



Научная статья | Original paper

Восприятие эмоционального состояния человека при трансформации видеоизображений его тела

О.А. Королькова¹ 

¹ Московский государственный психолого-педагогический университет,

Москва, Российская Федерация

 olga.kurakova@gmail.com

Резюме

Контекст и актуальность. Восприятие эмоциональных состояний, проявляющихся в экспрессиях тела человека, связано с использованием информации как о форме, так и о кинематике движения. Несмотря на значительное число исследований, остается не вполне понятным, какого минимального количества информации о движении и форме достаточно для распознавания эмоций, а также как соотносятся движение и форма при восприятии различных категорий эмоций. **Цель:** определить, в какой степени исключение информации о форме и движении затруднит распознавание эмоциональных экспрессий тела. **Гипотеза.** По сравнению с полноцветными видеофрагментами точность распознавания эмоций по динамическим точечным паттернам снизится; дальнейшее уменьшение числа кадров приведет к еще большему снижению точности. **Методы и материалы.** В исследовании приняли участие 233 человека (209 женщин и 24 мужчины в возрасте от 22 до 64 лет, медиана 37 лет). Стимульный материал представлял собой черные точки на белом фоне, которые в динамике воспринимались как движения тела человека, выражающего эмоции радости, страха, отвращения, возбуждения, разочарования, скуки и спокойствия. Он демонстрировался разным группам участников либо с плавным (23 кадра/с), либо со стробоскопическим (2 кадра/с) движением. Участники определяли эмоциональное состояние натурщиков. Полученные данные сопоставлялись с результатами предыдущего исследования с использованием полноцветных видеофрагментов. **Результаты.** Стробоскопическая экспозиция снижает точность распознавания всех экспрессий, кроме радости и спокойствия, а также меняет структуру категориальных полей восприятия эмоций. Основной вклад в восприятие радости вносят особенности формы и текстуры изображения человека и контекст ситуации. Для экспрессий отвращения и разочарования роль динамической информации повышается. Оценка экспрессии страха зависит как от статических, так и от динамических признаков и преимущественно опирается на экспрессию тела, но не лица. Восприятие скуки и возбуждения в первую очередь связано с кинематическими паттернами. Для распознавания спокойного состояния ключевую роль играет не статика или динамика, а информация о микроэкспрессии лица. **Выводы.** Выявлены дифференцированные паттерны вклада информации о форме и движении в восприятие эмоциональных экспрессий тела в зависимости от категории эмоции.

Ключевые слова: восприятие эмоций, биологическое движение, экспрессия тела, эмоциональные состояния, невербальное поведение

Финансирование. Исследование выполнено за счет гранта Российского научного фонда (проект № 24-18-00904 «Механизмы восприятия эмоционального состояния человека в процессах невербальной коммуникации», <https://rscf.ru/project/24-18-00904/>).

Благодарности. Автор благодарит Е.Г. Хозе за помощь в сборе данных для исследования.



Для цитирования: Королькова, О.А. (2025). Восприятие эмоционального состояния человека при трансформации видеоизображений его тела. *Экспериментальная психология*, 18(4), 64–79. <https://doi.org/10.17759/exppsy.2025180404>

Perception of emotional state of a person based on the transformed video recordings of the body

O.A. Korolkova¹ 

¹ Moscow State University of Psychology and Education, Moscow, Russian Federation
 olga.kurakova@gmail.com

Abstract

Context and relevance. The perception of emotional states from human body expressions involves the use of both shape and motion kinematics information. Despite a significant number of studies, it remains unclear what minimum amount of motion and shape information is sufficient for emotion recognition, as well as how motion and shape are related in the perception of different emotion categories. **Objective:** to determine the extent to which excluding shape and motion information will hinder the recognition of emotional body expressions. **Hypothesis.** Compared to full-color video clips, the accuracy of emotion recognition from dynamic dot patterns will decrease; further reduction in the number of frames will lead to an even greater decrease in accuracy. **Methods and materials.** The study involved 233 participants (209 women and 24 men, aged 22 to 64 years, median 37 years). The stimuli consisted of dynamic black dots on a white background perceived as human body movements expressing emotions of happiness, fear, disgust, excitement, disappointment, boredom, and neutral state. The stimuli were shown to different groups of participants with either smooth (23 frames/s) or stroboscopic (2 frames/s) motion. Participants determined the emotional state of the posers. The obtained data were compared with the results of a previous study using full-color video fragments. **Results.** Stroboscopic exposure reduces the recognition accuracy of all expressions except happiness and neutral state and also changes the structure of categorical fields of emotion perception. The main contribution to the perception of happiness is made by the shape and texture of the human image, as well as the context of the situation. For expressions of disgust and disappointment, the role of dynamic information increases. The assessment of fear expression depends on both static and dynamic features and is primarily based on body, not facial expressions. The perception of boredom and excitement is primarily associated with kinematic patterns. For recognizing a neutral state, it's not static or dynamic features that play a key role, but rather information about facial microexpressions. **Conclusions.** We identified differentiated patterns of the contribution of shape and movement information to the perception of emotional body expressions that depend on the emotion category.

Keywords: emotion perception, biological motion, body expression, emotional states, non-verbal behavior

Funding. The research was supported by the Russian Science Foundation (project No. 24-18-00904 “Mechanisms of person’s emotional state perception in the processes of nonverbal communication” <https://rscf.ru/project/24-18-00904/>).

Acknowledgements. The author is grateful to E.G. Hoze for assistance in data collection.

For citation: Korolkova, O.A. (2025). Perception of emotional state of a person based on the transformed video recordings of the body. *Experimental Psychology (Russia)*, 18(4), 64–79. (In Russ.). <https://doi.org/10.17759/exppsy.2025180404>



Введение

При восприятии эмоциональных состояний других людей мы опираемся не только на выражения их лица, но и на движения тела — жесты, походку, изменения позы и др. По сравнению с лицом экспрессия тела различима на большем расстоянии, что позволяет наблюдателю заранее определять намерения и предвосхищать возможные действия другого человека, корректируя в зависимости от этого свое поведение (например, избегая прямого контакта лицом к лицу). В частности, такие признаки, как опущенная голова или выдвинутые вперед плечи при ходьбе, оцениваются как состояние печали либо гнева, тогда как поднятая вверх голова, расправленные плечи и максимально прямая осанка сигнализируют о переживании радости (Venture et al., 2014). Несмотря на интенсивное развитие данной области в последние годы, усилившееся в том числе благодаря распространению автоматизированных способов анализа естественного экспрессивного поведения (Bazarevsky et al., 2020; Ikeda et al., 2025; Ildei et al., 2024), механизмы распознавания аффективных проявлений по движениям тела на сегодняшний день изучены недостаточно полно (De Gelder, Poyo Solanas, 2022).

Выделяют три типа информации, связанной с восприятием движений тела (Atkinson et al., 2007): статическую (форма, текстура, объем), кинематическую (скорость, ускорение, перемещение) и собственно динамическую (описание движений через действующие на тело силы и смещение центра масс). Согласно ряду исследований, статическая информация о позе связана с валентностью (знаком) выраженной эмоции больше, чем особенности кинематики. В свою очередь, в кинематике содержится информация, ключевая для определения уровня возбуждения и интенсивности эмоции (Atkinson et al., 2004; Dael et al., 2013; Fourati, Pelachaud, 2018).

Для разделения влияний формы и кинематики традиционно используются динамические точечные паттерны, отражающие взаимное расположение основных суставов тела и создающие впечатление движения человека или иного живого существа (Johansson, 1973; Runeson, Frykholm, 1983). Аткинсон с коллегами разработали одну из первых баз, которая содержала как видеозаписи актеров, выражающих базовые эмоции через движения тела, так и основанные на этих видеозаписях динамические паттерны, исключающие информацию о форме. Показано, что таких паттернов достаточно для распознавания базовых эмоциональных экспрессий тела (Atkinson et al., 2004). О наличии специализированных механизмов анализа экспрессий по биологическому движению, не сводящихся к восприятию формы, говорит тот факт, что поворот точечного паттерна на 180° в плоскости изображения или его демонстрация с обратным порядком кадров снижают точность распознавания выраженных эмоций сильнее, чем те же манипуляции, проделанные с исходным видеоизображением, в котором присутствует как движение, так и форма, а степень снижения точности при нарушении типичных паттернов движения варьирует в зависимости от категории эмоции (Atkinson et al., 2007).

Существуют две основные модели восприятия биологического движения: первая утверждает, что анализ информации о форме и движении идет параллельно по вентральному и дорзальному путям головного мозга, а интеграция происходит только на поздних этапах обработки (Giese, Poggio, 2003); вторая предполагает возможность распознавания действий на основе совокупности последовательных статичных «срезов», без привлечения информации о локальном движении. В более общем виде первая модель характеризуется



двуихфакторным пространством распознавания действий: первое измерение связано с оценкой движущих сил, второе — с концептуальным представлением о действии (Gardenfors, Warglien, 2012). Одна из методологических трудностей разделения вклада формы и движения заключается в том, что динамические паттерны могут, в том числе, содержать информацию о форме, а статичные изображения — имплицитно предполагать наличие движения (Presti et al., 2022).

В пользу первой модели говорит двойная диссоциация областей мозга, обеспечивающих восприятие формы и движения тела (Vangeneugden et al., 2014). В частности, информация, необходимая для различения формы, коррелирует с паттернами распределения активации в экстрапиарной области тела (ЕВА), тогда как информация о движении — с паттернами активации в задней части верхней височной борозды (pSTS), которая рядом авторов рассматривается как область, обеспечивающая интеграцию мультимодальной информации о биологическом движении, эмоциональной экспрессии, социальных действиях и др. Непосредственное воздействие на эти области при помощи транскраниальной магнитной стимуляции дифференцированно снижает точность распознавания признаков формы либо движения.

В пользу второго подхода свидетельствуют данные об ухудшении восприятия характеристик биологического движения (в частности, определения направления ходьбы — вперед/назад) при увеличении времени экспозиции отдельных кадров с 40 до 200 мс (Beintema et al., 2006). Авторы связывают данные результаты с тем, что в условиях снижения частоты кадров происходят более сильные изменения позы между последовательными кадрами, а также уменьшается общее число экспонируемых статичных поз. При этом введение дополнительного интервала между кадрами не приводит к дальнейшему понижению точности, что говорит о менее значимой роли локального движения в восприятии биологических стимулов. Таким образом, предполагается, что основная роль локального движения состоит в выделении движущейся фигуры из фона, а распознавание направления ходьбы связано преимущественно с временной суммацией информации о форме.

Вместе с тем, несмотря на значительное число проведенных исследований, остается не вполне понятным, какого минимального количества информации о движении и форме достаточно для распознавания эмоций по экспрессиям тела, а также как соотносятся движение и форма в восприятии различных категорий эмоций. Показано, что, по сравнению с полноцветными видеофрагментами, при демонстрации точечных паттернов снижается точность распознавания таких базовых эмоциональных экспрессий, как гнев, страх и отвращение, тогда как для радости и печали результаты для двух типов стимулов не различаются (Atkinson et al., 2004).

В настоящем исследовании нас интересовало, в какой степени исключение информации о форме и текстуре экспрессий тела и контексте ситуации, а также уменьшение количества доступной наблюдателю информации о движении затруднят распознавание эмоциональных состояний. Для этого был проведен эксперимент с использованием динамических точечных паттернов, демонстрируемых с плавной сменой кадров либо со снижением частоты кадров (аналог стробоскопа). Если кинематическая информация важна для распознавания эмоций, мы ожидали снижения точности оценок при более низкой частоте кадров. Полученные данные сопоставлялись с результатами нашего предыдущего эксперимента, в котором наблюдатели распознавали эмоции по полноцветным ви-



деофрагментам, сохраняющим информацию как о движении, так и о форме и контексте (Королькова, 2024). Если информация о форме важна для восприятия эмоциональных экспрессий, будет наблюдаться снижение точности оценок при экспозиции точечных паттернов. В отличие от более ранних исследований, в которых демонстрировались преимущественно экспрессии базовых эмоций, мы использовали как базовые, так и сложные ситуативные состояния.

Материалы и методы

Участники исследования: 233 человека (209 женщин и 24 мужчины в возрасте от 22 до 64 лет, медиана 37 лет) — студенты первого и второго высшего образования психологических вузов. Они случайным образом распределялись в 2 группы, выполнившие разные серии исследования:

- серия 1 (условие «плавная экспозиция») — 96 человек (90 женщин и 6 мужчин в возрасте от 22 до 64 лет, медиана возраста — 38 лет);
- серия 2 (условие «стробоскопическая экспозиция») — 137 человек (119 женщин и 18 мужчин в возрасте от 22 до 63 лет, медиана возраста — 37 лет).

Стимульный материал создавался на основе видеоклипов, входящих в базу EU-ESM (O'Reilly et al., 2016). Из базы отбирались видеоклипы, демонстрирующие экспрессивное поведение трех натуралистов (двух женщин и одного мужчины), снятых в полный рост. В своих движениях они выражали семь эмоциональных состояний: радость, страх, отвращение, возбуждение, разочарование, скуку и спокойствие (всего 21 видеокlip). Ранее мы использовали эти видеоклипы в нашей предыдущей работе (Королькова, 2024). В настоящем исследовании видеоклипы подвергались трансформации для того, чтобы полностью исключить информацию о контуре и текстуре изображения тела натуралиста, контексте ситуации и окружающих предметах, а также об экспрессии лица, оставив только информацию о движении тела. При помощи библиотеки Mediapipe Pose Landmarker (Bazarevsky et al., 2020) в каждом видеоклипе выделялись 33 ключевые точки, соответствующие основным сочленениям тела натуралиста. Из них 11 точек были расположены на голове (нос, глаза, рот, уши), что обеспечивало информацию о положении, размере и повороте головы и лица, однако не давало возможности оценить его экспрессию. Оставшиеся 22 точки были распределены следующим образом: по 1 точке на левые и правые плечевой, локтевой и запястный суставы, пальцы рук (большой, указательный, мизинец), тазобедренный, коленный и голеностопный суставы, пятку и указательный палец ноги.

Итоговый стимульный материал представлял собой черные точки на белом фоне, которые в динамике воспринимаются как движения тела человека (рис. 1). Для выявления роли количества кинематической информации стимульный материал демонстрировался с плавным движением (23 кадра/с, что соответствует длительности каждого кадра 43,5 мс) либо в условиях стробоскопического движения (2 кадра/с, что соответствует длительности каждого кадра 500 мс). Такая длительность обеспечивает возникновение кажущегося движения в первом случае и разрушает впечатление о движении во втором (Baker, Braddick, 1985; Mather et al., 1992; McDonnell et al., 2007).

Процедура исследования была аналогична использованной ранее (Королькова, 2024). Экспозиция стимульного материала и регистрация ответов выполнялись в удаленной форме через веб-браузер с использованием ПО jatos 3.5.8 и jspysch 6.3.0 на персо-



нальных компьютерах/ноутбуках участников, мобильные устройства не поддерживались. Участникам предоставлялось краткое описание исследования, объяснялись условия участия, предлагалось ввести свой пол и возраст и заполнить Торонтскую шкалу алекситимии TAS-20 в адаптации Е.Г. Старостиной и др. (Старостина и др., 2010). Методика TAS-20 применялась для выявления и последующего исключения из анализа данных тех участников, которые испытывают трудности с пониманием и вербализацией эмоций. Процедура была одобрена Этическим комитетом ФГБОУ ВО МГППУ.

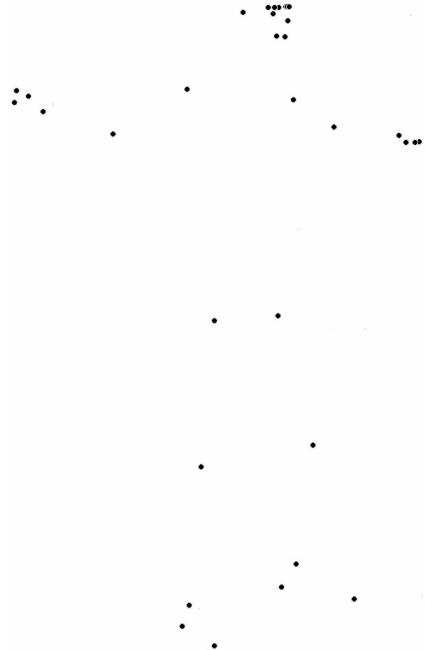


Рис. 1. Пример кадра из динамического стимульного материала
Fig. 1. An example frame from dynamic stimuli

В основной части исследования каждая проба начиналась с экспозиции фиксационного креста в центре экрана, его длительность случайно варьировалась в диапазоне 500–1200 мс. Затем в случайном порядке демонстрировались динамические стимулы длительностью 5–29 с (среднее 12 с), масштабированные до размера 1245×700 пикселей. Каждый стимул предъявлялся один раз без возможности повторного просмотра, после чего участникам предлагалось ответить на следующие вопросы:

1. В каком эмоциональном состоянии находится человек из видеоролика?
2. По каким признакам Вы определили эмоциональное состояние?
3. Как Вы думаете, что вызвало это состояние?

Вопрос № 1 имел пять вариантов ответа, а также вариант «Другое», при выборе которого участники могли вписать ответ в свободной форме в текстовое поле. Вопросы №№ 2 и 3 предполагали свободный ввод текста. Все вопросы являлись обязательными и предъявлялись последовательно. После ответа на третий вопрос участник переходил к просмотру следующего стимула. По окончании исследования участнику присваивался уникальный код и предлагалась форма обратной связи.



Анализ данных проводился в среде статистической обработки R 4.4.0 (R Core Team, 2024) с использованием пакетов lme4 1.1.35.5, lmerTest 3.1.3 и multcomp 1.4.26. Данные участников, имеющих балл по шкале TAS-20 выше 65, исключались из анализа. Строилась обобщенная линейная модель, в которую включались случайные эффекты испытуемого, его пола и возраста, балла по шкале TAS-20 и стимула. Фиксированными факторами служили условия экспозиции (межгрупповой фактор с двумя уровнями: плавная экспозиция в серии 1 и стробоскопическая экспозиция в серии 2), модальность экспрессии (внутригрупповой фактор, 7 уровней) и их взаимодействие. Зависимой переменной являлась точность распознавания выраженных эмоций (правильно/неправильно), которая оценивалась по результатам ответов на вопрос № 1. Апостериорные контрасты с поправкой Бенджамина и Хохберга применялись для сравнения точности распознавания в различных условиях. Дополнительно частоты категорий верных и «ошибочных» ответов для каждой экспрессии сравнивались между условиями экспозиции при помощи точного теста Фишера с поправкой Бенджамина и Хохберга.

Для сравнения с результатами предыдущего исследования строилась обобщенная линейная модель, включающая данные обоих экспериментов. В нее были включены те же случайные факторы, а в качестве фиксированных факторов использовались условия экспозиции (межгрупповой фактор, 6 уровней: полная информация; расфокусированное лицо; расфокусированное тело; расфокусированное тело \times 2; плавная экспозиция фигуры из точек; стробоскопическая экспозиция фигуры из точек), модальность экспрессии и их взаимодействие. Характеристики различных условий экспозиции полноцветных видеофрагментов были детально описаны ранее (Королькова, 2024). В условиях полной информации демонстрировались исходные видеофрагменты, не подверженные каким-либо трансформациям или окклюзиям; в условиях ограничения видимости лица либо тела на часть изображения накладывалась расфокусирующая маска, скрывающая соответствующие признаки экспрессии.

Результаты

Вклад кинематической информации в распознавание экспрессий тела

На основании анализа результатов участников по шкале алекситимии TAS-20 были исключены данные 7 человек. В оставшейся выборке медианный балл по шкале TAS-20 составил 43. Характеристики качества построенной регрессионной модели: псевдо- R^2_c Найджелкерка – 0,24; информационный критерий Акаике (AIC) – 5798,5. Величина случайных эффектов: испытуемого – 0,14; стимула – 0,35. Случайные эффекты пола и возраста испытуемых и балла по TAS-20 были исключены из финальной модели по причине близости к нулю.

В условиях плавной экспозиции фигуры из точек средняя по всем экспрессиям точность распознавания составила 0,61. При стробоскопической экспозиции она значимо снижалась ($z = 10,94; p < 0,001$) до 0,4. Сравнение между условиями по каждой из экспрессий показало, что значимое снижение наблюдается для всех экспрессий, за исключением радости и спокойствия (табл. 1; рис. 2). При плавной экспозиции максимальная точность распознавания характерна для скуки (0,75) и страха (0,73), тогда как для остальных экспрессий точность уменьшается (0,64–0,49). При стробоскопической экспозиции соотношение точности меняется: наиболее эффективно распознаются спокойствие (0,67) и скука (0,56), а для других экспрессий наблюдается падение точности до 0,39–0,17.



Таблица 1 / Table 1

Точность распознавания экспрессий в зависимости от условий

экспозиции и категории эмоции

Accuracy of expression recognition depending on exposure

conditions and emotion category

Категория эмоции / Emotion category	Предсказанная точность / Predicted accuracy		Сравнение между условиями / Comparison between conditions	
	Плавная экспозиция / Smooth exposure	Стробоскопическая экспозиция / Stroboscopic exposure	z	p
Радость / Happiness	0,51	0,39	2,59	0,065
Страх / Fear	0,73	0,35	8,84	< 0,001
Отвращение / Disgust	0,6	0,29	7,32	< 0,001
Спокойствие / Neutral state	0,64	0,67	-0,76	0,983
Возбуждение / Excitement	0,55	0,39	3,66	0,002
Скука / Boredom	0,75	0,56	4,69	< 0,001
Разочарование / Disappointment	0,49	0,17	8,48	< 0,001

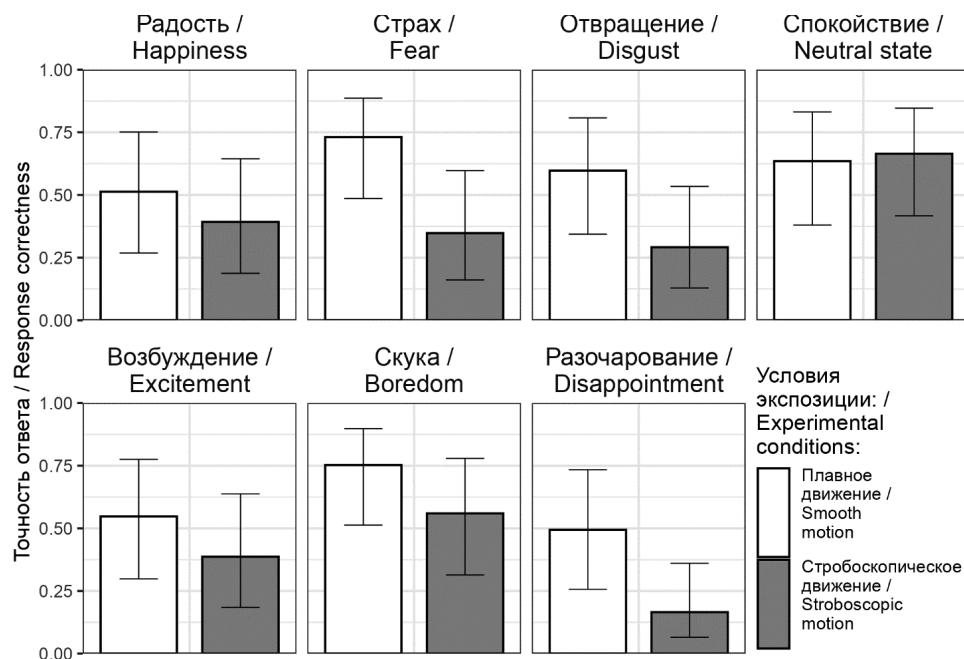


Рис. 2. Предсказанная точность и доверительные интервалы оценки экспрессий
по динамическим точечным паттернам

Fig. 2. Predicted accuracy and confidence intervals of expression evaluation by point-light displays



При восприятии динамических точечных паттернов в значительной доле проб наблюдались «ошибочные» ответы, не совпадающие с эмоциональным состоянием, которое выражали натуращики при съемке видеоклипов. Они характеризуют широкое категориальное поле восприятия экспрессий – те состояния, с которыми ядро эмоциональной категории наиболее часто смешивается. Так, при плавной экспозиции экспрессии радости она оценивается как «интерес» (частота оценок 0,18) и реже – как «удивление» (0,1). Страх в небольшом числе случаев распознается как «стыд» (0,08), а скука – как «возбуждение» (0,09). Спокойствие может восприниматься как «расстроенное» состояние (0,16); отвращение – как «интерес» (0,18) и «разочарование» (0,1); возбуждение – как «интерес» (0,23); разочарование – как «беспокойство» (0,35).

В условиях стробоскопической экспозиции точечных паттернов структура категориальных полей меняется. По сравнению с плавным предъявлением значимо возрастают частоты следующих «ошибочных» ответов (табл. 2): экспрессия радости чаще оценивается как «удивление» и «злость»; экспрессия страха – как «стыд», «разочарование» и «доброта»; экспрессия отвращения – как «испуг», «грусть» и «интерес»; экспрессия возбуждения – как «интерес», «обида» и «скука»; экспрессия разочарования – как «беспокойство», «испуг» и «подщечивание»; экспрессия скуки – как «расстройство». При плавной экспозиции экспрессии спокойствия участники чаще выбирали опцию свободного ввода ответа и описывали данное состояние как «испуг» или «напряжение».

Таблица 2 / Table 2

Сравнение частот ответов в зависимости от условий экспозиции и категории эмоции
Response frequency comparison depending on exposure conditions and emotion category

Категория эмоции / Emotion category	Ответ / Response	Отношение шансов / Odds ratio	p
Радость / Happiness	Радостный / Happy	1,44 [1,05; 1,99]	0,047
	Заинтересованный / Interested	1,23 [0,80; 1,90]	0,492
	Удивленный / Surprised	0,47 [0,26; 0,82]	0,015
	Скучающий / Bored	0,99 [0,43; 2,21]	1,000
	Злой / Angry	0,43 [0,19; 0,91]	0,047
	Другое / Other	0,96 [0,63; 1,47]	0,989
Страх / Fear	Испуганный / Afraid	4,55 [3,23; 6,46]	< 0,001
	Стыдящийся / Ashamed	0,24 [0,12; 0,43]	< 0,001
	Недружелюбный / Unfriendly	0,61 [0,30; 1,20]	0,256
	Разочарованный / Disappointed	0,00 [0,00; 0,27]	< 0,001
	Добрый / Kind	0,17 [0,08; 0,35]	< 0,001
	Другое / Other	1,07 [0,67; 1,72]	0,905
Отвращение / Disgust	Испытывающий отвращение / Disgusted	3,42 [2,45; 4,79]	< 0,001
	Испуганный / Afraid	0,23 [0,09; 0,54]	< 0,001
	Разочарованный / Frustrated	0,89 [0,52; 1,50]	0,825
	Грустный / Sad	0,24 [0,06; 0,70]	0,014
	Заинтересованный / Interested	0,49 [0,33; 0,71]	0,001
	Другое / Other	0,70 [0,39; 1,23]	0,356
Спокойствие / Neutral state	Спокойный / Neutral	0,88 [0,63; 1,23]	0,605



Категория эмоции / Emotion category	Ответ / Response	Отношение шансов / Odds ratio	p
Возбуждение / Excitement	Скучающий / Bored	0,55 [0,28; 1,03]	0,110
	Добрый / Kind	0,82 [0,21; 2,76]	0,897
	Удивленный / Surprised	1,37 [0,54; 3,46]	0,640
	Расстроенный / Frustrated	1,17 [0,75; 1,84]	0,640
	Другое / Other	2,05 [1,06; 4,02]	0,049
Скука / Boredom	Возбужденный / Excited	1,75 [1,27; 2,42]	0,001
	Заинтересованный / Interested	0,66 [0,45; 0,95]	0,047
	Подшучивающий / Joking	0,00 [0,00; +∞]	1,000
	Обиженный / Hurt	0,35 [0,13; 0,85]	0,030
	Скучающий / Bored	0,00 [0,00; 0,58]	0,012
	Другое / Other	1,09 [0,73; 1,62]	0,825
Разочарование / Disappointment	Скучающий / Bored	2,17 [1,54; 3,07]	< 0,001
	Расстроенный / Frustrated	0,41 [0,23; 0,72]	0,003
	Грустный / Sad	0,36 [0,10; 1,00]	0,067
	Обиженный / Hurt	0,38 [0,14; 0,93]	0,051
	Возбужденный / Excited	0,79 [0,45; 1,37]	0,589
	Другое / Other	0,97 [0,50; 1,83]	1,000

Соотношение информации о форме и кинематике при распознавании экспрессий тела

Мы сопоставили полученные результаты с данными ранее проведенного исследования, в котором демонстрировались полноцветные видеозаписи тех же натурщиков либо с расфокусированным лицом, либо с расфокусированным телом, либо без какой-либо ретуши изображения (Королькова, 2024). Для построенной регрессионной модели, включающей все данные, псевдо- R^2 Найджелкерка составил 0,32; информационный критерий Акаике (AIC) — 12663,8. Величина случайных эффектов: испытуемого — 0,22; стимула — 0,17. Случайные эффекты пола и возраста испытуемых и балла по TAS-20 также были исключены из финальной модели, поскольку приближались к нулевым значениям.

Сравнение показало, что точность распознавания экспрессии радости значимо снижалась при демонстрации фигуры из точек с любой частотой кадров по сравнению со всеми типами полноцветных видео ($z > 7,26; p < 0,01$), тогда как дальнейшее снижение точности при уменьшении частоты кадров не достигло уровня статистической значимости. Экспрессия отвращения, представленная в виде фигуры из точек с плавным движением, также распознавалась значимо хуже, чем все полноцветные видео ($z > 4,22; p < 0,01$). Точность еще более снижалась при уменьшении частоты кадров ($z > 7,21; p < 0,01$). Для разочарования характерен тот же паттерн ответов: высокий уровень распознавания полноцветных видео как с



ретушью, так и без нее; значимое снижение точности при демонстрации фигуры из точек ($z > 4,83; p < 0,01$) и дальнейшее снижение при ее стробоскопической экспозиции ($z = 8,35; p < 0,01$). Экспрессия страха одинаково хорошо распознавалась по полноцветному изображению без ретуши либо с расфокусированным лицом, но значимо хуже ($z > 4,04; p < 0,01$) – в условиях расфокусированного тела либо фигуры из точек с плавным движением, между которыми различий не выявлено ($z < 1,46; p = 1$). Стробоскопическая экспозиция фигуры из точек значимо снижала точность оценок страха по сравнению со всеми другими условиями ($z > 5,52; p < 0,01$). Спокойное (нейтральное) состояние распознавалось по фигурам из точек с той же точностью, что по видеофрагментам без ретуши ($z < 1,09; p = 1$). При этом максимальный среди всех условий уровень точности распознавания достигался при расфокусировании лица ($z > 3,54; p < 0,034$), а условия, в которых было открыто только лицо, напротив, приводили к наименьшей точности распознавания; в ряде случаев различия достигали статистической значимости. Для экспрессии возбуждения единственным значимым изменением было снижение точности ответа при уменьшении количества кадров фигуры из точек (см. табл. 1); во всех остальных условиях точность сохранялась на среднем уровне, который был самым низким относительно других экспрессий. Точность оценок скуки значимо не менялась при демонстрации плавного движения фигуры из точек по сравнению с полноцветным видео без ретуши либо с ретушью лица ($z < 1,1; p = 1$), но превышала точность, полученную в условиях расфокусированного тела ($z > 4,03; p < 0,01$). При снижении частоты кадров точность распознавания экспрессии фигуры из точек также уменьшалась ($z > 4,47; p < 0,01$).

Таким образом, полученные результаты позволили выявить основные паттерны вклада формы и движения в восприятие эмоциональных экспрессий: 1) первичная роль формы при минимальном вкладе движения (экспрессия радости); 2) значимый вклад как движения, так и формы лица либо тела (экспрессии отвращения и разочарования); 3) значимый вклад обоих источников информации о теле, но не о лице (экспрессии страха и скуки); 4) ключевая роль кинематики движения (экспрессия возбуждения); 5) разделение вклада лица и тела при относительно статичной позе (спокойствие).

Обсуждение результатов

Проведенное исследование частично подтвердило выдвинутые гипотезы. В частности, при уменьшении частоты кадров динамического точечного паттерна мы выявили снижение точности оценок выражений страха, отвращения, возбуждения, скуки и разочарования. Однако различия для радости и спокойствия были не значимы. Также в условиях стробоскопической экспозиции для всех экспрессий менялся паттерн ошибочных идентификаций. Сравнение с данными нашего предыдущего исследования показало, что при удалении информации о форме, текстуре и контексте ситуации снижается точность оценок экспрессий радости, отвращения, разочарования, страха, но для экспрессий возбуждения и спокойствия она сохраняется. Полученные результаты свидетельствуют о дифференцированном вкладе различных источников информации в зависимости от категории воспринимаемой эмоциональной экспрессии.

Резкое падение точности распознавания экспрессии радости по динамической фигуре из точек (0,51) по сравнению с полноцветными видеофрагментами (0,89–0,96) позволяет предположить, что основной вклад в восприятие данного состояния вносят особенности



формы и текстуры изображения человека, а также контекст ситуации. Если эти признаки недоступны наблюдателю, распознавание радости разрушается. Она ошибочно идентифицируется прежде всего как интерес (0,18) и удивление (0,1) и в меньшей степени как другие состояния — возбуждение, волнение, тревога. Дальнейшее уменьшение количества кинематической информации приводит к еще большему снижению точности до 0,39, которое, однако, не достигает уровня статистической значимости. Таким образом, кинематические паттерны движения сами по себе не позволяют надежно отличить выражение радости от других состояний с высоким уровнем активации, таких как интерес либо удивление. Данные результаты не согласуются с более ранним исследованием, в котором не было получено различий для экспрессии радости, распознаваемой по полноцветным или точечным динамическим стимулам (Atkinson et al., 2004), однако говорят в пользу первичной роли кинематики в различении высоко- и низкоинтенсивных экспрессий (Dael et al., 2013; Fourati, Pelachaud, 2018). В отличие от нашего исследования, в работе Аткинсона и др. (Atkinson et al., 2004) видеосъемка актеров происходила без использования какого-либо реквизита, создающего контекст ситуации, однако этим нельзя объяснить рассогласованность результатов, поскольку в условиях полноцветного видео были получены результаты, сопоставимые с нашими (87% верных ответов).

В случае экспрессий отвращения и разочарования наблюдаются сходные паттерны соотношения оценок, а различия между плавной и стробоскопической экспозицией достигают значимости. Можно говорить о том, что для надежного распознавания обеих экспрессий важна информация как о форме, так и о кинематике. Так, в нашем предыдущем исследовании (Королькова, 2024) отвращение оценивалось по полноцветным видеофрагментам на высоком уровне точности (0,82–0,9), в то время как в условиях доступности только кинематических паттернов точность падает до 0,6, а при минимальной информации о движении — до 0,29. Эти данные сопоставимы с более ранним исследованием (Atkinson et al., 2004), в котором получена точность 0,75 для полноцветного видео и 0,63 — для точечных паттернов.

Разочарование при наличии полной информации распознавалось на уровне 0,78–0,92; при исключении формы и текстуры точность составила 0,49, а при снижении частоты кадров — 0,17, что сопоставимо со случаем уровнем. Мы не нашли в литературе исследований, посвященных распознаванию экспрессии разочарования по точечным паттернам, однако данные автоматизированного распознавания разочарования по видеоизображениям движений тела и жестов (Vu et al., 2012), а также данные валидизации базы EU-ESM, видео из которой мы использовали (O'Reilly et al., 2016), демонстрируют точность на уровне 0,97 и более. При распознавании по полноцветным видео состояния фruстрации, которая является семантически близкой разочарованию, получены результаты на уровне 78% (Smekal et al., 2011).

Для страха и скуки наличие либо отсутствие информации об экспрессии лица существенно не влияет на точность распознавания, тогда как оба типа информации об экспрессии тела — его статическая и кинематическая характеристики — а также ситуативный контекст вносят значимый вклад в восприятие. Точность оценки страха, полученная в нашем исследовании (0,92 для полноцветных видео и 0,73 для точечных паттернов), согласуется с результатами Аткинсона и др. (Atkinson et al., 2004) (0,91 для полноцветных видео и 0,79 для точечных паттернов страха).

Распознавание экспрессии скуки также максимально в тех условиях, в которых присутствуют естественные кинематические паттерны тела (неретушированный видеофраг-



мент – 0,76; видео с расфокусированным лицом – 0,8; фигура из точек с плавной сменой кадров – 0,75). Снижение точности происходит при отсутствии достаточной информации о движениях тела (0,56).

Точность оценок экспрессии возбуждения была самой низкой среди других состояний, которые мы изучали (0,37–0,49 для полноцветных видео; 0,55 для плавного движения точечного паттерна и 0,39 – для его стробоскопической экспозиции), а единственное значимое различие было выявлено при снижении количества информации о движении. Наличие формы и текстуры в незначительной степени даже затрудняло распознавание. В 23% и более случаев возбуждение оценивалось как интерес.

Для распознавания спокойного состояния ключевую роль играет не статика или динамика сами по себе, а наличие либо отсутствие информации о микроэкспрессии лица. В ее отсутствие точность распознавания максимальна (0,83), а когда у наблюдателя есть возможность рассмотреть лицо, он обращает внимание на микроэкспрессии спокойного лица, находя в них признаки других состояний – расстройства либо скуки.

Проведенное исследование имеет ряд ограничений. Во-первых, остается открытым вопрос о вкладе контекста ситуации в оценку эмоциональных состояний. В используемом стимульном материале – динамических точечных паттернах – контекст как таковой отсутствовал. Однако он был представлен в полноцветных видеофрагментах, с распознаванием которых мы сопоставляли данные текущего эксперимента (Королькова, 2024). Для выявления дифференцированной роли контекста, не сводящейся к характеристикам формы тела и манипулированию натуралистов с различными предметами, требуется проведение дополнительной серии исследования. Во-вторых, исследование было проведено на выборке, преимущественно включавшей женщин. Опираясь на результаты других исследований, выявивших особенности распознавания действий и состояний по динамическим точечным паттернам у мужчин и женщин (Alaerts et al., 2011), при более сбалансированной по полу выборке можно ожидать общего снижения точности ответов.

Перспективы дальнейших исследований состоят в более детальном изучении способов проявления эмоциональных переживаний в невербальном поведении: какие именно особенности кинематики жеста, походки, действий с предметами вызывают у наблюдателя впечатление той или иной эмоции. В частности, интерес представляет анализ движений, типичных для той или иной переживаемой эмоции, и их характеристик. С использованием системы кодирования движений, разработанной Р. Лабаном и его последователями, было показано, что с восприятием и переживанием состояния радости связаны свободные, легкие движения, открытые позы тела и его подъем, прыжки и ритмичность движений, а с состоянием страха – мышечный контроль движений, отступление назад, сжатие и скручивание тела, уменьшающее занимаемый объем пространства (Melzer et al., 2019; Shafit et al., 2016). Количественный анализ кинематики движений также будет способствовать развитию области автоматизированного распознавания эмоций по движениям тела (Ahmed, Bari, Gavrilova, 2020).

Заключение

На материале динамических точечных паттернов, соответствующих экспрессивному невербальному поведению натуралистов, изучена роль формы и движения в оценке эмоциональных состояний радости, страха, отвращения, возбуждения, разочарования,



скуки и спокойствия. Выявлены дифференцированные паттерны вклада двух источников информации в восприятие экспрессий тела в зависимости от категории выраженной эмоции. Для экспрессии радости первичную роль играет форма при минимальном вкладе движения. Для экспрессий отвращения и разочарования продемонстрирован значимый вклад как движения, так и формы, и ситуативного контекста. На восприятие экспрессий страха и скуки значимое влияние также оказывают оба источника информации о теле, но не о лице. Выявлена ключевая роль движения в распознавании экспрессии возбуждения. При оценке спокойствия происходит разделение вклада лица и тела. Ключевую роль играет не статика или динамика сами по себе, а наличие либо отсутствие информации о микроязыкажных лицах.

Список источников / References

1. Королькова, О.А. (2024). Восприятие эмоционального состояния коммуниканта на основе информации об экспрессиях его лица и тела. *Экспериментальная психология*, 17(4), 28–43. <https://doi.org/10.17759/exppsy.2024170402>
Korolkova, O.A. (2024). Perception of Emotional State of a Communicant Based on Information about His Facial and Body Expressions. *Experimental Psychology (Russia)*, 17(4), 28–43. (In Russ.). <https://doi.org/10.17759/exppsy.2024170402>
2. Старостина, Е.Г., Тэйлор, Г., Квилти, Л., Бобров, А., Мошняга, Е., Пузырева, Н., Боброва, М., Ивашкина, М., Кривчикова, М., Шаврикова, Е., Бэгби, М. (2010). Торонтская шкала алекситимии (20 пунктов): Валидизация русскоязычной версии на выборке терапевтических больных. *Социальная и Клиническая Психиатрия*, 20(4), 31–38.
Starostina, E.G., Taylor, G., Quilty, L., Bobrov, A., Moshnyaga, E., Puzyreva, N., Bobrova, M., Ivashkina, M., Krivchikova, M., Shavrikova, E., Bagby, M. (2010). Toronto Alexithymia Scale (20 items): Validation of the Russian version on a sample of medical patients. *Social and Clinical Psychiatry*, 20(4), 31–38. (In Russ.).
3. Ahmed, F., Bari, A.S.M.H., Gavrilova, M.L. (2020). Emotion recognition from body movement. *IEEE Access*, 8, 11761–11781. <https://doi.org/10.1109/ACCESS.2019.2963113>
4. Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., Wenderoth, N. (2011). Action and emotion recognition from point light displays: An investigation of gender differences. *PLoS ONE*, 6(6), 1–9. <https://doi.org/10.1371/journal.pone.0020989>
5. Atkinson, A.P., Dittrich, W.H., Gemmell, A.J., Young, A.W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, 33(6), 717–746. <https://doi.org/10.1088/p5096>
6. Atkinson, A.P., Tunstall, M.L., Dittrich, W.H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition*, 104(1), 59–72. <https://doi.org/10.1016/j.cognition.2006.05.005>
7. Baker, C.L.J., Braddick, O.J. (1985). Temporal properties of the short-range process in apparent motion. *Perception*, 14, 181–192. <https://doi.org/10.1088/p140181>
8. Bazarevsky, V., Grishchenko, I., Raveendran, K., Zhu, T.L., Zhang, F., Grundmann, M. (2020). BlazePose: On-device Real-time Body Pose tracking. *ArXiv*, abs/2006.1, 1–4.
9. Beintema, J.A., Utrecht, U., Georg, K., Lappe, M. (2006). Perception of biological motion from limited-lifetime stimuli. *Perception & Psychophysics*, 68(4), 613–624.
10. Dael, N., Goudbeek, M., Scherer, K.R. (2013). Perceived gesture dynamics in nonverbal expression of emotion. *Perception*, 42(6), 642–657. <https://doi.org/10.1088/p7364>
11. De Gelder, B., Poyo Solanas, M. (2022). What postures communicate. *Enfance*, 3(3), 353–365. <https://doi.org/10.3917/ enf2.223.0353>
12. Fourati, N., Pelachaud, C. (2018). Perception of Emotions and Body Movement in the Emilya Database. *IEEE Transactions on Affective Computing*, 9(1), 90–101. <https://doi.org/10.1109/TAFFC.2016.2591039>



13. Gardenfors, P., Warglien, M. (2012). Using Conceptual Spaces to Model Actions and Events. *Journal of Semantics*, 29(4), 487–519. <https://doi.org/10.1093/jos/ffs007>
14. Giese, M.A., Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4(3), 179–192. <https://doi.org/10.1038/nrn1057>
15. Ikeda, E., Destler, N., Feldman, J. (2025). The role of dynamic shape cues in the recognition of emotion from naturalistic body motion. *Attention, Perception, and Psychophysics*, 87(2), 604–618. <https://doi.org/10.3758/s13414-024-02990-8>
16. Ildei, C.B., Benahmed, O., Bouidaine, D., Francisco, V., Decatoire, A. (2024). SmartDetector: Automatic and vision-based approach to point-light display generation for human action perception. *Behavior Research Methods*, 56, 8349–8361.
17. Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211. <https://doi.org/10.3758/BF03212378>
18. Mather, G., Radford, K., West, S. (1992). Low-Level Visual Processing of Biological Motion. *Proceedings of the Royal Society B: Biological Sciences*, 249, 149–155. <https://doi.org/10.1098/rspb.1992.0097>
19. Mcdonnell, R., Newell, F., Sullivan, C.O. (2007). Smooth Movers: Perceptually Guided Human Motion Simulation. In: D. Metaxas & J. Popovic (Eds.), *SCA '07: Proceedings of the 2007 ACM SIGGRAPH/Eurographics symposium on Computer animation* (pp. 259–269). ACM.
20. Melzer, A., Shafir, T., Tsachor, R.P. (2019). How Do We Recognize Emotion From Movement? Specific Motor Components Contribute to the Recognition of Each Emotion. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.01389>
21. O'Reilly, H., Pigat, D., Fridenson, S., Berggren, S., Tal, S., Golan, O., Bölte, S., Baron-Cohen, S., Lundqvist, D. (2016). The EU-Emotion Stimulus Set: A validation study. *Behavior Research Methods*, 48(2), 567–576. <https://doi.org/10.3758/s13428-015-0601-4>
22. Presti, P., Ruzzon, D., Galasso, G.M., Avanzini, P., Caruana, F., Vecchiato, G. (2022). The Avatar's Gist: How to Transfer Affective Components From Dynamic Walking to Static Body Postures. *Frontiers in Neuroscience*, 16, 1–14. <https://doi.org/10.3389/fnins.2022.842433>
23. R Core Team (2024). R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing, Vienna, Austria*. URL: <https://www.r-project.org/>
24. Runeson, S., Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, 112(4), 585–615. <https://doi.org/10.1037/0096-3445.112.4.585>
25. Shafir, T., Tsachor, R.P., Welch, K.B. (2016). Emotion regulation through movement: Unique Sets of Movement Characteristics are Associated with and Enhance Basic Emotions. *Frontiers in Psychology*, 6, 1–15. <https://doi.org/10.3389/fpsyg.2015.02030>
26. Smekal, V., Solanas, M.P., Fraats, E.I.C., Gelder, B. de. (2011). Differential contributions of body form, motion, and temporal information to subjective action understanding in naturalistic stimuli. *Frontiers in Integrative Neuroscience*, 18, 1–12. <https://doi.org/10.3389/fnint.2024.1302960>
27. Vangeneugden, J., Peelen, M.V., Tadin, D., Battelli, L. (2014). Distinct Neural Mechanisms for Body Form and Body Motion Discriminations. *The Journal of Neuroscience*, 34(2), 574–585. <https://doi.org/10.1523/JNEUROSCI.4032-13.2014>
28. Venture, G., Kadone, H., Zhang, T., Grèzes, J., Berthoz, A. (2014). Recognizing Emotions Conveyed by Human Gait. *International Journal of Social Robotics*, 6, 621–632. <https://doi.org/10.1007/s12369-014-0243-1>
29. Vu, H.A., Yamazaki, Y., Dong, F., Hirota, K. (2012). Emotion Recognition based on Human Gesture and Speech Information using RT Middleware. In: *IEEE International Conference on Fuzzy Systems* (pp. 787–791).

Информация об авторе

Ольга Александровна Королькова, кандидат психологических наук, ведущий научный сотрудник Института экспериментальной психологии, Московский государственный психолого-педагогиче-



ский университет (ФГБОУ ВО МГППУ), Москва, Российская Федерация, ORCID: <https://orcid.org/0000-0003-4814-7266>, e-mail: olga.kurakova@gmail.com

Information about the author

Olga A. Korolkova, PhD in Psychology, Leading Research Associate, Institute of Experimental Psychology, Moscow State University of Psychology and Education, Moscow, Russian Federation, ORCID: <https://orcid.org/0000-0003-4814-7266>, e-mail: olga.kurakova@gmail.com

Конфликт интересов

Автор заявляет об отсутствии конфликта интересов.

Conflict of interest

The author declares no conflict of interest.

Декларация об этике

Исследование было рассмотрено и одобрено Этическим комитетом ФГБОУ ВО «Московский государственный психолого-педагогический университет» (протокол № 2 от 25.04.2025).

Ethics statement

The study was reviewed and approved by the Ethics Committee of Moscow State University of Psychology and Education (report no. 2, 2025/04/25).

Поступила в редакцию 10.11.2025

Received 2025.11.10

Поступила после рецензирования 20.11.2025

Revised 2025.11.20

Принята к публикации 20.11.2025

Accepted 2025.11.20

Опубликована 30.12.2025

Published 2025.12.30