

東日本大震災被災地を主な対象とする 市街地の時空間モデリング

岡谷 貴之^{1,a)}

概要：われわれは、2011年3月の東日本大震災以降、同震災の津波被災地を主要な対象に、市街地の時空間モデリング手法の研究を行ってきた。その被害の大きさ（被災地の広さおよび被害の甚大さ）および、都市機能の整った市街地が被災した点など、同震災は人類史上も未曾有出来事であり、これを対象に研究を行うことには多面的な意義があると考えている。モデリングに使用する主要なデータは、車両で該地域の道路を走行し、屋根上にとりつけた全方位カメラで得た大量の多視点画像である。本稿では、これまでにやってきた取り組みのうち、代表的なものを述べる。具体的には、多視点画像と地図を用いて建物の現存状況を推定する方法、異時刻の多視点画像を用いて市街地の形状変化を推定する方法、がれきの有無など市街地の状態を画像認識し、さらに上空視点画像と組み合わせて市街地全体の状態推定を行う方法、可視化のためのストリートパノラマの生成方法などである。

1. はじめに

2011年3月11日に東日本沿岸部を襲った大津波は、同地域の市街地に甚大な被害をもたらした。われわれ^{*1}は、同被災地を車両を用いて定期的に撮影し、この地域の状況を時系列で記録するとともに、こうして得られる大量の画像データを元に、市街地の時空間変化をモデリングする方法を研究してきた。目指すところは、この地域の震災直後の被害の状況、および長期にわたる復興の過程を可視化することにある。一例として、図1は、ある街の津波前後の様子を可視化したものを示す（3.1節の方法による）。

同震災は、その被害の大きさ（被災地の広さおよび被害の甚大さ）および都市機能の整った市街地が被災したことなどの点で、人類史上も未曾有の出来事である。その被害や復興過程を可視化する技術を開発できれば、そこから貴重な知見を色々得ることができると考えられる。また同震災に限らず、近年、気候変動等に起因すると考えられる広域災害が全世界規模で発生しており、そのような技術への期待と需要は高まりつつある。

このような用途に、コンピュータビジョンを活用可能な余地はいまだ大きいと考えている。広い視点から見ると東日本大震災の被災地は、(いささか不謹慎な物言いとなるこ

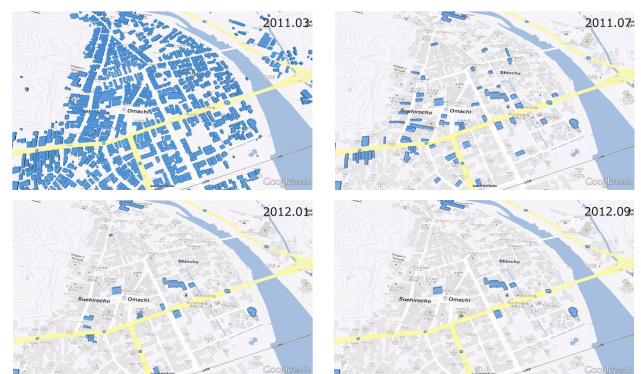


図1 市街地の時間変化の可視化例。3.1節の方法を使って、4つの日時の撮影画像から地図上の建物の有無を推定したもの。被災前（左上）の建物が、津波によって破壊され（右上）、その後の復旧工事により残った建物が撤去されてゆく様子（下段）が見て取れる。

とを恐れず言えば）この種の研究を行う上で適した対象であると言える。被災地と一口にいても、南北にきわめて広い地域にまたがっており、被った被害やそこからの復興の様子は、町ごとに規模や度合いなどの点で大きく異なっている。この多様さは、新たな技術の開発に理想的なテストケースを提供してくれる。

われわれの研究の目標は、経時に伴う市街地の幾何学的な変化を捉えることはもちろん、変化の意味を理解することにある。すなわち、街のどこがどのように変わったのか、そのとき何が起きたのか。例えば、復旧時にどれだけの量のがれきがいつ、どの場所で取り除かれたか、復興過程で

¹ 東北大学大学院情報科学研究科
宮城県仙台市青葉区荒巻字青葉 6-6-01

^{a)} okatani@vision.is.tohoku.ac.jp

^{*1} 東北大学情報科学研究科イメージ解析学講座（出口研究室および岡谷研究室）

の建築の完成状況から、ひいては経済活動の推移に至るまで、それぞれ捉えて可視化できるようにしたい。

なお、市街地の空間形状を計測する方法は、リモートセンシングやロボティクスなどで活発に研究開発が行われてきた(例えば [3], [5], [11])。われわれの研究は、コンピュータビジョンの手法を用い、画像から市街地の変化を意味的に理解しようとする点で、これらと一線を画すものと考えている。レンジセンサや上空視点情報のみを用いた「空間形状計測」では上述の目標は達成できない。

2. 被災地の映像記録

2.1 活動の概要

震災発生後約1か月後、われわれは東京大学池内研究室と合同チームを作って被災地に入り、被害の状況を車載全方位カメラにより記録した。開始後約一か月の撮影期間で、岩手、宮城、福島 の3県にまたがる被災地の撮影を完了した。その後も、数か月ごとの撮影を定期的に行い、復旧および復興の過程を記録することを継続している。これまでの活動で蓄積した映像データは、約40テラバイトに達する。この間撮影した画像の一例を図2に示す。

2.2 撮影システム

車両のルーフ上に全方位カメラ (Point Grey 社製 Lady-Bug3 あるいは Ladybug5) と GPS アンテナを取り付け、これらを車内に設置した PC に接続し、映像および GPS データを記録する。同カメラは 1616 × 1216 画素のカメラ6台が、鉛直方向の一部を除く全方位視野の映像を撮影するように配置され、毎秒10~20フレームで一枚の全方位分の画像を撮影・記録できる。われわれのシステムでは、車両のエンジン制御用コンピュータから車速信号を取り出し、これを用いて路上約2mごとに1枚の画像を撮影している。

3. 市街地時空間モデリング

3.1 SfM と市街地地図を用いた建物現存有無推定

最初に、市街地の地図と上述の方法で市街地をくまなく撮影した画像群を用いて、地図上の各建物の存在の有無を自動で推定する方法 [19] を説明する。この方法の処理のフロー図を図4に示す。市街地の全方位画像列、車両走行のGPSデータ、および地図の3つを入力にとり、同地図に記載されている各建物の存在の有無、あるいはその存在確率を出力する。

使用する地図は2次元のもので、建物を地上面に投影したものがポリゴンとして表現された一般的なものである。そこには3次元の情報や建物の高さの情報はない。類似の研究に市街地の精緻な3次元モデル(3次元地図)と多視点画像を用い、時間変化の有無を推定する方法の研究がある [17], [18]。われわれの方法はそのような3次元モデルを

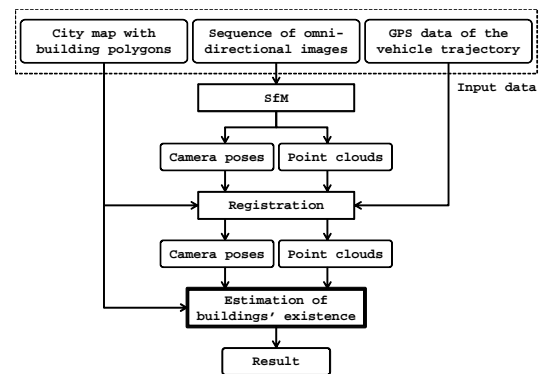


図4 Data Flow diagram of the proposed method.

前提としておらず、被災前の被災地のようにそういった3次元モデルが存在しない場所も対象になり得る。地図上のポリゴンが実際の建物の外壁に一致し、壁の高さは最低5m以上あると仮定して、各建物について粗い3次元モデルを生成し、その粗さを考慮した利用を行う。

図4にあるように、入力の1つの全方位画像列を使ってSfM [4], [7], [20] を実行し、市街地の様々な構造物の点群およびカメラ軌跡を得る。SfMは既に確立された技術だが、われわれは精度の鍵を握るバンドル調整の計算部分をロバスタ化 (= 対応点の誤りを自動的に排除できる) し、推定精度を向上を実現している (復元例を図3に示す)。こうして復元した点群とカメラ位置を、カメラ位置 (画像の撮影位置) のGPSデータを元に地図に対してレジストレーションする。そして、位置合わせされた点群、カメラ軌跡と地図 (から生成した建物の3次元モデル) を元に、地図に記載された各建物の存在を推定する (図4中太い線の箱)。

ここでは、点群と地図から起こした建物の構造をマッチングし、比較を行い、各建物の存在を推定する。もし建物があれば、その壁から多数の点が復元されるはずであり、そうでなければその周囲には点はないはずである。上述のように各建物は、建物を構成する壁を一辺とするポリゴンで表されている。そこで各建物を複数の壁に分解し、それぞれについて点群とのマッチングおよび比較を行う。その結果を統合することで建物の存在を推定する。

方法の考え方はこのように単純だが、そこには少なくとも二つの困難さがある。第一に、SfMの点群は、必ずしも建物壁面のみから生じるわけではない。電柱やフェンス、あるいは撤去された建物の基礎部分からも、紛らわしい点群が多数発生する。第二に、各建物 (とその壁) の可視性の問題がある。各建物がカメラにどのように写るかは、車両の走行位置に応じて変化する (図5)。他の建物に遮られれば映らないし、遮蔽するものがなくてもカメラから遠くになれば小さくしか映らない。また撮影時、車両は可能な限り小さな路地まで走行し、市街地の全貌を記録できるようにしてはいるものの、限界もある。被災直後には多くの道路が物理的に通行不可能であったし、その後も大規模な



図 2 車載全方位カメラで撮影した被災地の画像の例．上から岩手県大槌町，同赤浜地区，宮古市田老．

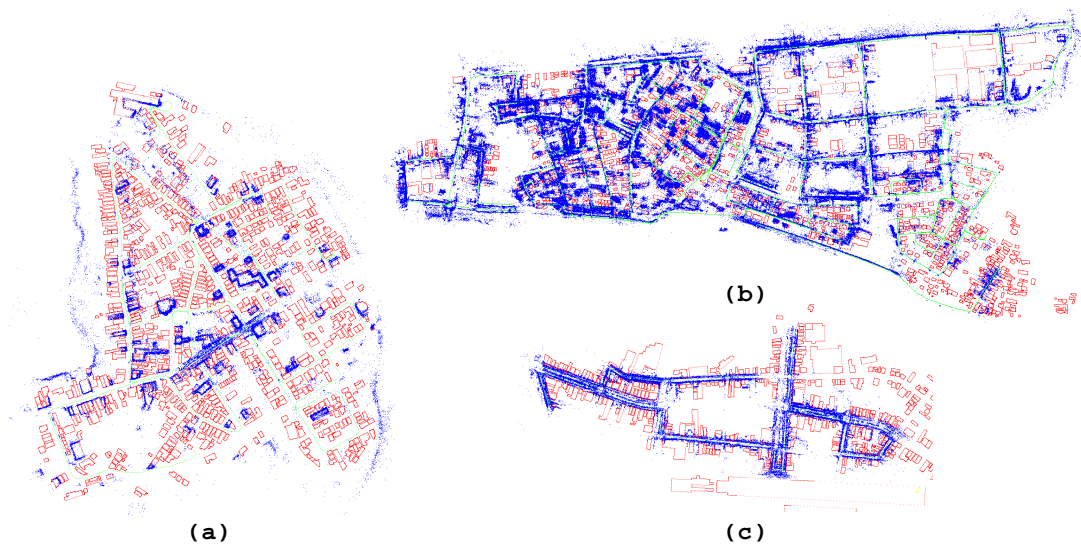


図 3 SfM による 3 次元復元の結果．復元された点群およびカメラ（車両）の軌跡を，それぞれ青と緑で，地図（赤）に重ねて表示している．(a) 岩手県大槌町（2,870 枚の画像から復元）．(b) 宮城野区の一部（画像数 18,002）．(c) 釜石市（画像数 1,266）．

復興工事により道路が閉鎖されていることも多いからである．したがって，すべての建物が至近距離から遮蔽なく撮影されていることはなく，そのような可視性の違いを考慮できなくてはならない．以上の困難さにより，ある建物の壁の周りに点群が多数発生しているというだけでは，その建物が存在するとは言えない．

このような問題を解決するため，二つの工夫を行った．一つは，SfM によって建物の壁からどのように点群が発生するかのモデルを構築し，利用することである．そこでは，遮蔽された壁や遠方の壁からは発生しないが，遮るものがない近くの壁からは多数発生することなど，各壁から発生する点群の規模を可視性と関連付けている．また，壁から



図 5 個々の建物の可視性の変化。カメラ軌跡（緑の曲線）に近い建物は大きく見えるが、他の建物に遮蔽されることもある。どのように画像に写るかは点群の規模を決定する。また建物の有無は、そのような遮蔽の有無を通じて相互依存の関係にあることに注意。

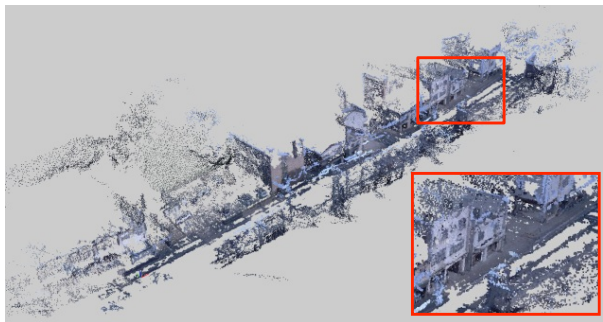


図 7 多視点ステレオで車載の多視点画像から生成した市街地の 3 次元モデル。復元不可能な部位も多く、精度もあまり高くない。

発生する点群の分布を捉えられるよう、点群を各点が小さな円盤になるように壁に投影したとき、その壁の被覆率を算出し、これを観測として使うこととした。壁から発生した点群は壁上で 2 次元的に広がり、一方壁近くにある電柱や建物の基礎から生じる点群はそうならないが、この観測量はそういったことも捉えることができる。最終的に、この観測量と建物の有無を確率的枠組みで結びつけ、事後確率最大化の問題として定式化した。

もう一つの工夫は、この枠組で建物の有無を推定する際、これを貪欲法による反復計算によって行うようにしたことである。可視性は各建物の有無に応じて変化するから、ある建物の有無を判定するには、周囲の建物の有無が分かっている必要があることになる。われわれの方法は、高い確信度をもって判定できる建物から順に判定を行って行くことで、建物間の相互依存関係を解決する。

以上の方法で得た、3 つの都市（の一部）の推定結果を図 6 に示す。

3.2 多視点幾何に基づく時間変化の検出

上の方法は、地図と画像を比較することで建物レベルの変化を検出するものである。この方法では、がれきのような地図上にはないものを対象とはできないし、任意の 2 時刻間の変化を捉えることもできない。一つの方法として、多視点画像を用いて多視点ステレオ (Multi-View Stereo, MVS) の方法で密にシーンの形状を得、これを 2 時刻で独



図 8 2 つの時刻の同じ場所を撮影した多視点 (ステレオ) 画像ペア。

立に行って結果を比較することが考えられる。MVS は、上空視点映像を元にした 3 次元地図の生成などに威力を発揮するが、車載カメラで撮影した画像列の場合、視点からの奥行きの変動が大きいこと、オクルージョンの多さ、市街地シーンのテクスチャの少なさが原因で、最新の手法を用いてもシーンの表面形状を正確に復元するのは難しい (図 7)。

文献 [12] の研究でわれわれは、3 次元形状を陽に求めることなく、シーンの形状変化を直接、推定することを考えた。そこでは図 8 に示すような、同じシーンを 2 つの時期に撮影した多視点画像列のペアを用いる。

画像からシーンの時間変化を自動認識する方法の研究は、サーベイランスを主なターゲットに、以前からよく行われてきた [11]。これらの多くは、シーンの変化部分を見つけるために、異時刻に撮影した画像を直接比較する。しかし図 8 のように、屋外を撮影する場合は、例えば天候による照明条件の変化など、撮影時の条件が大きく異なることが多い。そのため、それらの画像の濃淡値を直接比較してシーンの変化を検出するのは容易でない。

われわれはこの困難さを回避するため、同時刻画像間でのみ濃淡比較を行い、異時刻の画像間では幾何学情報 (視点間の幾何および対象シーンの奥行き) のやりとりだけをするようにした。つまり、異時刻の画像間で濃淡比較を行わないので、上述の撮影条件の問題とは無縁となる。

われわれの方法は次のように要約される。まず、一方の時刻の画像列を用いて、シーンの奥行きを画素ごとに確率分布として求める。そこでは画像の局所パッチの濃淡比較を行うが、MVS とは違って一意に奥行きを決定することはない。次にもう一方の時刻の画像列を用いて、同様に局所パッチの類似度を計算し、画像の各画素におけるシーンの奥行きが変化したかどうかを、先に求めておいた異時刻の奥行き確率分布と統合し、確率的に推定する。この方法を使って、震災直後とその数か月の 2 つの時期の間の変化を検出した。図 9 に結果の一例を示す。震災直後には

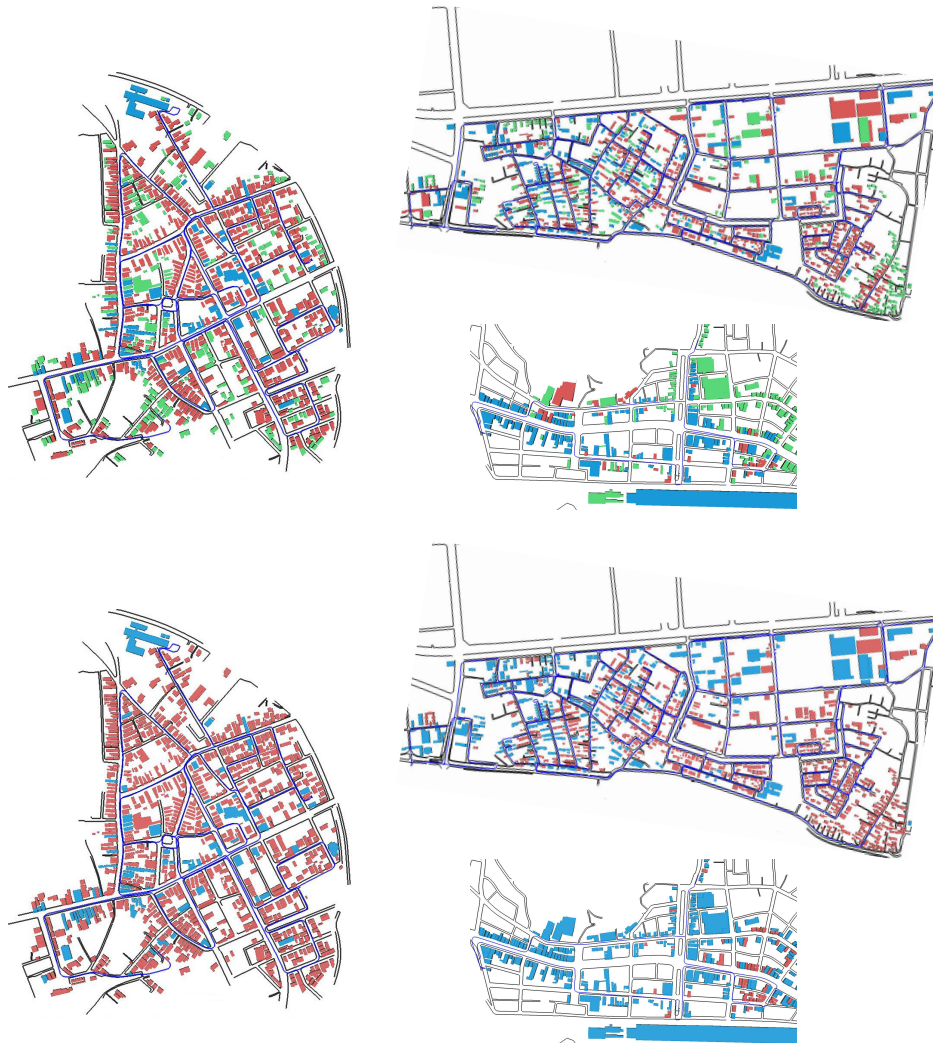


図 6 地図と画像を用いた地図上の建物の存在有無推定の結果．上が推定結果，下が真の状況．青が存在，赤が消失，緑が不明（十分な確信度では判定できない）をそれぞれ示す．

あったがれきが，復旧が一段落して除去されていることを，正しく検出できているのが分かる．

3.3 画像認識に基づく市街地状態推定

以上の方法はいずれも，市街地の空間形状変化を幾何学的に検出する方法であった．これとは異なるアプローチとして，機械学習ベースの画像認識を用いることが考えられる．市街地を撮影した多視点画像列に対し，色々な要素の認識を行うことで，空間形状に留まらない，市街地の状態やその変化を推定できる可能性がある．この考え方に基づいて行った研究が，文献 [13] である．

この研究は同時に，地上視点映像に上空視点映像と組み合わせる方法も提案している．地上視点映像と上空視点映像は，互いに直交する視線方向から，異なる解像度で市街地を撮影しており，相補的に活用できる，異なる種類の情報を含んでいる．上空視点映像は市街地の広範囲にわたる状態についての情報を保持する一方，視線方向の制約により，建物のひさしの影になる部分や，高架道路などで遮蔽

される地表面の映像は全く得られない．また，遮蔽されない部分でも地上視点映像に比べると解像度が低く，また天候に左右される欠点もある．一方で地上視点映像は，上述のような遮蔽領域であってもその近辺を走行できるかぎり撮影することが可能で，しかも，解像度は上空視点映像より高くなる．ただし撮影可能な地表面の範囲は上空視点映像よりずっと狭い．また，市街地のすべての路上を走行するには一定のコストがかかる．

以上を踏まえて，次のようなステップでがれきの認識を行う（図 10）．まず，地上視点映像を使って，各種画像特徴を用いた画像認識を行い，対象物（がれきなど）を検出する．次に，検出したがれき（の存在確率）を地表面に投影する．このとき，3.1 節で用いたのと同じ地図を用いて，建物による遮蔽を考慮した高精度な投影を行う．最後に，こうして得られたがれきの存在確率のマップを，上空視点映像およびその他の情報（digital elevation map など）とガウス過程に基づいて融合する．直感的には，地上視点映像がカバーする地表面については上空視点映像等を元に存

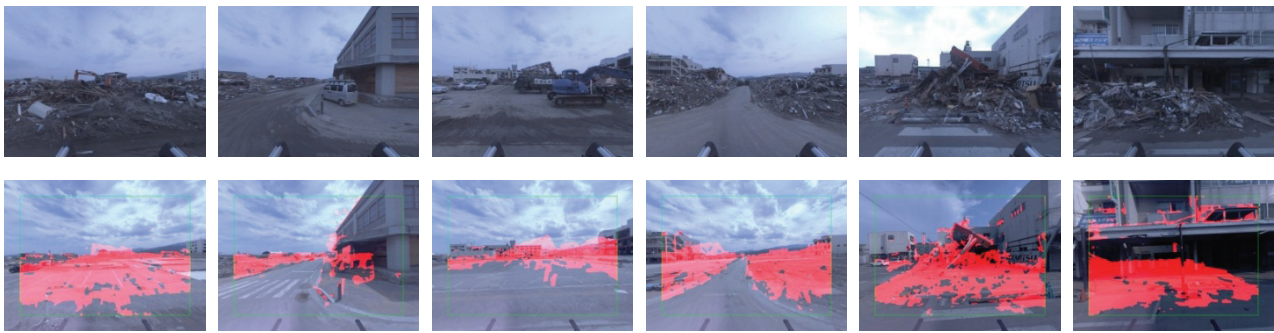


図 9 ある通りの被災直後に対するその数ヶ月後の変化の様子．復旧工事でがれきを取り除かれており，そのことが検出されている．

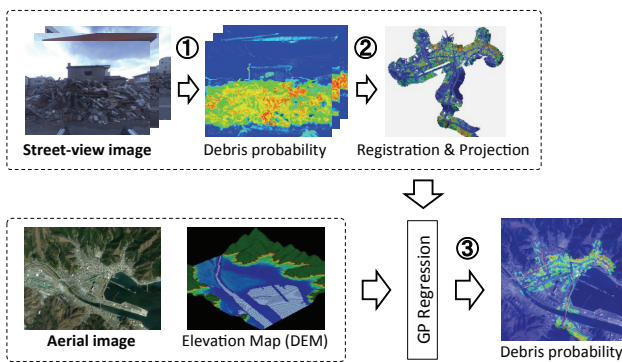


図 10 画像認識および地上視点映像と上空視点映像の融合に基づく市街地の状態推定方法の概要．推定する状態には，がれきの有無や植生など様々考えられる．

在確率のマップが平滑化され，そうでない領域については，上空視点映像等を元に，同マップが外挿されるように働く．こうして，地上視点映像がカバーできない地表面についても，一定の精度で推定（がれきの存在有無）が行われる．

岩手県釜石市を対象に以上の方法で得たがれきの推定結果を図 11 に，植生の推定結果を図 12 に示す．

4. 情報提示・可視化技術

4.1 ストリートパノラマ画像

被災地の被害は広い範囲にわたるため，広域空間の損害状況をわずかに一瞥するだけで把握できるような画像表示方法があると良い．Google 社のストリートビューに代表される全方位画像の提示は，シーンへの高い没入感を与えることはできる反面，広い範囲を見ようとすると頻りに視点を変える必要があり，閲覧時間が長くなってしまふ．

そのためわれわれは，長いシーンを一枚の横長のパノラマ画像として表示する方法を開発した [8]．多視点画像列を元にこのようなパノラマ画像を生成する方法は，長い間研究されてきた [1] が，これらのほとんどが，通りに面する建物の壁面が横方向に連続し，なるべく途切れる箇所が少ないこと，言い換えれば視点からの奥行き変化がなるべく小さいことを要請している．一方，ここで対象とするシーンは，通り沿いに建物がまばらにしか存在せず，空き地や



図 14 スマートフォンを用いた被災地の拡張現実表示．

がれきが散乱したような通り沿いの情景であり，この要請が満たされない．このことはパノラマ画像生成を本質的に難しくする．

これを克服するため，通りに面する建物の表面および地面のみを鮮明に描画し，それ以外の物体はぼかして提示する．それには，画像の各画素をラベリング（対応するシーンの点が建物表面，地面およびそれ以外の 3 種類のどれであるかを特定すること）すること，並びに「それ以外の物体」の像をぼかすことが必要である．前者はマルコフ確率場の最適化問題を解くことで，後者をライトフィールドフォトグラフィーの原理を応用することで，それぞれ解決した．図 13 に，生成したパノラマ画像の例を示す．

4.2 スマートフォンを用いた AR

最後に，身近なスマートフォンを用いて，被災地の現場でその場所の情報をユーザに提示する拡張現実 (AR)[14]を紹介する．図 14 にあるように，ある街の通りでスマートフォンをかざすと，仮想的な景色（例えば可視化された情報や過去の街並）が，スマートフォンの画面上に表示される．画面表示は実際に今ユーザが向いている方向に連動するように制御する．そのため，ユーザ（スマートフォン）の自己位置を GPS ならびにカメラで取得した画像によりその動きを検知し，切り出し領域をリアルタイムで変化させ，ユーザは画面を窓として仮想的な情景を見ていると感じられるようにする．ユーザは同じ方向を向いているのに端末の向きだけが変わった場合も考慮し，端末の画面側のカメラを使ってユーザが見たい方向を検出し，その方向の景色を表示することを実現した．

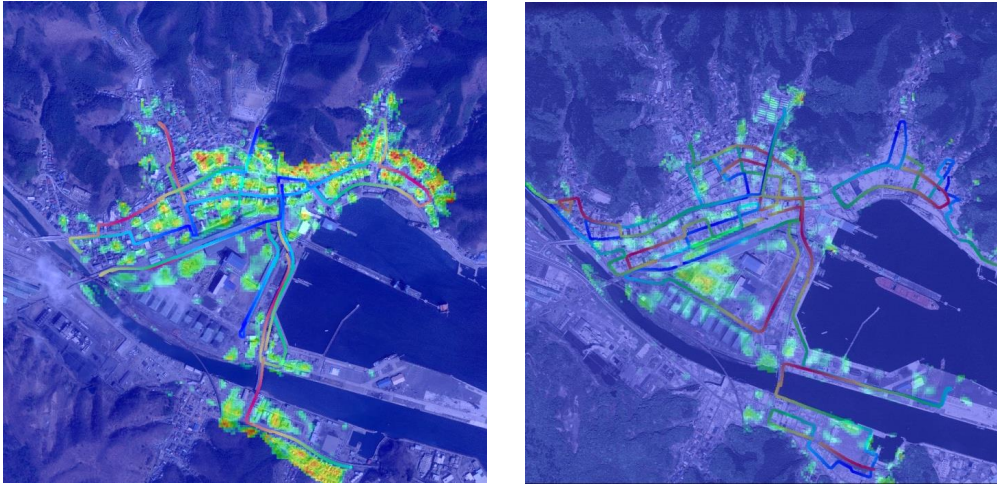


図 11 がれきの分布 (岩手県釜石市). 左: 2011 年 4 . 右: 2013 年 8 月 .

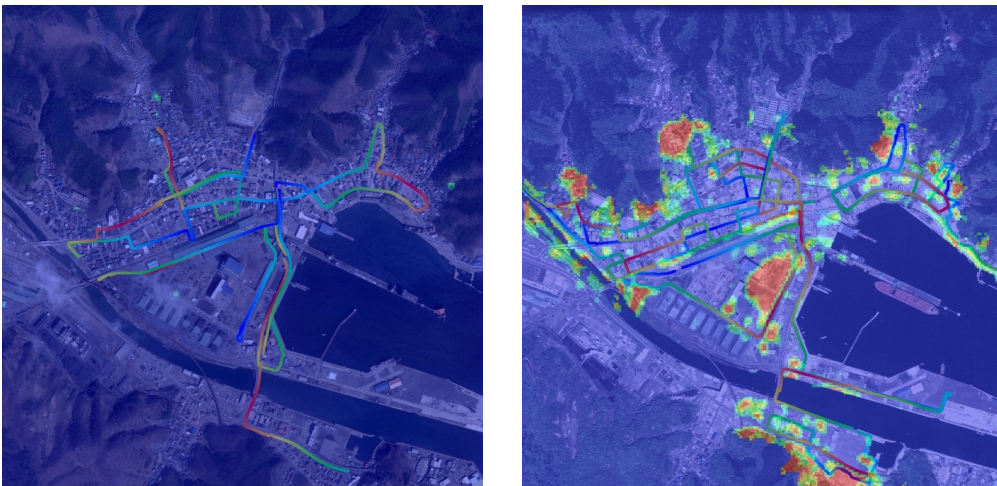


図 12 植生の分布 (岩手県釜石市). 左: 2011 年 4 . 右: 2013 年 8 月 .

5. まとめ

本稿では、われわれが研究してきた市街地の時空間モデリングおよび可視化のための手法のいくつかを紹介した。これらは、東日本大震災の津波被災地を主要ターゲットに据えている一方で、手法の汎用性や解くべき問題の普遍性も重視しつつ、研究開発を進めてきている。

被災地の復興が道半ばであると同様、われわれの研究もまたゴールはずっと先である。被災地の中には、将来の津波被害を避けるため、大規模な土地の嵩上げを行う地域も少なくない。中には、旧来の町並みをほとんど保存しない形で行われている再開発もある。その変化の大きさは、研究開始時点でのわれわれの想定を超えたものになりつつある。このような点を踏まえ、研究計画を軌道修正しながら、研究を継続しているところである。

謝辞

本稿で紹介した研究のいくつかは JSPS 科研費 25135701

ならびに 25280054 の助成を受けています。

参考文献

- [1] A. Agarwala, M. Agrawala, M. Cohen, D. Salesin, R. Szeliski. Photographing long scenes with multi-viewpoint panoramas. *ACM Transactions on Graphics*, 25(3):853–861, 2006.
- [2] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool. SURF: Speeded Up Robust Features. *Computer Vision and Image Understanding*, 110(3):346–359, 2008.
- [3] D. Crispell, J. Mundy, and G. Taubin. A Variable-Resolution Probabilistic Three-Dimensional Model for Change Detection. *Geoscience and Remote Sensing*, 50(2):489–500, 2012.
- [4] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision Second Edition*. Cambridge University Press, 2004.
- [5] A. Huertas and R. Nevatia. *Detecting Changes in Aerial Views of Man-Made Structures*. In *Proc. International Conference on Computer Vision*, pages 73–80, 1998.
- [6] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, P. Hanrahan. *Light field photography with a hand-held plenoptic camera*. Tech. rep., Stanford University Computer Science, 2005.
- [7] D. Nistér. An Efficient Solution to the Five-Point Rela-

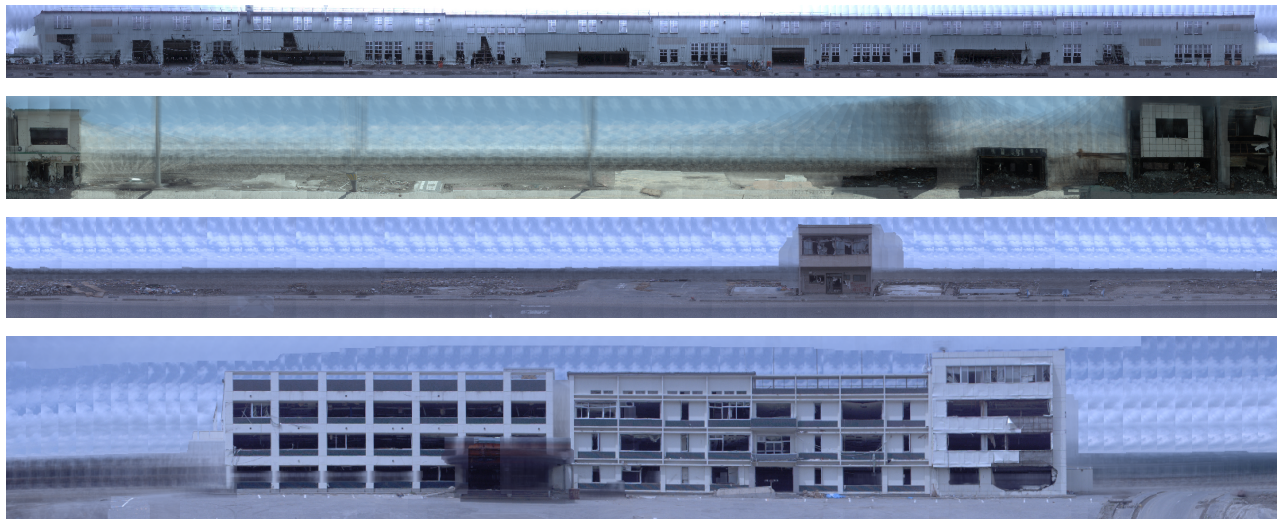


図 13 ストリートパノラマの生成結果．奥行きの変動が激しい被災地の通りを画像のボケを活用することで見やすく表示する．上から，釜石市に存在した倉庫，大槌町，田老，陸前高田市庁舎．

- tive Pose Problem. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 26(6):766–777, 2004.
- [8] T. Okatani, J. Yanagisawa, D. Tetsuka, K. Sakurada, and K. Deguchi. Creating multi-viewpoint panoramas of streets with sparsely located buildings. In *Proc. International Conference on Field and Service Robotics*, 2012.
- [9] M. Pollefeys, D. Nistér, J.-M. Frahm, A. Akbarzadeh, P. Mordohai, B. Clipp, C. Engels, D. Gallup, S.-J. Kim, P. Merrell, C. Salmi, S. Sinha, B. Talton, L. Wang, Q. Yang, H. Stewénius, R. Yang, G. Welch, and H. Towles. Detailed Real-Time Urban 3D Reconstruction from Video. *International Journal of Computer Vision*, 78(2-3):143–167, 2008.
- [10] T. Pollard and J. L. Mundy. Change Detection in a 3-D World. In *Proc. Computer Vision and Pattern Recognition*, pages 1–6, 2007.
- [11] R. J. Radke, S. Andra, O. Al-Kofahi, and B. Roysam. Image Change Detection Algorithms: A Systematic Survey. *IEEE Trans. Image Processing*, 14(3):294–307, 2005.
- [12] K. Sakurada, T. Okatani, and K. Deguchi. Detecting Changes in 3D Structure of a Scene from Multi-view Images Captured by a Vehicle-Mounted Camera. In *Proc. Computer Vision and Pattern Recognition*, pages 137–144, 2013.
- [13] K. Sakurada, T. Okatani, and C. Kitani. Massive City-scale Surface Condition Analysis using Ground and Aerial Imagery. In *Proc. Asian Conference on Computer Vision*, 2014.
- [14] 佐藤慎也，岡谷貴之，出口光一郎. 過去の街並みを可視化するスマートフォンを用いた拡張現実. 情報処理学会研究報告 Vol.2012-CVIM-182-No.3, 2012.
- [15] G. Schindler and F. Dellaert. Probabilistic temporal inference on reconstructed 3D scenes. In *Proc. Computer Vision and Pattern Recognition*, pages 1410–1417, 2010.
- [16] S. Seitz, B. Curless, J. Diebel, D. Scharstein, and R. Szeliski. A Comparison and Evaluation of Multi-View Stereo Reconstruction Algorithms. In *Proc. Computer Vision and Pattern Recognition*, pages 519–528, 2006.
- [17] A. Taneja, L. Ballan, and M. Pollefeys. Image based detection of geometric changes in urban environments. In *Proc. International Conference on Computer Vision*, pages 2336–2343, 2011.
- [18] A. Taneja, L. Ballan, and M. Pollefeys. City-Scale Change Detection in Cadastral 3D Models using Images. In *Proc. Computer Vision and Pattern Recognition*, 2013.
- [19] D. Tetsuka and T. Okatani. Detecting Building-level Changes of a City Using Street Images and a 2D City Map. In *Proc. IEEE Winter Conference on Applications of Computer Vision*, 2015.
- [20] B. Triggs, P. McLauchlan, R. Hartley, and A. Fitzgibbon. Bundle Adjustment - Modern Synthesis. In *Proc. International Conference on Computer Vision*, pages 298–372, 1999.