

УДК 159.9.07; 159.942.31; 159.942.33

АНАЛИЗ ВАРИАТИВНОСТИ ДИНАМИКИ ЭКСПРЕССИИ РАДОСТИ

А.В. Жегалло

Институт психологии Российской академии наук (ИП РАН)

Россия, 129366, Москва, Ярославская ул., 13, к.1.

E-mail: zhegalloav@ipran.ru

Ключевые слова: эмоциональные экспрессии, ADFES, нейрокультурная теория эмоций, BECV, Action Units, Principal Component Analysis, Facial Landmarks, Point Distribution Model.

Аннотация: выполнен анализ видеозаписей динамики экспрессии радости из базы ADFES. Для первоначальной разметки использовалось ПО OpenFace 2.0. Априорная индивидуальная динамика экспрессии радости характеризуется согласованным изменением интенсивности AU6 и AU12, а также согласованным изменением интенсивности 3-х фиксированных главных компонент. Эмпирически выделяемые индивидуальные паттерны характеризуются высокой вариативностью объяснительной способности и относительно низкой согласованностью между собой. Обобщенный динамический паттерн экспрессии радости характеризуется низкой объяснительной способностью, что противоречит представлениям об эмоциональных экспрессиях как универсальных коммуникационных сигналах.

В рамках нейрокультурной теории эмоций П. Экмана [1] проблема распознавания эмоционального состояния человека по выражению лица решается соотношением с нормативными эталонами. При этом эмпирическое обоснование нормативов отсутствует. Экман признает: “Ideally there should be various sources of evidence to demonstrate that an AU combination is a sign of an emotion. This would include cross-cultural and developmental data which showed how the occurrence of the AU combination was related to environmental circumstance, subjective experience, psychophysiological changes, other motor behavior, vocalizations, speech, and the behavior of other interactants who observe the AUs. No such complete evidence exists for any of the entries in Table 10-1. There is evidence for each entry, but it is partial” [2, P. 173].

В противовес экмановской теории Алан Фридлунд сформулировал теорию The Behavioral Ecology View of facial displays (BECV), рассматривающую выражения лица как «социальные инструменты», имеющие смысл в рамках контекста взаимодействия с другими людьми. Для выражений лица наиболее важно, как они функционируют в социальных взаимодействиях [3]. Фридлунд предлагает использовать «подход, основанный на данных», то есть изучать феноменологию выражений эмоциональных экспрессий в социальных коммуникациях.

Проведенное исследование направлено на разработку методики анализа динамических выражений лица. В качестве стимульного материала использовались видеозаписи из базы ADFES [4], 12 фронтальных видеозаписей экспрессии радости (Joy), разрешение 720×576 px, 25 fps, продолжительность записи 6 сек, что составляет 150 кадров. Натурщики: F01, F02, F03, F04, F05, M02, M03, M04, M06, M08, M11, M12. Первичная обработка видеозаписей выполнялась с помощью Open Source ПО OpenFace 2.0 [5]. В результате обработки для каждого кадра видеозаписи были получены: X, Y, Z – координаты 68 ключевых точек (facial landmarks); градуальные значения

интенсивности 17 Action Units по П. Экману (AU01, AU02, AU04, AU05, AU06, AU07, AU09, AU10, AU12, AU14, AU15, AU17, AU20, AU23, AU25, AU26, AU28, AU45); градуальные значения 6 подстроечных параметров, соответствующих масштабированию, поворотам и линейным смещениям; градуальные значения 34 non-rigid shape parameters, задающих трансформации PDM – point distribution model, используемой в программе для распознавания лица (в дальнейшем обозначаются как NRSP). Фактически данные параметры представляют собой набор априорных главных компонент, выделенных на репрезентативной выборке изображений лиц. Автор не приводит вербальные описания соответствующих трансформаций, но наличие исходного кода позволяет выполнить их графическую визуализацию.

Согласно П. Экману отличительными признаками экспрессии радости являются наличие AU6 (Cheek Raiser) и AU12 (Lip Corner Puller). Параллельная визуализация временной развертки интенсивности соответствующих AU показывает, что в диапазоне 1 сек – 2 сек происходит увеличение их интенсивности, см. рис. 1.

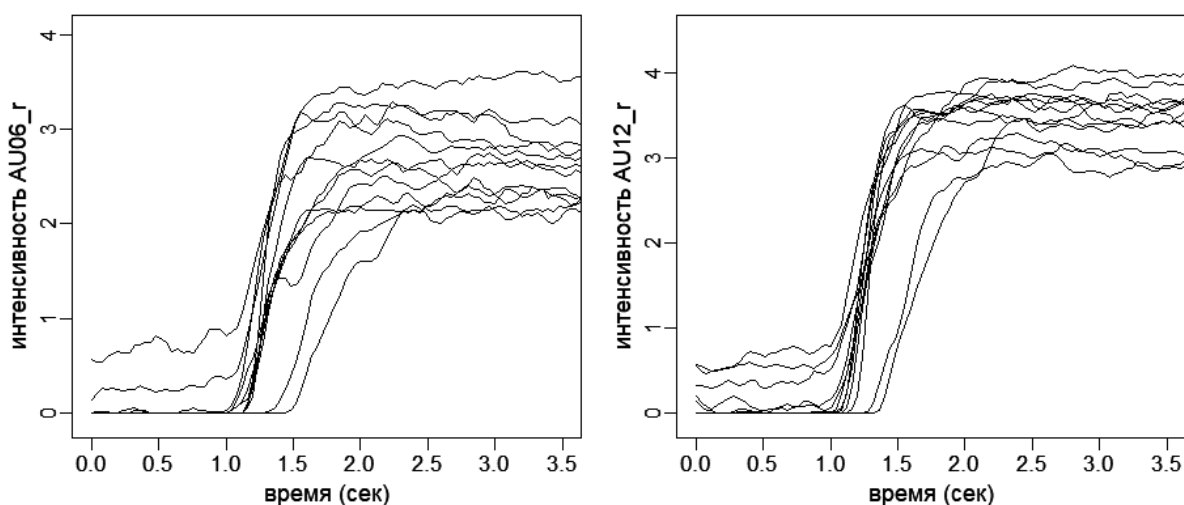


Рис. 1. Временная развертка интенсивности AU6 и AU12 для 12 наборов данных.

Визуальный анализ показывает согласованный характер динамики при наличии некоторых различий. Для выполнения численной оценки подсчитывались коэффициент конкордации Кендалла W и медианное значение коэффициента корреляции Пирсона между временными рядами. Для AU06 показатели согласованности составляют 0.81 и 0.98, для AU12 – 0.71 и 0.99. Коэффициент корреляции Пирсона между временными рядами AU6 и AU12 одного и того же натурщика составляет 0.98 – 0.99.

Присутствуют ли среди NRSP параметры, описывающие соответствующие изменения выражения лица? Анализ визуализаций временной развертки (рис. 2) показал, что на соответствующем временном диапазоне происходит рост NRSP_2 и снижение интенсивности NRSP_3 и NRSP_8. Показатели согласованности составляют NRSP_2 – 0.74, 0.99; NRSP_3 – 0.89, 0.93; NRSP_8 – 0.85, 0.83. Анализ графических визуализаций (рис. 3) показывает, что NRSP_2 и NRSP_3 совместно описывают поднятие уголков губ, NRSP_3 и NRSP_8 – сужение глаз.

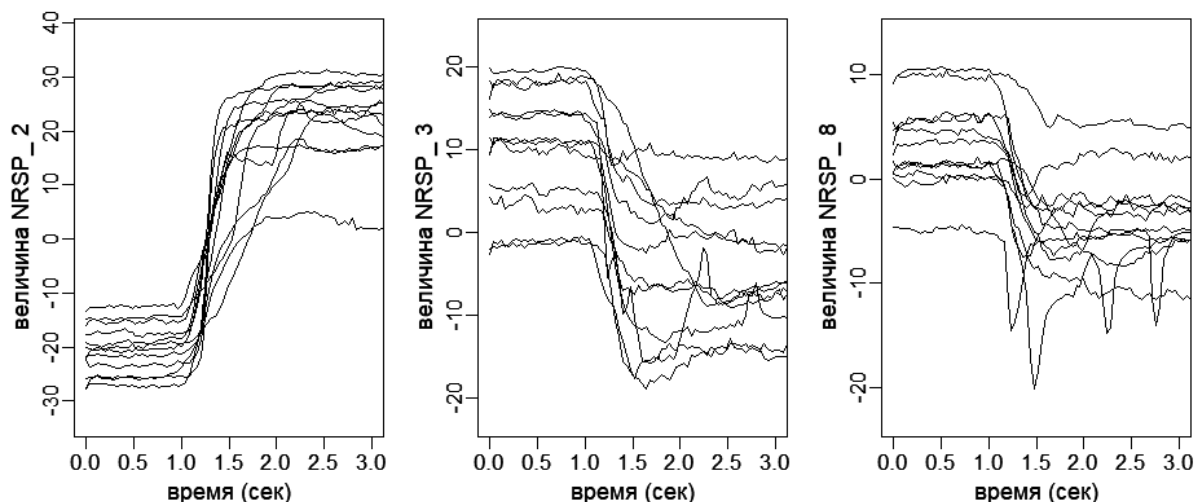


Рис. 2. Временная развертка величины NRSP_2, NRSP_3, NRSP_8 для 12 наборов данных.

Указанные NRSP объясняют соответственно: 34.6%, 9.1%, 2.7% суммарной вариативности всех NRSP или 25.7%, 6.7%, 2.0% суммарной вариативности всех подстроечных параметров модели.

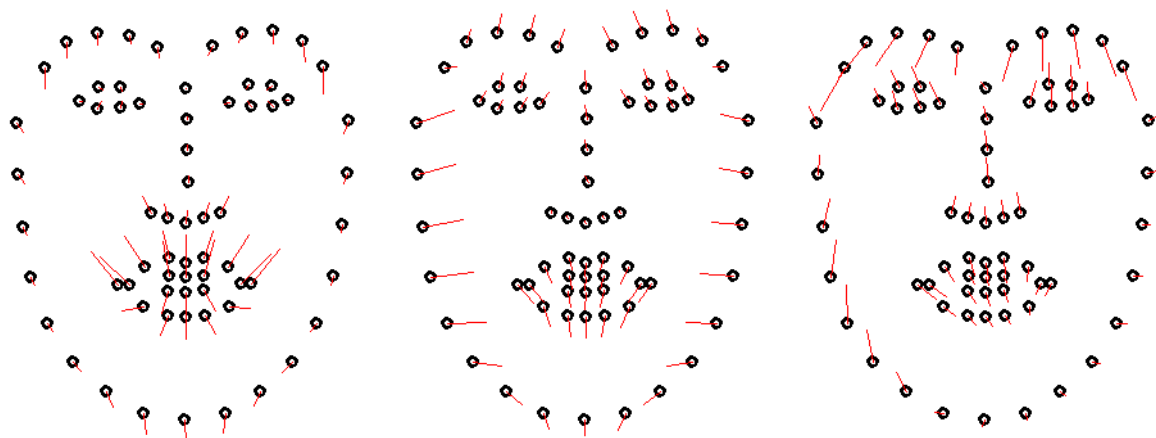


Рис. 3. Графическое представление NRSP_2, NRSP_3, NRSP_8. Черные точки – среднее положение ключевых точек. Выходящие из точек отрезки – направление изменения координат, соответствующее положительной величине NRSP.

В какой степени возможно выделение общих интегральных признаков экспрессии радости? Раздельный анализ главных компонент для каждого из 12 наборов данных показал, что временная динамика первой главной компоненты для разных наборов данных визуально согласованным образом изменяется на интервале 1 сек – 2 сек. При этом согласованность временных рядов в смысле W Кендалла низкая: .06, в смысле медианного значения коэффициента корреляции – высокая: 0.97. Также анализ главных компонент позволяет получить вектора, кодирующие каждую компоненту для каждого из наборов данных. Согласованность между ними составляет 0.08 и 0.75 соответственно. Вклад первой компоненты в объяснение общей вариации сильно варьирует: 0.87, 0.69, 0.73, 0.77, 0.75, 0.49, 0.34, 0.91, 0.88, 0.69, 0.91, 0.67 (значения соответствуют перечисленным выше кодам натурщиков).

При проведении анализа главных компонент на объединенном наборе данных согласованное изменение временной динамики на интервале 1 сек – 2 сек наблюдается

для компонент PC4 и PC5, объясняющих 3.4% и 1.3% вариативности соответственно, см. рис. 4. Согласованность PC4 составляет 0.90, 0.88; согласованность PC5 0.83, 0.97.

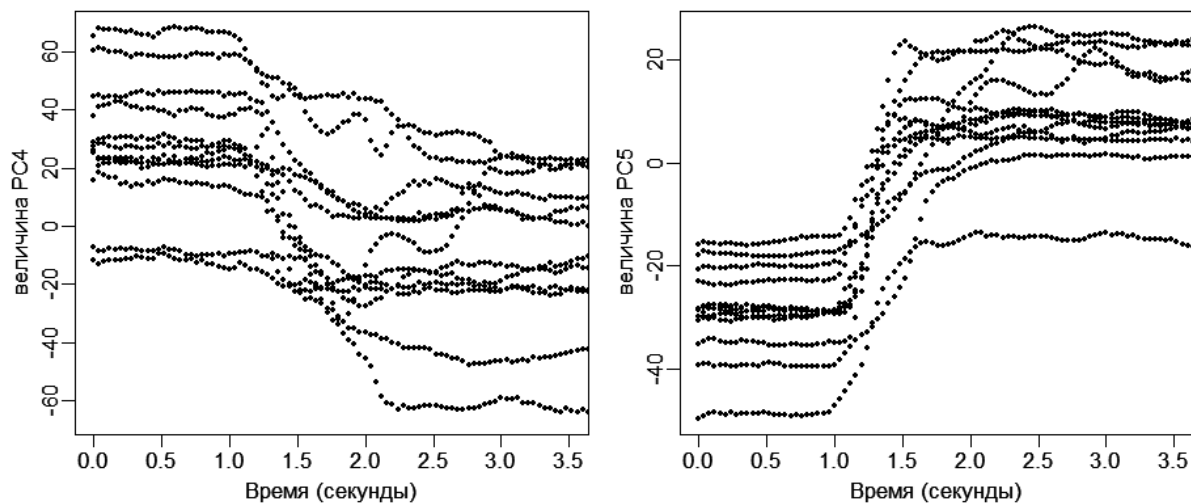


Рис. 4. Временная развертка величины PC3 и PC4 для 12 наборов данных.

Графическое представление компонент, см. рис. 5, показывает, что PC5 связана с поднятием уголков рта, а PC4 выполняет частично компенсирующую функцию.

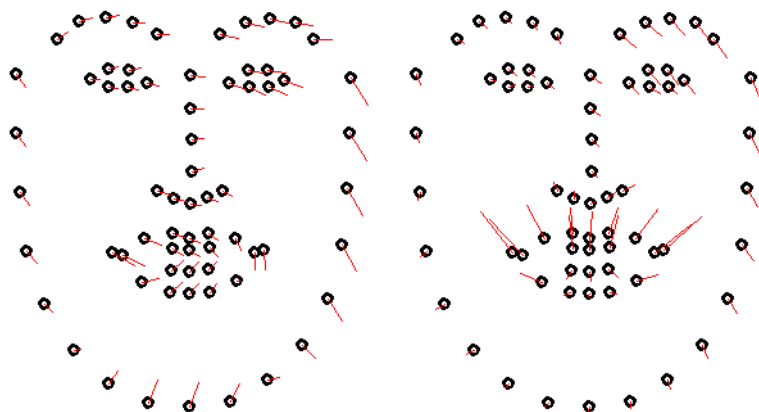


Рис. 5. Графическое представление PC4 и PC5.

Проведенный анализ показывает, что индивидуальная динамика экспрессии радости в проанализированных записях может быть приближенно описана первой главной компонентой. При этом объясняемый ей процент вариативности, может сильно различаться для разных натурщиков, что означает, что фактический состав динамики лица значительно различается. Индивидуальная динамика AU6 и AU12 хорошо коррелирует, что согласуется с общепринятыми представлениями о выражении «искренней радости». В то же время полное описание переходной динамики на уровне априорных генерализованных паттернов требует использования 2 – 3 взаимно дополняющих компонент. В этой части результаты противоречат «дискретной» нейрокультурной теории эмоций (эмоциональная экспрессия рассматривается как универсальный коммуникационный сигнал). Также полученный результат противоречит «многомерным» теориям эмоций, в частности теории Core Affect Дж. Расселла [6], в которой валентность эмоции рассматривается как одно из фундаментальных «измерений».

Априорные главные компоненты (NRSP) имеют лучшую объяснительную способность, чем эмпирические главные компоненты, получаемые по результатам анализа эмпирических данных относительно небольшого объема.

Планируемые направления дальнейших исследований:

- дальнейшая отработка методов анализа согласованности данных
- анализ видеозаписей других эмоциональных экспрессий из базы ADFES
- уточнение психофизиологического смысла NRSP.
- отработка методов искусственной генерации видеоматериала, содержащего выделенные динамические паттерны
- отработка методов анализа главных компонент с частично зафиксированными паттернами отдельных компонент

Исследование выполнено в рамках госзадания «Познавательные процессы в экологически валидных контекстах» № 0138-2024-0012.

Список литературы

1. Ekman P. Basic Emotions. Sussex, U.K.: John Wiley & Sons. Handbook of Cognition and Emotion, 1999. P. 45-60.
2. Ekman P., Friesen W.V., Hager J.C. Facial Action Coding System. Investigator's Guide. Salt Lake City, UT: Research Nexus. 2002. 185 p.
3. Crivelli C., Fridlund A.J. Inside-out: From basic emotions theory to the behavioral ecology view // Journal of Nonverbal Behavior. 2019. Vol. 43, No 2. P. 161-194.
4. Schalk J., Hawk S., Fischer A., Doosje B. Moving Faces, Looking Places: Validation of the Amsterdam Dynamic Facial Expression Set (ADFES) // Emotion. 2011. Vol. 11, No. 4. P. 907-920.
5. Baltrušaitis T., Zadeh A., Lim Y. C., Morency Y.-P. OpenFace 2.0: Facial behavior analysis toolkit // Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FG'18). 2018. P. 59-66.
6. Russel J.A. Core Affect and the Psychological Construction of Emotion // Psychological Review. 2003. Vol. 110, No. 1. P. 145-172.