

## МЕТОДИКА И МОДЕЛЬ КЛАСТЕРИЗАЦИИ ПАТТЕРНОВ ДВИГАТЕЛЬНОЙ АКТИВНОСТИ ЛИЦА КАК ПРЕОБРАЗОВАНИЙ МЕТАГРАФОВ

**Б.А. Князев, В.М. Черненький**

МГТУ им. Н.Э. Баумана, Москва, Российская Федерация  
e-mail: bknyazev@bmstu.ru; chernen@bmstu.ru

*Предложена методика кластеризации паттернов двигательной активности лица по последовательностям изображений на основе представления изображений в виде метаграфов и их преобразований. Отличительной особенностью работы является интеграция знаний из нескольких доменов в одну иерархическую структуру для вычисления таких преобразований. Предложены функции поиска паттерна и добавления нового паттерна, а также процедура обучения данных функций с использованием тренировочных баз, размеченных экспертами. Приведены результаты исследования алгоритма сравнения паттернов во временной и частотной областях. Рассмотрен алгоритм реорганизации кластеров, необходимый для оптимизации коллекции паттернов. Реализация представленной методики и модели позволит повысить производительность экспертов при работе с видеоархивами, снятыми не в лабораторных условиях, экспериментально проверить соответствие полученных кластеров паттернам системы кодирования двигательной активности лица FACS, применяемой во многих актуальных приложениях.*

**Ключевые слова:** паттерны поведения, активность лица, метаграф, иерархическая модель, домены преобразований, видеоизображение, кластеризация.

## METHOD AND MODEL FOR CLUSTERING FACIAL ACTIVITY PATTERNS USING METAGRAPH TRANSFORMATIONS

**B.A. Knyazev, V.M. Chernenkiy**

Bauman Moscow State Technical University, Moscow, Russian Federation  
e-mail: bknyazev@bmstu.ru; chernen@bmstu.ru

*A method for clustering facial activity patterns from image sequences is proposed that is based on image representation as metagraphs and on their transformations. The distinctive feature of this work is integration of knowledge from several domains into a single hierarchical structure to compute these transformations. The functions of searching for a pattern and adding of a new one as well as the procedure for learning these functions by exploiting the training datasets annotated by experts are suggested. Experimental data for the algorithm, which compares patterns as temporal sequences applying time and frequency warping, are presented. The algorithm for cluster reorganization that is necessary for optimization of a collection of patterns is discussed. Implementation of the presented method and model is expected to improve performance of experts working with human videos recorded in more challenging conditions than in a lab. The presented work can also be used to experimentally compare the extracted clusters with the patterns defined in the Facial Action Coding System, which is employed in many up-to-date applications.*

**Keywords:** behavioral patterns, facial activity, metagraph, hierarchical model, transformation domains, video, clustering.

Исследования внешности и поведения человека показывают, что существуют определенные паттерны (шаблоны): статические геометрические (физиогномика) и поведенческие. Эти паттерны несут в

себе разнородную информацию, полезную для решения актуальных проблем [1–3], которая содержится как в явном виде (количественные и качественные параметры паттернов), так и в неявном виде, требующем интерпретации. В отличие от контролируемых паттернов, в основном используемых для передачи явной информации по невербальному каналу, например, в системах управления жестами [4], произвольные паттерны сложнее поддаются интерпретации и кодированию с помощью конечных *алфавитов* (наборов паттернов).

Однако вопрос интерпретации параметров паттернов не рассматривается в настоящей работе, так как он относится к изучению психологических, психических, физиологических, социальных и других областей знаний, касающихся процессов формирования внешности и поведения индивида в конкретный момент времени. Актуальные области применения методов интерпретации параметров паттернов поведения лица можно найти в работах [5, 6].

Для описания количественных и качественных параметров как произвольных, так и произвольных *паттернов поведения лица* (Action Unit, AU) часто используют систему кодирования FACS [7]. Эта система содержит экспертные вербальные описания конечного набора паттернов поведения лица, следовательно, возникает вопрос о ее полноте, объективности и валидности интерпретации данных на основе указанной системы. Поэтому одной из задач настоящей работы является самоорганизованная кластеризация паттернов и их сопоставление с известным набором из системы FACS.

Количественный параметр паттерна — интенсивность (А-Е). Однако в данной работе также добавлены длительность и частотные составляющие паттерна. Качественный параметр — его вербальное описание, заимствованное из работы [7].

Цель настоящей работы — предложить методические основы кластеризации паттернов поведения лица по последовательностям двумерных изображений для более эффективного исследования видеоархивов с поведением человека и экспериментальной проверки паттернов системы кодирования FACS. Более строгое и подробное описание отдельных моделей методики будет рассмотрено в последующих работах.

**Краткое описание методики.** Распознавание паттернов происходит покадрово (рис. 1) [8]. Для каждого  $i$ -го изображения из последовательности

$$\Phi = [\mathbf{I}_1, \dots, \mathbf{I}_N] \in \mathbb{R}^{a \times b \times N}$$

формируется неориентированный метаграф  $\mathbf{G}_i$ , так что последовательность изображений преобразовывается в последовательность метаграфов:

$$\Phi \mapsto [\mathbf{G}_1, \dots, \mathbf{G}_N],$$

где  $a$  и  $b$  — размеры изображения (кадра);  $N$  — общее число кадров;  $\mathbf{G}_i = f(\mathbf{I}_i)$ . Аналогично работе [9], интервал времени  $dt = t(i+1) - t(i) = \text{const} \forall i \in \{1, \dots, N-1\}$ . Входные и выходные данные алгоритма формирования метаграфа изложены ниже, его подробное описание будет представлено в последующих работах. Для оперирования в матричной форме метаграф  $\mathbf{G}_i$  представляется в виде разреженного вектор-столбца  $\mathbf{g}_i = [g_1, \dots, g_n]^T$ .

Граф первого кадра  $\mathbf{G}_1$  — базовый (reference graph), его преобразованием формируют последующие графы  $[\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_N] \mapsto [\mathbf{G}_1, T_1(\mathbf{G}_1), \dots, T_{N-1}(T_{N-2}(\dots T_1(\mathbf{G}_1)))]$ . При этом в дальнейшем для кластеризации паттернов сохраняются только сами преобразования  $\mathbf{T} = [T_1, \dots, T_{N-1}]$  и граф элементов лица  $\mathbf{G}_{\text{лицо}}$ . При вычислении преобразований и сопоставлении вершин метаграфа учитываются незначительные изменения между соседними кадрами, используются знания между доменами (модальностями) для расчета согласованных преобразований. Операторы агрегирования различных модальностей использовались, например в работе [4], для распознавания жестов. Таким образом, последовательность изображений представляет собой метаграф лица и последовательность преобразований:  $\Phi \mapsto \langle \mathbf{G}_{\text{лицо}}, \mathbf{G}_{\mathbf{T}} \rangle$ .

Как показано далее, паттерн поведения лица ( $AU$ ) будет представлять собой подмножество преобразований на подмножестве изображений в одном из доменов преобразований. В предыдущих работах [10, 11] модели обучаются для распознавания конечного набора  $[AU_1, AU_2, \dots, AU_M]$  из системы кодирования FACS. Альтернативный подход — кластеризация неконечного алфавита. Однако в таких рабо-

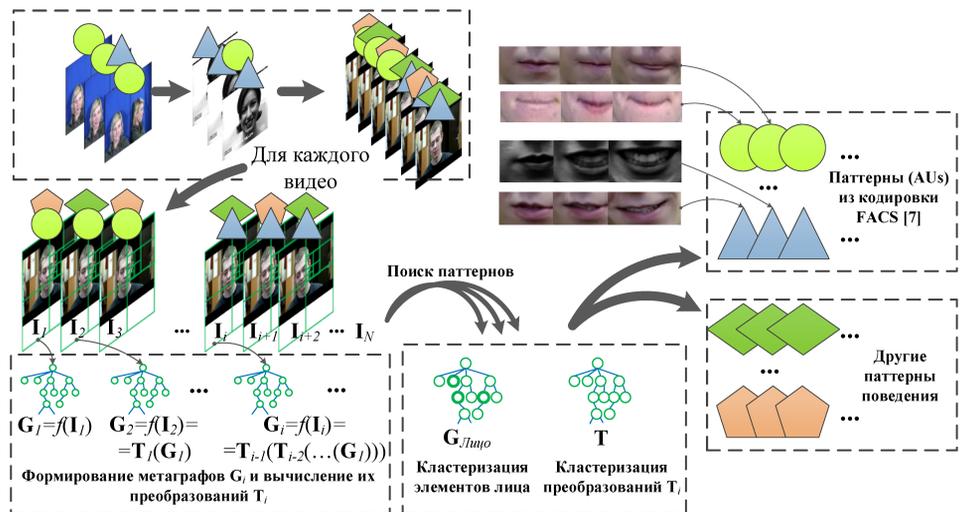


Рис. 1. Схема, иллюстрирующая методику кластеризации паттернов поведения лица (группы элементов одной формы представляют собой экземпляры паттернов в пределах одного кластера)

тах используются жесткие граничные условия, включая длительность паттерна, встречаемость паттерна у других людей и др. [8]. При анализе видеоархивов, снятых не в лабораторных условиях (см. рис. 1), становится понятно, что с каждым новым видеоматериалом появляются новые паттерны в связи с индивидуальными особенностями людей и изменившимся контекстом ситуации.

Поэтому далее для выполнения цели настоящей работы формулируются две задачи:

1. Обучение функции поиска и добавления паттерна в кластер.
2. Реорганизация кластеров паттернов.

Отметим, что класс — это тип паттерна, кластер — набор экземпляров паттернов одного класса, т.е. число кластеров равно числу классов, поэтому в некоторых контекстах понятия “класс” и “кластер” имеют одинаковый смысл.

Первая задача сводится к задаче обучения бинарной функции добавления нового кластера паттернов  $J_{\text{add}}$ , которая возвращает 0 в случае, если паттерн представляет собой разновидность или комбинацию других, или 1 в случае, если необходимо добавить новый кластер, центром которого является указанный паттерн. Вне зависимости от этого значения выполняется реорганизация кластеров, т.е. решается вторая задача.

Часть баз последовательностей изображений MMI [12] и СК+ [13] предложено использовать для обучения модели (функций добавления кластера и их реорганизации), часть — для тестирования модели и экспериментальной проверки соответствия полученных кластеров паттернам кодировки системы FACS. Также предложено применять базу видеоматериалов (“База1”), аннотированную с помощью инструмента автоматизированной обработки видеоизображений [14], для тестирования других паттернов, не представленных в базах MMI/СК+.

В целом методика объединяет концепцию метаграфа и методы пространственно-временного и визуального анализа с помощью мультимедийной взаимосвязи и функций оптимизации.

**Входные и выходные данные алгоритма формирования метаграфа статического изображения лица.** Входными данными алгоритма является  $i$ -е изображение  $I_i \in \mathbb{R}^{a \times b}$  последовательности  $\Phi = [I_1, \dots, I_N] \in \mathbb{R}^{a \times b \times N}$ , выходными данными — граф  $G_i = f(I_i)$ .

Для кластеризации паттернов поведения лица необходимо найти представление изображения, более устойчивое к искажениям объектов на нем, чем пиксельное представление. Такое представление должно быть более компактным:  $n < ab$ , где  $n$  — число ненулевых компонентов представления. Так, в работе [8] в качестве описания изображения использованы геометрические параметры лица на основе модели активного представления (ААМ) ( $n_1 = 8$ ) и дескриптор

SIFT ( $n_2 = \text{const} = 128$ ), совместно формирующие вектор-дескриптор  $\mathbf{x}_i = f(\mathbf{I}_i)$ ,  $\mathbf{x}_i \in \mathbb{R}^{(n_1+n_2)}$ . Известны и другие варианты, например, на основе фильтров Габора [15].

Графовые модели и их расширения успешно применяют во многих приложениях распознавания образов [16]. В настоящей работе для представления изображения создана модель, являющаяся частным случаем обобщенной модели метаграфа [17], но имеющая несколько специфичных для решаемой задачи свойств и ограничений, описанных далее. Введение метаграфа обусловлено наглядностью описания вложенностей и дуальностей значений из разных доменов (областей знаний, модальностей)  $D_q \in \{D_1, \dots, D_Q\}$  и связей (ребер) элементов иерархии. Доменами  $D_q$  могут быть как вербальные (семантика), так и визуальные (текстурные свойства, геометрические и цветовые свойства объектов) свойства. Кроме того структура может быть адаптивной, и, следовательно, обладать меньшей избыточностью по отношению к изображению.

Алгоритмы формирования ( $f$ ) и обучения метаграфа статического изображения для каждого домена сходны, но требуют детального и объемного описания, поэтому будут рассмотрены в отдельных работах. В настоящей статье требуется, чтобы выходом алгоритма  $f$  являлся иерархический граф (рис. 2):

$$\mathbf{G}_i = f(\mathbf{I}_i) = \langle V, E \rangle,$$

где  $V = \{v_{jkD_q}\} \in D_q$  — набор вершин графа со значениями в домене  $D_q$ ;  $j \in \{1, \dots, h\}$  — уровень иерархии вершины;  $h = h_{11}$  — глубина графа, т.е. число его уровней;  $h_{jkD_q}$  — глубина вершины  $v_{jkD_q}$ , условие  $h_{jk_1D_q} = h_{jk_2D_q}$  не гарантировано для всех  $j$  и  $k$  (иерархия является несбалансированной);  $k \in \{1, \dots, K^{j-1}\}$ ;  $K \geq 2$  — арность графа, или число дочерних вершин;  $E = \{w_{(j+1)mD_q}\} \in \mathbb{R}$  — набор ребер, соединяющих  $k$ -ю вершину уровня  $j$  с  $m$ -й вершиной уровня  $j+1$  в пределах домена  $D_q$ ;  $m$  и  $k$  связаны соотношениями  $m = \{K(k-1) + 1, \dots, Kk\}$ ,  $k = \lceil m/K \rceil$ . Значение  $K$  определяет число блоков, на которое разбивается текущий блок изображения, и является предметом дальнейших эмпирических исследований.

Перечислим требования, предъявляемые к домену  $D_q$ .

*Требование I.* Возможность декомпозиции значений домена  $D_q$  и определение операции умножения на скаляр:  $v_{jkD_q} = \bigcup_m w_{(j+1)mD_q} \times v_{(j+1)mD_q}$ , где  $\sum_m w_{(j+1)mD_q} = 1$ . Таким образом, вершины более верхнего уровня  $j$  могут содержать вершины более низкого уровня  $j+1$  с учетом весовых скалярных коэффициентов (значений ребер). Подобное представление имеет тесную взаимосвязь с представлениями, приведенными в работах по иерархическим моделям [18,19], и

представляет собой специальный случай метаграфов, рассмотренных в работе [17]. Однако здесь для формирования метаграфа заданы значения его вершин в нескольких доменах  $D_q$ :  $v_{jk} = \bigcup_q v_{jkD_q}$ . Кроме того, к домену предъявляют еще три требования.

*Требование II.* Декомпозиция существует хотя бы для одного уровня  $j$  и число вариантов таких декомпозиций для одной и той же вершины больше или равно единице.

*Требование III.* Существует пустое значение  $v_{jkD_q} = \emptyset$ , которое объявляется как недекомпозируемое и для которого не определена операция умножения на скаляр.

*Требование IV.* Определена мера расстояния между непустыми вершинами:  $\rho_{j_1k_1j_2k_2D_q} = (v_{j_1k_1D_q} \ominus v_{j_2k_2D_q}) \in \mathbb{R}$ .

Примером домена, удовлетворяющего перечисленным требованиям  $I - IV$ , могут стать пространственные позиции вершины (например, ее центроид):  $D_q = S$ . Значения в таком домене, принадлежащем пространству  $\in \mathbb{R}$ , следовательно:  $v_{jkS} = \sum_m w_{(j+1)mS} v_{(j+1)mS} = \mathbf{w}_{(j+1)S}^T \mathbf{v}_{(j+1)S}$ , где  $\mathbf{w}_{(j+1)S}$ ,  $\mathbf{v}_{(j+1)S} \in \mathbb{R}^K$ , число вариантов декомпозиций больше 1 для  $K \geq 2$ . В качестве пустого элемента может использоваться  $\infty$ , меры расстояния — разность.

Примерами доменов могут быть: текстурное описание с помощью набора двумерных вейвлет-фильтров с ориентациями  $\theta$  и масштабами  $\sigma$  [15]; распределения в цветовых каналах [4]; ссылка на концепт из модели знаний. Доказательство требований  $I - IV$  для этих и других доменов будет рассмотрено в последующих работах.

Предложенный метаграф можно представить как набор отдельных графов (см. рис. 2):

$$\mathbf{G}_i = f(\mathbf{I}_i) = \{\mathbf{G}_{iD_1}, \dots, \mathbf{G}_{iD_Q}\}.$$

**Формирование разряженного вектора метаграфа.** Вычислительные системы обычно работают с векторами и матрицами, поэтому, как было отмечено ранее, во многих приложениях [8, 10, 11] изображения  $\mathbf{I}_i \in \mathbb{R}^{a \times b}$  преобразовываются в некоторые вектор-дескрипторы  $\mathbf{x}_i \in \mathbb{R}^n$  фиксированной размерности  $n$ . Это позволяет работать с последовательностью изображений  $[\mathbf{I}_1, \dots, \mathbf{I}_N] \in \mathbb{R}^{a \times b \times N}$  как с матрицей  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{n \times N}$ .

В таких случаях графы обычно представляются в виде инцидентных и смежных матриц. Однако в иерархическом графе связи присутствуют только на соседних уровнях, поэтому элементы на определенных индексах всегда будут нулевыми, что неэффективно. В настоящей работе структура, а следовательно, и размерность, формируемого метаграфа для каждого изображения зависит от самого изображения. В связи с этим для приведения метаграфа к одной размерности, а также сохранения его пространственной структуры

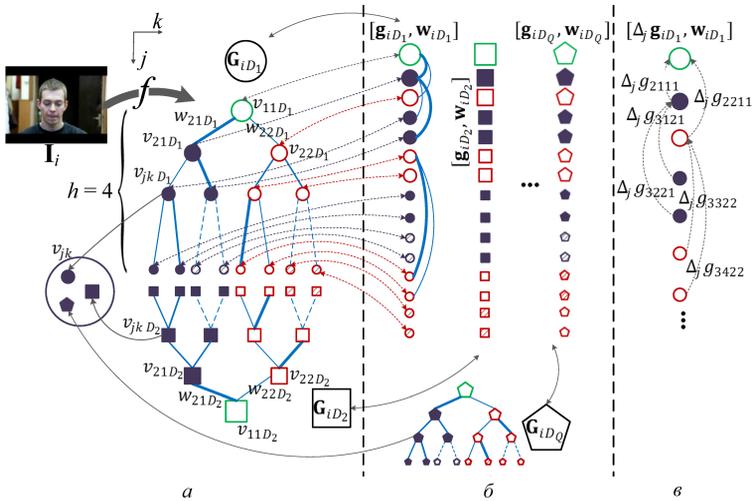


Рис. 2. Формирование метаграфа (алгоритм  $f$ ) для трех доменов (а), разреженных векторов (б) и разностного вектора для домена  $D_1$  (в), частный случай для  $K = 2$  (заштрихованные вершины формально отсутствуют в метаграфе ( $=\emptyset$ ), в векторе они присутствуют как пустые элементы для сохранения структуры графа; каждый тип фигуры означает отдельный домен  $D_q$ , толщина ребра ассоциирована с его весовым значением; стрелкой показано направление вычисления разности)

и других свойств применяется преобразование набора вершин  $V$  в разреженный вектор-столбец  $g_{iD_q} = [g_1, \dots, g_n]^T \in D_q^n$ , а набора ребер (весов)  $E$  в разреженный вектор-столбец  $w_{iD_q} = [0, w_1, \dots, w_n]^T \in \mathbb{R}^n$ ,

где  $m(j, k) = \sum_{i=0}^{j-1} K^i + k, m \in \{1, \dots, n\}$  — индекс вершины  $v_{jkD_q}$  и

ребра  $w_{jkD_q}$  в векторах  $g_{iD_q}$  и  $w_{iD_q}$  соответственно. Такое представление является обратимым, т.е. возможно восстановить граф  $G_i = f(I_i)$

(см. рис. 2). Размерность векторов  $n = \sum_{i=0}^h K^i$ , где  $h = \log_c \max(a, b)$  —

высота дерева, которая будет идентична для всех изображений размером  $ab$  в независимости от домена;  $c > 1$  — отношение размеров блоков верхнего уровня к нижним (обычно равно 2). При этом число нулей будет отличаться в зависимости от содержания изображения и домена и будет больше у менее информативных изображений. Особенностью является то, что, применяя разреженные типы хранения данных, в которых для хранения нулевых значений не выделяется память, можно решить проблему хранения вектора такой размерности. Отметим, что, если вектор весов известен для всех  $j$  и  $k$ , то значения вершин являются функцией значений весов и вершин уровня  $h$ :  $\Delta g_{iD_q} = \varphi(\Delta w_{iD_q})$ . Таким образом, последовательность изображений переходит в матрицу разреженных метаграфов:

$$\Phi \mapsto \left[ \left[ \mathbf{g}_{1D_q}, \mathbf{w}_{1D_q}; \dots; \mathbf{g}_{1D_Q}, \mathbf{w}_{1D_Q} \right]; \dots; \left[ \mathbf{g}_{ND_q}, \mathbf{w}_{ND_q}; \dots; \mathbf{g}_{ND_Q}, \mathbf{w}_{ND_Q} \right] \right] \in \mathbf{D}^{n \times 2Q \times N},$$

где  $\mathbf{g}_{iD_q}$  —  $i$ -й метаграф домена  $D_q$  в форме разряженного вектора размерностью  $n$ ;  $\mathbf{w}_{iD_q}$  — вектор его весовых коэффициентов. Вектора  $\mathbf{g}_{iD_q}$  могут обобщаться и формировать граф  $\mathbf{G}_{\text{лицо}}$ , необходимый для определения элементов лица.

Недостаток такого представления заключается в сложности обработки абсолютных значений вершин в разных доменах, а также в отсутствии связей между доменами. Для устранения указанных недостатков и кластеризации паттернов поведения предлагаются преобразования метаграфов.

**Сопоставление метаграфов последовательностей изображений и формирование последовательности переходов.** Цель настоящего раздела — обучить модель сопоставлять два последовательных метаграфа  $\mathbf{G}_i$  и  $\mathbf{G}_{i+1}$  и вычислять оптимальный вектор преобразований  $\mathbf{T}$  одного к другому. Такие преобразования будут основой алгоритма кластеризации паттернов поведения.

Пусть имеется два метаграфа  $\mathbf{G}_i$  и  $\mathbf{G}_{i+1}$ , такие что  $\exists j, k : h_{ijk} \neq h_{(i+1)jk}$ ,  $K_1 = K_2 = K$ . Каждая вершина следующего графа сопоставляется с предыдущей по следующим принципам.

1. При вычислении преобразований и сопоставлении вершин метаграфа подразумеваются изменения  $|\Delta \mathbf{g}| < \tau$  на каждом следующем кадре. Минимальное значение  $\tau$  ограничено техническими возможностями. Максимальное значение  $\tau$  остается предметом эмпирического исследования и зависит от области применения. Для исследования двигательной активности человека это значение соответствует не менее 20 кадр/с.
2. При сравнении двух метавершин приоритет отдается более близкой вершине в пространстве и времени, а не более похожей визуально. Поскольку графы приведены к идентичной структуре в каждом домене  $[\mathbf{g}_{iD_q}, \mathbf{w}_{iD_q}]$  и  $[\mathbf{g}_{(i+1)D_q}, \mathbf{w}_{(i+1)D_q}]$ , рассматриваем преобразования только между вершинами с одинаковыми индексами  $j$  и  $k$  в пределах одного домена.

Вектор преобразования весов определим как

$$\Delta \mathbf{w}_{iD_q} = (\mathbf{w}_{(i+1)D_q} - \mathbf{w}_{iD_q}) \in \mathbb{R}^n,$$

вектор преобразования значений вершин — как

$$\Delta \mathbf{g}_{iD_q} = \left( \Delta_j \mathbf{g}_{(i+1)D_q} \ominus \Delta_j \mathbf{g}_{iD_q} \right) \in \mathbb{R}^n, \quad (1)$$

где

$$\begin{aligned} \Delta_j \mathbf{g}_{iD_q} &= [v_{11}, v_{21} \ominus v_{11}, v_{22} \ominus v_{11}, \dots, v_{jk} \ominus v_{(j-1)k_{par}}, \dots] = \\ &= [v_{11}, \rho_{2,1,1,1}, \rho_{2,2,1,1}, \dots, \rho_{jk(j-1)k_{par}}, \dots]. \end{aligned}$$

Здесь  $k_{par} = \lceil k/K \rceil$  — индекс родительской вершины;  $v_{jk} \in \mathbf{g}_{iD_q}$ . Вектор  $\Delta_j \mathbf{g}_{iD_q} \in \mathbb{R}^n$  в силу *требования IV* домена к мере расстояния  $\rho$ .

Операция  $\Delta_j$  аналогична операции вычисления дискретной производной, но ей не является, так как индексы, для которых определяются меры расстояния, идут не последовательно, и вместо разности в области вещественных чисел используется мера расстояния. Отметим также, что в (1) первым элементом вектора  $\Delta_j \mathbf{g}_{iD_q}$  будет мера расстояния между корневыми элементами ( $v_{11}$ ), тогда как остальные элементы будут равны разностям соответствующих мер  $\rho$  в пространстве  $\mathbb{R}$ .

Вектора весовых коэффициентов  $\Delta_j \mathbf{w}_{iD_q}$  необходимы только для согласования доменов и более корректного вычисления преобразований  $\Delta_j \mathbf{g}_{iD_q}$ , о чем будет изложено далее. Для последующей кластеризации паттернов определяется вектор преобразования между графами  $\mathbf{G}_i$  и  $\mathbf{G}_{i+1}$  в домене  $D_q$ :  $\mathbf{T}_{iD_q} = \Delta_j \mathbf{g}_{iD_q} \in \mathbb{R}^n$ , и матрица преобразований для всех доменов (рис. 3):

$$\mathbf{T}_i = [\Delta_j \mathbf{g}_{iD_1}; \dots; \Delta_j \mathbf{g}_{iD_Q}] \in \mathbb{R}^{n \times Q}. \quad (2)$$

Последовательность изображений представляет собой аналогию второй производной по  $i$  и  $j$ :  $\Phi \mapsto \Delta_{i,j}^2 \mathbf{g}_i$ .

Отметим, что основная проблема решения задачи корректного вычисления значений  $\mathbf{T}_i$  заключается в том, что последовательность  $\Phi$  представляет собой дискретный набор матриц  $\mathbf{I}_i$ , истинный вид функции перехода между которыми неизвестен так же, как неизвестны и значения матриц  $\mathbf{I}$  между  $i$ -м и  $(i+1)$ -м изображениями.

Еще одной особенностью является то, что значения вершин метаграфов  $\mathbf{G}_i$  и  $\mathbf{G}_{i+1}$  вычисляются для областей, локализованных в

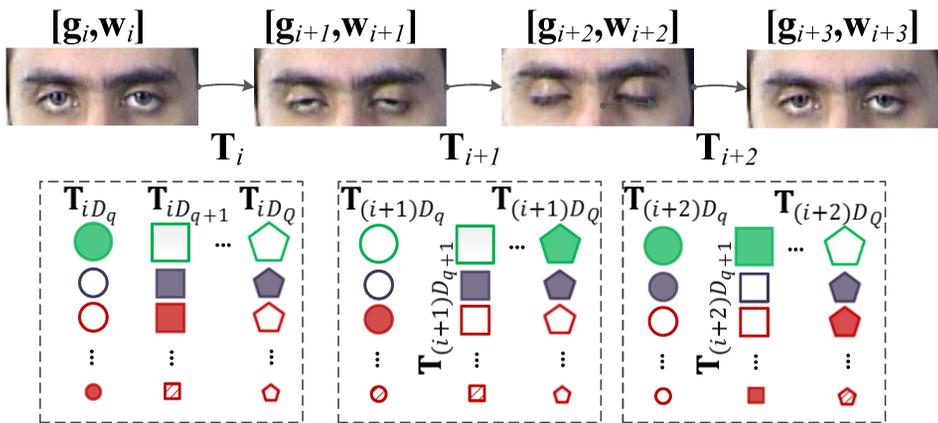


Рис. 3. Преобразования для вырезанных изображений глаз из базы ММІ для одного из экземпляра паттерна “Моргание” (показаны только некоторые кадры)

пространстве и времени, поэтому при переходе от одного графа к другому изменения этих значений могут быть несовместимы с изменениями значений других вершин и в других доменах. Таким образом, при сопоставлении двух вершин неизбежно наличие нескольких *преобразований-кандидатов*, один из которых (или их уникальная комбинация) является истинной, т.е. наиболее приближенной к реальным изменениям в терминах некоторой метрики или функции ошибки. Иными словами, на практике от одного изображения к другому можно перейти различными способами, поэтому необходимо определить функции оптимизации и предложить ее решение.

Экспериментальные исследования показали следующее. Несмотря на то, что на  $i$ -м кадре ориентация некоторого объекта могла быть вычислена как  $\theta = 0^\circ$ , а на кадре  $i + 1$  как  $\theta = 30^\circ$ , истинное значение  $\Delta\theta$  могло быть неравно  $30^\circ$ , так как объект был изменен в других доменах (переместился в пространстве, изменилось распределение яркости, добавились новые свойства текстуры, такие как складки, или морщины во время улыбки или нахмуривания, каждая из которых обладает своими углами ориентации и др.).

Обычно в случаях неоднозначности преобразований используют вероятностные характеристики или методы линейной (нелинейной) оптимизации для поиска оптимальных значений, которые требуют выпуклости целевой функции. В настоящей работе для поиска оптимальных преобразований предложено применять взаимосвязь знаний из разных доменов и предположение об их согласованности.

Значения векторов изменения весов и значений вершин принадлежат пространству  $\mathbb{R}$ , но весовые коэффициенты являются нормализованными, что позволяет сопоставлять домены более корректно. Следовательно, связь любой пары доменов ( $D_q, D_{q+1}$ ) можно представить матрицей соответствия  $\mathbf{W}_{iq(q+1)} \in \mathbb{R}^{n \times n}$  между изменениями весов (рис. 4):

$$\Delta \mathbf{w}_{iD_q} = \mathbf{W}_{iq(q+1)} \Delta \mathbf{w}_{iD_{q+1}}. \quad (3)$$

Каждая строка матрицы  $\mathbf{W}$  будет отражать связь одного из элементов домена  $D_q$  с элементами всех уровней домена  $D_{q+1}$ , первая строка — связь корневого элемента и т.д.

Поскольку матрица  $\mathbf{W}$  содержит  $n^2$  неизвестных, а каждое матричное уравнение (3) формирует  $n$  уравнений, ее решение в явном виде будет возможно через  $n + 1$  кадров и потребует  $O(n^2(n + 1))$ , что для больших  $n$  практически нереализуемо, т.е. требуется применение методов аппроксимации или распараллеливания вычислений.

Предполагается, что существует такая оптимальная характеристическая матрица  $\mathbf{W}_{q(q+1)} = \text{const}$ , которая будет постоянна в независимости от номера кадра  $i$ , и тем самым отражает связь доменов  $D_q$  и

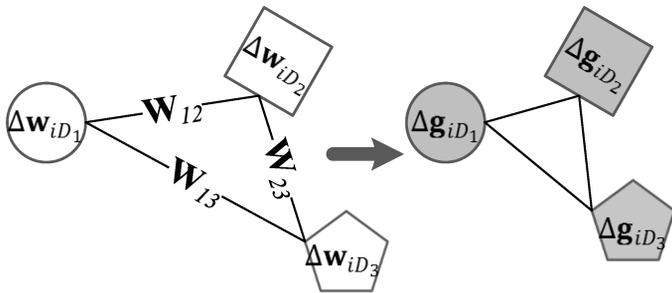


Рис. 4. Схема, иллюстрирующая идею вычисления преобразований на основе согласования доменов (частный случай для трех доменов)

$D_{q+1}$ . Ее вычисление возможно, так как число вариантов декомпозиций каждого элемента больше или равно единице и, соответственно, можно подобрать эти веса. Если число вариантов равно единице, то изменения весов будут отражаться в измененных значениях вершин на всех уровнях выше. Если число вариантов больше единице, то необходимо поочередно фиксировать значения весов.

Одним из решений задачи является ее сведение к квадратичному программированию [20]. Допустим, что число  $n$  небольшое и матрица  $\mathbf{W}$  вычислена после получения  $n + 1$  кадра. Рассмотрим случай, когда число полученных кадров больше  $n + 1$  и найти решение, удовлетворяющее (3) для всех предыдущих кадров, невозможно без корректировки изменений весовых коэффициентов  $\Delta \mathbf{w}_{iD_q}$ . Тогда формулируется задача обучения данных матриц так, чтобы определяемые на их основе весовые коэффициенты и преобразования вершин были максимально оптимально совместимы во всех доменах для каждого кадра в терминах следующей функции:

$$\arg \min_{\mathbf{w}, \mathbf{W}} \left( \left\| \Delta \mathbf{w}_{iD_q} - \Delta \mathbf{w}'_{iD_q} \right\| + \left\| \mathbf{W}_{q(q+1)} - \mathbf{W}'_{q(q+1)} \right\| \right),$$

где  $\|\mathbf{x}\| = \left( \sum_i x_i^2 \right)^{1/2}$  — норма (длина) вектора  $\mathbf{x}$  или  $\|\mathbf{X}\| = \left( \sum_{i,j} x_{i,j}^2 \right)^{1/2}$

в случае матрицы;  $\Delta \mathbf{w}_{iD_q}, \Delta \mathbf{w}_{iD_{q+1}}$  — изначальные вектора изменения весовых коэффициентов, вычисленные на основе формирования метаграфов для  $i$ -го кадра;  $\Delta \mathbf{w}'_{iD_q} = \mathbf{W}_{q(q+1)} \Delta \mathbf{w}_{iD_{q+1}}$  — вектор изменения весовых коэффициентов, полученный для текущего кадра на основе вычисленной матрицы  $\mathbf{W}_{q(q+1)}$ ;  $\mathbf{W}'_{q(q+1)}$  — одно из решений  $\Delta \mathbf{w}_{iD_q} = \mathbf{W}'_{q(q+1)} \Delta \mathbf{w}_{iD_{q+1}}$ . Такая задача представляет собой частный случай задачи поиском седловой точки [20, 21]. Ее экстремум может быть усредненным или локальным минимумом [21], если функция невыпукла, так как нормы разности на всех уровнях метаграфа (аргументов функции) идентично суммируются в функцию ошибки, что не всегда корректно аппроксимирует реальную функцию.

Поиск оптимальной матрицы  $\mathbf{W}_{(q+1)}$  требует теоретических и практических исследований и будет предложен в следующих работах. Далее подразумевается, что значения матрицы найдены и стабильны, на их основе можно рассчитать значения вершин и сформировать согласованную во всех доменах матрицу преобразований  $\mathbf{T}_i \in \mathbb{R}^{n \times Q}$ .

**Кластеризация паттернов.** Задача этого раздела — имея матрицу преобразований для  $N$  кадров  $\mathbf{T} \in \mathbb{R}^{n \times 2Q \times (N-1)}$ , предложить модель, позволяющую с помощью описания (1) определять факт наличия паттерна поведения, а с помощью матрицы (2) относить его к одному из известных типов или к некоторому новому типу, т.е. обобщать накопленные знания на новые последовательности изображения.

Важное допущение — постоянный или априори известный интервал между изображениями, измеряемый во временных единицах или номерах кадров. Это допущение аналогично постоянному интервалу между двумя точками изображения в пространстве, равному одному пикселю (или в других единицах). Для всех анализируемых последовательностей также подразумевается одинаковый размер изображения.

**Паттерн поведения как преобразования метаграфа.** Необходимо учитывать, что последовательности изображений могут иметь различную частоту съемки (fps), каждый человек обладает индивидуальными особенностями воспроизведения паттерна во времени и паттерны инерционны.

С позиции преобразований метаграфов паттерн поведения лица  $\mathbf{AU}_{C_1} \in C_1$  уникально определяется через непрерывную последовательность векторов преобразований на фиксированных позициях на интервале  $[i, i + n_{C_1}]$  из последовательности  $\Phi$  ( $i \in \{1, \dots, N - n_{C_1}\}$ ) в домене  $D_q$ :

$$\mathbf{AU}_{C_1 D_q} = \left[ \mathbf{T}_{i D_q}, \dots, \mathbf{T}_{(i+n_{C_1}-1) D_q} \right] \in \mathbb{R}^{n \times 2Q \times n_{C_1}},$$

где  $\mathbf{T}_{t D_q} = [\mathbf{T}_1, \dots, \mathbf{T}_n]^T \in \mathbb{R}^n$  — вектор преобразований от кадра  $t$  к  $t+1$  в домене  $D_q$ ;  $t \in \{i, \dots, i + n_{C_1}\}$ ;  $\mathbf{m}_{C_1} = [m_1, \dots, m_n]$  — вектор индексов, соответствующих индексам элементов векторов  $\Delta \mathbf{g}_{i D_q}, \Delta \mathbf{w}_{i D_q}$  (см. (2)), такой, что  $m_i = 0$ , если  $|\mathbf{T}_i| < \tau_0, 1$  иначе, вектор  $\mathbf{m}_{C_1}$  фиксирован в пределах кластера  $C_1$  с учетом функции ошибки;  $n_{C_1}$  — длительность паттерна, которая может варьироваться между экземплярами одного класса.

Синтезированный пример последовательности преобразований, в котором каждый паттерн порождает преобразования только в одном из доменов, показан на рис. 5, а. Однако на практике большинство паттернов порождают изменения как минимум в двух доменах (например, распределение яркости и текстура). Это тем более касается таких сложных паттернов, как “улыбка”, “выражение агрессии” и др.

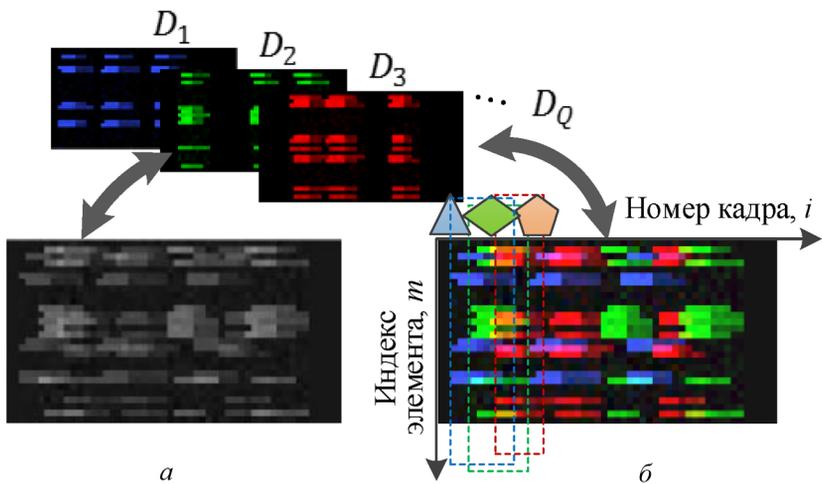


Рис. 5. Синтезированный пример векторов преобразований вершин  $g$  в одном из доменов с добавлением шума, вычисленных на основе изменений весовых коэффициентов (три класса паттернов, каждому из которых присвоен свой цвет; паттерны разных классов могут существовать одновременно на одном кадре на разных уровнях (индексах  $j, k$ ) метаграфа):

$a$  – однодоменная структура;  $b$  – мультидоменная структура

Вопрос определения “мультидоменных паттернов” оставим открытым для дальнейших исследований.

Сложность разделения паттернов в однодоменных моделях вызвана тем, что на интервале  $[i, i + n_{C_1}]$  могут существовать другие паттерны или их части  $AU_{C_2}$ , такие, что часть их индексов пересекается:  $m_{C_1} m_{C_2} \neq \mathbf{0}$  (рис. 5, б). Кроме того, на ограниченном интервале все индексы преобразований одного паттерна могут быть частью индексов другого  $m_{C_1} m_{C_2} = \{m_{C_1} \text{ или } m_{C_2}\}$  или полностью совпадать  $m_{C_1} m_{C_2} = \{m_{C_1} \text{ и } m_{C_2}\}$ . В таких случаях разделение паттернов требует дополнительных данных от других доменов. Поэтому домены должны быть выбраны с учетом того, что пересечения не могли происходить в пределах одного домена.

Однако следует учитывать шумовые составляющие при принятии окончательного решения о принадлежности паттерна к классу. Диапазон (амплитуда) абсолютных значений преобразований определяет интенсивность паттерна  $I_{C_1}$  (в случае системы FACS есть пять градаций  $A - E$  [7]) и не влияет на его принадлежность к тому или иному классу, знак преобразования может характеризовать класс.

**Поиск паттерна в последовательности изображений.** Поскольку структура метаграфа фиксирована для фиксированного размера изображения, для сравнения паттернов можно применять метрики на основе евклидовой нормы.

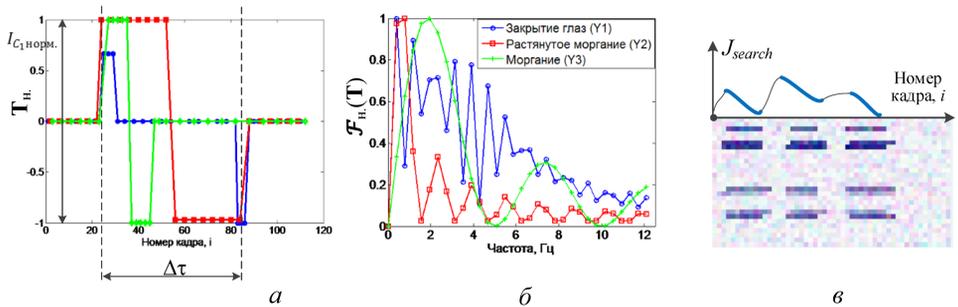
Функция поиска границ паттерна  $AU_{C_0}$  вычисляется для каждого  $n_{C_0} \in \{1, \dots, N - 1\}$  и номера кадра  $i \in \{1, \dots, N - n_{C_0}\}$  из последо-

вательности  $\Phi$ , пока не будет минимизирована следующая функция:

$$J_{search}(i, D_q, n_{C_0}) = \left\{ \frac{1}{n_{C_0 D_q}} + \sum_{i, n_{C_0 D_q}} \left( \alpha \|\Delta \mathbf{m}_{i D_q}\|_{N_1} + \beta \|\Delta \mathbf{T}_{i D_q}\| \right) \right\}, \quad (4)$$

где  $\Delta \mathbf{x}_{i D_q} = (|\mathbf{x}|_{(i+n_{C_0 D_q}) D_q} - |\mathbf{x}|_{i D_q})$ ; используется модуль вектора, так как преобразование может быть отрицательным;  $\|\mathbf{x}\|_{N_1} = \sum_i |x_i| / i$  — мера нормы, учитывающая предположение о том, что разность индексов более низкого уровня вносит меньший вклад в разность  $\Delta \mathbf{m}_{D_q}$ , чем разность индексов более высокого уровня;  $\alpha = (0, 1], \beta > 0$  — коэффициенты регуляризации. Функция (4) будет минимизирована, когда будет найден компромисс между длительностью паттерна и стабильностью значений преобразований и их позиций в метаграфе. Задание  $\alpha = 1$  запрещает любое отклонение индексов между двумя преобразованиями, коэффициент  $\beta$  имеет аналогичный смысл для разности значений преобразований и требует настройки или обучения (рис. 6) так же, как и порог абсолютного значения преобразования  $\tau_0$ .

Как только значения функции  $J_{search}$  начинают увеличиваться, поиск прекращается, и последнее значение  $n_{C_0 D_q}$  считается оптимальным. Далее находятся кандидаты паттернов в каждом домене (пары  $\langle i_{D_q}, n_{C_0 D_q} \rangle$ ), затем итоговое значение границ паттерна выбирается исходя из максимальной длительности среди всех доменов:  $(i, n_{C_0}) = \max_{D_q} [n_{C_0, D_1}, \dots, n_{C_0, D_Q}]$ . Если для следующих номеров  $i$  функция  $J_{search}$  начинает уменьшаться, то определяются границы следующего паттерна и т.д. (см. рис. 6).



**Рис. 6.** График нормализованных значений преобразований ( $T_n$ ) для паттернов  $Y_1, Y_2$  и  $Y_3$  (а), их Фурье-образы  $\mathcal{F}_n(T)$  (б), график функции поиска границ паттерна, построенного на основе синтезированной последовательности паттернов (в):

$\Delta\tau$  — интервал, на котором  $|\Delta T|_2 \approx 0$ ,  $I_{C_1 \text{ норм}}$  — нормализованная интенсивность паттернов  $Y_2$  и  $Y_3$ ;  $Y_1$  — паттерн “закрытие глаз”,  $Y_2$  — “растянутое моргание”,  $Y_3$  — “моргание”; на графике  $Y_1$  четко выделена фаза “глаза закрыты” ( $T = 0$ )

**Функция принятия решения о принадлежности паттерна к кластеру.** После того, как найдены такие границы, что  $n_{C_0} > 0$  для каждого номера кластера  $j$  рассчитывается бинарная функция добавления текущего паттерна  $\mathbf{AU}_{C_0}$  в кластер  $C_j$ :

$$J_{add}(\mathbf{AU}_{C_0}, j, D_q) = \text{dist}(\mathbf{AU}_{C_0}, j, D_q) - \tau_{jD_q}; \quad (5)$$

$$\text{dist}(\mathbf{AU}_{C_0}, j, D_q) = \left( \left\| \Delta \mathbf{m}_{C_0jD_q} \right\|_{N_1} + \gamma \left\| \varphi(\mathbf{T}_{C_0D_q}) - \varphi(\mu_{jD_q}) \right\| \right), \quad (6)$$

где  $\Delta \mathbf{m}_{C_0jD_q} = (\mathbf{m}_{C_0D_q} - \mathbf{m}_{jD_q})$ ,  $\varphi(\mathbf{x})$  — некоторое преобразование вектора  $\mathbf{x}$ . В качестве преобразования  $\varphi$  предложено использовать оператор Фурье. Если  $(J_{add} > 0) \equiv (J_{add} = 1)$  для текущего номера  $j$ , то набор преобразований и частотные характеристики отличаются от паттернов в коллекции на значение, большее порогового значения  $\tau_{jD_q}$ . Следовательно, необходимо добавить текущий паттерн  $\mathbf{AU}_{C_0}$  как экземпляр нового кластера  $C_0$ . Если  $(J_{add} \leq 0) \equiv (J_{add} = 0)$ , то  $\mathbf{AU}_{C_0} \in C_j$ , паттерн добавляется в кластер  $C_j$ , где он используется для реорганизации кластеров. Пороговые значения  $\tau_{jD_q}$  и коэффициенты нормализации  $\gamma$  определяют экспериментально.

Временные последовательности успешно сравниваются с помощью нечетких конечных автоматов [4, 9]. В настоящей работе последовательности преобразований полагаются уже структурированными, поэтому предложено использовать алгоритм, аналогичный сопоставлению осей времени (dynamic time warping, DTW), примененный в работе [8] с радиально-базисной функцией ядра, но в частотной области (dynamic frequency warping, DFW). Это обусловлено тем, что процессы, имеющую разную продолжительность и содержащие повторяющиеся паттерны поведения, целесообразнее исследовать частотными методами. Для проверки гипотезы рассмотрим два примера последовательностей (сессия № 119 и 205) базы MMI, на которых присутствуют паттерны двух разных классов ( $Y_1$  — “закрытие глаз”;  $Y_3$  — “моргание”). Без частотного анализа эти паттерны могут быть классифицированы как одинаковые. Нормированные значения синтезированных векторов преобразований  $\mathbf{T}$  для вершины метаграфа “глаза”, полученные по кадрам двух сессий, представлены на рис. 6, а, б.

Результаты исследования временного и частотного алгоритмов показывают, что они проявляют себя по-разному в зависимости от нормализации данных, В таблице в столбцах  $\mathbf{g}$  и  $\mathbf{g}_n$  представлены результаты алгоритма сравнения значений вершин и их нормализованных значений соответственно, в столбцах  $\mathbf{T}(\mathbf{g})$ ,  $\mathbf{T}(\mathbf{g}_n)$  и  $\mathbf{T}_n$  — результаты сравнения преобразований векторов  $\mathbf{g}$  или  $\mathbf{g}_n$  и их нормализованных значений соответственно,  $\mathcal{F}$  и  $\mathcal{F}_n$  — результаты сравнения преобразований Фурье и их нормализованных значений. При сравнении  $\mathbf{T}_n$  и  $\mathcal{F}_n$

нормализация вектора  $g$  и векторов  $g$  и  $T$  не влияет на результат. В целом, несмотря на результаты, гипотезу о преимуществе алгоритма частотного сравнения следует проверить на бóльшем объеме данных и с нелинейными функциями ядра. Для решения задачи сравнения паттернов следует рассмотреть возможность формирования метаграфов во временном и частотном доменах. Также возможно встраивание частотных значений для фильтрации паттернов и шумов в функцию  $J_{search}$ .

Результаты исследования алгоритмов временного и частотного сопоставления последовательностей

Сравниваемая пара паттернов (см. рис. 6)	Временной алгоритм					Частотный алгоритм						
	$g$	$g_n$	$T(g)$	$T(g_n)$	$T_n$	$\mathcal{F}(g)$	$\mathcal{F}(g_n)$	$\mathcal{F}(T(g))$	$\mathcal{F}(T(g_n))$	$\mathcal{F}(T_n)$	$\mathcal{F}_n(g)$	$\mathcal{F}_n(T)$
$Y_1 - Y_2$	17	0,24	6,3	0,35	33,7	0,31	0,31	8,30	0,0013	0,68	0,11	1,23
$Y_2 - Y_3$	10	<b>0,05</b>	<b>0,005</b>	<b>0,14</b>	<b>0,03</b>	0,68	<b>0,25</b>	8,39	<b>0,0003</b>	0,52	<b>0,07</b>	<b>0,47</b>
$Y_1 - Y_3$	<b>0,07</b>	0,07	3	0,39	12,7	<b>0,04</b>	1,02	<b>0,80</b>	0,0012	<b>0,07</b>	0,15	0,96

Примечание. Минимальные значения в столбцах, которые при правильном выборе алгоритма должны быть в строке  $Y_2 - Y_3$  выделены полужирным.

**Обучение функций  $J_{search}$  и  $J_{add}$ .** Одним из вариантов определения пороговых значений и коэффициентов в выражениях (5) и (6), т.е. обучения функции  $J_{add}$ , является использование баз последовательностей изображений, аннотированных экспертами. Среди таких баз следует выделить базы MMI [12] и СК+ [13]. База MMI содержит 2876 видеоизображений по несколько секунд ( $\sim 2 \dots 6''$ ) с  $fps = 25$  кадр/с ( $\sim 50 \dots 150$  кадров), на каждом из которых присутствует 1 – 6 из около 30 паттернов в зависимости от сложности выражения лица. База СК+ включает в себя приблизительно 700 последовательностей изображений, в каждой из которых около 20 кадров. Суммарное число представленных субъектов, выполняющих паттерны, составляет приблизительно 200 человек.

Часть базы MMI (329 видеоизображений) также имеют пок кадровые описания фаз паттернов: нарастание (onset); пик (apex); спад (offset). Данные материалы могут применяться для обучения функции поиска границ  $J_{search}$  (4).

**Реорганизация паттернов.** Под реорганизацией понимается выделение общих паттернов, вычисление новых центров кластеров  $\mu_{jD_q}$  и пороговых значений  $\tau_{jD_q}$ , определяющих их границы, в каждом домене (рис. 7). Для разделения паттернов на классы (кластеры) применяются алгоритмы  $k$ -средних [4] и его нелинейные и взвешенные расширения [8].

Центр кластера  $C_j$  в домене  $D_q$  рассчитывается как паттерн  $\mu_{jD_q} = AU_{C_j}$ , расстояние от которого до других паттернов минимально в пределах кластера:  $\arg \min_{AU_{C_j}} \text{dist}(AU_{C_j}, j, D_q)$ .

Пороговое значение определяется как

$$\tau_{jD_q} = 2 \max_{AU_{C_j}} \text{dist}(AU_{C_j}, j, D_q),$$

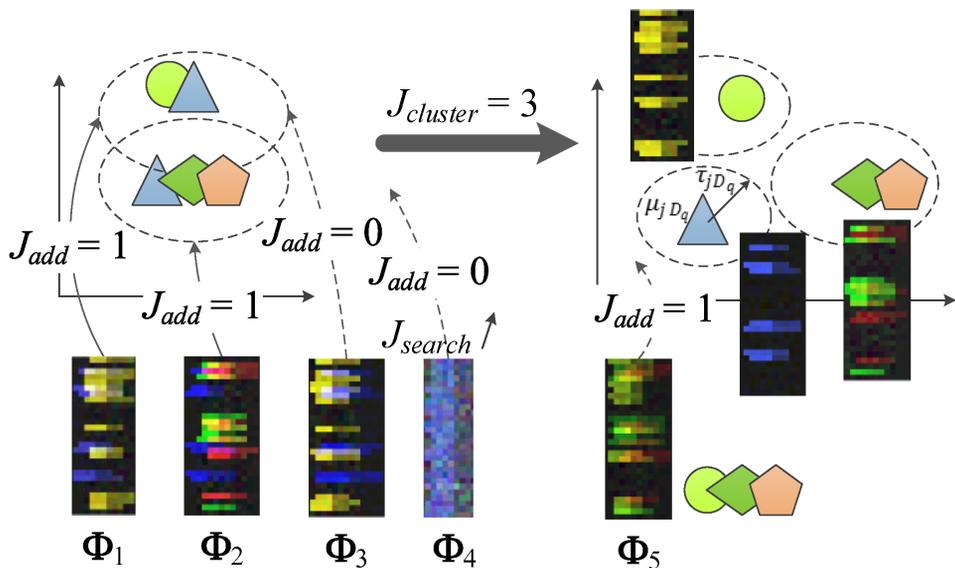


Рис. 7. Схема добавления нового кластера и реорганизации кластеров

но так, чтобы кластеры не пересекались. Паттерн выделяется как общий с центром  $\mu_{iD_q}$  и пороговым значением  $2\tau_{iD_q}$ , если для двух паттернов  $\mathbf{AU}_{C_i}, \mathbf{AU}_{C_j}, i \neq j$ , существует функция

$$\text{dist}(\mathbf{AU}_{C_i}, \mathbf{AU}_{C_j}, D_q) = 2\tau_{iD_q} < \min(\text{dist}(\mathbf{AU}_{C_i}, k, D_q), \text{dist}(\mathbf{AU}_{C_j}, m, D_q)) \quad \forall k, m,$$

где  $k, m$  — номера кластеров, т.е. два паттерна ближе друг к другу, чем каждый из них к любому паттерну из другого кластера.

**Заключение.** Получена методика, объединяющая концепцию метаграфа и методы пространственно-временного и визуального анализа с помощью мультидоменной взаимосвязи и функций оптимизации, для решения задачи более эффективного исследования видеоархивов с поведением человека и экспериментальной проверки паттернов системы кодирования FACS. Основная идея методики — преобразования метаграфов и их взаимосвязь между различными доменами — используется впервые для решения указанной задачи. Программная реализация методики позволит устранить или сгладить недостатки, частично или полностью присутствующие в предыдущих работах (недостаточная точность для использования на реальных видеоархивах, вызванная чувствительностью алгоритмов к сложному фону, поворотам лица, неравномерной освещенности и другим искажениям объекта).

Недостатки методики: сложность обучения модели по последовательностям изображений с частотой менее 20 кадр/с, отсутствие прямой возможности распознавания паттернов на статическом изображении; необходимость экспертного труда на начальных стадиях обу-

чения модели и на стадии фильтрации; сопоставление полученных паттернов.

В дальнейшем интерес представляет определение оптимальной арности и глубины графа, исследование характеристик матрицы соответствия доменов, функций поиска паттерна, его добавления в кластер и реорганизации паттернов, а также экспериментальная проверка соответствия полученных кластеров системе кодирования FACS, после которой можно будет сделать вывод о полноте и объективности кодировки FACS. Результаты настоящей работы могут быть полезны не только для анализа поведения лица человека, но и для анализа паттернов в других областях знаний.

## ЛИТЕРАТУРА

1. *Ануашвили А.Н.* Основы объективной психологии. Международный институт управления, психологии и психотерапии. М.; Варшава, 2005. [Электронный ресурс]. Режим доступа: <http://anuashvili.ru> (дата обращения 02.05.2014).
2. *Kanade T.* Visual Processing and Understanding of Human Faces and Bodies, 9th International Conference (ICVS 2013), Jul. 2013, Keynote Talk.[Электронный ресурс] Режим доступа: [http://workshops.acin.tuwien.ac.at/ICVS/downloads/Kanade\\_ICVS2013.pdf](http://workshops.acin.tuwien.ac.at/ICVS/downloads/Kanade_ICVS2013.pdf). (дата обращения 19.02.2013).
3. *Ekman P., Rosenberg E.L.* What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System. N.Y.: Oxford University Press, 2005. 639 p.
4. *Алфимцев А.Н.* Разработка и исследование методов захвата, отслеживания и распознавания динамических жестов: Дис... канд. техн. наук. М., 2008, 167 с.
5. *Bartlett M.S., Whitehill J.* Automated facial expression measurement: Recent applications to basic research in human behavior, learning, and education // Oxford Handbook of Face Perception, Oxford University Press, 2011. P. 489–514.
6. *Романова Н.М., Рытик А.П., Самохина М.А., Скрипаль А.В., Усанов Д.А.* Особенности глазодвигательных реакций человека при произнесении истинной и ложной информации // Психология. СГУ. 2008. С. 65–73.
7. *Ekman P., Friesen W.* Facial Action Coding System: A Technique for the Measurements of Facial Movements. Consulting Psychologists Press, 1978.
8. *Zhou F., Simon T., de la Torre F., Cohn J.F.* Unsupervised discovery of facial events// Technical Report CMU-RI-TR-10-10, Carnegie Mellon University, 2010. P. 1–20.
9. *Девятков В.В., Лычков И.И.* Моделирование и анализ ситуаций в виртуальной среде движущихся объектов // Вестник МГТУ им. Н.Э. Баумана. Сер. Приборостроение. 2013. // № 3. С. 26–42.
10. *Sikka K., Dykstra K., Sathyanarayana S., Littlewort G., Bartlett M.* Multiple kernel learning for emotion recognition in the wild // Proceedings of the 15th ACM on International conference on multimodal interaction (ICMI'13). ACM, New York, USA. P. 517–524.
11. *Koelstra S., Pantic M., Patras I.* A Dynamic Texture Based Approach to Recognition of Facial Actions and Their Temporal Models // IEEE Trans. on Pattern Analysis and Machine Intelligence. 2010. P. 1940–1954.
12. *Valstar M.F., Pantic M.* Induced Disgust, Happiness and Surprise: an Addition to the MMI Facial Expression Database // Proc. International Language Resources and Evaluation Conference. Malta. 2010. P. 65–70.

13. Lucey P., Cohn J., Kanade T., Saragih J., Ambadar Z., Matthews I. The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression // Proc. 2010 IEEE Computer Society Conference on CVPR Workshops. 2010. P. 94–101.
14. Knyazev B. Human nonverbal behavior multi-sourced ontological annotation // Proc. International Workshop on Video and Image Ground Truth in Computer Vision Applications (VIGTA '13). Jul. 2013. Article 2. P. 1–8.
15. Кауанова Л.Х., Латышева Е.Ю., Спиридонов И.Н. Алгоритм распознавания эмоционального состояния по изображениям лица с использованием дискриминантного анализа и фильтров Габора // Медицинская техника. № 3. 2012. С. 1–4.
16. Wiskott L., Fellous J.-M., Kruger N., von der Malsburg C. Face Recognition by Elastic Bunch Graph Matching // IEEE Trans. on Pattern Analysis and Machine Intelligence. Vol. 17. No. 7. Jul. 1997. P. 775–779.
17. Basu A., Blanning R.W. Metagraphs and Their Applications // Integrated Series in Information Systems. 2007. Vol. 15. No. VIII. 172 p.
18. Скурихин А.В. Рекурсивно-иерархическое представление одномерных фракталоподобных сигналов // Тр. СПИИРАН. 2003. № 1 (3). С. 107–117.
19. Болотова Ю.А., Спицын В.Г., Фомин А.Э. Применение модели иерархической временной памяти в распознавания изображений // Известия Томского политехнического университета. 2011. Т. 318. № 5. С. 60–63.
20. Интрелизатор М. Математические методы оптимизации и экономическая теория. М.: Прогресс, 1975. 607 с.
21. Christopher M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

## REFERENCES

- [1] Anuashvili A.N. Osnovy ob'ektivnoj psihologii. Mezhdunarodnyy institut upravleniya, psihologii i psikhoterapii. 4-oe izd [Fundamentals of objective psychology. Int. Inst. Management, Psychol. Psychoth. 4th ed.]. Moscow, Warsaw, 2005. Available at: <http://anuashvili.ru> (accessed 02.05.2014).
- [2] Kanade T. Visual Processing and Understanding of Human Faces and Bodies. 9th Int. Conf. (ICVS 2013), 2013, Keynote Talk. Available at: [http://workshops.acin.tuwien.ac.at/ICVS/downloads/Kanade\\_ICVS2013.pdf](http://workshops.acin.tuwien.ac.at/ICVS/downloads/Kanade_ICVS2013.pdf) (accessed 19.02.2013).
- [3] Ekman P., Rosenberg E.L. What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System. N.Y., Oxford University Press, 2005. 639 p.
- [4] Alfimcev A.N. Razrabotka i issledovanie metodov zahvata, otslezhivaniya i raspoznavaniya dinamicheskikh zhestov. Diss. kand. tekhn. nauk [Development and study of methods of capture, tracking and recognition of dynamic gestures. Cand. tech. sci. diss.]. Moscow, 2008. 167 p.
- [5] Bartlett M.S., Whitehill J. Automated facial expression measurement: Recent applications to basic research in human behavior, learning, and education. Oxford Handbook of Face Perception, Oxford University Press, 2011, pp. 489–514.
- [6] Romanova N.M., Rytik A.P., Samohina M.A., Skripal' A.V., Usanov D.A. The eye-moving reaction features of the person due to pronouncing true and false information. *Psihologija* [Psychology], SGU, 2008, pp. 65–73 (in Russ.).
- [7] Ekman P., Friesen W. Facial Action Coding System: A Technique for the Measurements of Facial Movements. Palo Alto, CA: Consulting Psychologists Press, 1978.
- [8] Zhou F., Simon T., de la Torre F., Cohn J.F. Unsupervised discovery of facial events. *Technical Report CMU-RI-TR-10-10*, Carnegie Mellon University, 2010, pp. 1–20.

- [9] Devjatkov V.V., Lychkov I.I. Simulation and the analysis of situations in the virtual environment of moving objects. *Vestn. Mosk. Gos. Tekh. Univ. im. N. E. Baumana, Priborost.* [Herald of the Bauman Moscow State Tech. Univ., Instrum. Eng.], 2013, no. 3, pp. 26–42 (in Russ.).
- [10] Sikka K., Dykstra K., Sathyanarayana S., Littlewort G., Bartlett M. Multiple kernel learning for emotion recognition in the wild. *Proc. 15th ACM on Int. Conf. on Multimodal Interaction (ICMI '13)*, ACM, New York, USA, pp. 517–524.
- [11] Koelstra S., Pantic M., Patras I. A Dynamic Texture Based Approach to Recognition of Facial Actions and Their Temporal Models. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2010, pp. 1940–1954.
- [12] Valstar M.F., Pantic M. Induced Disgust, Happiness and Surprise: an Addition to the MMI Facial Expression Database. *Proc. Int. Language Resources and Evaluation Conf.*, Malta, 2010, pp. 65–70.
- [13] Lucey P., Cohn J., Kanade T., Saragih J., Ambadar Z., Matthews I. The Extended Cohn-Kanade Dataset (CK+): a complete dataset for action unit and emotion-specified expression. *Proc. IEEE Comp. Soc. Conf. on CVPR Workshops*, 2010, pp. 94–101.
- [14] Knyazev B. Human nonverbal behavior multi-sourced ontological annotation. *Proc. Int. Workshop on Video and Image Ground Truth in Comp. Vision Appl. (VIGTA '13)*, 2013, Article 2, pp. 1–8.
- [15] Kashapova L.H., Latysheva E.Ju., Spiridonov I.N. Discriminant Analysis of Two-Dimensional Gabor Features for Facial Expression Recognition. *Meditinskaya Tekhnika* [Biomedical Engineering, 2012, iss. 3, vol. 46, pp. 89–92], 2012, vol. 46, no. 3, pp. 1–4 (in Russ.).
- [16] Wiskott L., Fellous J.-M., Kruger N., von der Malsburg C. Face Recognition by Elastic Bunch Graph Matching. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 1997, vol. 17, no. 7, pp. 775–779.
- [17] Basu A., Blanning R.W. *Metagraphs and Their Applications. Integrated Series in Information Systems*. Springer, 2007, vol. 15, no. VIII, 172 p.
- [18] Skurihin A.V. Recursive and hierarchical representation of one-dimensional fractal-like signals. *Tr. S.-Peterburgskogo Ins. Inf. i Avtomat., RAN (SPIIRAN)* [Proc. SPb Inst. Inf. and Autom., Russ. Ac. Sc. (SPIIRAS)], 2003, no. 1 (3), pp. 107–117 (in Russ.).
- [19] Bolotova Ju.A., Spicyn V.G., Fomin A.Je. Model application of hierarchical temporary memory in recognitions of images. *Izv. Tomskogo Politehnicheskogo Un.* [Bulletin of the Tomsk Polytechnic Un.], 2011, vol. 318, no. 5, pp. 60–63 (in Russ.).
- [20] Intriligator M.D. *Mathematical Optimization and Economic Theory. Series: Classics in Applied Mathematics (Book 39)*. Prentice-Hall, 1971. 571 p. (Russ. Ed.: Intriligator M. *Matematicheskie metody optimizacii i jekonomicheskaja teorija*. Progress Publ., 1975. 607 p.).
- [21] Bishop Chr. M. *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Springer-Verlag, New York, Inc., Secaucus, NJ, USA, 2006.

Статья поступила в редакцию 24.02.2014

Борис Александрович Князев — аспирант кафедры “Системы обработки информации и управление” МГТУ им. Н.Э. Баумана, инженер НИИЦ БТ МГТУ им. Н.Э. Баумана. Автор девяти научных работ в области искусственного интеллекта, компьютерного зрения, параллельных вычислений.  
МГТУ им. Н.Э. Баумана, Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5.

B.A. Knyazev — post-graduate of “Systems of Information Processing and Management” department of the Bauman Moscow State Technical University, engineer of the Research and Testing Center for Biometric Technology of the Bauman Moscow State Technical University. Author of nine publications in the field of artificial intelligence, computer vision, and parallel computing.

Bauman Moscow State Technical University, Vtoraya Baumanskaya ul. 5, Moscow, 105005 Russian Federation.

Валерий Михайлович Черненький — д-р техн. наук, заведующий кафедрой “Системы обработки информации и управления” МГТУ им. Н.Э. Баумана. Автор 105 научных работ в области теории описания параллельно-последовательных процессов, имитационного моделирования информационных систем, проектирования автоматизированных систем организационного управления.

МГТУ им. Н.Э. Баумана, Российская Федерация, 105005, Москва, 2-я Бауманская ул., д. 5.

V.M. Chernenkiy — Dr. Sci. (Eng.), head of “Systems of Information Processing and Management” department of the Bauman Moscow State Technical University. Author of 105 publications in the field of description theory of parallel and serial processes, simulation of information systems, designing of automated systems of organizational management.

Bauman Moscow State Technical University, Vtoraya Baumanskaya ul. 5, Moscow, 105005 Russian Federation.