文化財の 3D モデリングに対する動画画像解析の応用*

國井 洋一[†]・近津 博文[†]

Application of Image Sequential Analysis for 3D Modeling of Cultural Heritage*

Yoichi KUNII[†] and Hirofumi CHIKATSU[†]

Recently, efficient spatial data acquisition and visualization have been receiving more attention from the view point of digital city, VR museum including digital archives. Generally, in order to perform object modeling through digital images, line or feature extraction and stereo matching are performed, and many matching methods such as area based matching, feature based matching have been proposed. In particular, line gives important information for object extraction, and satisfied 3D results depend on rigorous line extraction and matching.

With this motive, the authors have developed robust line matching which is performed by line extraction and line tracking. The line extraction was performed by Canny operator, and line matching was performed using optical flow and trifocal tensor.

In order to apply the line matching method for modeling of historical object, 3D modeling for the “Koma house” through image sequences was investigated in this paper. The Koma house was built in 17th century, and designated as national important cultural assets in 1971.

1. はじめに

近年、歴史的建造物に対するデジタル・アーカイブ化が行われているが、全国で重要文化財に指定されている建造物は 3,000 棟以上にもおよぶため [1], 建造物の効率的 3 次元モデリング手法が期待されている。とくに、データ取得の簡便さや装置の価格面および 3 次元表現までを可能にする手法として動画画像の利用が注目されており、動画画像を用いた多くのモデリング手法が提案されている。たとえば、消失点や複数の視点に対する幾何学関係をを用いたモデリング手法 [2] や実スケールのボクセル空間内においてモデリングを行う手法 [3] などであるが、前者は対象物の相似モデルの復元であるため実モデルの復元が課題であり、さらに高精度な実モデルを復元するためにはカメラの焦点距離や主点位置あるいはレンズ歪み係数といった内部標定要素を考慮する必要がある。また、後者の場合には前述の内部標定要素や 3 次元座標が

既知である複数の基準点をあらかじめ取得する必要があるため、モデリングの効率化のボトルネックとなっている。したがって、動画画像を利用する場合には内部標定要素取得や基準点測量などの拘束条件の簡略化、マッチングの自動化、実モデルの構築などが課題となる。

ところで、点の連結で構成される線分は建造物を抽出するための重要な情報であることを考慮すると線分抽出とラインマッチングとを同時に行うことが建造物のモデリングに効率的であると判断される。すなわち、一本の線分がマッチングされたということは線分抽出とその線分の両端点に対するマッチングが自動的に行われたことになる。この特徴を用いたラインマッチング手法として、従来から動画画像の各フレーム画像に対するエッジポラ幾何を用いた線分の追跡手法が提案されているが [4], 歴史的建造物等の対象物に対する 3D モデリングを行うためにはより多くの線分に対して正確にラインマッチングを行う必要があり、さらにワイヤフレームモデルやテクスチャマップモデルの作成を考えた場合、ラインマッチングされた線分によって構成される面情報の取得が課題となる。

そこで本研究では、まず歴史的建造物に対してビデオカメラを水平方向に移動させて動画画像を撮影し、撮影し

* 原稿受付 2004 年 2 月 16 日

[†] 東京電機大学 理工学部 建設環境工学科 Department of Civil and Environmental Engineering, Tokyo Denki University; Ishizaka, Hatoyama, Hiki-gun, Saitama, 350-0394, JAPAN

Key Words: 3D modeling, line matching, image sequential analysis, digital archive, digital photogrammetry.

た動画像に対するオプティカルフロー推定, トリフォーカル・テンソル [5] およびエビポーラマッチングを組み合わせた効率的なラインマッチング手法についての検討を行った. また, トリフォーカル・テンソルはトリプレット画像における幾何関係を表すものであるが, 本研究ではより精度の高いトリフォーカル・テンソルを求めるためにトリプレット画像に対する効率的なカメラキャリブレーション手法を開発し, 本手法によって得られるカメラの内部標定要素を用いて正確なトリフォーカル・テンソルの算出について検討を行った.

一方, 本研究では上記の手法によってラインマッチングされた線分の各端点に対する3次元実座標を効率的に算出するために, 結合調整法 [6] によって線分の各端点に対する3次元実座標を算出することとした. 結合調整法とは, カメラから測定対象における任意の一点 (原点) までの距離 (撮影距離) を拘束条件とすることにより, 地上基準点を必要とせずに対象物の3次元実座標を算出することが可能となる手法である. また, 結合調整法の際に必要なカメラの内部標定要素は, 上述のトリプレット画像によるカメラキャリブレーションにてすでに求められているため, より効率的に対象物の3次元実座標を算出することが可能となる. すなわち, 本手法は計測対象に対してあらかじめ撮影距離のみを測定することにより, 従来より課題となっていた実モデルの復元および基準点測定の簡略化を図り, さらに内部標定要素の自動取得も可能とするものである. また, ラインマッチングされた線分情報によって構成される面の認識をモルフォロジーの Opening 処理によって行い, 建造物に対する3Dモデリングの効率化についても検討を行った.

なお, 本研究では1971年に重要文化財として国からの指定を受けた [7] 埼玉県日高市に所在する国指定重要文化財 “高麗家” を対象とし, ラインマッチングおよびモデリング手法の検討を行った.

2. オプティカルフローによるラインマッチング

本研究では動画像中のフレーム間に対して線分抽出を行った後にオプティカルフロー推定を行い, 一つ手前のフレームにおいて抽出された線分を推定されたオプティカルフローを用いて移動させ, 移動させた線分と現フレームの線分との対応を類似度によりラインマッチングを行うものである. 本研究では, オプティカルフロー推定には精度が良好でかつ処理が高速な (1), (2) 式に示す Lucas-Kanade 法 [8] を採用した. 本手法の詳細を以下に示す.

$$u = \frac{\sum_w \frac{\partial I}{\partial x} [J(p) - I(p)]}{\sum_w \left(\frac{\partial I}{\partial x} \right)^2} \quad (1)$$

$$v = \frac{\sum_w \frac{\partial I}{\partial y} [J(p) - I(p)]}{\sum_w \left(\frac{\partial I}{\partial y} \right)^2} \quad (2)$$

ここに,

p : 画面座標 (x, y)

I : 一つ手前のフレームにおける濃淡値

J : 現フレームにおける濃淡値

w : 局所領域

2.1 第1フレームに対する処理

動画像中の第1フレームに対して Canny 演算子 [9] によってエッジ抽出を行い, 抽出されたエッジに対する直線化を行う. Canny 演算子とはエッジ方向の計算に対してガウス関数を導入し, エッジ高さとエッジに対する信頼度をしきい値としてエッジ抽出を行うものであり, ここではエッジ高さおよび信頼度に対するしきい値をそれぞれ10, 0.1とした. つぎに, 抽出されたエッジ端点をそれぞれ直線で結んだ直線化画像中においてワイヤフレームモデルを作成する上で必要な線分のみに着目し, 着目した必要な線分のうち最も短い線分の長さ (10 画素) をしきい値と定め, 長さがしきい値未満の直線は自動的に除去し, しきい値以上の直線のみを抽出線分として以後の処理に用いることとした. Fig. 1 は高麗家に対する動画像の第1フレームであり, Fig. 2 は上記に述べたエッジ抽出, 直線化およびしきい値処理の結果得られた線分画像である.

Fig. 1 Koma house

Fig. 2 Lines by Canny and threshold procedure

しかしながら, Fig. 2 から明らかなように, 対象物のモデリングに不要な背景部分等の線分が画像中に含ま

れるため、動画中の第1フレームにおいてのみ手動によってモデリングに不要な線分の除去を行うこととした。この結果第1フレームにおける線分本数は85本となり、その結果をFig.3に示す。

Fig. 3 Lines by manual correction

2.2 フレーム間に対する処理

動画中の第2フレームに対しては自動で線分抽出を行った後、第1、第2フレーム間におけるオプティカルフロー推定を行う。すなわち、第1フレームにおける全線分の端点を推定したオプティカルフローの量だけ移動させ、移動させた線分と第2フレームにおいて抽出された線分との対応付けを類似度[10]を用いて行い、対応がとれた線分のみを残して対応がとれなかった線分は消去する。類似度とはFig.4に示すようなフレーム間において注目する2本の線分 L_i 、 L_j がどれだけ類似しているかを表す指標であり、(3)式により算出される。ここで、 L_m は L_i 、 L_j の両端点(A, BおよびC, D)に対するそれぞれの中点(E, F)を結んだ直線である。また、類似度は0~1の実数となるが、本研究では類似度がしきい値以上の線分のみを対応線分としてラインマッチングを行った。すなわち、背景やオクルージョン等の影響により線分が消失した場合は、類似度が0となるため注目線分に対するオプティカルフロー推定によるラインマッチングは打ち切れ、以後の処理は行われないこととなる。なお、ここでの(3)式における重み係数は、 $k_d=0.3$ 、 $k_p=0.7$ とした。Fig.5は第1、第2フレームによって推定されたオプティカルフローの画像である。

$$s = \frac{1}{k_d D + k_p P} \quad (3)$$

ここに、

$$D = D_i + D_j,$$

$$P = \frac{P_i}{P_j} (P_j < P_i) \text{ または } P = \frac{P_j}{P_i} (P_i < P_j)$$

k_d, k_p : D, P に対する重み係数

($k_d + k_p = 1$ となるように設定する.)

なお、本研究では類似度のしきい値を決定するために、類似度に対して様々なしきい値を設定して動画中の数フレームに対してラインマッチングを試験的に行い、最も効率的なラインマッチングが行われた値0.7を採用し

た。ところで、本手法では第2、第3フレームに対しても同様な処理を行うが、濃淡値変化の少ない部分においてはエッジ抽出およびオプティカルフロー推定の精度が低下するため第3フレームに対してラインマッチングが行われない場合も推測される。したがって、ここでは対応がとれずに第3フレームでは検出されなかった第2フレーム上の線分の端点位置を以下に述べるトリフォーカル・テンソルによる補完処理のために保存しておく。したがって、ここまでの処理では第1、第2、第3フレームの画像間で共通に検出された対応線分のみが取得されることになる。

Fig. 4 Similarity of 2 lines

Fig. 5 Optical flow estimation

3. トリフォーカル・テンソルによるラインマッチング

トリフォーカル・テンソルはトリプレット画像における幾何関係を表すものであるため、濃淡値情報を利用せずに対応点の算出が可能であり、オプティカルフロー推定による手法によって得られなかった対応線分を効率的に補完することが可能である。したがって、ここでは上記のオプティカルフロー推定によりマッチングされた第1、第2、第3フレームにおける対応線分の各端点を3フレーム間における対応点として抽出した後、抽出された

対応点によって 3 フレーム間に対するトリフォーカル・テンソルを算出し、トリフォーカル・テンソルを用いてオブティカルフロー推定でも対応が得られなかった第 2 フレームの線分端点に対する第 3 フレーム上での対応点を算出するものである。すなわち、背景等の影響により第 3 フレームでは対応線分の抽出が不十分であった場合においてもトリフォーカル・テンソルによって線分端点の対応点が算出し、算出された対応点の端点同士を結ぶことにより対応線分を取得しようとするものである。

ところで、トリフォーカル・テンソルは 3 フレーム間の画像において 7 点以上の共通の対応点が存在すれば、対応点の画面座標の値を用いて算出することが可能であるが、より精度の高いトリフォーカル・テンソルを求めるためにはレンズ歪みや主点位置のずれ等を考慮し、かつ対応点の画面座標をセンサ座標に変換する必要がある。そこで、本研究ではトリプレット画像に対して地上基準点を必要としないカメラキャリブレーション手法を開発し、本手法によって 3 枚の画像に共通の内部標定要素を算出した後、算出された内部標定要素を用いて対応点のセンサ座標の算出を行った。以下にカメラキャリブレーションおよびトリフォーカル・テンソルの算出について述べる。

3.1 トリプレット画像によるカメラキャリブレーション

本研究では動画像中の連続する 3 フレームの画像を用いて、以下に示す手法によりカメラキャリブレーションを行った。

- (1) 前述のオブティカルフロー推定によるラインマッチングにより、第 1, 第 2, 第 3 フレームの 3 枚の画像間において共通に対応付けされた線分を取得し、その対応線分の各端点を 3 枚の画像間における対応点として抽出する。
- (2) 第 1 および第 2 フレームの 2 枚の画像に対し、抽出した対応点全ての画面座標を用いて共面条件による相互標定を行い、相互標定によって算出された相互標定要素を用いて全ての対応点に対する相対 3 次元座標を算出する。共面条件とは左右カメラの撮影点と左右のカメラに撮影されている同一点に対する像点の 4 点が同一平面状に含まれるという条件より誘導される条件式である (付録 1. 参照)。また相互標定とは相対的に相似な空間モデルを形成することである。
- (3) 第 3 フレームに対し、相対 3 次元座標が求められた対応点のうちの任意の 9 点を用いて空間後方交会法によりセルフキャリブレーション [11] を行う。空間後方交会法とは 3 次元座標が与えられている点を用いて撮影時における外部標定要素 (カメラの位置・姿勢) を求めることであり、セルフキャリブレーションとは内部標定要素と外部標定要素を同

時に求めることである。

- (4) 第 1, 第 2, 第 3 フレームの 3 枚の画像に対し、(3) のセルフキャリブレーションにおいて得られたカメラの内部標定要素を初期値として 3 枚同時にカメラキャリブレーションを行う。すなわち、相対 3 次元座標が求められている対応点のうち 3 枚の画像に共通な 9 点を用いて、共線条件式により以下に示す 51 個の未知量を算出する。51 個の未知量の内訳は 3 枚の画像に対する外部標定要素 18 個 (6×3)、3 枚の画像に共通の内部標定要素 6 個と 9 点に対する 3 次元座標 27 個 (9×3) の 51 個である。これらの未知量に対して 3 枚の画像に共通な 9 点より導かれる 54 個 ($9 \times 2 \times 3$) の共線条件式と、カメラの焦点距離に対する拘束条件 1 個、三つのカメラの姿勢角に対する拘束条件 9 個 (3×3) の合計 64 個の条件式を最小二乗法により同時に解くこととなる。なお、ここで用いた 9 点は (3) で設定した 9 点と同じ点を採用した。また、共線条件式とは撮影点、地上点および地上点に対する像点とは一直線上にあるという条件から誘導される条件式である (付録 2. 参照)。

Table 1 に本手法によって得られた第 1, 第 2, 第 3 フレームのトリプレットに対する内部標定要素の値を示す。なお、スケールファクタ ($a_1 \sim a_4$) とはセンサ座標 (u, v) (単位:mm) を画面座標 (x, y) (単位:pixel) へ変換するための係数であり、変換式は主点位置 (x_0, y_0) を用いて (4), (5) 式にて表される。

$$x = x_0 + a_1 u + a_2 v \quad (4)$$

$$y = y_0 + a_3 u + a_4 v \quad (5)$$

通常 a_2 および a_3 は 0 であり、また a_1 および a_4 は CCD センサにおける 1 画素当たりの縦横の長さ (セルサイズ) から算出可能である。しかし、一般的に CCD センサでは受光した光を電気信号へ変換する際に若干の水平駆動の同期ずれが生じるため、 a_1 および a_2 を未知量として取り扱うことにより同期ずれの補正を行う。なお、垂直駆動に関しては同期ずれを生じることがないため、 a_3 および a_4 は既知量として取り扱う。

Table 1 Interior parameters

x_0	337.3750	a_1	261.1891
y_0	254.2406	a_2	0.6943
f	2.1471	a_3^*	0
k_1	6.9603E-06	a_4^*	289.8551

x_0, y_0 :Principal point[pixel], f :Focal length[mm],

k_1 :Lens distortion coefficient, $a_1 \sim a_4$:Scale factor

*: a_3 and a_4 are set as known parameters.

3.2 トリフォーカル・テンソル

トリフォーカル・テンソルは異なる視点から同一の物体が撮影されたトリプレット画像の幾何関係を $3 \times 3 \times 3$ のテンソル、すなわち 27 個のパラメータで表すものでありエピポーラ幾何に基づいた理論によって算出される。トリフォーカル・テンソル T_{ijk} を構成する 27 個のパラメータを $t_{111} \sim t_{333}$ 、動画像中の第 1, 第 2, 第 3 フレームにおける一組の対応点の各画像に対するセンサ座標を (x_1, y_1, z_1) , (x_2, y_2, z_2) , (x_3, y_3, z_3) とすると、(6)~(9) 式に示す 4 個の関係がトリプレット画像における対応点一組毎に得られることとなる。

$$-z_2 z_3 g_{22} + z_2 y_3 g_{23} + y_2 z_3 g_{32} - y_2 y_3 g_{33} = 0 \quad (6)$$

$$z_2 z_3 g_{21} - z_2 x_3 g_{23} - y_2 z_3 g_{31} + y_2 x_3 g_{33} = 0 \quad (7)$$

$$z_2 z_3 g_{12} - z_2 y_3 g_{13} - x_2 z_3 g_{32} + x_2 y_3 g_{33} = 0 \quad (8)$$

$$-z_2 z_3 g_{11} + z_2 x_3 g_{13} + x_2 z_3 g_{31} - x_2 x_3 g_{33} = 0 \quad (9)$$

ここに、

$$g_{ij} = x_1 t_{1ij} + y_1 t_{2ij} + z_1 t_{3ij}$$

したがって、27 個のパラメータを全て求めるためには 7 点以上の対応点を取得する必要があるが、4 本以上の対応線分がオプティカルフローによって得られればよいことになる。また、算出された 27 個のパラメータにより第 3 フレームにおいて対応が得られなかった点に対しても (6)~(9) 式によって対応関係が求められることになるため、これらの点に対する第 3 フレームにおけるセンサ座標を算出することが可能となる。

一方、本研究では算出された第 3 フレームでの対応点の位置が正しいか否かをオプティカルフローの推定量を利用して判別を行った。すなわち、トリフォーカル・テンソルによって算出された第 3 フレームでの対応点の位置が、第 2 フレームから何画素移動したものを求めた後に、第 2, 第 3 フレーム間において推定された全てのオプティカルフローの推定量に対する最小値と最大値を求め、算出した対応点の移動量がオプティカルフローの推定量の最小値から最大値までの範囲内である場合は推定された対応点は正しいと判断し、範囲外である場合は推定された対応点は誤っていると判断し対応点としては認めないこととする。以上の結果、第 3 フレーム上で対応が認められた端点同士を直線で結ぶことにより、オプティカルフロー推定のみでは対応が得られなかった線分に対しても対応線分が取得されることとなる。

ここまでの処理を第 2, 第 3, 第 4 フレーム、さらに第 3, 第 4, 第 5 フレームと最終フレームまで自動的に繰り返すことにより、第 1 フレームと最終フレームにおける対応線分が自動取得されることとなる。

3.3 精度検証

本研究では動画像に対するオプティカルフロー推定とトリフォーカル・テンソルとを組み合わせた手法の有用性を

検証するために、最小二乗マッチング法 (Least squares matching method) [12], 確率的弛緩法 (Probabilistic relaxation method) [13] および面積相関法 (Area correlation method) [14] によってラインマッチングを行い比較検討を行った。Table 2 は高麗家正面に対する動画像 300 フレームに対して各手法による最終フレームに対するラインマッチングの結果である。オプティカルフローとトリフォーカル・テンソルとを組み合わせた手法からは、他の手法よりも高いマッチング率が得られた結果となった。また、オプティカルフロー推定のみによるラインマッチング結果との比較においてもマッチング率の向上が認められた。したがって、オプティカルフローはカメラの移動速度が大きいほど精度が低下する特徴を持っているが、トリフォーカル・テンソルは移動速度が大きい場合はフレーム間の基線比が大きくなるためトリプレット画像におけるカメラキャリブレーションの精度が向上し、トリフォーカル・テンソルによる対応点算出の精度も向上することが確認された。

Table 2 Results of line matching

Method	Number of lines	Matching rate[%]
(a)	41	48.2
(b)	20	23.5
(c)	28	32.9
(d)	9	10.6

(a) Combined method

(b) Least squares matching method

(c) Probabilistic relaxation method

(d) Area correlation method

4. エピポーラマッチング

ここまでのラインマッチング手法は動画像の各フレームに対して連続的に処理を行うものであった。そのため、第 1 フレームと最終フレーム間に注目した場合、両画像間において対応線分が存在するにもかかわらず、オプティカルフロー推定およびトリフォーカル・テンソルによるラインマッチングの過程において端点の追跡が途切れ、ラインマッチングが達成されない場合が想定される。このような線分に対するラインマッチングを補完するために、本研究では第 1 フレームと最終フレームにおいてエピポーラマッチングを行うこととした。すなわち、オプティカルフローおよびトリフォーカル・テンソルによるラインマッチングによって得られた第 1 フレームと最終フレームに対する対応線分の端点の画面座標を用いて相互標定を行った後にエピポーララインを推定することにより対応が得られていない第 1 フレーム上の線分の各端点に対する最終フレームにおける対応点をエピポーララインに沿って探索し、探索された最終フレームにおける端点同士を再び直線で結ぶことにより対応線分を取得

するものである。

また、最終フレームに対して Canny 演算子によるエッジ抽出および直線化処理をあらかじめ行っておき、取得した対応線分の良否判別を行った。すなわち、エピポーラマッチングにより最終フレームにおいて取得した対応線分の両端点から半径 20 画素以内の範囲に端点が存在する直線に着目し、着目した直線と対応線分との類似度を算出し、算出した類似度がしきい値 0.7 以上であれば対応線分に相当する直線が存在するために取得した対応線分は正しいと判断して採用し、しきい値未満である場合は対応線分が得られなかったと判断して線分を消去することとした。なお、エピポーラマッチングには前述のラインマッチングと同様に最小二乗マッチング法、確率的弛緩法、面積相関法の 3 手法による対応点探索を試みた結果、Table 3 に示すように最小二乗マッチング法がもっとも高いマッチング率 91.8[%] を示した。また、ラインマッチングが行われた 78 本の線分に対して第 1 フレームと最終フレームの線分抽出結果から目視によるマッチングの正確性の検証を行い、78 本全ての線分に対するマッチングが正確に行われたことを確認した。したがって、オプティカルフロー推定、トリフォーカル・テンソルおよび最小二乗マッチング手法を用いたエピポーラマッチングを組み合わせることによりマッチング率が向上し、さらに対応が得られなかった線分が自動的に消去されることにより誤対応を防ぐこともできるため、マッチングの効率化が図られると判断される。以上の提案手法により得られた最終フレームを Fig. 6 に示す。

Table 3 Results of epipolar matching

Method	Number of lines	Matching rate[%]
(a)	78	91.8
(b)	73	85.9
(c)	68	80.0

- (a) Least squares matching method
- (b) Probabilistic relaxation method
- (c) Area correlation method

Fig. 6 Result of proposed matching

5. 3D モデリング

上述のラインマッチング手法により得られた線分はすべて正確にマッチングが行われており、対応が得られなかった線分はすでに消去されている。そのため、本研究ではラインマッチングされた線分を用いて以下に示す手法により歴史的建造物に対する 3D モデリングを効率的に行うことを検討した。まず、歴史的建造物に対するワイヤフレームモデルを作成し、さらに作成されたワイヤフレームに対してテクスチャマッピングを行うためには各線分がそれぞれどのような面を構成しているかを把握する必要があり、各線分によって構成される建造物の壁面や屋根面といった面情報を効率的に認識する必要がある。そこで、本研究ではまずモルフォロジーの Opening 処理 [15] を用いて、画像中の高麗家に対する壁面および屋根面の自動抽出を検討することとする。Opening 処理によって画像中の特徴域を抽出する場合抽出すべき特徴域と類似した色情報を持つ構造要素が必要となるが、本研究では構造要素の形状を円形とし、さらに円形構造要素の直径を決定するために、動画像の第 1 フレームに対して様々な直径を設定して壁面および屋根面の抽出を試験的にを行い、最も効率よく屋根面および壁面が抽出された値 (7 画素) を採用した。また、構造要素に対する色情報の取得ためには抽出すべき特徴域から手動にて色情報を取得するのが一般的であるが、本研究ではラインマッチングされた線分に対する周辺領域の色情報 (HSI 値) を用いて構造要素に対する色情報の自動取得を行った [16]。以上の作業をラインマッチングされた全ての線分に対して行うことにより、高麗家の屋根面および壁面の自動抽出が達成される。Fig. 7 に壁面および屋根面の自動抽出した結果を示す。

Fig. 7 Result of surface extraction

つぎに、抽出された高麗家に対する壁面および屋根面とラインマッチングされた線分との対応づけを行うために、Fig. 7 に対するコーナー検出を画像処理により行い [16]、検出された各コーナーに対して最も近い距離に存在する線分の端点を検出し、各端点が高麗家の壁面および屋根面におけるコーナーに対応するものとした。以上の処理により、検出された端点を用いて壁面および屋根面それぞれに対して TIN モデルを発生させることにより、壁面および屋根面が認識されることとなる。

さらに、TINモデルにおける各端点に対する3次元座標を求めるために、第1フレームと最終フレームの画像に対してカメラキャリブレーションを行う。ここでのカメラキャリブレーションには地上基準点を必要としない結合調整法を用いた。結合調整法とは基準点測定の簡略化を目的に筆者らにより開発された手法であり、2枚の画像に対する撮影点および共面条件[14]により算出される相似モデルを初期値とし、カメラから測定対象上における任意の一点（原点）までの撮影距離を拘束条件とした空間後方交会法によりカメラキャリブレーションと絶対標定を同時に行うものである。これにより、TINモデルにおける各線分端点は動画画像における第1フレームと最終フレームとの間でマッチングが行われている点であるため、3次元座標の算出を行うことが可能となる。なお、絶対標定とは作成されたモデルを実空間上での座標系に変換することである。

以上の処理を高麗家の右側面に対する動画画像に対しても同様に行い、正面および右側面の動画画像において共通に写しこまれている点の3次元座標を用いて座標系の統合を行った。Fig. 8に統合されたワイヤフレームモデルを示す。さらに、作成したワイヤフレームの各面に対する色情報を第1フレーム画像より取得し、対応する各面に貼り付けることによりテクスチャマッピングを行った。以上の処理によって作成されたテクスチャマップモデルをFig. 9に示す。

Fig. 8 Wire frame model

Fig. 9 Texture mapped model

本研究では以上の手法の有効性および作成されたワイヤフレームモデルおよびテクスチャマップモデルの信頼性を評価するために、撮影高度を1351[mm]として、焦点距離3.4[mm]のビデオカメラを用いて3次元情報が既

知であるテストモデル（縦：347[mm]，横：491[mm]，奥行き：28[mm]）の動画画像を室内において撮影し、本研究で示した提案手法によりラインマッチングを行った後に線分の端点に対して求められる3次元座標に対して精度検証を行った（線分本数：65本，端点数：130点）。その結果、平面座標に対する精度は±1.56[mm]，奥行き座標に対しては±5.27[mm]という値を得た。ところで、各端点に対する平面座標および奥行き座標の平均二乗誤差はカメラの撮影高度，焦点距離，基線長，画面座標の読取り精度を用いて(10)，(11)式により推定されるが、

$$\sigma_{xy} = \left(\frac{H}{f} \right) \sigma_p \quad (10)$$

$$\sigma_z = \left(\frac{H}{f} \right) \left(\frac{H}{B} \right) \sigma_p \quad (11)$$

ここに、

H : 撮影高度, f : 焦点距離, B : 基線長,

σ_p : 画面座標の読取り精度

ここでの実験における第1フレームと最終フレームに対する撮影点間の距離を基線(400[mm])とし、また端点位置の最小読取り値は1画素であるため、読取り精度±1画素(±3.4[μm])として(10)，(11)式より算出されるテストモデルの3次元座標に対する推定値は平面±1.59[mm]，奥行き±5.37[mm]となり、本提案手法によって得られた計測精度とほぼ同等であることが確認された。

さらに、高麗家の正面において3次元座標が得られた各線分端点間の直線距離をレーザレンジファインダ（測定精度：±1.5[mm]）によって計9本測定し、直接距離測定の結果得られた値に対して本手法によって得られた各端点の3次元座標より算出される間接距離の値との比較を行った結果、9本の直線距離に対する平均二乗誤差は±52[mm]であったが、室内実験より得られた平面・奥行き精度を採用し、画面座標の読取り精度を±1画素とした場合の距離の精度に対する推定値は誤差伝播の法則より±27[mm]となる。したがって、高麗家のモデリングに対する読取り精度は±2画素程度に低下したものと推測されるが、この要因としては屋外における背景の複雑さ、日照の影響、エッジ部分の材質や形状の複雑さ等により線分抽出の精度が低下したものと考えられる。しかしながら、本手法は動画画像の利用によりラインマッチングの自動化、さらには実空間での3Dモデリングの自動化を可能にしたものであり、文化財のデジタル・アーカイブ化において貢献するものと期待される。

6. おわりに

本研究では歴史的建造物に対する効率的な3Dモデリングを目的にラインマッチングの効率化、さらにはOpening処理およびTINモデルを用いた3Dモデリング

の効率化について検討を行った。その結果、ラインマッチングはオプティカルフロー推定、トリフォーカル・テンソルおよびエピポーラマッチングを組み合わせる用いることによって処理はほぼ自動化され、手動にて行われる処理は第1フレームに対する不要線分の除去ならびにCanny演算子および類似度に対するしきい値設定のみであり、そのうえで高いマッチング率(91.8%)が達成されることが確認された。また、3Dモデリングに関してはラインマッチングによって得られた線分情報からTINモデルを作成し、作成したTINモデルの各端点に対する3次元座標を結合調整法を用いて求めることによりワイヤフレームモデルが効率的に作成され、さらにOpening処理において構造要素に対する直径の設定のみを手動にて行うことにより面情報の認識が効率的に行われ、テクスチャマッピングに関しても効率化されることが確認された。以上の結果より、本手法は歴史的建造物に対する3次元モデルの作成に有効であると判断されるが、今後の課題としてはオクルージョン部分が多く存在するような複雑な形状を有する文化財に対する応用が挙げられる。

参考文献

- [1] 文化庁ホームページ: <http://www.bunka.go.jp/> (2004)
- [2] P. H. S. Torr, A. W. Fitzgibbon and A. Zisserman: The problem of degeneracy in structure and motion recovery from uncalibrated image sequences; *International Journal of Computer Vision*, Vol. 32, No. 1, pp. 27-44 (1999)
- [3] 佐藤, 神原, 横矢, 竹村: マルチベースラインステレオ法を利用した動画からの屋外環境の三次元モデル化; *VR学会論文誌*, Vol. 7, No. 2, pp. 275-282 (2002)
- [4] D. D. Morris and T. Kanade: A unified factorization algorithm for points, line segments and planes with uncertainty models; *Proc. 6th International Conference on Computer Vision*, pp. 696-702 (1998)
- [5] R. I. Hartley: A linear method for reconstruction from lines and points; *Proc. 5th International Conference on Computer Vision*, pp. 882-887 (1995)
- [6] H. Chikatsu and Y. Kunii: Performance evaluation of recent high resolution amateur cameras and application to modeling of historical structure; *Internationals Archives of Photogrammetry and Remote Sensing*, Vol. 34, Part 5, pp. 337-341 (2002)
- [7] 高麗家パンフレット (1980)
- [8] B. D. Lucas and T. Kanade: An iterative image registration technique with an stereo vision; *DARPA Image Understanding Workshop*, pp. 121-130 (1981)
- [9] J. Canny: A computational approach to edge detection; *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. PAMI-8, No. 6, pp. 679-697 (1986)
- [10] N. Chiba and T. Kanade: A tracker for broken and closely-spaced lines; *Internationals Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 5, pp. 676-681 (1998)
- [11] 解析写真測量委員会編: 解析写真測量; 日本写真測量学会 (1983)
- [12] A. Gruen: Adaptive least squares correlation: A powerful image matching technique; *South Africa Journal of Photogrammetry, Remote Sensing and Cartography*, pp. 175-187 (1985)
- [13] 高木, 下田: 画像解析ハンドブック; 東京大学出版会 (1991)
- [14] 動体計測研究会: イメージセンシング; 日本測量協会 (1997)
- [15] 小畑: モルフォロジー; コロナ社 (1996)
- [16] 國井, 近津: 動画画像解析による効率的建造物の抽出と三次元都市モデルの構築について; 写真測量とリモートセンシング, Vol. 42, No. 1, pp. 12-20 (2003)

付 録

付録 1. 共面条件

共面条件とはFig. A1に示すとおり, 点Pに対して撮影された一組のステレオ画像に対する撮影点 $O_1(X_{01}, Y_{01}, Z_{01})$, $O_2(X_{02}, Y_{02}, Z_{02})$ および像点 $P_1(X_1, Y_1, Z_1)$, $P_2(X_2, Y_2, Z_2)$ が同一平面 (エピポーラ面: epipolar plane) 上に存在するという条件であり, 共面条件式は (A1) 式のように平面の方程式にて示される。

$$\begin{vmatrix} X_{01} & Y_{01} & Z_{01} & 1 \\ X_{02} & Y_{02} & Z_{02} & 1 \\ X_1 & Y_1 & Z_1 & 1 \\ X_2 & Y_2 & Z_2 & 1 \end{vmatrix} = 0 \quad (A1)$$

ここで, 左カメラの撮影点 O_1 を座標系の原点とする。また, O_1 と O_2 とを結ぶ直線を X 軸とし, 基線長 B_x は単位長さをとる。これにより, 相互標定において求める未知量, すなわち標定要素は左カメラの姿勢角 ϕ_1 , κ_1 および右カメラの姿勢角 ω_2 , ϕ_2 , κ_2 の5個となり, (A1) 式の共面条件式は (A2) 式のように変形される。

$$F(\phi_1, \kappa_1, \omega_2, \phi_2, \kappa_2) = \begin{vmatrix} Y_1 & Z_1 \\ Y_2 & Z_2 \end{vmatrix} = 0 \quad (A2)$$

以上より, 左右の画像における対応点1点につき (A2) 式の共面条件式が1個得られるため, 5点以上の対応点が得られれば最小二乗法により各標定要素の算出, すなわち相互標定を行うことが可能となる。

付録 2. 共線条件

共線条件とは, Fig. A2に示すように点 $P(X, Y, Z)$ を撮影した際の写真上における像点 $p(x, y, z)$ およびカメラの投影中心 $O(X_0, Y_0, Z_0)$ の3点が一直線上に存在する条件のことであり, (A3) 式で示される。

$$\begin{pmatrix} X - X_0 \\ Y - Y_0 \\ Z - Z_0 \end{pmatrix} = \lambda \mathbf{R}^{-1} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \quad (A3)$$

Fig. A1 Coplanarity condition

ここに、 λ は定数である。また、 $\mathbf{R}$ は姿勢角 (ω, ϕ, κ) に対する回転行列であり、(A4)式で示される。

$$\mathbf{R} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad (\text{A4})$$

ここに、

$$a_{11} = \cos\phi\cos\kappa, a_{12} = -\cos\phi\sin\kappa, a_{13} = \sin\phi,$$

$$a_{21} = \cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa,$$

$$a_{22} = \cos\omega\cos\kappa - \sin\omega\sin\phi\sin\kappa,$$

$$a_{23} = -\sin\omega\cos\phi,$$

$$a_{31} = \sin\omega\sin\kappa - \cos\omega\sin\phi\sin\kappa,$$

$$a_{32} = \sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa,$$

$$a_{33} = \cos\omega\cos\phi$$

点 p の Z 座標はカメラの焦点距離 f に相当するため、 $z = -f$ を代入して展開すると (A5), (A6) 式の共線条件式が得られることとなる。

$$x = -f \frac{a_{11}(X - X_0) + a_{12}(Y - Y_0) + a_{13}(Z - Z_0)}{a_{31}(X - X_0) + a_{32}(Y - Y_0) + a_{33}(Z - Z_0)} \quad (\text{A5})$$

$$y = -f \frac{a_{21}(X - X_0) + a_{22}(Y - Y_0) + a_{23}(Z - Z_0)}{a_{31}(X - X_0) + a_{32}(Y - Y_0) + a_{33}(Z - Z_0)} \quad (\text{A6})$$

Fig. A2 Collinearity condition