Automated Detection of Damaged Areas after Hurricane Sandy using Aerial Color Images

Shi Ye[1], Seyed Hossein Hosseini Nourzad[2], Anu Pradhan[3], Ivan Bartoli[4] and Antonios Kontsos[5]

[1] Graduate Student, Department of Civil, Architectural and Environmental Engineering, Drexel University, Philadelphia, PA 19104: PH (215) 350-9732; FAX (215) 895-1363; email: sy394@drexel.edu
[2]Ph.D. Candidate, Department of Civil, Architectural and Environmental Engineering, Drexel University, Philadelphia, PA 19104: PH (215) 429-0700; FAX (215) 895-1363; email: hnourzad@drexel.edu
[3]Assistant Professor, Department of Civil, Architectural and Environmental Engineering, Drexel University, Philadelphia, PA 19104; PH (215) 571-3540; FAX (215) 895-1363; email: arp69@drexel.edu
[4]Assistant Professor, Department of Civil, Architectural and Environmental Engineering, Drexel University, Philadelphia, PA 19104; PH (215) 895-2087; FAX (215) 895-1363; email: ib77@drexel.edu
[5]Assistant Professor, Department of Mechanical Engineering and Mechanics, Drexel University, Philadelphia, PA 19104; PH (215) 895-2297; FAX (215) 895-1478; email: akontsos@coe.drexel.edu

ABSTRACT

Rapid detection of damaged buildings after natural disasters, such as earthquakes and hurricanes, is an urgent need for first response, rescue and recovery planning. In this context, post-event aerial images which could be collected right after disasters are valuable sources for damage detection. However, manual analysis process of the acquired imagery could be both time consuming and costly. To address this issue, a series of classification models for post-hurricane automated detection of damaged buildings is presented in this paper. First, five feature sets were generated through feature extraction and transformation. Then, several classifiers were trained using two groups of classification methods: (1) the Minimum-distance and (2) the Support Vector Machine (SVM) methods. The effectiveness of these classifiers was evaluated in terms of classification accuracies and testing time. The results demonstrated the combination of feature sets and classification methods can provide the best performance. Furthermore, optimal classifiers were selected for future automated real-time damaged building detection. The observed performances of these optimal classifiers indicate promising application for a wide variety of image-based classification tasks.

INTRODUCTION

Hurricane Sandy is the second-costliest hurricane in the United States history, and a total of 24 states in the USA have been affected by this disaster, according to

the National Hurricane Center (Eric S. Blake, 2013). Buildings, roads, vegetation and power systems in several regions were destroyed by Sandy. For rescue and recovery planning, it is urgent and important to rapidly collect damage information for affected areas. The analysis of high-resolution satellite, aerial and Synthetic Aperture Radar (SAR) imagery is a quick way to obtain information related to damaged buildings in large areas. Various automated methods for damage detection after natural disasters have been proposed (Hasegawa, et al., 2000, Mitomi, et al., 2001, Radhika, et al., 2013, Turker and Cetinkaya, 2005). However, the analysis of aerial imagery by the Federal Emergency Management Agency (FEMA) still relies on a group of experts that typically scan hundreds of thousands of aerial images to evaluate actual damage levels (Sagara, 2012). Obviously, the manual process of analyzing disaster damages is both time consuming and costly. Therefore, this research focuses on rapid detection of post-hurricane damages from aerial images based on a novel automated detection method.

There are two common methods to estimate damages in buildings using high-resolution images. The first method is called "change detection" and is based on identifying changes on an object at different times (Singh, 1989). Several change detection techniques were employed in previous works for detecting changes in buildings from pre- and post- disaster images (Brunner, et al., 2010, Matsuoka and Yamazaki, 2004, Matsuoka and Yamazaki, 2004, Myint, et al., 2008, Radhika, et al., 2013, Turker and Cetinkaya, 2005, Womble, et al., 2007). For example, tornado affected areas were identified using remote-sensing images and three change detection approaches, including the Principal Component Analysis (PCA), Image Differencing, and Object-oriented Classification (Myint, et al., 2008). The major limitation of these change detection methods is that the pre- and post- disaster images are needed to be acquired and matched for the whole area.

The second method is based on object classification. Relevant investigations focused on training classifiers using different classification methods, e.g., minimum distance classifier, decision tree algorithm and support vector machine (SVM) using various features, such as wavelet, Histograms of Oriented Gradients and texture features (Balz and Liao, 2010, Barnes, et al., 2007, Mitomi, et al., 2001, Radhika, et al., 2012, Saito, et al., 2004, Sumer and Turker, 2005, Yamazaki, et al., 2007). Moreover, a context-based detection method was proposed based on the combination of edge information, multi-spectral gray tone, and spatial relationship to identify damage levels of buildings from satellite images (Yamazaki, et al., 2007). However, they did not provide information about classification accuracies and specific damage conditions. Also, to rapidly identify tornado-borne debris, wavelet detection of edges and texture-wavelet analysis were used in (Radhika, et al., 2012), but the information about running time of the algorithms is not available.

This paper compares the performance of several classification methods (i.e., minimum distance classifier and five SVM classifiers) on different feature sets (e.g., filter responses, SURF descriptors and k-means based histograms), based on classification accuracies and running times.

RESEARCH APPROACH

In this paper, several classifiers are trained using a subset of labeled datasets. Then, the effectiveness of these classifiers is estimated on a testing dataset. The overall proposed approach for detecting damaged buildings is shown in Figure 1.

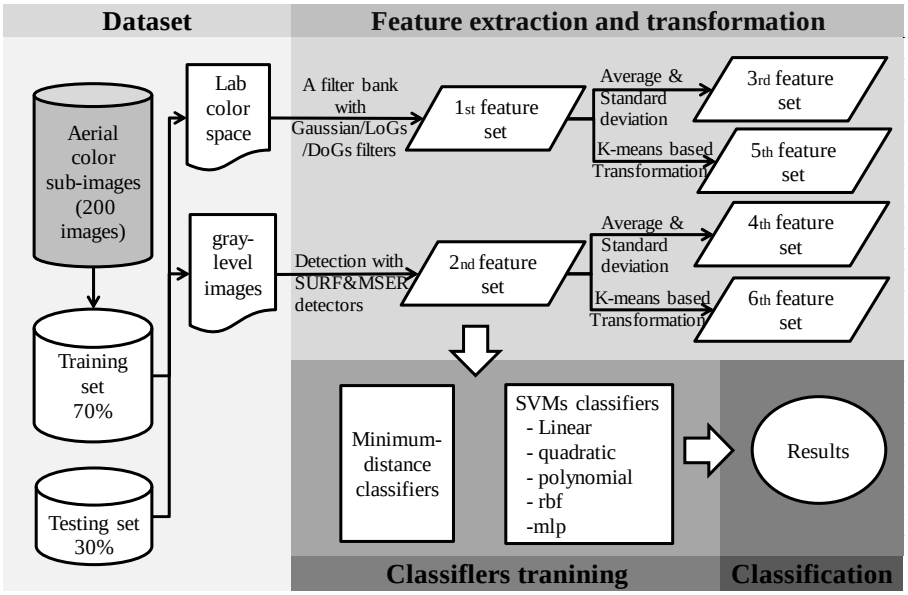

Figure 1. Overall approach for damaged buildings detection

DATASET

A dataset containing 200 aerial images of damaged and undamaged buildings (e.g., see Figure 2), was first created by manually extracting post-disaster images of the Long Beach Barrier Island in New Jersey. The overall view of the study area is shown in Figure 2(g).

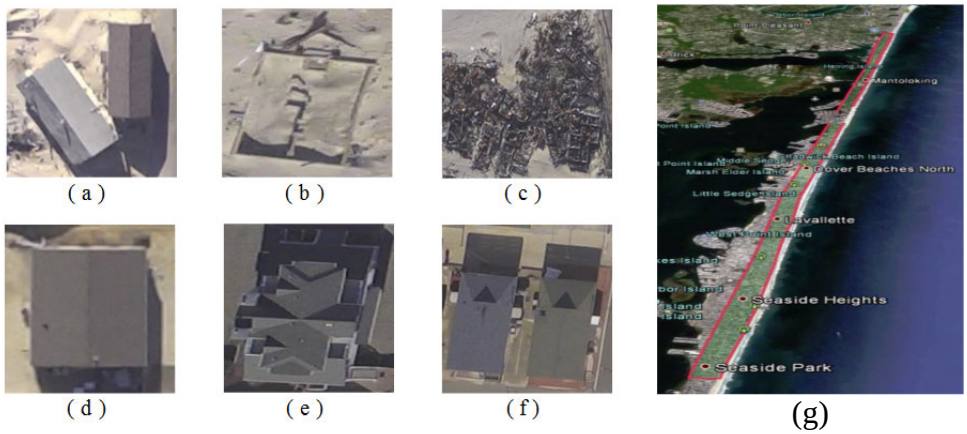

Figure 2. Study area (Long Beach Barrier Island in New Jersey) images: sample damaged images (a-c), sample undamaged images (d-f), and Satellite image of the Study area (g).

In order to measure classification accuracy, the dataset was divided into two sets: (a) 70% of the images were used as the training images, and (b) 30% of the

images were used as the testing images. Each image in the training and testing sets was downsampled to 200×200 pixels.

TRAINING PHASE

The training phase of the proposed approach has two steps: (1) feature extraction and transformation, to extract and determine informative features from the images, and (2) classifier training, to train the selected classification algorithms by using the extracted features from the training dataset.

Feature extraction. Two groups of feature sets were generated through feature extraction. The first feature set (i.e., Gaussian, Laplacian of Guassian (LOG) and first order derivatives of Gaussian (DoG) filter responses) was obtained from Lab color space which was derived from the "master" CIE 1931 XYZ color space, by applying a filter bank proposed in (Winn, et al., 2005). The filter-bank was made of three Guassian filters, four LoG filters and four DoG filters. It was applied to the CIE Lab color space with the dimension **L** for lightness and the other two dimensions (a and b) for the color-opponent dimensions. The three Gaussian kernels (with standard deviation σ=1,2,4) were applied to each L, a, and b channels, and the four LoGs (with σ=1,2,4,8) and four DoGs (with σ=1,2,4,8) were applied only to the **L** channel. Thus, 17 filter responses were produced for each pixel in an image, leading to a 40,000 (200×200) by 17 feature matrix. Finally, the first feature set was obtained by randomly selecting 10,000 rows of the feature matrix of each image.

The second feature set includes Speed-up Robust Features (SURF) and Maximally Stable Extremal Regions (MSER) features extracted from gray-level images by using SURF and MSER detectors. SURF features, first presented in (Bay, et al., 2006), are based on the determinant of the Hessian matrix and second order Gaussian derivative approximations. Integral images were used to improve calculation speeds. SURF points of interest were detected from gray-level images, each point associated with 64 descriptors, which represent the 4×4 sub-regions (each sub-region has a four dimensional descriptor vector $v=[\sum dx,\sum dy,\sum|dx|,\sum|dy|]$, where dx and dy are wavelet responses), around the point. For each image, the size of SURF features matrix was M-by-64, where M was the number of extracted points of interest. MSER was proposed by (Matas, et al., 2004). MSER areas were selected by choosing regions with unchanged shapes over a large set of thresholds. All pixels inside the MSER had higher or lower intensities than surrounding regions. MSER ellipse centers were obtained using MSER detectors from gray-level images. SURF feature descriptors were used to represent the characters of these centers. So, an N-by-64 MSER feature matrix was obtained for each image, where N was the number of extracted regions of interest. However, the number of detected points and regions of interest (i.e., M and N) varied in different images, and the minimum observed number for all the images in the dataset was 6. Therefore, the second feature set was defined by randomly selecting 6 points/regions to keep the same dimensions. To reduce the uncertainty due to this random selection, the process was repeated ten times, and classification accuracies were calculated by taking the average.

Feature Transformation. Four new feature sets were generated by: (i) calculating the average and standard deviation of features, or (ii) K-means based transformation. Average and standard deviation were used to reduce the dimensions of these features. The third feature set represents the average and standard deviation of the first feature set. The fourth feature set represents the average and standard deviation of the second feature set.

The K-means clustering method with K = 10 was applied to the feature matrix of the training dataset. K-means partitioned the M-dimensional feature vectors of the feature matrix (M = 17 in first feature set, M= 64 in the second feature set) into k M-dimensional clusters which minimized the within-cluster distances. Squared Euclidean distance method was used to calculate the within-cluster distances. Each feature vector belongs to the cluster with the smallest distance. Thus, a histogram of the distribution of indices for each image was formed.

CLASSIFIER TRAINING

We selected two groups of classification methods to train the classifiers: (1) the Euclidean minimum distance method and (2) the support vector machine method. In the first method, the Euclidean distances between an unlabeled sample image x and labeled images y (L_j) were calculated. Then, any given unlabeled image x was classified as the closest labeled images. The label **L** of the image x was found by:

$$L = \arg\min_{L_j} \sqrt{\sum_{i=1}^{n} \left(x_i - y_i(L_j)\right)^2} \quad (1)$$

In the second method, various SVM algorithms were used to train classifiers. Five kernel functions were used in this study, such as Linear, Quadratic, Polynomial, Gaussian Radial Basis (rbf), and Multilayer Perceptron (mlp). A hyperplane separating two-class data is obtained by maximizing the margin between two classes and minimizing training error. An image was classified according to Eq. (2). In the equation, w_i is scaling factor, k is the kernel function, S_i is the support vectors and b is the bias.

$$c = \sum_i w_i k(s_i, x) + b \quad (2)$$

where x is the feature vector representing the image. The value of c determines the class of x: if c > 0, then x is classified as a member of the first group, otherwise it is classified as a member of the second group.

Testing phase. The trained methods (i.e., minimum distance and SVM classifiers) were used to classify the testing dataset. Then, the classification accuracies were computed as the ratio between correctly classified images and the total number of images.

RESULTS

In this section, the classification accuracies and running times of different methods are presented, based on their application on the 6 proposed feature sets.

For validation purposes, we employed several methods similar to the method proposed by Winn et al. (2005). In their research, Winn et al. (2005) generated the features (i.e., filter responses) on every pixel of the image. Then, a nearest neighbor classifier was trained on the k-means (k=2000) clustered filter responses. In this

research, we developed k-means (k=10) clustered filter responses as the 6th feature set, and then trained a Minimum-distance classifier and five SVMs classifiers on it. With Minimum-distance classifier, performance (average accuracy 72.83%) comparable to the result (accuracy 76.3%) obtained by Winn et al. (2005) was achieved. The results also claimed that the Minimum-distance classifier (average accuracy 72.83%) performed better than SVMs classifiers (average accuracy from 50% to 57.5%) on this feature set. These classifiers were also trained on other feature sets, and the results are shown in Figure 3.

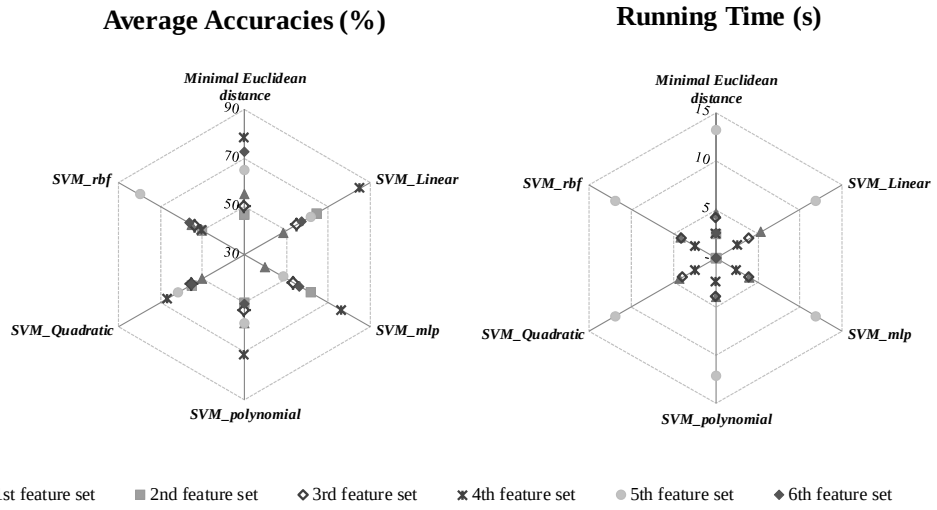

Figure 3. Accuracies and testing time of each feature set on different classification classifiers

Based on the results in Figure 3 the following remarks can be made:

1. For the Gaussian/LoGs/DoGs features, applying the k-means based transformation (applied on the 5th feature sets), enhanced the performance compared to using the features themselves (i.e. 1st feature sets). For SURF/MSER features, taking the average and standard deviation (applied on the 4th feature set) significantly improved the performance of most classifiers compared to using the features themselves (i.e. 2nd feature set).

2. The highest accuracy (i.e., 85%) was obtained by using linear SVM classifier on the 4th feature set, performing better than Winn et al. (2005) method. The linear SVM classifier performed well on 2nd, 3rd and 4th feature sets. Meanwhile, for all classifiers except the SVM rbf, the 4th feature set (i.e., average and standard deviation of SURF and MSER features) appears to be the best feature set.

3. Regarding the testing time, features sets of SURF/MSER (2nd, 4th and 6th feature sets) required less time than Guassian/LoGs/DoGs features (1st, 3rd and 5th feature sets). Moreover, the 5th feature set was the most time-consuming one. In general, the minimum distance classifier necessitates more time than the SVM classifiers.

CONCLUSIONS

The main goal of this paper was to compare various approaches for automated detection of damaged buildings using aerial imaging. In terms of features, results

showed that the SURF/MSER features performed better than the Guassian/LoGs/DoGs features. In terms of the classifiers, the linear SVM classifier appeared to be the best classifier for the proposed feature sets. The combination of the linear SVM method and the average and standard deviation of SURF/MSER features (4^{th} feature set) obtained the highest accuracy (85%). In general, for future real-time detection of damage in buildings, the usages of the linear SVM method and Minimum-distance method on 4^{th} feature set are recommended. For future work, we will continue to investigate other classification methods (e.g., ensemble methods), and will use more features (e.g., wavelet features). To develop real-time detection methods, the performance of different methods on various feature sets will be investigated, in terms of accuracies, running time, and associated storage memory.

ACKNOWLEDGMENT

The National Science Foundation support (CMMI # 1313863) is gratefully acknowledged. Any opinions, findings, conclusions, or recommendations presented in this paper are those of the authors and do not necessarily reflect the views of the National Science Foundation.

REFERENCES

Balz, T., and Liao, M. (2010). "Building-Damage Detection Using Post-Seismic High-Resolution Sar Satellite Data." *International Journal of Remote Sensing*, 31(13), 3369-3391.

Barnes, C. F., Fritz, H., and Jeseon, Y. (2007). "Hurricane Disaster Assessments with Image-Driven Data Mining in High-Resolution Satellite Imagery." *Geoscience and Remote Sensing, IEEE Transactions on*, 45(6), 1631-1640.

Bay, H., Tuytelaars, T., and Gool, L. (2006). "Surf: Speeded up Robust Features." *Computer Vision – Eccv 2006*, A. Leonardis, H. Bischof, and A. Pinz, eds., Springer Berlin Heidelberg, 404-417.

Brunner, D., Lemoine, G., and Bruzzone, L. (2010). "Earthquake Damage Assessment of Buildings Using Vhr Optical and Sar Imagery." *Geoscience and Remote Sensing, IEEE Transactions on*, 48(5), 2403-2420.

Csurka, G., Dance, C., Fan, L., Willamowski, J., and Bray, C. d. "Visual Categorization with Bags of Keypoints." *Proc., In Workshop on Statistical Learning in Computer Vision, ECCV*, 1-22.

Eric S. Blake, T. B. K., Robert J. Berg, John P. Cangialosi and John L. Beven II (2013). "Tropical Cyclone Report Hurricane Sandy (Al182012) 22-29 October 2012." National Hurricane Center.

Hasegawa, H., Aoki, H., Yamazaki, F., Matsuoka, M., and Sekimoto, I. "Automated Detection of Damaged Buildings Using Aerial Hdtv Images." *Proc., Geoscience and Remote Sensing Symposium, 2000. Proceedings. IGARSS 2000. IEEE 2000 International*, 310-312 vol.311.

Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). "Robust Wide-Baseline Stereo from Maximally Stable Extremal Regions." *Image and Vision Computing*, 22(10), 761-767.

Matsuoka, M., and Yamazaki, F. "Building Damage Detection Using Satellite Sar Intensity Images for the 2003 Algeria and Iran Earthquakes." *Proc.,*

Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International, 1099-1102.

Matsuoka, M., and Yamazaki, F. (2004). "Use of Satellite Sar Intensity Imagery for Detecting Building Areas Damaged Due to Earthquakes." *Earthquake Spectra*, 20(3), 975-994.

Mitomi, H., Saita, J., Matsuoka, M., and Yamazaki, F. "Automated Damage Detection of Buildings from Aerial Television Images of the 2001 Gujarat, India Earthquake." *Proc., Geoscience and Remote Sensing Symposium, 2001. IGARSS '01. IEEE 2001 International*, 147-149 vol.141.

Myint, S., Yuan, M., Cerveny, R., and Giri, C. (2008). "Comparison of Remote Sensing Image Processing Techniques to Identify Tornado Damage Areas from Landsat Tm Data." *Sensors*, 8(2), 1128-1156.

Radhika, S., Tamura, Y., and Matsui, M. (2012). "Use of Post-Storm Images for Automated Tornado-Borne Debris Path Identification Using Texture-Wavelet Analysis." *Journal of Wind Engineering and Industrial Aerodynamics*, 107–108(0), 202-213.

Radhika, S., Tamura, Y., and Matsui, M. "Texture-Wavelet Analysis for Automating Wind Damage Detection from Aerial Imageries." *Proc., Advance Computing Conference (IACC), 2013 IEEE 3rd International*, 1246-1250.

Sagara, E. (2012). "Hurricane Sandy's Destruction: Aerial Assessment Shows Nearly 72k Buildings Damaged in N.J." *New Jersey News*.

Saito, K., Spence, R. J. S., Going, C., and Markus, M. (2004). "Using High-Resolution Satellite Images for Post-Earthquake Building Damage Assessment: A Study Following the 26 January 2001 Gujarat Earthquake." *Earthquake Spectra*, 20(1), 145-169.

Singh, A. (1989). "Review Article Digital Change Detection Techniques Using Remotely-Sensed Data." *International Journal of Remote Sensing*, 10(6), 989-1003.

Sumer, E., and Turker, M. "Building Damage Detection from Post-Earthquake Aerial Imagery Using Building Grey-Value and Gradient Orientation Analyses." *Proc., Recent Advances in Space Technologies, 2005. RAST 2005. Proceedings of 2nd International Conference on*, 577-582.

Turker, M., and Cetinkaya, B. (2005). "Automatic Detection of Earthquake-Damaged Buildings Using Dems Created from Pre- and Post-Earthquake Stereo Aerial Photographs." *International Journal of Remote Sensing*, 26(4), 823-832.

Winn, J., Criminisi, A., and Minka, T. "Object Categorization by Learned Universal Visual Dictionary." *Proc., Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, 1800-1807 Vol. 1802.

Womble, J. A. P., Mehta, K. C. P., and Adams, B. J. P. (2007). "Automated Building Damage Assessment Using Remote-Sensing Imagery." 1-10.

Yamazaki, F., Vu, T. T., and Matsuoka, M. "Context-Based Detection of Post-Disaster Damaged Buildings in Urban Areas from Satellite Images." *Proc., Urban Remote Sensing Joint Event, 2007*, 1-5.