

令和元年度 卒業論文

画像列からの3次元形状復元への
M 推定導入による外れ値耐性の評価

宮崎大学 工学部 情報システム工学科

森 慎太郎

指導教員 椋木雅之

目次

1	はじめに.....	1
2	画像列からの3次元形状復元.....	3
2.1	Structure from motion.....	3
2.2	再投影誤差最小化法.....	4
2.3	最急降下法.....	6
3	M推定導入による外れ値対策.....	7
3.1	外れ値.....	7
3.2	M推定.....	8
4	M推定を利用した画像列からの3次元形状復元.....	11
5	実験.....	13
5.1	実験データ.....	13
5.2	パラメータの初期値設定.....	15
5.3	実験手順.....	17
5.4	実験結果.....	18
6	おわりに.....	25
	参考文献.....	26

1.はじめに

画像列が与えられた時、その中の物体の3次元形状を復元する手法として Structure from Motion(SfM)が知られている。SfMは、入力として画像間で対応する特徴点座標が与えられた時、画像を撮影したカメラの位置・姿勢と特徴点の3次元座標を同時に計算するものである。SfMは、撮影時のカメラの位置・姿勢の情報なしに3次元形状復元が行えるという利点がある。

SfMでは、画像間での対応点が与えられる必要がある。対応点を自動で求めることを特徴点マッチングと呼ぶ。特徴点マッチングでは、対応点に誤りが含まれることがある。この誤りは2種類に分類できる。1つは、特徴点の位置が画像間で多少ずれて検出されることにより生じる誤差で、これは形状復元においてはノイズとみなせる。もう1つは、画像間で本来異なる3次元点が誤って対応づけられる誤対応で、形状復元においては、外れ値とみなせる。SfMでは、ノイズは平均化されそれほど影響しないが、外れ値があると形状復元の精度が低下する原因となる。

そこで本研究では、外れ値が入っている場合でも復元の精度を向上させるために、外れ値に強いロバスト推定の1つであるM推定を

利用する。M 推定を導入することで、3次元形状復元をする際に外れ値の影響を低減できるか調査する。

2.画像列からの3次元形状復元

2.1 Structure from Motion

本研究では、SfM を用いて3次元形状復元を行う。SfM による3次元形状復元は、入力として画像間で対応する特徴点座標が与えられた時、カメラの位置・姿勢と特徴点の3次元座標を計算する問題である。

特徴点とは、画像中の角や線の交わり、エッジなどの他の点と区別できる固有性を持った点のことを指す。画像から、特徴点座標を検出する手法には、SIFT[1]やAKAZE[2]等がある。

検出した特徴点を画像間で対応づける処理を特徴点マッチングと呼ぶ。特徴点マッチングでは、同じ物体を写した2枚の画像が与えられる。それぞれの画像から検出した特徴点の内、同じ物体上の同じ点に対応する特徴点同士を対応づける。

対応づけが与えられたら、次節で説明する再投影誤差最小化法を利用して、カメラの位置・姿勢と特徴点の3次元座標を求める。

2.2 再投影誤差最小化法

再投影誤差最小化法とは、画像の特徴点とその特徴点の推定した3次元座標を画像上に再投影した点との誤差（再投影誤差）が最小となる様なパラメータ推定を行う SfM の手法である(図 1)。正規分布の誤差を仮定した下での、最尤推定の考え方に沿った SfM の定式化である[3]。処理の流れを以下に示す。

1. カメラの位置・姿勢と特徴点の3次元座標の初期値を仮に与える。
2. そのカメラの位置・姿勢で特徴点を透視投影により画像上に再投影する。
3. この再投影された点と入力された画像上での特徴点の位置との再投影誤差が最小になる様にパラメータを逐次更新していく。

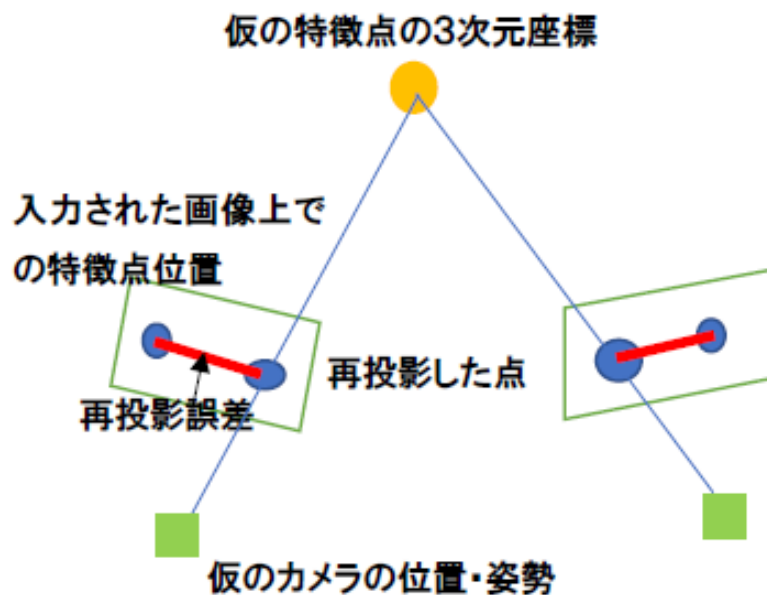


図 1 再投影誤差最小化法

再投影誤差最小化法は、以下の式で表せる。

$$\mathit{arg\ min} E(\vec{w}) \quad (1)$$

$$E(\vec{w}) = \sum_{f,p} |p[R_f \vec{s}_p + \vec{t}_f] - \vec{u}_{fp}|^2 \quad (2)$$

$$p(x, y, z)^\top := \frac{l}{z} (x, y)^\top \quad (3)$$

E を誤差関数と呼ぶ。カメラの位置・姿勢と特徴点の3次元座標は、ある世界座標系で表される。各画像を撮影したカメラの位置・姿勢は世界座標系における3次元の回転 R_f と並進 \vec{t}_f で表す。 R_f は、自由度3の回転行列である。 R_f のパラメータを並べたベクトルを \vec{q}_f とする。 t_f は、3次元の並進ベクトルである。 \vec{s}_p は、推定した特徴点の3次元座標である。 f はフレーム番号、 p は点番号を表す。世界座標系で表された3次元座標 \vec{s}_p は、 $R_f \vec{s}_p + \vec{t}_f$ によってカメラ座標系に変換され、透視投影により画像上に再投影される。透視投影は式(3)で表される。ここで、 l は焦点距離である。 u_{fp} は、特徴点の画像座標とその画像間での対応である。異なるフレーム番号で、同じ点番号のものが対応している。推定するパラメータ $\vec{q}_f, \vec{t}_f, \vec{s}_p$ を並べたベクトルを \vec{w} とする。再投影誤差最小化法は、 u_{fp} が与えられた時、 E を最小化するパラメ

ータ \vec{w} を求めることが目的である。

2.3 最急降下法

誤差関数 E を最小化するために、最急降下法を用いる。最急降下法とは、勾配 ∇E に基づいて、 E が小さくなる方向へパラメータを更新する手法である。最急降下法は以下の式で表せる。

$$\vec{w}^{(t+1)} = \vec{w}^{(t)} - \alpha \nabla E(\vec{w}^{(t)}) \quad (4)$$

$\vec{w}^{(t)}$ と $\vec{w}^{(t+1)}$ は更新前後のパラメータを表す。 α はパラメータ更新の大きさを決める微小な定数である。

3.M 推定導入による外れ値対策

3.1 外れ値

外れ値とは、あるデータ集合が与えられた時、そのデータ集合中に含まれるデータの大部分と統計的に傾向が異なる一部のデータのことである。外れ値は、必ずしも誤りとは限らないが、データの集計や分析の際にその存在が結果の精度を悪化させる可能性がある[4]。

本研究では、特徴点マッチングでの誤対応を外れ値と呼ぶ。SfMでは、画像間の対応点が与えられる必要がある。対応点を自動で取ることを、特徴点マッチングと呼ぶ。特徴点マッチングには、誤差と誤対応の2種類の誤りが含まれる。誤差は、特徴点が画像間で多少ズレて検出されることにより生じる。この誤差は、正しい対応づけに対して比較的小さなズレであり、再投影誤差最小化法の最尤推定の考え方で平均化されるため、結果に大きな影響を及ぼさない。一方誤対応は、画像間で、本来異なる3次元点が誤って対応づけられたものである。この誤りは、大きなズレとなる場合があり、最尤推定における正規分布誤差には当てはまらない。そのため、復元結果に大きく影響する可能性がある。

3.2 M 推定

M 推定とは、外れ値に強い評価関数を設定してパラメータを推定する手法である。データ $\{(x_i, y_i)\}$ が与えられ、 x と y の間に以下の式が成り立つとする。

$$\mathbf{y} = \mathbf{f}(\mathbf{x}; \boldsymbol{\theta}) \quad (5)$$

ここで $\boldsymbol{\theta}$ は関数 $f(x)$ のパラメータを明示したものである。

一般の最小 2 乗法では、以下の式を最小にするパラメータ $\boldsymbol{\theta}$ を推定する。

$$L = \sum \rho(\mathbf{d}_i) \quad (6)$$

$$\rho(\mathbf{d}) = \mathbf{d}^2 \quad (7)$$

$$\mathbf{d}_i = \mathbf{y}_i - \mathbf{f}(\mathbf{x}_i; \boldsymbol{\theta}) \quad (8)$$

この方法では、データに外れ値 (x_o, y_o) が存在すると、 \mathbf{d}_o の値が大きくなるので、正しいパラメータの推定ができなくなる。

M 推定によるパラメータ推定は、関数 $\rho(\mathbf{d})$ を変更することで外れ値の影響を低減して、パラメータ推定を行う。本研究では、式(9)を用いる。

$$\rho(\mathbf{d}) = \mathbf{1} - \exp\left(-\frac{\mathbf{d}^2}{\sigma^2}\right) \quad (9)$$

図 2,3 に M 推定、最小 2 乗法での $\rho(d)$ のグラフを示す。M 推定での $\rho(d)$ は、 d の絶対値 $|d|$ が大きくなると一定値に漸近している。外れ値は、 $|d|$ の大きい範囲にあるので、外れ値があったとしても影響を一定に抑えられる。一方、最小 2 乗法は、 $|d|$ が大きくなるほど、 $\rho(d)$ が大きな値をとるため、外れ値が大きな影響を与える。

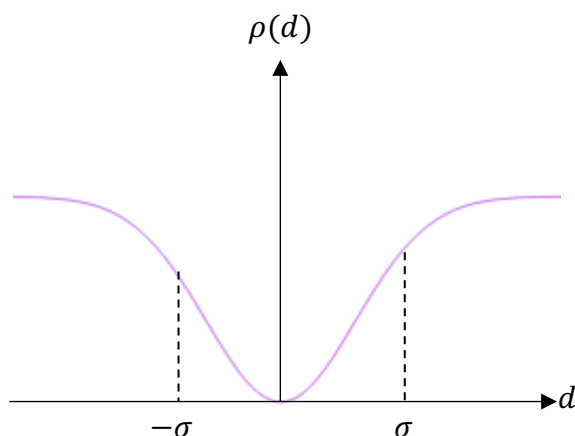


図 2 M 推定の $\rho(d)$

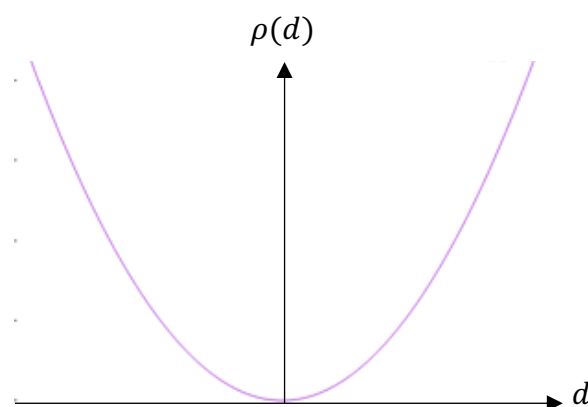


図 3 最小 2 乗法の $\rho(d)$

図 4,5 に M 推定、最小 2 乗法の $\rho(d)$ の微分 $\rho'(d)$ のグラフを示す。M 推定では、 $|d|$ が大きくなると 0 に漸近している。最急降下法を使う場合、 $|d|$ が大きくなると勾配 ∇E は 0 に漸近するので、パラメータ推定に大きな影響を与えない。一方、最小 2 乗法では、 $|d|$ が大きくなると ∇E も大きくなるため外れ値の影響を受けやすい。

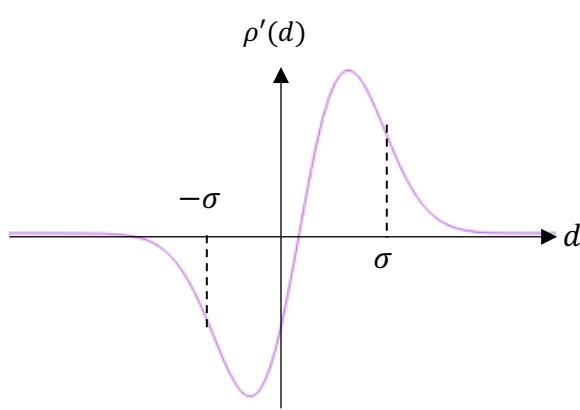


図 4 M 推定の $\rho'(d)$

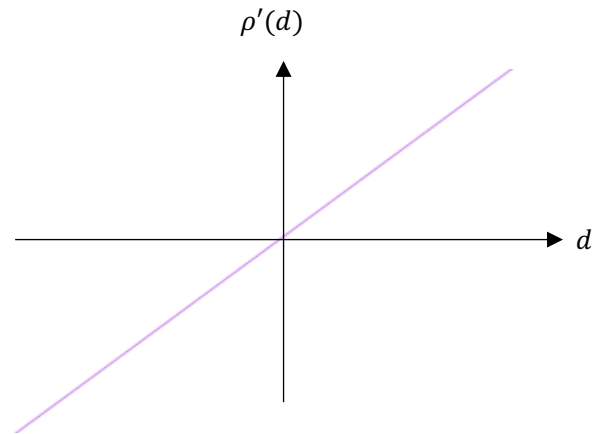


図 5 最小 2 乗法の $\rho'(d)$

4. M 推定を利用した画像列からの 3 次元形状復元

再投影誤差最小化法に M 推定を導入して、外れ値に強い形状復元を行う。式(2)の誤差関数 E に式(9)の $\rho(d)$ を組み合わせた次式を最小化するパラメータを求める。

$$F(\vec{w}) = \sum_{f,p} \left[1 - \exp\left(-\frac{d^2}{\sigma^2}\right) \right] \quad (10)$$

$$d(\vec{w}) = p[R_f \vec{s}_p + \vec{t}_f] - u_{fp} \quad (11)$$

ここで σ は、外れ値の範囲を決めるパラメータである。図 2 に示した様に、再投影誤差が σ に比べて大きい点は M 推定でのパラメータ推定にあまり影響を及ぼさない。 σ の値は、最尤推定における正規分布誤差の標準偏差程度が妥当と考えられるが、その値はわからない。また、再投影誤差の値は収束計算における解の収束の度合いにも影響を受ける。そのため、収束計算において最初から σ の値を小さな値にすると再投影誤差が大きいため、ほとんどのデータが外れ値とみなされてしまう。

そこで、本研究では、収束計算の最初の段階では σ を比較的大きな値 σ_0 と設定し、収束計算を一定回数 $\ell = 1000$ 回繰り返す毎に σ を $1/\sqrt{2}$ 倍することで、徐々に小さな値にしていく。

式(10)の最適化には、2.3 節の最急降下法を用いる。Fの偏微分は式(12)で表される

$$\frac{\partial F}{\partial w} = 2d \frac{\partial d}{\partial w} \frac{1}{\sigma^2} \exp\left(-\frac{d^2}{\sigma^2}\right) = \frac{1}{\sigma^2} \exp\left(-\frac{d^2}{\sigma^2}\right) \frac{\partial E}{\partial w} \quad (12)$$

したがって、

$$\nabla F = \frac{1}{\sigma^2} \exp\left(-\frac{d^2}{\sigma^2}\right) \nabla E \quad (13)$$

となる。即ち、M 推定での最急降下法は、最小 2 乗法での最急降下法での更新式に、重み $\frac{1}{\sigma^2} \exp\left(-\frac{d^2}{\sigma^2}\right)$ をかけたものである。また、この式から σ を小さくすると $\frac{1}{\sigma^2}$ により重みが大きくなることがわかる。M 推定の最急降下法では、更新量を一定に抑えるために、 σ を $1/\sqrt{2}$ 倍した際には、式(4)の α を 1/2 倍する

5.実験

M 推定の導入により、外れ値の影響を低減できるか調査するために精度評価の実験を行った。

5.1 実験データ

精度評価のために、3次元モデルから正確な対応点データを作成した。図6に利用したデータの3次元的な配置を示す。復元対象となる3次元モデルは、半径1の球とした。球面上に一定間隔で特徴点を300点とった。この球を撮影する仮想的なカメラを30台設置した。カメラは球の中心を斜め上から見下ろす姿勢で、等間隔で全周から撮影する様に配置した。3次元空間の各特徴点を各カメラに投影し、特徴点の2次元座標を求めた。図7に投影された2次元点を示す。2次元座標は $[-1,1]$ の間の範囲内になる様に正規化した。また、特徴点の対応関係も正確に与えた。

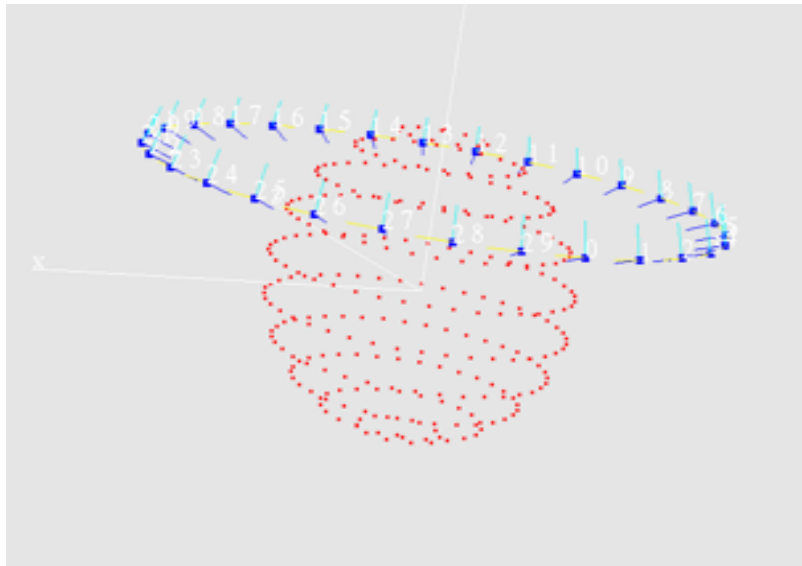


図6 正解データ

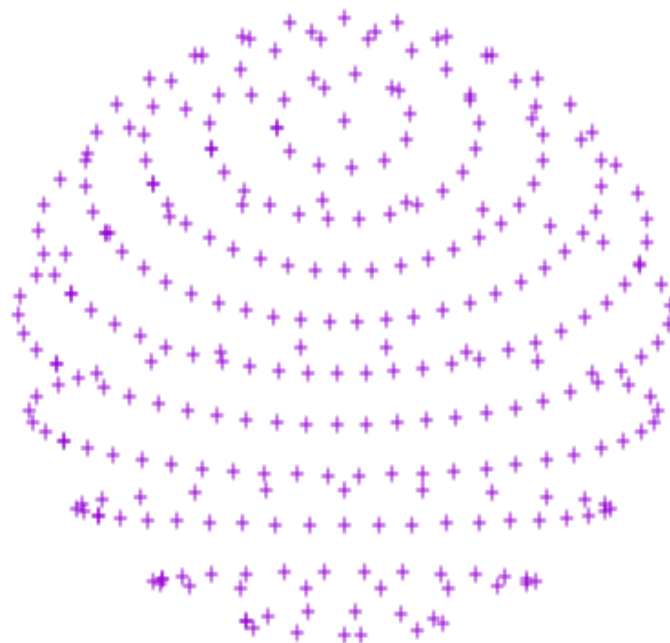


図7 投影された2次元点

こうして得られた 30 画像×300 点=9000 点の二次元座標に対して、人工的に外れ値とノイズを付加した。

外れ値は 2 通りの方法で付加した。外れ値 1 は、3 次元特徴点の内 100 点について、投影された 2 次元座標の x, y の値を入れ替えた点を新たに作成し、追加した。2 次元座標の x, y の値の入れ替えは、元のデータとは鏡像の関係にあり、3 次元の剛体変換では整合性が取れないため、外れ値となる。外れ値 2 は、3 次元特徴点を新たに 100 点追加し、各画像の 2 次元座標を $[-1, 1]$ の一様乱数により与えた。これらの点は 3 次元的な関係を見捨てており、外れ値とみなせる。

ノイズは、標準偏差 0.05 の正規分布ノイズを発生させ、各特徴点の 2 次元座標に加えることで、付加した。

5.2 パラメータの初期値設定

再投影誤差最小化法では、カメラの位置・姿勢と特徴点の 3 次元座標の初期値を与える必要がある。本論文では、注視回転による初期値 (図 8[3]) を用いる。特徴点の 3 次元座標 \vec{s}_p の初期値は、全て原点 $(0, 0, 0)$ とした。カメラの位置は、原点を中心とした円周上に配置する。

カメラの姿勢は、全てのカメラが原点方向を向くとする。初期値を式で表すと次式となる。

$$\begin{aligned}
 \mathbf{R}_f &= \mathbf{R}_{\omega f}^y, \mathbf{t}_f = (\mathbf{0}, \mathbf{0}, r)^T \text{ for all } f \\
 \mathbf{s}_p &= (\mathbf{0}, \mathbf{0}, \mathbf{0}) \text{ for all } p
 \end{aligned}
 \tag{14}$$

$\mathbf{R}_{\omega f}^y$ は y 軸 (図 8 の紙面鉛直方向) 周りの角度 ωf の回転を表す行列とする。 ω は角速度で、30 台のカメラが円周上で等間隔とみなせるように、 $\omega = 2\pi/30$ とする。 r はカメラの円軌道の半径である。

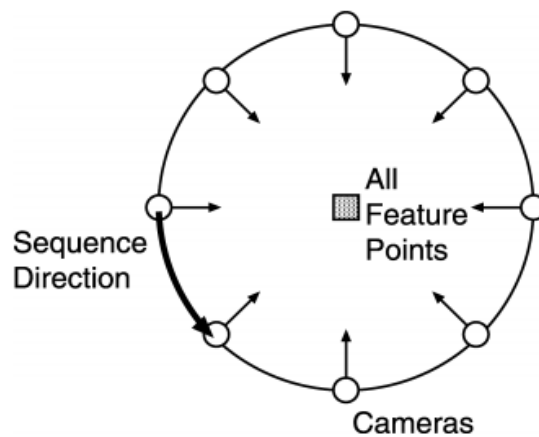


図 8 注視回転による初期値 [3]より引用

5.3 実験手順

外れ値の種類 2 通り × ノイズの有無 2 通り × M 推定の有無 2 通りの計 8 通りの実験を行った。M 推定なしは、式(9)の代わりに式(7)を用いた方法である。M 推定ありでは、 $\sigma_0 = \sqrt{0.5}$ とした。いずれの方法でも最急降下法を利用し、合計 5000 回のパラメータ更新(式(4))を行った。5000 回終了時には、式(2)の値はほとんど変化しなくなり、計算は収束していた。最急降下法で使用する α の初期値は、 $\alpha = 0.001$ と設定した。

5.4 実験結果

作成した2次元データを用いて8種類の復元を行った。図9-16に、上記で説明した8通りの復元結果を示す。

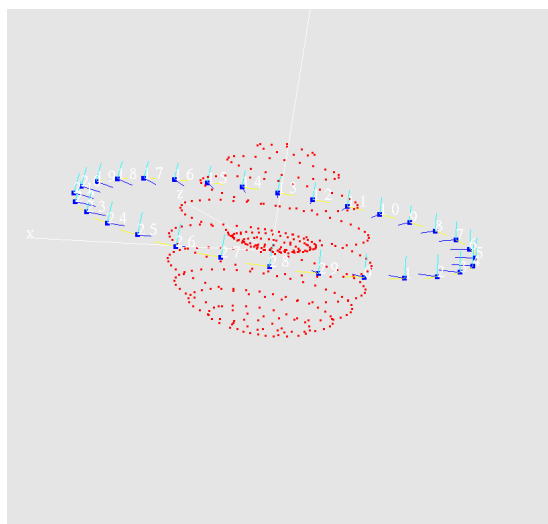


図 9 外れ値1 ノイズなし M推定なし

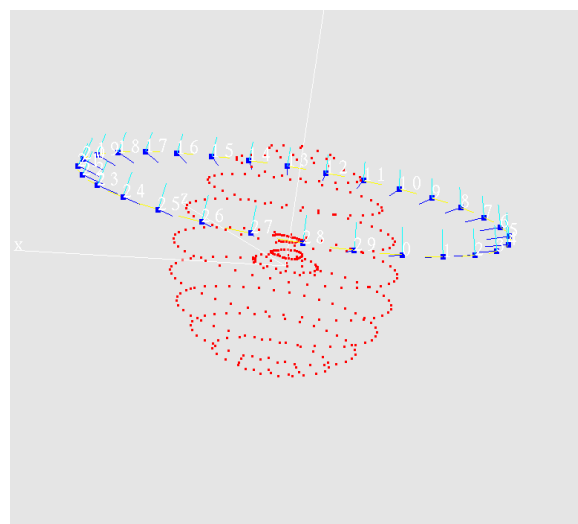


図 10 外れ値1 ノイズなし M推定あり

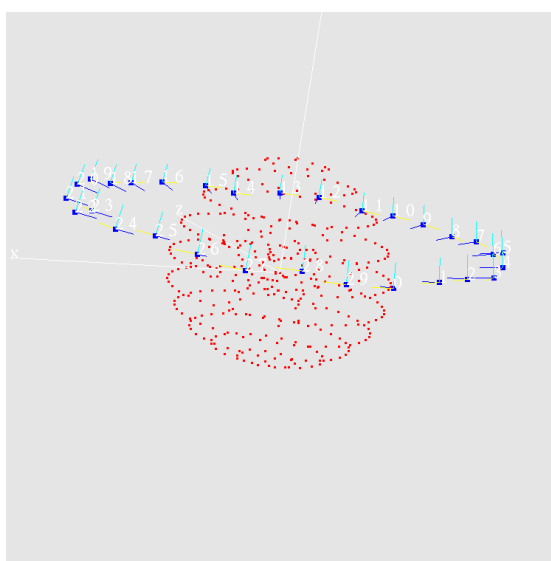


図 11 外れ値2 ノイズなし M推定なし

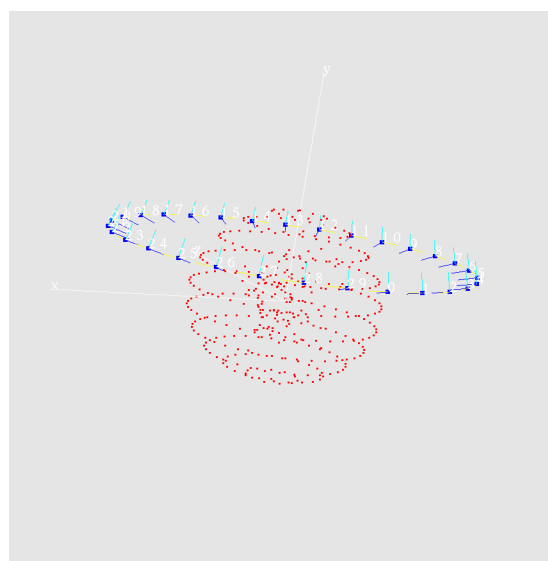


図 12 外れ値2 ノイズなし M推定あり

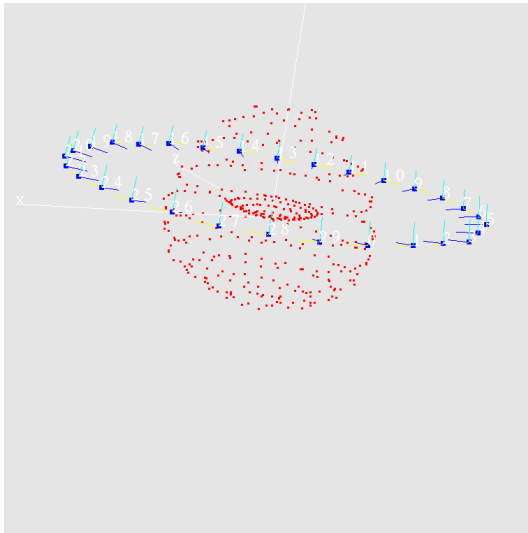


図 13 外れ値1 ノイズあり M推定なし

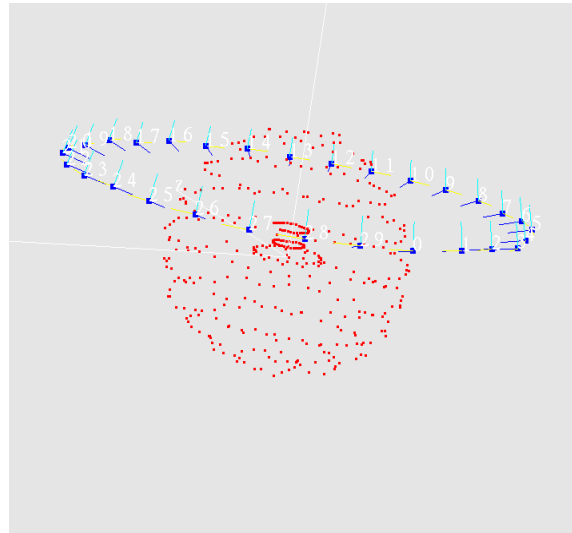


図 14 外れ値1 ノイズあり M推定あり

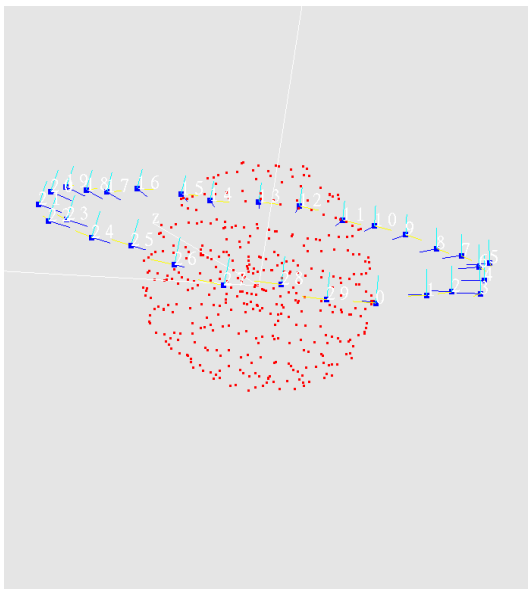


図 15 外れ値2 ノイズあり M推定なし

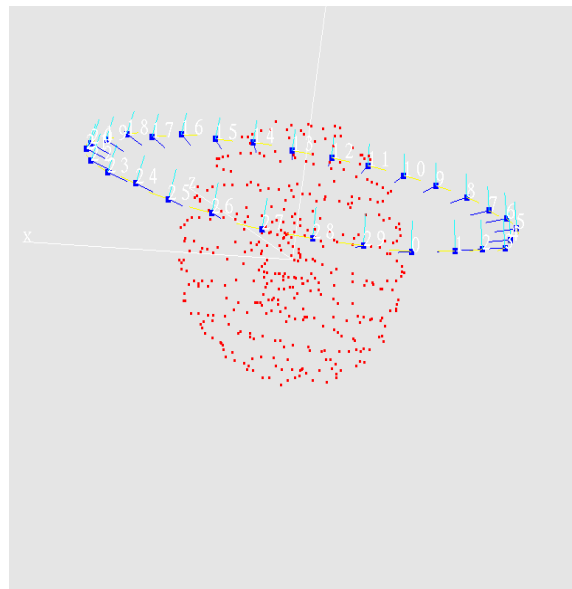


図 16 外れ値2 ノイズあり M推定あり

図の赤い点が特徴点、青い点がカメラ位置、青い点から出ている線がカメラの視線方向を示す。図 9,11 のノイズなしに関しては、球面の形が歪んで復元されている。また、カメラの位置も上から見下ろす様になっておらず、外れ値の影響を受けたと考えられる。一方、図 10,12 に関しては、図の中心付近に外れ値が固まっているが、球面上の点は良好に復元できている。図 13,14 も同様になっており M 推定を使った方がより良く復元できている。一方、図 15,16 では外れ値の影響を受けているが、図 16 の M 推定ありの方が球面上に点が集まっており、図 15 の M 推定なしに比べて高精度な復元が行われていると言える。

復元結果を定量的に評価するために、復元結果での再投影誤差の値を比較した。結果を表 1,2 に示す。表中の E_2 は外れ値以外の 300 点についての再投影誤差の和、 E は全ての点の再投影誤差の和である。いずれの結果でも、 E の値は M 推定の有無に関わらず同程度となっており、外れ値も含めて同程度に解が収束していることがわかる。一方、 E_2 の値は、M 推定ありでは M 推定なしに比べて最大 1/13 程度小さくなっている。これは、外れ値以外の点が外れ値の影響を受けずに正確な位置に復元できていることを示している。

表 1 再投影誤差 (外れ値 1)

M 推定	ノイズあり		ノイズなし	
	E_2	E	E_2	E
なし	134.8	511.1	92.9	469.0
あり	52.1	608.3	6.7	569.5

表 2 再投影誤差 (外れ値 2)

M 推定	ノイズあり		ノイズなし	
	E_2	E	E_2	E
なし	70.0	1952.1	28.2	1910.5
あり	51.1	2033.7	6.4	1989.5

外れ値の影響低減をより詳細に確認するために、外れ値以外の点について、再投影誤差の頻度分布を調べた。結果を図 17-24 に示す。グラフの横軸は、再投影誤差の 2 乗であり、縦軸は再投影誤差の 2 乗がその区間にある画像上の特徴点数である。図 17-20 がノイズなしの結果である。図 17,19 の M 推定なしの場合では、外れ値によって再投影誤差の 2 乗が大きい範囲に分布しているが、図 18,20 の M 推定ありの場合では、全て再投影誤差が小さい範囲に分布している。よって、外れ値の影響を全く受けていないことがわかる。

図 21-24 がノイズありの結果である。図 21,22 では、ノイズによって M 推定ありの場合でも再投影誤差の 2 乗が大きい範囲のデータもあるが、M 推定なしに比べては小さい範囲に分布している。図 23,24 については、M 推定なしの場合でも再投影誤差の 2 乗が小さい範囲に多く分布しているが、M 推定ありの方がより再投影誤差の 2 乗が小さい範囲に分布していることがわかる。以上のことから M 推定の導入により外れ値の影響を低減できていることがわかる。

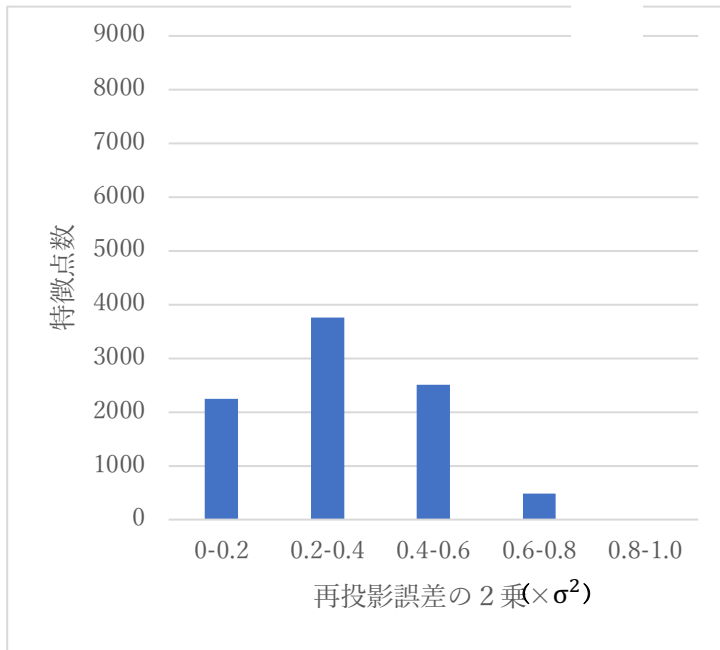


図 17 外れ値 1 ノイズなし M 推定なし

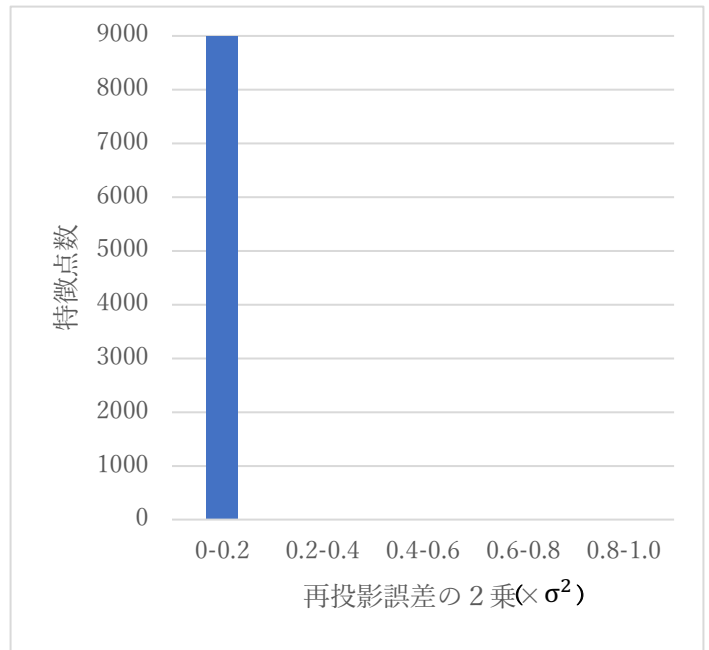


図 18 外れ値 1 ノイズなし M 推定あり

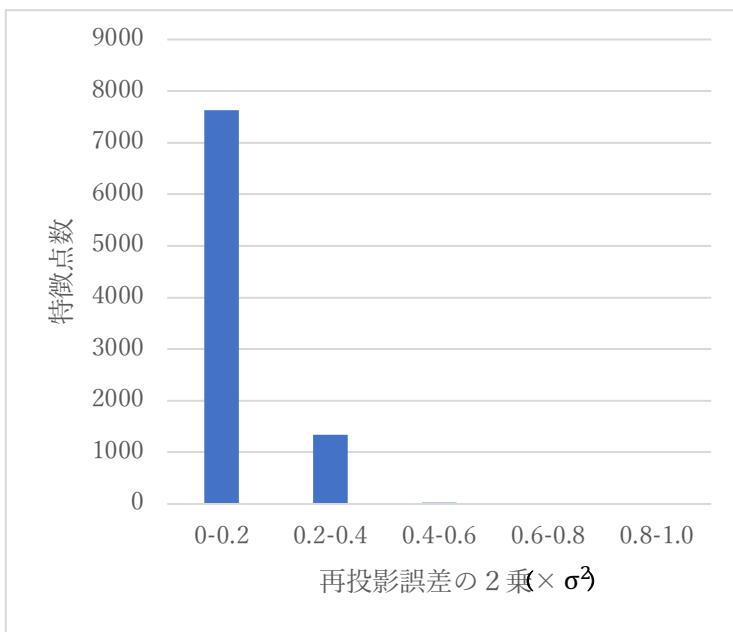


図 19 外れ値 2 ノイズなし M 推定なし

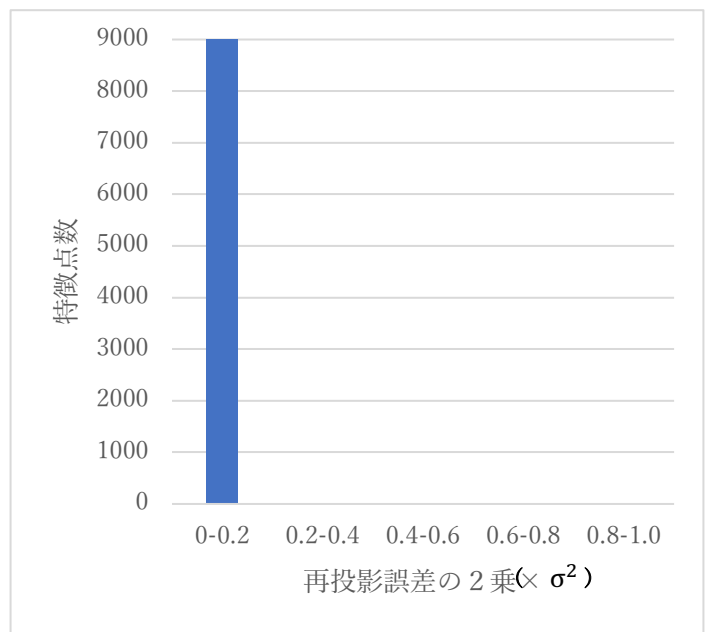


図 20 外れ値 2 ノイズなし M 推定あり

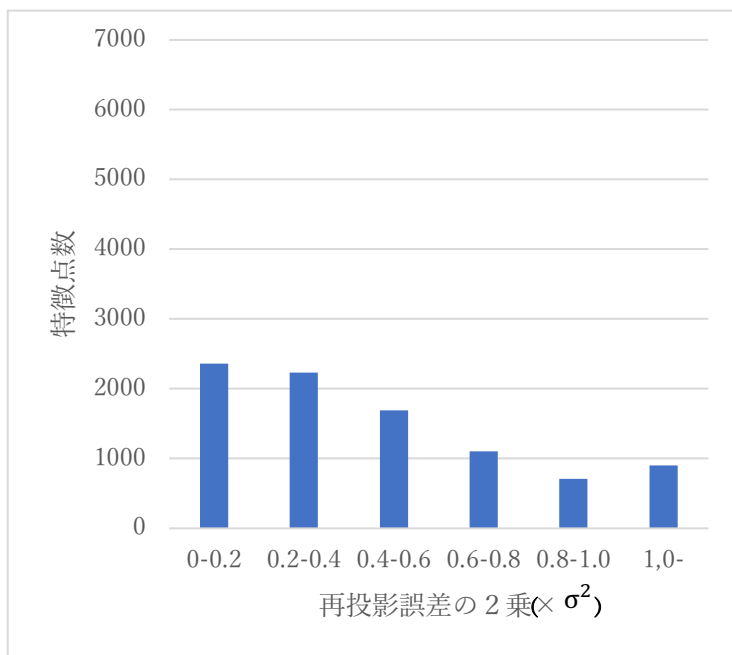


図 21 外れ値 1 ノイズあり M 推定なし

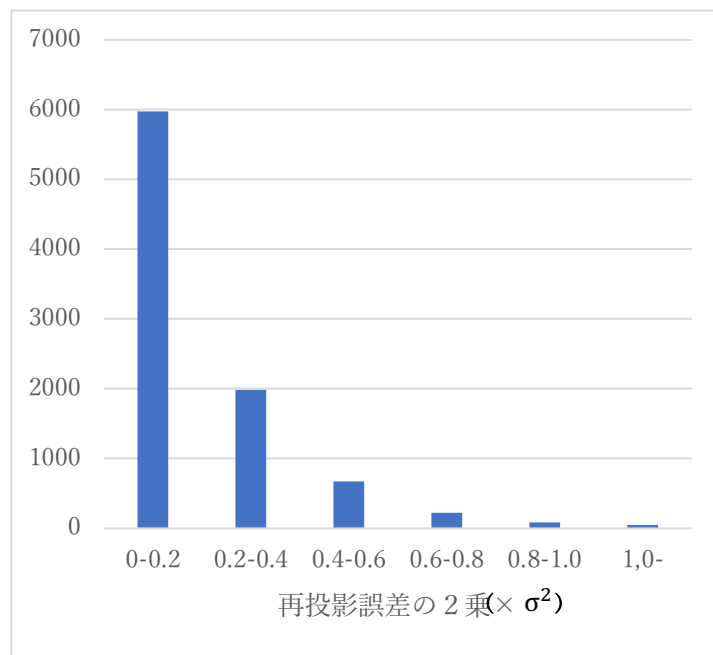


図 22 外れ値 1 ノイズあり M 推定あり

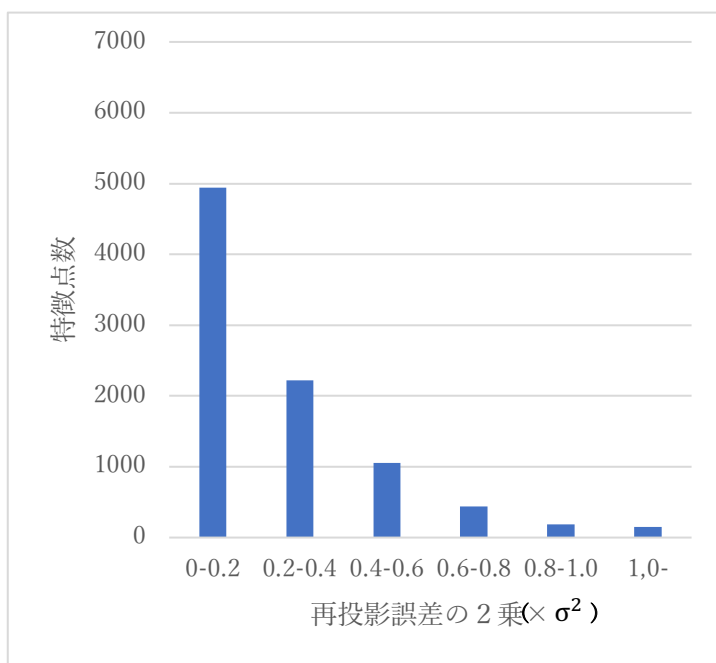


図 23 外れ値 2 ノイズあり M 推定なし

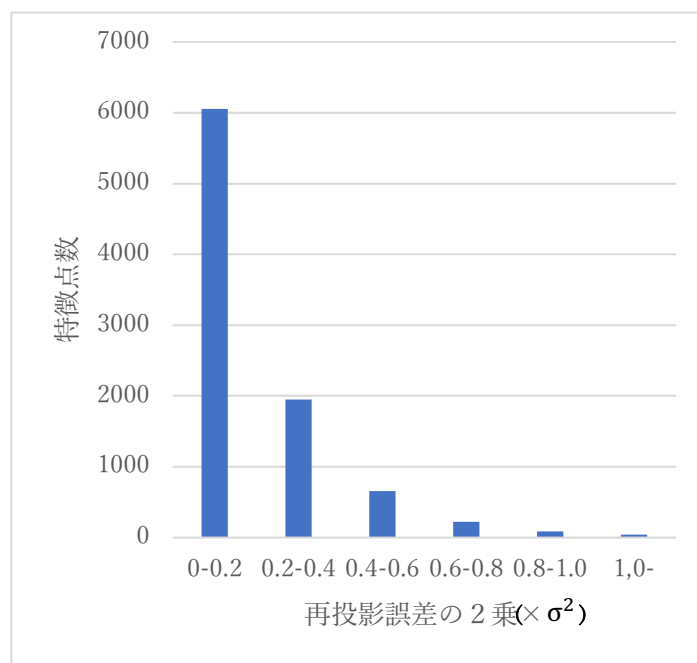


図 24 外れ値 2 ノイズあり M 推定あり

6. おわりに

本研究では、SfM に M 推定を導入することで、外れ値があった場合でも、その影響を低減できるか調査した。人工的に生成した対応点データに、外れ値、ノイズを付加して M 推定ありとなしで形状復元を行った。復元結果の形状から、M 推定ありでは外れ値の影響をそれほど受けず、定性的に正しい形状復元が行えることを確認した。また、外れ値以外の点の再投影誤差を定量的に評価し、M 推定ありの方が精度良く形状復元が行えていることを確認した。これらの実験を通して、M 推定を利用することによる外れ値の低減効果を示した。

今後の課題としては、実画像で 3 次元形状復元を行なった際でも M 推定を利用した方が外れ値に強いかを検証する必要がある。また、本論文では M 推定を利用しているが、RANSAC などの他のロバスト推定を使った場合の復元結果はどうなるかを検証する必要がある。さらに、本論文では、再投影誤差最小化を行う際に最急降下法を使用しているが、ニュートン法などを使った場合も外れ値の影響を低減できるかを検証する必要がある。

謝辞

最後に本研究を行うにあたり、指導教員である棕木雅之教授には、お忙しい中研究のアイデアやプログラムの実装、論文に関する助言など、丁寧に指導していただいたこと大変感謝しています。また、夜遅くまで質問の解答や論文等の添削をしていただき本当に感謝しています。棕木研究室の皆様もわからないことなどあった際に助けていただき感謝しております。最後に、家族に心からの御礼を申し上げます。

参考文献

- [1] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, IJCV, (2004), pp.91-110.
- [2] Pablo F, Alcantarilla, Jesús Nuevo, Adrien Bartoli, “Fast Explicit Diffusion for Accelerated Features in Nonlinear Scale Spaces”, BMVC, (2013), pp.13.1-13.11.
- [3] 棕木 雅之他 , “非線形最適化による建物画像列からの全集形状一括復元のための初期値設定”,情報処理学会論文誌 CVIM , (2004) , pp.64-73.
- [4] 呂 竜夫他,“統計実務におけるレンジチェックのための外れ値検出方法”, 統計研究彙報 第 72 号,(2015) , pp 41-54.