

Задача распознавания ситуаций насилия с применением автоматизированных систем и методов искусственного интеллекта¹

С. Н. Ениколопов, кандидат психологических наук, доцент, заведующий кафедрой криминальной психологии факультета юридической психологии Московского городского психолого-педагогического университета (enikolopov@mail.ru)

Ю. М. Кузнецова, кандидат психологических наук, старший научный сотрудник Института системного анализа РАН

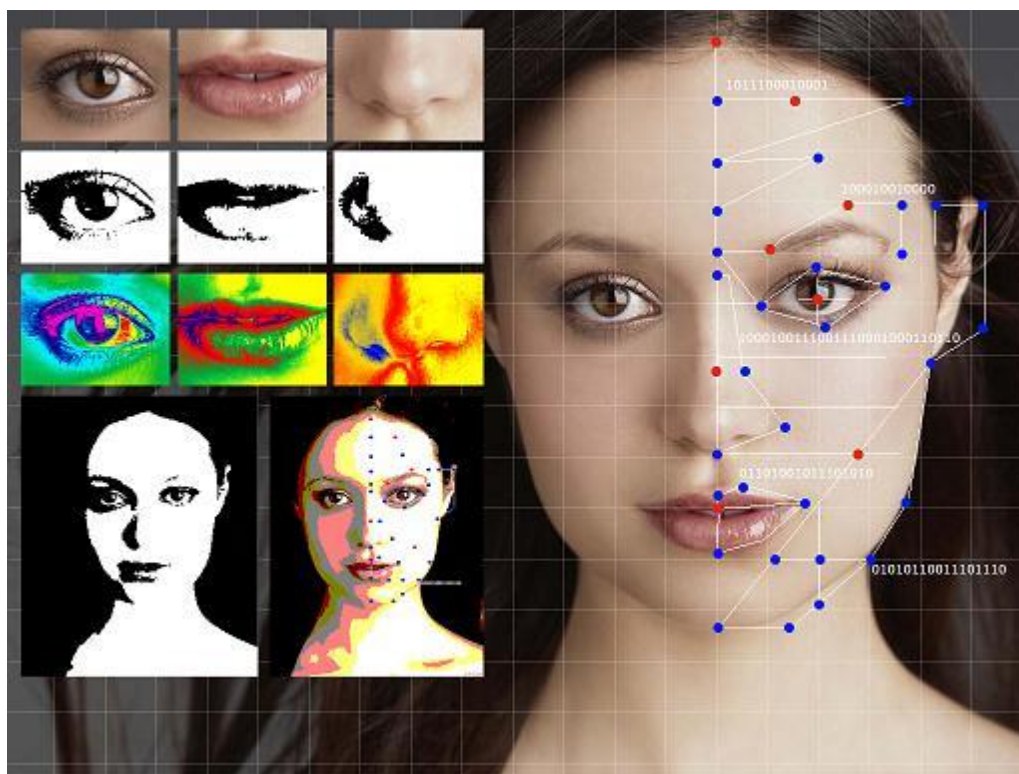


Рисунок опубликован на сайте ООО «Т-Инжиниринг» (<http://www.ti-eng.ru/technology/imagerecognition/>)

¹ Работа поддержана грантами РФФИ 10-06-00236а и РГНФ 10-06-00834а..

В настоящей работе рассматривается задача создания методами искусственного интеллекта компьютерного оборудования, позволяющего автоматизировать стадию выделения признаков насилия и служить средством, обеспечивающим предварительно обработанной информацией оператора, который исследует сложившуюся ситуацию и принимает решение. Преимуществом работы с такими системами является значительное расширение поля, доступного для восприятия оператора, поскольку система привлекает его внимание только к тем точкам, в которых назревает или наблюдается некоторое нарушение общественного порядка. Оператор получает возможность в режиме реального времени отслеживать ситуацию, складывающуюся в многочисленных точках наблюдения, и своевременно принимать необходимые меры по предотвращению эскалации угрозы или устранению последствий происшествия. Современной тенденцией развития таких автоматизированных систем является переход от использования визуальной информации к мультимодальному анализу, основанному на объединении видео- и аудиопотоков, поступающих с места действия. Показано, что одновременная обработка информации должна начинаться на первых стадиях анализа, т. е. целесообразно не суммировать данные независимых систем обработки, а «сливать» потоки звуковой и зрительной информации и обрабатывать их вместе как единый поток. Таким образом, в современных разработках систем распознавания поведения получает свою реализацию модель, приближенная к психологическим представлениям о восприятии человека.

Ключевые слова: распознавание эмоций, мультимодальные системы, агрессивное поведение.

Распознавание лицевой экспрессии и эмоциональных компонентов речи

Важным направлением исследований являются работы, направленные на создание алгоритмов автоматического распознавания шести прототипических эмоций по эмоциональной мимике и звуковым характеристикам.

Известно, что выражение эмоций может быть весьма разнообразным и различаться как в зависимости от индивидуальных особенностей, так и от ситуативного контекста. Культурный контекст играет важную роль, если речь идет об эмоциях, не относящихся к категории базовых, поскольку принятые в определенной общности способы выражения сложных эмоциональных состояний различаются. Кроме того, проявления, соответствующие «смешанным» эмоциям, трудны для распознавания, поскольку за ними скрывается целый комплекс испытываемых человеком чувств. Еще один фактор, влияющий на эффективность распознавания эмоций, – возможность фальсифицировать эмоциональное поведение. Культурные требования, личные представления о допустимом или любые соображения, вытекающие из восприятия человеком актуальной ситуации (желание что-то скрыть или демонстрация чувства, которого на самом деле нет), – эти и другие факторы оказывают влияние на выражение человеком своих эмоций.

Основой для разработки систем автоматизированного распознавания эмоций являются исследования лица и лицевой экспрессии. С этой специфической точки зрения лицо определяется как «мультиканальная система, производящая множественные послания» [6]. Сигналы, исходящие от этой мультиканальной системы, делятся на статические, медленные и быстрые. К статическим сигналам (стабильным признакам) относятся такие постоянные характеристики лица, как цвет кожного пигмента, форма лица, структура лицевого черепа, локализация и толщина жировых отложений, характерные особенности рта, носа, глаз и бровей. Медленными сигналами являются изменения, возникающие в течение определенного периода: постоянные морщины, мышечный тонус, текстура кожи, ее оттенок. Быстрые сигналы определяются деятельностью лицевых мышц, могут длиться секунды или доли секунды и производят временные изменения выражения лица, рисунка морщин, взаимного расположения и форм отдельных частей. Как система, продуцирующая множественные послания, лицо является источником информации о возрасте, гендерной и этнической принадлежности, чувствах, настроении, отношениях, характере, интеллекте и т. д.

Ошибки при определении эмоционального выражения связаны в плане организации наблюдения с такими факторами, как расположение воспринимаемого лица, освещенность и наличие предметов, загораживающих лицо; со стороны наблюдателя выделяется его способность к пониманию и идентификации эмоциональных состояний (чувствительность к выражениям эмоций и наличие навыков их категоризации) и контекста. Однако основным препятствием для правильной идентификации является намерение у наблюдаемого скрыть истинные чувства.

При разработке автоматизированных бимодальных систем распознавания эмоций решается ряд технических задач, основными из которых являются [6]:

1. разработка средств для эффективного обнаружения и сегментирования единичного изображения лица;
2. разработка средств отслеживания изменений выражения лица на видеозаписи;
3. определение совокупности индикаторов эмоциональной мимики, обеспечивающих анализ единичного изображения и последовательности изображений; моделирование на их основе мимической экспрессии, соответствующей базовым эмоциям;
4. выявление закономерностей изменений в соотношении отдельных визуальных признаков, соответствующих базовым эмоциям;
5. определение способа расчленения аудиоряда и выделения в получаемых сегментах акустических признаков, релевантных задаче распознавания эмоций;
6. распознавание эмоций на основе анализа унимодальной информации и моделирование алгоритмов би- и мультимодальной классификации, воспроизводящих внутренние механизмы распознавания звуковых и визуальных признаков отдельно и совместно;
7. достижение необходимого уровня надежности распознавания в условиях вариативности звуковых и визуальных характеристик;
8. разработка схемы категоризации эмоций, основанной на оценке мимической активности; определение адекватной структуры базы данных и комментариев;
9. определение способов объединения звуковой и зрительной информации, адекватных задачам бимодального распознавания эмоций; решение задачи синхронизации при разных уровнях интенсивности унимодальных сигналов;
10. разработка средств и методов, необходимых для использования целостной функционирующей программной системы автоматического распознавания эмоций и т. д.

В качестве информационной основы, позволяющей решать эти и другие технические задачи, используются различные базы данных, разработанные для нужд развития информационных технологий. Элементами таких баз данных являются многочисленные визуальные, аудиальные или бимодальные образы, специальным образом полученные, прошедшие компьютерную обработку и классифицированные в соответствии с задачами, для которых создавалась та или иная база. На специализированном сайте [10] размещены ссылки, с помощью которых можно ознакомиться, например, с такими базами данных в сфере эмоциональной экспрессии, созданными в последние годы, как The Color FERET Database, USA; SCface – Surveillance Cameras Face Database; Multi-PIE; The Yale Face

Database; Project – Face In Action (FIA) Face Video Database, AMP, CMU; Cohn-Kanade AU Coded Facial Expression Database; MIT-CBCL Face Recognition Database; Image Database of Facial Actions and Expressions – Expression Image Database; Face Recognition Data, University of Essex, UK; NIST Mugshot Identification Database; Japanese Female Facial Expression (JAFFE) Database; BioID Face DB – HumanScan AG, Switzerland; Indian Face Database; GavabDB: 3D face database, GAVAB research group, Universidad Rey Juan Carlos, Spain; BJUT-3D Chinese Face Database; The Basel Face Model (BFM); The Iranian Face Database (IFDB) и др.

В такой широко применяемой базе данных, как Система кодирования лицевых движений (Facial Action Coding System) [10], после выполнения процедуры оптимального сегментирования изображения лица было выделено около 100 признаков, позволяющих описать паттерны мимической экспрессии, например: Внутренние концы бровей приподняты; Внешние концы бровей приподняты; Глаза прикрыты; Верхняя губа приподнята; Носогубная складка углублена; Уголок губ опущен; Подбородок приподнят и т. д.

Каждому признаку здесь присвоен определенный номер, и сочетание номеров позволяет «кодировать» различные мимические выражения. Существенным фактором, который необходимо учитывать, является наличие индивидуального своеобразия выражения эмоции, поэтому в FASC для каждой эмоции существуют списки типичных кодов, что расширяет возможности ее уверенного автоматического распознавания (рис. 1).

Остановимся на задаче проведения сегментирования изображения лица. Согласно определению, данному в [4], «под сегментацией понимают процесс разбиения изображения на непересекающиеся области, семантически соответствующие объектам. Сегментация изображения обеспечивает снижение объема обрабатываемой информации. В процессе распознавания образов сегментация занимает одно из основополагающих мест по причине зависимости качества решения, получаемого в результате работы системы распознавания образов в целом, от правильно выделенных объектов. Ошибочное определение положения и размеров объектов на изображении, причиной которого может служить избыточная или недостаточная сегментация, в значительной степени усложняет получение приемлемого решения задачи и приводит к ошибочным результатам».

Технически достижение удовлетворительного результата при сегментировании распадается на несколько этапов:

1. предобработка исходного изображения, т. е. выравнивание яркостей, приведение к нужному виду и размеру;
2. выделение области лица с помощью анализа данной фотографии;
3. выделение черт лица (глаза, нос, рот) с помощью анализа области изображения, выделенной на предыдущем этапе.

Алгоритм выделения области лица на изображении с помощью комбинации методов цветовой и яркостной сегментации состоит из следующих этапов:

1. выделение границ на изображении;
2. выделение границ, имеющих цвет кожи;
3. выделение области, которая содержит границы, выделенные на этапе 2 [3].

Как выглядит результат сегментирования изображения, можно увидеть из приводимой ниже иллюстрации (рис. 2). Здесь можно увидеть, какое значение имеет для распознавания доступный для наблюдателя ракурс – в условиях естественного наблюдения камеры, разумеется, не всегда обеспечивают идеальный видеоряд. Снижать эффективность деятельности автоматизированной системы может также неоптимальный уровень освещенности.

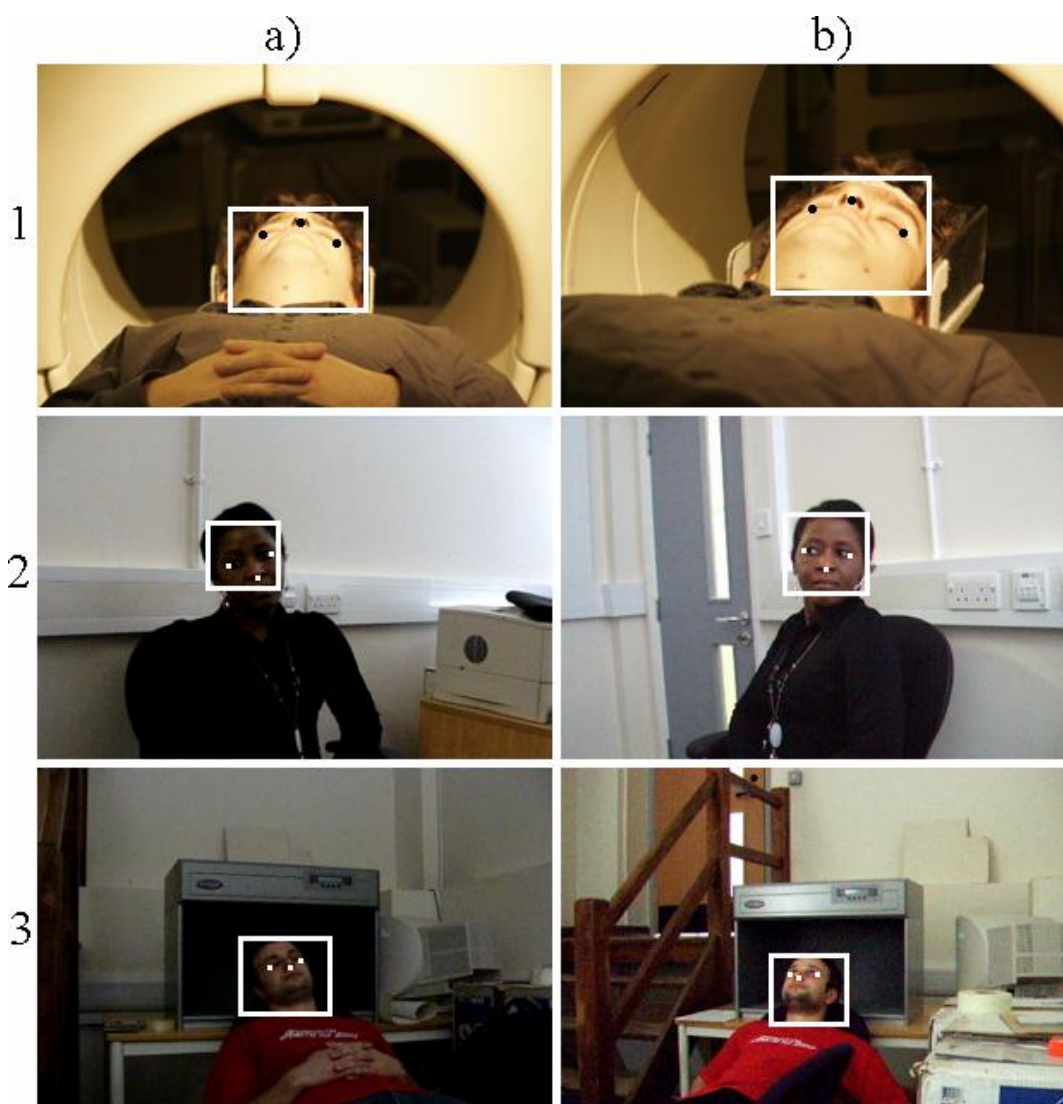


Рис. 2. Пример сегментации изображений (область лица выделена белым квадратом): а) фронтальная камера, б) левая камера (из: [1])

Аналогичным образом производится анализ эмоциональных компонентов в аудиопотоке. Звуковая информация сегментируется – выделяется из фона и разбивается на отдельные «звуковые кадры», которые затем преобразуются в цифровую модель, подвергающуюся дальнейшей обработке: выделению таких признаков, которые достоверно связаны с испытываемыми человеком эмоциями. Так, считается, что существуют следующие связи тембра и интенсивности голоса с эмоциональными состояниями говорящего:

- явно высокий тон – энтузиазм, радость, заинтересованность;
- высокий, в широком диапазоне силы, тональности и высоты – гнев и страх, неуверенность;
- чрезмерно высокий, пронзительный – беспокойство;
- мягкий и приглушенный, с понижением интонации к концу каждой фразы – печаль, усталость;
- форсирование звука – напряжение, намерение обмануть;
- в состоянии эмоционального возбуждения обычно возрастает сила голоса, изменяются его высота и тембр, но иногда сильное возбуждение может, наоборот, проявляться в уменьшении силы голоса (человек «шипит от ярости») [7].

Технически анализ звуковой информации, позволяющей вычленять признаки эмоциональных состояний, может быть описан как последовательность следующих этапов:

- запись голоса и формирование звукового файла в формате .wav;
- выделение единичного отрезка с определением начала записи и его длины;
- расчет $\max A$ (максимальной амплитуды), E – энергии, F – образа частотной составляющей;
- распознавание проводится с предыдущим обучением интеллектуальной составляющей [2].

Визуально результат обработки голосовой информации можно видеть на рис 3.



Рис. 3. Графическая запись речевого сообщения (опубликован на сайте <http://infgeoservice.narod.ru/publik2.html>)

Полученная в ходе сегментирования изображения и звукового потока информация используется для дальнейшей обработки движущегося изображения, но описание этого и последующих процессов (таких как слияние и одновременный анализ потоков разной модальности) уже не может обойтись без привлечения гораздо более сложного математизированного аппарата.

Распознавание потенциально опасного поведения

Описанные системы автоматизированного распознавания эмоций потенциально могут быть востребованы в разных сферах экономической и социальной деятельности, готовых для восприятия так называемого «аффективного компьютеринга» – термин, предложенный Розалин Пикар для описания обращенной к аффективной сфере человека компьютерной системы. В возглавляемой Р. Пикар группе Affective Computing разрабатываются модели внедрения эмоциональных компьютеров в такие виды деятельности, как психотерапия фобий (для автоматического оценивания эмоциональной реакции пациента на конкретное терапевтическое воздействие и рекомендаций терапевту по поводу выбора стратегии дальнейшей работы); дистанционное обучение (для интенсификации взаимодействия между учителем и учеником); обслуживание пожилых людей или детей в домашних условиях; маркетинг (для определения реакции покупателей на определенный товар); проекты «умного дома»; повышение безопасности дорожного движения (путем разработки автоматических средств, анализирующих выражение лица водителя, для организации его внимания в различных ситуациях на дороге). Одной из самых важных функций таких эмоционально «продвинутых» компьютеров должно стать их участие в создании более безопасной среды, и в данном направлении можно ожидать прогресса систем, создаваемых для идентификации и предотвращения террористических атак в местах массового скопления – вокзалах, аэропортах, поездах, автобусах, самолетах [8]. Дело в том, что своевременное распознавание агрессивных намерений или действий с необходимостью подразумевает способность к определению эмоциональных состояний наблюдаемых. Разработчики автоматизированных систем безопасности используют известное в психологии разделение агрессии на инструментальную и аффективную. Для нужд создания компьютерных технологий важны их особенности, приведенные в табл. 1.

Таблица 1.

Типы агрессии и способы их идентификации

Агрессия	Характеристики	Распознавание
Инструментальная	Целенаправленность, стремление достичь очевидную цель	Отслеживание движений и жестов, имеющих диагностическую значимость. Категоризация поведения как агрессивного в соответствии с известными паттернами. Выявление опасной или провоцирующей среды.
Аффективная	Сильно выраженные чувства; доминирует, как правило, гнев.	Гнев (наряду с другими эмоциями) распознается по выражению лица, речи, жестам.

Как видно из приведенных описаний, признаки, позволяющие решать задачу распознавания аффективной агрессии, выявляются благодаря устройствам, которые были описаны выше. Таким образом, возникает необходимость их интеграции с системами, способными распознавать инструментальную агрессию.

Упрощенная схема разворачивания агрессивного поведения, удовлетворительно отвечающая потребностям создания автоматизированных систем распознавания, такова:

- фаза инициации: резкое повышение уровня тревоги;
- фаза эскалации: повышение уровня тревоги, изменение телесных состояний (например, широко раскрытые глаза, опущенная голова, нахмуренные брови, стиснутые зубы, учащенное дыхание, красное лицо) и способа коммуникации (например, брань, угрозы и т. д.);
- кризис: если ситуация не может быть разрешена, она переходит в фазу неконтролируемого насилия;
- деэскалация: может наступать естественным образом, но чаще требуется внешнее вмешательство. Ситуация становится нормальной.

В результате сбора и анализа информации, предоставленной экспертами в области охраны порядка на железных дорогах в Нидерландах, была описана следующая процедура обнаружения и контроля агрессивного поведения (инструментальной агрессии):

- выделить: при осмотре мониторов внимание может быть привлечено к определенному изображению. Импульс может переживаться как какое-либо непривычное наблюдение, событие, внешний тревожный «звонок», предчувствие и т. п.;

-
- ориентироваться: выделить наиболее значимый компонент в изображении и выдвинуть одно или несколько предположений. Одной из гипотез является предположение, что все нормально и тревога ложная. Если она не подтверждается – переход к следующему шагу;
 - рассмотреть: искать другие признаки, которые будут поддерживать или опровергать гипотезы. Чем больше подтверждений, тем более вероятной становится соответствующая гипотеза;
 - сделать вывод: при достижении вероятностного порога гипотеза принимается. Величина порога зависит от знаний, опыта и индивидуальных представлений;
 - действовать: предпринять соответствующие меры для решения проблемы [12].

При разработке автоматизированных систем распознавания агрессии в качестве необходимой стадии выделяется работа с экспертами – носителями эмпирически эффективных знаний, какие признаки в поведении человека могут служить индикаторами агрессивных действий или намерений.

В экспертном знании выделяются его уровневые компоненты – признаки и концепты, а также описываются алгоритмы рассуждений экспертов. Опросы экспертов показывают, что вербализуемые представления отражают уровень концептов («человек встревожен»; «двое ссорятся»), которые являются сложными синтетическими образованиями, содержащими в себе результаты деятельности категоризации. Для нужд создания компьютерных программ данный уровень представления знания является итоговым, но не начальным, и в этом направленность анализа разработчиков интеллектуальных систем и представителей психологической науки совпадают. Возникает необходимость вычленения отдельных признаков агрессивности, т. е. таких элементов восприятия, наличие которых организует и направляет процесс категоризации коммуникативной ситуации как потенциально или актуально угрожающей.

Поиск признаков агрессивного поведения, их обнаружение, синтез, соотношение с общим состоянием среды (контекстом) у специалиста-практика протекает без контроля сознания, поэтому, с одной стороны, обнаружение угрозы им субъективно переживается как своеобразный инсайт, с другой – задача объяснения, что нужно сделать для обнаружения угрозы, вызывает значительные затруднения. Действительно, известно, что экспертное знание, позволяющее своему носителю решать профессиональные задачи на высоком уровне эффективности, не предъявлено ему самому в виде отрефлексированной и четко структурированной системы. Именно поэтому среди методов искусственного интеллекта существует специальный арсенал средств, помогающий экспертам различных специальностей актуализировать, вербализовать и структурировать имеющиеся у них представления. После такой обработки экспертные знания не только становятся доступными для использования другими людьми, но и приобретают свойство системности. Результатом взаимодействия с экспертами является построение специальной онтологии, т. е. такого схематического описания определенной области

реальности, которое состоит из структуры специальным образом обработанных данных, содержащей все релевантные классы объектов, их связи и правила, действующие в этой области. Для интеллектуальной системы онтология служит семантическим основанием производства рассуждений, воспроизводящих познавательные процессы специалиста. В конкретном исследовании [12] в качестве элементов онтологии агрессивного поведения в поездах были определены такие визуально и аудиально выявляемые показатели, как, например, следующие: Физический контакт; Человек бежит; Взгляд в упор; Человек быстро оглядывается; Крики; Быстрые движения рук; Кто-то показывает пальцем; У человека нож/оружие; У человека бутылка/сигарета; Оставленный багаж и т. д.

После составления удовлетворяющего цели создания системы автоматизированного восприятия перечня первичных признаков с помощью методов искусственного интеллекта было разработано компьютерное обеспечение, которое оказалось в состоянии воспроизводить ментальные действия, которые обслуживают процесс принятия решения о наличии агрессии у человека. Были составлены типовые сценарии, соответствующие возникающим в реальности ситуациям проявления агрессии, и компьютер «научился», анализируя видео- и аудиоинформацию, привлекать внимание оператора к ситуациям, которые могли требовать вмешательства и пресечения.

Примером отечественной разработки в данной сфере является проект, представленный на сайте ООО «Т-Инжиниринг» [11]. Объявлено, что используя собственный продукт – систему технического зрения УРСА™, – данная организация работает над поведенческо-ситуационным модулем, дорабатывая систему до уровня распознавания тревожных ситуаций применительно к классам объектов и правонарушений. В частности, поставлены задачи дифференцирования и прогнозирования агрессивного и дружественного поведения, определения «свой – чужой» на основе анализа и накопления моделей поведения, антропометрических данных и классификации описанных сценариев (по данным правоохранительных структур), а также для определения «забытых» или «оставленных» предметов, несанкционированного проникновения и любых других атипичных ситуаций, ведущих к правонарушениям. Благодаря расположенному на сайте рисунку (рис. 4), можно видеть, как выглядит экран монитора, с которым работает оператор автоматизированной системы. Согласно приведенному выше алгоритму, получив сигнал от системы, оператор должен будет выполнить свою часть работы – убедиться в достоверности тревоги и принять решение, что нужно сделать в создавшейся ситуации. Таким образом, автоматизированные системы не являются полностью автономными, их функция заключается в том, чтобы интенсифицировать процессы сканирования мультимодальной информации и проводить ее предварительную категоризацию в соответствии с имеющимися алгоритмами, имитирующими рассуждения эксперта. Важной характеристикой таких систем, закладываемой в их устройство, является возможность широкой адаптации, подстройки и модернизации аппаратурной части и способность к обучению, свойственная программному обеспечению. В последние годы в рамках работ по искусственному интеллекту проектирование таких адаптивных и активных компьютерных систем является одним из приоритетов деятельности специалистов Института системного анализа РАН в области создания динамических интеллектуальных систем и когнитивного моделирования. Создание таких систем

отвечает требованиям одного из Приоритетных направлений развития науки, технологии и техники РФ – безопасность и противодействие терроризму.



Рис. 4. Изображение, предъявляемое оператору автоматизированной системы распознавания атипичного поведения [опубликован на сайте ООО «Т-Инжиниринг» <http://www.ti-eng.ru/technology/imagerecognition/>]

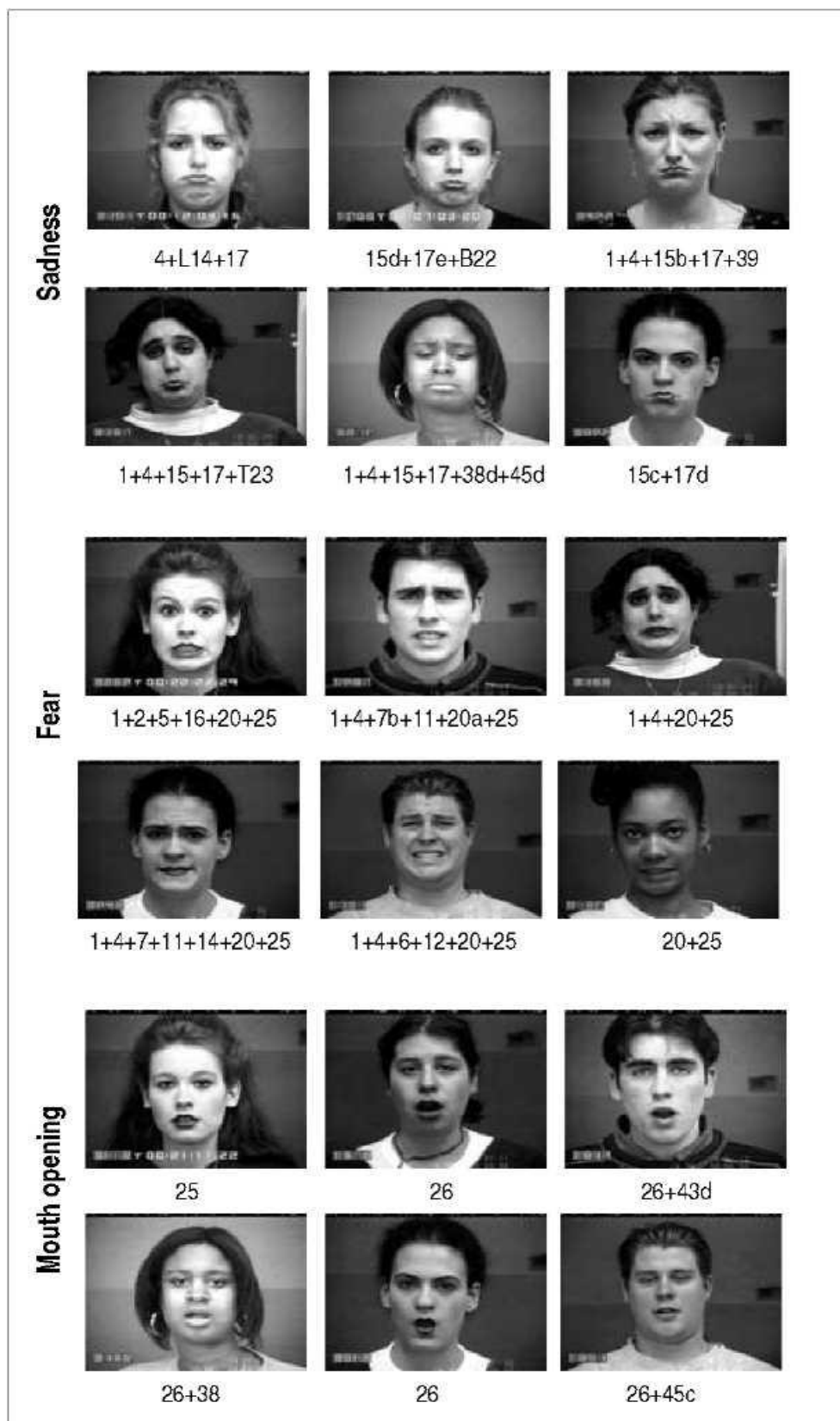


Рис. 1. Пример из базы данных FACS, иллюстрирующий индивидуальные варианты выражения эмоций: приведены числовые «коды» и соответствующие им варианты мимики (по:[7])

Литература

1. Анищенко С., Шапошников Д., Подладчикова Л., Камли Р., Сухоленцев К., Гао К. Мониторинг движений головы с помощью фовеального подхода и детектирования локальных лицевых опорных точек // http://nisms.krinc.ru/papers/PRIA_9_rus.pdf
2. Гафуров А. О. Алгоритмы оцифровки звука и нейросетевые методы распознавания слов и эмоций человека или живого существа в интеллектуальной нейроинформационной системе «НейроКибер» // <http://infgeoservice.narod.ru/publik2.html>
3. Лопатина А. Д. Выделение области лица с помощью комбинации методов цветовой и яркостной сегментации // Вестник УГАТУ. Управление, ВТ и И. 2009. Т. 13. № 2 (35) // [http://www.ugatu.ac.ru/publish/vu/stat/ugatu-2009-2\(35\)/24.pdf](http://www.ugatu.ac.ru/publish/vu/stat/ugatu-2009-2(35)/24.pdf)
4. Полякова М. В., Ищенко А. В., Худайбердин Э. И. Порогово-пространственная сегментация цветных текстурированных изображений на основе метода JSEG // ААЭКС. 2010 №1(25) // <http://aaecs.org/polyakova-mv-ishenko-av-hudaiberdin-ei-porogovo-prostranstvennaya-segmentaciya-cvetnih-teksturovannih-izobrazhenii-na-osnove-metoda-jseg.html>
5. Сайт «Речевые Технологии» // <http://speetech.by/press/analytics/1>
6. Datcu D. Multimodal Recognition of Emotions // Wöhrmann Print Service, 2009.
7. Fasel B., Monay F. & Gatica-Perez D. Latent Semantic Analysis of Facial Action Codes for Automatic Facial Expression Recognition // <http://www.idiap.ch/~gatica/publications/FaselMonayGatica-acmmm-mir04.pdf>
8. <http://affect.media.mit.edu/>
9. <http://www.face-and-emotion.com/dataface/facs/description.jsp>
10. <http://www.face-rec.org/databases/>
11. <http://www.ti-eng.ru/technology/imagerecognition/>
12. Yang Zh. Multi-Modal Aggression Detection in Trains. Delft: TU Delft Mediamatica, 2009.

The Task of Recognition of Violent Situations Using Automatic Systems and Methods of Artificial Intelligence¹

S. N. Enikolopov, Ph.D in Psychology, assistant professor, the head of the Criminal psychology chair of the Juridical psychology faculty, Moscow State University of Psychology and Education (enikolopov@mail.ru)

Yu. M. Kuznetsova, Ph.D in Psychology, senior researcher at the Institute of system analysis of the Russian Academy of Science

The article discusses the task of development of AI computer software with the purpose of automatizing the stage of indication of violence markers that provides an operator with processed information and helps him to analyze the situation and make a decision. The advantage of such systems is dramatic widening of the scope perceived by an operator since the system draws his attention only to the points of emerging or possible breaches of public order. An operator is provided with the opportunity to monitor the situation in numerous observed locations in the real time and to take necessary measures for prevention of a threat escalation or eliminating consequences of an accident. The current trend in development of such automatic systems is the transition from visual information processing to multimodal analysis based on combined audio- and video- streams translated from the scene of action. It is shown that simultaneous processing should begin at the first stages of the analysis since it is most rational not to summarize data from independent processing systems but to “merge” streams of audial and visual information and process them together as a single stream of data. Thus in modern developments of behavior recognition systems the model, close to psychological concepts of human perception, is implemented.

Keywords: emotions recognition, multimodal systems, aggressive behavior.

¹ With the support of the Russian Fund of Fundamental Research (grant 10-06-00236a) and the Russian Humanitarian Scientific Fund (grant 10-06-00834a).

References

1. *Aniwenko S., Shaposhnikov D., Podladchikova L., Kamli R., Suholencev K., Gao K.* Monitoring dvizhenij golovy s pomow'ju foveal'nogo podhoda i detektirovanija lokal'nyh licevyh opornyh toчек // http://nisms.krinc.ru/papers/PRIA_9_rus.pdf
2. *Gafurov A. O.* Algoritmy ocifrovki zvuka i nejrosetevye metody raspoznavanija slov i jemocij cheloveka ili zhivogo suwestva v intellektual'noj nejroinformacionnoj sisteme «NejroKiber» // <http://infgeoservice.narod.ru/publik2.html>
3. *Lopatina A. D.* Vydelenie oblasti lica s pomow'ju kombinacii metodov cvetovoj i jarkostnoj segmentacii // Vestnik UGATU. Upravlenie, VT i I. 2009. T. 13. № 2 (35) // [http://www.ugatu.ac.ru/publish/vu/stat/ugatu-2009-2\(35\)/24.pdf](http://www.ugatu.ac.ru/publish/vu/stat/ugatu-2009-2(35)/24.pdf)
4. *Poljakova M. V., Iwenko A. V., Hudajberdin Je. I.* Porogovo-prostranstvennaja segmentacija cvetnyh teksturovannyh izobrazhenij na osnove metoda JSEG // AAJeKS. 2010 №1(25) // <http://aaecs.org/polyakova-mv-ishenko-av-hudaiberdin-ei-porogovo-prostranstvennaya-segmentaciya-cvetnih-teksturovannyh-izobrazhenii-na-osnove-metoda-jseg.html>
5. Sajt «Rechevye Tehnologii» // <http://speetech.by/press/analytics/1>
6. *Datcu D.* Multimodal Recognition of Emotions // Wöhrmann Print Service, 2009.
7. *Fasel B., Monay F. & Gatica-Perez D.* Latent Semantic Analysis of Facial Action Codes for Automatic Facial Expression Recognition // <http://www.idiap.ch/~gatica/publications/FaselMonayGatica-acmmm-mir04.pdf>
8. <http://affect.media.mit.edu/>
9. <http://www.face-and-emotion.com/dataface/facs/description.jsp>
10. <http://www.face-rec.org/databases/>
11. <http://www.ti-eng.ru/technology/imagerecognition/>
12. *Yang Zh.* Multi-Modal Aggression Detection in Trains. Delft: TU Delft Mediamatica, 2009.