

授業分析支援のための受講者の三次元可視化手法の検討

A Method for 3D-Visualization of Students for Supporting Lecture Analysis

西口 敏司^{*1}, 村上 正行^{*2}
Satoshi NISHIGUCHI^{*1}, Masayuki MURAKAMI^{*2}

^{*1}大阪工業大学 情報科学部

^{*1}Faculty of Information Science and Technology, Osaka Institute of Technology

^{*2}京都外国語大学 マルチメディア教育研究センター

^{*2}Kyoto University of Foreign Studies, Research Center for Multimedia Education

Email: nishigu@is.oit.ac.jp

あらまし：授業の分析手法の一つとして、ビデオで受講者の様子を撮影しておき、後で分析者が映像から受講者の様子を確認して姿勢変化に関する情報などを付与する方法が一般的に行われている。しかしながら、ビデオで撮影した映像では、受講者の傾きや向きなどの姿勢を判断することが困難であるという問題がある。そこで本稿では、受講者の様子を、深度センサーを用いて三次元情報として獲得し、分析しやすいように提示する手法について検討する。

キーワード：授業分析，三次元可視化，深度センサー

1. はじめに

授業分析においては、授業を受ける受講者や講師の様子をビデオで撮影しておき、後でビデオ映像を確認しながら受講者の姿勢や講師の位置などに関する情報を付与して分析することで、授業改善に役立てるとい手法が採られている⁽¹⁾。一方、ビデオで撮影した映像に対して手動でこのような情報を付与することは、非常に手間がかかる作業であるという問題がある。一方、計算機を用いて自動的に人物の姿勢や位置を推定する研究も活発に行われているが、画像認識技術を用いた人物の姿勢推定手法は、現段階では認識率が低く、完全な自動化は困難な状況である。

そこで本稿では、授業分析のために受講者の姿勢に関する情報（以下、姿勢情報）を付与する作業において、従来のようなビデオ映像だけではなく、深度センサーを用いて授業シーンの三次元情報を獲得し、可視化して提示することで、どのような授業分析支援が可能となるかについて検討した。

2. 深度センサーを用いた授業シーンの獲得

現在、受講者の顔の上げ下げや姿勢を記録するために授業の様子を映像として記録する際には、ビデオカメラを用いることが一般的である。このような映像は、教室内の受講者がよく見え、かつ、受講者を俯瞰する位置に設置されることが多い。このようにして撮影された映像は、後で分析者が視聴し、顔の上げ下げなどの姿勢変化に関する情報を付与することとなる。

しかしながら、このような映像は固定された視点からのみの映像となるため、視点位置によっては、例えば受講者の前後方向の動きや、体幹の傾き具合などを判断することが難しいことがある。また複数台のビデオカメラで撮影したビデオ映像でも、時間同期の手間や複数の映像を見比べながら姿勢情報を

付与しなければならないという問題がある。

そこで本稿では、受講者の様子を、深度センサーを用いて獲得し、提示する枠組みについて検討する。深度センサーとしては、マイクロソフト社の Kinect センサーを用いる。このセンサーは、本来はゲーム機の入力デバイスとして人間のジェスチャを検出するために開発されたものであり、赤外線パターンを対象に照射しつつ赤外線カメラで観測し、照射したパターンとのずれを計測することで対象までの深度を検出する。このセンサーの水平視野角は約 58 度、垂直視野角は約 45 度で、深度情報を深度マップとして奥行約 4 メートルまで獲得でき、映像も同時に獲得できる。模擬授業における受講者のシーンの例を図 1 に示す。また、同時に獲得される深度マップの例を図 2 に示す。図 2 では、画像の明度で深度を表現し、明るいほど被写体が近いことを示している。



図 1 模擬授業のシーンの例



図 2 模擬授業の深度マップの例

3. 受講者の三次元可視化

獲得された深度マップから、Point Cloud Library⁽²⁾と呼ばれるライブラリソフトウェアを利用して、シーンの三次元点群を求めて可視化する。図1に示した画像情報及び図2に示した深度マップに基づいて、獲得できた全てのデータを三次元可視化した例を図3に示す。左下の座標軸は、深度センサーを原点とする座標系を表す。分析者はマウス操作によって任意視点からのシーンを見ることができ、1章で述べたような受講者の姿勢変化などの観測が容易になると考えられる。

一方、受講者が存在している部屋の壁などの提示は、授業シーンの分析には直接必要ではないと考えられることや、受講者の顔の上げ下げや体幹の前後移動を分析するには、椅子や机の天板などを基準として、それらからのずれが分かることが有用であると考えられる。

そこで本研究では、受講者に対応する点群のみを可視化するために、まず、予めセンサー座標系と机座標系との間の対応関係を求めておき、求められた三次元点群の全てを一旦机座標系に変換する。同時に、机座標系における机や椅子の位置を計測しておくことで、机が配置された領域かつ、机の天板より上の領域に存在する点群を受講者に対応する点群として獲得することができる。図4に、受講者部分のみを切り出し、各受講者の前の机の縁に垂直な平面を同時に提示した例を示す。机の縁を基準として、前のめりになっているか、あるいは、そうでないかの判断が容易になるのではないかと考えられる。

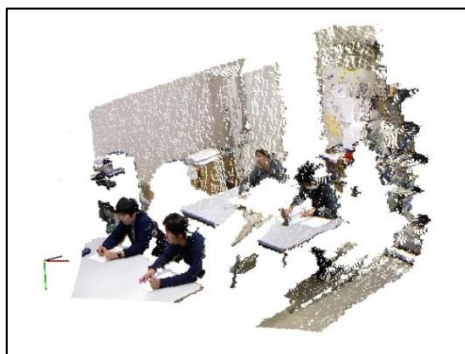


図3 三次元点群に基づく提示の例

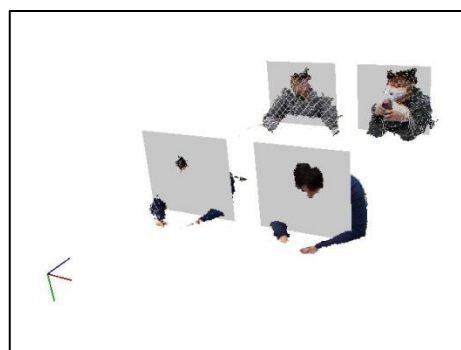


図4 受講者部分のみ切出して提示した例

また、定量的な分析の可能性について調べるために、図1の状況における各受講者が座っている椅子の位置から一定範囲内に存在する点群を、各受講者を表す点群であるとみなし、その重心位置の時間変化を示したグラフを図5に示すこの図の横軸は時間経過を表し、縦軸は各受講者の前後方向の位置を表す。このグラフから、例えば受講者Dについては姿勢変化の回数は少ないが変動量は大きいなど、各受講者の姿勢変化の様子を分析するのに役立つのではないかと考えられる。

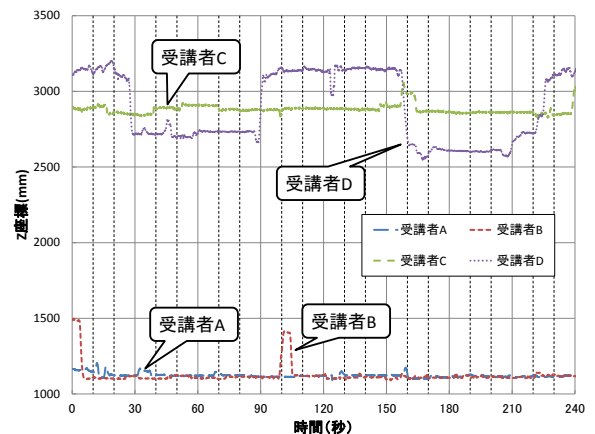


図5 各受講者の動きの例

4. おわりに

本稿では、受講者の姿勢情報の分析を容易にすることを旨として、深度センサーを用いて授業映像から受講者を三次元可視化する手法及び定量的な分析の可能性について検討した。一般的なビデオ映像とは異なり、深度センサーでは奥行き情報も獲得できることから、より分析しやすい可視化が期待できる。

一方、今回用いた深度センサーは、コストも安く容易に入手可能であるというメリットはあるが、視野角が狭いこと、空間解像度が低いこと、および、ビデオカメラと同様にオクルージョンの問題があることから、1台の深度センサーのみでは観測範囲が狭いのではないかと考えられる。そこで、複数台の深度センサーを用いた三次元シーンの合成や、深度センサーで得られた情報に基づく、高解像度なビデオ映像に対する情報の付与手法などが今後の検討課題として挙げられる。

謝辞

本研究の一部は、科学研究費補助金基盤研究(B)(23300311)の補助を受けた。ここに記して感謝する。

参考文献

- (1) 村上正行, 角所考, 美濃導彦: “マルチメディア一斉講義における内容に基づく受講生の注視行動の分析,” 人工知能学会誌 Vol.17, No.4, pp473-480, 2002.
- (2) Radu Bogdan Rusu and Steve Cousins, “3D is here: Point Cloud Library (PCL),” IEEE International Conference on Robotics and Automation (ICRA), 2011