

# 全天球カメラ映像を用いた複数受講者の姿勢推定

## Pose Estimation of Multiple Students using Spherical Camera Images

川北 亨<sup>\*1</sup>, 西口 敏司<sup>\*1</sup>, 村上 正行<sup>\*2</sup>

Toru KAWAKITA<sup>\*1</sup>, Satoshi NISHIGUCHI<sup>\*1</sup>, Masayuki MURAKAMI<sup>\*2</sup>

<sup>\*1</sup> 大阪工業大学 情報科学部

<sup>\*1</sup> Faculty of Information Science and Technology, Osaka Institute of Technology

<sup>\*2</sup> 京都外国語大学 外国語学部

<sup>\*2</sup> Faculty of Foreign Studies, Kyoto University of Foreign Studies

Email: e1b14032@st.oit.ac.jp

**あらまし**：本稿では、大学などの講義における授業リフレクションへの応用を目指し、講義中の受講者の姿勢を推定する手法を提案する。一般的なカメラでは撮影可能な範囲が限定されるため、講義室空間全体をカバーするには複数台のカメラを設置して撮影する必要があり、またカメラ映像間での時間同期をとる必要も生じるため非常に手間のかかる作業となる。そこで本研究では、1台で全周を撮影することが可能な全天球カメラを使用して撮影した映像を対象として受講者の姿勢を推定する手法について検討する。

**キーワード**：授業リフレクション、骨格検出、受講者姿勢推定

### 1. はじめに

大学などの講義において、講師が自らの授業を撮影したビデオを視聴し、受講者や講師自身を観察・分析するなどして振り返ることで授業改善を目指す授業リフレクションという取り組みがある。この取り組みでは、講義室に複数台のカメラを設置し、撮影した映像を視聴して受講者の振る舞いを分析し、講義の改善を試みる<sup>1)</sup>。この分析において講師と受講者の講義中のかかわりを観察するには、受講者と講師の双方を撮影するために2台以上のカメラを用意する必要があり手間もかかることから、1台のカメラで受講者と講師の様子を同時に撮影可能な全天球カメラを使用して講義を撮影し、視聴する取り組みもある<sup>2)</sup>。しかしながら、1回分の講義映像全体をすべて見返すと分析にかかる時間や労力の観点から、継続的な実践は困難であるという問題もある。

そこで本研究では、これらの負担を軽減するために、撮影した全天球カメラ映像から、講義の状況に基づく映像要約等に利用可能な情報の一つとして、各受講者の姿勢を推定する手法について提案する。

### 2. 全天球カメラによる講義の撮影

本研究では、全天球カメラ映像を使用するため、通常のカメラで撮影される映像とは異なり、図1(a)に示すように、一般的には魚眼映像として撮影される。そのため、通常の映像のような見えとするためには、図1(b)に示すような正距円筒図法に基づくパノラマ展開をする必要がある。一般的なスクール形式の机配置の講義室において、講師に近い位置の座席上に、視点位置を他の受講者と等しい高さに配置して映像を撮影すると、その視点からの複数の受講者、および講師を同時に撮影することが可能であり、また、画像の中央付近の高さの歪みは比較的小さい。この映像に対して、例えばVR用のヘッドマウントディスプレイを介して視聴すれば、その座席に座つ

たような感覚で、他の受講者および講師を観察することが可能となる。

一方、ある視点から見たパノラマ映像では、図1に示すように、視点固定のため、カメラからみた受講者の相対的な位置の違いにより、同じ姿勢の受講者であっても見えが異なるため、単純な姿勢推定手法を適用することができない。



(a) パノラマ展開前 (b) パノラマ展開後

図1 全天球カメラ映像の例

### 3. 全天球カメラ映像を用いた姿勢推定方法

2章で述べた問題に対処するために、本稿では、識別したい姿勢を様々な角度から撮影したモデル姿勢画像を用意し、各時刻で観測された受講者の姿勢画像との類似度を評価することで、当該受講者の姿勢を推定する手法を提案する。

#### 3.1 モデル姿勢の獲得

いくつかの講義映像を観察した結果、本稿では、受講者の姿勢として以下の4つの姿勢を取り上げる。

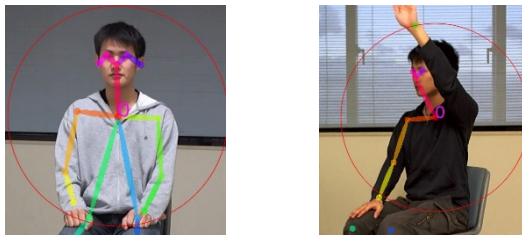
- 姿勢1：前を見ている（講師の方を見ている）
- 姿勢2：俯いている（ノートをとっている）
- 姿勢3：肘をついている（少し気がそれている）
- 姿勢4：手を挙げている（質問している）

予め、それぞれの姿勢をとりつつ、正面方向を含め左右に30度ずつ、最大90度まで向きを変え、姿勢毎に計7枚の画像を一般的なカメラで撮影する。さらに、撮影した画像に対して人物の関節に相当する画像上の位置が検出可能な骨格推定アルゴリズム

を適用し、画像に映った人物の両手首、首元、腰、鼻の位置に相当する座標を獲得する。図 2(a)に姿勢 1 を正面から撮影した人物の画像例、図 2(b)に姿勢 4 を左 30 度から撮影し、骨格検出アルゴリズムを適用した人物の画像例を示す。関節上の点が検出された位置を示し、各部位を結ぶ線がそれぞれの点の相対的な位置関係を表す。すべての画像上の姿勢に対し、検出された首元と鼻の位置の長さを基準とし、首元的位置を原点とする正規化座標を求めておく。

### 3.2 モデル姿勢を用いた姿勢推定

前節で獲得したモデル姿勢を、講義室を撮影した全天球カメラ画像上に映った受講者に適用することで、各受講者の姿勢を推定する。



(a) 姿勢 1 のモデル (b) 姿勢 4 のモデル

図 2 モデル姿勢の例

まず、全天球カメラで撮影した映像の各フレームに対し骨格検出アルゴリズムを適用し、検出された受講者毎の骨格情報を求める。得られた骨格情報(両手首、首元、腰、鼻)に対し、3.1 節で述べた方法と同様に、各受講者に対応する正規化座標を求める。

次に、獲得された各受講者の骨格情報と、3.1 節で予め用意しておいたモデル姿勢の骨格情報との類似度を求めることで、当該受講者の姿勢を推定する。

本稿ではスクール形式の机配置を想定し、基本的に各受講者は講義室の正面方向に体を向けて座っているという仮定と、全天球カメラは受講者用座席のうち前列に設置してあるという仮定を置く。このとき、全天球カメラで撮影した映像をパノラマ化した画像の中心を正面方向である 0 度とすると、パノラマ画像の中心から左右 1/4 の位置がカメラから見て ±90 度に対応する位置であることから、検出された各受講者の首元の画像上の位置に応じた角度に対応するモデル姿勢との類似度を求めればよい。これにより、当該受講者の姿勢を効率的に推定することが可能となる。

## 4. 実験

前章で述べた手法に基づき、どの程度正確に受講者の姿勢が推定できるか検証した。高低差のない 40 名程度の受講者を収容できる講義室に全天球カメラとして Ricoh Theta S を 1 台設置した。骨格検出アルゴリズムの適用には、OpenPose ライブラリ<sup>3)</sup>を利用した。このライブラリを講義映像に適用した結果の例を図 3 に示す。

約 5 分の講義映像 3 本に対して、10 秒毎に人物検

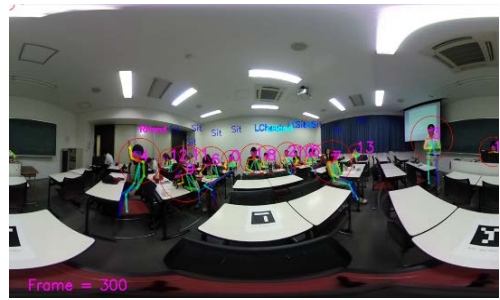


図 3 受講者検出結果の例

出及び骨格推定し、骨格推定できた受講者群に対して、手動で付与した正解姿勢と、推定した姿勢の結果を比較し、推定精度を求めた。その結果を表 1 に示す。各列は、正解姿勢に対する提案手法による姿勢判定の割合を(%)で示している。

表 1 姿勢推定結果(%)

		正解姿勢						
		正面	俯き	挙手		頬杖		
				右	左	右	左	
正解姿勢の割合		76.0	6.8	0.2	0.1	7.3	9.6	
推定姿勢の割合	正面	69.4	42.6	0.0	0.0	23.8	14.2	
	俯き	6.4	46.9	0.0	0.0	0.9	7.0	
	挙手	右	1.2	3.0	50.0	0.0	0.0	0.0
		左	4.4	0.0	50.0	0.0	1.8	7.6
	頬杖	右	5.5	3.3	0.0	0.0	58.5	11.5
		左	13.2	4.2	0.0	100.0	15.0	59.8

この結果から、正面や頬杖(右・左)については、比較的推定精度は高かったが、俯きについては正面と間違えることが多かった。モデル姿勢のより詳細な検討が必要であると考えられる。

## 5. おわりに

本研究では、大学などの講義における受講者の姿勢推定を目的とし、全天球カメラで撮影した映像に対してモデル姿勢との類似度を求める手法による受講者の姿勢推定を試みた。実験の結果、正面や頬杖を向いた姿勢については比較的高い精度で推定可能であった。今後の課題としては、その他の姿勢について、よりよいモデル姿勢を検討することなどで推定精度を向上させることなどが挙げられる。

### 参考文献

- (1) 米谷 淳, “授業観察事始め —授業というフィールドにおける本格的な行動研究を目指して—,” 大学授業のフィールドワーク —京都大学公開実験授業—, 京都大学高等教育教授システム開発センター編, 玉川大学出版部, p.74 - p.99, 2001.
- (2) 西口 敏司, 豊浦 正広, 村上 正行, “没入型 HMD を用いた臨場感の高い授業体験,” 教育システム情報学会 全国大会, 11-13, 2016.
- (3) Zhe Cao, Tomas Simon, Shih-En Wei, Yaser Sheikh, “Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields,” CVPR2017.