digital surface models are used to compute the texture features. Extensive datasets consisting of building damage states that have been visually inspected by local authorities and research teams are used to set up the training and testing subsets. Furthermore, the proposed texture features are compared with features commonly used to identify collapsed buildings. The study concludes that an SVM trained with 3DGLCM-based features identifies collapsed buildings with high accuracy and outperforms an SVM trained with common features used in previous studies.

1. Introduction

Texture features describe the spatial contextual information of pixel values within a defined area. This information is widely used in image analysis because it has been proved that texture information increases the performance of classification algorithms. The applications of texture in fields such as medicine, agriculture, disaster assessment and geology are well reported. Although the first publications on texture data date several decades ago (Haralick et al., 1973), their applications, modifications and further improvements prevail (Hall-Beyer, 2017b; Cavalin and Oliveira, 2017; Antel et al., 2003; Boulkenafet et al., 2016). There are numerous approaches to retrieving texture information such as gray-level co-occurrence matrices (GLCM), local binary patterns, convolutional neural networks, multi-scale patch-based recognition, wavelets, and pseudo cross variograms.

The use of texture in remotely sensed data has been reported in previous studies. For example, Soares et al. (1997) studied texture features to characterize agricultural land use. The spatial distribution of CO_2 emission rates was examined using the texture-based classification of Landsat-TM images in Soegaard and Møller-Jensen (2003). Brenning et al. (2012) evaluated the use of texture filters applied to IKONOS orthoimages in detecting rock glacier flow structures. Culbert et al. (2012) explored the use of texture measures to estimate the variability in avian species richness. Zakeri et al. (2017) used texture measures for

* Corresponding author.
E-mail address: lmoyah@irides.tohoku.ac.jp (L. Moya).

https://doi.org/10.1016/j.isprsjprs.2019.01.008
Received 25 May 2018; Received in revised form 10 January 2019; Accepted 11 January 2019
Available online 19 January 2019

land cover classification using polarimetric synthetic aperture radar (SAR) images. Hall-Beyer (2017a) presented guidelines to choose GLCM textures for landscape classification. Sun et al. (2016) evaluated the application of texture features in the characterization of collapsed buildings using a post-event SAR image.

Unlike the work of Sun et al. (2016), most studies on the detection of buildings damaged due to large-scale disaster use two or more images. It has come to our attention that those studies use spatial information as well. Although the techniques employed have not been categorized as texture-based methods, they are based on a similar fundamental basis: contextual variations in pixel values in time. For instance, the variations in the averages of pixels among a pair of images have been demonstrated useful to detect building damage. Another common features are the correlation coefficient and coherence between the pixels of two images. These features can be computed using a square moving window or an object-based approach, that is, predefined regions recognized to be homogeneous. Using a pair of images, recorded before and after a disaster, such features have been extensively used to characterize damage to buildings (Chini et al., 2009; Matsuoka and Nojima, 2010; Liu et al., 2013; Uprety et al., 2013; Gokon et al., 2016; Nakmuenwai et al., 2016; Miura et al., 2016; Wieland et al., 2016; Watanabe et al., 2016; Liu and Yamazaki, 2017; Karimzadeh and Mastuoka, 2017; Moya et al., 2018b,a; Anniballe et al., 2018; Ferrentino et al., 2018a,b). In general, the characterization of damage to buildings is often based on changes detected between a pair of images. Thus, change detection might be interpreted as the use of spatial variations in pixels within a three-dimensional domain. The generalization of texture definition to multi-layered images may perhaps provide new insights or at least offer a different perspective of features that have been already applied.

The concept of a three-dimensional (3D) texture has been used for multi-layered and multi-temporal images in previous studies, where the third dimension represents a spectral coordinate or the dates in which the same area was recorded. Most 3D-texture approaches resemble an extension of texture methods applied to a single image, that is, a bi-dimensional domain. Jin et al. (2012) proposed the application of pseudo cross variogram (PCV) between images recorded at different times and applied it to land cover classification. The 3D Fourier transform is another approach to characterizing 3D textures. Such a method was proposed to create synthetic images with a desired texture in Sarkar and Healey (2010). Similarly, Qian et al. (2013) used wavelet coefficients from a 3D discrete wavelet transform to simultaneously characterize the spatial and spectral structure of hyperspectral images. The GLCM method, one of the most popular texture methods, has been extended to a three-dimensional space as well. Lazaroff and Brennan (1993) constructed the GLCM from the spatio-temporal co-occurrence of digital numbers from Landsat TM data and used it to evaluate changes in forest canopies. Gautama and Heene (1998) used a similar approach with ERS SAR images and performed land cover classification. Unfortunately, these remarkable and pioneering applications of GLCM in a three-dimensional domain, hereafter referred to as 3DGLCM, did not greatly impact subsequent research on satellite remote sensing data. Nevertheless, 3DGLCM has been widely used in other fields such as biomedical imaging (Depeursinge et al., 2014). A recent application of 3DGLCM in hyperspectral data can be found in Tsai et al. (2007) and Tsai and Lai (2013). Regarding multitemporal remote sensing data, typically, the application of GLCM is performed separately on each individual image, and then, the variations in GLCM-based features are used for classification (Bignami et al., 2011; Anniballe et al., 2018). The GLCM is used to compute a number of features, and if GLCM is calculated over each input layer separately, a tremendous volume of data will be produced; see Fig. 1a for the case of only two images. This unexpected pitfall makes additional processing, such as principal component analysis, necessary to reduce the amount of data (Zakeri et al., 2017; Hall-Beyer, 2017a). In contrast, 3DGLCM can reduce the number of output layers (Fig. 1b).

In this paper, we explore the potential of 3DGLCM for identifying collapsed buildings. A comprehensive interpretation of 3DGLCM-based texture features in the context of building damage detection is presented. The remainder of this paper is structured as follows. Section 2 introduces the fundamental basis of 3DGLCM. Since the early stages in this study was influenced only by the idea of extending the standard GLCM to multi-layered images, we have mainly followed the original nomenclature introduced in Haralick et al. (1973). Thus, our work presents some differences with the works of Lazaroff and Brennan (1993) and Gautama and Heene (1998). In Section 3, the performance of collapsed building classification using 3DGLCM-based features is evaluated. Two different disaster events with different remote sensed data are used as cases studies here. Finally, the conclusions are drawn in Section 4.

2. Three-dimensional Gray Level Co-occurrence Matrix (3DGLCM)

2.1. Definition

Suppose that a set of N_z images (layers) to be analyzed are rectangular, with N_x pixels in the horizontal direction and N_y pixels in the vertical direction. In addition, the digital value of each pixel is quantized to N_g gray levels. Let $L_x = \{0, 1, ..., N_x - 1\}$, $L_y = \{0, 1, ..., N_y - 1\}$ and $L_z = \{0, 1, ..., N_z - 1\}$ be the X, Y and Z domains. Recall that the Z domain might represent dates in multi-temporal image analysis, polarization for SAR images, or spectral coordinates in multispectral/hyperspectral images. Furthermore, let $G = \{0, 1, ..., N_g - 1\}$ be the set of N_g gray levels. The set $L_x \times L_y \times L_z$ is the set of pixels of the imagery ordered by their column-row-layer designations. The imagery I can be represented as a function that assigns some digital value in G to each pixel in $L_x \times L_y \times L_z$;$I$: $L_x \times L_y \times L_z \rightarrow G$.

In the original definition of co-occurrences in Haralick et al. (1973), a pair of pixels were identified by the relative position of one pixel, here referred as *neighbor pixel*, with respect to the other, hereafter referred as *reference pixel*. The relative distance was defined by a scalar d and an orientation angle α. In Fig. 2, the reference pixels are colored black, and the neighbor pixels are colored gray. Fig. 2a shows an instance of a pair of pixels with $d = 1$ and $\alpha = 0$, and Fig. 2b shows all possible neighboring pixels ($\alpha = 0, 45, 90, 135$). In this paper, the relative distance is defined by a vector $\boldsymbol{d} = [d_1, d_2, d_3]$, in which the vector is pointing from the reference toward the neighbor pixel. Fig. 2c depicts a pair of pixels separated by $\boldsymbol{d} = [0, 0, 1]$, and Fig. 2d shows all the closest neighbor pixels for a space consisting of two images. The 3DGLCM is specified by the matrix of relative frequencies, $P_{i,j}$, of pairs of pixels separated by a vector $\boldsymbol{d}$ that occur in the $L_x \times L_y \times L_z$ domain, in which the digital value of the reference and neighboring pixels equal i and j, respectively. The non-normalized matrix of frequencies $\boldsymbol{P}$ is defined by

$$P(i, j, \boldsymbol{d}) = \#\{((a, b, c), (d, e, f)) \in (L_x \times L_y \times L_z) \times (L_x \times L_y \times L_z) | [d - a, e - b, f - c] = \boldsymbol{d}, I(a, b, c) = i, I(d, e, f) = j\} \quad (1)$$

where # denotes the number of elements in the set and (a, b, c) and (d, e, f) are coordinates in the three-dimensional space. The definitions presented above represent the general structure of the framework, which present some contrast with the original definition in Haralick et al. (1973). With the introduction of the vector distance $\boldsymbol{d}$ a clear distinction in the direction is defined, that is, the vector is always pointing toward the neighbor pixel. That is not the case in the original definition. The reason behind this decision is clarified in the next subsection. The texture features computed from the 3DGLCM are grouped according to Hall-Beyer (2017b) in the following form. A first group emphasizes the distance of each element $P_{i,j}$ from the 3DGLCM diagonal (i.e., $|i - j|$). *Contrast*, *dissimilarity*, and *homogeneity* belong to this group:

$$\text{Contrast} = \sum_{i,j=0}^{N-1} p_{i,j}(i - j)^2 \quad (2)$$

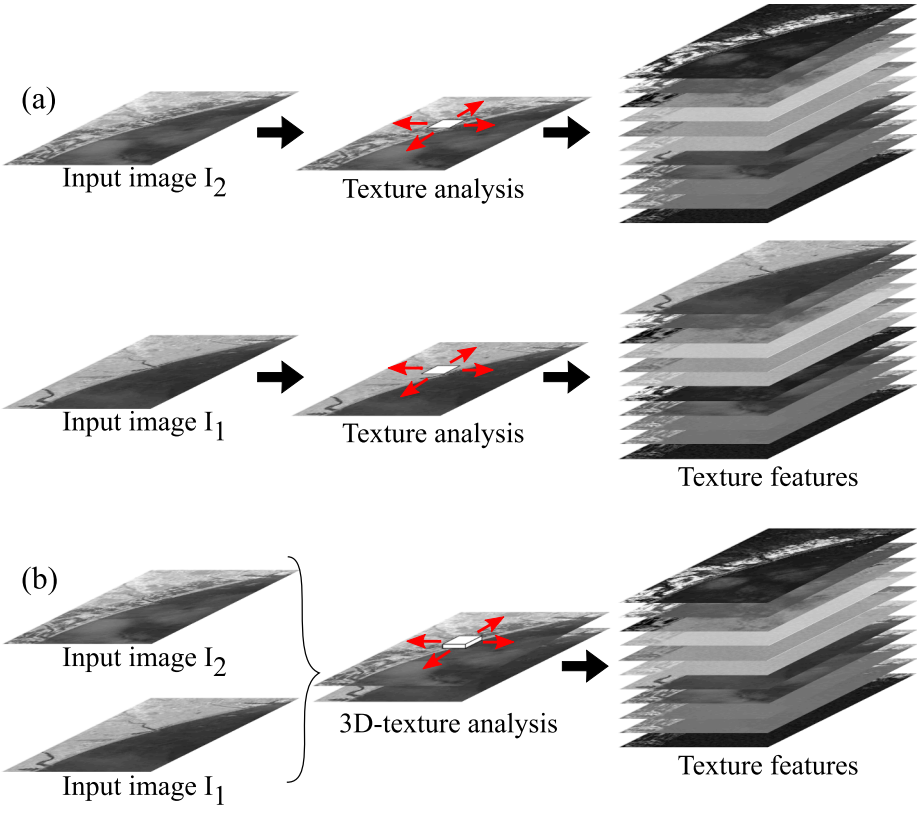

Fig. 1. Scheme of the texture images computed on each image separately (a) and that computed together within a three-dimensional domain (b).

$$\text{Dissimilarity} = \sum_{i,j=0}^{N-1} p_{i,j}|i-j| \tag{3}$$

$$\text{Homogeneity} = \sum_{i,j=0}^{N-1} \frac{p_{i,j}}{1+(i-j)^2} \tag{4}$$

A second group of features emphasize the orderliness. That is, how regular an arbitrary pair of neighbor pixels occurs. Among them, the *Angular second moment* (ASM), *energy* and *entropy* are found:

$$\text{ASM} = \sum_{i,j=0}^{N-1} p_{i,j}^2 \tag{5}$$

$$\text{Energy} = \sqrt{\text{ASM}} \tag{6}$$

$$\text{Entropy} = \sum_{i,j=0}^{N-1} p_{i,j}(-\ln p_{i,j}) \tag{7}$$

A third group represents the descriptive statistics of the GLCM and it includes the *mean reference* (μ_i), *mean neighbor* (μ_j), *standard deviation reference* (σ_i), *standard deviation neighbor* (σ_j) and *correlation* (r):

$$\mu_i = \sum_{i,j=0}^{N-1} i(p_{i,j}) \tag{8}$$

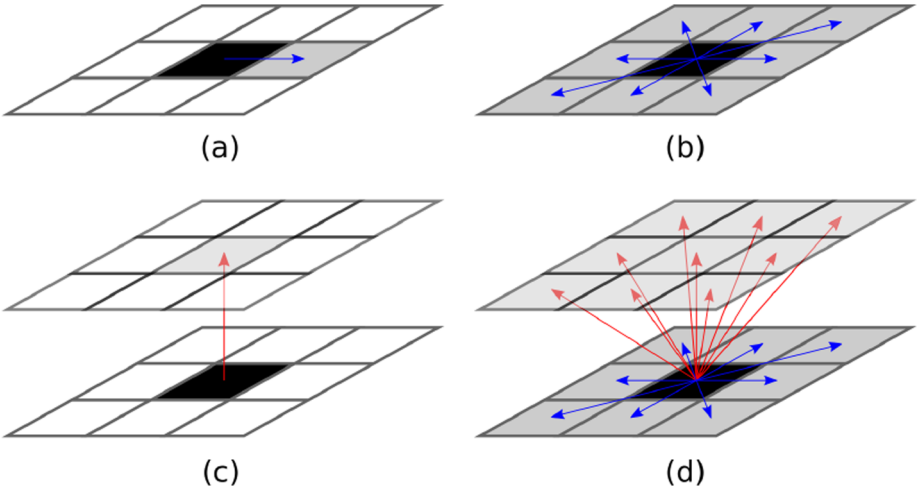

Fig. 2. Examples of spatial relations between reference (black) and neighboring (gray) pixels; (a) pair of pixels defined by a distance $d = 1$ and an angle $\theta = 0$; (b) all neighboring pixels at a distance $d = 1$ in a bi-dimensional domain; (c) pair of pixels within a three-dimensional domain defined by the vector distance $\boldsymbol{d} = (0, 0, 1)$; (d) all neighboring pixels in a three-dimensional domain.

$$\mu_j = \sum_{i,j=0}^{N-1} j(p_{i,j}) \tag{9}$$

$$\sigma_i^2 = \sum_{i,j=0}^{N-1} p_{i,j}(i-\mu_i)^2 \tag{10}$$

$$\sigma_j^2 = \sum_{i,j=0}^{N-1} p_{i,j}(j-\mu_j)^2 \tag{11}$$

$$r = \sum_{i,j=0}^{N-1} p_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2\sigma_j^2}}\right] \tag{12}$$

where $p_{i,j}$ denotes the normalized GLCM, that is $P_{i,j}/N$, with N denoting the number of neighboring pixel pairs.

2.2. Technical interpretation

In this subsection, we focus on one of the simplest but still practical case of Eq. (1), that is, the construction of $\boldsymbol{P}$ from a pair of images using vector $\boldsymbol{d} = [0, 0, 1]$ (Fig. 2c). It is our intention to make an interpretation of the 3DGLCM-based textures and evaluate its potential for detecting damage-induced changes in the aftermath of a large-scale disaster. Under this configuration, one image will represent a pre-event image, and the other will represent a post-event image. We used the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ depicted in Fig. 3a to set up two cases. Let us assume that the matrix $\boldsymbol{A}$ denotes a region with a building at the center. It might represent a digital surface model in certain units; however, the interpretation is valid for data recorded from any other type of sensor as well. In the first case, it is assumed that the building did not experience any damage during the disaster, and therefore, it should present the same values in both the pre- and post-event image. This setting is shown in Fig. 3b. In contrast, in the second case, the matrix $\boldsymbol{B}$ shows a possible situation of the same region if the building would have collapsed. This case is shown in Fig. 3c. Notice that when using the vector $\boldsymbol{d} = [0, 0, 1]$, the reference pixels will be located in the pre-event image, and the neighboring pixels will be in the post-event image. The pre-event and post-event images together constitute a three-dimensional domain whereby

$$\begin{aligned} \boldsymbol{I}^{pre} &= \{(a, b, c) \in (L_x \times L_y \times L_z) | c = 0\} \\ \boldsymbol{I}^{post} &= \{(d, e, f) \in (L_x \times L_y \times L_z) | f = 1\} \end{aligned} \tag{13}$$

The un-normalized 3DGLCMs constructed from these two cases are

$$\boldsymbol{P}^{AA} = \begin{pmatrix} 16 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 9 \end{pmatrix}$$

$$\boldsymbol{P}^{AB} = \begin{pmatrix} 13 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 4 & 4 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{14}$$

where $\boldsymbol{P}^{AA}$ and $\boldsymbol{P}^{AB}$ are the 3DGLCM constructed from the first and second case, respectively. To clarify the construction process of the matrix $\boldsymbol{P}$, Fig. 3b shows all the vectors $\boldsymbol{d}$ that start in a pixel with value zero in the pre-event image (reference pixel) and point to a pixel with value zero in the post-event image (neighbor pixel). There are 16 such vectors in total, which is the value of $P_{0,0}^{AA}$. In the second case study, shown in Fig. 3c, there are only 13 vectors with this configuration; therefore, $P_{0,0}^{AB} = 13$. Notice that in both cases, the sum of the elements in row i, $\sum_j P_{i,j}$, is the number of pixels in the pre-event image whose digital value is equal to i. Likewise, the sum of elements in column j, $\sum_i P_{i,j}$, is the number of pixels whose digital value is equal to j in the post-event image. Another straightforward observation is that if there are no changes between the pre- and post-event images, as in the first case, the matrix $\boldsymbol{P}$ will be diagonal. When changes occurs, the amount that was concentrated in the diagonal will be distributed along the rows, that is, $P_{i,i}^{AA} = \sum_j P_{i,j}^{AB}$.

Notice that although the two cases only use 6 gray levels, the number of zero elements in both $\boldsymbol{P}^{AA}$ and $\boldsymbol{P}^{AB}$ is quite large. In general, the matrix $\boldsymbol{P}$ is inherently sparse when it is constructed from images whose number of gray levels, N_g, is larger than the number of pixels within the window-size. This issue might represent a problem regarding unnecessary memory requirements and operations during the implementation of the method. A more efficient representation of the gray level of co-occurrences is a list in which only significant numbers are stored. For instance, the list representation for the second case is

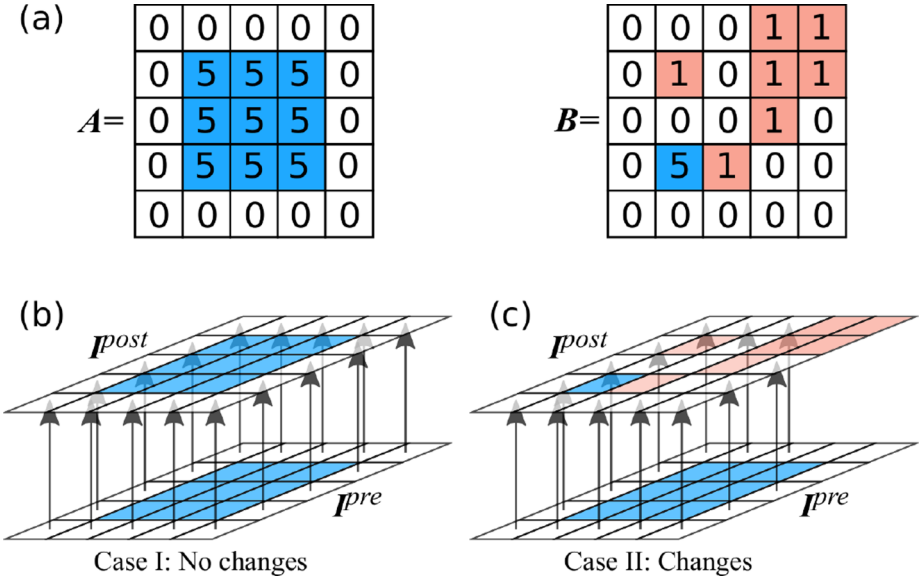

Fig. 3. (a) Matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ used to set up two simple cases of three-dimensional domains composed of 2 layers I^{pre} and I^{post}; (b) Case I: No changes occurred; then, $\boldsymbol{I}^{pre} = \boldsymbol{I}^{post} = \boldsymbol{A}$; (c) Case II: Changes occurred; then, $\boldsymbol{I}^{pre} = \boldsymbol{A}$ and $\boldsymbol{I}^{post} = \boldsymbol{B}$.

Table 1
Texture features computed from the 3DGLCM constructed from cases I and II shown in Fig. 3b and c, respectively.

Group	Feature	Case I No changes	Case II Changes
Contrast	Contrast	0.00	6.68
	Dissimilarity	0.00	1.56
	Homogeneity	1.00	0.58
Orderliness	ASM	0.54	0.34
	Energy	0.74	0.58
	Entropy	0.65	1.31
Statistics	μ_i	1.80	1.80
	μ_j	1.80	0.48
	σ_i	5.76	5.76
	σ_j	5.76	1.05
	r	1.00	0.38

$$\boldsymbol{P}^{list} = \begin{pmatrix} 0 & 0 & 13 \\ 0 & 1 & 3 \\ 5 & 0 & 4 \\ 5 & 1 & 4 \\ 5 & 5 & 1 \end{pmatrix} \tag{15}$$

Here, the first column denotes the row index i of $\boldsymbol{P}$, the second column denotes the column index j of $\boldsymbol{P}$, and the third column is $P_{i,j}$. Eqs. (2)–(12) can be redefined using $\boldsymbol{P}^{list}$. For example, Eq. (2) is expressed as

$$Contrast = \sum_{k=0}^{M} p_{k,2}^{list}(p_{k,0}^{list} - p_{k,1}^{list})^2 \tag{16}$$

where M denotes the number of rows in $\boldsymbol{P}^{list}$. An additional advantage of P^{list} is that the order of the rows is irrelevant; specifically, the rows can be interchanged without affecting the feature values. This means that there is no need to sort the pixels according to their digital value.

The features computed from both cases (Fig. 3b and c) using Eqs. (2)–(12) are shown in Table 1. As mentioned previously, texture features can be grouped by their weights or by their degree. Weights refer to the factors multiplying $P_{i,j}$ in Eqs. (2)–(12). For the contrast group (Eqs. (2)–(4)), the weights have a geometrical interpretation. They reflect the position of the element $P_{i,j}$ relative to the element of the diagonal: $P_{i,i}$. This group has clear interpretation for the case of change detection. The weights are the changes in intensity between the pre- and post-event images. Going back to the definition of $\boldsymbol{P}$ in Eq. (1), Eqs. (2)–(4) can be rearranged and expressed in a more comprehensive form:

$$\begin{aligned} Contrast &= \frac{1}{N}\sum_{i,j=0}^{N}(I_{i,j}^{pre} - I_{i,j}^{post})^2 \\ Dissimilarity &= \frac{1}{N}\sum_{i,j=0}^{N}|I_{i,j}^{pre} - I_{i,j}^{post}| \\ Homogeneity &= \frac{1}{N}\sum_{i,j=0}^{N}\frac{1}{1+(I_{i,j}^{pre} - I_{i,j}^{post})^2} \end{aligned} \tag{17}$$

Thus, contrast is actually the average of the squared differences between elements of the pre- and post-event images. Dissimilarity is the average of the absolute value of the differences between elements of the pre- and post-event images. Similarly, Homogeneity is the average of the inverse of the squared differences between elements of the pre- and post-event images, where a value of 1 is added in the denominator to avoid division by zero. Observe that contrast and dissimilarity are zero when there are no changes; otherwise, they are positive values. In addition, homogeneity is equal to 1 when there are no changes and less than 1 when changes occur. A rather similar feature extensively used for damage detection is the difference in the average between the two images (Wieland et al., 2016; Liu and Yamazaki, 2017; Moya et al., 2018b; Matsuoka and Yamazaki, 2004), whose computation for the second case is

$$\bar{\boldsymbol{B}} - \bar{\boldsymbol{A}} = \frac{1}{25}\sum_{i,j=0}^{4}(B_{i,j} - A_{i,j}) = -1.32 \tag{18}$$

The difference in the average has the same exponential degree as the dissimilarity; however, dissimilarity, computed in the second case, is greater in magnitude. When the difference in the average is computed, positive and negative changes can cancel. In contrast, changes are always accumulating in the computation of the dissimilarity. It seems that dissimilarity emphasizes changes better than the difference in averages, and thus, it might perform better in damage characterization.

The second group of texture features, referred as *orderliness* in Hall-Beyer (2017b), includes the angular second moment (ASM), energy and entropy. As mentioned before, if there are no changes, then the matrix $\boldsymbol{P}$ is diagonal. However, when changes occur, the number in the diagonal is distributed horizontally. For example, $P_{0,0}^{AA} = 16$ is distributed to $P_{0,0}^{AB} = 13$ and $P_{0,1}^{AB} = 3$. Likewise, $P_{5,5}^{AA} = 9$ is distributed to $P_{5,0}^{AB} = 4$, $P_{5,1}^{AB} = 4$ and $P_{5,5}^{AB} = 1$. It is well known that from a set of positive numbers $\{x_i\}$, the sum of squared elements is less than or equal to the square of the sum $\sum_i x_i^2 \leqslant (\sum_i x_i)^2$. Therefore, from Eq. (5), the ASM is expected to be larger when there are no changes than that computed when changes occur, as confirmed in the results of the first and second cases. The same trend should be observed for Energy and the opposite for Entropy. However, in theory, there are certain change patterns that do not affect the quantities in ASM, energy, and entropy. Consider the elements in the diagonal of $\boldsymbol{P}^{AA}$ moving horizontally until a certain column j, that is, $P_{i,i}^{AA} = P_{i,j}^{AB}$. This occurs when all pixels with the same digital number in the pre-event image increase or decrease the same amount in the post-event image. Given such changes, the orderliness-type features will have the same magnitude as in the case in which no changes occurred. However, in real practice, such uniform changes do not occur. Changes due to the collapse of buildings have non-uniform patterns. Therefore, for the case of building damage detection, high values of *ASM* and *Energy* and low values of *Entropy* are more likely to be observed for non-damaged buildings.

Regarding to the third group of texture features, descriptive statistics, features from Eqs. (8)–(12) are included. Back once again to the definitions of 3DGLCM with $\boldsymbol{d} = [0, 0, 1]$, μ_i and σ_i represent the average and standard deviation of the pre-event image. Similarly, μ_j and σ_j represent the average and standard deviation of the post-event image. Furthermore, r is the correlation coefficient between both the pre- and post-event matrices. Similar to the difference in averages (Eq. (18)), the correlation coefficient has been extensively used for damage detection (Moya et al., 2018b,a; Gokon et al., 2016; Liu et al., 2013; Liu and Yamazaki, 2017; Yamazaki and Matsuoka, 2007; Matsuoka and Yamazaki, 2004).

Before concluding this section, note that by using a vector $\boldsymbol{d}$ in the construction of the matrix $\boldsymbol{P}$, the direction of the vector makes a clear distinction between a reference and neighboring pixel. If no distinction between the reference and neighboring pixel is desired, as in the original publication of Haralick et al. (1973), $\boldsymbol{P}$ should have been constructed using the vectors $\boldsymbol{d}$ and $-\boldsymbol{d}$. In that case, $\boldsymbol{P}$ will always be symmetric, and most of the interpretation presented here should be redefined. However, under this option, the resemblance to previous studies on building damage detection would be obscure. Eq. (17) would no longer be correct, and the comparison of 3DGLCM-based features with features used in previous studies might not be very clear. Therefore, it was decided to continue using our modification in the following sections.

3. Empirical evaluation

The identification of collapsed buildings in the aftermath of a large-scale disaster is important for disaster management. Hence, the

Table 2
Building damage situation due to the 2011 Tohoku earthquake and tsunami.

Damage state	Description	Number of buildings
DS0	No damage	219,000
DS1	Minor damage	23,000
DS2	Moderate damage	40,000
DS3	Major damage	36,000
DS4	Complete damage	8,000
DS5	Collapsed	34,000
DS6	Washed away	78,000

capabilities of 3DGLCM-based features for collapsed building detection need to be evaluated. Two disaster events were used for this purpose: (i) the earthquake-tsunami of March 11, 2011 that occurred in Tohoku Prefecture, Japan captured by TerraSAR-X images and (ii) the earthquake event that occurred on April 16, 2016 in the prefecture of Kumamoto, Japan, for which LIDAR data were recorded.

3.1. The 2011 Tohoku earthquake and tsunami

The 11 March 2011 Tohoku earthquake, with Mw 9.0, is one of the largest well-recorded earthquakes ever. The maximum record of strong-motion acceleration was 2.7*g*, and the coseismic deformation was 5 m. The earthquake triggered a tsunami that caused extensive damage in the coastal area of Tohoku. The maximum measured tsunami height was 40 m (Mori et al., 2011). The epicenter was 150 km off the coast of Honshu, the largest island of Japan. The Ministry of Land, Infrastructure, Transport and Tourism (MLIT) (2011) conducted field surveys and provided a building damage inventory. Seven levels of damage, from no-damage to washed away, were defined in the mentioned survey. Table 2 shows the number of surveyed buildings separated by their damage level. In this study, we focus on the buildings located within the inundated areas on the coast of Miyagi Prefecture for which TerraSAR-X images were available.

3.1.1. Data used

Fig. 4a shows the location of the study area, one of the zones most severely affected by the induced tsunami. The technical specifications of the satellite images are denoted in Table 3. Fig. 4b and c shows the TerraSAR-X images recorded on 21 October 2010 and 13 March 2011, respectively. Both images were acquired in the StripMap mode with HH polarization on a descending path with 37.3° incident angle. The azimuth and ground range resolutions were approximately 3.3 m. The images were orthorectified, transformed to sigma naught, and speckle noise reduced. Sigma naught represents the radar reflectivity per unit area in the ground range. The enhanced Lee filter method (Lopes et al., 1990) with window size of 3 × 3 was employed for the speckle noise reduction.

3.1.2. Results and discussion

To perform the texture analysis, the images were quantized into 256 levels. Using a moving window, the 3DGLCM was constructed, and the texture features were subsequently computed and allocated in the pixel located at the center of the window. Thus, 11 texture images were constructed. Fig. 5b-l shows the textural images using a moving window of size 13 × 13. The procedure was repeated using window sizes of 5 × 5, 9 × 9, and 17 × 17 as well. For the accuracy assessment, the buildings surveyed by the MLIT located within the study area were used to construct a dataset. The spatial distribution of the surveyed buildings is shown in Fig. 5a. In order to provide further details, the 3DGLCM and texture features computed at the center of individual buildings are reported. Four buildings were selected (Fig. 6), two non-damaged buildings (DS0) and two washed away buildings (DS6). The non-zero elements of the 3DGLCM is depicted in the right column of Fig. 6 and the computed texture features is shown in Table 4. From the dataset constructed from the MLIT's field survey, each sample represents a building and is composed of 11 features and a class label. Each feature is computed as the average of a texture feature located within the building footprint. The class label is assigned according to the MLIT's survey. Fig. 7 shows the boxplot of the samples, according to the class label, computed from the texture images shown in Fig. 5b-l. Each boxplot shows five coordinates representing the upper limit of 5%, 25%, 50%, 75%, and 95% of the samples. The contrast and dissimilarity show an ascending trend according to the damage level, while homogeneity shows a descending trend, which confirm the interpretation presented in the previous section. Regarding the orderliness group (ASM, energy and entropy), there is no distinction between buildings in different damage states, including DS6. Recall that ASM would have a large value if all pixels values remained unchanged or experienced the same change in magnitude, which is effectively impossible because of the presence of speckle noise in SAR images. The speckle effect is the result of the cancellation and amplifications of the electromagnetic waves reflected from the objects within a pixel. Thus, it seems that, when SAR images are used to identify collapsed buildings, ASM, energy and entropy are not suitable features. This issue is clearly observed in the 3DGLCM computed from the samples shown in Fig. 6. For the case of non-damaged buildings (Fig. 6a and b), the non-zero elements of the 3DGLCM should be located at the diagonal. However, because of the speckle noise, several non-zero elements are located off the diagonal. Nevertheless, they are closer to the diagonal than those computed from the washed away buildings (Fig. 6c and d). That is, features from the contrast group performs better to characterize washed away buildings

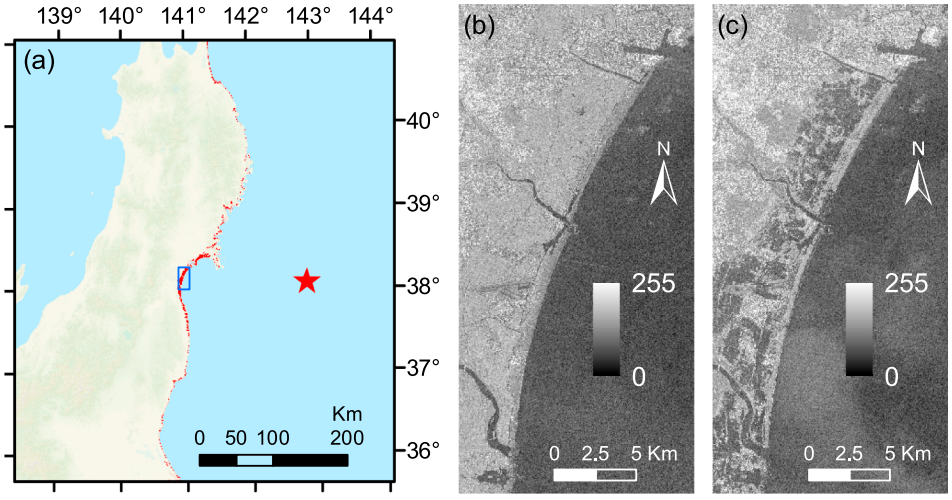

Fig. 4. The 2011 Tohoku earthquake and tsunami, area of interest and results. (a) Location of the study area in northern Japan: the blue rectangle shows the location of the study area, the red areas were inundated by the tsunami, and the red star shows the location of the epicenter; (b) TerraSAR-X image recorded on 21 October 2010; (c) TerraSAR-X image recorded on 13 March 2011. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 3
Specifications of the satellite images used in the first case study, the 2011 Tohoku earthquake and tsunami.

Sensor	Acquisition	Incidence angle	Pixel resolution	Path	Polarization	Band
TerraSAR-X	20/09/010	37.3	1.25	Descending	HH	X
TerraSAR-X	12/03/2011	37.3	1.25	Descending	HH	X

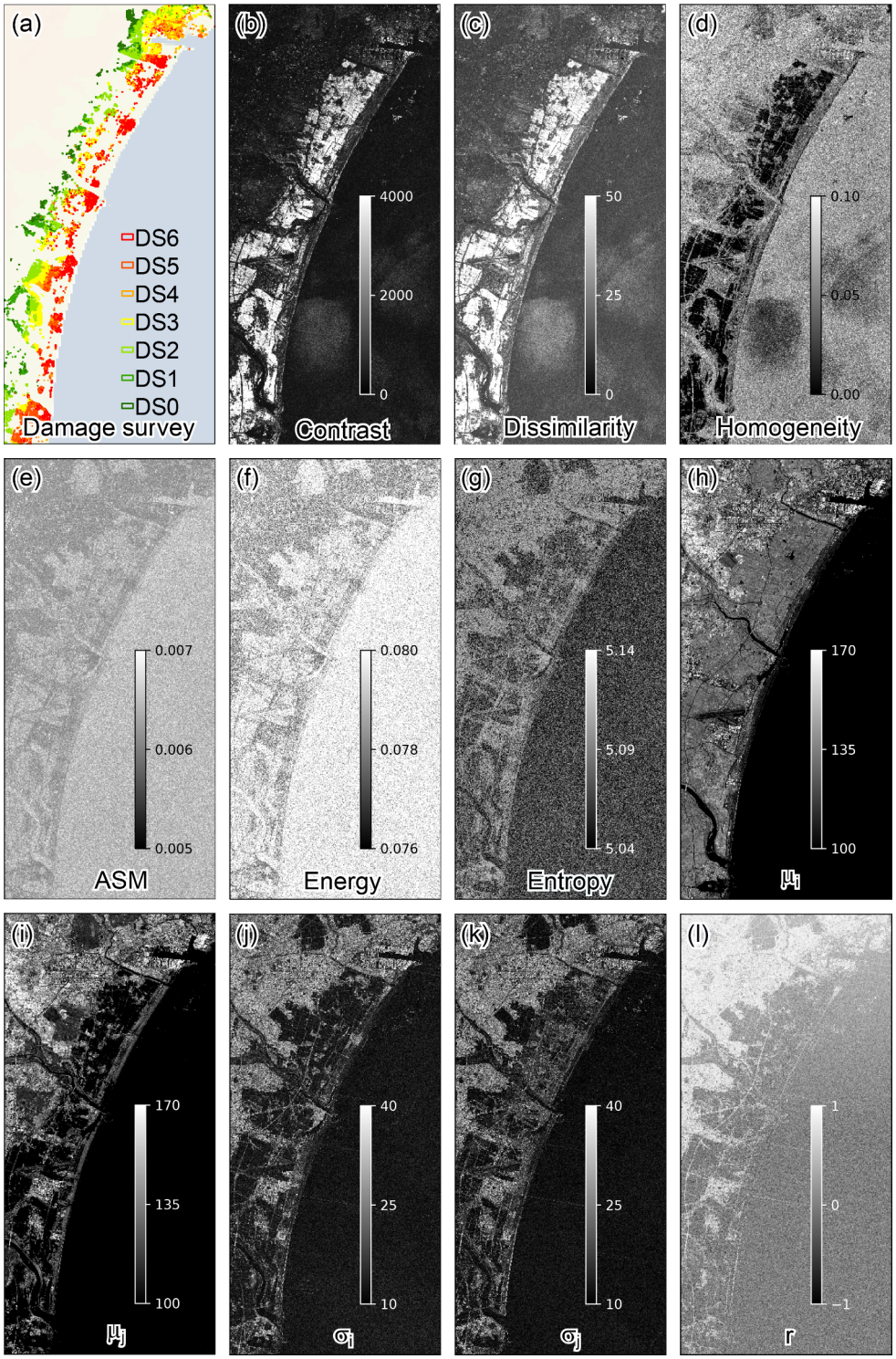

Fig. 5. (a) Buildings surveyed by the MLIT. (b)-(l) 3DGLCM-based texture images obtained using the pair of TerraSAR-X datasets showed in Fig. 4.

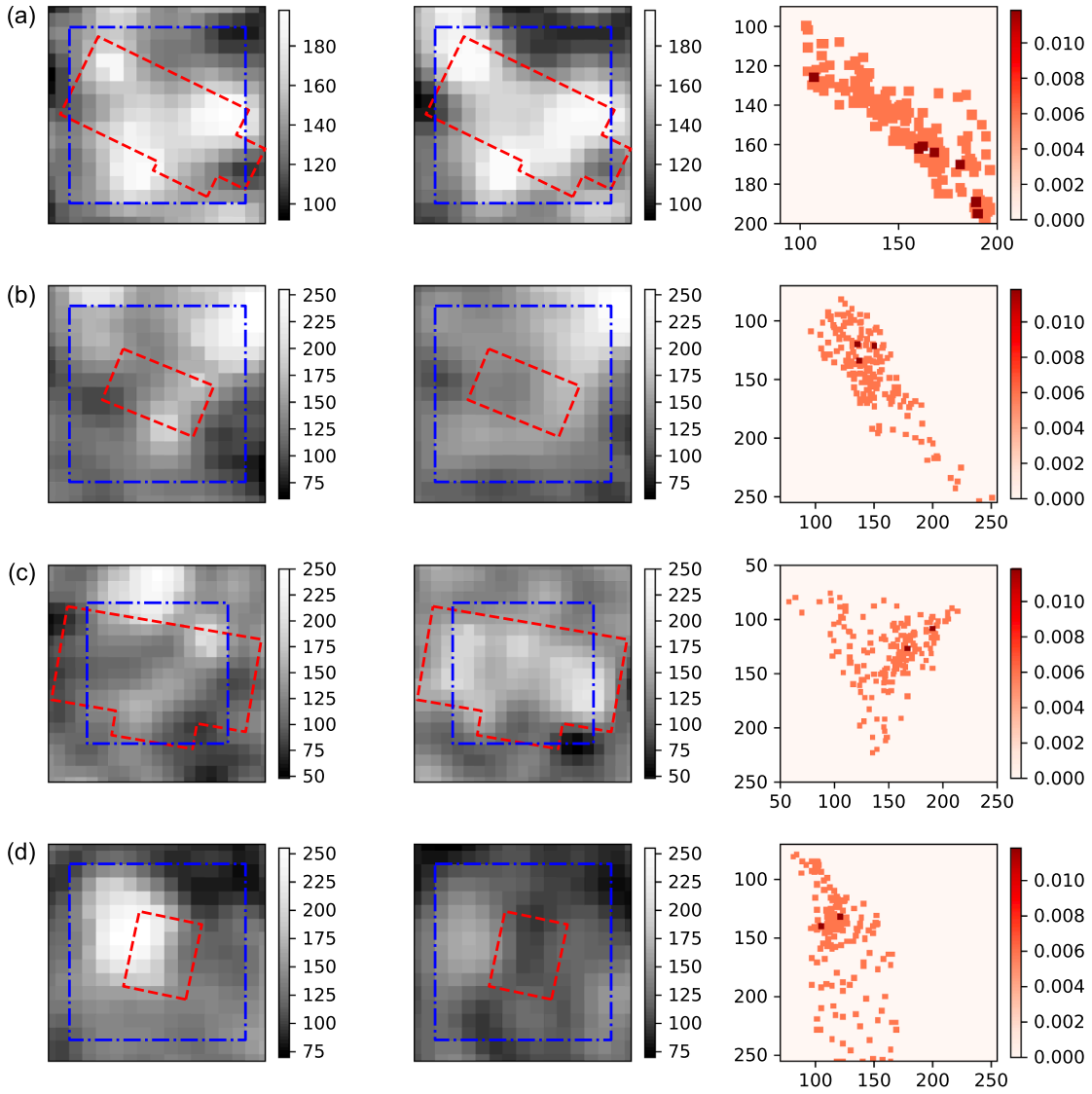

Fig. 6. Illustrations of the construction of the 3DGLCM using a window size of 13 × 13 at the location of individual buildings. (a) and (b) are non-damaged buildings. (c) and (d) are washed away buildings. Images of the pre-event (left) and post-event (middle) are shown in grey levels. The constructed 3DGLCM (right) is shown as an image. The red dashed polygon superimposed in the satellite images represents the building footprint and the blue dash-dotted polygon denotes a window of size 13 × 13 from which the 3DGLCM was constructed (right). Note that for the sake of better visualization, only the region with non-zero elements of the 3DGLCM is shown; furthermore, the size of pixels with non-zero values were increased five times. The texture features computed from the 3DGLCM is reported in Table 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 4
Texture features computed from the 3DGLCMs depicted in Figs. 6 and 10.

Feature	Samples from Fig. 6				Samples from Fig. 10			
	(a)	(b)	(c)	(d)	(a)	(b)	(c)	(d)
Contrast	163.30	422.99	2339.85	2314.00	0.27	0.00	31.15	79.64
Dissimilarity	9.11	16.89	39.70	33.91	0.25	0.00	5.08	8.76
Homogeneity	0.13	0.05	0.02	0.05	0.88	1.00	0.11	0.02
ASM	0.01	0.01	0.01	0.01	0.13	1.00	0.04	0.03
Energy	0.08	0.08	0.08	0.08	0.37	1.00	0.20	0.17
Entropy	5.07	5.11	5.11	5.11	2.32	0.00	3.53	3.75
μ_i	153.46	143.98	130.47	151.16	65.18	79.00	57.76	74.38
μ_j	156.46	147.57	150.80	120.67	65.12	79.00	53.28	65.62
σ_i	23.13	35.35	29.97	43.05	1.59	0.00	3.73	3.79
σ_j	26.44	26.48	30.30	18.98	1.55	0.00	1.87	3.43
r	0.88	0.82	−0.06	0.51	0.95	–	0.46	0.89

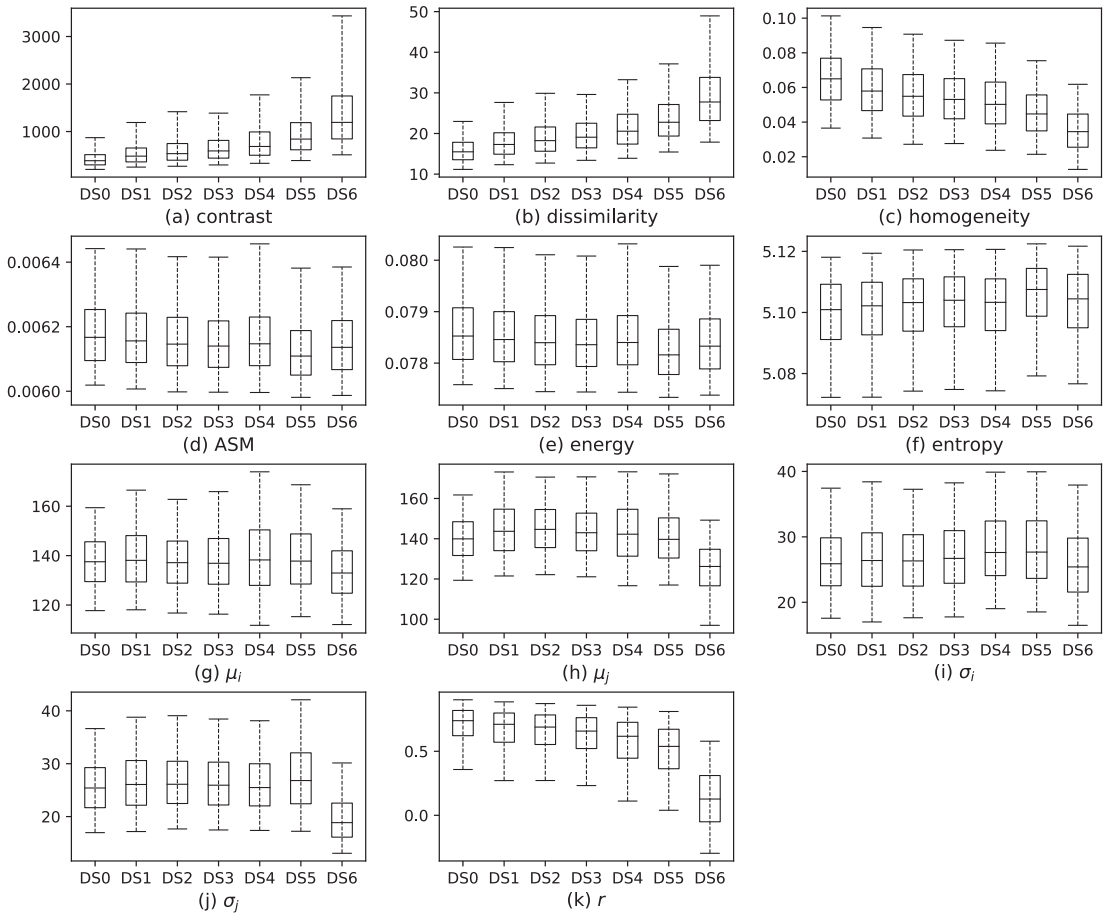

Fig. 7. Box plot of 3D-texture features of buildings damaged due to the 2011 Tohoku earthquake and tsunami. The 3D-GLCM was calculated using the vector $\boldsymbol{d} = [0, 0, 1]$ and a window size of 13×13.

(see Table 4). With respect to the descriptive statistics group, μ_i and σ_i denotes respectively the average and the standard deviation of the pre-event SAR image; and thus, they cannot record any damage pattern. On the other hand, μ_j and σ_j represent the average and standard deviation of the post-event SAR image, which do contain information of collapsed buildings. Furthermore, buildings with low damage states show higher correlation coefficient than those with large damage states (Fig. 7k and Table 4).

A significant overlapping between buildings with damage states of less than or equal to DS5 is observed. This issue has been discussed before in Wieland et al. (2016). One of the main reasons concerns the resolution of the TerraSAR-X images. The damage classification performed by MLIT is according to the damage in the structural system. A structural damage is assumed when, for instance, a lateral story drift greater than about 2–3% is observed (Moehle, 2015; FEMA 356, 2000). That is, a drift of 12–18 cm for a two-story building. Cracks of width greater than 1 mm in structural elements are indicators of damage as well (IAEA, 2002). As can be noted, the pixel resolution (1.25 m) is much larger than these damage indicators. Hence, it is very unlikely that SAR images can detect damage levels lower than collapse.

Fig. 7 provides useful information about the potential of the features for classification purposes. However, it is not conclusive because it does not show the distribution of the samples in the hyper-spatial domain. The boxplots should be interpreted as information of the samples projected onto a single axis. It is, therefore, of interest to investigate whether 3DGLCM-based features can be used for the classification of collapsed buildings. For this purpose, the support vector machine (SVM) method is selected to classify the feature space. The SVM classifies data by constructing a hyperplane to separate them into two classes with a possible maximal margin. The margin defines the distance between the hyperplane and the closest training sample. Further details on the basis of SVM theory can be found elsewhere (Vapnik, 1999). As a baseline reference, the results reported in Wieland et al. (2016) are used for comparison to our results. Among several experiments, Wieland et al. (2016) assessed the trained SVM using the same SAR images as in this case study but with a different feature space. The aforementioned study use three change indexes, the averaged difference, the correlation coefficient and a combination of both features. Additional statistical features, such as the mean, mode, standard deviation, minimum and maximum of the backscatter coefficient, were included as well. Every feature were computed per building footprint. The accuracy assessment procedure is performed as follows: First, buildings with damage states DS0-DS4 were merged and labeled as non-changed samples; similarly, buildings with damage states DS5 and DS6 were merged and labeled as changed samples. From the updated dataset with the binary class label, 3000 samples, 1500 from non-changed and 1500 from changed buildings, were extracted randomly. Then, a 10-fold cross-validation (10-FCV) was used to test the classification performance. In a 10-FCV, the dataset is divided into 10 subsets. For the evaluation, 9 subsets were used to train an SVM and then tested on the remained subset. This procedure is repeated 10 times, in which a different subset is used to test the SVM classifier. Three standard accuracy measures, F1, recall and precision, are reported as the average over the 10 evaluations.

Table 5
Accuracy assessment of the first experiment. SVM-classification of collapsed buildings based on 3DGLCM-based features and that reported in Wieland et al. (2016) (Baseline). Samples with damage levels from DS1 to DS4 were labeled as non-changed, while samples with damage levels DS5 and DS6 were labeled as changed. The 3DGLCM were computed with window sizes of 5, 9 × 9, 13 × 13 and 17 × 17. NC: non-changed samples, C: changed samples, Av: average.

Window size	F1			Recall			Precision		
	NC	C	Av.	NC	C	Av.	NC	C	Av
Baseline	0.75	0.76	0.76	0.74	0.77	0.76	0.76	0.75	0.76
5 × 5	0.80	0.79	0.80	0.83	0.76	0.79	0.77	0.82	0.80
9 × 9	0.81	0.79	0.80	0.86	0.74	0.80	0.77	0.84	0.81
13 × 13	0.83	0.81	0.82	0.87	0.77	0.82	0.79	0.86	0.83
17 × 17	0.84	0.83	0.84	0.87	0.80	0.84	0.81	0.86	0.84

Table 6
Accuracy assessment of the second experiment. SVM-classification of collapsed buildings using 3DGLCM computed with different window sizes. Samples with damage levels from DS1 to DS5 were labeled as non-collapsed, while samples with damage level DS6 were labeled as collapsed. NC: non-collapsed building, C: collapsed building, Av: average.

Window size	F1			Recall			Precision		
	NC	C	Av.	NC	C	Av.	NC	C	Av
5 × 5	0.85	0.85	0.85	0.84	0.87	0.86	0.86	0.84	0.85
9 × 9	0.87	0.87	0.87	0.86	0.87	0.87	0.87	0.87	0.87
13 × 13	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
17 × 17	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89	0.89

Table 7
Accuracy assessment of the third experiment. SVM-classification of collapsed buildings using 3DGLCM computed with different window sizes. Samples with damage levels from DS1 to DS4 were labeled as non-collapsed, while samples with damage level DS6 were labeled as collapsed. Samples with damage level DS5 were not included. NC: non-collapsed building, C: collapsed building, Av: average.

Window size	F1			Recall			Precision		
	NC	C	Av.	NC	C	Av.	NC	C	Av
5 × 5	0.87	0.87	0.87	0.86	0.87	0.87	0.87	0.86	0.87
9 × 9	0.89	0.89	0.89	0.90	0.88	0.89	0.89	0.89	0.89
13 × 13	0.90	0.90	0.90	0.91	0.89	0.90	0.90	0.91	0.91
17 × 17	0.91	0.90	0.91	0.92	0.89	0.91	0.90	0.91	0.91

Table 8
Technical specifications of the remote sensing data used for the second case study, the 2016 Kumamoto earthquake.

Sensor	Instrument	Acquisition	Point density (points m^2)	DSM resolution (m)
LiDAR	Leica ALS50II	15/04/2016	1.5–2	0.5
LiDAR	Leica ALS50II	23/04/2016	3–4	0.5

$$recall = \frac{TP}{Positive}$$
$$precision = \frac{TP}{TP + FP}$$
$$F1 = \frac{2}{1/precision + 1/recall} \quad (19)$$

where *Positive* denotes the number of samples of the class under evaluation, which in this study equals 1500 for both changed and non-changed samples, *TP* denotes *true positive* and indicates the number of samples correctly classified as the class under evaluation. *FP* denotes *false positive* and indicates the number of samples incorrectly classified as the class under evaluation (Fawcett, 2006). The parameters express high (low) accuracy when its values are close to one (zero). The accuracy results are highlighted in Table 5. It is observed that the classification using 3DGLCM-based features outperforms the classification using basic statistics of the backscatter coefficient. Furthermore, it is observed that the accuracy improves when the window size increases, which indicates that spatial contextual information is very effective in SAR images.

Recall that the class labels were assigned according to the MLIT survey, where the damage levels were categorized considering the condition of the main structural system. In MLIT, the damage state DS5 is classified as collapsed, with the description "*Main structure is damaged. It is difficult to reuse it as originally intended*". Thus, it is very likely that a building classified as DS5 was standing right after the tsunami arrived and by the time the post-event SAR image was recorded. Accordingly, the category DS5 contains buildings that collapsed and did not collapse by the time of the disaster. Under these uncertainties, two additional experiments were performed. In the second experiment, samples with DS5 were labeled as non-changed samples. Table 6 summaries the accuracy results, in which an improvement in the scores is observed. In the third experiment, shown in Table 7, samples labeled as DS5 were neglected in the 10-FCV of the SVM. There, an even better performance is observed.

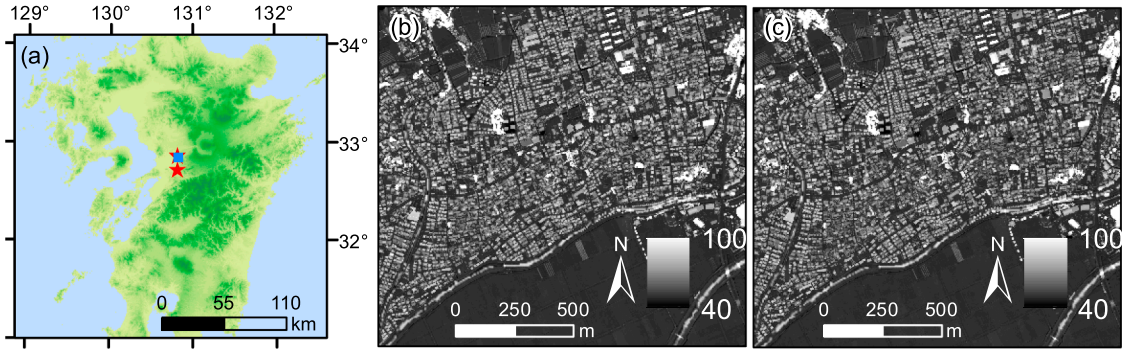

Fig. 8. The 2016 Kumamoto earthquake, area of interest and results. (a) Location of the study area, Mashiki town, on Kyushu island, Japan: The blue rectangle shows the location of the study area, and the red stars show the location of the epicenter of the foreshock (bottom) and mainshock (top). (b) Digital surface model (DSM) recorded on 15 April 2016. (c) DSM recorded on 23 April 2016. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

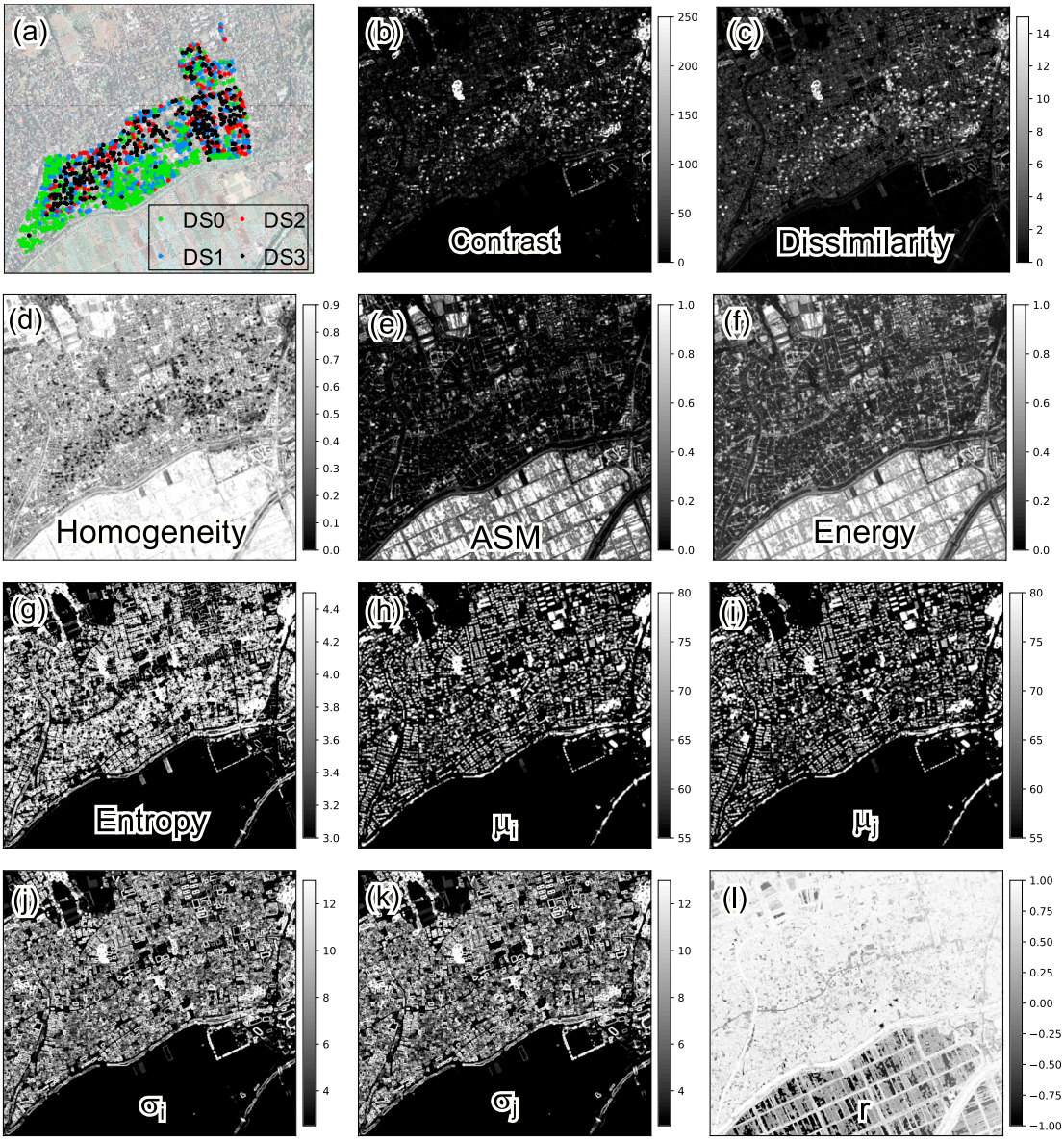

Fig. 9. (a) Building damage inventory performed by Yamada et al. (2017a). DS0: no damage, DS1: partially collapsed, DS2: totally collapsed, DS3: story collapsed. (b)-(l) 3DGLCM-based texture images obtained using the DSMs.

3.2. The 2016 Kumamoto earthquake

On 14 April 2016, an Mw 6.2 earthquake struck Kumamoto Prefecture, Japan. Then, approximately 28 h later, another earthquake with Mw 7.0 occurred. The first event was designated as the foreshock, and the second was the mainshock. In both events, the largest recorded shaking intensity was 7, the highest score in the Japanese Meteorological Agency (JMA) intensity rating system. Both events occurred in the town of Mashiki, with a population of approximately 33,000. Substantial damage to buildings, lifeline systems, and transportation infrastructure was reported. Over 8000 residential buildings were severely damaged or collapsed. The causes of the extended damage were attributed to the fault surface rupture, subsurface soil amplification, and the seismic performance of the buildings (Yamada et al., 2017a).

3.2.1. Data used

Fig. 8a shows the location of the study area on Kyushu Island, the third largest island of Japan located to the southwest. In this case study, digital surface models (DSMs) constructed from LIDAR data are employed. Their technical specifications are shown in Table 8. The first DSM data were recorded on April 15, after the foreshock and before the mainshock, and the second DSM was recorded on April 23, after the mainshock. Both LIDAR datasets were recorded by the Asia Survey (2016). The point density of the first LIDAR dataset was 1.5–2 points per square meter, and that of the second LIDAR dataset was 3–4 points per square meter. The resolution of both DSMs was 50 cm. As preprocessing, a digital elevation model (DEM) provided by the Geospatial Information Authority of Japan (GSI) was used to remove the ground surface elevation from the DSMs. The DSMs, after removing the ground elevation, corresponding to the study area are depicted in Fig. 8b and c.

A joint team from Kyoto University, NEWJEC Inc., and the Building Research Institute conducted a field survey and prepared a building damage inventory (Yamada et al., 2017a,b), as shown in Fig. 9a. In the referred inventory, damaged buildings were classified into four damage states (DS): no damage (DS0), partially collapsed (DS1), totally collapsed (DS2), and story failure (DS3). Again, the concept of collapsed here refers to the state of the structural system. Yamada et al. (2017a) describes DS2 as buildings with serious damage to structural elements such as the structure tilting.

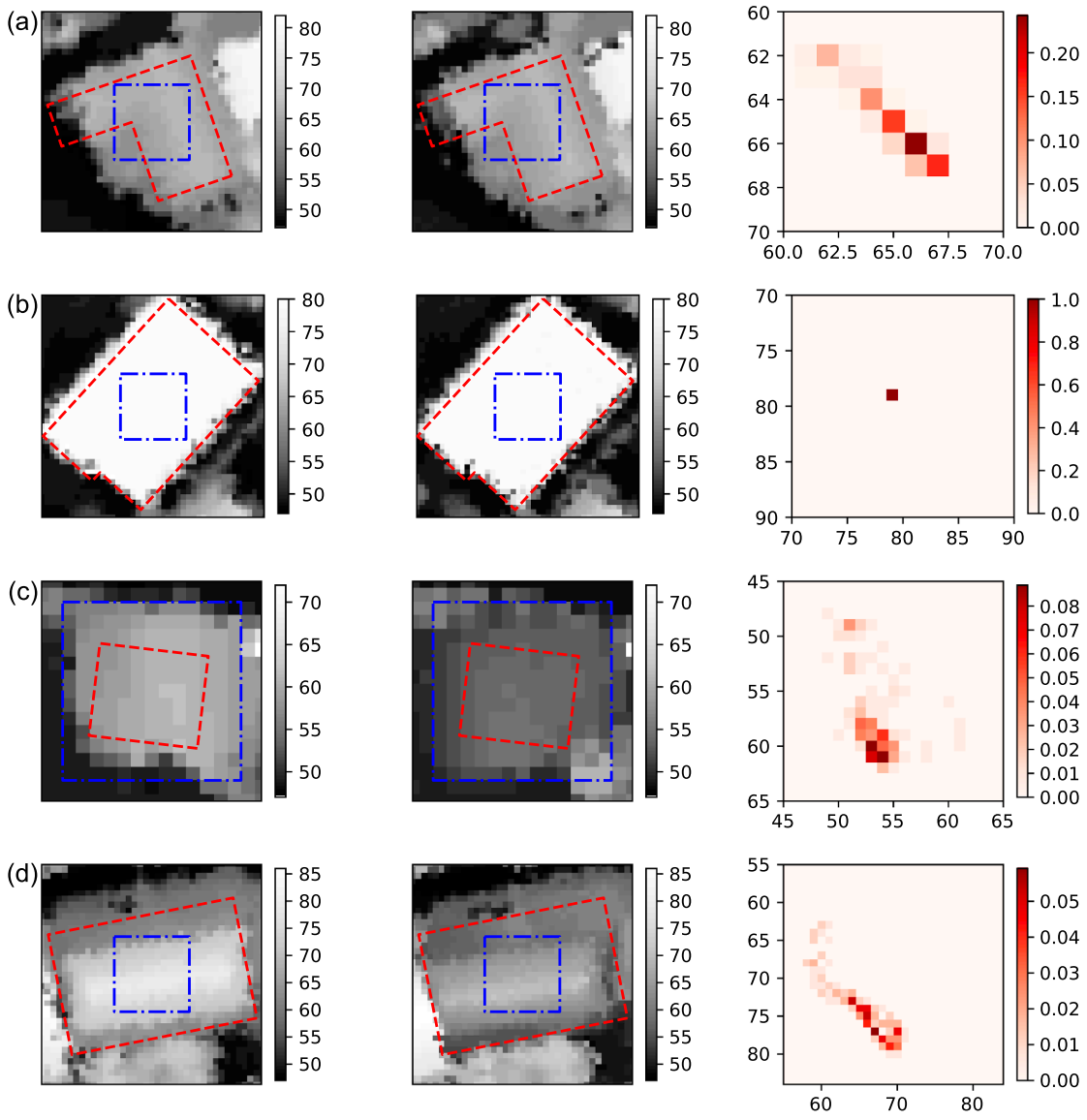

Fig. 10. Illustrations of the construction of the 3DGLCM using a window size of 13 × 13 at the location of individual buildings. (a) and (b) non-collapsed buildings. (c) and (d) collapsed buildings. DSMs of the pre-event (left) and post-event (middle) are shown in grey levels. The constructed 3DGLCM is shown as an image (right). The red dashed polygon superimposed in the DSMs represents the building footprint and the blue dash-dotted polygon denotes the window size from which the 3DGLCM was constructed (right). For the sake of visualization, only the region with non-zero elements of the 3DGLCM is shown. The texture features computed from the 3DGLCMs are reported in Table 4. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

3.2.2. Results and discussion

Following the same procedure as in Section 3.1, the DSMs were quantized into 256 levels. Using a moving window, the 3DGLCM was computed, and using Eqs. (2)–(12), texture features were subsequently calculated. The procedure was repeated for the following window sizes: 5 × 5, 9 × 9, 13 × 13 and 17 × 17. Fig. 9b-l shows the texture images computed using a window size of 13 × 13. For the accuracy assessment, the field survey performed by Yamada et al. (2017a) was employed. Again, each sample in the dataset represents a building and is composed of 11 features and a category label. Each feature is computed as the average value of the pixel value of the texture images located within the building footprint. The category is assigned according the building damage level. Fig. 10 shows four individual buildings, from which the 3DGLCM computed. The texture features were then computed and reported in Table 4. Fig. 11 shows the percentile box for each feature computed using a window size of 13 × 13. As observed previously, features of the contrast-group that belongs to collapsed buildings are distributed in a wider range than those belonging to non-collapsed buildings. On the other hand, unlike the results from the SAR images, features of the orderliness-group provide different patterns between buildings with low and high damage levels. Here, the non-zero elements of the 3DGLCM constructed from non-collapsed buildings (Fig. 10a and b) are mainly located in the diagonal; and in consequence high value of ASM and energy is produced, whereas low value of entropy is produced. Note the effect of the roof type on the 3DGLCM, a gable roof (Fig. 10a) produces many non-zero terms; whereas a flat roof (Fig. 10b) produces only one non-zero element. Regarding the collapsed buildings (Fig. 10c and d), the non-zero elements from the 3DGLCM are mostly off the diagonal. Samples of DS0 and DS1 show that their orderliness-like features are distributed in a wider range than those from samples of the DS2 and DS3 classes. Regarding the group of statistics, $\mu_i \approx \mu_j$ and $\sigma_i \approx \sigma_j$ for samples with DS0, DS1 and DS2, while $\mu_i \neq \mu_j$ and $\sigma_i \neq \sigma_j$ for samples in DS3. Furthermore, r from DS3 samples seems to be lower than the other samples.

Because we decided to perform the same accuracy assessment as shown in the case of the 2011 Tohoku earthquake, the feature-space

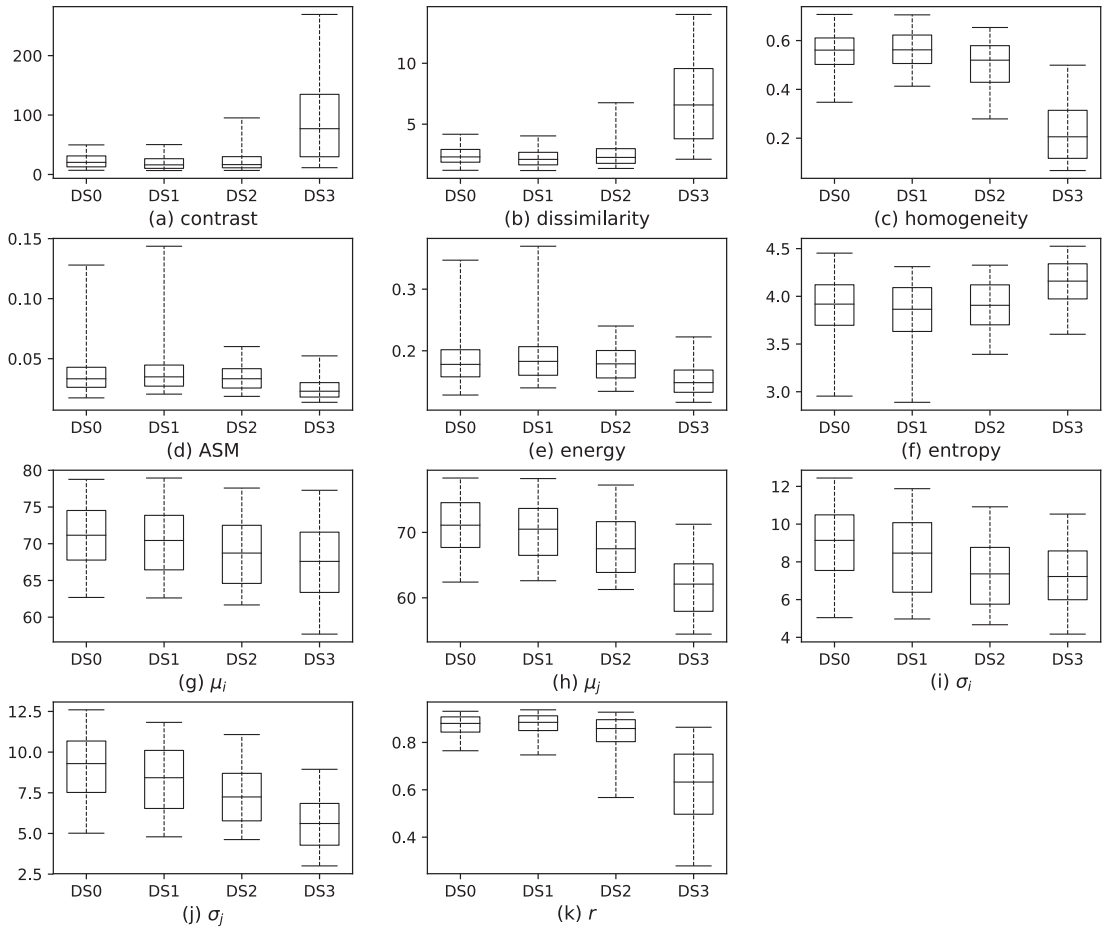

Fig. 11. Box plot of 3D-texture features of buildings damaged due to the 2016 Kumamoto earthquake. The 3D-GLCM was calculated using vector $\boldsymbol{d}$ = [0, 0, 1] and a window size of 13 × 13.

Table 9
Accuracy assessment of SVM trained with 3DGLCM-based features using different window sizes. The accuracy assessment from an SVM trained with the dataset provided in Moya et al. (2018b) is reported as a baseline as well. NC = Non-collapsed; I = inclined; C = collapsed; Av = average.

Window size	F1			Recall			Precision		
	NC	C	Av	NC	C	Av	NC	C	Av
Baseline	0.88	0.86	0.87	0.95	0.80	0.88	0.83	0.94	0.89
5 × 5	0.94	0.94	0.94	0.96	0.92	0.94	0.93	0.96	0.95
9 × 9	0.91	0.91	0.91	0.91	0.90	0.90	0.91	0.91	0.91
13 × 13	0.91	0.91	0.91	0.93	0.89	0.91	0.90	0.93	0.92
17 × 17	0.93	0.93	0.93	0.95	0.92	0.94	0.92	0.94	0.93

constructed in Moya et al. (2018b) was used as baseline rather than the reported accuracy results. The feature space used in Moya et al. (2018b) was constructed as follows: First, the pixels located within the building footprint and whose distance to the building's boundary is greater than 1 m were used to compute three parameters: the averaged differences between both DSMs, the standard deviation of the differences, and the correlation coefficient between both DSMs. In the initial accuracy assessment, samples classified as DS0, DS1, and DS2 were merged and labeled as non-changed samples, while samples classified as DS4 were labeled as changed. Then, the 10-fold cross-validation procedure was performed. As in the first case study, the SVM method was used for the classification. Table 9 shows the F1, recall and precision computed from the baseline dataset and from the 3DGLCM-based features. It is

Table 10
Accuracy assessment of SVM trained with 3DGLCM-based features using different window sizes. The accuracy assessment from an SVM trained with the dataset provided in Moya et al. (2018b) is reported as a baseline as well. NC = Non-collapsed; I = inclined; C = collapsed; Av = average.

Window size	F1				Recall				Precision			
	NC	I	C	Av	NC	I	C	Av	NC	I	C	Av
Baseline	0.62	0.16	0.79	0.52	0.86	0.12	0.76	0.58	0.49	0.41	0.84	0.58
5 × 5	0.67	0.55	0.83	0.68	0.69	0.55	0.82	0.69	0.65	0.56	0.86	0.69
9 × 9	0.70	0.51	0.80	0.67	0.77	0.48	0.78	0.68	0.65	0.55	0.84	0.68
13 × 13	0.60	0.52	0.76	0.63	0.64	0.55	0.74	0.64	0.59	0.54	0.81	0.65
17 × 17	0.65	0.51	0.85	0.67	0.70	0.50	0.83	0.68	0.64	0.55	0.88	0.69

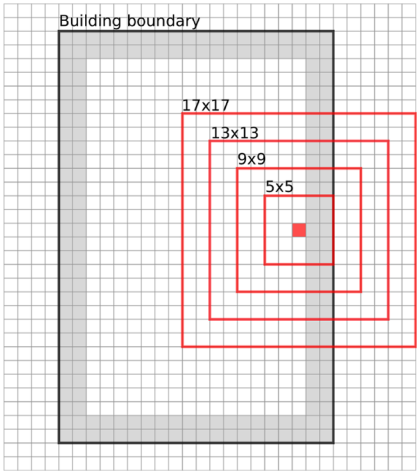

Fig. 12. Display grid used to build up the DSM of a region where a building is located. The thicker black rectangle denotes the building boundary, the gray pixels show the region where the elevation suffers from large errors, the red squares denote the different window sizes used to compute the 3DGLCM, and the red pixel shows the location where the 3DGLCM-based feature is allocated. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

observed that classification using 3D-texture outperforms that obtained using the features proposed in Moya et al. (2018b). From Fig. 11, the contrast group features of samples DS2 seem similar to those from samples DS0 and DS1; on the other hand, the orderliness group features of sample DS2 are similar to those from the DS3 sample. Therefore, an additional experiment with three classes was performed. Here, DS0 and DS1 were merged and labeled as *non-collapsed* samples, DS2 samples were labeled as *inclined*, and DS3 samples were labeled as *collapsed*. Table 10 shows the accuracy measures, where a fairly low accuracy of approximately 0.50 is observed for the samples labeled as inclined. However, it shows a significant improvement compared with the results from the baseline.

Another peculiar observation in Table 9 is that the accuracy decreases when the window size increases from 5×5 to 13×13 but then increases again when the window size is 17×17. It is our belief that the reason for this effect is the same reason pixels close to the building boundary were neglected in Moya et al. (2018b) when the feature space was constructed. As stated previously, the pre-event and post-event DSM were constructed from LIDAR data with different density points. Thus, the elevations constructed every 50 cm are not exactly equal in both DSMs. In particular, this effect produces large errors at the building boundaries, where a certain pixel denotes the roof elevation in one DSM but gives the elevation of the ground surface in the other DSM. To understand the effect of such errors in the 3DGLCM, Fig. 12 illustrates a small area of the grid used to construct the DSMs. Specifically, each square denotes a pixel and contains an averaged elevation. Let us assume that the black rectangle represents the boundary of a building and that the pixels in gray are those that might contain significant errors. The red squares denote the different window sizes used to construct the 3DGLCM, from which the texture features were computed and stored in the same location as the red pixel shown in Fig. 12. Recall that the features for each sample are the average of the texture pixels located within the building footprint. Thus, when a small window is used, such as 5×5, only a few pixels within the building footprint are strongly affected and will not compromise the average. On the other hand, the average value will be affected when the window size increases. However, when the window size is sufficiently large, the ratio of affected pixels within the window size is very small, and thus, the 3DGLCM is not strongly affected. We believe that is why the accuracy increases when a window size of 17×17 is used.

4. Conclusion

In this study, the Gray Level of the Co-Occurrence Matrix, known as GLCM, is constructed using images arranged in a three-dimensional domain in a method referred to as 3DGLCM. Because building damage detection is the major interest of the authors, a comprehensive interpretation of the 3DGLCM applied to non-collapsed and collapsed buildings is performed. Furthermore, the common features computed from GLCM were computed from the 3DGLCM as well. It was found that the group of features referred to as *contrast group features* outperform the average of the differences between a pair of images, a feature extensively used for building damage detection. Moreover, the *orderliness group features* have not been used before in the way used in this paper and it was found they may provide additional information to identify damaged buildings. Empirical evaluations of the capabilities of these new sets of features in identifying collapsed buildings were performed as well. Two cases were studied: the 2011 earthquake and tsunami that occurred in the Tohoku Prefecture, Japan and the 2016 earthquake that occurred in the Kumamoto prefecture, Japan. In the first case, synthetic aperture radar images were used to construct the 3DGLCM, and in the second case, LIDAR-based DSMs were used. Support Vector Machines (SVMs) were trained using the 3DGLCM-based features and compared with results reported elsewhere. The accuracy observed from our results were superior to those reported in previous studies.

Acknowledgements

This research was partly funded by the Japan Science and Technology Agency (JST) through the SICORP project number J150002645 and the CREST project number JP-MJCR1411, and the Japan Society for the Promotion of Science (JSPS) Kakenhi (17H06108). We also acknowledge the Asia Air Survey Co., Ltd., Japan for providing the LiDAR data.

References

Anniballe, R., Noto, F., Scalia, T., Bignami, C., Stramondo, S., Chini, M., Pierdicca, N., 2018. Earthquake damage mapping: an overall assessment of ground surveys and vhr image change detection after L'Aquila 2009 earthquake. Remote Sens. Environ. 210, 166–178.

Antel, S.B., Collins, D., Bernasconi, N., Andermann, F., Shinghal, R., Kearney, R.E., Arnold, D.L., Bernasconi, A., 2003. Automated detection of focal cortical dysplasia lesions using computational models of their mri characteristics and texture analysis. NeuroImage 19 (4), 1748–1759. http://www.sciencedirect.com/science/article/pii/S105381190300226X.

Asia Air Survey Co. Ltd., 2016. 2016 Kumamoto Earthquake. Available at http://www.ajiko.co.jp/article/detail/ID5725UVGCD/ (Online; accessed 02-April-2018).

Bignami, C., Chini, M., Stramondo, S., Emery, W.J., Pierdicca, N., 2011. Objects textural features sensitivity for earthquake damage mapping. In: 2011 Joint Urban Remote Sensing Event, pp. 333–336.

Boulkenafet, Z., Komulainen, J., Hadid, A., 2016. Face spoofing detection using colour texture analysis. IEEE Trans. Inform. Forens. Secur. 11 (8), 1818–1830.

Brenning, A., Long, S., Fieguth, P., 2012. Detecting rock glacier flow structures using Gabor filters and ikonos imagery. Remote Sens. Environ. 125, 227–237.

Cavalin, P., Oliveira, L.S., 2017. A review of texture classification methods and databases. In: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials (SIBGRAPI-T), pp. 1–8.

Chini, M., Pierdicca, N., Emery, W.J., 2009. Exploiting sar and vhr optical images to quantify damage caused by the 2003 bam earthquake. IEEE Trans. Geosci. Remote Sens. 47 (1), 145–152.

Culbert, P.D., Radeloff, V.C., St-Louis, V., Flather, C.H., Rittenhouse, C.D., Albright, T.P., Pidgeon, A.M., 2012. Modeling broad-scale patterns of avian species richness across the midwestern united states with measures of satellite image texture. Remote Sens. Environ. 118, 140–150.

Depeursinge, A., Foncubierta-Rodriguez, A., Van De Ville, D., Mller, H., 2014. Three-dimensional solid texture analysis in biomedical imaging: review and opportunities. Med. Image Anal. 18, 176–196.

Fawcett, T., 2006. An introduction to roc analysis. Pattern Recogn. Lett. 27 (8), 861–874 rOC Analysis in Pattern Recognition.

FEMA 356, 2000. Prestandard and commentary for the seismic rehabilitation of buildings. Federal Emergency Management Agency, Washington, D.C.

Ferrentino, E., Marino, A., Nunziata, F., Migliaccio, M., 2018a. A dual?polarimetric approach to earthquake damage assessment. Int. J. Remote Sens. 0 (0), 1–21.

Ferrentino, E., Nunziata, F., Migliaccio, M., Vicari, A., 2018b. A sensitivity analysis of dual-polarization features to damage due to the 2016 central-Italy earthquake. Int. J.

Remote Sens. 39 (20), 6846–6863.
Gautama, S., Heene, G., 1998. Multitemporal texture analysis using co-occurrence matrices in sar imagery. In: Proceedings of the IASTED International Conference - Signal Image Processing. Las Vegas, Nevada - USA, pp. 403–407.
Gokon, H., Koshimura, S., Matsuoka, M., 2016. Object-based method for estimating tsunami-induced damage using TerraSAR-X data. J. Disaster Res. 11 (2), 225–235.
Hall-Beyer, M., 2017a. GLCM Texture: A Tutorial v. 3.0 March 2017. Available at https://prism.ucalgary.ca/handle/1880/51900 (Online; accessed 14-March-2018).
Hall-Beyer, M., 2017b. Practical guidelines for choosing glcm textures to use in landscape classification tasks over a range of moderate spatial scales. Int. J. Remote Sens. 38 (5), 1312–1338.
Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural features for image classification. IEEE Trans. Syst., Man, Cybernet. SMC-3 (6), 610–621.
IAEA, 2002. Guidebook on non-destructive testing of concrete structures. International Atomic Energy Agency, Vienna.
Jin, H., Li, P., Cheng, T., Song, B., 2012. Land cover classification using chris/proba images and multi-temporal texture. Int. J. Remote Sens. 33 (1), 101–119.
Karimzadeh, S., Mastuoka, M., 2017. Building damage assessment using multisensor dual-polarized synthetic aperture radar data for the 2016 m 6. 2 Amatrice earthquake, Italy. Remote Sens. 9 (4).
Lazaroff, M.B., Brennan, M.W., 1993. Multitemporal texture analysis of features computed from remotely sensed imagery. In: Digital Image Processing and Visual Communications Technologies in the Earth and Atmospheric Sciences II, p. 10 pages.
Liu, W., Yamazaki, F., 2017. Extraction of collapsed buildings in the 2016 kumamoto earthquake using multi-temporal palsar-2 data. J. Disaster Res. 12 (2), 241–250.
Liu, W., Yamazaki, F., Gokon, H., Koshimura, S., 2013. Extraction of tsunami-flooded areas and damaged buildings in the 2011 Tohoku-Oki earthquake from TerraSAR-X intensity images. Earthq. Spectra 29 (S1), S183–S200.
Lopes, A., Touzi, R., Nezry, E., 1990. Adaptive speckle filters and scene heterogeneity. IEEE Trans. Geosci. Remote Sens. 28 (6), 992–1000.
Matsuoka, M., Nojima, N., 2010. Building damage estimation by integration of seismic intensity information and satellite l-band sar imagery. Remote Sens. 2 (9), 2111–2126.
Matsuoka, M., Yamazaki, F., 2004. Use of satellite sar intensity imagery for detecting building areas damaged due to earthquakes. Earthq. Spectra 20 (3), 975–994.
Ministry of Land, Infrastructure, Transport and Tourism (MLIT), 2011. Results of the survey on disaster caused by the great east japan earthquake (first report). http://www.mlit.go.jp/report/press/city07_hh_000053.html (published on 2011-08-04, accessed on 2017-12-22).
Miura, H., Midorikawa, S., Matsuoka, M., 2016. Building damage assessment using high-resolution satellite sar images of the 2010 Haiti earthquake. Earthq. Spectra 32 (1), 591–610.
Moehle, J., 2015. Seismic Design of Reinforced Concrete Buildings. McGraw-Hill Education, USA.
Mori, N., Takahashi, T., Yasuda, T., Yanagisawa, H., 2011. Survey of 2011 Tohoku earthquake tsunami inundation and run up. Geophys. Res. Lett. 38 (7).
Moya, L., Marval Perez, L.R., Mas, E., Adriano, B., Koshimura, S., Yamazaki, F., 2018a. Novel unsupervised classification of collapsed buildings using satellite imagery, hazard scenarios and fragility functions. Remote Sens. 10 (296).
Moya, L., Yamazaki, F., Liu, W., Yamada, M., 2018b. Detection of collapsed buildings from lidar data due to the 2016 Kumamoto earthquake in Japan. Nat. Hazards Earth Syst. Sci. 18 (1), 65–78.
Nakmuenwai, P., Yamazaki, F., Liu, W., 2016. Multi-temporal correlation method for damage assessment of buildings from high-resolution sar images of the 2013 Typhoon Haiyan. J. Disaster Res. 11 (3), 577–592.
Qian, Y., Ye, M., Zhou, J., 2013. Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features. IEEE Trans. Geosci. Remote Sens. 51 (4), 2276–2291.
Sarkar, S., Healey, G., 2010. Hyperspectral texture synthesis using histogram and power spectral density matching. IEEE Trans. Geosci. Remote Sens. 48 (5), 2261–2270.
Soares, J.V., Renn, C.D., Formaggio, A.R., da Costa Freitas Yanasse, C., Frery, A.C., 1997. An investigation of the selection of texture features for crop discrimination using sar imagery. Remote Sens. Environ. 59 (2), 234–247 spaceborne Imaging Radar Mission.
Soegaard, H., Møller-Jensen, L., 2003. Towards a spatial CO2 budget of a metropolitan region based on textural image classification and flux measurements. Remote Sens. Environ. 87 (2), 283–294.
Sun, W., Shi, L., Yang, J., Li, P., 2016. Building collapse assessment in urban areas using texture information from postevent sar data. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. 9 (8), 3792–3808.
Tsai, F., Chang, C.-K., Rau, J.-Y., Lin, T.-H., Liu, G.-R., 2007. 3d computation of gray level co-occurrence in hyperspectral image cubes. In: Yuille, A.L., Zhu, S.-C., Cremers, D., Wang, Y. (Eds.), Energy Minimization Methods in Computer Vision and Pattern Recognition. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 429–440.
Tsai, F., Lai, J.S., 2013. Feature extraction of hyperspectral image cubes using three-dimensional gray-level cooccurrence. IEEE Trans. Geosci. Remote Sens. 51 (6), 3504–3513.
Uprety, P., Yamazaki, F., Dell'Acqua, F., 2013. Damage detection using high-resolution sar imagery in the 2009 L'Aquila, Italy, earthquake. Earthq. Spectra 29 (4), 1521–1535.
Vapnik, V.N., 1999. An overview of statistical learning theory. IEEE Trans. Neural Networks 10 (5), 988–999.
Watanabe, M., Thapa, R.B., Ohsumi, T., Fujiwara, H., Yonezawa, C., Tomii, N., Suzuki, S., 2016. Detection of damaged urban areas using interferometric sar coherence change with Palsar-2. Earth, Planets Space 68 (1), 131.
Wieland, M., Liu, W., Yamazaki, F., 2016. Learning change from synthetic aperture radar images: performance evaluation of a support vector machine to detect earthquake and tsunami-induced changes. Remote Sens. 8 (10).
Yamada, M., Ohmura, J., Goto, H., 2017a. Wooden building damage analysis in Mashiki town for the 2016 Kumamoto earthquakes on April 14 and 16. Earthq. Spectra 33 (4), 1555–1572.
Yamada, M., Yamada, M., Hada, K., Fujino, Y., Mori, J., Sakaue, H., Hayashida, T., Fukatsu, S., Nishihara, E., Ouchi, T., Fujii, A., 2017b. Investigation of building damage in mashikitown for the 2016 kumamoto earthquake-effect of local soil conditions. J. Japan Soc. Civ. Eng., Ser. A1 (Structural Engineering & Earthquake Engineering (SE/EE)) 73 (4), I_216–I_224.
Yamazaki, F., Matsuoka, M., 2007. Remote sensing technologies in post-disaster damage assessment. J. Earthq. Tsunami 1 (3), 193–210.
Zakeri, H., Yamazaki, F., Liu, W., 2017. Texture analysis and land cover classification of Tehran using polarimetric synthetic aperture radar imagery. Appl. Sci. 7 (452).