

МЕТОДЫ СКЕЛЕТНОГО ТРЕКИНГА И АНАЛИЗА МИМИКИ В СИСТЕМАХ ВИДЕОМОНИТОРИНГА

Торопов А. Г.¹

Научный руководитель – канд. физ.-мат. наук, доцент Самарин А. В.¹

¹Университет ИТМО

agtoropov@itmo.ru

Работа выполнена в рамках темы НИР №640113 «AI Proctor-автоматизированная система прокторинга на основе методов искусственного интеллекта».

Введение

В последние годы прогресс в области компьютерного зрения и технологий искусственного интеллекта существенно расширил возможности анализа видеопотока. Особое внимание уделяется методам скелетизации и распознавания мимики, которые находят применение в системах автоматизированного прокторинга и поведенческого мониторинга пользователей [1, 2, 4]. Оценка позы и анализ выражений лица способствуют повышению объективности и надежности контроля соблюдения установленных регламентов при проведении экзаменов и иных дистанционных мероприятий. Настоящее исследование посвящено анализу существующих подходов к скелетизации и распознаванию мимики, их экспериментальной проверке и определению наиболее эффективных решений для практического применения.

Основная часть

Разработан широкий спектр методов анализа видеоданных, подтвердивших свою эффективность в задачах биометрии и поведенческого мониторинга [3, 5–7]. Вместе с тем ручной просмотр экзаменационных записей остается ресурсоёмким и затратным процессом, нередко требующим привлечения нескольких прокторов. В связи с этим актуальной является разработка автоматизированных средств оценки мимики и двигательной активности обучающихся, позволяющих выявлять аномальные поведенческие паттерны и потенциально подозрительные действия. Использование подобных решений способствует снижению нагрузки на прокторов и повышению объективности анализа видеоматериалов. В работе представлен обзор современных методов обработки видеопотока и проведена сравнительная оценка их эффективности. Рассмотрены подходы, основанные на сверточных нейронных сетях и трансформерных архитектурах. По результатам экспериментального тестирования наилучшие показатели в задачах автоматизированного прокторинга продемонстрировали MediaPipe Pose для скелетизации движений и EmoNet для анализа мимики. Указанные методы обеспечивают оптимальное соотношение точности и вычислительной эффективности, а также сохраняют устойчивость при вариациях условий съемки.

Выводы

Проведено исследование и экспериментальная апробация методов оценки позы и распознавания мимики, применимых к анализу видеопотока. По результатам сравнительного анализа в качестве наиболее эффективных решений выбраны MediaPipe Pose и EmoNet. Полученные выводы могут быть использованы при дальнейшем развитии систем автоматизированного прокторинга и повышения качества контроля соблюдения экзаменационного регламента.

Литература

1. D. J. Bora and A. K. Gupta, "Clustering approach towards image segmentation: An analytical study," *Int. J. Res. Comput. Appl. Rob.* 2 (7), 115–124 (2014). <https://doi.org/10.48550/arXiv.1407.8121>
2. Samarin, A., Savelev, A., Toropov, A. et al. Segmentation of the Iris and Pupil of the Human Eye in Images from an Infrared Camera. *Pattern Recognit. Image Anal.* 34, 855–862 (2024). <https://doi.org/10.1134/S1054661824700743>
3. L.-C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," *arXiv Preprint* (2017). <https://doi.org/10.48550/arXiv.1706.05587>
4. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *arXiv Preprint* (2015). <https://doi.org/10.48550/arXiv.1505.04597>
5. O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Ed. by N. Navab, J. Hornegger, W. Wells, and A. Frangi, *Lecture Notes in Computer Science*, Vol. 9351 (Springer, Cham, 2015), pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28
6. Gerczuk, Maurice & Amiriparian, Shahin & Ottl, Sandra & Schuller, Björn. (2021). EmoNet: A Transfer Learning Framework for Multi-Corpus Speech Emotion Recognition. *IEEE Transactions on Affective Computing*. PP. 1-1. 10.1109/TAFFC.2021.3135152.
7. Sengar, Sandeep & Kumar, Abhishek & Singh, Owen. (2024). Efficient Human Pose Estimation: Leveraging Advanced Techniques with MediaPipe. 10.48550/arXiv.2406.15649.