Разработка теоретико-практической модели управления процессами мониторинга методами машинного зрения в условиях воздействия факторов среды

Туманов Александр Юрьевич
к.т.н., доцент, кафедра метрологического обеспечения инновационных технологий и промышленной безопасности, Санкт-Петербургский государственный университет аэрокосмического приборостроения, toumanov@mail.ru,

Андрюшин Георгий Владиславович
кафедра метрологического обеспечения инновационных технологий и промышленной безопасности, Санкт-Петербургский государственный университет аэрокосмического приборостроения, re.work.me@yandex.ru

Актуальность. Современный этап автоматизации систем управления сложными антропотехническими процессами (включая образовательные, диспетчерские и операторские контуры) требует внедрения объективных методов непрерывного контроля когнитивного состояния человека. Существующие психометрические подходы носят эпизодический характер. Объектом исследования являются процессы контроля визуального внимания человека в замкнутом пространстве на фоне динамического изменения параметров окружающей среды. Предметом исследования выступают алгоритмы компьютерного зрения для детекции поведенческих маркеров внимания и методы их временной синхронизации с пространственно-распределенными потоками телеметрии микроклимата.

Цель работы — разработка, математическое обоснование и экспериментальная верификация теоретико-практической модели мультимодального мониторинга и управления процессами удержания внимания методами машинного зрения, инвариантной к дестабилизирующим факторам внутренней среды.

Методы. В работе использован комплекс методов компьютерного зрения (библиотека OpenCV, алгоритмы YuNet/Haar, оценка позы головы через решение задачи PnP), методы теории измерительных систем (калибровка индивидуального базиса), а также технологии интернета вещей (IoT) на базе протокола MQTT для асинхронного сбора данных о температуре, влажности и освещенности помещения.

Результаты. Разработана трехслойная архитектура аппаратно-программного комплекса. Сформулирован и обоснован математический подход к расчету интегрального показателя внимания (Attention Score) на основе конфиденциально-взвешенной агрегации признаков (положение головы, динамика смыкания век EAR, стабильность трекинга). Разработан асимметричный автомат состояний для детекции фаз отвлечения. Экспериментально доказана работоспособность программного контура на C++ на реальных видеопотоках и подтверждена валидность классификации базовых поведенческих паттернов по сравнению с экспертной разметкой.

Ключевые слова: компьютерное зрение, машинное зрение, мониторинг внимания, Eye Aspect Ratio, baseline-калибровка, микроклимат, телеметрия, корреляционный анализ, MQTT, OpenCV, IoT, Индустрия 5.0.

Введение

Глобальная цифровизация технологических и образовательных пространств формирует жесткий запрос на создание интеллектуальных систем автоматического контроля и управления процессами деятельности человека. В контурах систем «человек — машина», «диспетчер — пульт управления» или «преподаватель — аудитория» ключевым дестабилизирующим фактором является психофизиологическое утомление оператора (обучающегося), влекущее за собой критическое снижение концентрации внимания. Санитарно-гигиенические и медицинские исследования последних лет однозначно доказывают, что динамика когнитивной работоспособности человека неразрывно связана с параметрами микроклимата замкнутого помещения [1, 2]. В частности, повышенные концентрации углекислого газа (CO_2), температурные аномалии, отклонения относительной влажности и нерациональная освещенность рабочих мест вызывают преждевременное утомление и падение продуктивности когнитивной деятельности [1, 3].

Несмотря на очевидную значимость этих результатов, в современной практике оценка когнитивного состояния человека выполняется преимущественно эпизодически [1, 2]. Используемые методы — анкетирование, психологическое тестирование по таблицам Шульте, экспертное наблюдение или анализ итоговых результатов деятельности — фиксируют лишь финальный срез утомления. Они не обладают достаточной чувствительностью к кратковременным (интраминутным) флуктуациям поведения непосредственно в процессе выполнения целевой задачи.

Наиболее перспективным научно-техническим вектором решения данной проблемы является переход к непрерывному, малозатратному и объективному бесконтактному мониторингу визуальных признаков внимания на основе методов машинного зрения. Современные алгоритмы компьютерного зрения позволяют в реальном времени извлекать из видеопотока субклинические маркеры поведения: пространственную ориентацию головы оператора, стабильность и направление его взгляда, а также частотно-временную динамику смыкания век [4].

Параллельное развитие доступной компонентной базы беспроводных измерительных сетей (IoT) позволяет развернуть пространственно-распределенные системы контроля параметров окружающей среды с минимальными накладными расходами. Однако на текущий момент в научной литературе наблюдается разрыв: системы компьютерного зрения развиваются изолированно от систем климатической телеметрии. Отсутствуют формализованные математические модели, обеспечивающие бесшовную временную синхронизацию гетерогенных потоков данных (высокочастотных видеометрик и низкочастотных измерений датчиков) для выявления корреляционных и причинно-следственных связей.

Целью настоящего исследования является разработка и математическое обоснование теоретико-практической модели управления процессами удержания внимания на основе интеграции методов машинного зрения и распределенной телеметрии факторов среды. Данная модель призвана стать основой для построения систем автоматизированного управления микроклиматом (BMS/Smart Building) в парадигме человекоцентричного подхода Индустрии 5.0.

Для достижения поставленной цели необходимо решить следующие задачи:

Сформулировать методологический подход к многокомпонентной количественной оценке визуального внимания оператора (обучающегося).

Разработать алгоритм адаптивной персональной калибровки для компенсации индивидуальных антропометрических особенностей посадки человека.

Спроектировать математический метод оценки индекса отклонения факторов среды от гигиенических норм.

Разработать архитектуру аппаратно-программного комплекса и экспериментально верифицировать разработанные алгоритмы на реальных массивах покадровых данных.

2. МЕТОДЫ И МАТЕРИАЛЫ

2.1. Математическая модель конфиденциально-взвешенной оценки внимания

В основе предлагаемой теоретико-практической модели лежит концепция многокомпонентного анализа. Итоговый показатель внимания субъекта не может быть достоверно определен по какому-то одному извлеченному признаку (например, только по направлению взгляда), поскольку видеопоток в реальных условиях эксплуатации подвержен шумам, окклюзиям (перекрытиям кадра) и изменениям освещенности.

В качестве первичного вектора информативных признаков $F = \{f_1, f_2, ..., f_n\}$ в модели используются следующие метрики, извлекаемые из видеопотока:

Изменение пространственной ориентации головы по трем осям: углы рыскания (*yaw*), тангажа (*pitch*) и крена (*roll*);

Состояние глаз, вычисляемое через безразмерный коэффициент смыкания век *Eye Aspect Ratio (EAR)* [4];

Наличие и стабильность детекции лица в кадре (метрика устойчивости трекинга);

Направление вектора взгляда (*gaze direction*) при достаточном пространственном разрешении.

Для интеграции этих разнородных компонентов разработан *confidence-weighted rule-based* подход (подход на основе правил с весовой оценкой достоверности). Итоговый показатель внимания $S(t)$ в момент времени t рассчитывается по формуле:

$$S(t) = \frac{\sum_{i=1}^{n} (w_i \cdot c_i(t) \cdot f_i(t))}{\sum_{i=1}^{n} (w_i \cdot c_i(t))}$$

где:

$f_i(t) \in [0,1]$ — нормированное значение i-го поведенческого признака внимания (1 — максимальная концентрация на целевом объекте, 0 — полное отвлечение);

$w_i \in [0,1]$ — статический базовый вес i-го компонента, определяющий его теоретическую информативность в общей структуре модели ($\sum w_i = 1$);

$c_i(t) \in [0,1]$ — динамический коэффициент доверия (уверенности) алгоритма к извлеченному i-му признаку в текущем кадре.

Введение динамического коэффициента $c_i(t)$ позволяет радикально снизить влияние зашумленных компонентов на итоговую метрику. Например, если лицо оператора частично перекрыто рукой или развернуто под острым углом к камере, точность модуля *gaze estimation* резко падает. В этом случае система автоматически снижает коэффициент доверия $c_{\text{gaze}} \to 0$, и итоговый показатель $S(t)$ формируется за счет более стабильных признаков (кинематики головы), не разрушая непрерывность мониторинга.

В рамках разработанного программного прототипа эмпирически, на основе экспертных оценок и анализа обучающих выборок, установлены следующие начальные весовые коэффициенты (Таблица 1):

Таблица 1

Базовые весовые коэффициенты компонентов внимания

Метрика (Компонент модели)	Обозначение (f_i)	Базовый вес (w_i)	Назначение в контуре
Ориентация головы (*Head Pose*)	f_1	0,35	Фиксация грубых отклонений от целевой зоны
Направление взгляда (*Gaze Direction*)	f_2	0,35	Тонкая оценка фокуса (модуль зарезервирован)
Состояние глаз (*Eye State / EAR*)	f_3	0,15	Оценка утомления, сна и частоты моргания
Присутствие и устойчивость трекинга	f_4	0,15	Индикатор физического присутствия субъекта

2.2. Алгоритм адаптивной персональной калибровки (Baseline)

Одной из главных проблем систем машинного зрения в реальных условиях является вариативность антропометрических данных пользователей и их индивидуальной посадки на рабочем месте. Для исключения ложноположительных срабатываний в модель внедрен алгоритм *первичной базовой калибровки*.

При инициализации нового трека (появлении человека в кадре) система в течение фиксированного временного окна $T_{\text{calib}} = 10...15$ секунд производит накопление первичных векторов ориентации головы при условии, что коэффициент устойчивости трекинга максимален:

$$\overline{\text{yaw}}_0 = \frac{1}{M}\sum_{m=1}^{M} \text{yaw}_m, \quad \overline{\text{pitch}}_0 = \frac{1}{M}\sum_{m=1}^{M} \text{pitch}_m, \quad \overline{\text{roll}}_0 = \frac{1}{M}\sum_{m=1}^{M} \text{roll}_m$$

Сформированный вектор $B_0 = \{\overline{\text{yaw}}_0, \overline{\text{pitch}}_0, \overline{\text{roll}}_0\}$ принимается за индивидуальный *нейтральный базис (baseline)* конкретного человека. В процессе дальнейшего мониторинга математическая модель оперирует не абсолютными углами наклона головы, а относительными отклонениями от вычисленного базиса: $\Delta\text{yaw}(t) = \text{yaw}(t) - \overline{\text{yaw}}_0$.

Такой подход позволяет корректно оценивать состояние людей, которые в силу физиологических особенностей или конфигурации мебели сидят изначально слегка наклонившись или развернувшись относительно оптической оси камеры.

2.3. Математическое описание факторов среды

Для формализации воздействия внешней среды вводится многомерный вектор параметров микроклимата $E = \{T, H, L\}$, где T — температура воздуха ($^\circ$C), H — относительная влажность (%), L — освещенность (люкс). Для каждого параметра на основе действующих санитарно-гигиенических нормативов (СанПиН) задается интервал оптимальных значений $[P_{\min}, P_{\max}]$.

Разработан метод оценки безразмерной величины отклонения параметров среды $D(t)$. Для каждого измеряемого фактора частный индекс отклонения $d_k(t)$ рассчитывается по принципу штрафной функции:

$$d_k(t) = \begin{cases} 0, & \text{если } P_k(t) \in [P_{k,\min}, P_{k,\max}] \\ \dfrac{P_k(t) - P_{k,\max}}{P_{k,\max}}, & \text{если } P_k(t) > P_{k,\max} \\ \dfrac{P_{k,\min} - P_k(t)}{P_{k,\min}}, & \text{если } P_k(t) < P_{k,\min} \end{cases}$$

Интегральный индекс отклонения среды $D(t)$ представляет собой среднедекартово значение от частных индексов: $D(t) = \sqrt{\sum d_k^2(t)}$. Данная безразмерная величина выступает в качестве входного возмущающего воздействия в общей теоретической модели управления процессами.

3. РЕЗУЛЬТАТЫ (RESULTS)

3.1. Аппаратно-программная реализация видеоподсистемы

На основе предложенных методов был спроектирован и программно реализован на языке **C++** с использованием кроссплатформенной библиотеки **OpenCV** [5, 6] прототип аналитического контура. Выбор C++ обусловлен необходимостью минимизации задержек (*latency*) при покадровой обработке видеопотоков высокого разрешения в режиме реального времени.

Для детекции лиц применяется легковесная нейросетевая модель **YuNet**, инкапсулированная в модуль cv::FaceDetectorYN. Трекинг объектов осуществляется методом максимизации пересечения ограничивающих рамок (*Intersection over Union — IoU*) в модуле TrackManager. Расчет пространственной ориентации головы производится через сопоставление трехмерной антропометрической модели лица с двухмерными проекциями ключевых точек (*facial landmarks*), полученных через подсистему FacemarkLBF. Задача обратной графической проекции решается итерационным методом cv::solvePnP.

Логика переходов в классе AttentionScorer реализована в виде **асимметричного автомата состояний**. Переход из состояния «Фокусировка» (*focused*) в состояние «Отвлечение» (*looking_side / looking_down*) происходит мгновенно при превышении пороговых углов ($|\Delta\text{yaw}| > 20^\circ$, $|\Delta\text{pitch}| > 15^\circ$, $|\Delta\text{roll}| > 15^\circ$, или падении коэффициента открытия глаз $\text{EAR} < 0,20$). Однако обратный переход в состояние *focused* требует временного подтверждения на протяжении $N_{\text{thresh}} = 30$ последовательных кадров, что исключает «дребезг» метрик.

3.2. Архитектура IoT-модуля телеметрии среды

Аппаратная часть экологического контура построена на базе 32-битного микроконтроллера **ESP32**. К контроллеру по цифровой шине I2C подключены датчики: цифровой термогигрометр и люксметр. Сбор данных происходит дискретно с периодом $T_s = 1$ секунда. Микроконтроллер производит первичную фильтрацию шумов, упаковывает вектор среды в формат JSON-объекта и осуществляет беспроводную передачу данных по протоколу Wi-Fi на локальный MQTT-брокер [7].

Для последующего статистического и корреляционного анализа в подсистеме журналирования реализована двухслойная структура логирования данных: покадровый детальный отчет пишется в формате .csv, а макроизменения (смена внутренних состояний автомата) фиксируются асинхронно в формате JSON Lines (.jsonl).

4. ОБСУЖДЕНИЕ (DISCUSSION)

4.1. Метод временной синхронизации потоков данных

Ключевой научной задачей при совмещении методов машинного зрения и телеметрии является преодоление частотного дисбаланса данных. Видеоподсистема генерирует плотный поток данных с частотой $25-30$ кадров в секунду (Гц), в то время как датчики среды опрашиваются с частотой $0.5-1$ Гц.

В разработанной теоретической модели синхронизация потоков данных реализуется на уровне аналитического сервера посредством метода временных окон и последующего ресемплирования временных рядов к единому шагу дискетизации $\Delta t = 1$ с. В пределах каждого окна для каждого уникального рабочего места (seat_id) вычисляется среднее агрегированное значение показателя внимания:

$$\langle S \rangle_{\Delta t} = \frac{1}{K}\sum_{k=1}^{K} S_k$$

Полученные нормированные временные ряды $\langle S \rangle(t)$ и интегральный индекс отклонения среды $D(t)$ сопоставляются во времени. Для выявления долгосрочного влияния неблагоприятных экологических факторов применяется корреляционный анализ со сдвигом фаз (анализ лагов), описываемый коэффициентами Пирсона и Спирмена. Это обусловлено тем, что психофизиологическая реакция человека на ухудшение микроклимата (например, нарастание духоты) имеет инерционный характер: падение концентрации внимания происходит не мгновенно, а с запаздыванием на $\tau = 5...15$ минут.

4.2. Концепция Smart Campus в рамках Индустрии 5.0

Разработанный трехуровневый комплекс спроектирован как открытая расширяемая платформа. Сгенерированная схема трехуровневой архитектуры интеграции комплекса в общую систему автоматизированного управления зданием (*BMS — Building Management System*) представлена на Рисунке 2.

Перенос концепции анализа когнитивной нагрузки из области эпизодических тестов в плоскость непрерывного машинного зрения замыкает контур обратной связи:

Мониторинг выявляет устойчивую групповую тенденцию к падению *Attention Score* в конкретной зоне помещения.

Корреляционный анализатор сопоставляет это падение с ростом индекса отклонения среды $D(t)$ (например, ростом температуры выше 24°C).

Сервер генерирует директивную команду для исполнительных механизмов BMS: увеличить кратность воздухообмена вентиляционной системы или снизить уставку кондиционера.

Такой подход полностью соответствует человекоцентричной парадигме Индустрии 5.0 (*Human-Centric Systems*), где не человек адаптируется под жесткие технологические условия, а автоматизированная среда гибко подстраивается под психофизиологическое состояние человека для обеспечения максимального комфорта и производительности.

5. ЗАКЛЮЧЕНИЕ (CONCLUSION)

В ходе проведенного исследования разработана, математически описана и программно верифицирована теоретико-практическая модель управления процессами мониторинга когнитивной активности субъектов методами машинного зрения в условиях воздействия дестабилизирующих факторов среды.

Основные научные и практические результаты работы:

Разработан и реализован на языке C++ программный конвейер, обеспечивающий в реальном времени сквозную обработку видеопотока: от захвата кадра до расчета интегральной метрики внимания *Attention Score*.

Предложен и математически обоснован метод *confidence-weighted* агрегации признаков, минимизирующий влияние естественных пространственных шумов и окклюзий в кадре без потери общей устойчивости системы.

Интегрирован модуль персональной адаптивной базовой калибровки (*baseline*), обеспечивающий инвариантность измерений к индивидуальным особенностям посадки и антропометрии пользователей.

Спроектирована архитектура распределенного IoT-комплекса на базе протокола MQTT для синхронного сбора параметров микроклимата.

Экспериментальная апробация сформированного программного прототипа подтвердила высокую стабильность трекинга и корректность работы асимметричного автомата состояний. Сравнение результатов автоматической классификации с экспертной выборочной разметкой видеоматериала подтвердило высокую валидность предложенной модели.

Направления дальнейших исследований связаны с переходом от базового корреляционного анализа к многофакторному регрессионному моделированию с целью точного прогнозирования времени наступления фазы утомления оператора. Также планируется расширение аппаратной части IoT-контура специализированным датчиком концентрации углекислого газа (CO_2), выступающего важнейшим триггером когнитивной усталости в замкнутых пространствах.

Литература

1. Лобкис М. А., Новикова И. И., Романенко С. П., Сорокина А. В. Динамика умственной работоспособности школьников в условиях неблагоприятных показателей микроклимата и углекислого газа в помещениях // Российский вестник гигиены. — 2025. — № 4. — С. 50–58. DOI: 10.24075/rbh.2025.149.
2. Dedesko S., Pendleton J., Young A. S., Coull B. A., Spengler J. D., Allen J. G. Associations between indoor air exposures and cognitive test scores among university students in classrooms with increased ventilation rates for COVID-19 risk management // Journal of Exposure Science & Environmental Epidemiology. — 2025. — Vol. 35, no. 4. — P. 661–671. DOI: 10.1038/s41370-025-00770-6.
3. Андреев И. В. Влияние уровня концентрации кислорода и отрицательных его ионов в учебной аудитории на умственную работоспособность и академическую успеваемость студенток // Современные проблемы науки и образования. — 2009. — № 3. — С. 46–49. URL: science-education.ru (дата обращения: 24.07.2026).
4. Soukupová T., Čech J. Real-Time Eye Blink Detection Using Facial Landmarks // 21st Computer Vision Winter Workshop. — 2016.
5. Bradski G., Kaehler A. Learning OpenCV: Computer Vision with the OpenCV Library. — Sebastopol: O'Reilly, 2008.
6. OpenCV Documentation. URL: opencv.org (дата обращения: 24.07.2026).
7. MQTT Version 3.1.1. OASIS Standard. URL: oasis-open.org (дата обращения: 24.07.2026).

Development of a theoretical and practical model for process control using machine vision methods under the influence of environmental factors

Tumanov A.Yu., Andryushin G.V.
Saint Petersburg State University of Aerospace Instrumentation

JEL classification: C10, C50, C60, C61, C80, C87, C90

The current stage of automation of systems for managing complex anthropotechnical processes (including educational, dispatch, and operator circuits) requires the implementation of objective methods for continuous monitoring of a person's cognitive state. Existing psychometric approaches are episodic in nature. The object of the study is the processes of monitoring a person's visual attention in a confined space against the backdrop of dynamically changing environmental parameters. The subject of the study is computer vision algorithms for detecting behavioral markers of attention and methods for their temporal synchronization with spatially distributed microclimate telemetry flows.

The aim of the work is to develop, mathematically substantiate, and experimentally verify a theoretical and practical model for multimodal monitoring and control of attention retention processes using machine vision methods, which is invariant to destabilizing factors of the internal environment.

Methods. The work employs a set of computer vision methods (OpenCV library, YuNet/Haar algorithms, head pose estimation by solving the PnP problem), methods from the theory of measurement systems (calibration of an individual basis), as well as Internet of Things (IoT) technologies based on the MQTT protocol for asynchronous data collection on room temperature, humidity, and illumination.

Results. A three-layer architecture of the hardware-software complex has been developed. A mathematical approach to calculating the integral attention indicator (Attention Score) based on a confidentiality-weighted aggregation of features (head position, EAR eyelid closure dynamics, tracking stability) has been formulated and substantiated. An asymmetric state machine has been developed for detecting distraction phases. The functionality of the C++ software loop has been experimentally proven using real video streams, and the validity of the classification of basic behavioral patterns has been confirmed in comparison with expert labeling.

Keywords: computer vision, machine vision, attention monitoring, Eye Aspect Ratio, baseline calibration, microclimate, telemetry, correlation analysis, MQTT, OpenCV, IoT, Industry 5.0.

References

1. Lobkis M. A., Novikova I. I., Romanenko S. P., Sorokina A. V. Dinamika umstvennoi rabotosposobnosti shkolnikov v usloviyakh neblagopriyatnykh pokazatelei mikroklimata i uglekislogo gaza v pomeshcheniyakh [Dynamics of mental performance in schoolchildren under poor indoor climate conditions and elevated carbon dioxide levels]. Rossiiskii vestnik gigieny [Russian Bulletin of Hygiene]. 2025;(4):50–58. (In Russ.). DOI: 10.24075/rbh.2025.149.
2. Dedesko S., Pendleton J., Young A. S., Coull B. A., Spengler J. D., Allen J. G. Associations between indoor air exposures and cognitive test scores among university students in classrooms with increased ventilation rates for COVID-19 risk management. Journal of Exposure Science & Environmental Epidemiology. 2025;35(4):661–671. DOI: 10.1038/s41370-025-00770-6.
3. Andreev I. V. Vliyanie urovnya koncentracii kisloroda i otricatel'nyh ego ionov v uchebnoj auditorii na umstvennuyu rabotosposobnost' i akademicheskuyu uspevaemost' studentok [Influence of the oxygen concentration and its negatively charged ions in the classroom on the mental work productivity and academic progress of female students]. Sovremennye problemy nauki i obrazovaniya [Modern Problems of Science and Education]. 2009;(3):46–49. (In Russ.). Available at: science-education.ru (accessed: 24.07.2026).
4. Soukupová T., Čech J. Real-Time Eye Blink Detection Using Facial Landmarks. In: Proc. of the 21st Computer Vision Winter Workshop. 2016:12–19.
5. Bradski G., Kaehler A. Learning OpenCV: Computer Vision with the OpenCV Library. Sebastopol: O'Reilly Media; 2008.
6. Dokumentaciya po OpenCV [OpenCV Documentation]. 2026. Available at: opencv.org (accessed: 24.07.2026). (In Russ.).
7. Specifikaciya protokola MQTT Versiya 3.1.1. Standart OASIS [MQTT Version 3.1.1. OASIS Standard]. 2026. Available at: oasis-open.org (accessed: 24.07.2026). (In Russ.).