

OS5A-S2 :

# ミリ波信号を用いた人物姿勢推定における 高精度化及び推定シーン汎用化

2023.7.27

†天谷幸太郎, †‡五十川麻理子

†慶應義塾大学大学院 理工学研究科 ‡JSTさきがけ



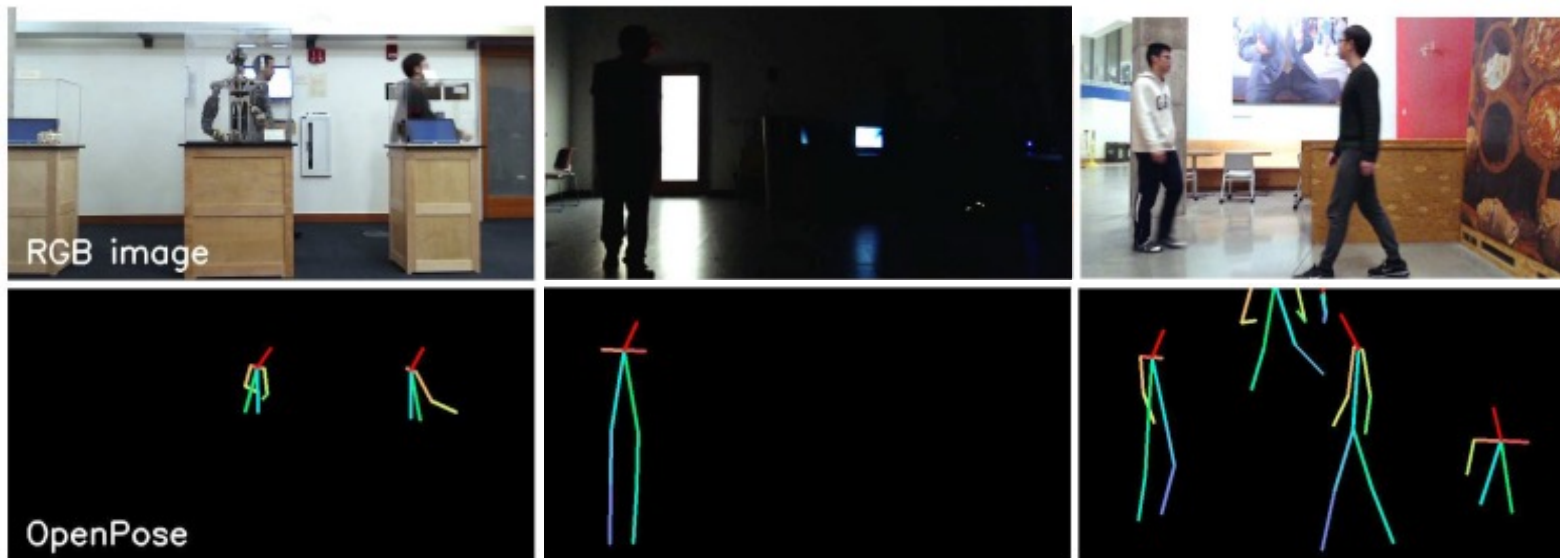
Keio University

1858

CALAMVS  
GLADIO  
FORTIOR

# RGB動画像ベース人物姿勢推定における課題点

- 可視光を通さない遮蔽に頑健でない
- 暗所では推定困難である
- プライバシーに配慮しにくい



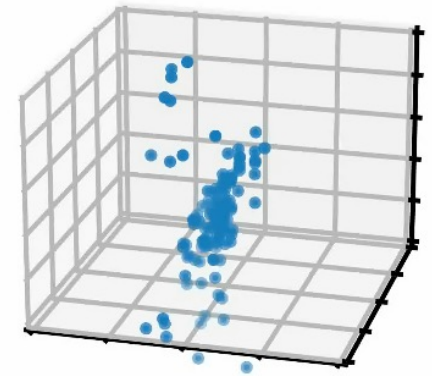
RGB動画像による推論の失敗例<sup>[1]</sup>

[1] Zhao, Mingmin, et al. "Through-wall human pose estimation using radio signals." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

# 無線信号を用いた人物姿勢推定

Wi-Fi(2~5GHz), ミリ波(76~81GHz)を使用した手法

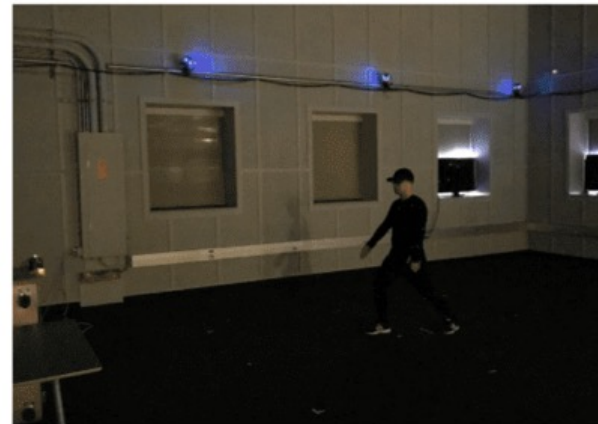
- 壁や布などの遮蔽を透過して推論可能
- 暗所でも使用可能
- プライバシーに配慮可能



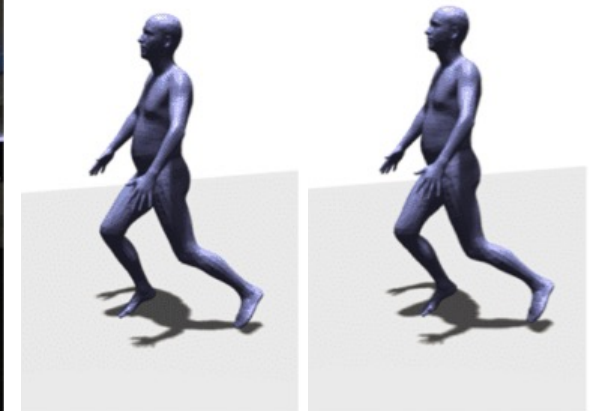
計測データの例



数GHzの無線信号を用いた骨格推定<sup>[1]</sup>



暗所におけるミリ波を用いた三次元メッシュ推定<sup>[2]</sup>



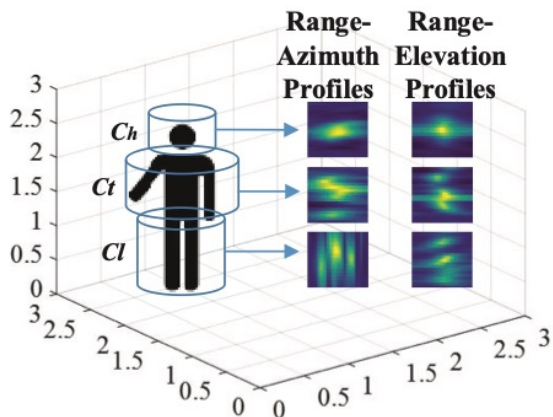
[2] Xue et al., “mmMesh: Towards 3D Real-Time Dynamic Human Mesh Construction Using Millimeter-Wave”, In MobiSys, pp.269-282, 2021.

# 先行研究

mmMesh<sup>[2]</sup>, m<sup>3</sup>Track<sup>[3]</sup> : ミリ波を使用した人物姿勢推定手法

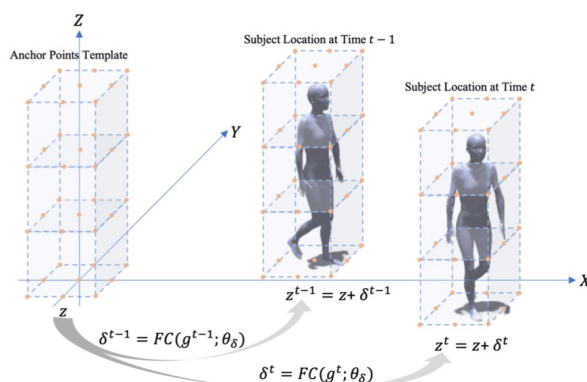
- mmMeshは直方体, m<sup>3</sup>Trackは円柱を特徴量抽出に使用  
→ 人物の姿勢が直立であることを前提としている
- 計測空間に人物のみが存在するデータを使用  
→ 人物の周りに人物以外の物体が存在する場合を考慮しない

m<sup>3</sup>Track



円柱による特徴量抽出

mmMesh



直方体による特徴量抽出



日常で一般的な姿勢, 環境

[3] Kong et al., "m3track: mmwave-based multi-user 3d posture tracking", in Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Services, pp. 491-503, 2022

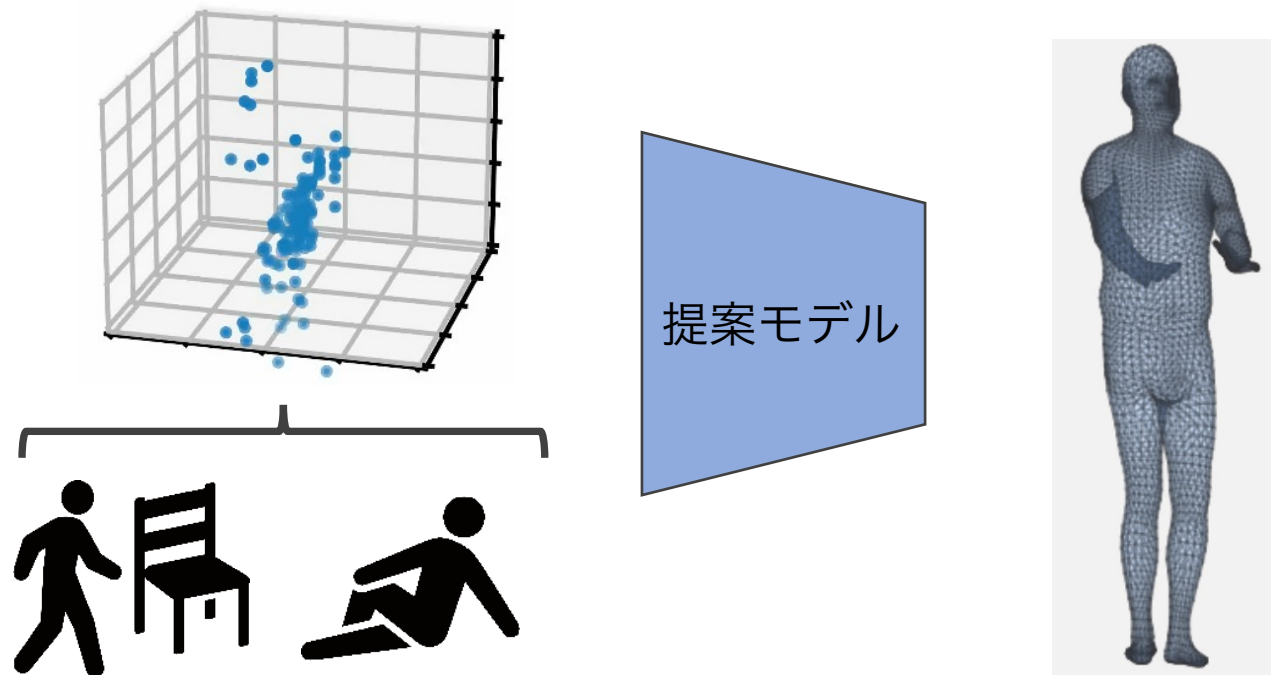
# 本研究の目的

## 目的

- 計測空間に人物以外の物体が存在する場合にも頑健な手法の実現
- 人物が直立以外の、水平に寝る、地面に座るなどの姿勢にも頑健な手法の実現

入力：ミリ波点群

出力：人物三次元メッシュ



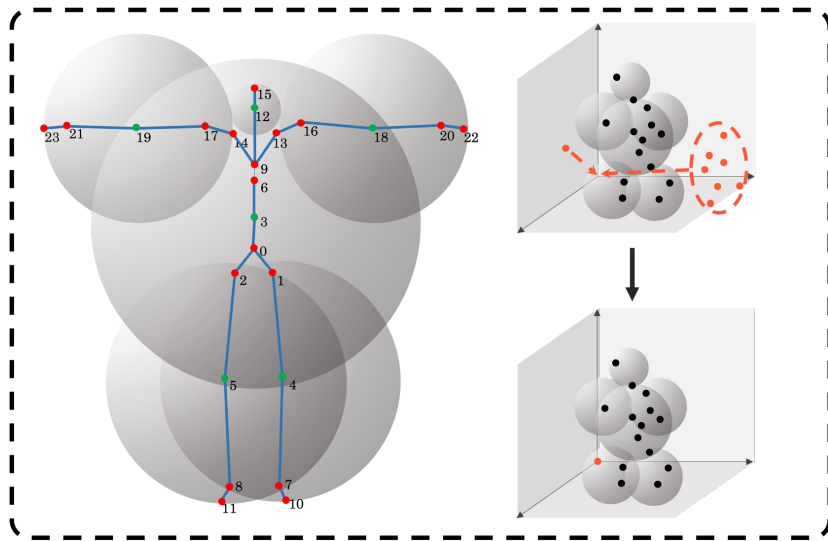
# 本研究の着眼点

## Key Point 1 :

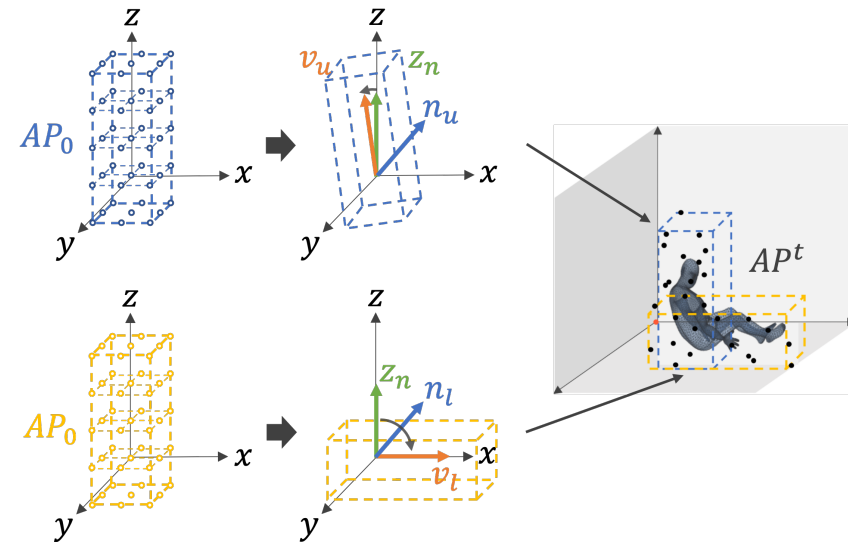
任意の姿勢で人間の可動域を完全に内包する球集合によるDenoise処理

## Key Point 2 :

上半身・下半身それぞれに分割された直方体を用いることによる，多様な姿勢に対応可能な特徴量抽出



球集合によるDenoise処理

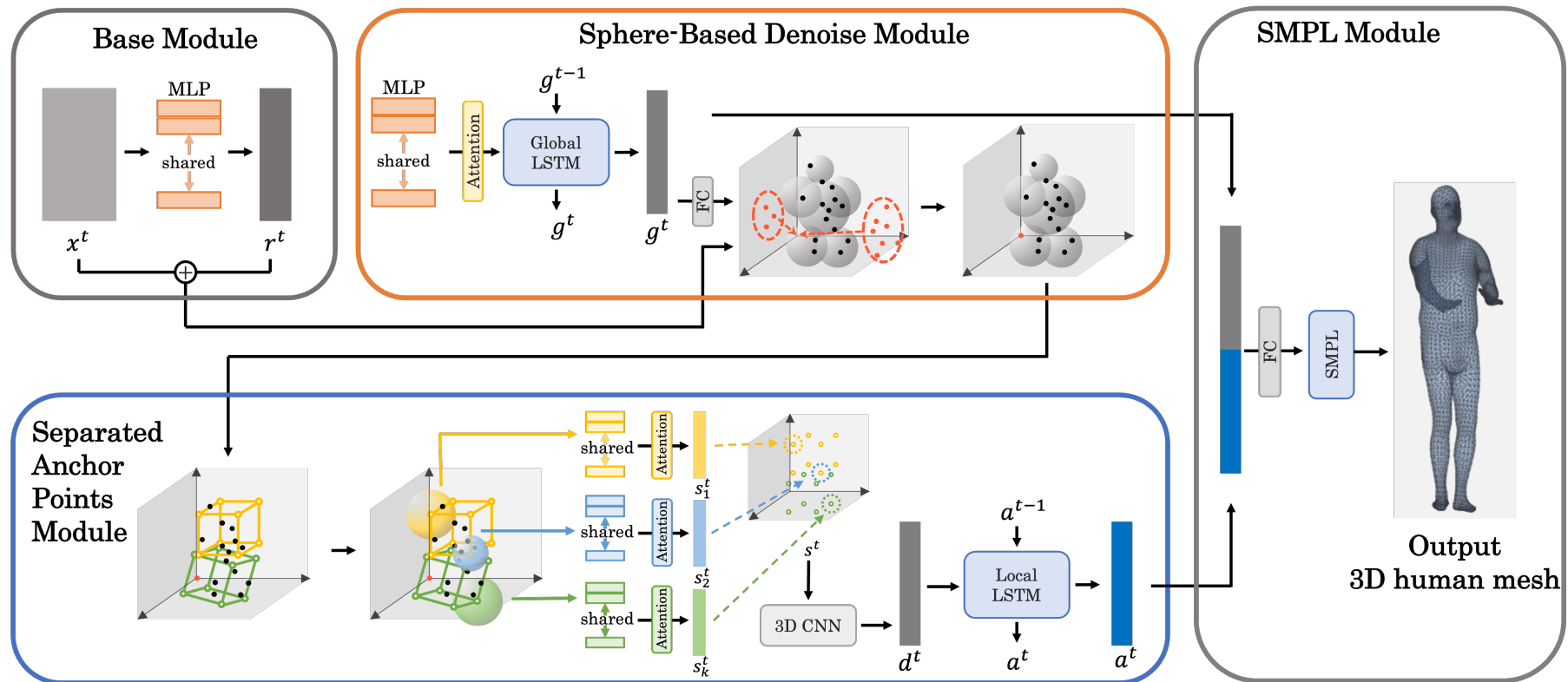


分割された直方体による特徴量抽出

# モデル構成

Base Module, Sphere-Based Denoise Module, Separated Anchor Points Module, SMPL Moduleで構成

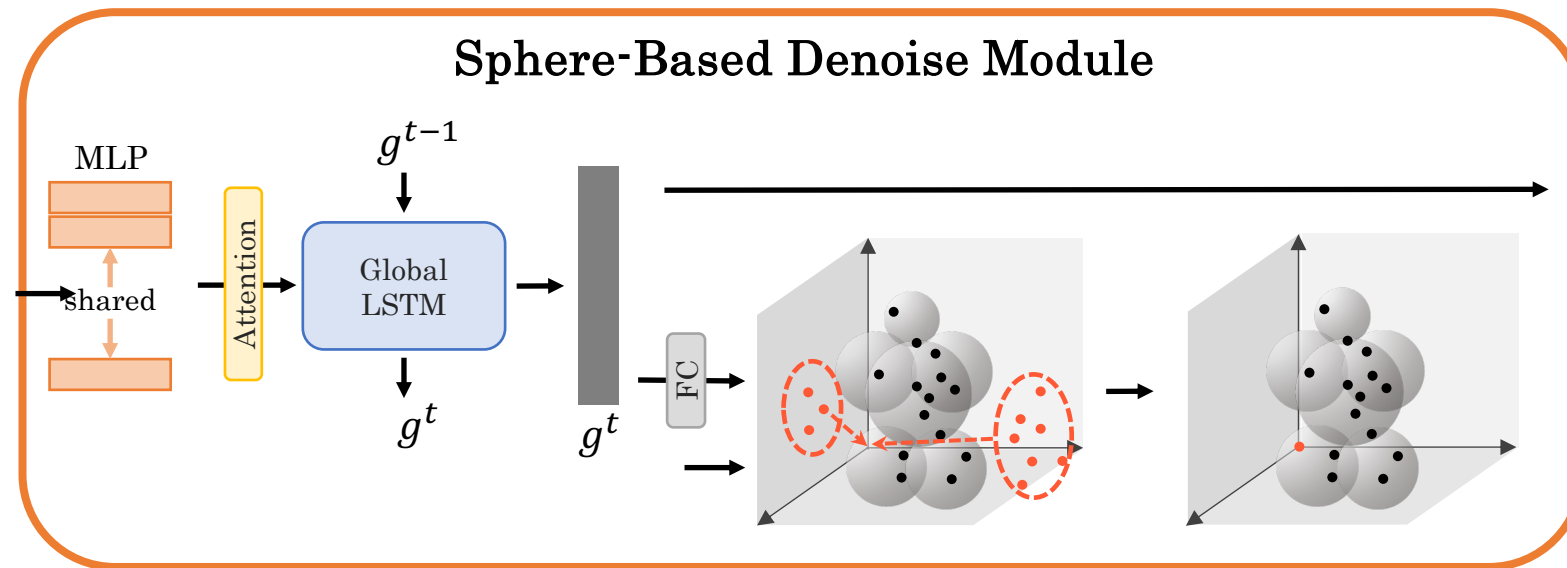
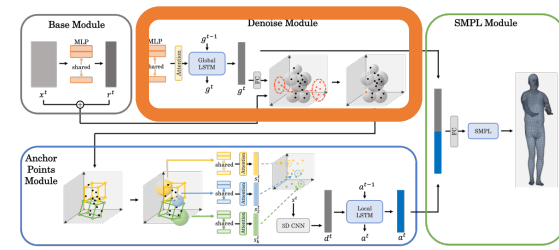
大域特徴量の抽出と球集合を用いたDenoise



分割された直方体を用いた局所特徴量の抽出

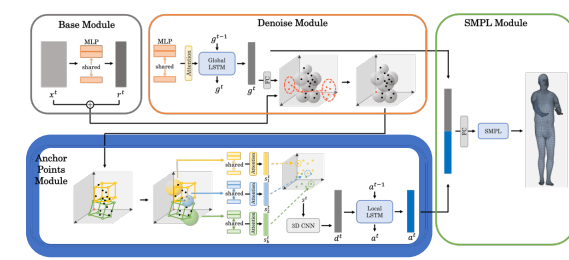
# Key Point 1: Sphere-Based Denoise Module

- Shared-MLP, Attention, LSTMにより特徴量を抽出  
→SMPL Moduleの入力とする
- さらに全結合層を通してそれぞれ6個の関節位置と半径rを推論する  
→6個の球体内に無い点群の値をノイズとして, 値を0にする

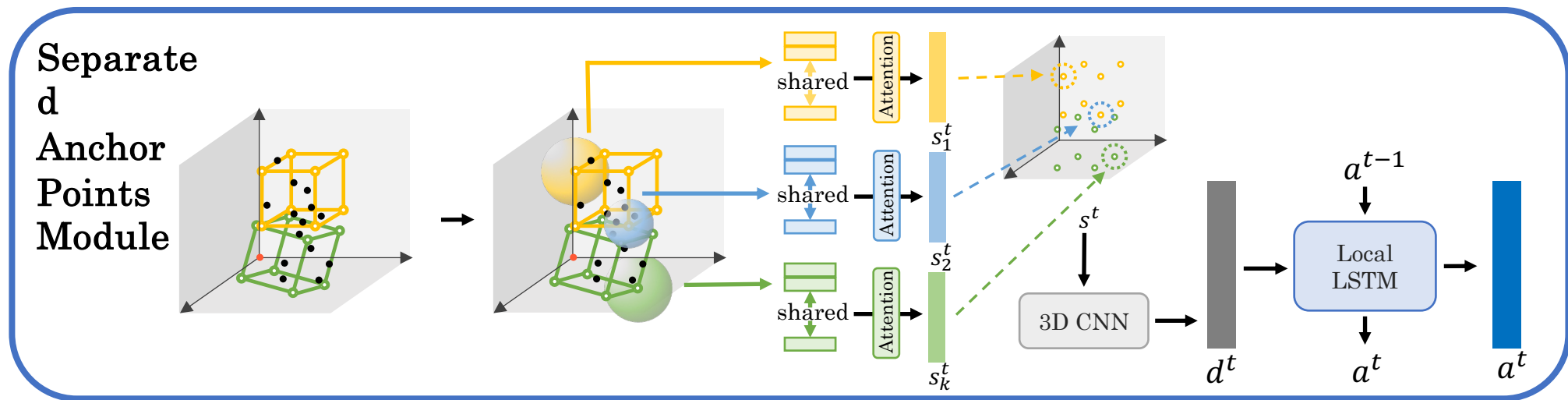


Sphere-Based Denoise Moduleのモデル図

# Key Point 2: Separated Anchor Points Module



- Denoise Moduleで求めた関節位置を用い， Separated Anchor Pointsを定義
- 二つの直方体を， 上半身， 下半身の向きに沿うように配置
- アンカーポイントごとに特徴量抽出後， 3DCNNで位置エンコーディング

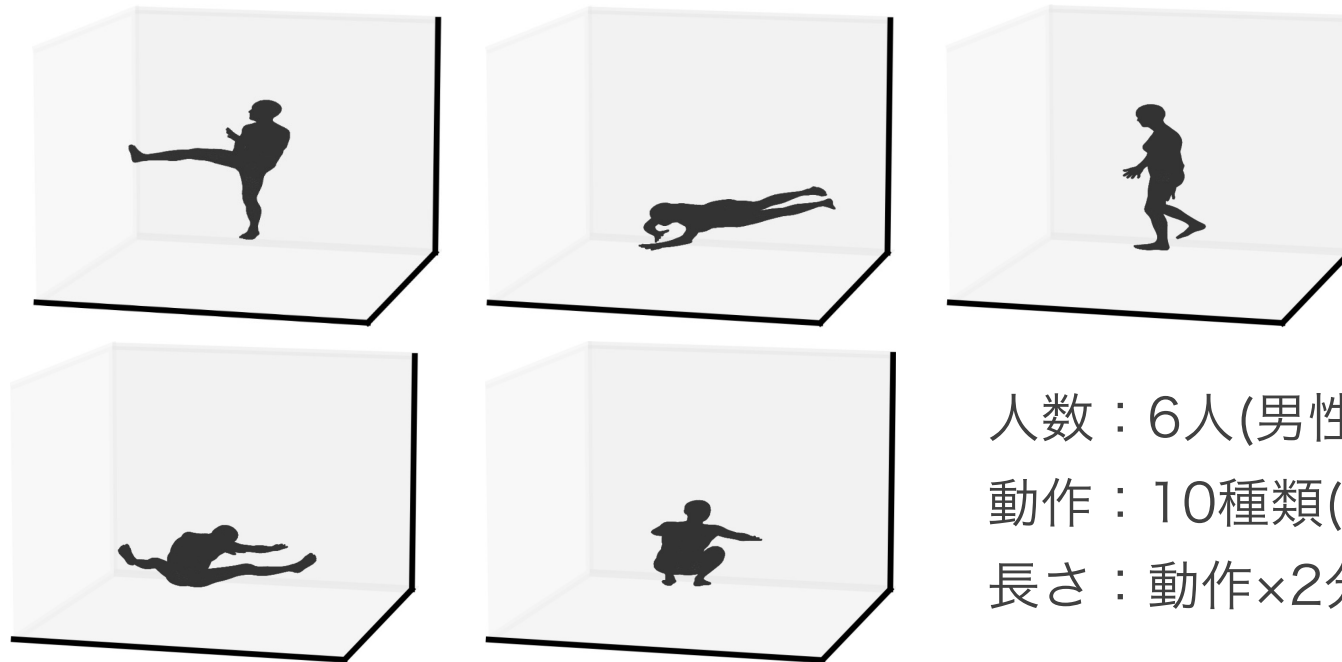


Separated Anchor Points Moduleのモデル図

# データセット

本研究用に、新たに以下の特徴を含むデータセットを独自に作成

- 計測空間に人物以外の物体が存在する場合
- 人物が直立以外の、水平に寝る、地面に座るなどの姿勢



人数：6人(男性4人，女性2人)  
動作：10種類(5種×箱あり,なし)  
長さ：動作×2分

データセットに含まれる姿勢

# 実験結果：定量的評価

全ての評価指標において、提案手法が同等または高精度

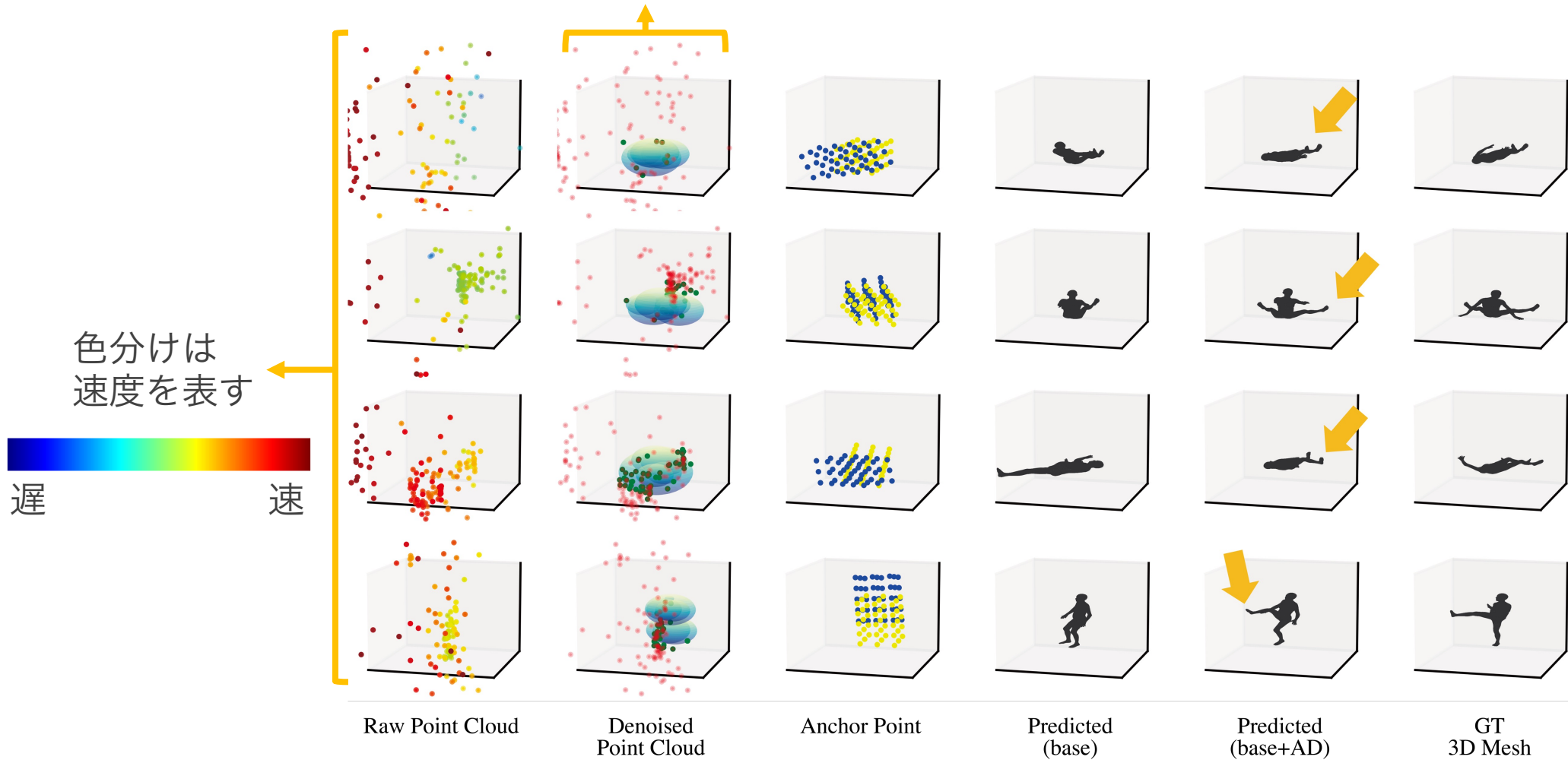
- 関節位置の誤差評価指標であるMPJPEにおいて、約10%精度が向上
- メッシュ頂点位置の誤差評価指標である $E_{ver}$ において約7%精度が向上

## 評価

Method	$E_{rot}$ (°)↓	$E_{ver}$ (cm)↓	MPJPE (cm)↓	PCKh @0.5↑	$E_{loc}$ (cm)↓	$E_{body}$ (cm)↓
Xue <i>et al.</i> [2]	33.8	35.7	33.2	0.23	30.1	<b>0.88</b>
Ours	<b>32.3</b>	<b>33.3</b>	<b>30.4</b>	<b>0.25</b>	<b>23.3</b>	<b>0.88</b>

# 実験結果：定性的な評価

赤：球体外部のDenoiseされた点群  
緑：球体内部の点群



# まとめ

暗所や遮蔽に頑健かつプライバシーに配慮できるミリ波ベース人物三次元メッシュ推定手法について、従来手法の課題点(環境ノイズに脆弱、推定可能な姿勢への汎用性の低さ)解決のため以下を提案

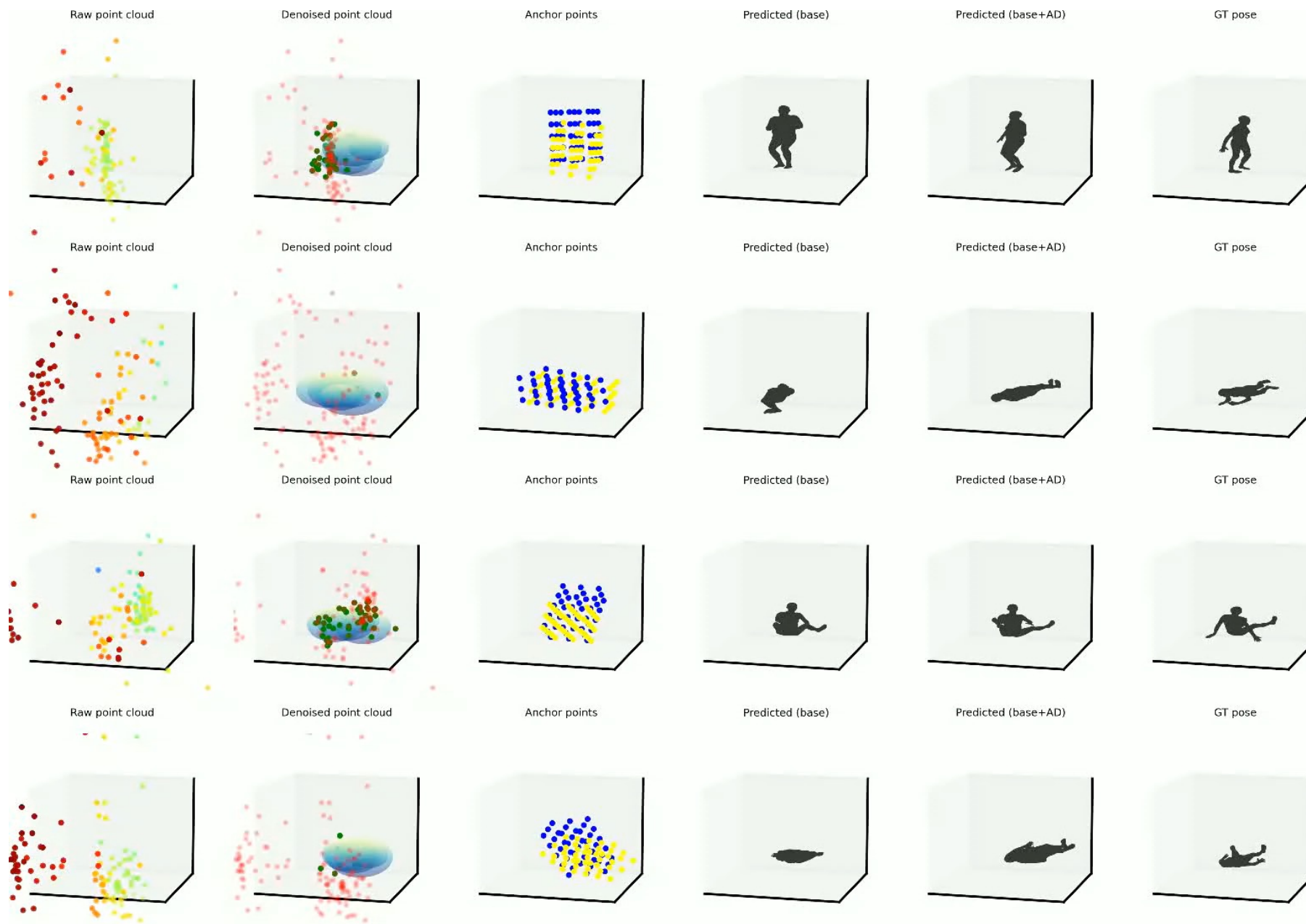
- 人物領域を任意の姿勢で抽出するSphere-based Denoise Module
- より多様な姿勢に対応するSeparated Anchor Points Module

今後の課題, 展望

- 精度良く推論可能な姿勢の汎用性のさらなる向上
- より表現力の高いAnchor Pointsの導入

# 補足資料

# 補足： 結果動画



Raw Point  
Cloud

Denoised  
Point Cloud

Anchor  
Point

Base+A+D  
(Ours)

Base  
(mmMesh)

Ground  
Truth

# 人物姿勢推定

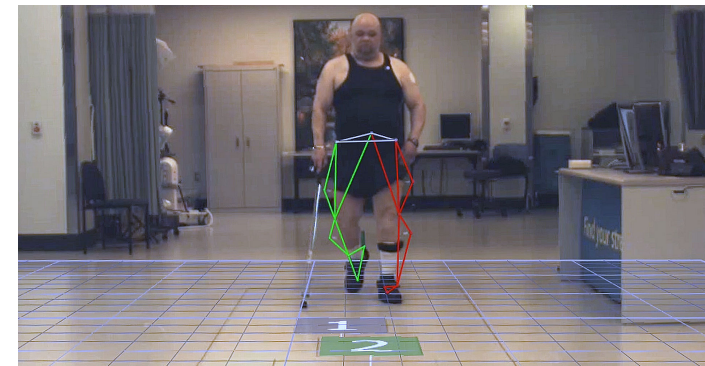
- 人物の姿勢を計測，推定しデータ化する手法
- 代表的な方式：光学式(赤外線センサ)，慣性センサ式，画像/動画式

## 人物姿勢推定の応用先

- 医療
- スポーツ科学
- AR / VR
- 映画，アニメ，ゲーム
- ロボット工学
- 自動運転



映画への応用<sup>[1]</sup>



医療への応用<sup>[2]</sup>

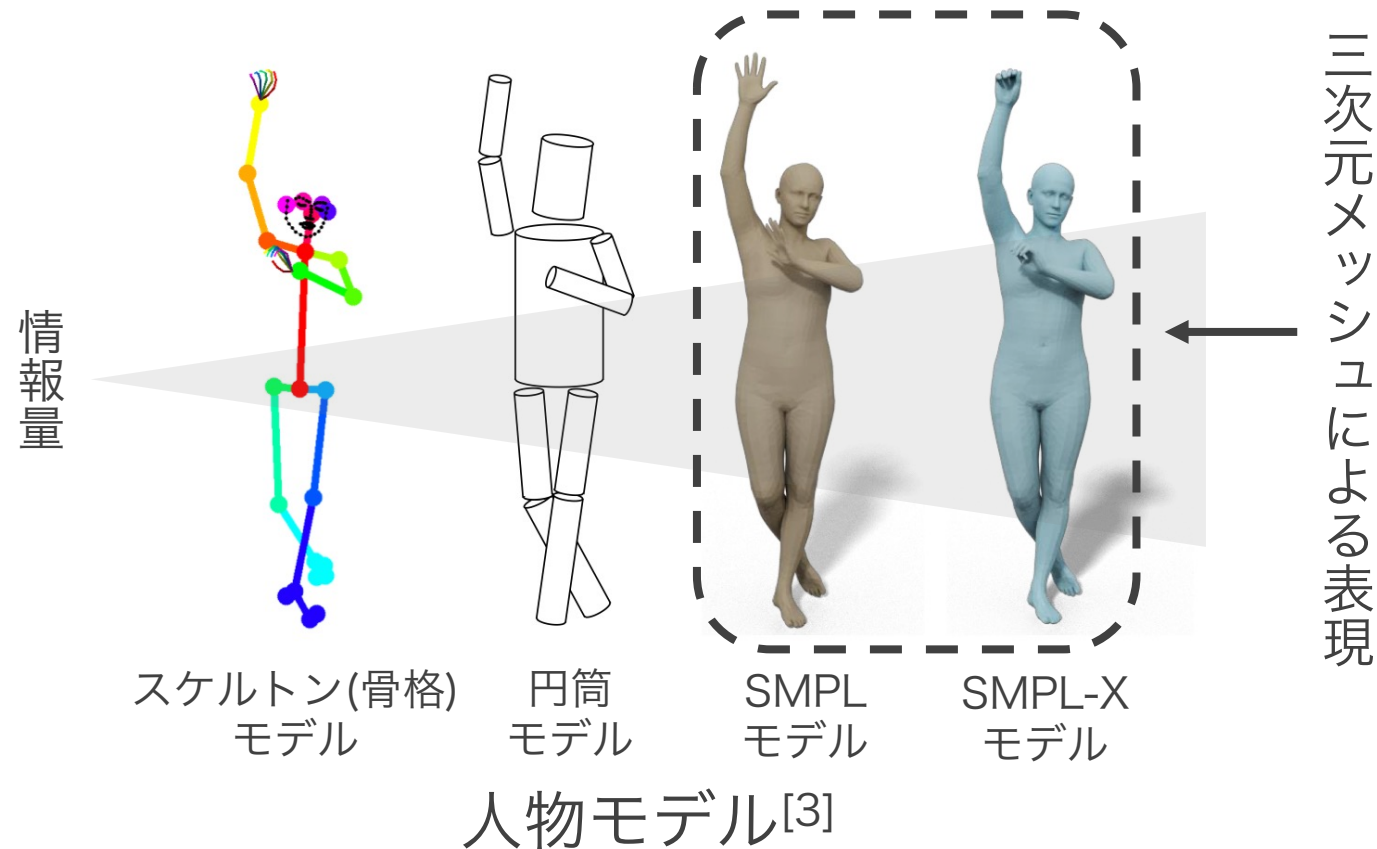
[1] Avatar: The Way of Water | Acting in The Volume, 2023, <https://www.youtube.com/watch?v=J3LMli3gG4>

[2] Watch: Hollywood motion capture technology finds a new role in hospital rehab, 2020,

<https://www.statnews.com/2020/02/12/motion-capture-technology-hospitals-physical-rehabilitation/>

# 姿勢推定に使われる人物モデル

- 人物モデルには, 骨格だけでなく形状も推定するモデルがある
- 詳細に人物を表せるほど応用先が広がる → 本研究ではSMPLモデルを使用

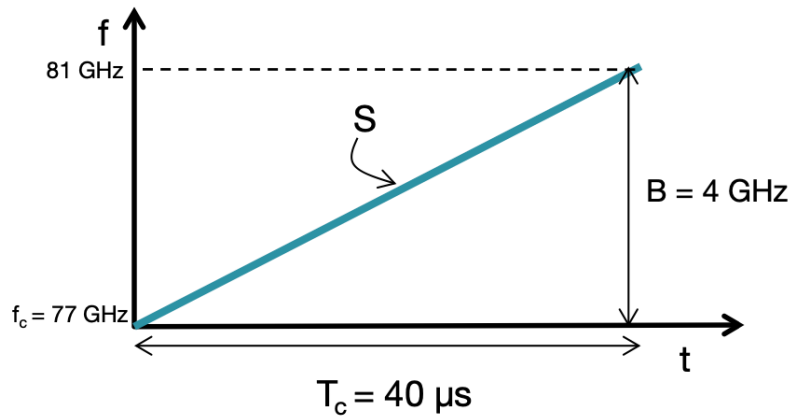


# 補足：ミリ波センサ

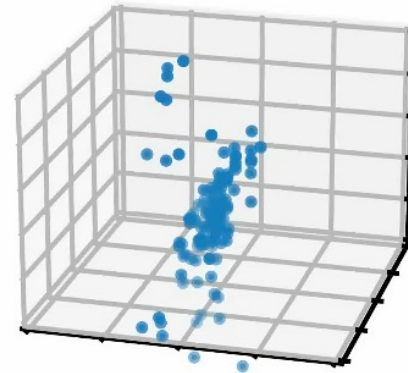
- 複数の送信アンテナからミリ波チャープ信号を送信
- 複数の受信アンテナで反射信号を受信



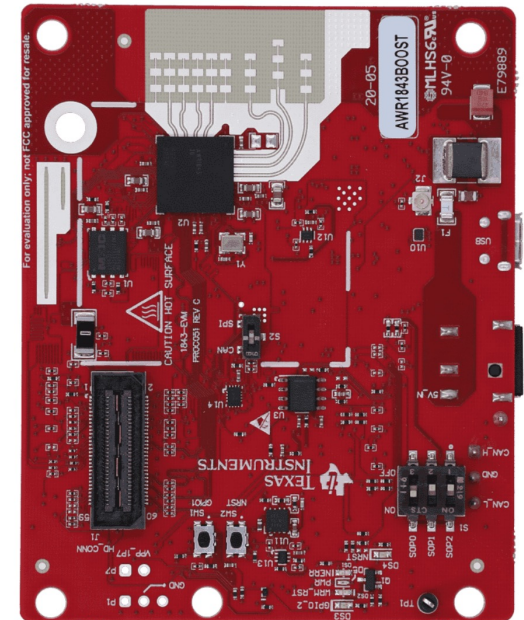
三次元位置, 速度, 反射強度, 範囲の情報が得られる



チャープ信号[6]



得られる点群



ミリ波センサ[7]

[7] Texas Instruments AWR1843BOOST, <https://www.ti.com/tool/ja-jp/AWR1843BOOST>, 2023年1月閲覧

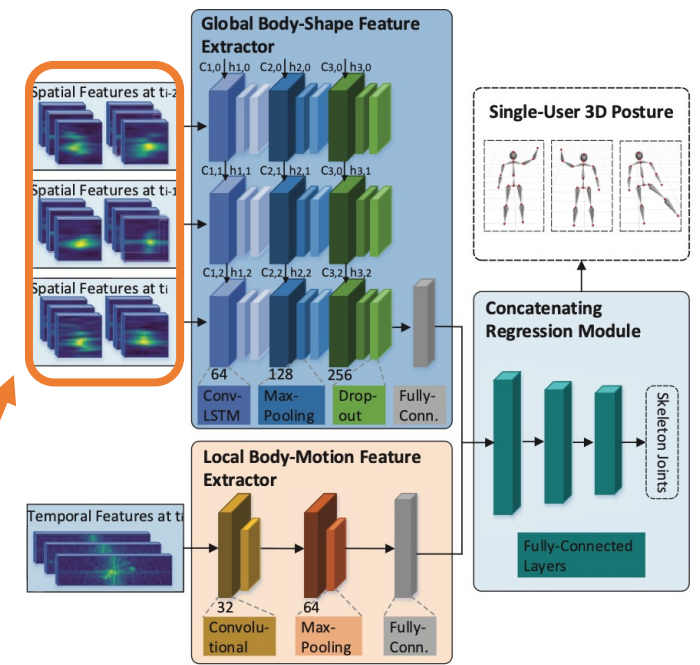
# 関連研究：m<sup>3</sup>Track<sup>[8]</sup>

複数人の人物三次元骨格の復元手法

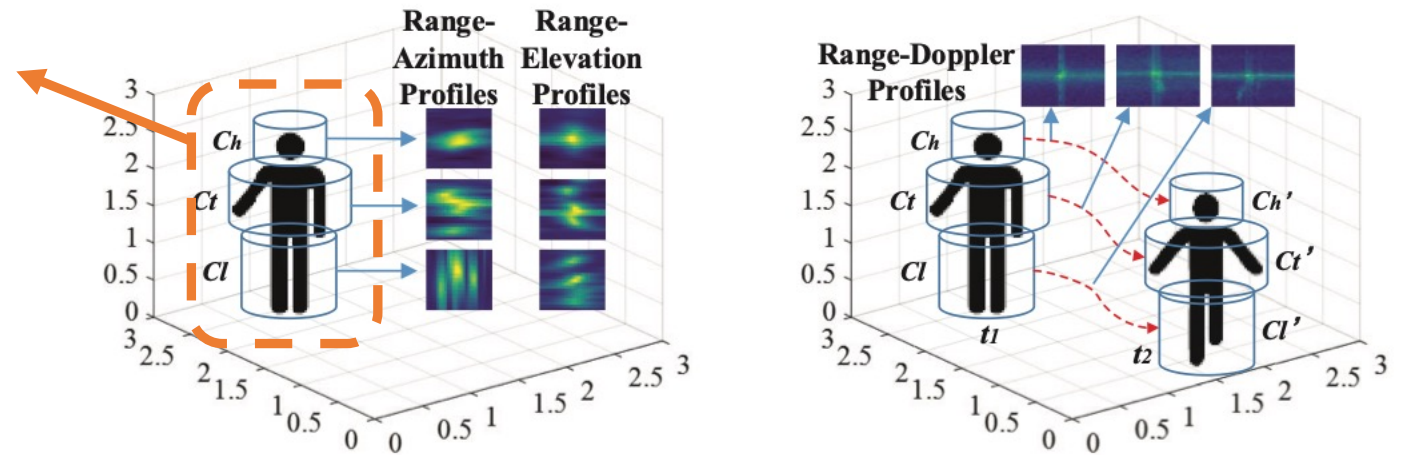
- データを頭, 胴, 下半身に分けた特徴量抽出
- 人物位置を利用したデータ抽出

課題点

- 人物の姿勢が直立であることを前提とした特徴抽出
- 計測空間に人物のみが存在することを前提とする



モデル構成



円柱による信号抽出

[8] Kong et al., “m3track: mmwave-based multi-user 3d posture track”, g,” in Proceedings of the 20th Annual International Conference on Mobile Systems, Applications and Ser-vices, pp. 491–503, 202

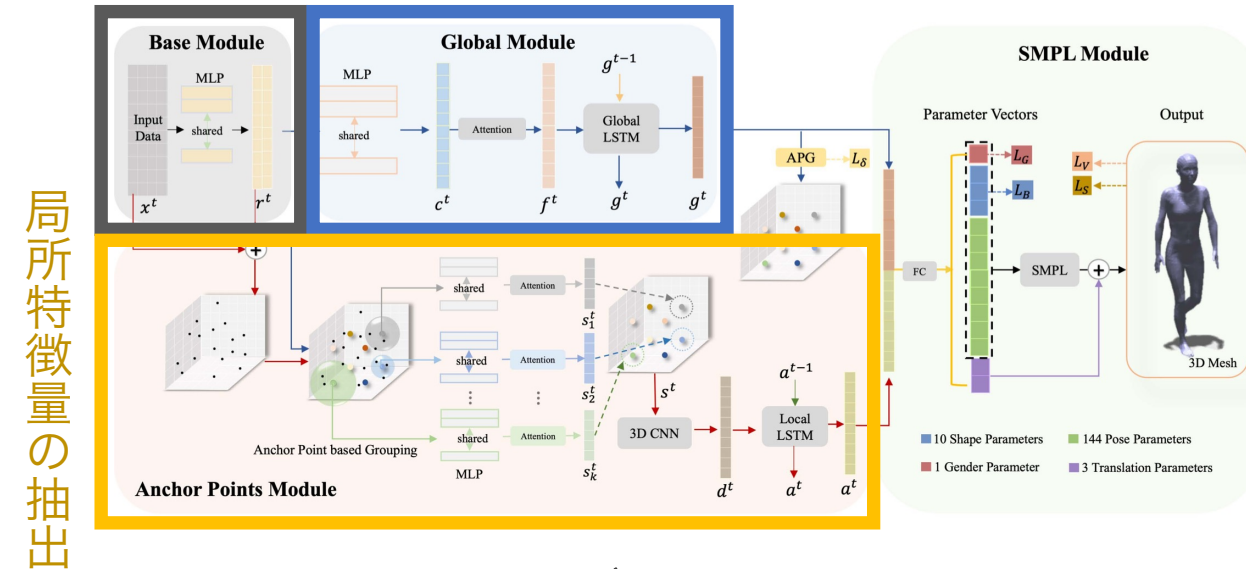
# 関連研究：mmMesh<sup>[6]</sup>

- 人物三次元メッシュの復元手法
- Shared-MLPを用いた点群処理
- 縦長の直方体状に配置されたアンカーポイントに点群を集約

## 課題点

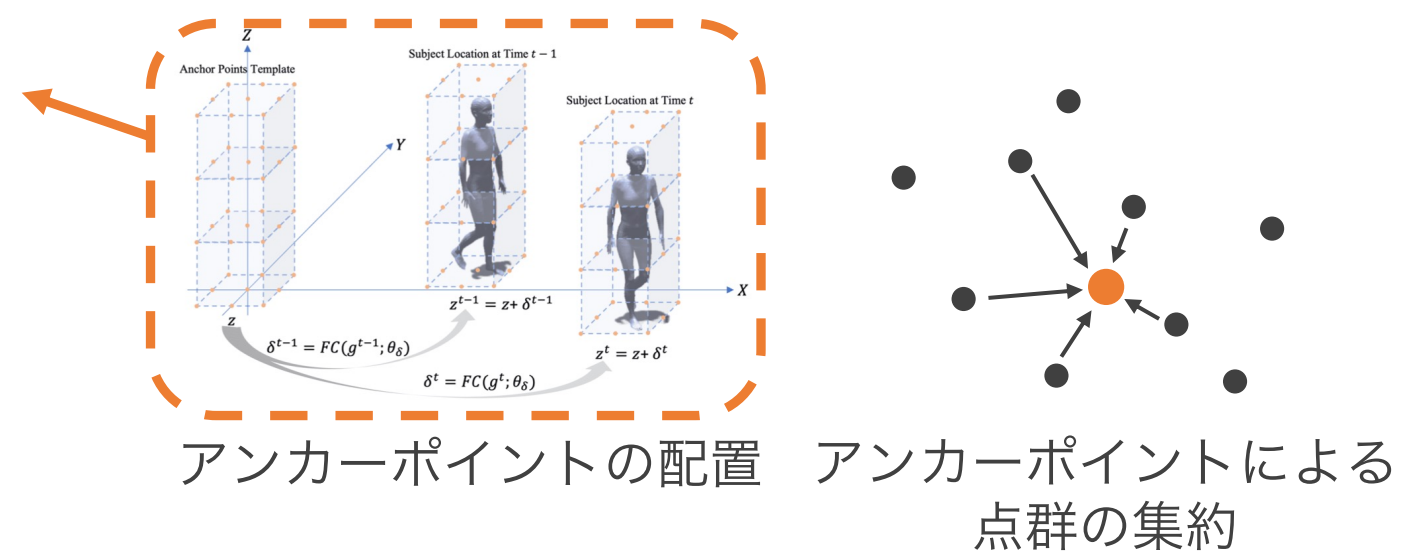
- 人物の姿勢が直立であることを前提とした特徴抽出
- 計測空間に一人の人物のみが存在することを前提とする

## 点群情報の拡張 大域特徴量の抽出



局所特徴量の抽出

## モデル構成

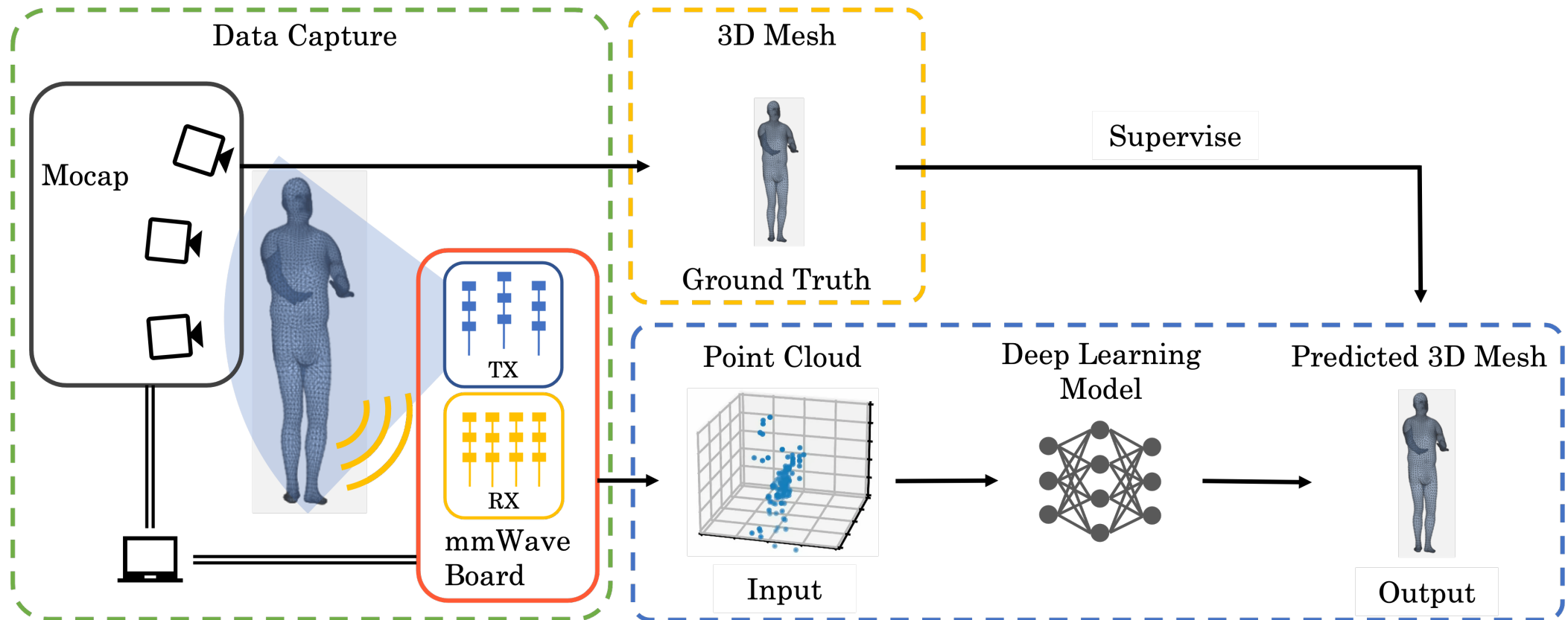


アンカーポイントの配置

アンカーポイントによる点群の集約

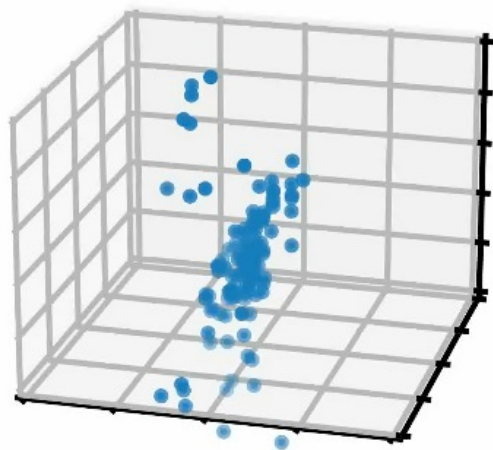
# 提案フレームワーク

- Mocapとミリ波センサを用いてデータセットを作成
- Mocapを教師データ, ミリ波を学習データとする

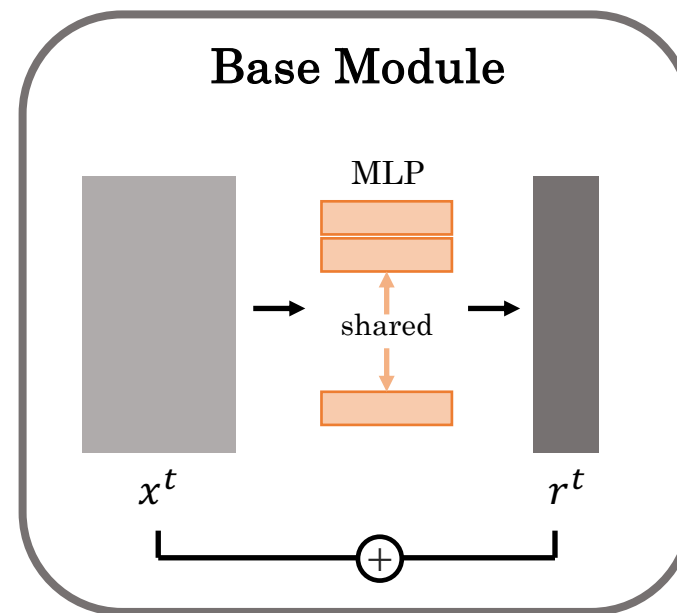


# 補足：Base Module

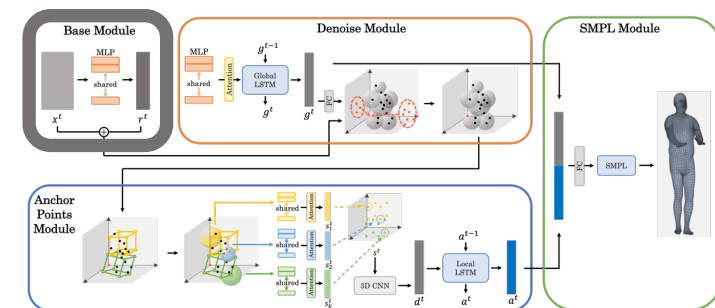
- $x_i^t$ は(x, y, z, 範囲, 速度, 反射強度)の情報をもつ
- ミリ波センサで計測された点群データ $x^t$ を Shared-MLPによって高次元の特徴ベクトル $r^t$ に変換
- $x^t$ と $r^t$ を結合させた点群を出力



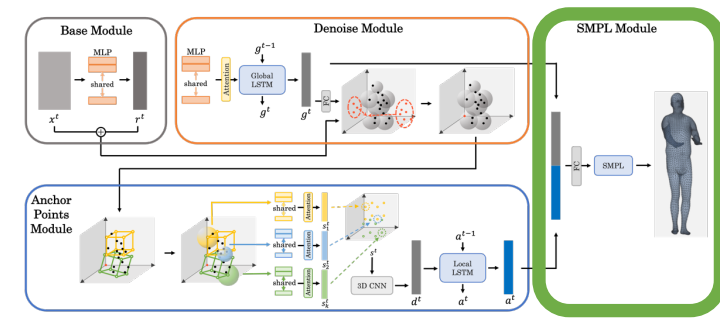
点群(x,y,z)により可視化した点群



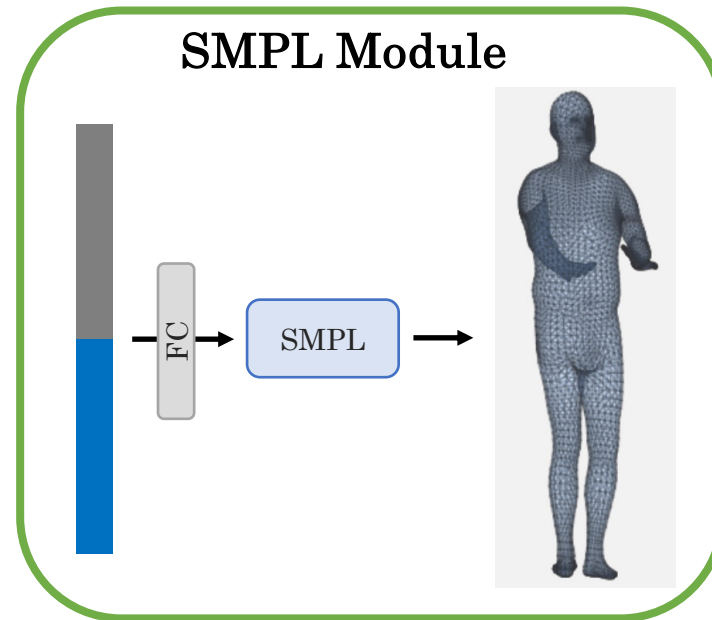
Base Moduleのモデル図



# 補足：SMPL Module



- Denoise Module, Anchor Point Moduleで求めた特徴量を連結
- 特徴量から全結合層によってSMPLパラメータを推論
- 学習済みのSMPLモデルを使用して三次元メッシュを推論



SMPL Moduleのモデル図

# 補足：損失関数

$$Loss = \sum_{K \in (V, S, B, j)} \alpha_K * \sum_t^T \|K^t - GT(K^t)\|_{L1} + \alpha_G * \sum_t^T H(G^t, GT(G^t)) + L_l + L_r$$

$V$ ：メッシュの頂点位置  
 $S$ ：骨格の位置  
 $B$ ：SMPLの体格パラメータ  
 $j$ ：骨盤位置  
 $G$ ：性別  
 $H$ ：Hinge Loss

$$L_l^t = \frac{1}{6} (\|j_{xyz}^t - T_1^t\|_{L1} + \sum_{i=2}^6 \|j_{xyz}^t - Sk_i^t\|_{L1})$$
$$L_r^t = \frac{1}{6} \sum_{i=1}^6 (r_i^{sk} - r_i^j)$$

Denoise Moduleで関節位置を教師あり学習

# 補足：評価指標

## ➤ 評価指標

- Average Vertex Error ( $E_{ver}$ )
  - 真値と推論の, **メッシュ頂点**のユークリッド距離誤差平均
- Average Joint Rotation Error ( $E_{rot}$ )
  - 真値と推論の, **関節角度**の誤差平均
- Mean Per Joint Position Error (MPJPE)
  - 真値と推論の, **関節位置**のユークリッド距離誤差平均
- Percentage of Correct Key Points (PCKh@0.5)
  - 真値の各関節から頭の大きさの半分の距離の範囲に, 推論した関節がある確率
- Mesh Localization Error ( $E_{loc}$ )
  - メッシュの**骨盤位置**のユークリッド距離誤差

# 補足：評価指標の数式

## ➤ 評価指標

以下， $T$ は時刻， $J$ は関節数， $v$ は頂点位置， $\theta$ はルート関節の回転ベクトル， $x$ は関節位置， $l$ はルート関節位置である。

- Average Vertex Error ( $E_{ver}$ )
  - $\frac{1}{TJ} \sum_{t=1}^T \sum_{j=1}^J \|v_t^j - \hat{v}_t^j\|_2$
- Average Joint Rotation Error ( $E_{rot}$ )
  - $\frac{1}{TJ} \sum_{t=1}^T \sum_{j=1}^J |\theta_t^j - \hat{\theta}_t^j|$
- Mean Per Joint Position Error (MPJPE)
  - $\frac{1}{TJ} \sum_{t=1}^T \sum_{j=1}^J \|x_t^j - \hat{x}_t^j\|_2$
- Mesh Localization Error ( $E_{loc}$ )
  - $\frac{1}{TJ} \sum_{t=1}^T \sum_{j=1}^J \|l_t^j - \hat{l}_t^j\|_2$