

DATA EXTRACTION OF STRUCTURAL BEHAVIORS AT THE MOMENT OF EARTHQUAKE OCCURRENCE FROM VIDEO CAPTURING CLIPS

Y. Mukai⁽¹⁾, T. Hattori⁽²⁾, Y. Mizushima⁽³⁾, Y. Nagano⁽⁴⁾

⁽¹⁾ Associate Professor, Kobe University, ymukai@port.kobe-u.ac.jp

⁽²⁾ Graduate Student, Kobe University, 155t048t@stu.kobe-u.ac.jp

⁽³⁾ Researcher, Takenaka Corporation, mizushima.yasunori@takenaka.co.jp

⁽⁴⁾ Professor, University of Hyogo, nagano@sim.u-hyogo.ac.jp

Abstract

When the Tohoku Taiheiyo-Oki Earthquake occurred at the east coast area of Tohoku region in Japan in 2011, lots of realtime video movies were recorded by mobile phone users and lots of video images of actual behaviors of buildings or infrastructures at just the moment of the natural disaster have been distributed through the web streaming site on the Internet. In this study, effective utilization of those video image resources for analyzing structural behaviors under the actual disasters is discussed. For this purpose, importance to provide data extracting techniques for quantification of the target's motions taken as the image information on the video screen is pointed out.

Emphasis is put on assuring the data accuracy and the degree of errors on the extracted data from the video photographs recorded by normal video cameras or mobile phones. Furthermore, this approach aims to utilize effectively the various kinds of video resources recorded at the moment of disasters as quantitative data resources to know structural motion of the target structures. As the actual example of the video movies at just the right moment of the disasters, the high-rise buildings' free vibrations after the earthquake in 2011 are selected. All the motion videos found for this study are available on the Internet streaming site, and those on-line videos have been uploaded by non-professional photographers who happened to take the video in personal by their mobile phone or handy camera devices. In this paper, investigations for the following items are carried out ^{[1], [2]}.

Firstly, exactness and resolution on the extracting data from motion pictures recorded by the normal video camera are evaluated. For this aim, the actual video images of the building model specimen on the shaking table tests are used. Through this investigation, the validity of the data extractions by using the normal video camera for sensing displacements is assured. The data property reproduced from the motion videos are investigated and the exactness of the identified motions are evaluated by comparing to the actual measurement data with the displacement sensors equipped directly on the testing specimen. In this study, quantitative results about the data accuracy and the available resolution on the motion captures from the normal video camera are shown.

Secondly, the actual video images streamed on the Internet about the free vibrations after earthquakes at the high-rise buildings in Tokyo while the Tohoku earthquake in 2011 are investigated. The data extractions of the displacement on the target building are evaluated. Through this study, as one of the importance for recording scenes, it is considered that the virtual fixed point should be caught on the screen with the target at the same time. It is assured that the probability to success extracting data can be improved, if there is some position to be able to deal with the fixed point on the video images. Furthermore, from the reproducing test supposing to the video recording with the mobile phones, importance of recording the fixed point on the video screen is also investigated and confirmed quantitatively.

Keywords: Motion Capture, Motion Identification, Mobile Phone, Motion Data Extraction, Displacement Sensor



1. Introduction

In general, observation of structural behaviors under earthquakes is carried out with the motion sensors equipped on the monitoring buildings permanently. It seems to be difficult to predict the exact timing that earthquakes will happen, so continuous monitoring for seismic motions has been operated on especially limited building structures. In 1995, when the Hyogoken-Nambu Earthquake occurred in Kobe in Japan, a few motion videos of buildings or their equipments at just the moment of earthquake happened to be recorded by the security video cameras that were fixed to the convenience stores or the ticket gate of the stations, etc. Those video images have been effective resources to evaluate and to consider the characteristics of local building responses during the earthquake, and it seems to be one of the feasible examples to use actually video cameras as monitoring devices for seismic observations.

More than a decade later, smart communication technology has let us wear mobile phones as ordinary everyday items and we can use easily video recorder applications in them at anytime we want. In 2011, when the Tohoku Taiheiyo-Oki Earthquake occurred at the east coast area of Tohoku region in Japan, lots of real-time video movies were recorded by mobile phone users and lots of actual performances at the moment of the natural disaster have been distributed through the streaming web site on the Internet. Those precious video records have been utilized as visual material for investigating various phenomena accompanying to the huge earthquakes, such as Tsunami, structural collapse, floatation of underground buried objects, outbreak of crack on ground surface, etc. On the other hand, it seems that those video photographs have not been used to convert into the form of the accurate motion data, but can be used only as visual resources for recognizing and understanding some mechanisms of the seismic behavior. So, it is thought to be important and beneficial to regard those streaming video movies as the real-time sensing data and to try to take available information for motion identifications out of those video photographs.

In this paper, we firstly estimate for the exactness and the resolution on the motion identification by using motion image records with the normal video camera. The video recordings at the shaking-table tests are selected for this aim. Those tests were carried out by using the compact-scale building specimen mounted on the shaking-table of Kobe University in Japan. This building specimen is composed by 4 lamped mass which are supported by the roller bearings. The video photographs taken by the normal video camera setting to the fixed position away from the shaking-table were picked up for the quantification of the structural dynamic motions. Through the operation of the 2D-motion capturing on the PC, the reproducing property from the video movies are investigated and the exactness of the identified motions are evaluated by the comparison to the actual measurement data from the displacement sensors equipped directly on the specimen.

As an application of those motion capturing techniques, the actual example of the system identification of the high-rise buildings is shown in this paper. Source movies are gotten from the Internet streaming site and those have been uploaded by the general persons. And those persons happened to stay on the upper floor of the high-rise building in Tokyo when the Tohoku earthquake in 2011 occurred. By using those video photographs, the displacement and the natural frequency of the target building are investigated. On the other hand, to consider accuracy and range of errors on the extracting data from the motion videos which are accidentally recorded on the photographing persons hands, reproducing tests are carried out by using small-scaled vibrating target. Simple moving mass specimen was used in this case study and the photographing person recorded the free vibration of the specimen with the video camera application of the mobile phone (iPhone 6S Plus). The extracted data from the motion videos are compared with the actual vibration of the target which is directly measured by the laser displacement sensor. The influence of the extracted data due to the hand shaking of the photographing person are also evaluated in this case study.

2. Data accuracy on displacement monitoring from video photographs

First of all, to confirm accuracy of the data extraction from monitoring with a normal video camera, the video recording of the shaking-table tests using a compact-scale building specimen is examined. Displacement of this building specimen is evaluated from video camera monitoring and those data are compared to the another data observed directly with displacement sensors.



2.1 Recording target for sensing test

4 stories of building specimen is used as the recording target for the video camera sensing tests. Dynamic behavior of this target is recorded by a normal video camera during the period that the seismic excitation is generated on the shaking-table. Fig.1 shows an outlook of the building specimen and its specifications. Each story with concentrating mass is placed up on its lower story and is supported by the roller bearings with very little friction. Every two neighboring stories are connected by the compression coil springs and the lowest story connected to the shaking table by the tension coil springs. This building specimen is supposed to be the seismic isolation building and the lowest story of this specimen is corresponded to the isolation layer by providing less stiffness than the upper stories. Natural period of the 1st dominant mode is designed to be about 3 second on this building specimen. As the input motion on the shaking table, one of the actual seismic wave data recorded at the Hyogo-ken Nambu Earthquake in Kobe in 1995 is used. The NS component of "JR Takatori" wave is the selected input data as seen in Fig.2 and it was recorded at the observation point near the railway station of Takatori by the Railway Technical Research Institute (RTRI) in Japan. The level of this original wave data was reduced into 60% and 68% and those were input to the target on the shaking table.



Fig. 1 – Outlook of monitoring targets (building specimen) and its specifications



Fig. 2 – input wave data (JR Takatori 1993 NS)

2.2 Data extracting process from the images recorded with the normal video camera

Standard Definition (SD) image movies are used for this evaluation, resolution of the screen image has 640 x 480 pixel length (having all 307,200 pixels). Photography frame per second is 30 fps. In this sensing test, it is supposed to use SD video camera for monitoring displacements of the target specimen. Fig.3 shows one scene in



movie recording the target. The vibration direction of the shaking table was set only single in right and left on monitoring screen.



Fig. 3 - Clipping frame of video screen and setting "marker" onto the target

Motion capturing from the video recordings is done by tracing the "marker" assigned on the target. To calculate and quantify actual motion of the marker on the screen frames, the 2D motion analysis software, "DIPP-Motion V 2D" (of the product made in Ditect co. in Japan) has been used in this study. This software can extract physical motion data i.e. displacement of the target on the video images by tracing the position of specified pixels area on every frames. In the followings, procedure of motion analysis is explained.

- Step-1) Set the marker including all required positions on the target to trace their motions. In which, the marker means the pixels area which is specified as identifiable pattern of image calling a "template". As seen in Fig.3, set the size of detective domain to search the position of the template on every continuous video frames. In this study, the template for the marker is set to 15 x 16 pixels rectangular area, and the detective domain is set to 29 x 30 pixels rectangular area (every 7 pixels around the template). 5 markers (from P0 to P4 as seen in Fig.3) are specified as the image pattern at the target on the screen to quantify every inter-story displacement of the building specimen.
- Step-2) Catch the motion of the template on every continuous video frames. This process is automatically operated by the software. At that time, if the area of the template had the deviation from its original point, it would be required to correct the template position manually.
- Step-3) Revise the pixel numbers into the actual length of the target. This can be done by assigning the actual order of length to the pixels between two points on the screen.
- Step-4) Set the reference point on the screen. The motion of every marker is output as the relative displacement between the marker and the reference point.
- 2.3 Comparison of sensing data from video to the displacement sensor's observation

Displacement data of the target building specimen are extracted from the video recordings and they are compared with the another data records by displacement sensors. Fig.4 shows the inter-story displacements of the target building specimen at the top story and the lowest story (the seismic-isolation layer), respectively. Video-sensing data on the top story are evaluated as the relative motion of the marker P4 for P3 and the lowest story's data are evaluated as the relative motion of the marker P1 for P0. As seen in those figures, it is assured that good reproducibility of the displacement data from the normal video photographs, and it is considered that the motion sensing with video camera can give moderate accuracy for the case such as tracing seismic motions



of buildings. Fig.5 shows relation between resolution of the monitoring range and the extracting data errors. The transverse axis of this figure is corresponding to the rate of the target's vibrating amplitude for the screen width. The ordinate axis indicates the error ratio for the extracting data from the video photographs vs. the direct sensing data with displacement sensors. As seen in this figure, when the amplitudes of vibration exceed about 0.15% of the screen width, it seems that error ratio of the extracting data from video photographs can keep to be around 10% or less.



(a) Inter-story displacements on the top (4th) story (b) Inter-story displacements on the lowest (1st) story

Fig. 4 – Comparison of displacement data (60% of JR Takatori wave input)



Fig. 5 – Relation between the resolution of the monitoring range and the extracting data errors

3. Data extraction from the actual building motion videos published on the Internet

Just after occurring the Tohoku Taiheiyo-Oki Earthquake in 2011, lots of videos catching various kinds of moments of disasters accompanying this earthquake have been published on the web streaming sites and most of them have been uploaded by general persons (non-professional photographer). Those uploaded resources can be accessed and watched in public without any permission, and movies of dynamic behaviors on some buildings or infrastructures can be also found out from those on-line resources. It seems that the photographing persons who recorded those kinds of scenes were motivated by getting extraordinary experience at that time accidentally, so that most of photographers could not assume the situation to shoot with video cameras beforehand. While most of uploaded video images provide lots of visual information which can effectively give us the situation of the disasters, it seems that those materials are difficult to utilize directly evaluating quantitative physical or mechanical behaviors caused by natural disasters on the targets.

However, it must be very useful for civil engineers and researchers to extract quantitative motion data about dynamic behavior of building structures or infrastructures during earthquakes from those resources of video photographs. While the earthquake ground motion continues, most of people must be desperate for



protecting their own safety against the disasters. So, it seems that most of them could not afford to be taking videos of what happened there immediately. Actually, it was very rare to record the moment of earthquake on videos, because photographing persons could get their action late a little after that right moment. During the earthquake in 2011, the long-period components of ground motions generated lots of high-rise buildings in the metropolitan area in Japan continuous vibrations for lots of minutes. Downtown in Tokyo in Japan places at about 380 km distance from the epicenter of this earthquake in 2011. Vibrations of high-rise buildings had been continued for a long time even after the ground motion converged. For such a reason, quite a few video photographs recording free vibration of high-rise buildings after earthquakes can be found on the web streaming site.

To search the video photographs on the Internet, two keyword "Tohoku Taiheiyo-Oki Earthquake" (in Japanese) and "High-rise building" (in Japanese) are set on the searching word entry in the Google ("https://www.google.co.jp/") and the corresponding video photographs are downloaded from the Youtube ("https://www.youtube.com/?gl=JP&hl=ja"). Among the available video photographs, some videos which don't have enough recording time of building's vibration are excepted, and then it is turned out that 36 movies can be available for this study. About 70% of those 36 movies recorded the high-rise building's vibrations and the others are of the low-rise buildings. Since the high-rise building's vibrations after earthquakes were continued for a long time, it seems that people could have more chance to record their vibrating behaviors. Various situations for the photographing people to record high-rise building's vibration were existed. Some were while walking and some were at stopping. Or, some were from the ground, or on the car, or some were from the building nearby. However, it seems that the videos recorded by moving photographer are not so suitable for the purpose to identify building motions on the video images.











In this section, some case studies to extract displacement data of the high-rise buildings from the streaming videos on the Internet are operated. As the sample of the video photographs of the high-rise building's vibrations to use for the data extractions, it was adopted one video movie which has comparatively long recording time and the target building has been kept in the screen with less hand shaking^[3]. In this paper, two scenes of the high-rise building's vibrations are evaluated. Fig.6 show the one of the screen shot from the capturing frames of the sample movies. This series of videos had been recorded from the person who placed at the other high-rise building nearby to the target building. It could be known that the target building on the screen as seen in Fig.6 is the steel construction building of 53 stories and that the photographing person had been placed on the 47th floor on the building nearby. The facade of the target building has 30 m span on the front bay and faces to photographing position mostly parallel. The distance between those two buildings are known as about 100 m. Fig.6(b) is the another scene recorded by the same photographing person. Total time of this movie has 105 seconds, and the first scene as seen in Fig.6(a) has been kept for about 16.5 seconds and the second scene has been kept for about 24.5 seconds.

The markers on the target building P1 are positioned on the first scene and the second scene as seen in Fig.6(a) and (b), respectively. In those movies, it seems that the video photographer handing his recording



devices and any equipment fixing it to the floors did not use for recording. So, it need to be considered that the target's vibration on this movie has been included the undesirable motion due to the photographer's hand shaking. And also, it seems that the building vibration recorded in those movies could not avoid containing the photographing position's movement because that the photographer placed on the floor of vibrating building. As seen in Fig.6, the unique building which can be identified its position on the backward far from the target building can be seen on the screen frame. The further marker P0 is set on this position of this unique building as seen in Fig.6, and this marker is considered as the motion of the approximate fixed point on the screen. This building is placed at about 5.1 km distance from the photographer's position. As the reason for this, both the video camera's motion due to the photographer's hand shaking or the floor's motions of the photographer's position must appear as the motion of this "virtual" fixed point on the screen.





(b) Scene 2 capturing the target





(a) Scene 1 capturing the target









Fig. 7 and 8 show the extracted displacements from the movie clips tracing the marker P1 and P0, respectively. Fig. 9 show the difference between those 2 markers P1 and P0. As seen in those figures, it is observed that the absolute motion on the screen of the markers include the video camera's motion due to the movement at the photographing place. On the other hand, as seen in Fig.9, the vibration depending on the target itself can be evaluated by considering the relative motion between the target position (P1) and the virtual fixed point (P0). Namely, it is assured that the displacement data of the free vibration of the target building can be



extracted well by using relative motion of the target position for the adequate fixed point. Fourier spectrum of the extracted displacement data for the scene 1 are shown in Fig.10. As seen in those figures, the spectrum curves by the extracted displacements from the movie clips tracing P1 include both motions of the target buildings and the video camera's motion due to the hand shaking of photographer or the movement at the photographer's position. However, by observing the Fourier spectrum curve from the relative displacement between P1 and P0 as seen in Fig.10 (c), the dominant natural frequency of the target itself are clearly recognized and it can be identified as about 0.18 Hz.



Fig. 10 – Fourier spectrum of the displacement data for the scene 1

4. Evaluation of video camera's movements by the photographer

By investigating the movie clips of the high-rise building's vibration after earthquakes which are available on the Internet, the situations of the target on the screen are quite different according to the photographing persons' position where he happened to be. In this section, to consider accuracy and range of errors of the extracting data from the mobile phone's video movies which are accidentally recorded by the photographing persons handing, reproducing tests are carried out by using small-scaled vibrating target. In this case study, simple moving mass specimen is used, and the photographing person records the free vibration of the specimen with the video camera application of the mobile phone (iPhone 6S Plus). Fig .11 show the experimental situation for this test. As seen in this figure, the size of the target width is 300 mm, and the video camera catches this target from the position of 1 m front away. At the rear side of the target, the other building' windows are recorded on the same screen to simulate existence of the approximate fixed point and the distance between the target and this rear building is 22 m. The actual vibration of the target is also directly measured by using the laser displacement sensor at the same time. The photographing person get the movies for the following 3 cases.

- Case-1) Video camera is fixed to the tripod stand, and the free vibration of the target is recorded under the condition without any hand shaking.
- Case-2) Video camera is swayed and rotated by the photographing person, but the target is not vibrated. Namely, this case study supposed to evaluate the ideal motion only due to hand shaking for the video camera.
- Case-3) Video camera is swayed and rotated by the photographing person, and the target is also actuated to perform free vibration at the same time.

Fig.12 show the comparison of the displacement data which are extracted from the video images and measured by the laser sensor for the Case-1. As seen in this figure, it is confirmed that those time histories of the displacements on sensing by using video images and actual displacement sensor show well agreement each other under the situation without hand shaking. Fig.13 (a) and (b) show the identified motion at the marker position P1 on the screen for the Case-2 comparing with the motions for the referential marker positions P0 and P2, respectively. In this Case-2, the target specimen was not vibrated, so the identified motion is all the "apparent" motion due to hand shaking of the video camera. Difference between the identified displacements for P1 and P2 are very little, because both marker positions are placed on almost same plane having the similar distance from the video camera (as seen in Fig.13(b)). On the other hand, the identified displacement for the P0 has some difference from the displacements for P1, because the position P0 is placed far from the target position P1 and the swaying motion including in hand shaking motion cause this difference (as seen in Fig.13(a)).





(a) Overview of testing situation

(b) Arrangement of the target





Fig. 12 – Comparison of displacement data (Case-1)







Fig. 14 - Comparison of relative displacement data extracted by using different fixed point (Case-3)



Fig.14 (a) and (b) show the comparison of relative displacement data extracted for the position P1 by using the different fixed point P0 and P2 for the Case-3, respectively. As seen in Fig.14 (b), it is found that errors on the identified displacement due to the hand shakings can be effectively reduced when the marker position for the fixed point can be placed on almost same plane with the marker position on the target. P1 and P2 have similar distance from the video camera as mentioned in Fig.11(b). On the other hand, as seen in Fig.14 (a), it seems that a certain extent of error on the identified displacement due to the hand shaking of the video camera cannot avoid at the situation that the fixed point is placed on the different plane for the position of the target about distance from the video camera.

Fourier spectrum of the extracted displacement data for the Case-3 are shown in Fig.15. In those figure, (a) is the spectrum curve calculated by direct sensing data with laser sensor, (b) and (c) are corresponding to the spectrum curves calculated by the extracted data of the relative displacement at P1 referred to P0 and P2, respectively. As seen in those figures, all the spectrum curves show the same dominant peak frequency about 1.2 Hz depending on the structural property of the vibrating model. Accordingly, it is assured that the structural property for the harmonic motion as like free vibration of the buildings can be technically extracted from the video images by using the FFT analysis, when the video camera's motion due to the hand shaking of photographer can be regarded as almost non-harmonic.



Fig. 15 - Fourier spectrum of the displacement data for the Case-3

5. Conclusions

As concluding remarks, accuracy and degree of error on the data extractions from normal video cameras are evaluated in this study. Main purpose of this study is put on quantifying the motion data from the various on-line videos recorded at the moment of disasters. As the actual example of the video movies of the disasters, the video images of the high-rise buildings' free vibrations after the earthquake in Tohoku Taiheiyo-Oki Earthquake in Japan in 2011 are selected. All the movie clips related to this study are available on the Internet streaming site, and those video movies have been uploaded by non-professional photographers who happened to take the video images in personal with their mobile phones or handy camera devices. It seems difficult to completely extract accurate motion data of the vibration's time history of the target caught on the video cameras, but those video movies must include valuable information for investigating the nature of the disasters. From such a point of view, the following investigations are carried out in this study.

At first, data accuracy of the normal video camera to use for sensing displacement on the target is discussed. For this aim, motion images of the building model specimen during the shaking table tests are evaluated. Through this investigation, the validity of the data extractions by using normal video cameras for sensing displacements is assured. In this case study, it is shown that motion capturing data can make quantity less than about 10% of errors and such level of accuracy can be gained unless the resolution of the target motions would become lower than 0.15% of amplitude for the screen width.

Secondly, from the investigation for the actual streaming videos on the Internet about the free vibrations of the high-rise buildings after earthquakes, the data extractions of the displacement of the specified target building are evaluated. Since various situation for the position of the photographing person placing at the timing of earthquake, most of video movies seems to have less information to operate the quantitative data extraction. In this study, as one of important items for recording scenes, it is mentioned that the virtual fixed point needs to



be recorded on the screen with the target at the same time. It can be assured that the probability to success for extracting data can be increased, if some position to be able to deal with the fixed point on the movies. Furthermore, even if the time history of the displacements include a certain extent of errors due to hand shaking by the photographer, it is observed that structural properties of the target building, such as natural periods, can be evaluated from the spectrum analyses.

At the last part in this study, data accuracy and errors included in the motion capturing from the video images recorded by mobile phones are examined with reproducing tests for video recording situations. This case study supposed to the situation that the video camera handed by the photographer, and that a certain extent of hand shaking cannot be avoided. For the extraction of the time history data of the displacement, even if the fixed point can be taken but when it is placed on the different plane with the position of the target, a certain extent of error on the identified displacement due to the hand shaking of the video camera cannot be avoided. However, through this study, it is assured that the structural property for the harmonic motion as like free vibration of the buildings can be technically extracted from the video images by using the FFT analysis.

6. Acknowledgements

This work was supported by JSPS KAKENHI Grant Number 26420552 and Grant Number 16H03124.

7. References

- [1] Hattori T., Mukai Y. (2015): A Study on Evaluations of Motion Reproducibility by Sensing on Video Capturing, *Summaries of technical papers of annual meeting 2015 Architectural Institute of Japan (AIJ)*, Structure-II, 837-838. (in Japanese)
- [2] Hattori T., Mukai Y. (2016): A Study for Utilization of Capturing from Video Clips of Actual Structural Behavior on a High-rise Building during Long-period Earthquake, *Trans. of Technical Reports of AIJ Kinki Branch*, 56 (Structural and Construction engineering), 213-216. (in Japanese)
- [3] https://www.youtube.com/watch?v=ACKMPD6MySs (2011): High-rise buildings in Shinjuku district during 3.11Tohku Taiheiyo-Oki Earthquake (Earthquake in Japan), *YouTube*.