Article

Crowdsourcing Rapid Assessment of Collapsed Buildings Early after the Earthquake Based on Aerial Remote Sensing Image: A Case Study of Yushu Earthquake

Shuai Xie [1,2], Jianbo Duan [1,]*, Shibin Liu [1], Qin Dai [1], Wei Liu [1], Yong Ma [1], Rui Guo [1] and Caihong Ma [1]

[1] Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing 100094, China; xieshuai@radi.ac.cn (S.X.); liusb@radi.ac.cn (S.L.); daiqin@radi.ac.cn (Q.D.); liuwei@radi.ac.cn (W.L.); mayong@radi.ac.cn (Y.M.); guorui@radi.ac.cn (R.G.); mach@radi.ac.cn (C.M.)

[2] University of Chinese Academy of Sciences, Beijing 100049, China

* Correspondence: duanjb@radi.ac.cn; Tel.: +86-10-8217-8151

Abstract: Remote sensing (RS) images play a significant role in disaster emergency response. Web2.0 changes the way data are created, making it possible for the public to participate in scientific issues. In this paper, an experiment is designed to evaluate the reliability of crowdsourcing buildings collapse assessment in the early time after an earthquake based on aerial remote sensing image. The procedure of RS data pre-processing and crowdsourcing data collection is presented. A probabilistic model including maximum likelihood estimation (MLE), Bayes' theorem and expectation-maximization (EM) algorithm are applied to quantitatively estimate the individual error-rate and "ground truth" according to multiple participants' assessment results. An experimental area of Yushu earthquake is provided to present the results contributed by participants. Following the results, some discussion is provided regarding accuracy and variation among participants. The features of buildings labeled as the same damage type are found highly consistent. This suggests that the building damage assessment contributed by crowdsourcing can be treated as reliable samples. This study shows potential for a rapid building collapse assessment through crowdsourcing and quantitatively inferring "ground truth" according to crowdsourcing data in the early time after the earthquake based on aerial remote sensing image.

Keywords: crowdsourcing; building collapse assessment; earthquake; aerial image; EM algorithm

1. Introduction

Building collapse is one of the most serious types of earthquake damage. Most casualties from earthquakes are associated with collapsing buildings [1]. The extent of buildings damage reflects seismic intensity, which is important information to assess the losses of life and property in an earthquake-hit area [2]. Rapid assessment of collapsed buildings early after the earthquake can be instrumental in search and rescue during an emergency. It is hard to obtain the whole in-situ information of building damage in a short time after the earthquake, because the earthquake-damaged zones are not accessible in most cases. However, the aerial or satellite remote sensing can provide the image of the whole disaster area, making it possible to estimate the building damage of large disaster-affected regions in the early time. Many methods to visually interpret or automatically extract the building damage after the earthquake were proposed based on high-resolution aerial or satellite remote sensing image over the past ten years, which made a great contribution in estimating damage

extent of buildings caused by earthquake using remote sensing data. Chiroiu uses post-earthquake Ikonos imagery to assess the collapsed buildings by visual interpretation [3]. Saito et al. demonstrated that more collapsed buildings were recognized using pre- and post-earthquake QuickBird imagery, because the pre-earthquake imagery is a good reference of the building outlines [4]. Vu et al. uses the region-independent edge detection algorithm to detect the collapsed buildings based on Ikonos imagery. The results of "very heavy damage" and "destroyed" are consistent with visual interpretation and site survey [5]. Huyck et al. uses texture change detection algorithm based on pre- and post-earthquake imagery of mono-sensor and multi-sensor, respectively, finding that the results are quite different and only "hardest hit area" is recognized consistently [6]. Hutchinson et al. firstly extracts building boundary based on pre- and post-earthquake satellite imagery, then calculated the boundary compactness index defined as the ratio of the number of boundary pixels in the post- and pre-earthquake house, finally identifies the damage buildings through a threshold [7]. Chen and Hutchinson proposed a probabilistic classification framework by means of a multiclass classifier based on bitemporal satellite images to address the major limitation in past attempts which is the use of deterministic approaches to classify damage levels [8]. Geiß et al. quantitatively evaluates the suitability of multi-sensor remote sensing to assess the seismic vulnerability of buildings showing potential for a rapid screening assessment of large areas [1].

Due to the complex image characteristics of post-disaster ground objects and the limitation of resources, automated damage detection techniques with remote sensing are still in the preliminary stage [9]. The last decade has seen a proliferation of sophisticated sensors and technology capable of capturing, transferring and storing immense amounts of data, like remote sensing image, increasing the importance and demand for fast and reliable methods of analysis [9]. Web2.0 technology has resulted in changes in the way that data are created. Individuals, who have the characteristics of large quantity, flexible time and uncertain location, now provide vast amounts of information to websites and online databases, much of which is spatially referenced [10]. This phenomenon called "crowdsourcing" is the product of the network society, which is an online and distributed pattern of problem-solving and producing. Therefore, remote sensing combined with crowdsourcing was used to quickly and accurately analyze large data sets by creating and leveraging a distributed network of human analysts. Crowdsourcing geographic information for disaster response has become a research frontier [11]. The Virtual Disaster Viewer, which was a pilot project following the May 2008 Wenchuan, China earthquake, provided damage assessments through crowdsourcing by having experts interpret pre- and post-event satellite imagery [9]. After the 2010 earthquake in Haiti, the GEO-CAN initiative utilized crowdsourcing through the recruitment of experts to make critical damage assessments based on high-resolution post-event satellite and aerial imagery [12]. Building upon the GEO-CAN effort in Haiti, damage assessment after the Christchurch 2011 was improved by asking participants, including non-experts, to delineate damaged buildings use a polygonal tool, in order to making crowdsourcing damage assessments of disaster areas faster and more accurate [13]. The above studies have clearly demonstrated the power of crowdsourcing for damage assessment to improve disaster response. As such, it offers substantial advantages, but suffers from a general lack of quality assurance [14]. The participants with different professional background and knowledge level have different understanding of remote sensing image. The interpretation results of the same image may be different. It is necessary to quantitatively evaluate the quality of crowdsourcing data to ensure the accuracy of damage assessment.
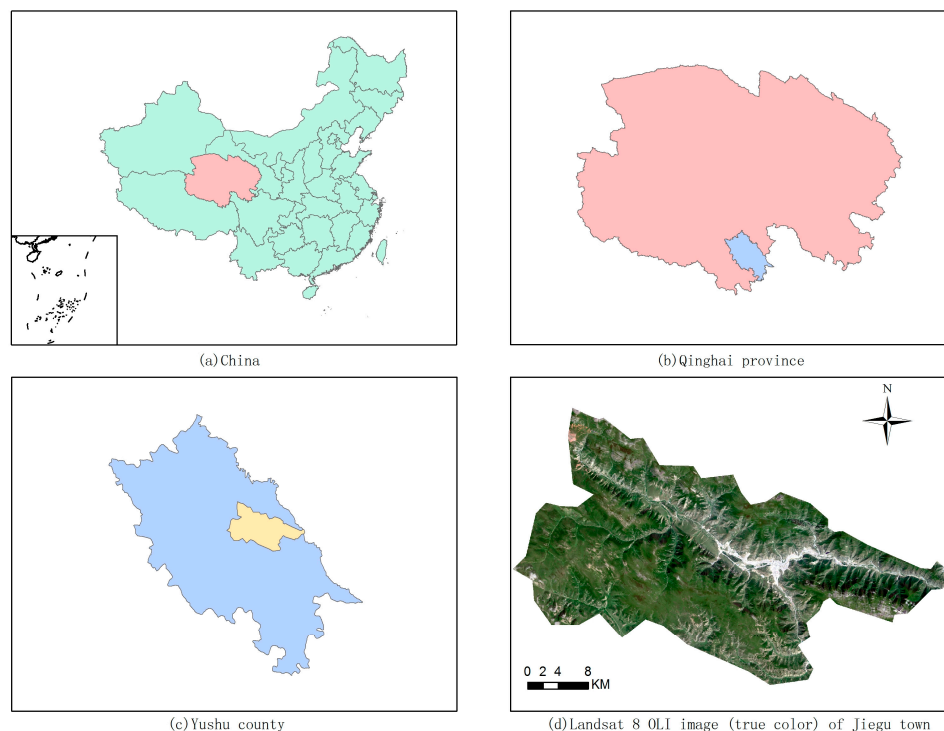
In this paper, Yushu earthquake is chosen as a case study. A web-based platform is built to collect the post-earthquake building damage assessment results contributed by public participants. High-resolution aerial remote sensing images are used due to the advantage of being captured and processed faster and higher spatial resolution than satellite imagery [15]. The problem we focus on is damage "extent" identification for buildings which is relatively straightforward and fast by means of RS images, instead of damage "level" assessment which is based on ground evaluations requiring a considerable amount of time and effort [16]. The next section describes the details of

study area. The data used for this research and overview of processing are introduced in the third section. The probabilistic model is applied to estimate individuals' error-rates and infer ground truth. The experiment results are presented in the fourth section. Following the results, some discussion is provided regarding the accuracy of EM algorithm compared to the "majority" method and the variation distribution of assessment results among participants on each building. In addition, the features of each crowdsourcing-derived damage type are analyzed, which can be regarded as reliable samples to train machine learning to recognize objects of interest. Our main conclusions are presented in the final section and future work is proposed.

2. Materials and Methods

2.1. Study Area and Data

On 14 April 2010, a 7.1 magnitude earthquake occurred near Yushu, China, at 7:49 a.m. local time. The epicenter was located at 33.1 degrees north latitude, 96.6 degrees east longitude and focal depth was 14 km. The terrain is mainly mountainous, with an average elevation of 4493 m. The earthquake caused a large number of casualties and collapsed houses. The site survey data of Chinese scholars after the earthquake demonstrate that houses in the central area of Jiegu Town are mainly with brick and concrete structure, while other houses are mainly with brick and civil structure. Some of the brick and concrete structures in the town center area suffered serious damage, and almost all the brick and civil structures in the western and southern regions of the town were totally damaged [17]. Figure 1 below shows the location of Jiegu Town, where the earthquake happened.

Figure 1. Maps about administrative areas and terrains: (**a**) the geographical location of Qinghai Province, China; (**b**) the geographical location of Yushu County in Qinghai Province; (**c**) the geographical location of Jiegu Town in Yushu County; and (**d**) the topographic map of Jiegu Town made by Landsat 8 OLI image (true color).

The data used for this application are a 0.4 m resolution multispectral aerial image of the damaged zones of Jiegu Town captured on 14 April 2010, provided by Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences (Figure 2).

Figure 2. The aerial image of Jiegu Town captured on 14 April 2010, which was the area severely damaged by the earthquake.

2.2. Methods

2.2.1. Architecture of Processing

The architecture consists of four parts, which are, in order, basic imagery preparation, damage assessment collection, data quality evaluation and damage map export (Figure 3).

(1) Basic imagery preparation

In this part, we take the post-event high resolution aerial image (Section 2.1 mentioned) as basic imagery, which crowdsourcing participants use to visually interpret collapsed buildings. We obtain two images that cover the different part of the study area, respectively, but having the overlapping zone. To generate the entire region image, image registration, mosaic and clipping are applied to the two images we obtain.

(2) Damage assessment collection

We publish the pre-processed image online that is accessible to participants. Each visible building in the image is assigned one of the damage types according to the observed damage.

(3) Data quality evaluation

Because of the different professional background of participants, the quality of the data provided by the individual is uneven. The term crowdsourcing has three distinct meanings proposed by Goodchild [14]. The first meaning refers to the solution of a problem by referring it to a number of people, without respect to their qualifications. The second meaning refers to the ability of a group to validate and correct the errors that an individual might make. The third interpretation refers to the ability of the crowd to converge on the truth. We could not get the field data in a short time after the earthquake. However, we could infer ground truth from subjective labeling of the post-event high-resolution aerial image by participants. In the next section, we introduce the details of how we use maximum likelihood estimation, Bayes' theorem and EM algorithm to estimate the ground truth and individual error rate.

(4) Damage map export

In the last part, the results integrated by the previous part are visualized, generating a damage map that shows the spatial distribution of damage types. The damage distribution map is generated with individual buildings rendered with colors representing the type of damage.
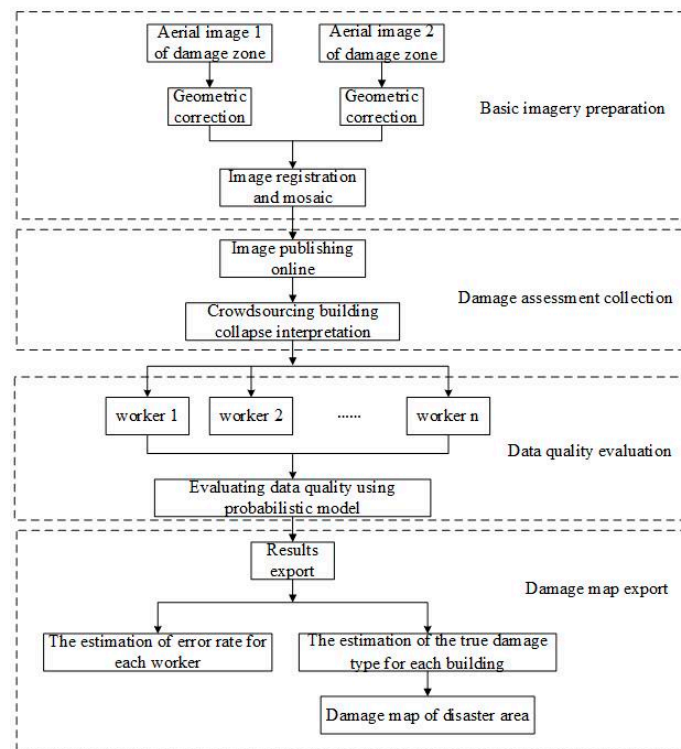
Figure 3. The architecture of processing.

The process above enables assessing the damage of buildings rapidly and flexibly through crowdsourcing without considering the constraint conditions when using algorithm to process image. Each participant is regarded as a "classifier", which classifies post-disaster buildings into three categories according to human's understanding of the image, and then we evaluate the accuracy of each person and integrate the multiple people's results into a final reliable result.

2.2.2. Probabilistic Model

Note that there are *W* participants to assess I buildings, which may be damaged using *K* damage types. It is assumed that all responses given by a single participant are independent and all the participants interpret independently. In addition, a participant may interpret the same building more than once. Note that $\alpha_{kl}^{(w)}$ ($k = 1, \ldots , K; l = 1, \ldots , K; w = 1, \ldots , W$) are the probability that a participant w will label *l* given *k* is the true type, which are called the individual error rate. $n_{il}^{(w)}$ ($i = 1, \ldots ,$ I; $l = 1, \ldots , K; w = 1, \ldots , W$) are the number of times participant w label building *i* as *l*, and p_k ($k = 1, \ldots , K$) are the probability that the true damage type of building is *k*. Let G_{ik} ($k = 1, \ldots , K$) be a binary variable of building *i*. If *t* is the true damage type of building *i*, then $G_{it} = 1$ and $G_{ik} = 0$ ($k \neq t$), namely, p$(G_{ik} = 1) = p_k$. We follow a general model for subjective labeling originally proposed by Dawid and Skene [18] and apply it to the building damage labeling problem. The data from all participants are assumed to be independent and all the true damage types of buildings are assumed to be available. Generally, the likelihood function for the full data is

$$\prod_{i=1}^{I} \prod_{k=1}^{K} \{ p_k \prod_{w=1}^{W} \prod_{l=1}^{K} (\alpha_{kl}^{(w)})^{n_{il}^{(w)}} \}^{G_{ik}} . \tag{1}$$

Using maximum likelihood estimation, and we obtain estimators

$$\hat{\alpha}_{kl}^{(w)} = \frac{\sum_i G_{ik} n_{il}^{(w)}}{\sum_l \sum_i G_{ik} n_{il}^{(w)}} . \tag{2}$$

When p_k ($k = 1, \dots, K$) are unknown, these can be estimated:

$$\hat{p}_k = \frac{\sum_i G_{ik}}{I}. \tag{3}$$

At this point, the true damage types of buildings are unknown. We using Bayesian theory to estimate the binary variable G_{ik} ($k = 1, \dots, K$),

$$\text{p}\left(G_{ik} = 1 \mid \text{data}\right) = \frac{\text{p}\left(\text{data} \mid G_{ik} = 1\right)\text{p}\left(G_{ik} = 1\right)}{\text{p}\left(\text{data}\right)} = \frac{\text{p}\left(\text{data} \mid G_{ik} = 1\right)\text{p}\left(G_{ik} = 1\right)}{\sum_{t=1}^{K} p\left(\text{data} \mid G_{it} = 1\right)\text{p}\left(G_{it} = 1\right)}. \tag{4}$$

Therefore,

$$\text{p}\left(G_{ik} = 1 \mid \text{data}\right) = \frac{\prod_{w=1}^{W}\prod_{l=1}^{K}\left(\alpha_{kl}^{(w)}\right)^{n_{il}^{(w)}} p_k}{\sum_{t=1}^{K}\prod_{w=1}^{W}\prod_{l=1}^{K}\left(\alpha_{tl}^{(w)}\right)^{n_{il}^{(w)}} p_t}. \tag{5}$$

Then, we use EM algorithm for finding maximum likelihood estimates of parameters in the model above, due to the dependency of the hidden variables G_{ik}. EM algorithm is short for Expectation Maximization algorithm, which was described by Dempster et al. in 1977 [19]. It is an iterative optimization method for maximum likelihood estimation of parameters, which can estimate the parameters from incomplete data set.

In this problem, we treat G_{ik} as missing data then the conditions of the EM algorithm are satisfied. The iterative procedure is as follows:

1. Give initial estimates of the Gs.
2. Use Equations (2) and (3) to obtain estimates of the ps and αs.
3. Use Equation (5) and the estimates of the ps and αs to calculate new estimates of the Gs.
4. Repeat Steps 2 and 3 until the results converge.

In Step 1, we use the equation below to calculate initial estimates of Gs,

$$\hat{G}_{ik} = \frac{\sum_w n_{ik}^{(w)}}{\sum_w \sum_l n_{il}^{(w)}}$$

3. Experiment Results

Our project asked the participants to classify the post-earthquake damage buildings into one of three damage types: (1) basically intact; (2) partially collapsed; and (3) completely collapsed. These type numbers are used in subsequent tables. Here, the Yushu earthquake case study was selected to illustrate the results. The experiment area was a sub-region of Jiegu Town, which had visible various types of damage extent, shown in Figure 4, and contained 3456 data points labeled by 27 volunteers, describing the damage buildings at 127 locations. As can be seen in Figure 4, "basically intact", "partially collapsed" and "completely collapsed" are represented by green, yellow and red points, respectively.
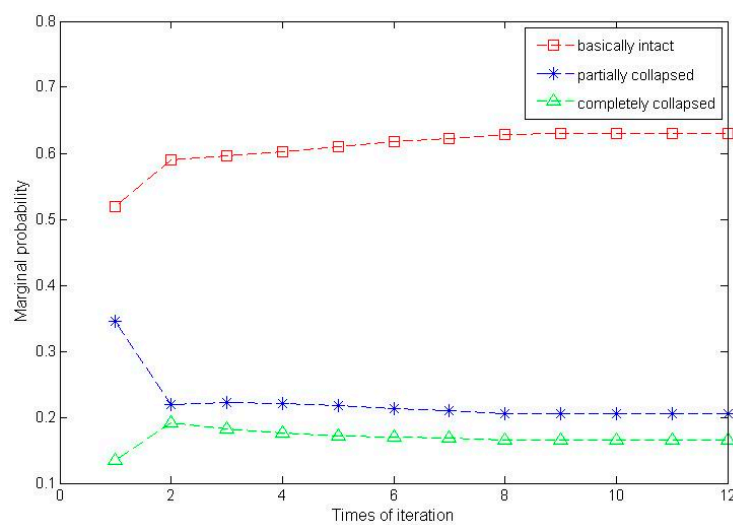
The system consists of a database for damage assessment accessed through a browser-based interface built using the ArcGIS API. Data from the participants are collected in the browser and transferred into the database through AJAX and PHP. One selects a damage type and then draws a point on the corresponding building. When a participant assesses a building's damage, the record is stored along with the longitude and latitude of the point he draw, the damage type he select and his id. The data we collected were used later to estimate the error-rate of each participant in identifying the damage extent of each building and infer the ground truth.

Figure 5 gives the variation tendency of marginal probabilities of the three damage types with the iteration of EM algorithm. As shown in Figure 5, the results converge when the iteration is

12 times. Table A1 (see Appendix A) gives the estimates of the individual error-rates (α) of the 27 participants. The diagonal elements of each matrix are the estimate of the probability of a correct allocation by a participant. Table A2 (see Appendix A) gives the estimated probabilities for the Gs for each building. For most buildings, the posterior probability is 1 for one damage type, and the consensus appears obvious.
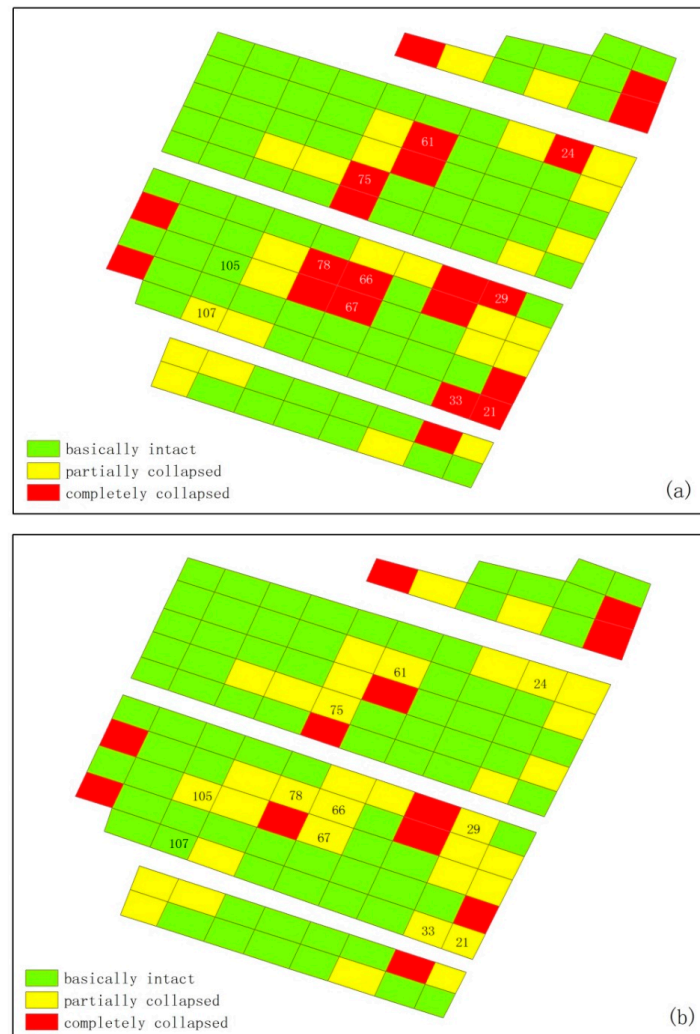

Figure 4. The experiment area and the location of the 3456 data points, which were contributed by 27 participants online. In the map, the green, yellow and red points indicate the damage type of "basically intact", "partially collapsed" and "completely collapsed", respectively.

Figure 5. The red, blue and green lines demonstrate the fluctuation of marginal probability of "basically intact", "partially collapsed" and "completely collapsed", respectively, with the increase of the number of iterations. When the number of iterations reaches 12, EM algorithm converges.

4. Discussion

Of the 3456 damage assessments received for the experiment region, we find that "basically intact" annotations made up 52.14%, "partially collapsed" made up 34.64% and "completely collapsed" made up 13.22%. There is no clear bias towards one or two damage types. However, if using the EMS-98 scale, the distribution of annotations reveals an overall bias to assess a building as "No Damage" or "Destroyed" [13]. In order to demonstrate the advantage of EM algorithm in terms of inferring ground true, we make a comparison with "majority" method. Figure 6a,b shows the assessment results of EM algorithm and "majority" method, respectively.

Figure 6. Comparison of the results between the two different methods: (**a**) the result of EM algorithm; and (**b**) the result of "majority" method, which are generated with individual buildings rendered with colors representing the type of damage. The rectangles represent the buildings, and green, yellow and red indicate "basically intact", "partially collapsed" and "completely collapsed", respectively. The rectangles with the id number on them have the different results between the two methods.

As shown in Figure 6, there are 11 buildings that have different results between the two methods: 21, 24, 29, 33, 61, 66, 67, 75, 78, 105 and 107. The first nine buildings are completely collapsed in EM results while are partially collapsed in "majority" method. The No. 105 building is basically intact in EM results while is partially collapsed in "majority" method. The situation of No. 107 building is opposite to No. 105. The "majority" method does not take into account the accuracy

of the participants, and EM may not choose the majority damage types as the final result of one building due to the low accuracy of participants who make the assessment. For example, the No. 33 building received the assessment results of 26 participants. Among them, there are 19 participants who gave the "partially collapsed" result and six participants who assessed the No. 33 building as "completely collapsed". Consequently, the "majority" method regards the No. 33 building as "partially collapsed". According to Table A1, we calculate the average accuracy of 19 participants who label as "partially collapsed" when the true is partially collapsed and the average accuracy of six participants who label as "completely collapsed" when the true is completely collapsed. The calculation results are shown in Table 1. We also calculated the corresponding incidence defined as the product of individual accuracy and marginal probability of damage type, as seen in Table 2. Obviously, the average accuracy of the latter is larger than the former and so is the average incidence, indicating that the participants who label the No. 33 building as "completely collapsed" have more "weight".

Table 1. The accuracy of 19 participants who label No. 33 building as "partially collapsed" and 6 participants who label No. 33 building as "completely collapsed". The averages are calculated in the last line of the table.

Participant ID	Accuracy of 2	Accuracy of 3
1	0.5021	
2	0.2699	
3	0.3825	
4		0.6200
5	0.6806	
6	0.7419	
8	0.5761	
9		0.8492
10	0.6132	
11	0.8465	
12	0.7924	
13		0.6200
14		0.5246
15	0.6887	
16	0.5912	
17	0.6015	
18	0.5011	
20	0.7608	
21	0.7306	
22		0.7602
23	0.1905	
24	0.3056	
25		0.9523
26	0.2712	
27	0.5359	
Average	0.5570	0.7211

To demonstrate the variation distribution of assessment results that participants give on each building, the percentage of each damage type on each building is presented in Figure 7, in which the building ids are sorted by the percentage of "basically intact" from small to large. Besides, the standard deviation of participants' assessment results on each building is shown in Figure 8. No clear agreement between participants on each building apart from the No. 99 that all participants label as "basically intact". This is because participants with different professional background have different cognition, or due to the limitation of image inherent characteristics such as spatial resolution and angle of view. Whatever limitations the professional faces naturally also apply to volunteers [20]. However, no obvious bias towards extreme value indicates that a majority of participants worked without malice.

Table 2. The incidence of 19 participants who label No. 33 building as "partially collapsed" and 6 participants who label No. 33 building as "completely collapsed". The averages are calculated in the last line of the table.
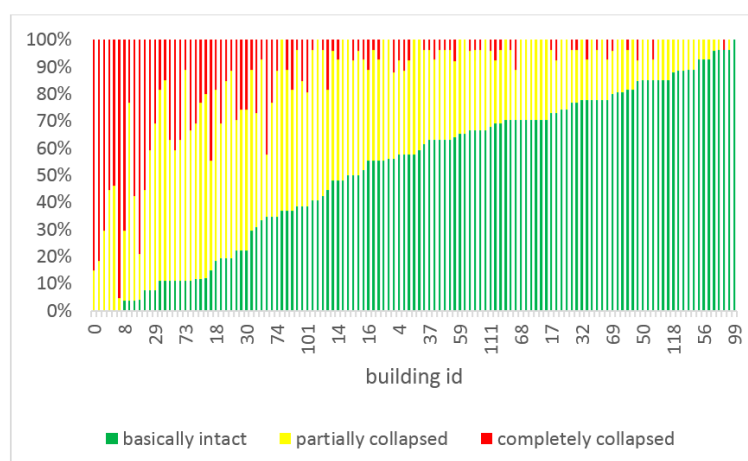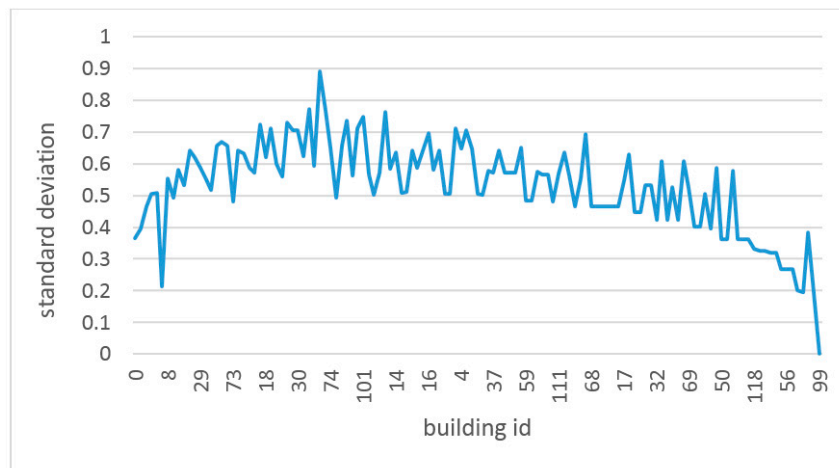
Participant ID	Incidence of 2	Incidence of 3
1	0.1030	
2	0.0553	
3	0.0784	
4		0.1024
5	0.1395	
6	0.1521	
8	0.1181	
9		0.1402
10	0.1257	
11	0.1736	
12	0.1625	
13		0.1024
14		0.0866
15	0.1412	
16	0.1212	
17	0.1233	
18	0.1027	
20	0.1560	
21	0.1498	
22		0.1255
23	0.0391	
24	0.0627	
25		0.1572
26	0.0556	
27	0.1099	
Average	0.1142	0.1190

In remote sensing applications, "ground truth" data are often used as the basis for training pattern recognition algorithms to detect objects of interest [21]. Many semiautomatic techniques have been designed to exploit Earth Observation (EO) data for earthquake damage assessment to the maximum possible extent, yet visual inspection still remains the best way to achieve meaningful results [16]. Combined with multiple participants, probability model-derived final results are used as the reliable samples to analyze the features of each damage type.

Figure 7. The percentage of each damage type on each building. The green, yellow and red bars represent the percentage of "basically intact", "partially collapsed" and "completely collapsed", respectively, on each building. The building ids depicted by *X*-axis are in ascending order according to the percentage of "basically intact" on each building.

Figure 8. The standard deviation of participants' assessment results on each building. The building ids depicted by *X*-axis are also in ascending order according to the percentage of "basically intact" on each building.

We select some samples of each damage type from the results of EM algorithm, as shown in Table 3. The number 1, 2 and 3 represent the damage type of basically intact, partially collapsed and completely collapsed, respectively. The common features of damage type 1 have a clear outline and regular shape, and an intact shadow. Damage type 2 has fuzzy boundaries, irregular shape, offset orientation, and loss of shadow effects. Damage type 3 has no visible characteristics of man-made objects.

Some limitations of our experiment are discussed below.

Table 3. The crowdsourcing-derived samples of each damage types.

Damage Types	Sample 1	Sample 2	Sample 3	Sample 4
1	11	98	109	110
2	15	19	36	96
3	0	2	8	76

The study area in this paper is one part of the whole damage area, which we used as an experimental area to apply our methods on it and present the results clearly. In addition, the approach is also applicable in the wide area, which is the most important purpose of the crowdsourcing damage assessment. Because the model is proposed for the general circumstance without the limits of area, we will extend the application to wide interpretation in future work. The raw data we collected show that most participants interpret only once, although the method allows participants to interpret multiple times. We could not determine the minimum number of the participants, because the web-based interface is open to the public and any case could appear. We could calculate a result based on a span from the start to the time we choose, for example a week. The method would give a result based on any data collection phase.

Although "actual" ground truths of building damage are necessary in order to truly discuss the applicability of the proposed method, we could not find ground photographs or broadcasted videos on

TV of the damaged buildings in the target area, only text information or papers are available. We focus on using the wisdom of the public to find more damaged buildings in the early stage of the earthquake. The assessment results of the target area in this paper are highly consistent with results of the existing study published by Dou et al. [17] in the corresponding area.

5. Conclusions and Future Work

Building damage assessment in the early time after an earthquake is a very crucial problem. Knowing where the collapsed buildings are and to what extent buildings have been damaged are closely related to life-saving for emergency response. However, it is hard to survey the whole in-situ information in a short time after an earthquake. Satellite or aerial remote sensing technology has the capability of earth observation, and becomes a useful tool for damage estimation without being physically present in disaster area. Aerial remote sensing images, which are captured and processed faster and have higher spatial resolution, make it possible to rapidly assess collapsed buildings early after the earthquake. A web-based interface was built. Anyone who accessed our website was required to assign one of the three damage types for each buildings based on the aerial image. The 3456 data points from 27 participants on the experimental area, which is a sub-region of Yushu, were collected. MLE, Bayes' theorem and EM algorithm were applied to estimate the individual error-rates and infer "ground truth" according to the 3456 data points of 27 participants. The results suggest that EM algorithm is better than "majority" method and there is no clear bias towards extreme value among damage types contributed by participants. This study shows that the variation of image understanding among participants exists, due to their different professional background. We demonstrated how to collect and store the data created by individuals online, how to make them contribute their results flexibly and easily, drawing a point instead of a polygon, and how to quantitatively estimate individuals' accuracy and the "ground truth" of each building, using a probabilistic model. By means of a sequential procedure of RS image pre-processing, publishing RS image online, collecting crowdsourcing building damage assessment data, evaluating the quality of data contributed by crowdsourcing and damage mapping, the building collapse can be rapidly assessed with viable accuracies in the early time after the earthquake. We conclude that RS data combined with crowdsourcing have a high capability to support large-area assessments of building collapse, meeting the need of disaster emergency response. A new processing framework is proposed to establish the connection between remote sensing image and crowdsourcing, demonstrating potential for crowdsourcing rapid assessment of building collapse early after the earthquake based on aerial remote sensing image. Future efforts will focus on providing multi-source and multi-temporal remote sensing image. Multi-source RS data, such as oblique images, can provide more views of buildings on the ground. Multi-temporal RS data, such as pre-earthquake images, are considered to be very useful as a reference in identifying the damage in the post-event image [4]. Although these adjustments will refine the final results towards rapid damage assessment, a problem should be considered: how to balance the operational complexity of system and the improvement of results, because ordinary people are more inclined to use easy-to-operate systems without spending too much time. In summary, the aim is that the rapid assessment results through crowdsourcing could meet the needs of deploying rescue forces in the early time after the earthquake. More details about the damage level of buildings surveyed on the spot afterwards are beyond the scope of this paper.

Author Contributions: Jianbo Duan conceived and designed the experiments; Jianbo Duan, Rui Guo and Caihong Ma built the experimental platform; Shuai Xie prepared and processed the aerial data, performed the experiments, analyzed the data and wrote the paper; Shibin Liu supervised the research; and Qin Dai, Wei Liu and Yong Ma gave comments and revised the manuscript.

Appendix A

Table A1. The estimates of the individual error-rates.

Participant 1				Participant 15			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.64	0.35	0.01	1	0.52	0.48	0
2	0.38	0.5	0.11	2	0.15	0.69	0.16
3	0	0.35	0.65	3	0.1	0.29	0.62

Participant 2				Participant 16			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.83	0.12	0.05	1	0.7	0.3	0
2	0.5	0.27	0.23	2	0.07	0.59	0.34
3	0	0.36	0.64	3	0.06	0.11	0.83

Participant 3				Participant 17			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.9	0.08	0.03	1	0.99	0.01	0
2	0.62	0.38	0	2	0.4	0.6	0
3	0.24	0.53	0.24	3	0.1	0.8	0.1

Participant 4				Participant 18			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.66	0.3	0.04	1	0.76	0.24	0
2	0.23	0.73	0.04	2	0.25	0.5	0.25
3	0.09	0.29	0.62	3	0	0.37	0.63

Participant 5				Participant 19			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.94	0.06	0	1	0.91	0.09	0
2	0.32	0.68	0	2	0.39	0.61	0
3	0.05	0.8	0.15	3	0.11	0.37	0.53

Participant 6				Participant 20			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.46	0.5	0.04	1	0.46	0.29	0.24
2	0.12	0.74	0.14	2	0	0.76	0.24
3	0	0.32	0.68	3	0	0.11	0.89

Participant 7				Participant 21			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.47	0.47	0.06	1	0.83	0.15	0.02
2	0.46	0.54	0	2	0.27	0.73	0
3	0.25	0.35	0.4	3	0.05	0.67	0.29

Participant 8				Participant 22			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.9	0.09	0.01	1	0.54	0.45	0.01
2	0.42	0.58	0	2	0	0.46	0.54
3	0.08	0.61	0.3	3	0	0.24	0.76

Table A1. *Cont.*

Participant 9				Participant 23			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.26	0.56	0.17	1	1	0	0
2	0.04	0.4	0.56	2	0.81	0.19	0
3	0	0.15	0.85	3	0.43	0.43	0.14

Participant 10				Participant 24			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.73	0.27	0	1	0.94	0.06	0
2	0.08	0.61	0.31	2	0.69	0.31	0
3	0	0.3	0.7	3	0.26	0.48	0.26

Participant 11				Participant 25			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.91	0.09	0	1	0.48	0.49	0.02
2	0.15	0.85	0	2	0	0.77	0.23
3	0	0.73	0.27	3	0	0.05	0.95

Participant 12				Participant 26			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.36	0.64	0	1	0.98	0.03	0
2	0	0.79	0.21	2	0.69	0.27	0.04
3	0.05	0.68	0.26	3	0.2	0.35	0.45

Participant 13				Participant 27			
	Observed				Observed		
True	1	2	3	True	1	2	3
1	0.77	0.18	0.05	1	0.76	0.19	0.05
2	0.23	0.5	0.27	2	0.35	0.54	0.12
3	0	0.38	0.62	3	0.05	0.38	0.57

Participant 14			
	Observed		
True	1	2	3
1	0.67	0.33	0
2	0.09	0.91	0
3	0.05	0.43	0.52

Table A2. The estimated probabilities for the Gs.

Building ID	Damage Types			Building ID	Damage Types		
	1	2	3		1	2	3
1	0	0	1	65	1	0	0
2	0.002	0.998	0	66	0.024	0.976	0
3	0	0	1	67	0	0	1
4	1	0	0	68	0	0.003	0.997
5	1	0	0	69	1	0	0
6	0.006	0.994	0	70	1	0	0
7	1	0	0	71	1	0	0
8	0.996	0.004	0	72	0.956	0.044	0
9	0	0	1	73	0.999	0.001	0
10	1	0	0	74	0	1	0
11	1	0	0	75	0	1	0
12	1	0	0	76	0	0	1
13	0.01	0.99	0	77	0	0	1

Table A2. *Cont.*

Building ID	Damage Types			Building ID	Damage Types		
	1	**2**	**3**		**1**	**2**	**3**
14	0	1	0	78	1	0	0
15	1	0	0	79	0	0	1
16	0	1	0	80	0	0	1
17	1	0	0	81	1	0	0
18	1	0	0	82	1	0	0
19	0	1	0	83	1	0	0
20	0	1	0	84	1	0	0
21	0	0	1	85	1	0	0
22	0	0	1	86	1	0	0
23	0	1	0	87	1	0	0
24	1	0	0	88	0	0.97	0.03
25	0	0.072	0.928	89	1	0	0
26	1	0	0	90	1	0	0
27	1	0	0	91	0	1	0
28	1	0	0	92	0.001	0.999	0
29	0	1	0	93	1	0	0
30	0	0	1	94	0	1	0
31	0	1	0	95	0	1	0
32	0	1	0	96	0	1	0
33	1	0	0	97	0	0.99	0.01
34	0	0	1	98	1	0	0
35	0	0	1	99	1	0	0
36	1	0	0	100	1	0	0
37	0	0.998	0.002	101	1	0	0
38	1	0	0	102	0	1	0
39	1	0	0	103	1	0	0
40	1	0	0	104	1	0	0
41	1	0	0	105	0.999	0.001	0
42	0	0	1	106	1	0	0
43	0	0	1	107	1	0	0
44	1	0	0	108	0	1	0
45	1	0	0	109	1	0	0
46	1	0	0	110	1	0	0
47	1	0	0	111	1	0	0
48	0	1	0	112	1	0	0
49	1	0	0	113	1	0	0
50	1	0	0	114	1	0	0
51	1	0	0	115	1	0	0
52	1	0	0	116	1	0	0
53	1	0	0	117	1	0	0
54	0	1	0	118	1	0	0
55	1	0	0	119	1	0	0
56	1	0	0	120	1	0	0
57	1	0	0	121	1	0	0
58	1	0	0	122	0	0	1
59	1	0	0	123	1	0	0
60	1	0	0	124	1	0	0
61	1	0	0	125	1	0	0
62	0	0	1	126	0	0	1
63	0	0	1	127	1	0	0
64	1	0	0				

References

1. Geiß, C.; Taubenböck, H.; Tyagunov, S.; Tisch, A.; Post, J.; Lakes, T. Assessment of seismic building vulnerability from space. *Earthq. Spectra* **2014**, *30*, 1553–1583. [CrossRef]

2. Lei, L.; Liu, L.; Zhang, L.; Bi, J.; Wu, Y.; Jiao, Q.; Zhang, W. Assessment and analysis of collapsing houses by aerial images in the wenchuan earthquake. *J. Remote Sens.* **2010**, *14*, 333–344.

3. Chiroiu, L. Damage assessment of the 2003 Bam, Iran, earthquake using Ikonos imagery. *Earthq. Spectra* **2005**, *21*, 219–224. [CrossRef]

4. Saito, K.; Spence, R.; de C Foley, T.A. Visual damage assessment using high-resolution satellite images following the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 309–318. [CrossRef]

5. Vu, T.T.; Matsuoka, M.; Yamazaki, F. Detection and animation of damage using very high-resolution satellite data following the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 319–327. [CrossRef]

6. Huyck, C.K.; Adams, B.J.; Cho, S.; Chung, H.-C.; Eguchi, R.T. Towards rapid citywide damage mapping using neighborhood edge dissimilarities in very high-resolution optical satellite imagery—Application to the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 255–266. [CrossRef]

7. Hutchinson, T.C.; Chen, Z.Q. Optimized estimated ground truth for object-based urban damage estimation using satellite images from the 2003 Bam, Iran, earthquake. *Earthq. Spectra* **2005**, *21*, 239–254. [CrossRef]

8. Chen, Z.; Hutchinson, T.C. Probabilistic urban structural damage classification using bitemporal satellite images. *Earthq. Spectra* **2010**, *26*, 87–109. [CrossRef]

9. Barrington, L.; Ghosh, S.; Greene, M.; Har-Noy, S.; Berger, J.; Gill, S.; Lin, A.Y.M.; Huyck, C. Crowdsourcing earthquake damage assessment using remote sensing imagery. *Ann. Geophys.* **2011**, *54*, 680–687.

10. See, L.; Comber, A.; Salk, C.; Fritz, S.; van der Velde, M.; Perger, C.; Schill, C.; McCallum, I.; Kraxner, F.; Obersteiner, M. Comparing the quality of crowdsourced data contributed by expert and non-experts. *PLoS ONE* **2013**, *8*, e69958. [CrossRef] [PubMed]

11. Goodchild, M.F.; Glennon, J.A. Crowdsourcing geographic information for disaster response: A research frontier. *Int. J. Digit. Earth* **2010**, *3*, 231–241. [CrossRef]

12. Ghosh, S.; Huyck, C.K.; Greene, M.; Gill, S.P.; Bevington, J.; Svekla, W.; DesRoches, R.; Eguchi, R.T. Crowdsourcing for rapid damage assessment: The global earth observation catastrophe assessment network (GEO-CAN). *Earthq. Spectra* **2011**, *27*, S179–S198. [CrossRef]

13. Huynh, A.; Eguchi, M.; Lin, A.Y.-M.; Eguchi, R. Limitations of crowdsourcing using the EMS-98 scale in remote disaster sensing. In Proceedings of the 2014 IEEE Aerospace Conference, Big Sky, MT, USA, 1–8 March 2014.

14. Goodchild, M.F.; Li, L. Assuring the quality of volunteered geographic information. *Spat. Stat.* **2012**, *1*, 110–120. [CrossRef]

15. Ofli, F.; Meier, P.; Imran, M.; Castillo, C.; Tuia, D.; Rey, N.; Briant, J.; Millet, P.; Reinhard, F.; Parkan, M. Combining human computing and machine learning to make sense of big (aerial) data for disaster response. *Big Data* **2016**, *4*, 47–59. [CrossRef] [PubMed]

16. Dell'Acqua, F.; Gamba, P. Remote sensing and earthquake damage assessment: Experiences, limits, and perspectives. *Proc. IEEE* **2012**, *100*, 2876–2890. [CrossRef]

17. Aixia, D.; Xiaoqing, W.; Xiang, D.; Xiaoxiang, Y.; Long, W.; Yanfang, D.; Dingjian, J. Quantitative methods of rapid earthquake damage assessment using remote sensing and its application in yushu earthquake. *J. Catastr.* **2012**, *27*, 75–80.

18. Dawid, A.P.; Skene, A.M. Maximum likelihood estimation of observer error-rates using the EM algorithm. *Appl. Stat.* **1979**, 20–28. [CrossRef]

19. Dempster, A.P.; Laird, N.M.; Rubin, D.B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Ser. B* **1977**, *39*, 1–38.

20. Kerle, N.; Hoffman, R.R. Collaborative damage mapping for emergency response: The role of cognitive systems engineering. *Nat. Hazards Earth Syst. Sci.* **2013**, *13*, 97–113. [CrossRef]

21. Smyth, P.; Fayyad, U.; Burl, M.; Perona, P.; Baldi, P. *Inferring Ground Truth from Subjective Labelling of Venus Images*; MIT Press: Cambridge, MA, USA, 1995.