

# Разработка метода навигации по коллекциям цифровых изображений\*

© Мясников Евгений Валерьевич

Самарский Государственный Аэрокосмический Университет  
mevg@smr.ru

## Аннотация

Центральным вопросом при построении систем навигации является способ отображения коллекции в двумерное навигационное пространство. Суть предлагаемого в работе метода состоит в реализации отображения в два этапа. На первом этапе по коллекции изображений строится система иерархически вложенных кластеров. На втором этапе производится отображение исходного пространства описания изображений в пространство навигации.

В работе дан обзор методов, используемых при построении систем навигации, приведены результаты исследований предложенного метода и сравнения его с известными методами. Полученные результаты позволяют сделать вывод о возможности успешного применения разработанного метода.

## 1 Введение

Поиск изображений по содержанию в последнее время является актуальной областью исследований. Однако поиск изображений по шаблону или в соответствии с какими-либо параметрами запроса далеко не всегда приводит пользователя к желаемому результату. Кроме того, пользователь часто желает просто посмотреть, какие изображения могут находиться в цифровой библиотеке. И в том, и другом случае пользователю необходимо переходить от одного изображения к другому. Тогда мы имеем дело уже не с поиском изображений по содержанию, а с навигацией на основе содержания. Система навигации может служить прекрасным дополнением к системе поиска изображений, использоваться отдельно в коллекциях цифровых изображений домашнего пользования или в рамках создания пользовательских интерфейсов интернет-поисковых систем.

Постановка самой задачи навигации по

коллекциям цифровых изображений является в значительной степени новой, оригинальной, и, с учетом постоянно растущих объемов цифровых данных, актуальной.

В настоящей работе для построения систем навигации предлагается использовать новый метод отображения исходного пространства описания изображений в пространство навигации. Метод основан на совместном использовании методов автоматической классификации и понижения размерности данных.

Работа организована следующим образом. Во втором разделе рассмотрены методы, используемые при построении систем навигации по коллекциям цифровых изображений, дается обзор известных систем навигации.

Третий раздел посвящен описанию предлагаемого подхода. Рассматриваются все исследуемые методы кластеризации и снижения размерности пространства.

Результаты экспериментального исследования предложенного метода приводятся в четвертом разделе. Полученные результаты содержат как оценки качества предложенного метода, так и результаты сравнения с известными методами.

Пятый раздел посвящен вопросу выбора признаков, используемых в данной работе.

В конце работы дается заключение, содержащее перечень основных ее результатов.

## 2 Обзор современного состояния исследований

Центральным вопросом при построении систем навигации по коллекциям цифровых изображений является способ построения отображения коллекции изображений в двумерное пространство навигации. Существует целый ряд средств, позволяющих отображать изображения коллекции, упорядочивая изображения по времени создания или другой сопутствующей информации. Альтернативный подход к отображению состоит в извлечении из изображений каких-либо признаков и размещении изображений на экране в соответствии со значениями признаков. Так как размерность пространства признаков может в десятки и сотни раз превышать размерности пространства отображения, то для создания двумерных

(трехмерных) отображений необходимо применять методы снижения размерности. Такие методы обычно подразделяют на линейные и нелинейные.

Линейные методы используют дискретный вариант разложения Карунена – Лоэва, называемый также методом главных компонент (PCA - Principal component analysis). В этом методе осуществляется поворот системы координат в исходном пространстве признаков таким образом, чтобы в проекции на новые оси – главные компоненты – дисперсия всего множества точек была максимальна. При этом дисперсия сосредоточена большей частью в первых компонентах, что позволяет рассматривать только их, отбрасывая остальные. Метод PCA был использован при создании относительно небольшого числа систем [8, 13].

Наиболее распространенным классом методов нелинейного снижения размерности пространства является многомерное масштабирование (шкалирование) – MDS (multidimensional scaling).

Под MDS понимают техники, с помощью которых производится отображение множества векторов многомерного пространства в пространство малой размерности (как правило, двух- или трехмерное) с сохранением, по возможности, расстояний между ними.

Все MDS алгоритмы пытаются минимизировать функцию потерь, характеризующую величину рассогласования расстояний между первоначальными и полученными векторами в пространстве малой размерности. Наиболее часто функция потерь определяется как

$$E = \sum_{i < j} \frac{(d_{ij} - d_{ij}^*)^2}{d_{ij}^2}, \quad (1)$$

где  $d_{ij}$  – расстояние между объектами  $i$  и  $j$  в исходном пространстве,  $d_{ij}^*$  – расстояние между объектами  $i$  и  $j$  в пространстве малой размерности.

Как правило, алгоритмы MDS имеют итерационный характер, и в качестве начального приближения используются случайные значения. Вследствие этого, MDS алгоритмам свойственно останавливаться в локальных минимумах, порождая каждый раз различные представления в навигационном пространстве.

В работе [16] было предложено запускать MDS алгоритм для каждого набора входных данных некоторое малое число раз (5-10) и выбирать наилучшую конфигурацию в соответствии с используемым критерием. В [16] выделяется 2 класса алгоритмов, в зависимости от того, является ли порождаемое пространство навигационное непрерывным или дискретным. К первому типу алгоритмов относят алгоритмы, использующие итеративную процедуру Ньютона-Рапсона, и алгоритмы, использующие процедуру имитации отжига.

Второй класс алгоритмов MDS работает на дискретной сетке, что, по мнению ряда

исследователей, больше удовлетворяет потребности пользователей, чем случай непрерывного пространства отображения, приводящий к перекрытию иконок на экране. Очевидным способом получения отображения на дискретную сетку является привязка к дискретной сетке результата работы, полученного в непрерывном пространстве. Однако в этом случае могут возникать конфликты, когда несколько изображений должны быть привязаны к одной ячейке сетки. Некоторые стратегии разрешения таких конфликтов предложены в [17]. Из работы [16] известен также “жадный” (greedy) алгоритм, позволяющий получить приемлемый результат за короткое время. Для получения отображения на дискретное пространство навигации был также использован генетический алгоритм. Этот алгоритм позволил исследователям получить существенно меньшие значения функции потерь, однако время его работы оказалось существенно выше, чем для “жадного” алгоритма.

Следует отметить, что на практике алгоритмы MDS применимы к небольшим множествам объектов (порядка сотни) из-за вычислительных ограничений. По той же причине рекомендуется вычислять матрицу расстояний между объектами заранее.

Существует отдельный класс алгоритмов многомерного масштабирования, использующих силовые методы укладки графов. Работа этих алгоритмов основана на математических моделях механических процессов. Наиболее известными являются модели Фрухтермана - Рейнгольда [7] и Камада-Кавайи [9].

В работе [20] метод MDS применен для размещения на плоскости результатов поиска изображений. Интересной стороной работы является использование цветовых сигнатур (color signatures) для описания изображений и введение экскаваторного расстояния (Earth Mover’s Distance - EMD) между сигнатурами, трактуемого как минимальное количество работы, необходимое для преобразования одной сигнатуры в другую. К сожалению, пространство, образуемое сигнатурами, не является линейным. Это создает препятствие использованию такого метода снижения размерности, как PCA.

В работе [12] MDS подход использован для размещения на экране различных версий одного изображения, сгенерированных автоматически путем варьирования вектора параметров. Система, описанная в работе, предназначена для принятия решения о том, какие наборы параметров более всего удовлетворяют пользователя – специалиста по компьютерной графике.

При навигации по большим коллекциям изображений возникают различные проблемы. В частности, попытка снизить размерность пространства признаков для большого количества объектов занимает непомерно большое количество времени. Кроме того, отображение всех объектов на

экране одновременно просто не представляется возможным. Вследствие этого, исследователями предпринимаются попытки построения иерархических структур, группирующих объекты в соответствии с их характеристиками. Наиболее очевидным способом выполнить такого рода разбиение является кластеризация.

В системе [3] реализован подход, основанный на иерархической кластеризации. На каждом шаге пользователю представляется набор изображений, каждое из которых является представителем кластера. Выбирая одно из представленных изображений, пользователь переходит к следующему уровню иерархии, и процесс повторяется. Организация базы данных имеет иерархическую структуру квадродерева. В работе использован иерархический агломеративный алгоритм кластеризации.

Возможен также подход, при котором на каждом шаге просмотра пользователь выбирает несколько кластеров. Далее, все объекты, входящие в данные кластеры, кластеризуются заново, представляя пользователю новый набор кластеров, и процесс повторяется. Описанный принцип (Scatter/Gather), был использован в работах [23, 4]. Изначально он был разработан для просмотра коллекций текстовых документов.

В работе [11] использованы нейронные сети Кохонена (так называемые самоорганизующиеся карты – SOM), организованные в иерархическом порядке. Полученная иерархическая структура сохраняется и используется при поиске в базе данных. Изображения располагаются в узлах дискретной сетки в соответствии с их схожестью.

В системе AutoAlbum [15] все изображения автоматически разбиваются на альбомы, образуя, таким образом, двухуровневую иерархию. В системе используется два вида кластеризации: кластеризация по времени создания фотографий и кластеризация по содержанию фотографий. Для кластеризации используется вероятностный алгоритм кластеризации (best-first model merging), в качестве вероятностной модели применяется скрытая марковская модель (HMM). Данный алгоритм позволяет работать на упорядоченных множествах данных и схож с агломеративными методами кластеризации. Отображение фотографий для пользователя внутри кластера осуществляется в соответствии со временем создания.

Система PhotoTOC [14] является дальнейшим развитием системы AutoAlbum. В PhotoTOC используются те же основные алгоритмы, улучшен пользовательский интерфейс и изменен алгоритм выбора представителя кластера (представитель выбирается на основании минимума дивергенции Кульбека-Либера (Kullback-Lieber) между гистограммами фотографии и центра кластера).

подавляющее большинство систем навигации по коллекциям изображений производят кластеризацию или строят двумерные отображения

исключительно на основе примитивных признаков. Очевидно, что это является их серьезным недостатком, так как визуальная схожесть изображений на примитивном уровне далеко не всегда соответствует семантической схожести, оцениваемой человеком. На этот недостаток указывает ряд исследователей. Следует отметить, что данный недостаток на сегодняшний день не может быть преодолен из-за так называемого семантического провала [5].

Однако можно привести ряд работ, в которых предприняты попытки совместить текстовые аннотации к изображениям с признаками. В работе [10] рассматривается возможность совместной кластеризации (метод порождает совместное распределение слов и элементов изображений). В системе ANVIL [19] пользователь может ввести текстовый запрос, обрабатываемый процессором естественного языка, в ответ на который выводится ряд изображений, расположенных следующим образом: в центре находится изображение, ранжированное выше всех остальных, другие результаты расположены вокруг него на расстояниях, отражающих их релевантность к запросу. В системе также используется кластеризация альтернативных результатов, в соответствии со схожестью заголовков.

Еще одним направлением исследований является постепенное уточнение запроса к поисковым системам. В работе [21] описывается архитектура и принципы функционирования поисковой системы El Nino. В системе совмещается поиск по визуальным признакам и текстовым аннотациям. Найденные изображения располагаются на плоскости в соответствии с визуальными характеристиками. Интересной особенностью системы является возможность улучшения запроса как путем пометки релевантных и нерелевантных элементов, так и путем изменения конфигурации элементов на плоскости в соответствии с их релевантностью.

### 3 Описание предлагаемого подхода

Как показал обзор литературы, существующие системы навигации по коллекциям изображений строятся на основе двух различных групп методов. К первой группе относятся методы кластеризации, позволяющие разбить всю коллекцию изображений на небольшие группы, внутри которых разница между изображениями невелика. Ко второй группе относятся методы снижения размерности пространства признаков, позволяющие расположить элементы коллекции на плоскости.

Для построения более эффективных средств навигации представляется целесообразным разработать и исследовать средства совмещения методов, относящихся к указанным группам. Однако именно этому вопросу исследователями пока не было уделено достаточно внимания.

В данной работе проведено исследование комбинированного алгоритма, функционирующего на основе широко известного метода снижения размерности Сэммона и нейросетевого метода кластеризации WTA. В разделе 3.1 приведено краткое описание известных методов кластеризации, используемых в настоящей работе. В разделе 3.2. приводятся описания известных линейного и нелинейного методов снижения размерности пространства, а также предлагаемого двухэтапного метода, функционирующего на их основе. В разделе 3.3 приведено описание предлагаемого комбинированного метода, функционирующего на основе методов снижения размерности и кластеризации. Экспериментальные исследования методов, описанных в настоящем разделе, приведены в разделе 5.

### 3.1 Методы кластеризации

Для реализации этапа кластеризации коллекции изображений и построения иерархии вложенных кластеров в процессе работы были рассмотрены как алгоритмы иерархической агломеративной кластеризации, так и нейросетевой алгоритм WTA.

В качестве алгоритмов иерархической кластеризации рассматривались алгоритмы Single Link, и Complete Link [22]. В общих чертах функционирование этих алгоритмов может быть описано следующей схемой:

1. Вычисление матрицы попарных расстояний между объектами.
2. Отнесение каждого объекта к отдельному кластеру.
3. Слияние пары элементов, расстояние между которыми минимально.
4. Удаление строки и столбца поглощенного кластера и пересчет матрицы.
5. Проверка остановочного критерия и переход к шагу 3.

В алгоритме Single Link расстояние между кластерами определяется как минимальное расстояние между парой объектов в кластерах. В алгоритме Complete Link – как максимальное расстояние. В качестве остановочного критерия обычно рассматривается максимальное число объектов в кластере. В данной работе остановка алгоритма производится при достижении заданного количества кластеров.

Нейросетевой кластеризационный алгоритм WTA функционирует на базе нейронной сети Кохонена [25], представляющей собой однослойную сеть, в которой каждый нейрон соединен со всеми компонентами входного вектора. При подаче на вход сети в момент времени  $t$  вектора  $\mathbf{x}(t)$  в конкурентной борьбе побеждает тот нейрон из общего числа  $K$ , веса которого в наименьшей степени отличаются от соответствующих компонент входного вектора. Для нейрона-победителя  $v$  выполняется соотношение

$$d(\mathbf{x}(t), \mathbf{w}_v(t)) = \min_{1 \leq i \leq K} d(\mathbf{x}(t), \mathbf{w}_i(t)).$$

Здесь  $d(\mathbf{x}(t), \mathbf{w}(t))$  - расстояние между входным вектором  $\mathbf{x}(t)$  и текущим состоянием синаптических весов  $\mathbf{w}_i(t)$  нейрона  $i$ . Выходной сигнал  $y_v$  выигравшего нейрона устанавливается в единицу, выходы остальных нейронов  $y_i$  становятся равными нулю. Нейрон-победитель подвергается адаптации, в ходе которой векторы весов изменяются в направлении вектора  $\mathbf{x}(t)$  по правилу коррекции:

$$\mathbf{w}_v(t+1) = \mathbf{w}_v(t) + \eta(t)[\mathbf{x}(t) - \mathbf{w}_v(t)].$$

Здесь  $\eta(t)$  - коэффициент обучения в момент времени  $t$ , значение которого со временем уменьшается.

Основными параметрами алгоритма являются [24]:

- начальное значение коэффициента обучения  $\eta_0$ ,  $\eta$  изменяется по формуле:

$$\eta(t) = \frac{\eta_0}{1 + \frac{t}{\tau}}, \quad \eta_0 \in (0,1)$$

- время действия механизма утомления, определяющее, какое количество циклов обучения действует штрафование часто побеждающих нейронов путем искусственного завышения расстояния пропорционально количеству побед ( $W_i := W_i * d(\mathbf{x}, \mathbf{w}_i)$ ) при определении нейрона-победителя,
- максимальное количество итераций служит критерием остановки независимо от того, произошла к этому моменту стабилизация весов сети или нет.

В работе было проведено исследование всех приведенных алгоритмов кластеризации. Результаты исследования приводятся в разделе 4.1.

### 3.2 Линейный и нелинейный методы снижения размерности пространства

Для *линейного преобразования* пространства в работе используется дискретный вариант преобразования Карунена-Лоэва, реализуемый нейронной сетью PCA [25] с применением правила Сенгера. Сеть PCA является однослойной сетью, количество нейронов в которой равно размерности выходного пространства. Нейроны располагаются в одном слое, функции активации нейронов – линейные. Каждый нейрон сети связан своими синаптическими связями со всеми компонентами входного вектора и при подаче на вход сети в момент времени  $t$  сигнала  $\mathbf{x}$  выполняет взвешенное суммирование данного сигнала:

$$y_i(t) = \sum_{j=0}^N w_{ij}(t) \cdot x_j(t)$$

Уточнение весов сети выполняется в соответствии с правилом Сенгера:

$$w_{ij}(t+1) = w_{ij}(t) + \eta(t) \cdot y_i(t) \cdot$$

$$\left( x_j(t) - \sum_{h=1}^i w_{hj}(t) \cdot y_h(t) \right)$$

При обучении сети используется адаптивный алгоритм изменения коэффициента обучения. В процессе обучения входная выборка предъявляется на вход сети многократно, вплоть до стабилизации весов.

Для *нелинейного преобразования* пространства в предлагаемом методе взят за основу алгоритм двумерного отображения, описанный в работе [26] и впервые примененный Сэммоном в 1969 году. Данный алгоритм позволяет минимизировать ошибку представления многомерных данных, выражаемую в виде:

$$\varepsilon = \frac{1}{\sum_{i < j} d_{ij}} \cdot \sum_{i < j} \frac{(d_{ij} - d_{ij}^*)^2}{d_{ij}} \quad (2)$$

Здесь  $d_{ij}$  и  $d_{ij}^*$  - расстояние между объектами  $i$  и  $j$  соответственно в многомерном и двумерном пространстве.

Работа алгоритма имеет итеративный характер, связанный со следующим рекуррентным соотношением для координат в двумерном пространстве  $y_{jk}$ :

$$y_{ik}(t+1) = y_{ik}(t) + \frac{2 \cdot \alpha}{N} \cdot \sum_{i < j} d_{ij} \cdot \sum_{\substack{j=1 \\ j \neq i}}^N \frac{d_{ij} - d_{ij}^*}{d_{ij} \cdot d_{ij}^*} \cdot (y_{ik}(t) - y_{jk}(t))$$

Настраиваемый параметр  $\alpha$  влияет на скорость работы и сходимость алгоритма.

Следует отметить, что возможно использование *двухэтапного метода* снижения размерности пространства. В таком методе на первом этапе с помощью нейронной сети РСА находится грубое приближение решения. На втором этапе полученное приближение уточняется описанным выше нелинейным методом снижения размерности пространства. Двухэтапный метод обладает потенциально более высокой точностью и более быстрой сходимостью. Результаты экспериментального исследования такого метода и входящих в его состав алгоритмов представлены в разделе 4.2.

### 3.3 Комбинированный метод снижения размерности пространства, функционирующий на основе методов снижения размерности и кластеризации

Приведенный в предыдущем разделе алгоритм нелинейного снижения размерности хорошо работает на небольших объемах данных, однако на больших объемах данных возможности применения алгоритма ограничивает вычислительная сложность. При объеме выборки  $N$  и числе итераций, сравнимом с объемом выборки, объем вычислений для получения отображения равен  $O[N^3]$ .

Решением данной проблемы может быть алгоритм, использующий аппроксимации приращений координат точек на каждой итерации. При этом, если вычислительная сложность для построения аппроксимационной оценки на каждой итерации  $O[k]$ , где  $k \ll N$ , то вычислительная сложность всего алгоритма может быть снижена до  $O[N^2]$ .

Среди существующих решений в качестве такой аппроксимации может быть использован метод, подобный предложенному *Чалмерсом* в работе [2]. В этом подходе на каждой итерации для каждого корректируемого элемента формируется 2 множества. В первом из множеств содержатся элементы, наиболее близкие к рассматриваемому в многомерном пространстве. Во втором множестве содержатся элементы, отбираемые на каждой итерации случайным образом. Такой подход был использован для минимизации ошибки, выражаемой в виде (1), однако он может быть применен и при минимизации ошибки Сэммона (2) (далее – *метод ЧС*).

В настоящей работе предлагается использовать новый *комбинированный метод (КМ)*. Он состоит в том, чтобы при аппроксимации использовать результаты иерархической кластеризации в рамках поэтапной процедуры следующего вида. На первом этапе для всех  $k$  кластеров верхнего уровня строится двумерное отображение центров этих кластеров. На втором этапе строится  $k$  отображений для подкластеров и объектов второго уровня. При этом для построения отображения каждого подкластера расположение координат центров суперкластеров фиксируется, а производится оптимизация только объектов, находящихся в рассматриваемом подкластере. Процесс повторяется для третьего уровня иерархии и так далее, пока не будет построено отображение всех объектов.

В случае, когда мощность кластера равна  $k$ , вычислительная сложность двумерного отображения элементов кластера равна  $O[k^3]$ . Далее, число кластеров при иерархической кластеризации в случае сбалансированного дерева кластеров составляет:

$$M = \frac{N}{k} + \frac{N}{k^2} + \dots + \frac{N}{k^{\log_k N}} = N \cdot \sum_{i=1}^{\log_k N} \frac{1}{k^i} = \frac{N-1}{k-1}.$$

Таким образом, вычислительная сложность в случае сбалансированного дерева кластеров равна  $\frac{N-1}{k-1} O[k^3] \sim O[k^2 N]$ . В случае сбалансированного дерева двухуровневой иерархии  $k = \sqrt{N}$ . Тогда сложность алгоритма, как и в случае предыдущего алгоритма, равна  $O[N^2]$ . Однако с ростом количества уровней иерархии сбалансированного дерева сложность алгоритма уменьшается. Для случая сбалансированного дерева высотой  $L = \log_k N$ , когда в каждом кластере оказывается  $k = N^{1/L}$  элементов, выражение сложности

принимает вид  $O\left[N^{1+\frac{2}{L}}\right]$ . Очевидно, по мере роста

числа уровней  $L$  это выражение стремится к  $O[N]$ .

В качестве *модификации комбинированного метода (МКМ)* можно предложить метод, в котором при построении отображения для любого подкластера в качестве фиксированных центров берутся центры кластеров верхнего уровня.

Результаты экспериментального исследования представленных в этом разделе методов приведены в разделе 4.2.

### 3.4. Описание метода навигации

Используя теоретические основы, представленные в данном разделе, можно предложить следующий метод навигации по коллекциям цифровых изображений.

На первом этапе для всей коллекции цифровых изображений производится расчет признаков (вопрос выбора системы признаков рассматривается в 5 разделе). На втором этапе производится иерархическая кластеризация изображений в многомерном пространстве признаков с использованием алгоритма WTA, описанного в разделе 3.1. Далее, применяется комбинированный метод отображения, предложенный в разделе 3.4, который для каждого изображения коллекции рассчитывает координаты его расположения на двумерной плоскости отображения (навигации).

В процессе навигации, переход к подмножеству изображений коллекции, соответствующему меньшему или большему уровню различий (операции приближения или удаления), соответствует отображению представителей кластеров (или самих изображений) из уровня иерархии, соответствующего выбранному уровню отображения. Далее, так как пространство навигации является двумерным, на нем естественным образом можно определить операции перемещения вверх, вниз, вправо и влево.

Следует отметить, что, помимо базовых операций навигации, таких как перемещение и изменение масштаба, разработанный метод позволяет определить операцию перехода к конкретному кластеру в дереве иерархии.

## 4 Результаты экспериментальных исследований

В данном разделе приводятся результаты экспериментальных исследований методов, изложенных в разделе 3.

### 4.1 Экспериментальные исследования алгоритмов кластеризации

Экспериментальные исследования алгоритмов кластеризации, носящие вспомогательный характер, проводились с целью сравнения качества кластеризации, получаемого различными

алгоритмами, и выбора алгоритма для дальнейшего использования. При проведении исследований качество кластеризации оценивалось по погрешности, выражаемой в виде:

$$\varepsilon_q = \frac{1}{U} \cdot \sum_{j=1}^U \|x_j - w_{v(j)}\|^2$$

где  $x_j$  - входной образ,  $w_{v(j)}$  - центр соответствующего кластера, рассчитанный как выборочное среднее входящих в него образов,  $U$  - количество входных векторов. Выбор данной формулы для оценки качества кластеризации, помимо простоты вычисления, был обусловлен спецификой выбранного в работе пространства признаков. Оно основано на цветовом пространстве CIE  $L^*a^*b^*$ , разработанном специально таким образом, чтобы степень рассогласования цветов в данном пространстве, рассчитанная как евклидово расстояние, соответствовало человеческому восприятию.

При исследовании использовались наборы по 1000 изображений. Такое относительно небольшое количество изображений представляется для вспомогательного исследования достаточным. В таблице 1 приведены средние значения ошибки квантования, получаемые в результате кластеризации на верхнем уровне иерархии для каждого из рассматриваемых алгоритмов. Были проведены эксперименты, в которых максимальное количество кластеров для верхнего уровня иерархии (а также для любой ветви иерархии) составляло 25, 50 и 100 кластеров.

Таблица 1. Средние значения ошибки квантования

Количество кластеров	Алгоритм		
	Single link	Complete link	WTA
25	0.387	0.193	0.169
50	0.359	0.150	0.139
100	0.293	0.116	0.112

Из приведенных результатов видно, что наилучшие значения погрешности обеспечивает алгоритм WTA. Следует отметить, что алгоритм Complete Link так же обеспечивает хорошие значения погрешности, худшие значения - у алгоритма Single Link.

Опираясь на результаты проведенного исследования, а также, учитывая, что для работы алгоритма WTA не требуется хранение матрицы попарных расстояний, при проведении описанных в следующих разделах экспериментов для кластеризации всюду выбирался алгоритм WTA.

Следует отметить, что максимальное число кластеров верхнего уровня (а также любой ветви иерархии) является настройкой алгоритма. При выборе данного значения следует учитывать как максимально возможное число объектов, участвующих в двумерном отображении, так и порождаемую глубину иерархии. При больших значениях этого параметра работа методов двумерного отображения может занять

неприемлемо большое время. При малых значениях возможно получение больших значений погрешности при применении комбинированных методов двумерного отображения вследствие недостаточного количества фиксируемых центров суперкластеров. Практически рассматриваемое значение следует выбирать из диапазона от нескольких десятков до сотни.

#### 4.2 Экспериментальные исследования методов двумерного отображения

Экспериментальные исследования линейного, нелинейного и двухэтапного методов снижения размерности проводились следующим образом. Из набора, содержащего порядка 10 тыс. различных цветных изображений, случайным образом производились выборки по 100 изображений. Далее для каждой такой выборки применялись линейный, нелинейный, а также двухэтапный методы снижения размерности, описанные в разделе 3.2. Малый объем выборки (100 шт.) обусловлен, прежде всего, высокой вычислительной сложностью алгоритма Сэммона снижения размерности. Целью экспериментов было определение целесообразности использования двухэтапного метода отображения, описанного в разделе 3.2.

Для оценки качества работы алгоритмов значение ошибки рассчитывалось в соответствии с критерием (2). При применении нелинейного метода снижения размерности Сэммона проводилось по 10 запусков, для каждого из которых начальное приближение выбиралось случайным образом. При этом фиксировалось также минимальное значение ошибки по 10 запускам.

В таблице 2 приведены средние значения ошибок по всем сгенерированным наборам изображений для каждого из трех рассмотренных методов. Для нелинейного метода снижения размерности Сэммона также указано среднее из полученных минимальных значений ошибки.

**Таблица 2. Сравнительный анализ линейного и нелинейного методов двумерного отображения**

	Количество итераций		
	100	200	300
Среднее значение ошибки для метода Сэммона	0.126	0.078	0.056
Среднее из минимальных значений ошибки для метода Сэммона	0.093	0.051	0.035
Среднее значение ошибки для нейросетевого алгоритма РСА	0.139		
Среднее значение ошибки для двухэтапного метода	0.044	0.041	0.039

Как видно из таблицы, значение ошибки (2) для алгоритма РСА является наихудшим. Однако использование результатов применения алгоритма РСА в качестве начального приближения дает возможность получить в двухэтапном методе достаточно малые значения ошибки уже при 100 итерациях. Использование нелинейного алгоритма не позволяет в среднем достичь результатов, полученных при использовании двухэтапного метода даже при достаточно большом количестве итераций.

С другой стороны, при использовании разумного количества попыток (до 10) и достаточного числа итераций (порядка 300 для наборов до 100 элементов) можно добиться меньших значений погрешности, чем при использовании двухэтапного метода. Однако с ростом количества попыток растет и время решения задачи.

Пример работы двухэтапного метода для отображения одной из сформированных выборок показан на рис. 1.



**Рис.1. Пример работы двухэтапного метода**

В следующей серии экспериментов было проведено исследование методов, описанных в разделе 3.3. Было проведено 100 экспериментов, в которых каждый из методов применялся к набору из 1000 изображений. Для оценки качества работы алгоритмов значение ошибки, как и в предыдущем случае, рассчитывалось в соответствии с критерием (2). Для получения начального приближения для метода ЧС использовался алгоритм РСА. Результаты экспериментов представлены в таблице 3.

Как видно из таблицы, метод КМ дал наибольшие значения ошибки. Значения ошибки для методов ЧС и МКМ не позволяют сделать выбор в пользу одного из них только на основе критерия качества (2).

Однако, учитывая, что метод МКМ имеет меньшую вычислительную сложность, при практическом применении следует отдать предпочтение именно этому методу.



Рис. 2. Пример работы модифицированной версии комбинированного метода

Таблица 3. Сравнительный анализ комбинированного метода отображения

Метод	Среднее значение ошибки	СКО ошибки	Среднее время работы, сек
РСА	0.1171	0.01449	2
ЧС	<b>0.02880</b>	0.005668	62
КМ	0.03407	0.002638	17
МКМ	<b>0.02767</b>	0.002047	19

Отметим также, что для хранения расстояний при применении метода Чалмерса желательно использовать матрицу размером  $N*N/2$ , что затрудняет использование метода в случае коллекций большого объема. При применении же модифицированной версии комбинированного метода требуется лишь хранение  $k*k/2$  расстояний. Учитывая указанные факторы, можно рекомендовать предложенную модификацию комбинированного метода к практическому использованию.

Пример работы модифицированной версии комбинированного метода для набора из 10000 изображений показан на рис. 2. Пример работы двухэтапного метода для отображения некоторых сформированных кластеров показан на рис. 3.

## 5 Выбор системы признаков

При построении систем распознавания и поиска изображений выбор системы информативных признаков является сложной задачей, решение которой во многом определяется личным опытом и предпочтениями разработчика. Следует отметить, что в настоящее время задача поиска может быть удовлетворительно решена лишь для узкоспециализированных областей. При построении систем поиска общего назначения, тем не менее, можно отметить склонность к использованию сложных систем признаков, построенных на основе цветовых, текстурных признаков, признаков формы и пространственных отношений между областями [1].

При решении задачи навигации одним из важнейших факторов, влияющих на выбор системы признаков, является понятность пользователю формируемого порядка расположения изображений.



Рис. 3. Примеры работы двухэтапного метода для сформированных кластеров

Даже довольно изощренная система признаков может оказаться неэффективной, если пользователь системы не понимает, каким образом распределяются изображения на экране, или по каким критериям формируются кластеры. Причиной такого эффекта в большой степени является значительная потеря информации при проецировании в пространства малой размерности [18].

В системах поиска и навигации цветовые признаки применяются наиболее часто, так как получаемые результаты наглядны и хорошо согласуются с ожиданиями пользователей. Поэтому в настоящей работе используются хорошо зарекомендовавшие себя гистограммы цветов, в качестве метрики взято евклидово расстояние между гистограммами.

При построении цветовых гистограмм было использовано цветовое пространство CIE  $L^*a^*b^*$ . Основным преимуществом данного цветового пространства является тот факт, что рассогласование цветов в данном пространстве (рассчитанное как евклидово расстояние между цветами) соответствует человеческому восприятию.

## 6. Направления будущих работ

В данной работе остались незатронутыми ряд важных вопросов. Один из них связан с проведением экспертной оценки разработанного метода и удобства его использования. Другим важным вопросом является дальнейшее исследование применяемой системы признаков, так как от качества используемой системы признаков зависит, в том числе, субъективная оценка пользователями результатов кластеризации и размещения изображений на мониторе. Так же планируется модификация предлагаемого метода, а

именно усовершенствование процедуры кластеризации изображений.

## 7. Заключение

В работе разработан метод навигации по коллекциям цифровых изображений. В основе подхода лежит применение как алгоритма кластеризации, используемого для иерархического представления коллекции изображений, так и двухэтапной процедуры снижения размерности пространства, используемой для отображения элементов иерархии на плоскости. Проведены экспериментальные исследования разработанного метода, показавшие его эффективность. Также выполнено экспериментальное сравнение известных методов кластеризации, использованных для иерархического представления коллекции изображений.

## Литература

- [1] V. Castelli, L. D.Bergman. Image Databases: Search and Retrieval of Digital Imagery. 2002.
- [2] M. Chalmers "A Linear Iteration Time Layout Algorithm for Visualising High-Dimensional Data", Proc. IEEE Visualization '96, San Francisco, pp. 127-132 (1996).
- [3] J.-Y. Chen, C. A. Bouman, and J. C. Dalton. Similarity Pyramids For Browsing And Organization Of Large Image Databases. In Human Vision and Electronic Imaging III, volume 3299 of Proceedings of SPIE, pages 563- 575, 1998.
- [4] F. Chen, U. Gargi, L. Niles, and H. Schutze. Multi-modal browsing of images in web documents. In Document Recognition and Retrieval VI, volume 3651 of Proceedings of SPIE, pages 122-133, 1999.
- [5] J.P.Eakins, M.E.Graham Content-based Image Retrieval. A Report to the JISC Technology Applications Programme. Institute for Data

- Research, University of Northumbria at Newcastle, 1999.
- [6] S.F.Everton. A Guide For The Visually Perplexed: Visually Representing Social Networks. Stanford University. January 2004.
- [7] T. Fruchterman, E.Reingold. Graph Drawing by Force-Directed Replacement. *Software--Practice and Experience* 21: 1129-1164, 1991.
- [8] A. Hiroike, Y. Musha, A. Sugimoto, and Y. Mori. Visualization of information spaces to retrieve and browse image data. In *Proceedings of the Third International Conference on Visual Information and Information Systems (VISUAL'99)*, volume 1614 of *Lecture Notes in Computer Science*, pages 155-162. Springer, 1999.
- [9] T.Kamada, S.Kawai. An Algorithm for Drawing General Undirected Graphs. *Information Processing Letters* 31: 7-15, 1989.
- [10] Kobus Barnard, Pinar Duygulu, David Forsyth. Clustering Art.
- [11] M. Koskela, J. Laaksonen, S. Laakso, and E. Oja. The PicSOM retrieval system: Description and evaluations. In *The Challenge of Image Retrieval. Electronic Workshops in Computing*, 2000.
- [12] J. Marks, B. Andalman, P. A. Beardsley, W. Freeman, S. Gibson, J. Hodgins, T. Kang, B. Mirtich, H. Pfister, W. Ruml, K. Ryall, J. Seims, and S. Shieber. Design Galleries: A general approach to setting parameters for computer graphics and animation. In *Proceedings of SIGGRAPH '97*, pages 389-400. ACM, 1997.
- [13] Z. Peccenovic, M. Do, M. Vetterli, and P. Pu. Integrated browsing and searching of large image collections. In *Proceedings of the Fourth International Conference on Advances in Visual Information Systems (VISUAL 2000)*, volume 1929 of *Lecture Notes in Computer Science*, pages 279-289. Springer, 2000.
- [14] J.C. Platt, Mary Czerwinski, Brent A. Field. PhotoTOC: Automatic Clustering for Browsing Personal Photographs. Technical Report. Microsoft Research. February 2002.
- [15] J.C. Platt. AutoAlbum: Clustering Digital Photographs using Probabilistic Model Merging. *Proc. IEEE Workshop on Content-Based Access of Image and Video Libraries*, pages 96-100, 2000.
- [16] K. Rodden. Evaluating Similarity-Based Visualisations As Interfaces For Image Browsing. Technical Report. University of Cambridge Computer Laboratory, 2002.
- [17] K. Rodden, W. Basalaj, D. Sinclair, and K. Wood. Evaluating A Visualisation Of Image Similarity As A Tool For Image Browsing. In *Proceedings of the IEEE Symposium on Information Visualization (InfoVis'99)*. IEEE, 1999.
- [18] K. Rodden, W. Basalaj, D. Sinclair, and K. Wood. A comparison of measures for visualising image similarity. In *The Challenge of Image Retrieval. British Computer Society Electronic Workshops in Computing*, 2000.
- [19] T. Rose, D. Elworthy, A. Kotcheff, A. Clare, and P. Tsonis. ANVIL: a system for the retrieval of captioned images using NLP techniques. In *The Challenge of Image Retrieval. Electronic Workshops in Computing*, 2000.
- [20] Y. Rubner, C. Tomasi, and L. J. Guibas. Adaptive color-image embeddings for database navigation. In *Proceedings of the Asian Conference on Computer Vision*, pages 104-111. IEEE, 1998.
- [21] S. Santini and R. Jain. Integrated browsing and querying for image databases. *IEEE Multimedia*, 7(3):26-39, 2000.
- [22] C. J. Van Rijsbergen. *Information Retrieval*. London, 1979.
- [23] J. Vendrig, M. Worring, and A. W. M. Smeulders. Filter image browsing: Exploiting interaction in image retrieval. In *Proceedings of the Third International Conference on Visual Information and Information Systems (VISUAL'99)*, volume 1614 of *Lecture Notes in Computer Science*, pages 147-154. Springer, 1999.
- [24] Мясников Е.В. Нейросетевые алгоритмы кластеризации отсчетов цветных изображений. Перспективные информационные технологии в научных исследованиях, проектировании и обучении (ПИТ-2006). Труды научно-технической конференции с международным участием. Том 2. – Самара, 2006. - стр. 118-123.
- [25] Осовский С. Нейронные сети для обработки информации. – М.: Финансы и статистика, 2002.
- [26] Фукунага К. Введение в статистическую теорию распознавания образов.- М.Наука. 1979 - 368стр.

## Digital image collection navigation based on automatic classification methods

E.V. Myasnikov

The central question of image navigation system construction is to build the projection of the collection into the two-dimensional space of navigation. The kernel of the method proposed is to build the projection in two steps. At the first step the hierarchical system of clusters is constructed. At the second step the initial space of image descriptions is projected into two-dimensional navigation space.

A survey of methods used for navigation system construction is given. The results of experimental analysis are present. The proposed method is compared to known methods. The results of this work allow to draw a conclusion of ability to successfully apply the method developed.

---

\* Работа подготовлена при поддержке ООО "ЯНДЕКС" (www.yandex.ru).