

<sup>1</sup>Е.Н. Амиргалиев, <sup>2</sup>И.Н. Букенова\*

<sup>1</sup>Международный университет информационных технологий, Алматы, Казахстан

<sup>2</sup>Алматинский технологический университет, Алматы, Казахстан

\*e-mail: ibukenowa@mail.ru

## АНАЛИЗ ПСИХОЭМОЦИОНАЛЬНОГО СОСТОЯНИЯ ЧЕЛОВЕКА ПО ВИДЕО

**Аннотация.** Распознавание действий наблюдаемого по видео, психологическое и эмоциональное состояния, в том числе студентов и школьников, - сложная и актуальная задача, имеющая социальную значимость. Темпы технологического развития и растущий интерес экспертов в стране и за рубежом показывают, что автоматизация определения психоэмоциональных реакций является практической и популярной областью исследований. В этой статье описаны различные методы по этой теме, кратко описаны и проанализированы современные системы, распознающие эмоциональное поведение человека. Проанализированы существующие методы распознавания человеческого поведения, эмоциональный искусственный интеллект, который позволяет компьютерам распознавать и интерпретировать человеческие эмоции и реагировать на них. Камера считывает состояние человека, а нейронная сеть обрабатывает данные для определения эмоций. Кроме того, включены недавние исследовательские статьи, опубликованные в течение 2019 и 2020 годов, чтобы представить последние достижения в этой области.

**Ключевые слова:** Компьютерное зрение, обработка изображения, кластерный анализ, обработка сигналов, нейросеть, фильтрация.

**Введение.** Важнейшей частью исследования в области взаимодействия человека и компьютера являются работы, направленные на распознавание психоэмоционального состояния человека. Поэтому, поскольку спрос на взаимодействие человека с компьютером продолжает расти, методы, которые автоматически распознают человеческие эмоции, будут привлекать все больше и больше исследователей и разработчиков. Решение этой задачи имеет большое научное значение для всех сфер фундаментальных исследований человека и информационных технологий. В последние годы явно усилился интерес к анализу видеонаблюдения с применением технологии искусственного интеллекта, рассматриваемого в качестве наиболее удобного объективного способа выявления эмоций, эмоционального состояния человека. Прежде всего, контролируйте эмоции ребенка каждую секунду обучения. Распознавание поведения человека и обобщение образа-сложные задачи компьютерного зрения.

В школе происходят трагедии, поэтому для выявления эмоций в школе нужно использовать технологию видеонаблюдения. Чтобы выявить молодых людей, которые могут представлять потенциальную опасность, мало их поймать, когда они используют оружие в школе в критической ситуации. Таким образом нельзя избежать возможного вреда. Поэтому чтобы повлиять на молодых людей и предотвратить эту критическую ситуацию, необходимо их определить заранее. Для этого устанавливают умное видеонаблюдение, где нейросеть следит. Сейчас это очень популярно: создаются специальные компьютерные программы, которые обладают способностью к обучению и восприятию определенных знаний. Смысл такой системы именно в том, чтобы сделать системный постоянный мониторинг большого количества людей, давая ценную информацию педагогам и психологам школы, сфокусировать их в направлении наибольшего риска.

Эмоции играют важную роль в жизни человека. Анализ психоэмоционального состояния человека позволяет отслеживать изменения в поведении людей и их отношении к событиям. Эмоции влияют на когнитивные процессы и принятие решений. Поэтому необходимость определения эмоциональных реакций становится к событиям. Эмоции влияют на когнитивные процессы и принятие решений. Поэтому необходимость определения эмоциональных реакций становится все более важной.

Чтобы обучить нейросеть, собирают выборку данных, вручную размечают изменение эмоционального состояния человека. Программа изучает шаблоны и понимает, какие признаки к какой эмоции относятся.

Имеется база данных изображений. Такими изображениями могут быть лица людей, образы, различное эмоциональное состояние, различные объекты трехмерного мира. Ставится задача поиска нужного изображения в базе данных. Причем задача может формулироваться как для поиска точного изображения, так и максимально приближенного к заданному. Важной задачей является подбор методов и алгоритмов обработки изображений, которые могут дать качественное решение задачи. Технологии обработки в этом случае зависят от многих параметров: объема базы данных, размеров на изображении, качества изображений, параметров яркости и контрастности изображений, наличия фона, ракурсы расположения объектов [1].

Проблема агрессивности, в том числе детской, давно привлекает внимание психологов. Одна из наиболее давних психологических теорий агрессивности – психоаналитическая – объясняет агрессию как один из видов природных инстинктов у человека, который в ходе социализации находит приемлемые в обществе каналы для выхода и способы выражения. В этом случае повышенная агрессивность ребенка объясняется недостаточной силой "Я", которое управляет поведением, а также недостаточным развитием психологических механизмов использования агрессивной энергии "в мирных целях". Усиливает агрессивность ребенка усвоенное им знание о себе как "неуправляемом", "злобном", "драчуне" и т. д. До сих пор психоаналитическая трактовка агрессивности остается одной из самых популярных.

Если мы видим в отчете видеонаблюдения, что у ребенка систематически наблюдается «агрессия» каждый день в течение недели, то мы знаем, что подросток находится в опасности. А завтра или послезавтра могут произойти какие-то небольшие инциденты, которые станут последней каплей, заставляющей разум взорваться и выйти из стабильного состояния. Внешний мир агрессии и печали, как правило, раскрывается после того, как событие произошло: попытка или самоубийство. Даже родители могут упустить эту ситуацию, не заметив, что с ребенком что-то происходит.

Цель данной работы заключается в обзоре методов и анализе психоэмоционального состояния наблюдаемых на основе бесконтактного видео наблюдения.

**Методы классификации эмоций.** Другим направлением, связанным с компьютерным зрением, является обработка сигнала. Многие методы обработки одномерных сигналов, обычно временных сигналов, могут быть естественным образом расширены для обработки двумерных или многомерных сигналов в компьютерном зрении. Однако из-за особого характера изображений существует множество методов, разработанных в области компьютерного зрения, не имеющих аналогов в области обработки одномерных сигналов. Отличительной особенностью этих методов является их нелинейность, которая вместе с многомерностью сигнала образует соответствующую подобласть в части обработки сигнала компьютерного зрения.

Система оценки человеческого риска предназначена для регистрации, анализа и изучения психоэмоционального состояния человека и оценки его потенциального уровня опасности. Человек обладает определенным диапазоном вибраций в своем теле, и форма этого спектра является одним из признаков психоэмоционального состояния или информации о возникновении эмоционального напряжения в другом канале информации.

Из статьи [2] "Обзор методов оценки психоэмоционального состояния человека" мы узнали, что одним из объективных методов разработанных учеными, в Санкт-Петербурге является система, предназначенная для регистрации, анализа и исследования психоэмоционального состояния человека, количественного определения уровней эмоций, детекции лжи, психофизиологической диагностики и дистанционного выявления потенциально опасных людей (Нгуен Д.К., & Южаков М. М., 2015). Система по мнен

потенциально опасных людей (Нгуен Д.К., & Южаков М. М., 2015). Система по мнению Нгуен Д.К., & Южаков М. М позволяет визуально и автоматически оценивать психофизиологическое состояние человека на основе вестибулярно-эмоционального рефлекса, с помощью программной визуализации. Ввод изображения объекта осуществляется с любого источника видео, например, цифровой телевизионной камеры, а программное обеспечение обрабатывает полученную информацию и предоставляет интерфейс для сохранения полученных результатов. Система Vibra Image производит автоматический мониторинг уровня эмоции, таких как стресс, агрессия, тревожность и другие, а также осуществляет детекцию лжи в режиме реального времени [3]. В данной статье приведен понятие психоэмоциональное состояние и его влияние на здоровье и деятельность человека. Представлен обзор методов и приборы, применяемые для оценки психоэмоционального состояния человека. Исследованы методы, используемые в лаборатории № 63 Института Неразрушающего Контроля Томского Политехнического Университета. В статье акцентируется внимание на применение более чувствительных электродов - наносенсоров для более точной оценки психоэмоционального состояния человека.

В [1] системы, определяющие эмоциональные реакции разбиты на группы по типу определения эмоциональных реакций: по физиологическим показателям; по мимике; по телодвижениям (пантомимика); по голосу.

В статье [3] авторы сосредоточились на методах суммирования видео, которые создают коллекцию коротких видеоклипов из длинной видеопоследовательности. В этой статье они представляют новый подход к обобщению видео, который работает непосредственно в сжатой области. Он основан на использовании визуальных функций, извлеченных из видеопотока, и на простом и быстром алгоритме суммирования видеоконтента. Предложенный метод был разработан для обработки видеопоследовательности и последовательного формирования выходного резюме.

Для выявления эмоциональных телодвижений необходима система распознавания действий человека.

Так в статье [4] авторы сосредоточены на создании эффективной системы распознавания действий человека путем извлечения дискриминационных патчей из обучающих видеороликов. Эти патчи являются наиболее репрезентативными частями, которые могут описать действие.

В статье [5] авторы представляют концептуальную модель движений человека, с помощью которой предлагается новый подход к признанию деятельности, связанной с движением человека. Путем глубокого анализа общего понимания движений предлагается концептуальная модель движения и авторы предлагают метод DFSA.

В статье [6] предлагается подход к обобщению видео, основанный на распознавании действий, когда действия человеческого тела сначала обнаруживаются и идентифицируются. Обнаруживая и распознавая многие действия человеческого тела предлагаемый авторами метод обеспечивает возможность также суммировать эти обнаруженные действия человека. Предлагаемый метод состоит в обнаружении, распознавании и обобщении нескольких действий, совершенных многими людьми в одном и том же видео. Предлагаемый метод позволяет обнаруживать и распознавать многие действия человеческого тела. Для этого предлагаются два метода: первый использует косинусную меру сходства HOG карты временной разницы (TDMar), а второй - классификацию действий CNN из изображений TDMar.

JIN CB и др [11] в своей статье представили дескриптор под-действия для детального обнаружения действия. Дескриптор вспомогательного действия состоит из трех уровней: позы, передвижения и уровня жеста. Три уровня дают три категории под-действий для одного действия, направленного на решение проблемы репрезентации. Предлагаемая модель обнаружения действий одновременно локализует и распознает действия нескольких людей

при видеонаблюдении с использованием временных функций на основе внешнего вида с несколькими CNN.

В другой работе AKULA A и др [12] демонстрируется использование ИК-камер в области AAL и обсуждается их эффективность в распознавании действий человека (HAR). Особое внимание обращается на одно из самых ответственных действий - падение. В работе был сгенерирован набор данных ИК-изображений, состоящий из 6 классов действий - ходьба, стояние, сидение на стуле, сидение на стуле со столом впереди, упал на стол впереди и упал / лежал на земле. Чтобы добиться надежного распознавания действий, авторы разработали контролируруемую архитектуру сверточной нейронной сети (CNN).

Система контроля психоэмоционального состояния человека (VibraImage) предназначена для выявления агрессивных и потенциально опасных людей, с помощью бесконтактного дистанционного сканирования с целью обеспечения безопасности в аэропортах и других охраняемых объектах [1]. Система позволяет автоматически вычислять и визуально оценивать психоэмоциональное состояние человека с помощью программной обработки телевизионного сигнала и его преобразования в виброизображение. Психоэмоциональное состояние человека характеризуется на основе запатентованных алгоритмов анализа вестибулярно-эмоционального рефлекса и макродвижений. Проект Natal от компании Microsoft. Программное обеспечение осуществляет полное 3-мерное распознавание движений тела, мимики лица и голоса. Natal способен распознавать эмоции по голосу и лицу.

В статье Kamiński Ł, Maćkowiak S, Domański M (2017) [7], представлен новый метод распознавания человеческой деятельности. Предлагаемое решение использует одну стационарную камеру для обнаружения обычных человеческих действий, таких как: размахивание руками, ходьба, бег и т. Д. В отличие от других методов, которые используют различные типы дескрипторов характеристических точек для описания человеческих поз, предлагаемое решение использует дескриптор CDVS который является частью стандарта MPEG-7. Это позволяет эффективно вычислять компактный дескриптор в камере.

**Результаты и анализ.** Если рассмотреть обнаружение и распознавание множественных действий человека с помощью этих вышеописанных методов, то можно сравнить полученные результаты степени точности для этих методов. В данном случае я взяла для сравнения методы описанные в [4,8,9,12].

Степень точности для современных методов, связанных с распознаванием нескольких действий человека, которые используют одну и ту же категорию представленных наборов данных показано в таблице 1.

Таблица 1. **Общая точность распознавания современных методов**

Методы	Точность %
Jin CB и др. 2017 [ 11 ]	83. 5% ICVL
Akula A и др. 2018 [ 12 ]	87,44% (инфракрасное видео)
CS-HOG-TDКарта [10]	92,5%
DT + BoVW [ [8] ]	85,4%
плотные траектории [ [9] ]	84,2 %
3D-патчи [4]	85,6%

Метод DFSA [5] работает лучше, чем традиционные методы, такие как C4.5, SVM, BP, и достиг общей истинной скорости классификации  $96,4\% \pm 0,025$ . Скорость распознавания метода [6] до 90%

**Заключение.** В описываемой работе я провела краткий анализ существующих методов распознавания действия человека, распознавания образов. Искусственный интеллект и ҚазҰТЗУ хабаршысы №4 2021

некоторые смежные области, такие как анализ сцены и машинное зрение, все еще находятся на начальных стадиях развития. Однако, описанные подходы действительно очень разнообразны и с помощью большинства методов можно решить практически любую задачу распознавания образов. Жизнь быстро информатизируется, и внедрение системы, которая отслеживает психоэмоциональное состояние по видеоряду, передаваемому по сетям 3G или через Интернет, позволит внедрять системы, которые могут контролировать жизнь общества, выявляя агрессию, гнев, депрессивные состояния.

Таким образом, существующие методики распознавания образов способны решать широкий круг задач, и в зависимости от ограничивающих факторов (бюджета разработки, скорости и точности распознавания образа) можно найти метод подходящий для используемого языка программирования.

#### ЛИТЕРАТУРА

- [1] Орлова, Ю.А. (2011). Обзор современных автоматизированных систем распознавания эмоциональных реакций человека. Волгоград, Россия: ВолгГТУ.
- [2] Нгуен Д.К., & Южаков М. М. (2015). Обзор методов оценки психоэмоционального состояния человека. Томск: VI Научно-практическая конференция.
- [3] Almeida J, Leite NJ, Torres R. d. S. (2013) Online video summarization on compressed domain. *J Vis Commun Imag Represent* 24(6):729–738
- [4] El-Masry M, Fakhr MW, Salem M. A. -M. Action recognition by discriminative EdgeBoxes. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, pp. 443–452
- [5] Xu C, He J, Zhang X (2017) DFSA: A classification capability quantification method for human action recognition. In: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). IEEE
- [6] Qazi HA, et al. (2017) Human action recognition using SIFT and HOG method. In: 2017 International Conference on Information and Communication Technologies (ICICT). IEEE
- [7] Kamiński Ł, Maćkowiak S, Domański M (2017) Human activity recognition using standard descriptors of MPEG CDVS. In: 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE
- [8] Wang, H., Klaser, A., Schmid, K., et al : "Dense trajectories and boundary motion descriptors for action recognition", *Int. J. Comput. Vis.* , 2013 , 103 , ( 1 ), pp. 60-79 Crossref Web of Science®
- [9] Wang, H., Kläser, A., Schmid, C., et al: ' Action recognition by dense trajectories'. 2011 IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2011, pp. 3169– 3176
- [10] El-Henawy I, Ahmed K, Mahmoud H (2018) Action recognition using fast HOG3D of integral videos and Smith–Waterman partial matching. *IET Imag Process* 12(6):896–908
- [11] JIN C-B, Shengzhe LI, Hakil et KIM (2017) Real-Time Action Detection in Video Surveillance using Sub-Action Descriptor with Multi-CNN. arXiv:1710.03383
- [12] AKULA A, SHAH AxK, et GHOSH R (2018) Deep learning approach for human action recognition in infrared images. *Cogn Syst Res* 50:146–154.

#### REFERENCES

- [1] Orlova, Yu.A. (2011). *Obzor sovremennykh avtomatizirovannykh sistem raspoznavaniya e`mocional`nykh reakcij cheloveka*. Volgograd, Rossiya: VolgGTU.
- [2] Nguen D.K., & Yuzhakov M. M. (2015). *Obzor metodov ocenki psikhoe`mocional`nogo sostoyaniya cheloveka*. Tomsk: VI Nauchno-prakticheskaya konferenciya.
- [3] Almeida J, Leite NJ, Torres R. d. S. (2013) Online video summarization on compressed domain. *J Vis Commun Imag Represent* 24(6):729–738
- [4] El-Masry M, Fakhr MW, Salem M. A. -M. Action recognition by discriminative EdgeBoxes. In: 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, pp. 443–452
- [5] Xu C, He J, Zhang X (2017) DFSA: A classification capability quantification method for human action recognition. In: 2017 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computed, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, SmartWorld/SCALCOM/UIC/ATC/CBDCOM/IOP/SCI). IEEE

- [6] Qazi HA, et al. (2017) Human action recognition using SIFT and HOG method. In: 2017 International Conference on Information and Communication Technologies (ICICT). IEEE
- [7] Kamiński Ł, Maćkowiak S, Domański M (2017) Human activity recognition using standard descriptors of MPEG CDVS. In: 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW). IEEE
- [8] Wang, H., Klaser, A., Schmid, K., et al : "Dense trajectories and boundary motion descriptors for action recognition", Int. J. Comput. Vis. , 2013 , 103 , ( 1 ), pp. 60-79 Crossref Web of Science®
- [9] Wang, H., Kläser, A., Schmid, C., et al: ' Action recognition by dense trajectories'. 2011 IEEE Conf. Computer Vision and Pattern Recognition (CVPR), 2011, pp. 3169– 3176
- [10] El-Henawy I, Ahmed K, Mahmoud H (2018) Action recognition using fast HOG3D of integral videos and Smith–Waterman partial matching. IET Imag Process 12(6):896–908
- [11] JIN C-B, Shengzhe LI, Hakil et KIM (2017) Real-Time Action Detection in Video Surveillance using Sub-Action Descriptor with Multi-CNN. arXiv:1710.03383
- [12] AKULA A, SHAH AxK, et GHOSH R (2018) Deep learning approach for human action recognition in infrared images. Cogn Syst Res 50:146–154

**<sup>1</sup>Е.Н. Амиргалиев, <sup>2</sup>И.Н. Букенова\***

<sup>1</sup>Халықаралық ақпараттық технологиялар университеті, Алматы, Қазақстан

<sup>2</sup>Алматы технологиялық университеті, Алматы, Қазақстан

\*e-mail: ibukenowa@mail.ru

## БЕЙНЕ АРҚЫЛЫ АДАМНЫҢ ПСИХОЭМОЦИОНАЛДЫ ЖАҒДАЙЫН ТАЛДАУ

**Андатпа.** Қазіргі уақытта видеодан байқалатын психологиялық және эмоционалдық жай - күйді, оның ішінде студенттер мен оқушылардың іс-әрекетін тану-әлеуметтік маңызы бар күрделі және өзекті міндет. Технологиялық даму қарқыны, шетелдік және отандық мамандардың өсіп келе жатқан қызығушылығы психоэмоционалды реакцияларды анықтауды автоматтандыру практикалық және танымал зерттеу саласы болып табылады. Бұл мақалада осы тақырыптағы әртүрлі әдістер сипатталған, адамның эмоционалды мінез-құлқын танитын заманауи жүйелер қысқаша сипатталған және талданған. Компьютерлерге адамның эмоциясын тануға және түсіндіруге және оларға жауап беруге мүмкіндік беретін адамның мінез-құлқын, эмоционалды жасанды интеллектін танудың қолданыстағы әдістері талданады. Камера адамның жағдайын оқиды, ал нейрондық желі эмоцияны анықтау үшін деректерді өңдейді. Сонымен қатар, осы саладағы соңғы жетістіктерді ұсыну үшін 2019 және 2020 жылдары жарияланған соңғы зерттеу мақалалары бар.

**Негізгі сөздер:** Компьютерлік көру, суретті өңдеу, кластерлік талдау, сигналдарды өңдеу, нейрондық желі, сүзгілеу.

**<sup>1</sup>Y.N. Amirgaliyev, <sup>2</sup>I.N. Bukenova\***

<sup>1</sup>International Information Technology University, Almaty, Kazakhstan

<sup>2</sup>Almaty Technological University, Kazakhstan, Almaty, Kazakhstan

\*e-mail: ibukenowa@mail.ru

## ANALYSIS OF THE PSYCHOEMOTIONAL STATE OF A PERSON BY VIDEO

**Abstract.** Recognizing a person's actions based on video, as well as their psychological and emotional state, including students, is a complex and urgent task that has social significance. The pace of technological development and the growing interest of experts at home and abroad show that the automation of determining psychoemotional reactions is a practical and popular area of research. This article describes various methods on this topic, briefly describes and analyzes modern systems that recognize human emotional behavior. The existing methods of recognizing human behavior, emotional artificial intelligence, which allows computers to recognize and interpret human emotions and respond to them, are analyzed. The camera reads the person's state, and the neural network processes the data to determine emotions. In addition, recent research papers published during 2019 and 2020 are included to present the latest advances in this field.

**Keywords:** Computer vision, image processing, cluster analysis, signal processing, neural network, filtering.