



## 研究の概要と1年間の進捗

M2 田川幸汰

### 1 概要

#### 1.1 研究背景

近年、デジタルツインの市場規模が急速に拡大しており、現実空間の高精度なデジタル再現技術が多様な分野で求められている[1]。特に、屋内外の空間情報をリアルタイムに反映し、ナビゲーションや現場管理に活用するニーズが高まっている。しかし、一般的な3次元モデルの構築では、多視点からの多数の画像撮影が必要となる上に、テクスチャを滑らかに割り当てるには高度な撮影技術や、専門のソフト等が必要となる。これらの問題はデジタルツインの現場導入を妨げる要因となっている[2]。

#### 1.2 研究目的

本研究では、全方位画像を活用して、簡易かつ効率的に高品質なテクスチャ付き3次元モデルを生成し、それを用いた実用的な自己位置推定手法の確立を目指す。全方位画像の活用により、多視点からの撮影や特別な撮影手段を必要とせず、簡易にデジタルツイン環境を構築することが可能となる。さらに、本手法はセンサやGPSに依存しないため、専用機材を新たに導入する必要がなく、屋内などセンサ情報の取得が困難な環境において、有効なテクスチャマッピング手法を提供できる。また、現実空間を3次元モデルとして再現しながら自己位置推定を行うことで、ナビゲーションや現場管理支援といった実用的なニーズに応える技術基盤を構築する。

#### 1.3 手法の概要

##### 1. 全方位カメラのパラメータ推定

- 2次元座標と3次元座標の対応を用いて全方位カメラの位置と角度を推定する。

##### 2. テクスチャの割り当て

- 全方位画像から取得したテクスチャと、建物の壁と床を単純な平面で表現した3次元ワイヤーフレームモデルを組み合わせてシーンを再現する。

##### 3. ユーザ画像との自動対応付けによる自己位置推定

- 入力画像と3次元モデルの対応点を自動で抽出し、位置を推定する。

### 2 全方位カメラのパラメータ推定

全方位画像から任意視線方向の透視投影画像を生成し、透視投影画像上の2次元座標と3次元モデルの頂点を対応付ける。カメラ位置姿勢の推定には、直交射影の共線性と共面性を用いた手法[3]を用いる。共線性とは、3次元点と画像面上の点が誤差なく対応する時、3次元点はカメラ中心から画像面上のベクトル上に存在することである。また共面性とは、3次元直線と画像面上の直線が誤差なく対応する時、3次元直線はカメラ中心と画像面上の直線が構成する平面上に存在することである。これらの誤差を最小化することで、カメラパラメータを推定する手法である。



## 2.1 推定結果

入力した全方位画像を図 1 に示す。



図 1: 全方位画像

パラメータ推定に用いた透視投影画像および2次元座標を図2に示す。

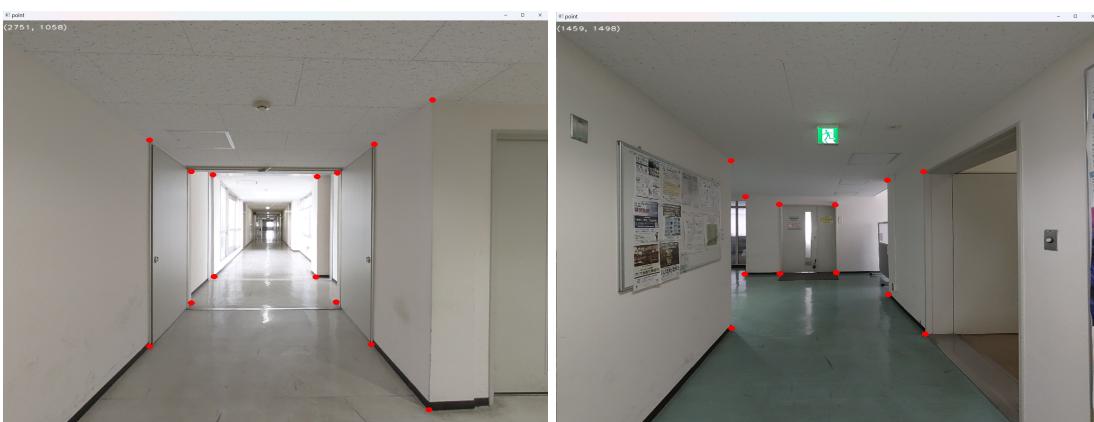


図 2: 透視投影画像

パラメータの推定結果を表3に示す。今回はカメラの位置を表す並進ベクトルのみ示す。

	推定結果 (cm)	実測値との誤差 (cm)
x	100.45	0.45
y	395.58	-4.41
z	136.82	1.82

図 3: パラメータの推定結果

実測値との精度比較では、いずれの条件においても誤差はおおむね 10cm 以内に収まることが多い。



### 3 テクスチャの割り当て

先ほど推定したカメラパラメータとともに、テクスチャを割り当てる。テクスチャを割り当てる際の工夫したポイントは以下の4点である。

#### 3.1 手法の概要

1. テクスチャ割り当ての条件
  - 複数の全方位カメラで同一領域のテクスチャを取得する際、重心がカメラ中心に近いテクスチャを優先する。
2. テクスチャの明るさ調整
  - 撮影環境による明るさの差を補正し、滑らかなシーンを再現する。
3. 側面部分のテクスチャ形状を四角形に変更
  - 店舗の壁面や正面の直線的な形状に自然に合うように変更し、透視投影の際に発生する歪みを射影変換により解消する。

#### 3.2 生成したテクスチャ及び3次元モデル

取得したテクスチャの一部を図4に示す。

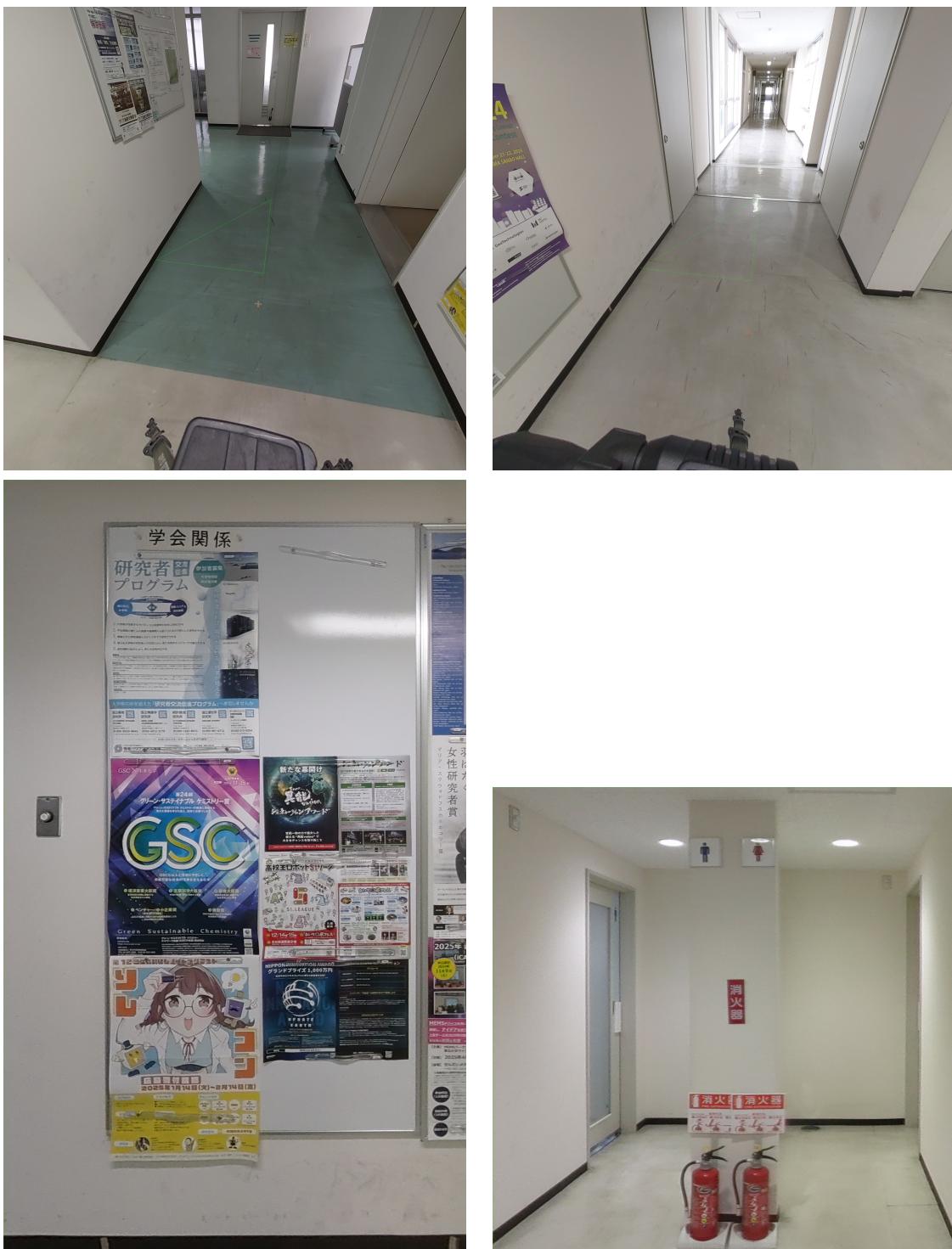


図4: テクスチャ画像

テクスチャを割り当てた3次元モデルを図5に示す。

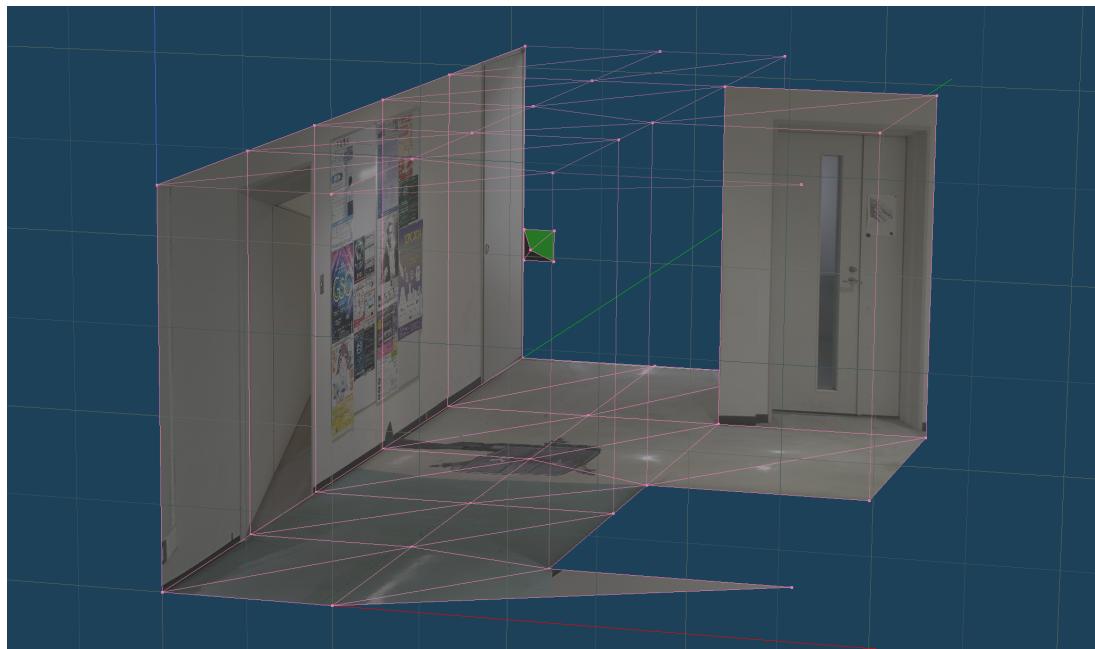


図 5: 3次元モデル

全方位カメラと面の角度が大きくなると歪みが発生してしまうが、ある程度正確にテクスチャを割り当てることができている。

## 4 ユーザ画像との自動対応付けによる自己位置推定

ユーザが入力した画像の特徴点と、3次元モデルのテクスチャの特徴点を自動で対応付けて、自己位置を推定する。特徴点のマッチングには AKAZE 特徴点検出器を使用するが、入力画像に対して多数のテクスチャ画像があるため、全ての画像に対してマッチングを行うと実行時間が大きくなり実用性に欠けてしまう。そのため、まず ResNet と呼ばれる大域特徴量抽出手法を用いて類似画像を絞り込み、その後、少數の画像に対して特徴点マッチングを行うことで、精度と実行時間の両立をねらう。

### 4.1 類似画像検索の結果

ResNet と AKAZE を用いた場合の類似画像検索結果を図 6 に示す。なお、複合スコア *CombinedScore* は以下の式で表される。

$$\text{CombinedScore} = \text{ResNet}_{score} \times 0.5 + (\text{AKAZE}_{matchcount}/100) \times 0.5 \quad (1)$$



図 6: 類似画像検索



実行時間について補足すると、AKAZE 特徴点検出器を用いたマッチングでは、画像1枚あたり約0.5秒の処理時間がかかるため、対象画像が多い場合は全体の処理時間が大幅に増加する。一方、ResNetによる大域特徴量の計算は、最初にすべての画像に対して特徴量を抽出する処理が必要だが、その後の類似画像検索は非常に高速であり、1回あたりの処理時間は AKAZE よりもはるかに短くなる。AKAZE と ResNet を組み合わせることで、精度を維持しながら、実行時間の面でも優れたパフォーマンスを実現している。

## 4.2 特徴点マッチングの結果

もっとも類似度の高い画像に対して AKAZE 特徴点検出器によるマッチングを行った結果を図 7 に示す。

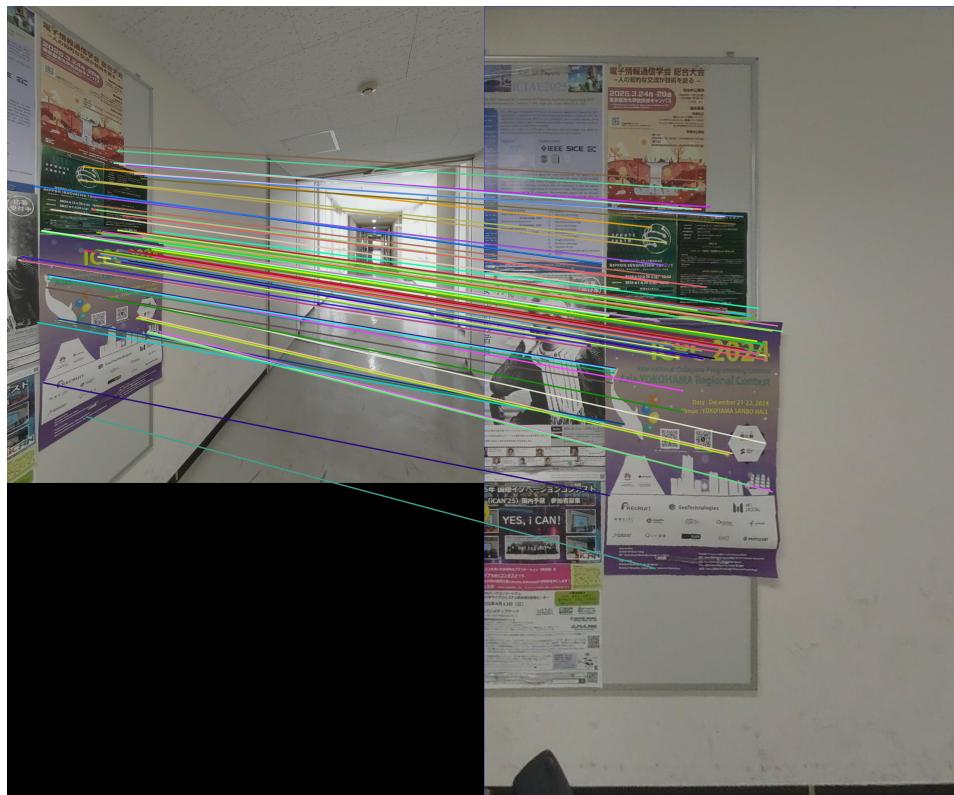


図 7: 特徴点マッチング

ユーザの入力画像と3次元モデルのテクスチャ画像間で、多くの特徴点を正しく対応付けできている。

## 4.3 自己位置推定の結果

2次元-3次元特徴点の対応付け結果をもとに、自己位置推定を行い、推定結果をもとに入力画像に3次元モデルのワイヤーフレームを描画した様子を図 8 に示す。



図8：自己位置推定の精度

実用上は問題ない程度の精度でモデルを描画することができている。

## 5 今後の計画

今後の研究計画を以下に示す。

### 1. 3次元モデルの適用範囲の拡張

- より広範囲の空間を再現し、テクスチャ画像がさらに増えた場合も正しく類似画像を推定できるか調査

### 2. 全方位動画を活用したテクスチャ自動取得

- 静止画ではなく全方位動画を入力として用い、適切なフレームを自動で抽出することで、効率的なテクスチャ収集を実現

### 3. 自己位置推定機能を応用した実用的なアプリケーションの開発

- 生成された3次元モデルと自己位置推定を組み合わせ、目的地までのルートを提示するナビゲーションシステムを構築

最後に、研究成果のバックアップは必ず取るようにしましょう。少なくとも月に一回は。



## 参考文献

- [1] MarketsandMarkets, 「Digital Twin Market Size, Share & Industry Trends Growth Analysis Report」
- [2] 国土交通省 不動産・建設経済局 情報活用推進課, 「屋内地図・屋内測位環境構築の手引き」 6p
- [3] 菅谷保之, 「直交射影の共線性と共面性を用いたカメラ姿勢の推定」