8 台の HD ビデオカメラ映像の Structure from Motion 解析による 実大振動台実験の動的三次元モデル作成

内山庄一郎¹⁾、井上 公²⁾、今井 弘³⁾

- 1) 非会員、独立行政法人防災科学技術研究所
e-mail : uchiyama@bosai.go.jp
- 2) 非会員、独立行政法人防災科学技術研究所
e-mail : inoue@bosai.go.jp
- 3) 非会員、独立行政法人防災科学技術研究所
e-mail : imai@bosai.go.jp

要 約

2014年6月5日につくば市の防災科学技術研究所で行われたインドネシア型レンガ組積造住宅の実大振動台実験において、試験体の壁面の崩壊を8台のフルハイビジョンカメラを用いて撮影し、得られた動画からStructure from Motion技術を用いて試験体の三次元モデルを生成した。その結果、加振開始から試験体が崩壊するまでの間、1/24秒間隔で合計399の三次元モデルを得ることができた。これにより、試験体の挙動を時系列で、任意の視点から観察できる。本手法は振動台実験等の新たな観察手法として有用である。

キーワード: Structure from Motion (SfM)、動画、実大振動台実験、三次元モデル

1. はじめに

本研究では、実大振動台実験の試験体を複数台のビデオカメラで撮影し、SfM (Structure from Motion) 処理によって変位の時系列計測が可能な三次元モデルとして再現することを最終的な目的とする。本研究はこの目的を達成するための撮影手法、撮影機材、処理技術に関する基礎的な知見を得るための準備と位置付けている。ここでは、複数台のフルハイビジョン（以下、フルHD）ビデオカメラを用いて試験体を撮影し、その映像から三次元モデルの作成を目指す。実大振動台実験における既存の観測手法には、LEDマーカーや加速度計、高速度カメラによる撮影などがあり、信頼性の高い手法として確立しているが、これらは機器を設置した一部分のデータを取得する技術であり、試験体全体の三次元的な動きを連続的に捉えることは困難であった。本研究で使用するSfM技術は、複数のカメラ画像から被写体の三次元形状を復元する技術であり、低コストで三次元情報が得られることが特徴である。この技術は、様々な自然災害の調査研究に活用されている（内山・ほか、2014）。

2. 解析対象と実大振動台実験の概要

今回の解析の対象は、2014年6月5日につくば市の防災科学技術研究所で行われたインドネシア型レンガ組積造住宅の実大振動台実験の二つの試験体うち、耐震補強が実施されていない試験体（インドネシアの典型的な庶民住宅）の面外壁の挙動である。実大振動台実験では、1995年の兵庫県南部地震で記録

された地震波形をベースに、振幅を低減したものから実際の110%の加振まで計8加振を行い、各試験体の挙動把握、耐震性能の比較を実施した。

3. 解析手法

解析は主として撮影処理とSfM解析処理の二つの流れに沿って実施した（図1）。撮影処理は次の手順で行った（図1左）。8台のフルHDカメラを用いて試験体を毎秒24フレームで動画撮影する。カメラは図2のように設置した。カメラ群と被写体との距離は6～7mである。8台の動画の解析開始・終了時刻（フレーム位置）を特定するため、撮影を手動で開始した後、全てのカメラに写り込む位置でフレーム合わせ用LEDを点灯・消灯する（図3左・右）。LEDはナノ秒単位で点灯・消灯が可能のため、毎秒24フレームを撮影する動画であっても、複数フレームにまたがる残光がほとんど無く記録できる。したがって、LEDの点灯・消灯が切り替わるフレームを抽出することで全動画の解析開始・終了フレームを正確に特定することができる。

図1 解析手法のフロー

図2 カメラの配置

図3左 フレーム合わせ用LEDが点灯中のフレーム（カメラ番号2、フレーム番号1260）

図3右 LEDが消灯した左写真直後のフレーム（カメラ番号2、フレーム番号1261）

SfM解析処理は次の手順で行った（図1右）。動画の全フレームを静止画に分解する。次に、各カメラの全フレームから、フレーム合わせ用LEDが点灯から消灯に切り替わったフレーム（図3左・右）を探し、これを解析開始フレームとする。同様に、加振終了後、LEDが消灯から点灯に切り替わったフレームを検出し、解析終了フレームとする。ここで決定された解析開始フレームから終了フレームの間には、加振前の準備期間から加振後の立ち入り許可までの解析に不要な映像も含まれているため、非常にデータ量が多い。したがって、動画に写り込んだ画像を確認しながら、試験体の倒壊前後の399フレームを抽出する。倒壊の途上にある試験体の状態を写す例として、図4に1秒間分24枚のフレームを示す。

図4 1秒間に撮影された24枚のフレーム（カメラ番号6、フレーム番号6821～6844）

次に、これらの静止画を基にSfM処理を行い、三次元モデルを作成する。SfM処理では8台のカメラで同時刻に撮影された合計8枚の静止画（図5）から一つの三次元モデルを生成する（図6）。

図5 同一時刻の各カメラのフレーム（左上からカメラ番号1・2…、右下が8。フレーム番号6839）

左：図6 作成された三次元モデル（青い四角は動画撮影位置。フレーム番号6839）

右：図7 レンズキャリブレーション作業の様子。画面のチェッカーボードを数枚撮影する。

SfMは写真測量技術をベースとしているため、三次元モデルの作成には各カメラのレンズの歪みに関する情報（レンズキャリブレーション）が必要である。レンズキャリブレーションにはいくつかの手法があるが、ここでは平面（液晶ディスプレイ）に表示したチェッカーボードを複数枚撮影し、その結果から推定する方法を用いた（図7）。なお、本研究では28mmに焦点距離を固定したカメラを使用しているが、レンズ歪みについてはレンズ個体差があるため、全カメラに対してキャリブレーション処理を行った。

ここまでの処理を全フレームに対して実施することで、加振から倒壊に至るまでの24分の1秒ごとの連続的な試験体の三次元モデルが得られることになる。

4. 機材およびソフトウェア

ここでは、本研究に用いた撮影機材、照明条件、フレーム合わせ用LED、SfMソフトウェア、レンズキャリブレーションソフトウェア、解析用ワークステーションについて述べる。

撮影機材は次の要件を満たすものとした。

条件1：フルHD動画が記録できること。解像度1920×1080。

条件2：焦点距離（対角画角）が28mm（75度）～35mm（63度）程度の広角レンズを持つこと。

条件3：24fpsでの記録が可能なこと

最終的にRicoh GR（28mm単焦点レンズ）を6台、Nikon D7000（18-55mmズームレンズを28mmに固定して使用）とNikon D7100（17-35mmズームレンズを28mmに固定して使用）各1台の合計8台を用意した。

今回の動画はフレームレートを24 fps (frame per second) の設定で撮影した。つまり秒間24枚の静止画を撮影していることと同義である。この設定における最適な照明条件を決定するにあたり、理想的には倒壊によって落下する壁材等がブレない程度的高速なシャッタースピードを割り出して設定すべきである。しかし、今回使用した撮影機材は、本来は静止画を撮影するためのカメラであり、ユーザーが動画撮影時の詳細な撮影設定を変更することができない。また、本研究は実験主催者のメイン実験の都合上、投光器などの照明設備の使用ができなかった。そこで、実験室内の全ての照明を点灯し、窓の暗幕を解放し、資材搬入用の大扉も全開にして、実験室内が最大の明るさになるようにした。なお、露出計を保有しておらず、室内の明るさに関する計測値は不明である。

フレーム合わせ用LEDは、1m×3段の測量用アルミ標尺の最下段パーツの表面に、車装飾用のLEDテープライト（幅8mm×1m）を3列貼り付けて作成した。電源には12Vカーバッテリーを用い、点灯・消灯のためのタンブラースイッチを取り付けた。

SfMソフトウェアには、Agisoft PhotoScan Professional 1.0.4を用いた。また、8台のカメラ全てについてレンズキャリブレーションを行うため、同社のレンズ歪み補正ソフトウェアAgisoft Lensを使用した（図7）。

5. 結果

上記の手法で撮影と解析を行った結果を述べる。撮影処理（図1左）についてはバッテリー切れや記録メディア容量不足などのトラブルもなく無事に終了した。SfM解析処理（図1右）については、以下に詳細を述べる。

5.1 動画のフレーム分解から解析開始・終了フレームの決定まで

8台のカメラでそれ約9分30秒程度の動画が撮影された。これを静止画フレームに分割すると、カメラ一台あたり約1.4万枚、約110GBのビットマップ画像が得られた。8台のカメラで合計約11万枚、0.9TBのディスク容量となった。次に、この静止画の中からフレーム合わせ用LEDが点灯・消灯するフレームを見つけだし、解析開始から終了までのフレームを抽出した。この結果、11,575フレームが抽出された。これは482.29秒間の映像に相当する。なお、カメラ番号6および8のみ11,576フレームが抽出された。その他の6台のカメラより1フレーム多いが、全フレーム数から見た誤差の割合は非常に小さく、解析に支障がないと判断したため、最終フレームをカットすることで対処した。また、解析開始フレームを0、終了フレームを11,575として静止画ファイル名にフレーム番号を付与した。

5.2 SfM処理

前節で抽出した11,575フレームの中からSfM処理を行うフレームを抽出した。具体的には加振により試験体に動きがみられたフレームの前200フレームと後200フレーム程度の合計399フレームを抽出した。今回の例ではフレーム番号6601～6999の範囲が相当する。なお、加振開始のタイミングを動画から正確に捉えることができないため、試験体の動きはフレームのサムネイルの目視で確認した。

この399フレームについてSfM処理を行い、399の三次元モデルを作成した。作成した三次元モデルには写真から切り出されたテクスチャを貼り付けた（図8）。今回は外壁一面のみの挙動を捉えることを目的としてカメラを設置したため、試験体全体の三次元モデルは生成できていない。同様に、複数の動画に写らない領域については空白となっている（図8）。

図8 同一時刻に撮影された三次元モデル（青い四角は動画撮影位置。いずれもフレーム番号6833）

また、被写体ブレが生じている（図9、図10）。フレームに写っている画像から判断するとシャッタースピードが遅いことが原因と推測されるが、使用したカメラは全て静止画撮影用カメラであり、動画撮影の詳細な設定ができないため、具体的なシャッタースピードの値は不明である。この現象は事前に把握していたが、予備的研究であるが故に専用の撮影機材を用意することができないため、避けることができなかった。いずれにせよ、ブレが生じると被写体の正確な三角測量が難しくなるため、三次元モデルの精度は落ちる。また、今回は全カメラのフレームが完全に同期していることが保証されていない。フレームのズレも写真測量精度を劣化させる。さらに、フルHD動画は画像一枚当たり200万画素程度であり、一般的なデジタルスチルカメラの1/8～1/12程度の画素数しか持たない。そして今回の実験では、カメラ設置台数が8台と少ない。生成された三次元モデルの形状を観察すると、現実の試験体の形状よりもレンガブロックの角や試験体の壁に生じたクラックなどの本来鋭利な部分が丸められた様相を呈している（図10）。この現象を改善するためには、シャッタースピードを高速にすること、具体的には動作撮影設定を自由に変更できる撮影機材の使用、投光器による補助照明の使用、全カメラをフレーム単位で同期するためのフレームカウンターの使用が必要と考えられる。

左：図9 被写体ブレが生じたフレーム。落下する壁面がブレている（フレーム番号6833）

右：図10 被写体ブレによる三次元モデルの精度低下。不自然な凹凸が多い。（フレーム番号6833）

6. まとめと課題

2014年6月5日につくば市の防災科学技術研究所で行われたインドネシア型レンガ組積造住宅における耐震補強工法の実大振動台実験において、未補強の試験体の壁面の8台のフルHDカメラで撮影した。これらの動画からSfM処理によって三次元モデルを作成した。この結果、試験体の外壁が面外倒壊するまでの399フレームについて、1/24秒ごとの三次元モデルを生成した。これにより、時系列で試験体の挙動を三次元的に観察できた。

課題としては、SfM処理の結果には処理に使用する画像の品質と枚数が影響するため、フルHD動画の画素数では解像度に不足があること、実大試験体の完全な三次元モデル化には相当多数のカメラを必要とすること、被写体ブレを防ぐためフレームあたりのシャッタースピードが高くなるよう、十分に明るい照明条件が必要であること、カメラ間のフレームの同期を取ること、加振開始のタイミングを画像で視覚的に検出できる工夫が必要ながことが挙げられた。これらの課題は比較的容易に解決可能である。一方、LEDマーカーで得られるような、試験体上の点の変位時系列を任意の点について得るためには、動的三次元モデル上の点をフレーム毎に追跡する処理が必要になる。これはテクスチャのマッチング処理によって実現可能である。

本手法は今後、振動台倒壊実験や静的破壊実験の際の新たな観測手法として利用可能になるものと考えられる。

謝辞

匿名査読者より本研究の課題について有益なコメントをいただきました。記して感謝申し上げます。

参考文献

- 1) 内山庄一郎・井上 公・鈴木比奈子：SfMを用いた三次元モデルの生成と災害調査への活用可能性に関する研究、防災科学技術研究所 研究報告、81、2014年、pp.37-60.
- 2) 内山庄一郎：SfM-MVSによる三次元モデリング、日本地球惑星科学連2014年大会、2014年、H-TT33-05.
(受理：2015年3月9日)
(掲載決定：2015年10月15日)

Creation of a dynamic three-dimensional model by “Structure from Motion” from multiple full HD movies of a full-scale shaking table test

UCHIYAMA Shoichiro ¹⁾, INOUE Hiroshi ²⁾ and IMAI Hiroshi ³⁾

1) National Research Institute for Earth Science and Disaster Prevention
e-mail : uchiyama@bosai.go.jp

2) National Research Institute for Earth Science and Disaster Prevention
e-mail : inoue@bosai.go.jp

3) National Research Institute for Earth Science and Disaster Prevention
e-mail : imai@bosai.go.jp

ABSTRACT

A full-scale shaking table test of Indonesian masonry houses conducted in NIED on June 5, 2014 was filmed with eight full HD movie cameras. 399 video frames were taken for each camera from the beginning of the shaking to the collapse of the building and processed by a SfM (Structure from Motion) software. The behavior of the model can be observed in chronological order from arbitrary viewpoints. SfM can be a powerful tool to observe the dynamic behavior of the whole surface of a model for shaking table tests.

Keywords: Structure from Motion, full HD movie, full-scale shaking table test, three-dimensional model