

## Система распознавания базовых эмоций на основе анализа двигательных единиц лица

© А.С. Бобе, Д.В. Коньшев, С.А. Воротников

МГТУ им. Н.Э. Баумана, Москва, 105005, Россия

*В статье описан вариант реализации системы распознавания эмоционального состояния человека для поддержки речевого общения с сервисными антропоморфными роботами. Выполнен обзор существующих подходов к распознаванию эмоций. Разработан и исследован новый алгоритм оценки базовых эмоций по П. Экману на основе расчёта 20 информативных признаков изображения лица. Степень выраженности каждой эмоции оценивается тремя независимыми классификаторами. Алгоритм реализован в среде Qt и протестирован на двух базах изображений, а также в режиме реального времени, показав качество распознавания около 85 %. Система предназначена для использования в составе приложений нейрокомпьютерного интерфейса в системах сервисной робототехники и психологической диагностики.*

**Ключевые слова:** сервисная робототехника, нейрокомпьютерный интерфейс, эмоциональное состояние, мимика, распознавание образов, машинное обучение.

**Введение.** Системы автоматического распознавания эмоционального состояния человека находят широкое применение в медицине при проведении психологической диагностики, в маркетинговых исследованиях для определения реакции клиентов на товары, в качестве канала информации в детекторах лжи и пр. Особенный интерес вызывает использование таких систем в робототехнике для решения задач поддержки речевого общения операторов с сервисными антропоморфными роботами [1, 2]. Для них, как правило, необходим естественный интерфейс управления и взаимодействия с человеком. Он реализуется, в том числе, путем использования так называемого *мимического аппарата* робота, представляющего собой комплекс программно-аппаратных средств, поддерживающих речевой диалог и имитирующих эмоциональное состояние лица.

Физической основой мимического аппарата сервисного робота является набор приводов-актуаторов, имеющих несколько устойчивых положений. Явно можно выделить минимальное, максимальное и нейтральное положения актуатора, но в общем случае этих положений может быть больше. Количество актуаторов также различается, оно составляет в различных роботах 16–30 штук.

В качестве примера рассмотрим сервисный робот «Алиса», разработанный в Зеленограде [1]. Для синтеза эмоций используют следующие группы мимических сервоприводов, которые деформируют силиконовую кожу головы:

- два непарных сервопривода путем сведения или подъема точек лица между бровями формируют эмоции «хмурость» и «жалость»;
- десять парных сервоприводов подниманием век и бровей формируют эмоции «радость», «грусть», «скепсис»;
- один сервопривод открывает рот, что используется при формировании нескольких эмоций.

При воспроизведении различных эмоциональных состояний используется несколько сервоприводов. Например, для создания эмоции «страх» задействуются актуаторы, ответственные за открывание рта, поднятие левой и правой брови.

Что касается разработок, связанных с *распознаванием эмоций*, то сегодня существует ряд готовых коммерческих решений и большое количество публикаций, посвящённых данной проблематике [3–8]. В большинстве работ реализован подход, основанный на выделении признаков изображения: *холистических* (рассчитываемых в целом по всему изображению), *локальных* (вычисленных в окрестности заданных контрольных точек), *темпоральных* (определяемых по движению частей лица во времени) и последующей классификации этих признаков с помощью методов машинного обучения. Для выделения признаков лица исследователи в основном используют так называемые *двигательные единицы лица* (Action Units, AU), входящие в Систему кодирования лицевых движений (Facial Action Coding System, FACS), разработанную П. Экманом [3]. Базовые эмоции для классификации также выбираются согласно FACS: радость, злость, грусть, отвращение, страх, удивление (иногда в список базовых включают эмоцию презрения).

В работе [4] комбинируется несколько типов признаков (оптический поток, SIFT, иерархические гауссовы) с последующей классификацией методом опорных векторов (Support Vector Machine, SVM). Авторы публикации [5] значительных результатов достигли при использовании в качестве признаков пространственно-временной модификации локальных бинарных шаблонов (LBP-TOP [6]). В материалах [7, 8] продемонстрирован алгоритм расчета интенсивности AU и сопоставление эффективности различных групп признаков и их объединений. Интересный подход к классификации предложен в [9], где *степени выраженности* AU преобразуются в маркеры наличия эмоций с помощью логических деревьев решений, специфичных для различных этнических групп.

Существенным препятствием при разработке систем распознавания эмоций является ограниченность доступных баз данных, а также высокая доля индивидуальных особенностей в проявлении той или иной эмоции у разных людей. Это несколько ограничивает возможности применения машинного обучения для классификации, по-

сколькx повышаются требования к обобщающей способности классификатора. Параметры изображения также существенно влияют на точность распознавания: изменение положения лица, наличие очков, макияжа или закрывающей брови причёски существенно затрудняют расчёт некоторых АУ, а иногда и вовсе делают его невозможным. Поэтому для успешной работы системы необходимо учитывать эти факторы. В данной работе благодаря применению локальной фильтрации изображения при расчете степени выраженности АУ, а также применению мультиклассификатора для оценки степени выраженности эмоций, удалось снизить влияние этих помех.

Авторы предлагают систему распознавания базовых эмоций на основе алгоритма расчета степени выраженности АУ ряда локальных признаков-маркеров по классификации П. Экмана. Полученные значения нормировали относительно нейтрального выражения лица того же субъекта, а затем рассчитывали степень выраженности каждой из базовых эмоций тремя классификаторами: нейронной сетью типа многослойного перцептрона (Multilayer Perceptron, MLP), вероятностным классификатором и системой логических правил.

**Расчет признаков изображения лица.** В исследуемом алгоритме рассчитываются локальные признаки лица, приведённые в табл. 1. Для этого на первом этапе проводят локализацию контрольных точек лица с использованием открытой библиотеки STASM [10].

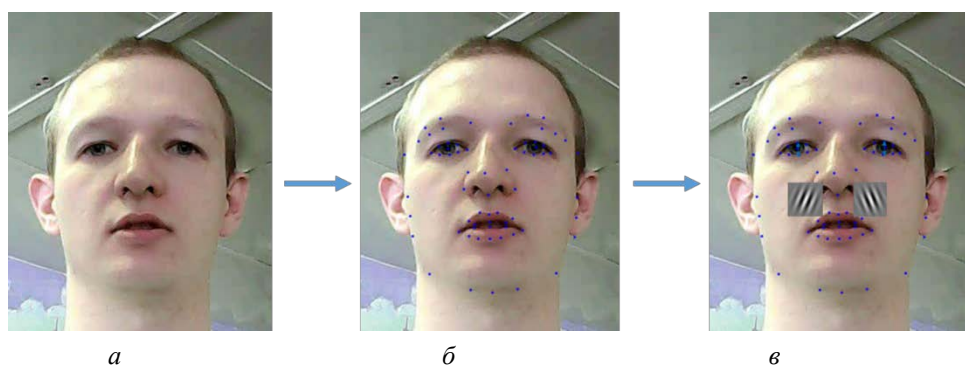
Таблица 1

**Признаки лица и связанные с ними АУ по классификации П. Экмана [3]**

Признак	Номер по П. Экману	Двусторонний признак	Примечание
1. Высота причёски над уровнем бровей	–	–	Контрольный признак
2. Морщины в области лба	1, 2, 38	–	
3. Складки между бровями	4, 9	–	
4. Ширина глазной щели	5, 6, 7, 43	+	Принимают среднее значение по двум глазам
5. Складки по сторонам носа	9, 10	–	
6. Подъём верхней губы	10	+	
7. Подъём бровей	2	+	Принимают среднее значение по двум бровям
8. Ориентация уголков рта	12–15, 17	+	
9. Радиус дуги нижнего века	6	+	Принимают среднее значение по двум глазам
10. Радиус дуги крыла носа	9	+	Принимают среднее значение по двум крыльям носа

Признак	Номер по П. Экману	Двусторонний признак	Примечание
11. Подъем внутренней части бровей	1, 9	+	Принимают среднее значение по двум бровям
12. Толщина верхней губы	9, 10, 23	+	
13. Складки в области под борода	17	–	
14. Степень открытия рта	25–27	–	
15. Складки в уголках рта	15	–	Принимают среднее значение по двум уголкам
16. Толщина нижней губы	17, 23	+	
17. Морщины под нижним веком	6, 9	–	Принимают среднее значение по двум глазам
18. Тип открытия рта	25–27	+	Видны ли зубы
19. Ориентация губ	12–15, 17	+	В случае закрытого рта
20. Форм-фактор области рта	20, 25–27	+	

Затем в окрестности контрольных точек вычисляются признаки лица: путем геометрического (для расстояний) или текстурного (для морщин) анализа изображения (рис. 1). При этом признаки разделяются на *односторонние* — те, которые характеризуются лишь степенью своего наличия (в основном это относится ко всем морщинам), и *двусторонние* — те, которые характеризуются знаком (например, брови могут быть как подняты, так и опущены). Как следует из табл. 1, набором из 20 признаков можно охарактеризовать также 20 AU по П. Экману.



**Рис. 1.** Изображение лица (а), локализация геометрических (б) и текстурных (в) признаков лица в окрестности контрольных точек

При расчете текстурных признаков областей лица можно использовать различные методы фильтрации изображений. Для определения степени выраженности мелких морщин (например, в области лба) предлагается вычислить меру среднеквадратического разброса  $S$  значений пикселей по области. Для области  $I$  размером  $m \times n$  пикселей получим

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^m \sum_{j=1}^n (I_{ij} - \mu)^2},$$

где  $\mu = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^n I_{ij}$ .

Для определения степени выраженности единичных крупных направленных морщин (например, в области носогубной складки) предлагается использовать двумерные фильтры Габора [11, 12]. Функция двумерного преобразования Габора представляет собой гармоническую функцию, сглаженную гауссовым окном. Комплексная гармоническая функция может быть представлена в виде

$$S(x, y) = \exp\left(j\left(2\pi(u_0x + v_0y) + P\right)\right),$$

где  $(u_0, v_0)$  и  $P$  — пространственные частоты и фаза гармонической функции соответственно.

Выражение для гауссова окна следующее:

$$w_r(x, y) = K \exp\left(-\pi\left(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2\right)\right),$$

где  $(x_0, y_0)$  — центр окна;  $a, b$  — параметры, характеризующие форму окна;  $K$  — амплитуда окна;  $\theta$  — угол поворота окна; индекс  $r$  означает оператор вращения:

$$(x-x_0)_r = (x-x_0)\cos\theta + (y-y_0)\sin\theta,$$

$$(y-y_0)_r = -(x-x_0)\sin\theta + (y-y_0)\cos\theta.$$

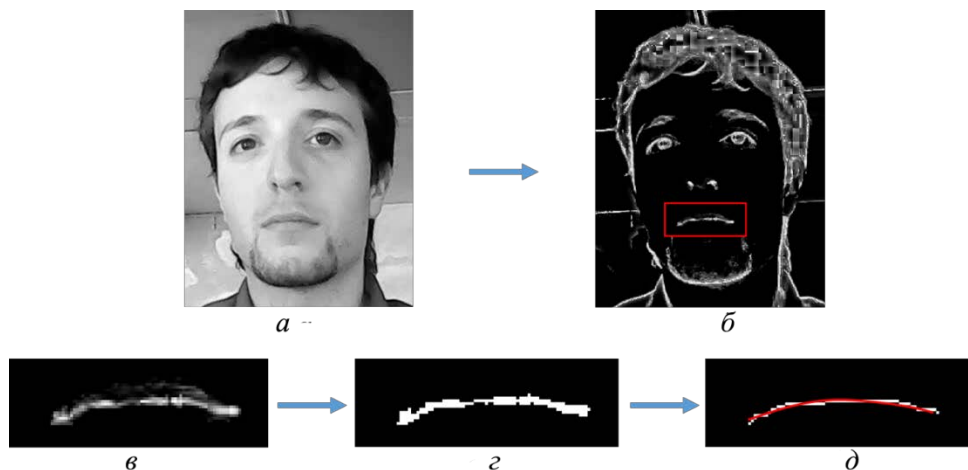
Тогда для итоговой комплексной функции Габора получим

$$g(x, y) = K \exp\left(-\pi\left(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2\right)\right) \exp\left(j\left(2\pi(u_0x + v_0y) + P\right)\right).$$

Параметры фильтра, включая пространственные частоты  $u_0, v_0$ , вычисляются исходя из предполагаемой ориентации морщин и их размеров.

Расчет геометрических признаков проводят на основе координат контрольных точек лица, но при этом также используют дополни-

тельные параметры, получаемые при фильтрации и анализе изображения. Необходимость в дополнительной информации при расчете признаков была выявлена в процессе вычислительных экспериментов и обусловлена тем, что даже небольшие погрешности в локализации контрольных точек лица существенно снижают качество работы алгоритма при расчете признаков исключительно на основе получаемой модели лица. В качестве примера на рис. 2 показано определение ориентации губ с помощью ранговой фильтрации изображения области рта.



**Рис. 2.** Принцип определения ориентации губ:

*a* — исходное изображение; *б* — результат ранговой фильтрации; *в* — изображение области рта; *г* — результат после бинаризации и удаления мелких фрагментов; *д* — контур ориентации и аппроксимирующая кривая

Фильтрация проводится следующим образом. Окно фильтра (в примере на рис. 2 —  $13 \times 13$  пикселей) проходит по всем точкам изображения. Для каждого положения окна выполняется бинаризация соответствующей области изображения по порогу, равному значению интенсивности центрального пикселя данной области, умноженному на некоторый коэффициент  $t$  (в данном примере  $t = 1,3$ ). Число пикселей со значениями интенсивности, превышающими пороговое, записывается в качестве значения пикселя выходного изображения:

$$Res_{i,j} = \sum_{k=i-hw}^{k=i+hw} \sum_{l=j-hw}^{l=j+hw} (I_{kl} > t I_{i,j}).$$

Результат такой фильтрации показан на рис. 2, *б*. Остальные этапы приведены на рис. 2, *в–д*.

Отметим, что контрольный признак высоты прически над уровнем бровей (табл. 1) рассчитывается в первую очередь и служит для

определения пригодности изображения для анализа в целом. В случае перекрытия области бровей сразу несколько важных АУ в этой области становятся недоступными, что существенно снижает качество работы системы. В наших экспериментах такие изображения отбраковывались и не учитывались в общей статистике.

**Нормировка признаков.** При разработке системы распознавания проводилась нормировка признаков двух видов: относительно опорного кадра и по диапазону значений.

*Нормировка относительно опорного кадра* используется для преобразования абсолютного значения признака в относительное. Оно характеризуется отношением абсолютного значения к значению, соответствующему изображению нейтрального выражения лица для одного и того же субъекта. На этапе обучения алгоритма за признаки опорного кадра принимаются усредненные значения признаков для нескольких имеющихся в базе изображений нейтрального выражения лица субъекта.

*Нормировку по диапазону значений* выполняют для приведения признаков к единой шкале. Нормировочная функция для значения признака  $x$  имеет вид

$$x_{normed} = \begin{cases} \operatorname{tg}^{-1} \left( \frac{x - \mu}{c - \mu} \right) \frac{2}{\pi} - 1, & \text{if } x \leq \mu, \\ - \left( \frac{x - c}{c - low} \right)^2, & \text{if } low < x < c, \\ - \left( \frac{x - c}{high - c} \right)^2, & \text{if } c < x < high, \\ \operatorname{tg}^{-1} \left( \frac{x - high}{high - center} \right) \frac{2}{\pi} - 1, & \text{if } x \geq high. \end{cases}$$

В данной работе значения параметров  $low$ ,  $c$ ,  $high$  для односторонних признаков определяются по статистическому анализу изображений обучающей выборки:

- $low = \mu = \frac{1}{N} \sum_{i=1}^n x_i$  — среднее значение данного признака по всем изображениям базы лиц;

- $c = \mu + \sigma = \mu + \sqrt{\frac{1}{N-1} \sum_{i=1}^n (x_i - \mu)^2}$  — среднее значение признака плюс его стандартное отклонение;

•  $high = \mu + 2\sigma = \mu + 2\sqrt{\frac{1}{N-1} \sum_{i=1}^n (x_i - \mu)^2}$  — среднее значение признака плюс его двойное стандартное отклонение.

Для двусторонних признаков используются две функции отдельно: для отрицательной выраженности и для положительной. Так как распределение значений таких признаков в общем случае несимметрично, нормированные значения рассчитываются отдельно для положительной и отрицательной областей (рис. 3).

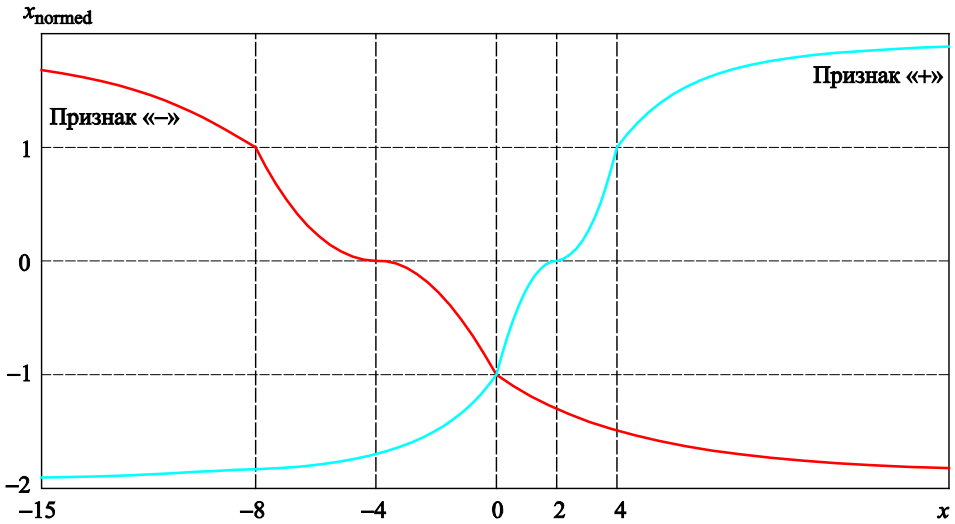


Рис. 3. Нормировочные функции для двустороннего признака  $x_{normed} = f(x)$  с параметрами  $\mu = 0$ ,  $\sigma_+ = 2$ ,  $\sigma_- = 4$

Из рис. 3 видно, что нормировка обеспечивает плавный отклик в области средних значений признака, когда степень его выраженности сомнительна (для значений с малым отклонением от среднего), и более резкий отклик в случае статистически значимых отклонений от среднего. Экстремальные значения признака с существенными отклонениями от среднего также сглаживаются данной нормировочной функцией, что уменьшает влияние артефактов на работу алгоритма.

**Классификация базовых эмоций.** В рассматриваемой системе распознавания используются классификаторы трех типов: вероятностный, нейронная сеть и набор логических правил.

В *вероятностном классификаторе* степень выраженности эмоции определяется как взвешенная сумма нормированных значений всех  $n$  признаков:



$$E_i = \frac{\sum_{k=1}^n w_k x_n}{\sum_{k=1}^n |w_k|}$$

Вес  $w_k$  каждого признака выбирается на основе значимости того или иного признака для данной эмоции (как в работе [13]). Признакам, не представляющим важность для данной эмоции, присваивается нулевой вес. Для двусторонних признаков значения для положительной и отрицательной выраженности интерпретируются как разные признаки (например, для одной эмоции требуется положительная выраженность признака, т. е. коэффициент, равный +1, для  $x_{n+}$ , а для другой — отсутствие отрицательной выраженности, т. е. коэффициент, равный -1, для  $x_{n-}$ ). Нейтральное выражение лица соответствует минимальной выраженности всех признаков.

Нейронная сеть типа многослойного перцептрона (MLP) также обучается на нормированных значениях всех признаков. Сеть имеет 20 входных нейронов, соответствующих количеству признаков, 12 нейронов в скрытом слое и 7 выходных нейронов, по числу распознаваемых эмоций.

Итоговая степень выраженности эмоции рассчитывается как сумма откликов двух классификаторов.

Набор логических правил представляет собой ряд различных вариантов реализаций эмоций, составленный на основе рекомендаций П. Экмана. Для каждого признака экспериментально выбирают минимальный порог — положительной либо отрицательной выраженности (обычно около  $0,3\sigma$  от среднего в ту или иную сторону), а затем составляют логическое правило. Например, один из вариантов эмоции «радость» можно представить в виде (см. табл. 1)

$$v = x_{5+} \ \&!(x_{2+}|x_{3+}) \ \&x_{14+} \ \&x_{15+} \ \&x_{18+} \ \&x_{4-},$$

где & — логический оператор «и»; | — логический оператор «или»; ! — логический оператор «не».

В этом варианте эмоция «радость» задается следующим образом: наличие складок по сторонам носа, рот приоткрыт и видны зубы, по уголкам рта есть складки, глаза сужены (признак 4 из табл. 1 со знаком «-»), нет морщин в области лба или складок между бровями.

Эксперименты показали, что данный классификатор сам по себе не обеспечивает достаточное качество работы системы, однако он обладает свойством независимости от абсолютных значений признаков, поэтому его использование целесообразно. С его помощью разрешаются спорные ситуации, когда несколько эмоций получают схожую высокую оценку степени выраженности. Эмоция, обладающая наибольшим соответствием набору логических правил, объявля-

ется доминирующей. Отклики «конкурирующих» эмоций уменьшаются до значения ниже порогового.

Итоговый классификатор — *мультиклассификатор* — строится как функция результатов работы нейросетевого и вероятностного классификаторов, с поправкой на основе применения логических правил (рис. 4). Поскольку значения на выходе обоих классификаторов укладываются в диапазон  $0...1$ , мультиклассификатор можно представить как простую сумму их откликов.

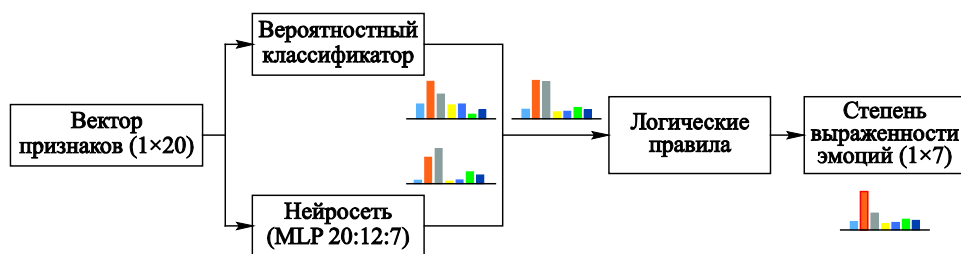


Рис. 4. Схема мультиклассификатора

**Экспериментальные исследования.** Разработанная система распознавания базовых эмоций реализована на языке С++ и протестирована на двух базах данных: Cohn-Kanade [14, 15] и ООО «Нейроботикс». База Cohn-Kanade является одной из наиболее популярных для исследования алгоритмов распознавания эмоций. В нее входит 123 субъекта, для каждого из которых даны серии изображений с переходом от нейтральной к одной из нескольких эмоций. База ООО «Нейроботикс» содержит 858 изображений 9 субъектов с метками различных эмоций. При тестировании половину пар изображений «нейтральное – соответствующая эмоция» включали в обучающую выборку, вторую половину – в тестовую. В эксперименте изображения нормировали по координатам центров глаз так, чтобы межзрачковое расстояние составляло 110 пикселей, а глаза находились на одном горизонтальном уровне. Степень выраженности каждой эмоции определяли в диапазоне  $[0...1]$ , при этом порог, при превышении которого эмоция считалась достаточно выраженной, принимали равным 0,3. При расчете вероятностных характеристик считалась распознанной та эмоция, для которой был получен максимальный отклик классификатора. Результаты представлены в табл. 2 и 4. В табл. 3 и 5 представлены усредненные значения откликов мультиклассификатора для каждой из эмоций (в диапазоне  $[0...1]$ ). Каждый столбец таблицы содержит данные по распознаваемой базовой эмоции, в строках — данные, полученные по всем эмоциям в результате работы алгоритма (рассматривать данные следует по столбцам).

Таблица 2

## Результаты распознавания эмоций, база Cohn-Kanade, %

Эмоция заданная	Эмоция при распознавании						
	Нейтральная	Радость	Удивление	Страх	Злость	Отвращение	Грусть
Нейтральная	82,3 91,1 92,5*	0,0 1,4 0,0	2,5 1,2 1,5	2,6 5,4 2,3	3,0 6,8 2,7	4,5 3,3 1,7	17,9 17,1 14,3
Радость	1,6 0,6 0,3	92,8 94,2 95,7*	0,0 0,0 0,0	2,3 1,3 0,4	4,5 2,2 0,0	7,9 15,3 8,6	4,3 0,7 0,0
Удивление	0,0 0,0 0,0	0,0 1,4 0,0	62,3 82,6 77,1*	25,0 18,9 20,2	0,0 0,0 0,0	0,0 1,7 0,0	3,6 3,6 0,0
Страх	0,0 1,2 0,0	0,0 2,9 0,0	35,1 14,8 20,4	59,3 65,7 71,6*	0,0 3,5 2,3	1,4 1,7 0,0	0,0 2,3 0,0
Злость	2,1 1,8 0,3	0,0 0,0 0,0	0,0 1,3 0,0	4,2 2,1 2,7	72,1 68,5 81,6*	12,6 9,4 8,6	23,0 14,3 10,7
Отвращение	0,0 1,8 0,6	7,2 0,0 4,3	0,0 0,0 0,0	6,6 5,3 2,8	13,6 9,8 8,2	66,8 66,9 77,6*	2,0 3,6 3,6
Грусть	14,1 3,4 6,3	0,0 0,0 0,0	0,0 0,0 0,0	0,0 1,3 0,0	6,8 9,1 5,2	6,8 1,7 3,4	49,3 58,4 71,4*

\* Первое значение — вероятностный классификатор; второе значение – нейросетевой классификатор; третье значение — мультиклассификатор.

Таблица 3

## Средние значения на выходе мультиклассификатора, база Cohn-Kanade

Эмоция, определенная алгоритмом	Распознаваемая эмоция						
	Нейтральная	Радость	Удивление	Страх	Злость	Отвращение	Грусть
Нейтральная	0,55	0,08	0,12	0,17	0,10	0,11	0,22
Радость	0,05	0,92	0,08	0,17	0,10	0,18	0,08
Удивление	0,09	0,14	0,83	0,22	0,01	0,06	0,16
Страх	0,02	0,14	0,42	0,56	0,05	0,07	0,13
Злость	0,11	0,10	0,04	0,09	0,72	0,32	0,30
Отвращение	0,04	0,24	0,08	0,18	0,26	0,65	0,12
Грусть	0,17	0,01	0,18	0,23	0,19	0,17	0,63

Таблица 4

## Результаты распознавания эмоций, база ООО «Нейроботикс», %

Эмоция заданная	Эмоция при распознавании						
	Нейтральная	Радость	Удивление	Страх	Злость	Отвращение	Грусть
Нейтральная	88,2 90,9 92,2*	5,2 0,2 1,0	2,0 0,0 0,0	12,6 2,8 3,9	12,6 3,5 7,8	1,2 0,8 0,8	24,3 6,4 9,0
Радость	0,0 0,0 0,0	76,1 91,2 96,0*	27,0 0,0 5,3	1,9 0,1 0,0	6,5 0,0 0,9	7,1 3,4 4,3	7,1 10,9 8,5
Удивление	4,3 0,0 0,0	0,5 0,6 1,3	54,2 70,8 79,0*	11,6 11,8 11,9	0,8 2,0 1,9	0,8 0,8 0,0	0,0 1,4 1,1
Страх	0,0 4,7 3,6	0,0 0,0 0,0	8,3 29,2 15,0	55,8 70,2 75,6*	3,7 3,7 2,8	7,1 9,3 5,2	1,4 1,4 1,6
Злость	6,5 3,5 3,6	9,4 4,7 1,0	0,0 0,0 0,0	2,8 1,9 1,9	45,1 60,6 64,4*	8,4 8,6 6,1	4,3 4,3 3,2
Отвращение	0,2 0,9 0,6	3,6 1,7 0,3	0,0 0,0 0,0	9,0 12,3 6,7	18,7 25,7 19,4	66,9 84,7 80,2*	0,0 2,9 1,6
Грусть	0,1 0,0 0,0	5,2 1,6 0,4	8,3 0,0 0,7	6,0 0,0 0,0	12,5 4,6 2,7	8,3 0,0 3,3	62,9 72,7 75,1*

\* Первое значение — вероятностный классификатор; второе значение – нейросетевой классификатор; третье значение — мультиклассификатор.

**Средние значения на выходе мультиклассификатора,  
база ООО «Нейроботикс»**

Среднее значение, определенное алгоритмом	Распознаваемая эмоция						
	Нейтральная	Радость	Удивление	Страх	Злость	Отвращение	Грусть
Нейтральная	0,52	0,14	0,12	0,15	0,16	0,13	0,16
Радость	0,06	0,83	0,21	0,07	0,13	0,21	0,19
Удивление	0,09	0,15	0,70	0,26	0,10	0,07	0,16
Страх	0,07	0,11	0,36	0,62	0,14	0,11	0,09
Злость	0,16	0,08	0,06	0,13	0,58	0,25	0,15
Отвращение	0,05	0,21	0,05	0,15	0,31	0,72	0,13
Грусть	0,15	0,07	0,12	0,14	0,18	0,19	0,62

Из приведенных данных видно, что процент верного распознавания составил в среднем 81,1 % для базы Cohn-Kanade и 80,4 % для базы ООО «Нейроботикс». Для обеих баз наибольшая точность распознавания получена по эмоциям «радость», поскольку АУ, связанные с этой эмоцией, проявляются наиболее выразительно. Эмоция «грусть» в среднем распознается хуже остальных, поскольку соответствующие АУ, напротив, выражены слабо, а многие из них неспецифичны (могут быть отнесены к проявлению других эмоций).

В табл. 3 и 5 желтым цветом выделены эмоции, которые выражены в достаточно большой степени (больше 0,25) вместе с определяемой базовой. По обеим базам эмоции «страх» и «удивление», а также «отвращение» и «злость» зачастую проявлялись вместе, что объясняется большим числом смешанных вариантов данных пар эмоций. Также по базе Cohn-Kanade совместно проявлялись эмоции «грусть» и «злость». Мультиклассификатор в большинстве случаев показал лучшие характеристики, чем каждый из классификаторов в отдельности.

**Заключение.** В основе рассмотренной системы визуального распознавания базовых эмоций человека лежит алгоритм расчета степени выраженности признаков лица, входящих в состав АУ по П. Экману. Признаки нормируются относительно опорного (нейтрального) кадра для каждого субъекта и приводились к общей числовой шкале на основе статистических данных. Для классификации признаков использован мультиклассификатор, состоящий из вероятностного классификатора, нейронной сети и набора логических правил. Система показала хорошие результаты при экспериментах на двух базах лиц. На основе разработанного алгоритма создан программный комплекс Neurobotics EmoDetect, способный оценивать эмоциональное состояние человека в реальном времени по видеопотоку. Комплекс использовался при разработке человеко-машинного интерфейса сервисного робота «Алиса». В будущем планируется усовершенствовать методы расчета признаков лица с целью повышения устойчивости работы

системы распознавания, а также добавить новые признаки лица для более полного соответствия алгоритма системе FACS П. Экмана.

Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, грант №16-07-01080.

## ЛИТЕРАТУРА

- [1] Конышев Д.В., Воротников С.А., Выборнов Н.А. Управление мимическим аппаратом сервисных роботов при синтезе эмоций. *Прикаспийский журнал «Управление и высокие технологии»*, 2014, № 3, с. 216–229.
- [2] Breazeal P. Robot Emotion: A functional perspective. Who Needs Emotions: The Brain Meets the Robot. *MIT Press*, 2004, pp.137–168.
- [3] Ekman P., Friesen W.V., Hager J.C. *Facial Action Coding System. The Manual*. Research Nexus division of Network Information Research Corporation, 2002.
- [4] Tariq U., Lin K., Li Z., Zhou Z., Wang Z., Le V., Huang T.S., Lv X., Han T.X. Emotion Recognition from an Ensemble of Features. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, 2012, vol. 42 (4), pp. 1017–1026.
- [5] Zhao G., Pietikäinen M. *Experiments With Facial Expression Recognition Using Spatiotemporal Local Binary Patterns*. University of Oulu, Finland, 2007.
- [6] Zhao G., Pietikäinen M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2007, vol. 29, pp. 915–928.
- [7] Bartlett M.S., Hager J.C, Ekman P., Sejnowskie T.J. *Measuring facial expressions by computer image analysis*. Cambridge University Press. Publ., 1999, pp. 253–263.
- [8] Chandrani S., Washef A., Soma M., Debasis M. Facial Expressions: A Cross-Cultural Study. *Emotion Recognition: A Pattern Analysis Approach*. Wiley, 2015, pp. 69–86.
- [9] Заболеева-Зотова А.В. Развитие системы автоматизированного определения эмоций и возможные сферы применения. *Открытое образование*, 2011, № 2, с. 59–62.
- [10] Milborrow S., Nicolls F. *Active Shape Models with SIFT Descriptors and MARS*. *VISAPP(2)*, 2014, pp. 380–387.
- [11] Gabor D. Theory of communication. Part 1: The analysis of information. *Journal of the IEE*, 1946, vol. 93, no. 26, pp. 429–441.
- [12] Бабий М.С, Чекалов А.П. Распознавание изображений на основе двумерного вейвлет-анализа. *Вісник Сумського державного університету. Серія Технічні науки*, 2012, № 1, с. 20–24.
- [13] Князев Б.А., Гапанюк Ю.Е. Распознавание аномального поведения человека по его эмоциональному состоянию и уровню напряженности с использованием экспертных правил. *Инженерный Вестник*, 2013, с. 509–524. URL: <http://engbul.bmstu.ru/doc/568250.html> (дата обращения 27.02.2015).
- [14] Kanade T., Cohn J. F., Tian Y. Comprehensive database for facial expression analysis. *Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, Grenoble, France*, 2000, pp. 46–53.
- [15] Lucey P., Cohn J. F., Kanade T., Saragih J., Ambadar Z., Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, IEEE*, 2010, pp. 94–101.

Статья поступила в редакцию 03.06.2016

Ссылку на эту статью просим оформлять следующим образом:

Бобе А.С., Коньшев Д.В., Воротников С.А. Система распознавания базовых эмоций на основе анализа двигательных единиц лица. *Инженерный журнал: наука и инновации*, 2016, вып. 9.

<http://dx.doi.org/10.18698/2308-6033-2016-09-1530>

**Бобе Анатолий Сергеевич** окончил МГТУ им. Н.Э. Баумана в 2013 г. Автор 10 научных работ в области разработки систем компьютерного зрения и распознавания образов.

**Коньшев Дмитрий Владимирович** окончил МГТУ им. Н.Э. Баумана в 2011 г. Аспирант кафедры «Робототехнические системы и мехатроника» МГТУ им. Н.Э. Баумана. Автор пяти научных работ в области разработки алгоритмов технического зрения роботов.

**Воротников Сергей Анатольевич** окончил МВТУ им. Н.Э. Баумана в 1978 г. Доцент кафедры «Робототехнические системы и мехатроника» МГТУ им. Н.Э. Баумана. Автор более 100 научных работ в области робототехники. e-mail: vorotn@bmstu.ru

## Emotion recognition system based on the facial motor units' analysis

© A.S. Bobe, D.V. Konyshev, S.A. Vorotnikov

Bauman Moscow State Technical University, Moscow, 105005, Russia

*The article describes the human emotion recognition system embodiment to support verbal communication with service anthropomorphic robots and considers existing emotion recognition approaches. We investigated a new algorithm for P. Ekman's estimation of main emotions based on 20 informative facial image features evaluation. Three independent classifiers calculated each emotion intensity. The algorithm is implemented in Qt medium and tested on two image databases, as well as in real time, showing average recognition rate of about 85 %. The system can be used in neurocomputer interface applications in robotics and psychological diagnosis systems.*

**Keywords:** service robotics, brain-computer interface, emotional state, mimics, pattern recognition, machine learning.

### REFERENCES

- [1] Konyshev D.V., Vorotnikov S.A., Vybornov N.A. *Prikaspiyskiy Zhurnal: Upravlenie i Vysokie Tekhnologii — Caspian Journal: Control and High Technologies*, 2014, no. 3, pp. 216–229.
- [2] Breazeal P. *Robot Emotion: a functional perspective. Who Needs Emotions: The Brain Meets the Robot*. The MIT Press Publ., 2004, pp. 137–168.
- [3] Ekman P., Friesen W.V., Hager J.C. *Facial Action Coding System. The Manual*. Research Nexus division of Network Information Research Corporation Publ., 2002.
- [4] Tariq U., Lin K., Li Z., Zhou Z, Wang Z., Le V., Huang T.S., Lv X., Han T.X., Emotion Recognition from an Ensemble of Features. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions*, 2012, vol. 42, no. 4, pp. 1017–1026.
- [5] Zhao G., Pietikäinen M. *Experiments with Facial Expression Recognition Using Spatiotemporal Local Binary Patterns*. University of Oulu Publ., Finland, 2007.
- [6] Zhao G., Pietikäinen M. Dynamic Texture Recognition Using Local Binary Patterns with an Application to Facial Expressions. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 2007, vol. 29, pp. 915–928.
- [7] Bartlett M.S., Hager J.C, Ekman P., Seynowskie T.J. *Measuring facial expressions by computer image analysis*. Cambridge University Press Publ., 1999, pp. 253–263.
- [8] Chandrani S., Washef A., Soma M., Debasis M. *Facial Expressions: A Cross-Cultural Study. Emotion Recognition: A Pattern Analysis Approach*. Wiley Publ., 2015, pp. 69–86.
- [9] Zaboleeva-Zotova A.V. *Otkrytoe obrazovanie — Open Education*, 2011, no. 2, pp. 59–62.
- [10] Milborrow S., Nicolls F. *Active Shape Models with SIFT Descriptors and MARS*, VISAPP (2) Publ., 2014, pp. 380–387.
- [11] Gabor D. Theory of Communication. Part 1: The analysis of information. *Journal of the IEE*, 1946, vol. 93, no. 26, pp. 429–441.

- [12] Babiy M.S., Chekalov A.P. *Vestnik Sumskogo Gosudarstvennogo Universiteta. Seriya Tekhnicheskie nauki — Bulletin of Sumy State University. Series Technical Sciences*, 2012, no. 1, pp. 20–24.
- [13] Knyazev B.A., Gapanyuk Yu.E. *Inzhenerny vestnik — Engineering Bulletin*, 2013, pp. 509–524. Available at: <http://engbul.bmstu.ru/doc/568250.html> (accessed 27 February, 2015).
- [14] Kanade T., Cohn J. F., Tian Y. Comprehensive database for facial expression analysis. *Proc. of the Fourth IEEE International Conference Automatic Face and Gesture Recognition. Grenoble, France, 2000*, pp. 46–53.
- [15] Lucey P., Cohn J.F., Kanade T., Saragih J., Ambadar Z., Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete expression dataset for action unit and emotion-specified expression. *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, 2010, pp. 94–101.

**Bobe A.S.** graduated from Bauman Moscow State Technical University in 2013. Author of 10 research publications in the field of computer vision systems and pattern recognition.

**Konyshev D.V.** graduated from Bauman Moscow State Technical University in 2011. Post-graduate student, Department of Robotic Systems and Mechatronics, Bauman Moscow State Technical University. Author of 5 research publications in the field of robotic vision algorithm development.

**Vorotnikov S.A.** graduated from Bauman Moscow State Technical University in 1978. Assoc. Professor, Department of Robotic Systems and Mechatronics, BMSTU. Author of over 100 research publications in the field of robotics. e-mail: vorotn@bmstu.ru