

АДАПТАЦИЯ СИСТЕМЫ КОДИРОВАНИЯ ДВИЖЕНИЙ ЛИЦЕВЫХ МЫШЦ К РАБОТЕ В НЕЙРОСЕТЕВОЙ МОДЕЛИ

Тюрин А.И.¹, Безыкорнов Д.С.¹

¹ГОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексева», Нижний Новгород, Россия (603950, ГСП-41, Н. Новгород, ул. Минина, д.24), e-mail: art-tyurin@mail.ru

Цель исследования: создание и тестирование системы метрик лицевой информации. **Методология:** исследование системы кодирования лицевых движений, сравнительное тестирование на созданном наборе лиц с эмоциями. **Результаты:** построение системы метрик, ее комплексное тестирование с учетом значения «уверенности», снижение общей ошибки распознавания на 20%; выявление эмоций, трудных для распознавания. **Построенная система** распознает эмоции страха и печали на 37–47% эффективнее. **Последующая работа:** повышение качества распознавания групп похожих эмоций. **Оригинальность и ценность:** полученная система метрик совмещает достижения современной психологии с возможностями существующих информационных технологий. **Построенная на ее основе система** распознавания эмоций позволяет получить высокую точность распознавания эмоций.

Ключевые слова: система кодирования лицевых движений, распознавание эмоций, нейронные сети

FACIAL ACTION CODING SYSTEM ADAPTATION TO WORK WITH NEURAL NETWORKS

Tyurin A.I.¹, Bezykornov D.S.¹

¹Nizhny Novgorod State Technical University n.a. R.E. Alekseev, Nizhniy Novgorod, Russia (603950, Nizhniy Novgorod, street Minina, 24), e-mail: art-tyurin@mail.ru

Purpose: Creating and testing face-metric information system based on previous work. **Design/methodology/approach:** Researching of Facial Action Coding System (FACS), testing selected set of metrics on created face images set. **Findings:** We have built metrics system, conducted comprehensive study with «confidence» value, decreased overall recognition error on 20%, found emotions hard to recognize. **Resulting system** provides better results than reference system. It can recognize fear and sadness about 37-47% more emotions right. **Research limitations/implications:** Further work may be done in area of increasing overall recognition quality of similar emotions. **Originality/value:** Resulting system uses last psychology researches and drafts in IT technology. **Recognition system** was been built allows us to get high quality of emotions recognition.

Keywords: Facial Action Coding System (FACS), emotion recognition, neural networks

Задача автоматического распознавания эмоций является одним из наиболее перспективных направлений развития компьютерного зрения. Обработка эмоций и различных сигналов «языка тела» происходит на уровне подсознания и без дополнительной осмысленной тренировки находится на невысоком уровне развития.

Помимо определенных выражений лица, человеческому мозгу доступен комплексный анализ – различные жесты, интонации голоса, изменение направления взгляда. При получении видеоизображения в информационной системе дополнительные данные недоступны, что приводит к повышению требований к качеству распознавания эмоций по лицу.

Одной из ключевых задач системы распознаваний эмоций является выбор метрик лица, которые возможно получить с учетом существующих алгоритмов и общего развития информационных технологий. В сфере исследования кодирования лицевых эмоций

существуют полноценные руководства, однако они не могут использоваться в автоматических системах распознавания без адаптации к ограничениям алгоритмов.

Данная работа предоставляет полноценную систему кодирования, основанную на классической схеме кодирования движений лицевых мышц, и методику комплексного анализа результатов распознавания эмоций автоматической системой.

Теоретическая часть

Для адаптации к использованию в нейросетевой модели была выбрана Система кодирования лицевых движений (*сокр.* СКЛид), разработанная Полом Экманом и Уоллесом Фризенном в 1978 г. [2]

6 эмоций были выбраны для распознавания в качестве базовых:

- радость;
- удивление;
- гнев;
- отвращение;
- печаль;
- страх.

Данный набор позволяет охватить наибольшее число реальных выражений лиц, так как другие эмоции являются естественным и более тонким развитием базовых.

Основой СКЛид являются двигательные единицы (*сокр.* ДЕ), которые позволяют закодировать любое анатомически возможное выражение лица. В проявлении эмоций участвуют не все мышцы, а лишь их небольшая часть. Основу каждой из базовых эмоций составляют 5–7 ДЕ, которые позволяют закодировать наиболее часто встречающиеся варианты выражения в так называемые прототипы:

$$1 + 2 + 5B + 26 \quad (1)$$

Формула (1) является прототипом эмоции удивление. ДЕ выражены числами. Латинские буквы от А до Е, стоящие после чисел, определяют интенсивность движения от минимальной до максимальной соответственно. Нейросетевая модель проходит обучение на большом количестве изображений людей. Человеческое лицо обладает большой вариативностью, поэтому слабо различимые (код А) и незначительные (код В) изменения лицевых мышц попадают в допуск возможных значений и являются труднорегируемые с достаточной долей достоверности. Поэтому для создания тестового набора нейронной сети были выбраны изображения эмоций со средней и выше степенью интенсивности.

Алгоритмы поиска ключевых лицевых точек могут выделять до 80 значащих областей на лице. Как правило, это границы глаз, рта и бровей. В представлении данных алгоритмов ДЕ основных прототипов эмоций были поделены на 3 группы:

- 1) статические ДЕ – возможно определение по фотографии;
- 2) динамические ДЕ – необходимо отслеживать изменение положения ключевых точек на лице или иметь среднее значение расстояния по данной ДЕ;

3) пустые ДЕ – активно участвуют в проявлении эмоций, однако не регистрируются алгоритмами поиска (например, ямочки на щеках).

Составляющие групп приведены в следующей таблице:

Таблица 1

Соответствие ДЕ группам определения алгоритмов поиска

	Статика	Динамика	Пустые
Радость	12	6	
Удивление	1,25	2,5,26	27
Гнев	4,25	5,23,24,26	7,10,22
Отвращение	15	9,26,10	16,17
Печаль	1,4,15		11
Страх	1	2,5,20,38	

Подробную расшифровку кодов ДЕ и основные прототипы эмоций можно увидеть в [3].

Для использования динамической группы ДЕ в нейронной сети к базовому набору была добавлена нейтральная эмоция без прототипа. Данное решение позволяет находить динамические ДЕ без отслеживания изменения состояния лица во времени.

После анализа соответствия выходных данных алгоритмов поиска ключевых лицевых точек и основных прототипов базовых эмоций была разработана система из 10 метрик.

1. Высота рта: верх верхней губы – низ нижней губы.
2. Высота открытого рта: низ верхней губы – верх нижней губы.
3. Уголок губ вниз: уголок рта – верх нижней губы.
4. Уголок губ вверх: уголок рта – верх верхней губы.
5. Ширина рта: левый уголок рта – правый уголок рта.
6. Высота подбородка: низ нижней губы – подбородок.
7. Ширина глаза: верх глаза – низ глаза.
8. Высота брови: верхний центр брови – середина глаз.
9. Внутренний уголок брови: внутренний угол брови – внутренний угол глаза.
10. Внешний уголок брови: внешний угол брови – внешний угол глаза.

Метрики под номерами 3, 4, 7–10 являются несимметричными и рассчитываются для левой и правой половины отдельно. Таким образом, на вход нейронной сети подается 16 значений. Нормирование данных показателей осуществляется по межзрачковому расстоянию. Оно минимально различается от человека к человеку и позволяет достаточно хорошо стандартизировать обучающую выборку.

Наглядно данные метрики представлены на следующем рисунке:

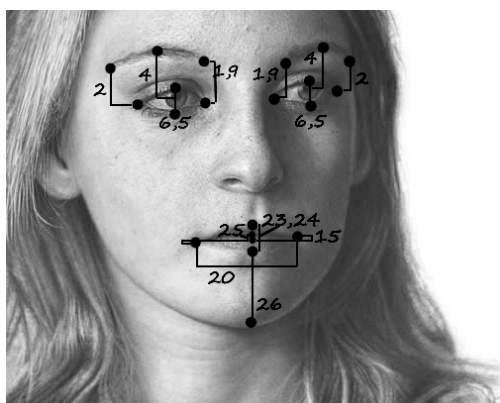


Рис. 1. ДЕ в адаптированной системе метрик

Как видно из рисунка 1, различные ДЕ перекрываются и для системы распознавания рассчитываются как одно расстояние. Данное представление обусловлено техническими особенностями и возможностями алгоритмов поиска ключевых точек, но тем не менее позволяет сохранить высокий процент распознавания.

Основные прототипы базовых эмоций зачастую имеют совпадающие ДЕ как по номеру, так и по их значению. Общая классификация с помощью нейронных сетей позволяет указать эмоцию с определенным процентом вероятности. Как правило, основная эмоция имеет намного более высокий процент (около 80–90%), чем подобная, но в некоторых случаях различие в индивидуальных особенностях и интенсивности проявления эмоции уменьшает разницу в показателях на выходах нейронной сети.

Для более полной оценки качества распознавания вводится значение «уверенности» распознавания:

$$C_{conf} = \frac{V_{em} - V_{max}}{N} \quad (2)$$

,где V_{em} – значение верно распознанной эмоции, V_{max} – следующее максимальное значение на данном изображении, N — количество верно распознанных эмоций определенного типа из общего количества этих эмоций (например, 70% верно распознанных эмоций из набора в 100 изображений с эмоцией «радость»).

Использовалось 4 диапазона значений метрики «уверенности»:

- 1) абсолютная: $C_{conf} > 25\%$;
- 2) твердая: $15\% < C_{conf} < 25\%$;
- 3) частичная: $5\% < C_{conf} < 15\%$;
- 4) самоуверенность: $C_{conf} < 5\%$.

Данный подход позволил комплексно протестировать созданную схему метрик и всю автоматическую систему распознавания эмоций в целом.

Экспериментальная часть

Для классификации эмоций была обучена нейронная сеть со следующими параметрами:

- Тип: многослойный персептрон
- Вход: 16 нейронов
- Выход: 7 нейронов
- Скрытый слой: 30 нейронов
- Количество эпох обучения: 100 000

Набор изображений формировался вручную из открытой базы Cohn-Kanade [4]. По каждой из 7 эмоций было отобрано 35 изображений. Для усложнения задачи распознавания общее количество людей было поднято до 50. Таким образом, повысилась вариативность лицевых черт на выборке.

Разметка ключевых лицевых точек происходила в автоматическом режиме с помощью алгоритма активных моделей внешнего вида [1]. Пример тестовых изображений с размеченными ключевыми точками представлен на рисунке 2.



Рис. 2. Пример распознавания ключевых точек: эмоция «радость» слева, «гнев» — справа

Для сравнения была выбрана описанная в [5] система кодирования. Обучение происходило на том же наборе данных, параметры нейронной сети не изменялись. Эмоция считалась распознанной, если на выходе нейронной сети, соответствующем данной эмоции, находилось наибольшее значение.

Диаграмма распознавания эмоций представлена на рисунке 3.

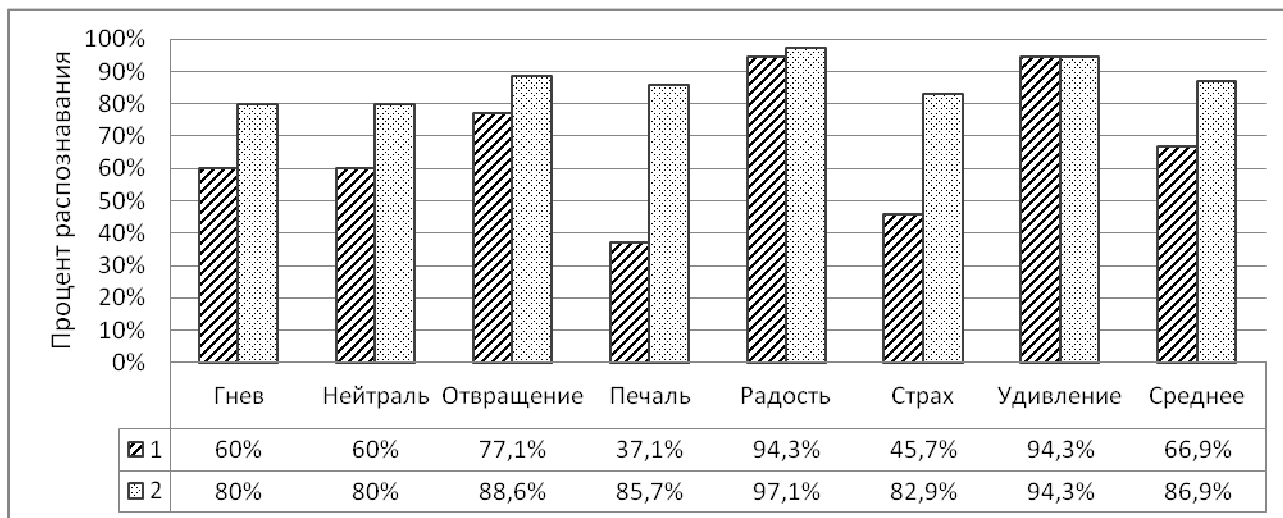


Рис. 3. Сводная таблица качества распознавания эмоций: 1 – алгоритм для сравнения, 2 – построенная система метрик

Как видно из статистических данных, использование системы метрик, основанной на СКЛиД, позволяет уменьшить общую ошибку распознавания на 20%. Также данная система показывает большую эффективность в распознавании наиболее трудных для классификации эмоций – страха и печали. Ошибка в данных эмоциях снижается на 37,2% и 48,6% соответственно.

При анализе построенной системы особое внимание уделялось значению C_{conf} . Сравнение алгоритмов можно произвести по рисункам 4 и 5:

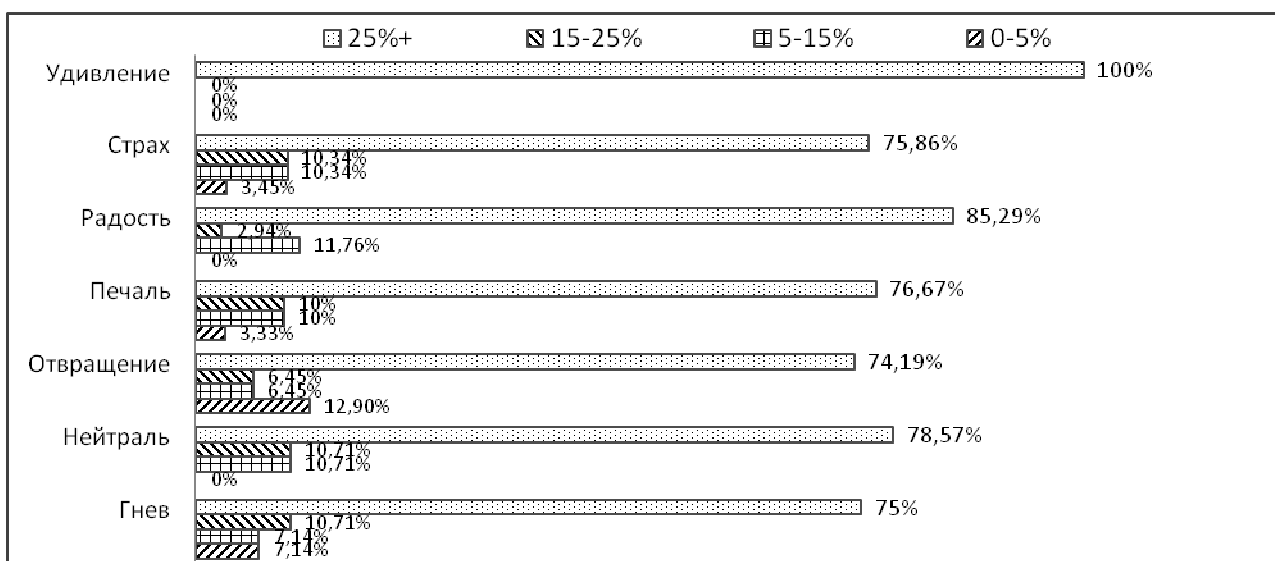


Рис. 4. Распределение значений C_{conf} построенной системы метрик

Как видно из рисунка 4, около 75% всех верно распознанных эмоций определены с наибольшей точностью. Максимальное количество верно распознанных изображений с минимальной разницей показывают эмоции «отвращение» и «гнев». В проявлении данных эмоций на лице человека участвуют схожие группы мышц. Снижение ошибки распознавания данных эмоций является одной из перспективных задач развития данной области.

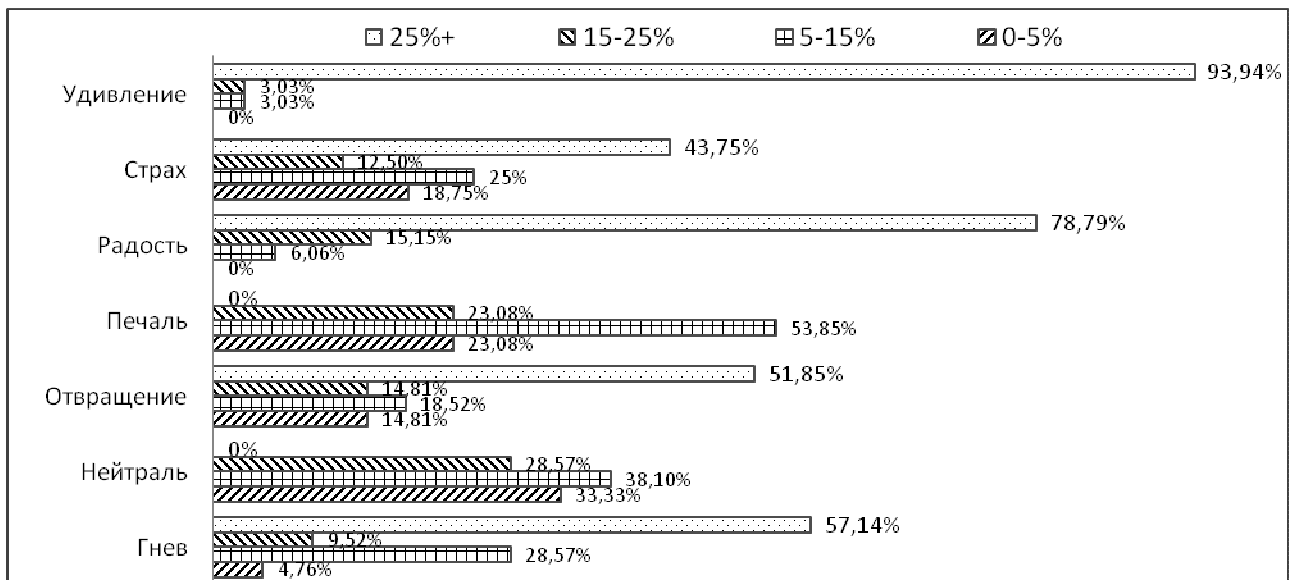


Рис. 5. Распределение значений C_{conf} стандартной системы

Использование описанной в [5] системы метрик наряду с увеличением ошибки распознавания эмоций позволяет получить неоднозначные результаты распознавания.

В результатах распознавания трех эмоций отсутствуют значения $C_{conf} > 25\%$. На нейтральной эмоции наблюдаются наиболее низкие значения уверенности. При определении эмоции в более сложных условиях, таких как перекрытие части лица и снижение общей интенсивности проявления эмоций, система, основанная на данной метрике, будет показывать быстро растущую ошибку по распознаванию.

Заключение

В данной работе была произведена адаптация системы кодирования лицевых движений к возможностям информационной системы распознавания эмоций. Полученные правила кодирования лицевой информации были объединены в общую систему метрик. Использование данной системы позволило снизить общую ошибку распознавания на 20% по сравнению с аналогичной. Расчет «уверенности» позволил комплексно оценить работу новой системы, а также выявить наиболее проблемные эмоции для распознавания.

Список литературы

1. Cootes T.F., Edwards G.J., Taylor C.J. Active appearance models // Proc. European Conference on Computer Vision 1998, Vol. 2, pp. 484–498, Springer, 1998.
2. Ekman P., Friesen W. Facial Action Coding System: A Technique for the Measurement of Facial Movement // Consulting Psychologists Press, Palo Alto, 1978.
3. Friesen W., Ekman P. EMFACS-7: Emotional Facial Action Coding System // Unpublished manual, University of California, California, 1983.

4. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression // Lucey P., Cohn J.F., Kanade T., Saragih J., Ambadar Z., Matthews I.J. //Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference, P. 94–101, San Francisco, CA, 2010
5. Widanagamaachchi W.N., Dharmaratne A. Emotion Recognizer: A neural network approach // 9th International Conference on Intelligent System Design and Applications, 2009.

Рецензенты:

Соколова Э.С., д.т.н., профессор, зав. кафедрой «Информатика и системы управления» ФГБОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексеева», г. Нижний Новгород;

Хранилов В.П., д.т.н., профессор кафедры компьютерных технологий в проектировании и производстве ФГБОУ ВПО «Нижегородский государственный технический университет им. Р.Е. Алексеева», г. Нижний Новгород.