

ОБНАРУЖЕНИЕ ЛОКАЛЬНЫХ ПРОСТРАНСТВЕННЫХ СТРУКТУР ДЛЯ РАСПОЗНАВАНИЯ ИЗОБРАЖЕНИЙ

Подход выделения признаков, широко применяющийся при решении прикладных задач компьютерного зрения, в настоящее время включает в себя использование алгоритмов обучения признакам без учителя, таким как ограниченная машина Больцмана и разреженный автоэнкодер, что позволяет извлекать признаки автоматически из доменно-специфичной общей выборки. Обучение без учителя особенно эффективно в сочетании с моделями глубокого обучения, такими как сверточные нейронные сети и глубокие сети доверия (DBN). Однако при обработке высокоуровневых и структурированных данных, подверженных инвариантным трансформациям (что особенно актуально для компьютерного зрения в 3D и в движении), построение словаря признаков для отдельно взятого элемента выборки может быть проблематичным.

Предложен альтернативный подход, включающий в себя обнаружение некоторых пространственных структур в изображениях, характерных для конкретной выборки. Подобные способы имеют историю применения в задачах распознавания для отдельных, узкоспецифичных доменов (таких как обнаружение взаимоотношений между признаками черт лица). Рассмотренный алгоритм предлагает общий подход к проблеме, эксплуатируя некоторые свойства ориентационных клеток первичной зрительной коры головного мозга, с их помощью обнаруживая обобщенные свойства изображений и агрегируя их для получения информативных статистических параметров, которые в дальнейшем могут использоваться непосредственно для распознавания и классификации.

ВЫДЕЛЕНИЕ ПРИЗНАКОВ; ЭКВИВАРИАНТНОЕ ПРЕОБРАЗОВАНИЕ; ЗРИТЕЛЬНАЯ КОРА; РАЗРЕЖЕННОЕ КОДИРОВАНИЕ; ФИЛЬТР ГАБОРА.

A.A. Khurshudov

DISCOVERING LOCAL SPATIAL RELATIONSHIPS FOR IMAGE RECOGNITION

Nowadays the feature detection approach widely used in lots of computer vision applications has been particularly improved by using unsupervised feature learning techniques such as Restricted Boltzman Machines and Sparse Autoencoders. It helps to learn features from large amounts of generally unrelated (or domain-related) data. Unsupervised feature learning has especially become useful combined with deep learning models like convolution neural networks and Deep Belief Networks. However, when dealing with complex and high-level structured data, as long as with data exposed to lots of invariant transformations (which is highly relevant to a computer vision in 3D and motion), it can become a problem to construct a bag-of-words like a feature dictionary to contain all the possible changes an object can take. Instead, the study offers a different approach involving discovering some relevant spatial relationships that appear across the dataset. The idea by itself is not a new one. The classical example of using spatial structures for recognition would be a distinct pattern of a human face, with two eyes and a mouth. However, most of the existing solutions are strictly limited by a certain domain. The study proposes an algorithm inspired by some properties of a primary visual cortex V1. It mimics the functionality of orientation cells, discovers general properties of a natural image and aggregates them together to extract some useful statistics to be further used in a classification algorithm.

FEATURE DETECTION; INVARIANT TRANSFORM; PRIMARY VISUAL CORTEX; SPARSE CODING; GABOR FILTER.

При изучении нейробиологических основ зрительной системы и эффективных алгоритмов представления визуальной информации для распознавания изображений был предложен подход объединения суще-

ственных признаков изображенного объекта в некоторую иерархическую структуру, где признаки низшего уровня, такие как границы и штрихи, комбинируются в более сложные, такие как углы, детали контуров



объектов, элементы фигуры или лиц, и т. д. В биологическом направлении подтверждением существования иерархии признаков являются клетки зрительной коры головного мозга, для которых неоднократно было показано [1, 2] существование ориентационной избирательности (чувствительности к границам под определенным углом), и отдельные классы «сложных» клеток, комбинирующих представления простых. Известно также о существовании клеток, активирующихся в ответ на исключительно узкоспециализированные и сложные объекты, такие как отдельный человек, причем способных при этом проявлять высокую степень толерантности к искажениям [3]. Вопрос о представлении в коре головного мозга сложных, функционально обособленных признаков до сих пор остается нерешенным.

В прикладных областях, таких как компьютерное зрение и машинное обучение, предлагались свои варианты построения иерархии признаков, наиболее известным из которых является обучение признакам на основе разреженного кода [4]. Подобный подход показал высокую эффективность на тестовых наборах данных, таких как выборки ImageNet [5] (массивная структурированная база изображений, де-факто стандарт сравнения алгоритмов распознавания), но при этом подвергается критике в отдельных исследованиях за статистическую неэффективность и избыточность получаемого кода. Отдельную сложность представляют собой пространственные преобразования, такие как масштабирование, трансляция и вращение, способные радикально изменить форму изображенного объекта и требующие отдельного набора признаков для разных возможных позиций и комбинаций, влияющих на восприятие характеристик изображения (таких как яркость, контраст и т. д.). Предположение об использовании разреженного кода в зрительной системе животных и человека представляется убедительным на некоторых (низших) уровнях обнаружения признаков, но для граничного случая несложно показать [6], что кодирование каждого представления отдельным элементом разреженного ансамбля нейро-

нов требует не соответствующего реальности количества нейронов.

Таким образом, для высших уровней иерархии признаков необходим отдельный механизм, устойчивый к пространственным преобразованиям.

Проблемы обучения высокоуровневым признакам

Этот вид обучения, как правило, проводится без учителя, не считая (при необходимости) организацию выборки таким образом, чтобы она соответствовала выбранному домену. При этом могут использоваться различные алгоритмы, такие как анализ главных компонент (Principal Component Analysis – PCA) [7], разреженный автоэнкодер [8], ортогональное согласованное преследование (Orthogonal Matching Pursuit – OMP) [9], решающие, принципиально, одну задачу: аппроксимацию выборки набором ограниченных, независимых и (как правило, желательно) разреженных компонент, которые, в конечном итоге, и представляют собой признаки изображения.

Для составления иерархии признаков проводится несколько этапов такого обучения на локальных фрагментах изображений различного размера и детализации. Возможно использование информации, полученной на первых этапах обучения, для фильтрации выборки. Так, обнаружив, что изображение не содержит низкоуровневых признаков, алгоритм может отбросить его при изучении признаков более высоких уровней.

Конечный результат схематично выглядит, как показано на рис. 1.

В примере признаки получены с помощью алгоритма разреженного автоэнкодера, но похожий результат может достигаться любым из перечисленных выше методов.

Очевидным следствием подобного подхода является то, что каждый следующий уровень иерархии будет состоять из большего количества признаков, чем предыдущий, в силу того, что признаки высоких уровней представляют собой некоторое подмножество комбинаций признаков низкого уровня. Кроме того, дополнительную

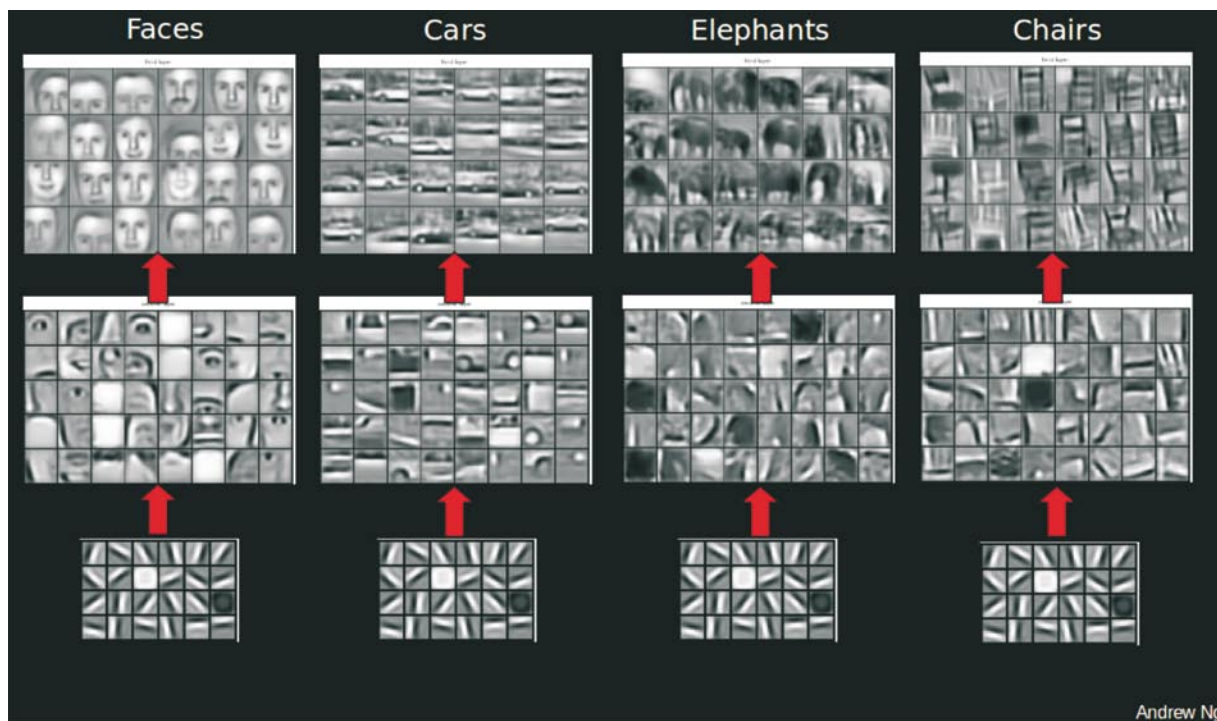


Рис. 1. Схема иерархии признаков для различных доменов
(слайд ECCV 10, Европейской конференции по проблемам компьютерного зрения)

вариативность вносят пространственные трансформации, особенно трехмерные, которые начинают играть роль на уровнях представления целых объектов. Поворот угла камеры в трехмерном пространстве, трансляция и окклюзия (частичное загоразивание) способны радикально изменить вид объекта. С точки зрения естественного зрения подобные трансформации считаются инвариантными, но алгоритмы разреженной декомпозиции при этом воспринимают каждый такой экземпляр выборки как различные изображения.

На практике обучение путем выделения разреженных компонент работает тем успешнее, чем более ограничено пространство инвариантных преобразований объектов выборки. Так, фотографии человеческих лиц (фиксированная часть фигуры, не имеющая составных изменяющихся элементов) успешнее подвержены классификации, чем фотографии собак (на фото изображено тело целиком, способное принимать множество различных положений и выглядящее по-разному с различных углов зрения).

Для решения поставленных задач было проведено тестирование алгоритма разреженного автоэнкодера на выборках человеческих лиц и силуэтов животных. Если сформулированное предположение о фундаментальной ограниченности подхода выделения разреженных компонент верно, то выборка, содержащая в себе большое количество трехмерных инвариантных преобразований, будет плохо поддаваться обучению. Результаты тестирования показывают, что этот вывод подтверждается (см. рис. 2).

Таким образом, представляется вероятным, что несмотря на эффективность при обучении низкоуровневым признакам, метод декомпозиции выборки на составляющие ключевые признаки неэффективен на высоких уровнях. Сама идея построения иерархии признаков, представляющих собой компоненты декомпозиции изображений выборки, неполна без учета пространственных закономерностей и характеристик, которые признаки приобретают с увеличением размера, сложности и детализации. Так, признаки первого уровня иерархии, обычно соотносящиеся с

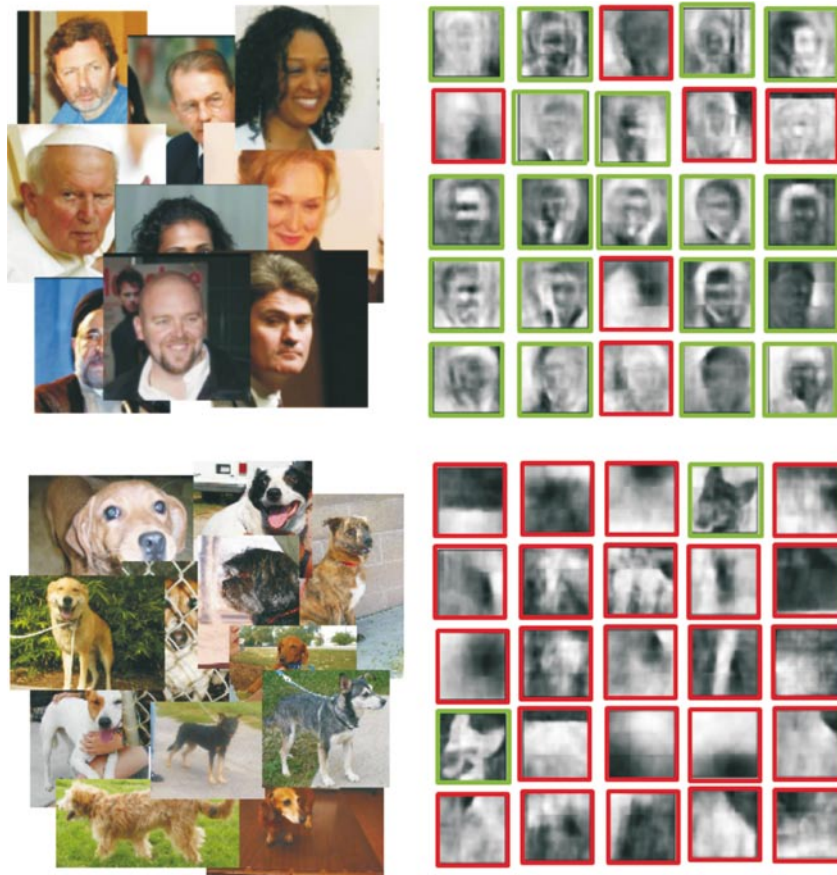


Рис. 2. Обучение высокоуровневым признакам на примерах LFW Face Database и Kaggle's Cats vs Dogs Challenge. Вторая выборка практически не поддается декомпозиции (зеленым отмечены удачные результаты)

рецептивными полями первичной зрительной коры (V1) и представляющие собой контрастные штрихи различной ориентации, узкоспецифичны и не подвержены инвариантным трансформациям: любое изменение ориентации штриха представляет собой переход к другому признаку, в отличие от рассмотренных признаков высоких уровней, где один и тот же силуэт лица может рассматриваться с различных углов, трансляцией в поле зрения или разным масштабом.

Ограниченная разреженная аппроксимация фильтрами Габора

Рассмотрим ограниченную версию алгоритма ортогонального согласованного преследования со следующими правилами:

в качестве словаря используется набор двумерных фильтров Габора, отличающихся

ся углом ориентации ϕ , последовательно меняющимся от нуля до 2π ;

количество коэффициентов, отличных от нуля, для результата работы алгоритма выберем равным единице.

Входные данные: сигнал b , матрица фильтров Габора A , критерий остановки (пороговое значение; варьируется в зависимости от приложения) n .

Выходные данные: вектор аппроксимации c , соответствующий линейному разложению сигнала b на компоненты матрицы A .

Алгоритм:

1. Установить остаток $r_0 = b$, счетчик итераций $t = 0$ и вектор $V = \emptyset$.

2. $v_i = a_i$, где a_i удовлетворяет условию $\max_k \langle r_i, a_k \rangle$, где a_k — строки матрицы A .

3. Добавить v_i в вектор $V_i : V_i = V_{i-1} \cup \{v_i\}$.

4. Для вектора c решить методом наименьших квадратов:

$$\min_{c \in C^t} \left\| b - \sum_{j=1}^t c(v_j) a_{v_j} \right\|.$$

5. Рассчитать новый остаток с использованием c :

$$r_t = r_{t-1} - \sum_{j=1}^t c(v_j) a_{v_j}.$$

6. $t = t + 1$.

7. Если критерий остановки достигнут, то для каждого значения $c_i \neq \max(c)$, $c_i = 0$. Вернуть c .

Изображения выборки кодируются следующим образом: каждый отдельный экземпляр разбивается на множество фрагментов, для каждого из которых выполняется алгоритм ограниченного ОМР. Таким образом, каждый фрагмент изображения аппроксимируется одним фильтром Габора, играющим роль локального рецептивного поля простой клетки зрительной коры. Единственный ненулевой коэффициент, полученный в результате ОМР, при этом определяет силу активации отдельной клетки. Разбивка изображения на фрагменты может происходить как простой продольно-поперечной сеткой, так и с перекрытием и последующим усреднением перекрывающихся активаций (рис. 3).

Тестовые результаты показывают, что несмотря на грубость аппроксимации (каж-

дый небольшой фрагмент изображения заменяется на максимально соответствующий ему прямой отрезок заданной ориентации), разбиение с перекрытием сохраняет большинство деталей, необходимых для распознавания, устраняя при этом часть шума, оттенки и текстуры поверхностей.

Полученный результат представляется моделью того, как воспринимает изображение совокупность клеток зрительной коры V1. Фильтры Габора в качестве признаков используются ввиду нескольких причин. Они неоднократно использовались в качестве моделирования рецептивных полей первичной зрительной коры [10, 11]. Дж. Бергстра и И. Бенджио [12] описывают статистические достоинства таких признаков по сравнению с признаками, обнаруженными с помощью обучения. Ключевым элементом в извлечении высокоуровневых признаков и основным преимуществом данного алгоритма по сравнению с классическим алгоритмом ОМР при этом оказывается тот факт, что фильтры Габора позволяют придать локальным участкам изображения пространственную семантику. За счет того что каждый из них соотносится с определенным, заранее известным углом ориентации ϕ , оказывается возможным выделить непрерывные участки одинаковой ориентации, соответствующие протяженным границам, и участки сочленений, где

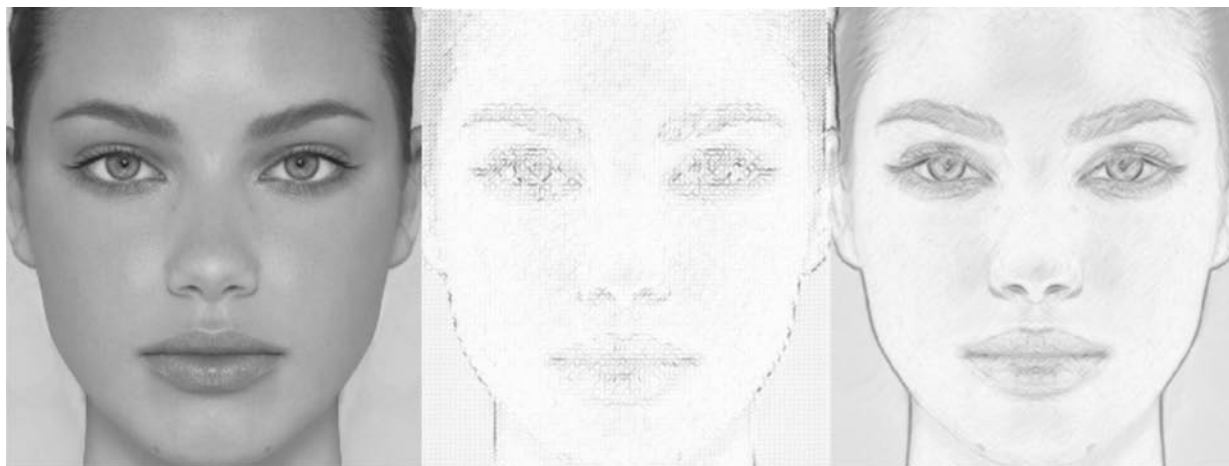


Рис. 3. Результат обработки изображения ограниченным ОМР с фильтрами Габора (карта активаций)

Слева направо: оригинал изображения, результат разбиения без перекрытия и с перекрытием соответственно. Наиболее темные участки изображения представляют наибольшие значения активации

на ограниченном пространстве встречаются активации клеток различных ориентаций. Более того, подобный подход позволяет ввести метрику схожести различных клеток, определив ее через разницу углов ориентации:

$$s = \|\varphi_1 - \varphi_2\| \bmod \varphi_{\max}, \quad (1)$$

где φ_1 – значение ориентации первой клетки; φ_2 – значение ориентации второй клетки; φ_{\max} – максимальное значение ориентации.

Таким образом, клетки, кодирующие ортогональные друг другу фильтры, будут восприниматься моделью как максимально различные.

Локальные пространственные структуры на базе групп ориентационных клеток

Дальнейшая обработка изображения строится на предположении, что устойчи-

вые комбинации локальных участков изображения определенной ориентации и участков-сочленений могут эффективно использоваться в качестве признаков для задачи распознавания. Рассмотрим в качестве примера базу данных рукописных цифр MNIST. Символы письменности в качестве объектов классификации особенно хорошо подходят для иллюстрации используемой идеи, поскольку, как представляется в современных работах [13], письменность изначально составлялась таким образом, чтобы различные ее знаки были эффективно различаемыми наборами штрихов и пересечений. При этом знаки письменности подвержены инвариантным преобразованиям (существует огромное количество шрифтов, вариаций почерка и материалов, влияющих на внешний вид), что объединяет этот пример с рассматриваемыми ранее инвариантными преобразованиями в 3D.

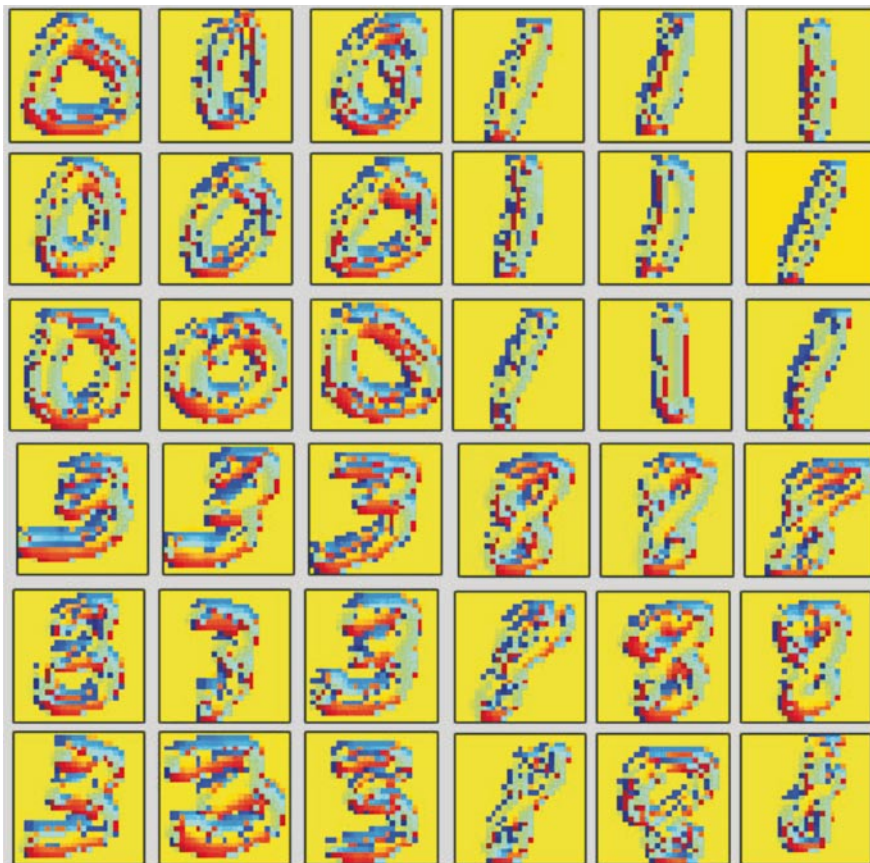


Рис. 4. Карты ориентаций цифр MNIST после преобразования ограниченным OMP. Каждый пиксель соответствует одной клетке, цвет определяет ориентацию. Близкие оттенки изображают схожие клетки

Карты ориентаций демонстрируют схожие цветовые участки в изображениях одного класса (рис. 4). Становится возможным выделять продолжительные «гладкие кривые», где ориентации соседних клеток изменяются незначительно, сочленения или переходы между однородными участками, а также концы линий. Однако разбиение изображений на фрагменты с перекрытием оставляет некоторое количество шума – случайных одиночных активаций. Для фильтрации и усиления позиций существующих групп активных клеток применим алгоритм, эксплуатирующий идею латерального ингибирования [13], соревновательного взаимодействия между соседними клетками, в ходе которого клетки с высоким значением активации и поддержкой со стороны схожих соседей распространяют свое влияние на отличающихся соседей, корректируя их активацию в сторону своей.

Алгоритм:

Входные данные: карта ориентаций O_{map} , карта активаций A_{map} , пороговое значение активации A_t , пороговое значение схожести O_t , максимальное значение ориентации φ_{max} .

Выходные данные: фильтрованные карты O'_{map} , A'_{map} .

1. $O'_{\text{map}} = O_{\text{map}}$, $A'_{\text{map}} = A_{\text{map}}$.
2. Установить счетчик изменений $c = 0$.
3. Для каждой клетки (i, j) в пределах карты ориентаций:

3.1. Получить $n_{i,j}$ – множество всех соседей клетки (i, j) ;

3.2. Если $A_{\text{map},i,j} > A_t$, и нет ни одного элемента $n_{i,j}$ такого, что $A_{\text{map},ni,nj} > A_t$ и $\|O_{\text{map},i,j} - O_{\text{map},ni,nj}\| \bmod \varphi_{\text{max}} < O_t$, то установить

$$O'_{\text{map},i,j} = \frac{\sum_{ni,nj} O'_{\text{map},ni,nj}}{\text{length}(n_{i,j})}, \quad A'_{\text{map},i,j} = \frac{\sum_{ni,nj} A'_{\text{map},ni,nj}}{\text{length}(n_{i,j})};$$

3.3. Если $A_{\text{map},i,j} < A_t$, и нет ни одного элемента $n_{i,j}$ такого, что $A_{\text{map},ni,nj} > A_t$ и $\|O_{\text{map},i,j} - O_{\text{map},ni,nj}\| \bmod \varphi_{\text{max}} < O_t$, то $A'_{\text{map},i,j} = 0$;

3.5. $c = c + 1$.

4. Повторять, пока $c \neq 0$.

После снижения шума латеральным ингибированием появляется возможность поиска локальных элементов с большей сложностью, таких как соединения однородных регионов и концы линий. В современной психофизиологии существует большое количество подтверждений тому, что именно такие детали изображения исследуются зрительной системой животных и человека с наибольшим вниманием и, вероятно, важны для распознавания.

Таким образом, изображение MNIST (28×28, размерность 784) представляется композицией точек пересечений и концов линий, которых для достоверного отображения любой цифры достаточно не более шести (по аналогии с символами почтового индекса). Такое представление, кроме того,



Рис. 5. Изображение, обработанное латеральным ингибированием, и дальнейшее выделение соединений и концов линий

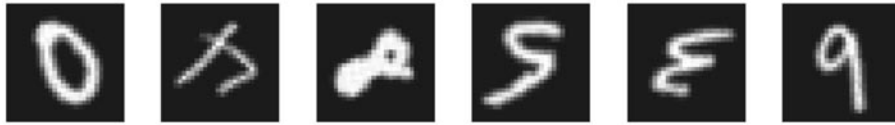


Рис. 6. Примеры экземпляров выборки MNIST с вращением

способно обеспечить высокий уровень ротационной устойчивости, если каждую цифру представить в виде маршрута обхода полученного графа, где каждый элемент маршрута будет представлять собой угол между двумя ребрами. Так, на рис. 5 угол $\pi/2$ между левой верхней чертой и горизонтальной поперечной сохранится, вне зависимости от того, на какой угол будет повернуто изображение. Инвариантность к масштабированию может достигаться, если длины ребер графа нормировать относительно средней или максимальной величины. При этом алгоритм распознавания будет воспринимать относительные расстояния между узлами, а не абсолютные.

Оценка результатов работы алгоритма

Для оценки результатов полученного алгоритма использована база MNIST с искусственным расширением выборки. Каждый образец был подвергнут нескольким пространственным преобразованиям, включающим в себя упомянутые трансляцию, вращение и масштабирование (рис. 6).

В качестве обучающего алгоритма использовалась софтмакс-регрессия (мультиномиальная регрессия) [14]. В качестве эталонных показателей в эксперименте применялись результаты работы классического многослойного перцептрона и сверточной нейронной сети архитектуры LeNet [15]. Для каждого алгоритма в статье приведено процентное значение успешно распознанных элементов выборки.

Результаты эксперимента демонстрируют преимущества рассмотренного алгоритма на выборках с большим количеством пространственных преобразований. При этом для отдельного класса преобразований трансляции эффективность алгоритма уступает сверточной нейронной сети, одним из ключевых свойств архитектуры которых как раз и является трансляционная устойчивость.

Для классификации изображений, имеющих большую структурную сложность, представляется целесообразным локальное выделение устойчиво повторяющихся графов в обособленных участках изображения (таких как глаза, нос и рот для распознавания человеческих лиц). Эти локальные структуры затем представляются единственным значением на следующем уровне масштаба, где алгоритм исследует пространственные связи уже между их сочетаниями и становится способен определить такие устойчивые паттерны, как взаимное расположение черт лица. Дополнительными свойствами локальных пространственных структур могут быть не только сочленения и концы линий, но и устойчивые детали кривых, простейшие геометрические фигуры, а также параметры выделенной области. Так, для цифр MNIST становится возможным определить преобладающую ось изображения цифры, что позволяет осуществлять ментальное вращение к стандартизированному положению.

Оценка результатов алгоритма

Выборка	Многослойный перцептрон, %	Сверточная сеть CNN-5, %	Рассмотренный алгоритм, %
MNIST классическая	98,74	98,91	95,3
MNIST вращение	61,23	73,23	85,14
MNIST трансляция	72,0	98,45	94,8
MNIST масштабирование	65,22	70,11	78,0

Достоинствами метода являются значительное снижение размерности признаков и использование пространственной семантики. Последнее представляется важным элементом системы зрительного восприятия и позволяет фиксировать инвариантные преобразования, которые во многих приложениях компьютерного зрения могут играть существенную роль. Так, зрительной системе необходимо знать, на какой именно угол повернут изображенный объект, где в поле зрения он находится, или уметь различать в объектах одного класса MNIST

несколько различных стилей написания.

Перспективным направлением дальнейшей работы является распространение метода на трехмерные объекты и комбинация локальных пространственных структур с предположительной 3D-моделью объекта. Такой подход сделает возможным ментальное вращение в трехмерном пространстве и даст возможность определять ориентацию и расположение предметов в пространстве, по аналогии с примером MNIST, не требуя наличия переполненного набора признаков.

СПИСОК ЛИТЕРАТУРЫ

1. **Olshausen B.A., Field D.J.** Sparse coding with an overcomplete basis set: A strategy employed by V1? // *Vision research*. 1997. Vol. 37. No. 23. Pp. 3311–3325.
2. **Riesenhuber M., Poggio T.** Hierarchical models of object recognition in cortex // *Nature neuroscience*. 1999. Vol. 2. No. 11. Pp. 1019–1025.
3. **Gross C.G.** Genealogy of the “grandmother cell” // *The Neuroscientist*. 2002. Vol. 8. No. 5. Pp. 512–518.
4. **Boureau Y. et al.** Sparse feature learning for deep belief networks // *Advances in neural information processing systems*. 2008. Pp. 1185–1192.
5. **Attneave F.** Some informational aspects of visual perception // *Psychological review*. 1954. Vol. 61. No. 3. P. 183.
6. **Jolliffe I.** *Principal component analysis*. John Wiley & Sons, Ltd, 2005.
7. **Bengio Y.** Learning deep architectures for AI // *Foundations and trends® in Machine Learning*. 2009. Vol. 2. No. 1. Pp. 1–127.
8. **Tropp J.A., Gilbert A.C.** Signal recovery from random measurements via orthogonal matching pursuit // *Information Theory, IEEE Transactions on*. 2007. Vol. 53. No. 12. Pp. 4655–4666.

9. **Marčelja S.** Mathematical description of the responses of simple cortical cells* // *JOSA*. 1980. Vol. 70. No. 11. Pp. 1297–1300.

10. **Daugman J.G.** Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters // *JOSA A*. 1985. Vol. 2. No. 7. Pp. 1160–1169.

11. **Larochelle H., Erhan D., Courville A., Bergstra J., Bengio Y.** An empirical evaluation of deep architectures on problems with many factors of variation // *Proc. of the 24th Internat. Conf. on Machine learning*. ACM, 2007. Pp. 473–480.

12. **Blakemore C., Carpenter R.H., Georgeson M.A.** Lateral inhibition between orientation detectors in the human visual system // *Nature*. 1970.

13. **Kwak C., Clayton-Matthews A.** Multinomial logistic regression // *Nursing research*. 2002. Vol. 51. No. 6. Pp. 404–410.

14. **LeCun Y. et al.** Comparison of learning algorithms for handwritten digit recognition // *Internat. Conf. on artificial neural networks*. 1995. Vol. 60.

15. **Tarr M. J., Pinker S.** Mental rotation and orientation-dependence in shape recognition // *Cognitive psychology*. 1989. Vol. 21. No. 2. Pp. 233–282.

REFERENCES

1. **Olshausen B.A., Field D.J.** Sparse coding with an overcomplete basis set: A strategy employed by V1?, *Vision research*, 1997, Vol. 37, No. 23, Pp. 3311–3325.
2. **Riesenhuber M., Poggio T.** Hierarchical models of object recognition in cortex, *Nature neuroscience*, 1999, Vol. 2, No. 11, Pp. 1019–1025.
3. **Gross C.G.** Genealogy of the “grandmother cell”, *The Neuroscientist*, 2002, Vol. 8, No. 5, Pp. 512–518.
4. **Boureau Y. et al.** Sparse feature learning for deep belief networks, *Advances in neural information*

processing systems, 2008, Pp. 1185–1192.

5. **Attneave F.** Some informational aspects of visual perception, *Psychological review*, 1954, Vol. 61, No. 3, P. 183.

6. **Jolliffe I.** *Principal component analysis*, John Wiley & Sons, Ltd., 2005.

7. **Bengio Y.** Learning deep architectures for AI, *Foundations and trends® in Machine Learning*, 2009, Vol. 2, No. 1, Pp. 1–127.

8. **Tropp J.A., Gilbert A.C.** Signal recovery from random measurements via orthogonal matching pursuit, *Information Theory, IEEE Transactions on*,



2007, Vol. 53, No. 12, Pp. 4655–4666.

9. **Marčelja S.** Mathematical description of the responses of simple cortical cells*, *JOSA*, 1980, Vol. 70, No. 11, Pp. 1297–1300.

10. **Daugman J.G.** Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters, *JOSA A*, 1985, Vol. 2, No. 7, Pp. 1160–1169.

11. **Larochelle H., Erhan D., Courville A., Bergstra J., Bengio Y.** An empirical evaluation of deep architectures on problems with many factors of variation, *Proceedings of the 24th International Conference on Machine learning*, ACM, 2007, Pp. 473–480.

12. **Blakemore C., Carpenter R.H., Georgeson M.A.** Lateral inhibition between orientation detectors in the human visual system, *Nature*, 1970.

13. **Kwak C., Clayton-Matthews A.** Multinomial logistic regression, *Nursing research*, 2002, Vol. 51, No. 6, Pp. 404–410.

14. **LeCun Y. et al.** Comparison of learning algorithms for handwritten digit recognition, *International Conference on artificial neural networks*, 1995, Vol. 60.

15. **Tarr M.J., Pinker S.** Mental rotation and orientation-dependence in shape recognition, *Cognitive psychology*, 1989, Vol. 21, No. 2, Pp. 233–282.

ХУРШУДОВ Артем Александрович — аспирант кафедры информационных систем и программирования Кубанского государственного технологического университета.

350072, Россия, Краснодарский край, г. Краснодар, ул. Московская, д. 2.

E-mail: art1783@gmail.com

KHURSHUDOV, Artem A. *Kuban State Technological University.*

350072, Moskovskaya Str. 2, Krasnodar, Krasnodar krai, Russia.

E-mail: art1783@gmail.com