

卒業論文概要書

Summary of Bachelor's Thesis

Date of submission: 01/30/2024 (MM/DD/YYYY)

学科名 Department	情報通信	氏名 Name	鳥羽祐哉	指導 教員 Advisor	渡辺 裕 印
研究指導名 Research guidance	オーディオビジュアル 情報処理研究	学籍番号 Student ID number	1W202234-3		
研究題目 Title	仮想マーカーによる人体 3D メッシュ作成モデルの損失関数の各要素の重みの調査 Investigation of Each Element's Weight of Loss Function for Human Body 3D Meshing Model with Virtual Markers				

1. 研究背景・目的

現在, 3D モデル作成技術は大きな注目を集めている. その中でも人体 3D モデルは, スポーツ分野や医療分野等, 多くの分野で重要視されている. 人体 3D モデル作成の中間表現として 3D 骨格情報を使用し, 人体 3D メッシュを再構成する手法が存在する.

骨格情報を中間表現として利用し人体 3D メッシュを推定する際の問題点として, 体型情報が失われることが挙げられる. そのため, やせ型や肥満型などの異なるタイプの体型を正確に推定することができない. 3D Human Mesh Estimation from Virtual Marker (HMEVM) [1] モデルは, この問題点を解決するために, 仮想的なマーカーを用いて人体 3D メッシュを作成する手法を提案した.

本研究では, HMEVM で用いられる損失関数の各要素が性能に及ぼす影響について調査する.

2. HMEVM

HMEVM は, 仮想マーカーを用いて人体骨格情報から人体 3D メッシュを作成する技術である. 入力された画像・動画から人物の骨格情報を抽出する. HMEVM の概要を図 1 に示す.

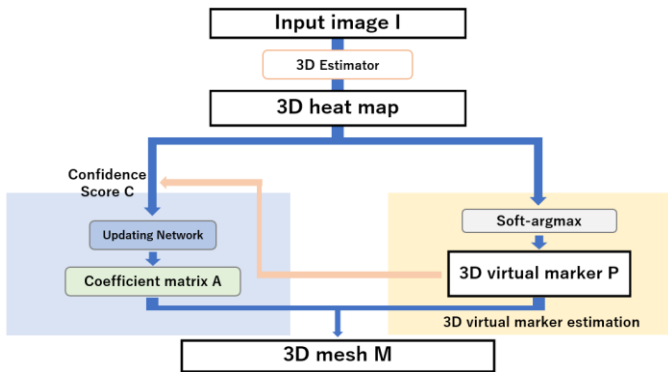


図 1 HMEVM の概要

HMEVM の構造は, 仮想マーカー生成, 仮想マーカー信頼度スコア推定, メッシュ作成の三つに分けられる. HMEVM の学習で用いられる損失関数を式 (1) に示す.

$$L = \lambda_{vm}L_{vm} + \lambda_cL_{conf} + \lambda_mL_{mesh} \quad (1)$$

L_{vm} は仮想マーカー生成の損失, L_{conf} は仮想マーカー信頼度スコア推定の損失, L_{mesh} はメッシュ作成の損失である. また, HMEVM の公式実装内で, 全体の損失関数の各要素に掛かっている重み λ を式 (2), (3), (4) に示す.

$$\lambda_{vm} = 100 \quad (2)$$

$$\lambda_c = 0.1 \quad (3)$$

$$\lambda_m = 0.1 \quad (4)$$

3. 重み λ と HMEVM の性能の関係調査

HMEVM の損失関数が持つ三つの要素の各重み λ を変化させたときのモデルの性能を評価する. 評価指標には Mean Per Vertex Error (MPVE) と Mean Per Joint Position Error (MPJPE) を用いる.

MPVE は生成されたメッシュの各頂点の位置と対応する実際のメッシュ頂点の距離を計算し, それらの距離の平均をとったものである. MPJPE はモデルが予測した各関節の位置と, 対応する実際の関節の距離を計算し, それらの距離の平均をとったものである. また, 各要素がそれらの値に及ぼす影響力を調査する.

4. 実験

4.1 データセット

本研究で用いたデータセットは, 3D Poses in the Wild (3DPW) [2] である. データセットには 60 個の動画が含まれており, フレーム数の合計は 51,000 枚である. また, これらのフレームに対応

した二次元骨格情報，三次元身体情報，三次元人物モデルが含まれている。

4.2 実験結果

各要素の重み λ を変化させたときの，MPVE，MPJPE の値を表 1，2，3 に示す。

実験結果より，MPVE，MPJPE の値が実験内で最良となる各要素の重み λ の値は，公式実装と同じものとなることが分かる。

4.3 各要素が評価に与える影響度

結果より，MPVE，MPJPE の値が実験内で最良となる際の損失関数は式 (5) となる。

$$L = 100L_{vm} + 0.1L_{conf} + 0.1L_{mesh} \quad (5)$$

また，各損失の最大値は $L_{vm_MAX} = 0.85$ ， $L_{conf_MAX} = 4.35$ ， $L_{mesh_MAX} = 364.73$ となった。

式 (5) の状態では損失範囲が考慮されていないため，損失範囲を $[0, 1]$ にしたものを式 (6) に示す。

表 1 各要素の重み λ の変化と評価 (1)

λ_{vm}	λ_c	λ_m	MPVE(↓)	MPJPE(↓)
0.1	0.1	0.1	380.5	347.3
1	1	1	450.5	415.9
10	10	10	348.4	309.1
100	100	100	289.3	271.1

表 2 各要素の重み λ の変化と評価 (2)

λ_{vm}	λ_c	λ_m	MPVE(↓)	MPJPE(↓)
100	1	1	129.2	113.8
1	100	1	448.6	413.7
1	1	100	309.2	281.6

表 3 各要素の重み λ の変化と評価 (3)

λ_{vm}	λ_c	λ_m	MPVE(↓)	MPJPE(↓)
100	0.01	0.01	123.7	108.7
100	0.1	0.1	118.0	103.9
100	1	1	129.2	113.8
100	10	10	244.7	275.1

表 4 仮想マーカー生成，仮想マーカー信頼度スコア推定，メッシュ作成それぞれの影響度

仮想マーカー生成	仮想マーカー信頼度スコア推定	メッシュ作成
69.9%	0.4%	29.9%

$$L = 100 \cdot 0.85 \cdot \frac{L_{vm}}{0.85} + 0.1 \cdot 4.35 \cdot \frac{L_{conf}}{4.35} + 0.1 \cdot 364.73 \cdot \frac{L_{mesh}}{364.73}$$

$$= 85L'_{vm} + 0.435L'_{conf} + 36.473L'_{mesh} \quad (6)$$

式 (6) における各要素の重みが損失の範囲を考慮した重みとなる。よって，この三つの重みの比率が MPVE，MPJPE の精度に与える影響度と考えられる。ここで影響度全体を 100 とした時の比率を表 4 に示す。

表 1 の重み $\lambda = 1$ の場合の MPVE，MPJPE と表 2 の MPVE，MPJPE を比較すると， λ_{vm} の値を大きくした時が最も MPVE，MPJPE が向上しており， λ_c を大きくした時は， $\lambda = 1$ の場合と差がほとんどない。表 4 より， λ_{vm} ， λ_m ， λ_c の順番で MPVE，MPJPE が向上しており，表 2 と同様の結果を示していると確認できる。したがって，表 4 の結果は妥当であると言える。

5. まとめ

本研究では，各要素の重み λ を変更した時の MPVE，MPJPE の値について確認した。また，各要素の損失の範囲を考慮し，各要素が最終的な評価指標に及ぼす影響度を求めた。その結果，MPVE，MPJPE の値が，実験内で最良となる重みの値は $\lambda_{vm} = 100$ ， $\lambda_c = 0.1$ ， $\lambda_m = 0.1$ であり，仮想マーカー生成の影響度は 69.9%，仮想マーカー信頼度スコア推定の影響度は 0.4%，メッシュ作成の影響度は 29.9% であった。

参考文献

- [1] X. Ma, J. Su, C. Wang, W. Zhu and Y. Wang, "3D Human Mesh Estimation from Virtual Markers," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 534-543, Jun. 2023.
- [2] T. V. Marcard, R. Henschel, M. J. Black, B. Rosenhahn and G. Pons-Moll, "Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera," European Conference on Computer Vision (ECCV), pp. 601-617, Sep. 2018.

2023 年度 卒業論文

仮想マーカーによる人体 3D メッシュ作成モデルの
損失関数の各要素の重みの調査

Investigation of Each Element's Weight of Loss Function for
Human Body 3D Meshing Model with Virtual Markers

指導教員 渡辺 裕 教授

提出日：2024 年 1 月 30 日

早稲田大学 基幹理工学部 情報通信学科

1W202234-3

鳥羽 祐哉

目次

第1章	序論.....	1
1.1	研究背景.....	1
1.2	研究目的.....	1
1.3	本論文の構成.....	2
第2章	関連技術.....	3
2.1	まえがき.....	3
2.2	HMEVM.....	3
2.3	MoSh.....	5
2.4	むすび.....	5
第3章	重み Λ とHMEVMの性能の関係調査.....	6
3.1	まえがき.....	6
3.2	概要.....	6
3.3	MPVE, MPJPEの概要.....	6
3.4	各要素の損失の範囲を考慮した損失関数.....	6
3.5	むすび.....	7
第4章	実験結果・考察.....	8
4.1	まえがき.....	8
4.2	データセット.....	8
4.3	実験結果.....	8
4.4	各要素の損失の範囲.....	9
4.5	各要素が評価に与える影響度.....	10
4.6	考察.....	11
4.7	むすび.....	11
第5章	結論・今後の課題.....	12
5.1	結論.....	12
5.2	今後の課題.....	12
謝辞	13
参考文献	14
図一覧	15
表一覧	16

第1章 序論

1.1 研究背景

現在, VR や 3D プリント技術の普及により, 3D モデル作成技術は大きな注目を集めている. その中でも特に動作解析や事前シミュレーションにおいて利用される人体 3D モデルは, スポーツ分野, 医療分野等, 多くの分野で重要視されている.

人体のような複雑な 3D モデルを扱う場合, 計算負荷軽減の為にメッシュの最適化が重要なステップとされている. 近年では, 人体 3D モデル作成の中間表現として 3D 骨格情報を使用し, 人体 3D メッシュを再構成する手法が提案されている. 例えば, 人体 3D メッシュを再構成するシステムの一つである Image-to-lixel meshnet [1] では全ての頂点の 3D 位置を回帰することを提案している. しかし, 人体 3D メッシュには数千以上の頂点があるため, 計算コストが非常に高くなる. この問題点を解決するために, Choi らは 2D 骨格情報から直接, 人体 3D メッシュ頂点座標を推定するシステムである Pose2Mesh [2] を提案した. しかし, この提案には 3D 骨格情報を抽出する際に, 体型に関する情報が失われてしまう問題点がある.

1.2 研究目的

骨格情報を中間表現として利用し, 人体 3D メッシュを推定する際の問題点として, 3D 骨格情報を抽出する際に, 体型情報が失われることが挙げられる. そのため, やせ型や肥満型などの異なるタイプの体型を正確に推定することができない. この問題を解決するための手法として, 身体表面に物理的なマーカーを配置する手法が存在する. しかし, この手法では物理的なマーカーを使用するため, 限られた環境での使用に制限される. また, 専門的な道具や知識が必要となる. そこで仮想的なマーカーを推定し, 3D 骨格情報を抽出することで体型情報を失わずに人体 3D メッシュを作成する 3D Human Mesh Estimation from Virtual Marker (HMEVM) [3] が提案された.

HMEVM は, トレーニングの際に仮想マーカー生成の評価, 仮想マーカー信頼度スコア推定の評価, メッシュ作成の評価の三つの評価を損失として利用する. しかし, 文献内では各損失がメッシュ作成性能に与える影響の比重について言及されていない. そのため, 本研究では, HMEVM で用いられる損失関数の各要素が HMEVM の評価指標に及ぼす影響について調査する.

1.3 本論文の構成

本論文の構成を以下に示す.

- 第1章 本章であり, 本研究の背景, 研究目的について述べる.
- 第2章 本研究の関連研究について述べる.
- 第3章 本研究の提案について述べる.
- 第4章 本研究における実験結果および考察について述べる.
- 第5章 結論について述べる.

第 2 章 関連技術

2.1 まえがき

本章では，実験対象である 3D Human Mesh Estimation from Virtual Marker (HMEVM) について述べる．また，関連技術である，Motion and Shape capture (Mosh) [4]についても述べる．

2.2 HMEVM

HMEVM は仮想マーカーを用いて人体骨格情報から人体 3D メッシュを作成する技術である．入力された画像・動画から人物の骨格情報を抽出する．

仮想マーカーは以下の二つの条件を満たすように作成される．第一に仮想マーカーの線形結合によって，メッシュを正確に再構成できること．第二に入力される画像・動画から容易に検出できるように，識別可能な視覚的パターンを持つこと．そのため，メッシュの点に近く，できるだけ体表面に位置するように学習される．

図 2.1 のマーカーは抽出された仮想マーカーを表す．青のマーカーは前方から目視できるマーカーで，グレーのマーカーは前方からは目視できないマーカーを表す．

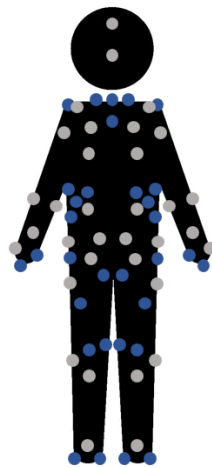


図 2.1 HMEVM による仮想マーカー

HMEVM は仮想マーカー生成・仮想マーカー信頼度スコア推定・メッシュ作成の三つの構造に分けられる．ここで HMEVM の概要を図 2.2 に示す．

入力画像 I が入力されると，まず仮想マーカーの三次元位置 P を推定する．次に仮想マーカー信頼度スコア推定である C に基づいて，仮想マーカーとメッシュ頂点との間の空間的関係を符号化するために用いる係数行列 A を更新する．仮想マーカーは人物

が向いている方向に限らず全身，360度存在するため，仮想マーカー位置にずれがあると，最終的なメッシュ作成に大きな誤差を与える可能性がある．そのため，仮想マーカー信頼度スコア推定 C に基づいて係数行列 A を更新する．

最後に，P と更新された A を掛け合わせて人体メッシュ作成 M を作成する．

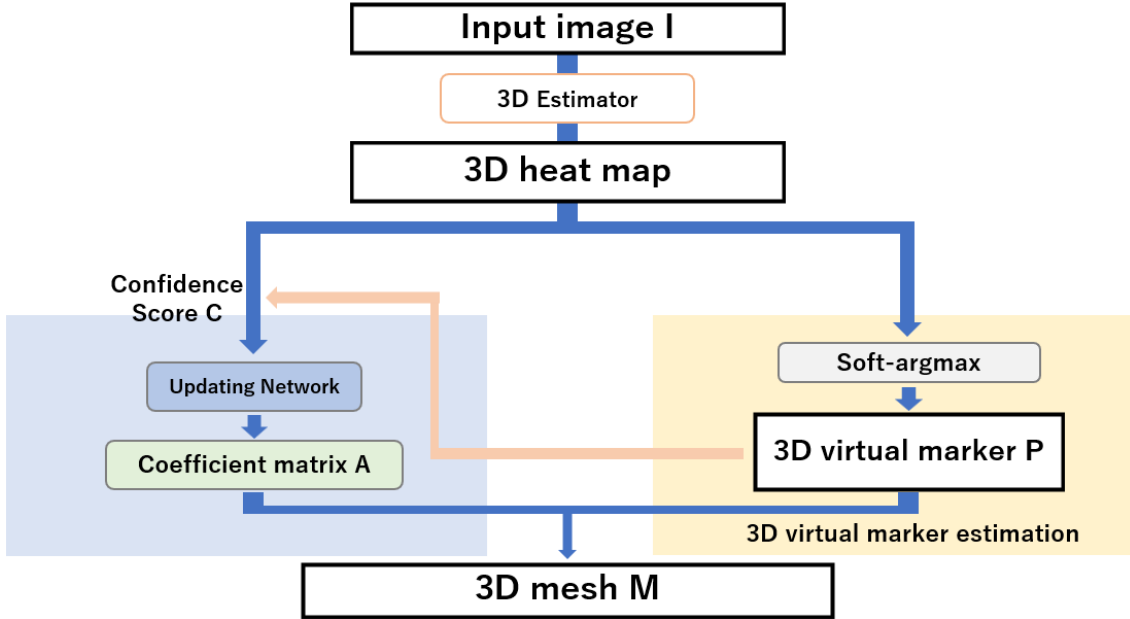


図 2.2 HMEVM の概要

HMEVM の学習で用いられる損失関数を式 (2.1) に示す．

$$L = \lambda_{vm}L_{vm} + \lambda_cL_{conf} + \lambda_mL_{mesh} \quad (2.1)$$

ここで L_{vm} は仮想マーカー生成の損失， L_{conf} は仮想マーカー信頼度スコア推定の損失， L_{mesh} はメッシュ作成の損失である．それぞれの損失関数を式 (2.2)，(2.3)，(2.4) に示す．

$$L_{vm} = \|P - P_{gt}\|_1 \quad (2.2)$$

$$L_{conf} = -\sum_{z=1}^K \log(H_z(P_{Zgt})) \quad (2.3)$$

$$L_{mesh} = L_{vertex} + L_{pose} + L_{normal} + \lambda_e L_{edge} \quad (2.4)$$

L_{mesh} の要素である四つの損失関数を式 (2.5)，(2.6)，(2.7)，(2.8) に示す．

$$L_{vertex} = \|M - M_{gt}\|_1 \quad (2.5)$$

$$L_{pose} = \|MJ - J_{gt}\|_1 \quad (2.6)$$

$$L_{normal} = \sum_f \sum_{\{i, j\} \subset f} \left| \left\langle \frac{M_i - M_j}{\|M_i - M_j\|_2}, n_{f_{gt}} \right\rangle \right| \quad (2.7)$$

$$L_{edge} = \sum_f \sum_{\{i, j\} \subset f} \left| \|M_i - M_j\|_2 - \|M_{i_{gt}} - M_{j_{gt}}\|_2 \right| \quad (2.8)$$

ここで、 M はメッシュ座標、 J は実際の3Dランドマーク関節、 J は $J \in \mathbb{R}^{M \times J}$ で、SMPLモデルであらかじめ定義されている行列である。 f 、 $n_{f_{gt}}$ はそれぞれ、メッシュの三角形の面、法線ベクトルを表す。

また、HMEVMの公式実装内で、全体の損失関数の各要素に掛かっている重み λ を式(2.9)、(2.10)、(2.11)に示す。

$$\lambda_{vm} = 100 \quad (2.9)$$

$$\lambda_c = 0.1 \quad (2.10)$$

$$\lambda_m = 0.1 \quad (2.11)$$

2.3 MoSh

MoShは、身体に装着された67個のマーカーと人体パラメトリックモデルを利用することで、3Dスキャンを使用せずに、正確な身体形状とそのポーズを推定する技術である。従来のマーカーベースのモーションキャプチャ技術では体表面の動きに関する重要な情報が、マーカーから骨格情報を抽出する際に失われていた。それを解決するために、モーションキャプチャデータから自動的に骨格情報を抽出するMoShが開発された。

MoShは体型を計測すると、その動きを直接キャプチャする。キャプチャしたデータが身体形状とポーズの精度に与える影響を評価し、動きをキャプチャするための新しいマーカーセットを作成する。その作成したデータから身体形状とポーズの動きを復元する。

2.4 むすび

本章では、実験対象であるHMEVMおよび、関連研究であるMoShについて述べた。

第3章 各要素の重み λ と HMEVM の評価指標との関係

3.1 まえがき

本章では、HMEVM の損失関数を解析する手法を提案する。本研究では評価指標として、Mean Per Vertex Error (MPVE) と Mean Per Joint Position Error (MPJPE) を用いる。

3.2 提案概要

本研究では2章2節で述べた HMEVM の損失関数が持つ三つの要素の各重み λ を変化させ、評価指標である MPVE, MPJPE の値の変化を比較する。また、各要素がそれらの値に及ぼす影響力を調査する。

手法としては、式 (2.9), (2.10), (2.11) で述べた各要素の重み λ を変化させ、MPVE, MPJPE との関係と比較し、評価する。また、各要素の損失の範囲を考慮するように損失関数を修正し、各要素が MPVE, MPJPE に及ぼす影響力を調査する。その後、各要素と MPVE, MPJPE との対応関係から、MPVE, MPJPE が最良の値となるような各要素の重み λ の値を検討する。

3.3 MPVE, MPJPE の概要

MPVE は生成されたメッシュの各頂点の位置と対応する実際のメッシュ頂点の距離を計算し、それらの距離の平均をとったものである。MPVE が小さいほど、生成されたメッシュ頂点位置と実際のメッシュ頂点の距離が近い。そのため、メッシュ生成が、より正確であると判断できる。

MPJPE はモデルが予測した各関節の位置と、対応する実際の関節の距離を計算し、それらの距離の平均をとったものである。MPJPE が小さいほど、予測された関節位置と実際の関節位置の距離が近い。つまり、骨格推定がより正確であると判断できる。

3.4 各要素の損失の範囲を考慮した損失関数

本研究では式 (2.1) で述べた全体の損失関数に L_{vm} (仮想マーカー生成の損失), L_{conf} (仮想マーカー信頼度スコア推定の損失), L_{mesh} (メッシュ作成の損失) が MPVE, MPJPE に及ぼす影響力を調査する。そのためには、各要素の重みの比率で各要素の影響力を推定する必要がある。しかし実際には各損失 L_{vm} , L_{conf} , L_{mesh} の取り得る範囲が異なる点を考慮しなければならない。これらの損失の理論上の範囲はそれぞれの損失関数の式 (2.2), (2.3), (2.4) より求められる。しかし、損失を求める上で、各項全ての予測値が正解値と極端な差を持つことは稀である。そのため、実際に損失が上限値をとる可能性は低い。そこで、実際に、各損失が取り得る最大値を実験により求める。それぞれの最大値をここでは L_{vm_MAX} , L_{conf_MAX} , L_{mesh_MAX} とする。各損失の範囲を式

(3.1), (3.2), (3.3) に示す.

$$0 \leq L_{vm} \leq L_{vm_MAX} \quad (3.1)$$

$$0 \leq L_{conf} \leq L_{conf_MAX} \quad (3.2)$$

$$0 \leq L_{mesh} \leq L_{mesh_MAX} \quad (3.3)$$

これらの損失の範囲を[0,1]に統一するために, 式 (3.1), (3.2), (3.3) についてそれぞれの最大値で除算する. このとき, それぞれの最大値は0でないと仮定する. 最大値で除算した結果を式 (3.4), (3.5), (3.6) に示す.

$$0 \leq \frac{L_{vm}}{L_{vm_MAX}} \leq 1 \quad (3.4)$$

$$0 \leq \frac{L_{conf}}{L_{conf_MAX}} \leq 1 \quad (3.5)$$

$$0 \leq \frac{L_{mesh}}{L_{mesh_MAX}} \leq 1 \quad (3.6)$$

この[0,1]の範囲にした各要素の損失関数を式 (2.1) に適用したものを式 (3.7) に示す.

$$L = \lambda_{vm} \frac{L_{vm}}{L_{vm_MAX}} + \lambda_c \frac{L_{conf}}{L_{conf_MAX}} + \lambda_m \frac{L_{mesh}}{L_{mesh_MAX}} \quad (3.7)$$

損失範囲を考慮した重みにするために, 式 (3.7) を変形した式を式 (3.8) に示す.

$$L = \lambda_{vm} L_{vm_MAX} \frac{L_{vm}}{L_{vm_MAX}} + \lambda_c L_{conf_MAX} \frac{L_{conf}}{L_{conf_MAX}} + \lambda_m L_{mesh_MAX} \frac{L_{mesh}}{L_{mesh_MAX}} \quad (3.8)$$

ここで $\lambda_{vm} L_{vm_MAX}$, $\lambda_c L_{conf_MAX}$, $\lambda_m L_{mesh_MAX}$ は各要素の損失の範囲の違いを考慮した各要素の重みとなる. そのため, 本研究では $\lambda_{vm} L_{vm_MAX}$, $\lambda_c L_{conf_MAX}$, $\lambda_m L_{mesh_MAX}$ の比率をもとに各要素が MPVE, MPJPE に与える影響力を調査する.

3.5 むすび

本章では, 最良な各要素の重み λ を求める手法を提案した. また, 本研究の評価指標である MPVE, MPJPE について述べた. そして, 損失の範囲を考慮した上で, 各要素が MPVE, MPJPE に及ぼす影響力を調査する手法を提案した.

第4章 実験結果・考察

4.1 まえがき

本章では，提案に基づく実験の概要，結果及び考察について述べる．実験では，各要素の重み λ の値と MPVE, MPJPE の対応関係を調査する．また，各要素が MPVE, MPJPE に及ぼす影響力を調査する．

4.2 データセット

本研究で用いたデータセットは，3D Poses in the Wild (3DPW) [5] である．データセットには 60 個の動画が含まれており，フレーム数の合計は 51,000 枚である．また，これらのフレームに対応した二次元骨格情報，三次元身体情報，三次元人物モデルが含まれている．

4.3 実験結果

各要素の重み λ を変化させた，MPVE, MPJPE の値を表 4.1, 4.2, 4.3, 4.4, 4.5 に示す．

表 4.1 各要素の重み λ の変化と評価 (1)

λ_{vm}	λ_c	λ_{mesh}	MPVE (↓)	MPJPE (↓)
0.1	0.1	0.1	380.5	347.3
1	1	1	450.5	415.9
10	10	10	348.4	309.1
100	100	100	289.3	271.1

表 4.2 各要素の重み λ の変化と評価 (2)

λ_{vm}	λ_c	λ_{mesh}	MPVE (↓)	MPJPE (↓)
100	1	1	129.2	113.8
1	100	1	448.6	413.7
1	1	100	309.2	281.6

表 4.3 各要素の重み λ の変化と評価 (3)

λ_{vm}	λ_c	λ_{mesh}	MPVE (↓)	MPJPE (↓)
1	1	1	450.5	415.9
10	1	1	263.8	236.8
50	1	1	139.7	119.7
100	1	1	129.2	113.8

表 4.4 各要素の重み λ の変化と評価 (4)

λ_{vm}	λ_c	λ_{mesh}	MPVE (↓)	MPJPE (↓)
100	0.01	0.01	123.7	108.7
100	0.1	0.1	118.0	103.9
100	1	1	129.2	113.8
100	10	10	244.7	275.1

表 4.5 各要素の重み λ の変化と評価 (5)

λ_{vm}	λ_c	λ_{mesh}	MPVE (↓)	MPJPE (↓)
100	0.1	0.1	118.0	103.9
200	0.1	0.1	119.8	103.8
300	0.1	0.1	120.0	103

実験結果より, MPVE, MPJPE の値が実験内で最良となる各要素の重み λ の値は公式実装と同じものとなることが分かる.

4.4 各要素の損失の範囲

提案に基づいて, 実験のなかで各要素の損失が取り得る値の範囲を式 (4.1), (4.2), (4.3) に示す.

$$0 \leq L_{vm} \leq 0.85 \quad (4.1)$$

$$0 \leq L_{conf} \leq 4.35 \quad (4.2)$$

$$0 \leq L_{mesh} \leq 364.73 \quad (4.3)$$

また, 式 (3.4), (3.5), (3.6) に基づいて式 (4.1), (4.2), (4.3) を変形すると式 (4.4), (4.5), (4.6) となる.

$$0 \leq \frac{L_{vm}}{0.85} \leq 1 \quad (4.4)$$

$$0 \leq \frac{L_{conf}}{4.35} \leq 1 \quad (4.5)$$

$$0 \leq \frac{L_{mesh}}{364.73} \leq 1 \quad (4.6)$$

4.5 各要素が評価に与える影響度

4章3節の結果より，MPVE，MPJPEの値が最良となる際の損失関数は式(4.7)となる。

$$L = 100L_{vm} + 0.1L_{conf} + 0.1L_{mesh} \quad (4.7)$$

この式(4.7)に各要素の損失の範囲を考慮した式(4.4)，(4.5)，(4.6)を適用した式を式(4.8)に示す。

$$L = 100 \times 0.85 \times \frac{L_{vm}}{0.85} + 0.1 \times 4.35 \times \frac{L_{conf}}{4.35} + 0.1 \times 364.73 \times \frac{L_{mesh}}{364.73} \quad (4.8)$$

ここで範囲を[0,1]に揃えた損失をそれぞれ $\frac{L_{vm}}{L_{vm_MAX}} = L'_{vm}$ ， $\frac{L_{conf}}{L_{conf_MAX}} = L'_{conf}$ ，

$\frac{L_{mesh}}{L_{mesh_MAX}} = L'_{mesh}$ とする。これを式(4.8)に適用したものを式(4.9)に示す。

$$L = 85L'_{vm} + 0.435L'_{conf} + 36.473L'_{mesh} \quad (4.9)$$

式(4.9)における各要素の重みが損失の範囲を考慮した重みとなる。よって，この三つの重みの比率がMPVE，MPJPEの精度に与える影響度となる。ここで影響度全体を100とした時の比率を表4.6，図4.1に示す。

表 4.6 仮想マーカー生成，仮想マーカー信頼度スコア推定，メッシュ作成
それぞれの影響度

仮想マーカー生成	仮想マーカー信頼度 スコア推定	メッシュ作成
69.7	0.4	29.9

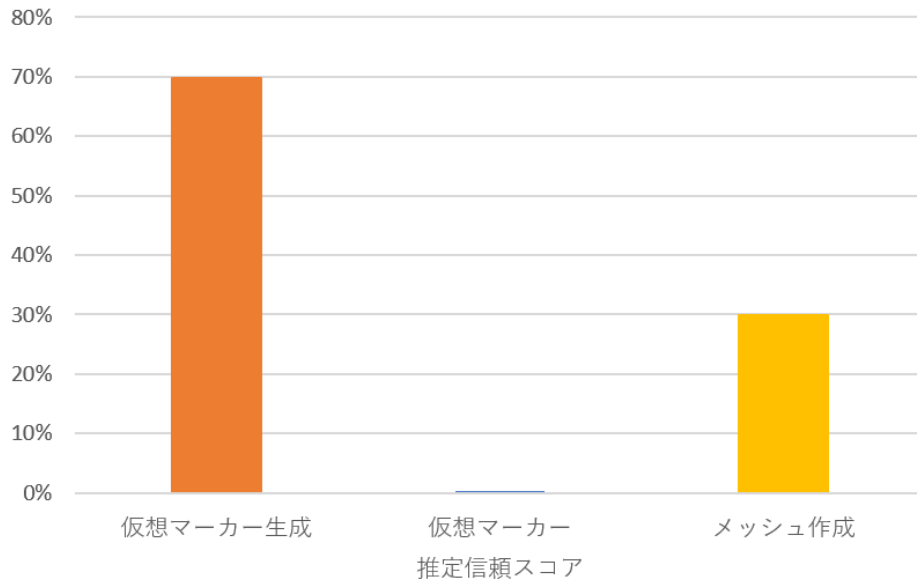


図 4.1 仮想マーカー生成，仮想マーカー信頼度スコア推定，メッシュ作成
それぞれの影響度

4.6 考察

表 4.1 の重み $\lambda = 1$ の場合の MPVE, MPJPE と表 4.2 の MPVE, MPJPE を比較すると， λ_{vm} の値を大きくした時が最も MPVE, MPJPE が向上しており， λ_c を大きくした時は， $\lambda = 1$ の場合と差がほとんどない。また，表 4.6 でも， λ_{vm} , λ_{mesh} , λ_c の順番で MPVE, MPJPE が向上しており，表 4.2 と同様の結果を示していると確認できる。したがって，表 4.6 の結果は妥当であると言える。

表 4.5 より， λ_{vm} の値を大きくすると，MPJPE は向上するが，MPVE は低下することが確認できる。 λ_{vm} は仮想マーカー生成の重みであり，かつ骨格情報推測の重みでもある。 λ_{vm} の値を大きくすると，全体損失における骨格情報推測の割合が増加するため，MPJPE が向上したと考えられる。また，MPVE はメッシュ頂点の精度の良さを表している。そのため， λ_{vm} の値を大きくすると，メッシュ作成の要素である λ_{mesh} の割合が減少する。したがって，MPVE が低下したと考えられる。

4.7 むすび

本章では提案に基づく評価実験を行い，その結果を示した。結果として，MPVE, MPJPE が最良となる重みの値は $\lambda_{vm} = 100$, $\lambda_c = 0.1$, $\lambda_{mesh} = 0.1$ であり，仮想マーカー生成の影響力は 69.9%，仮想マーカー信頼度スコア推定の影響力は 0.4%，メッシュ作成の影響力は 29.9% であると確認できた。

第5章 結論・今後の課題

5.1 結論

本研究では NMEVM の評価指標である MPVE, MPJPE が最良になる損失関数の要素の重みを求めた。また、各要素の損失の範囲を考慮し、各要素が最終的な評価に及ぼす影響力を求めた。その結果、MPVE, MPJPE の値が最良となる重みの値は $\lambda_{vm} = 100$, $\lambda_c = 0.1$, $\lambda_{mesh} = 0.1$ であり、仮想マーカー生成の影響力は 69.9%, 仮想マーカー信頼度スコア推定の影響力は 0.4%, メッシュ作成の影響力は 29.9%であった。

HMEVM 内で用いられていた各要素の重みは $\lambda_{vm} = 100$, $\lambda_c = 0.1$, $\lambda_{mesh} = 0.1$ であるため、HMEVM 内では最良な重みを用いていると言える。

5.2 今後の課題

本研究内で変化させた各要素の重み λ の値は、より細分化できる余地がある。そのため、本研究内では調査できなかった値で MPVE, MPJPE を求めると $\lambda_{vm} = 100$, $\lambda_c = 0.1$, $\lambda_{mesh} = 0.1$ の場合より良い値が出力される可能性がある。

謝辞

本論文の執筆に当たり，実験室の環境を整えてくださり，テーマや研究方針にご指導して下さった渡辺裕教授に感謝いたします。

また日頃から研究に関する相談や助言・ご指摘をくださった同研究室の皆様に心より感謝いたします。

最後に金銭面や生活面で支えてくださった家族に深く感謝いたします。

参考文献

- [1] G. Moon and K. M. Lee. “I2l-meshnet: Image-to-lixel Prediction Network for Accurate 3D Human Pose and Mesh Estimation from a Single Rgb Image,” European Conference on Computer Vision (ECCV), pp. 752–768, Aug. 2020.
- [2] H. Choi, G. Moo and K. M. Lee, “Pose2Mesh: Graph Convolutional Network for 3D Human Pose and Mesh Recovery from a 2D Human Pose,” European Conference on Computer Vision (ECCV), Aug. 2020.
- [3] X. Ma, J. Su, C. Wang, W. Zhu and Y. Wang, “3D Human Mesh Estimation from Virtual Markers,” IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 534-543, Jun. 2023.
- [4] M. Loper, N. Mahmood and M. J. Black, “MoSh: Motion and Shape Capture from Sparse Markers,” Transactions on Graphics (TOG), 33(6), Jan. 2014.
- [5] T. V. Marcard, R. Henschel, M. J. Black, B. Rosenhahn and G. Pons-Moll, “Recovering Accurate 3D Human Pose in The Wild Using IMUs and a Moving Camera,” European Conference on Computer Vision (ECCV), pp. 601-617, Sep. 2018.

図一覧

図 2. 1	HMEVM による仮想マーカ	4
図 2. 2	HMEVM の概要	5
図 4. 1	仮想マーカ生成, 仮想マーカ信頼度スコア推定, メッシュ作成 それぞれの影響度	11

表一覧

表 4. 1	各要素の重み λ の変化と評価 (1)	9
表 4. 2	各要素の重み λ の変化と評価 (2)	9
表 4. 3	各要素の重み λ の変化と評価 (3)	10
表 4. 4	各要素の重み λ の変化と評価 (4)	10
表 4. 5	各要素の重み λ の変化と評価 (5)	10
表 4. 6	仮想マーカー生成, 仮想マーカー信頼度スコア推定, メッシュ作成 それぞれの影響度.....	11