

Аннотация

В данной работе рассматривается задача реконструкции трёхмерной позы человека в системе координат, связанной с человеком, по видеоданным, полученным в полуконтролируемой среде, то есть без предварительной калибровки камер и при минимальных требованиях к оборудованию. Основное внимание уделяется построению пайплайна, включающего детекцию людей, оценку двумерной позы, восстановление трехмерной позы на основе двумерных наблюдений и слияние данных с нескольких несинхронизированных камер с использованием SOTA нейросетевых моделей. Рассматривается метод адаптивного ансамблирования предсказаний с временно-зависимыми весами, учитывающими контекстную информацию. Для слияния предсказаний и синхронизации камер в мультикамерной конфигурации предлагается метод, основанный на прокрустовом преобразовании. В рамках работы создан программный инструмент на языке Python, реализующий предложенный способ восстановления позы, и проведена его экспериментальная оценка на реальных видеоданных.