

Методы искусственного интеллекта

УДК 004.032.26

DOI: 10.25729/ESI.2023.32.4.001

Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 1. Стратифицированные модели самоподобных нейронных сетей и быстрых преобразований

Дорогов Александр Юрьевич

ПАО «Информационные телекоммуникационные технологии» («Интелтех»),
Санкт-Петербургский государственный электротехнический университет, Россия,
Санкт-Петербург, vaksa2006@yandex.ru

Аннотация. В работе показано, что в основе построения быстрых преобразований (подобных БПФ) лежат самоподобные структуры, которые в равной степени можно использовать и для построения быстрых нейронных сетей (БНС). Показано, что класс быстрых преобразований определяется системными инвариантами морфологического уровня и может быть описан, как морфогенез терминальных проекций нейронных модулей. Предложены лингвистические модели для описания морфологии, структуры и топологии регулярных самоподобных нейронных сетей. Модели легко обобщаются на многомерные варианты нейронных сетей данного типа. Благодаря своей структуре БНС обладают особыми алгоритмами обучения, которые принципиально отличаются от классического ErrorBackPropagation отсутствием механизма обратного распространения ошибки. Алгоритмы обучения основаны на предложенных в работе методах мультипликативной факторизации образов и быстрых преобразований. Разработанные алгоритмы завершаются за конечное число шагов с гарантируемой сходимостью (будет показано в части 2 статьи). Последовательное развитие концепции самоподобия приводит к разработке методов создания быстрых нейронных сетей с глубокой степенью обучения. Самоподобные нейронные сети обладают уникальной возможностью дообучения к новым данным без потери ранее приобретённых знаний. Показано, что БНС могут быть использованы для создания быстродействующей памяти образов с произвольным доступом и сложных устройств комбинационной логики. В работе представлены результаты исследований автора по следующим вопросам: биологические предпосылки самоподобия нейронных сетей; самоподобные многослойные структуры, морфогенез, стратификация модельных представлений; алгоритмы быстрых преобразований, быстрые нейронные сети (БНС), методы настройки; обучение БНС к эталонным функциям; пластичность БНС; пирамидальные нейронные сети глубокого обучения; многоканальные корреляторы; реализация памяти и комбинационной логики на пирамидальных структурах. Результаты исследований будут представлены статьями в трёх частях.

Ключевые слова: самоподобные структуры, быстрые преобразования, морфогенез, быстрые нейронные сети, пирамидальные структуры, память образов, комбинационная логика

Цитирование: Дорогов А.Ю. Быстрые преобразования и самоподобные нейронные сети глубокого обучения. Часть 1. Стратифицированные модели самоподобных нейронных сетей и быстрых преобразований / А.Ю. Дорогов // Информационные и математические технологии в науке и управлении. – 2023. – № 4(32). – С. 5-20. – DOI: 10.25729/ESI.2023.32.4.001.

Введение. Теория быстрых перестраиваемых спектральных преобразований появилась примерно в то же время, что и теория многослойных нейронных сетей. Обе теории развивались параллельно. Различная терминология, различные теоретические основы, различные области применения развели два направления достаточно далеко друг от друга, хотя общие черты – многослойность и перестраиваемость – были вполне очевидны.

Создание алгоритма быстрого преобразования Фурье (БПФ), безусловно, можно считать одним из выдающихся достижений второй половины 20 века. Время его рождения знаменательно совпало с начальным этапом развития вычислительной техники, когда быстродействие вычислительных машин было ещё крайне низким. Появление алгоритма БПФ стимулировало интерес и к другим видам спектральных преобразований. В задачах фильтрации, сжатия и выделения информативных признаков широкое применение нашли такие преобразования, как

Адамара-Уолша, Хаара, Виленкина-Кристенсона, Хартли, Наклонное, Вейвлет и другие, также обладающие быстрыми алгоритмами. Несмотря на отличия по видам функций, оказалось, что большинство алгоритмов быстрых преобразований имеют подобную структуру и отличаются друг от друга не более, чем значениями коэффициентов базовых операций. Осознание этого факта привело к идее построения обобщённых перестраиваемых спектральных преобразований, наделённых быстрым алгоритмом.

Возможность перестройки значений весовых коэффициентов и многослойная структура алгоритма роднят быстрые перестраиваемые преобразования с многослойными нейронными сетями прямого распространения. Первые подходы к обучению быстрых ортогональных преобразований были развиты в работах А.И. Солодовникова и его научной группы [1]. В то время подобный класс спектральных преобразований называли приспособленными быстрыми преобразованиями. Условие ортогональности влечёт за собой совпадение размерностей входного и выходного вектора для каждого слоя алгоритма. Если отказаться от этого условия, то снимается ограничение и на равенство размерностей, причём, как оказалось, структуру быстрого алгоритма при этом можно сохранить. Отсюда был прямой путь к построению быстрых нейронных сетей (БНС) [2, 3].

В рамках данной парадигмы быстрые перестраиваемые преобразования можно рассматривать, как многослойные нейронные сети, которые отличаются от последних линейными функциями активации и нулевыми смещениями в нейронах. По структурной организации БНС следует отнести к модульным нейронным сетям с инъективными связями. Роль модуля выполняют базовые операции малой размерности, например, для БПФ это базовая операция типа «бабочка». Дальнейшие исследования позволили доказать, что структура быстрых преобразований подчиняется системным инвариантам, основанным на принципе самоподобия.

Идея использования топологии быстрых преобразований БПФ для построения нейронных сетей многократно повторяется и в научных работах современных исследователей, но топология БПФ используется, как правило, только для проектирования нейросетевого модуля в составе свёрточной нейронной сети [4, 5]. Экспериментально показано, что это приводит к увеличению точности в задаче классификации и к увеличению быстродействия сетей. Для уменьшения числа вычислительных операций в свёрточных сетях, начиная с работы ЛеКуна [6], предлагается свёртку вычислять в частотной области на основе алгоритма БПФ. Эта же решение используется в ряде статей [7-9] для разных видов свёрточных сетей. Во всех выше перечисленных работах для построения алгоритма БПФ используется метод, предложенный Кули и Тьюки ещё в 60-х годах прошлого века для быстрого преобразования с размерностью, кратной степени двойки. Теоретически этот метод может быть использован и для других составных размерностей, но аналитика выражений оказывается весьма громоздкой. Упрощение способа построения преобразования является одной из целей настоящей работы. Далее будет показано, что это достигается за счёт стратификации модельных представлений быстрых алгоритмов.

Оказалось, что модульность и самоподобие характерны и для биологических нейронных сетей. В 20 веке идеи модульной самоподобной организации естественных нейронных сетей были последовательно высказаны нейробиологом Рафаэль Лоренте де Но (Lorente de No) [10], нейрофизиологами Вернон Бенджамин Маунткаслем (Vernon Benjamin Mountcastle) и Джералд Морис Эдельманом (Gerald Maurice Edelman) [11, 12], исследовавших строение новой коры головного мозга человека (неокортекса). Нейробиологи исторически называют нейронные модули нейронными колонками. Про новую кору известно, что вся она гораздо однороднее по своему строению, чем думали раньше; стремительное увеличение её в процессе филогенеза произошло путём умножения колонок, одинаковых в своей основе, а не путём развития

новых типов нейронов или разных способов внутренней организации. Внутренняя структура коры везде одинакова и состоит из повторяющихся многоклеточных единиц.

Лавинообразное увеличение неокортекса является важной чертой эволюции млекопитающих; степень этого увеличения отличает приматов от остальных млекопитающих, а человека – от остальных приматов. Если период появления ветви человека оценивают в 20 млн. лет, то развитие неокортекса – всего 2 млн. лет. Нейрофизиологи объясняют такое быстрое развитие многократным повторением структуры коры через однотипные многоклеточные единицы, фактически эволюционно обосновывая самоподобное строение неокортекса. Косвенным подтверждением этого уже в новом столетии явились результаты исследований [13], доказавшие, что все нейронные клетки в пределах одной мини-колонки связаны с одним и тем же рецептивным полем, а смежные мини-колонки могут быть связаны с разными рецептивными полями, что подтверждает принцип инъективной связи, являющийся необходимым условием самоподобия.

1. Стратификация модельных представлений. Математические модели служат средством изучения биологических нейронных сетей и проектирования искусственных нейронных сетей. От модели требуется, чтобы она была простой, но функционально достаточной. Основной проблемой математического моделирования является нахождение приемлемого компромисса между детализацией и простотой описания. Один из путей решения этой проблемы заключается в формировании иерархически вложенных семейств моделей, где каждый уровень иерархии соответствует уровню разумного абстрагирования свойств системы, что ведёт к упрощению каждой частной модели. Такое многоуровневое представление модели принято называть стратификацией [14], а каждый уровень модельного представления – стратой. Стратификация моделей позволяет разделить проектирование нейронной сети на относительно независимые этапы и использовать для каждого этапа специфичный математический аппарат.

Далее будет показано, что страты модельных представлений самоподобной нейронной сети упорядочены по степени абстракции. Высший уровень абстракции соответствует морфологическому представлению, а низший – параметрическому. В математической формулировке стратификация связана с выделением на каждом уровне иерархии эквивалентных отношений и переходом к факторным моделям, описывающим следующий уровень.

2. Морфологические модели модульных нейронных сетей. На морфологическом уровне игнорируется внутреннее строение модулей, значение имеет только их присутствие и наличие связей между ними. Характеристики связей также не определены. Ключевыми моментами концепции морфологического строения нейронных сетей являются многослойность, модульность и самоподобие. В этом разделе будут построены математические модели, связывающие эти понятия воедино.

2.1. Самоподобие и морфогенез. По определению, компактное топологическое пространство X самоподобно, если существует конечное множество S , индексирующее набор несюръективных отображений $\{f_s\}_{s \in S}$ для которых

$$X = \bigcup_{s \in S} f_s(X).$$

Пример. Компактом X является замкнутый отрезок $[0,1]$. Позиция точки на отрезке (кроме крайне правой точки) в десятичной системе выражается в виде $z = 0.z_{-1}z_{-2} \dots z_{-n} \dots$ в общем случае с бесконечным числом разрядов, где $z_i = 0, 1, 2, \dots, 9$.

Ограничим в позиционном представлении число разрядов уровнем n , и будем считать, что число $0.z_{-1}z_{-2} \dots z_{-n}$ соответствует отрезку длиной 10^{-n} на компакте X , а значение числа указывает начальную позицию этого отрезка на интервале $[0,1]$. В частности, для $n = 1$ число

$0.z_{-1}$ будет соответствовать отрезку длиной 0.1, а значение разряда z_{-1} указывает его позицию на X . Введём для этого уровня множество функций $f_{z_{-1}}(X)$, отображающих компакт X на отрезок длиной 0.1 в позиции $z = 0.z_{-1}$, очевидно, что

$$X = \bigcup_{z_{-1}} f_{z_{-1}}(X).$$

В этом примере символ объединения можно заменить прямой суммой, поскольку при разбиении отрезки не пересекаются. Таким образом, можно записать:

$$X = \bigoplus_{z_{-1}} f_{z_{-1}}(X).$$

Процесс можно итеративно продолжить, выбирая в качестве компактов отрезки уже построенного разбиения. Тогда для уровня n получим:

$$X = \bigoplus_{z_{-1} z_{-2}} \dots \bigoplus_{z_{-n}} f_{\langle z_{-1} z_{-2} \dots z_{-n} \rangle}(X).$$

В данном выражении кортеж $\langle z_{-1} z_{-2} \dots z_{-n} \rangle$ определяет многомерный индекс множества несюръективных отображений, разбивающих компакт X на части. Для позиционной системы счисления соответствие между значением числа и его поразрядным представлением $z \leftrightarrow \langle z_{-1} z_{-2} \dots z_{-n} \rangle$ является взаимно-однозначными.

Рассмотренное выше классическое понятие самоподобия ориентировано на обслуживание математических моделей фракталов, и не вполне подходит для представления самоподобных объектов не фрактальной природы, поэтому существует необходимость введения обобщающего определения, включающее в себя фрактал как частный случай. Если обратиться к биологии, то подходящим понятием может служить морфогенез живых систем. Под морфогенезом в биологии понимается процесс возникновения новых структур и изменения их формы в ходе индивидуального развития организмов. Не ставя перед собой задачу построения математической модели морфогенеза, введём рабочее определение, достаточное для задач морфологического синтеза самоподобных сетей.

Определение 1. Будем говорить, что на популяции индексированных объектов $X = \{A_z\}$ определён морфогенез, если для каждого объекта популяции найдётся точное соответствие «родитель-потомок», однозначно индексирующее дочерние объекты.

Например, пусть $X = X_0 = \bigoplus_{z_0} A_{z_0}$ – начальная индексированная популяция, тогда направленное развитие популяции по поколениям определяется последовательностью индексированных популяций вида:

$$X_0 = \bigoplus_{z_0} A_{z_0}, \quad X_1 = \bigoplus_{z_1} \bigoplus_{z_0} A_{\langle z_1 z_0 \rangle}, \dots, X_{n-1} = \bigoplus_{z_{n-1}} \dots \bigoplus_{z_1} \bigoplus_{z_0} A_{\langle z_{n-1} \dots z_1 z_0 \rangle}, \dots$$

Если X – это компактное топологическое пространство, заданное отрезком единичной длины, и объектами популяций являются части отрезка, то все поколения объектов популяции совпадают с этим пространством, так что

$$X = X_0 = X_1 = \dots = X_{n-1} = \dots,$$

и если ещё потребовать конечности выбора правил индексации объектов популяции, то понятие морфогенеза трансформируется к определению самоподобия на компакте. Заметим теперь, что если ограничиться только условием конечности выбора правил индексации, то это приводит к самоподобным растущим популяциям, не связанным с фракталами. Если снять условие конечности выбора правил индексации, то процесс роста популяции вследствие морфогенеза в общем случае не будет самоподобным.

При неопределённости числа поколений затруднительно ответить на вопрос о конечности выбора правил индексации и самоподобия. В этом случае необходимо вводить дополнительные ограничения, выделяющие самоподобный процесс, например, если для всех поколений морфогенеза отображения «родитель-потомок» совпадают, то, очевидно, морфогенез тривиально самоподобен. Однако класс популяций тривиального самоподобного морфогенеза достаточно узкий, и для практических целей целесообразно его несколько расширить.

Определение 2. Морфогенез назовём *регулярным*, если в каждом поколении отображение «родитель-потомок» совпадает для всех объектов и однозначно определяется номером поколения.

Очевидно, что тривиальный самоподобный морфогенез является частным случаем регулярного, когда отображения «родитель-потомок» для всех поколений совпадают. Дальнейшее изложение будет в основном связано с регулярным морфогенезом. Проследивается явная автомодельность в описании регулярного и тривиального самоподобного морфогенеза, поэтому и для регулярного морфогенеза мы будем использовать термин самоподобный, делая необходимые пояснения в тех случаях, когда требуется выделить отличия.

2.2. Морфогенез многослойной самоподобной сети. Пусть многослойный сетевой граф Γ имеет n слоёв. Обозначим через $\{A_i^m\}$ – множество вершин в слое m , где $m = 0, 1, \dots, n-1$. Множество вершин входного (нулевого) слоя назовём *афферентом* сети и обозначим $Aff(\Gamma)$, а множество вершин конечного слоя назовём *эфферентом*¹ сети и обозначим $Eff(\Gamma)$. Пусть A_i^m – некоторая вершина сети в слое m . Назовём афферентом вершины (далее обозначается $Aff(A_i^m)$) подмножество вершин афферента сети, связанных дугами с вершиной A_i^m , таким образом, имеем $Aff(A_i^m) \subset Aff(\Gamma)$. Аналогично введём понятия эфферента вершины $Eff(A_i^m)$, как подмножество вершин эфферента сети, связанного дугами с вершиной A_i^m , таким образом, $Eff(A_i^m) \subset Eff(\Gamma)$. Афференты и эфференты вершин будем также называть *терминальными проекциями*.

Обозначим через $\Gamma^{-1}(A_i^m)$ рецепторную окрестность вершины A_i^m , а через $\Gamma(A_i^m)$ её аксоновую окрестность. Правило построения графа зададим следующими выражениями:

$$\begin{aligned} Aff(A_i^m) &= \bigoplus_{A_k^{m-1} \in \Gamma^{-1}(A_i^m)} Aff(A_k^{m-1}), \\ Eff(A_i^m) &= \bigoplus_{A_k^{m+1} \in \Gamma(A_i^m)} Eff(A_k^{m+1}). \end{aligned} \quad (1)$$

Эти правила были введены автором в 2004 году [15] и названы условиями слабой связанности, поскольку их реализация порождает слабосвязанные сети. Символ прямой суммы подчёркивает, что для любой вершины сети терминальные проекции вершин её окрестности не пересекаются.

В данном случае объектами популяции морфогенеза являются проекции вершин слоёв на терминальные множества. Номер слоя соответствует номеру поколения, причём для афферентов номер поколения возрастает от конечного слоя к начальному, а для эфферентов – от начального к конечному. Начальными популяциями при этом являются вершины терминальных слоёв. Условия слабой связанности справедливы как для регулярного, так и не регулярного морфогенеза.

¹ В биологии под афферентами понимают чувствительные нервные окончания (чувствительные рецепторы), которые расположены в органах и тканях и способны воспринимать разного рода раздражения. Эфферентами называют выходы нейронных подсистем, воздействующие на двигательные нервные волокна.

По отношению к проекциям на терминальное поле сети множество вершин слоя разбивается на афферентные и эфферентные классы. Афферентному классу принадлежат вершины нейронного слоя, имеющие общее афферентное множество, назовём эти классы афферентными доменами и обозначим $Dom_p(A_{z^m}^m) = X_{i^m}^m$, здесь $A_{z^m}^m$ - один из представителей домена, i^m - порядковый номер домена в слое m . Аналогично, отношение проекции вершин слоя m на терминальное эфферентное поле сети разбивает множество вершин конечного слоя на классы – эфферентные домены, которые обозначим $Dom_g(A_{z^m}^m) = Y_{j^m}^m$, где j^m - порядковый номер эфферентного домена в слое m .

Будем полагать, что n – слойный граф сети состоит из одной компоненты связности и удовлетворяет условиям (1). В этом случае на основании первого условия каждая вершина выходного слоя будет связана со всеми вершинами входного слоя, это означает, что в выходном слое существует всего один афферентный домен. Пусть этот домен имеет номер 0. При регулярном морфогенезе размеры всех рецепторных окрестностей вершин выходного слоя одинаковы, положим их равными p_{n-1} . Афференты вершин окрестностей вершины конечного слоя (т.е. слоя $n - 1$) не пересекаются и поэтому множество вершин слоя $n - 2$ распадается на домены, которые пронумеруем индексом i_{n-1} . Индекс будет изменяться в диапазоне от 0 до $p_{n-1} - 1$. Таким образом, единственный домен конечного слоя в результате первого шага морфогенеза трансформировался в p_{n-1} доменов слоя $n - 2$, которые обозначим $Dom_p^{n-2} \langle i_{n-1} \rangle$. Будем полагать, что каждая вершина слоя $n - 2$ имеет рецепторную окрестность, состоящую из p_{n-2} вершин. В слое $n - 3$ домены в результате морфогенеза получают номера, определяемые двойным индексом $Dom_p^{n-3} \langle i_{n-2} i_{n-1} \rangle$. Таким образом, достигнув начального слоя, получим, что нумерация доменов этого слоя будет определяться сложным индексом $Dom_p^0 \langle i_1 i_2 \dots i_{n-2} i_{n-1} \rangle$. Но в начальном слое каждый афферентный домен состоит только из одной вершины, поэтому построенный сложный индекс будет определять и номер вершины начального слоя. Афферентные домены промежуточных слоев индексируются выражением $Dom_p^m \langle i_{m+1} i_{m+2} \dots i_{n-2} i_{n-1} \rangle$.

Аналогично, каждая вершина входного слоя связана со всеми вершинами выходного слоя, это означает, что во входном слое существует всего один эфферентный домен. Для регулярного морфогенеза зададим размеры аксоновых окрестностей по слоям числами g_0, g_1, \dots, g_{n-2} и будем использовать индексы $j_0, j_1, \dots, j_{n-3}, j_{n-2}$ для нумераций доменов. Повторяя последовательность регулярных разбиений, получим, что для конечного слоя номер домена будет определяться сложным индексом $Dom_g^{n-1} \langle j_0 j_1 \dots j_{n-3} j_{n-2} \rangle$. Но в конечном слое каждый эфферентный домен состоит только из одной вершины, поэтому построенный сложный индекс будет определять и номер вершины конечного слоя. Эфферентные домены промежуточных слоев индексируются выражением $Dom_g^m \langle j_0 j_1 \dots j_{m-2} j_{m-1} \rangle$. Следующая теорема является обоснованием использования термина «слабосвязанная сеть».

Теорема. О параллельных путях. В слабосвязанных сетях отсутствуют параллельные пути.

Доказательство. Предположим противное. Пусть существуют два параллельных отрезка путей, которые начинаются в вершине A' и сходятся в вершине B , проходя через её окрестные вершины A_1 и A_2 . Так как A' – общая точка параллельных отрезков и она предшествует вершинам A_1 и A_2 , то афференты окрестных вершин A_1 и A_2 пересекаются так, что $Afr(A_1) \cap Afr(A_2) = Afr(A')$, но это противоречит условию слабой связности. Поскольку утверждение справедливо для любой вершины, включая терминальные, то отсюда следует невозможность параллельных путей для всей сети.

Теорема о морфологии слабосвязанной сети. В каждом слое слабосвязанной сети с одной компонентой связности афферентные и эфферентные домены попарно пересекаются во всех возможных сочетаниях точно по одной вершине, т.е. для каждой парной комбинации афферентного и эфферентного доменов существует единственная вершина $A_{z^m}^m$, такая, что

$$Dom_p(A_{z^m}^m) \cap Dom_g(A_{z^m}^m) = A_{z^m}^m.$$

Доказательство. Нужно доказать, что все парные сочетания разнотипных доменов имеют не пустое пересечение и что эти пересечения состоят только из одной вершины слоя.

Докажем первое. Предположим, что в слое m существует пара разнотипных доменов, которые не пересекаются по вершинам слоя. В этом случае вершины афферентного домена не связаны с частью вершин эфферента сети и поэтому в сети выделяется две правых конуса вершин, не связанных между собой во всех последующих слоях. С другой стороны, вершины эфферентного домена не связаны с частью вершин афферента сети и в сети выделяются два левых конуса, не связанных между собой в предшествующих слоях. Поскольку домены не пересекаются, то в этом случае в сети существует две не связанные между собой компоненты, что противоречит условию теоремы об однокомпонентности сети.

Докажем второе. Предположим, что домены пересекаются по двум вершинам A_1^m и A_2^m . Тогда эта пара вершин имеет общую афферентную окрестность, т.е. в слое $m-1$ существует, по крайней мере, одна вершина A_k^{m-1} , связанная с вершинами A_1^m и A_2^m , и её эфферент по условию морфогенеза (1), должен быть равен прямой сумме эфферентов вершин ее аксоновой окрестности. В эту прямую сумму входят и вершины A_1^m и A_2^m , но поскольку, по предположению, они принадлежат одному эфферентному домену слоя, то для них выполняется $Eff(A_1^m) \cap Eff(A_2^m) \neq \emptyset$ (см. рис. 1). Таким образом, мы пришли к нарушению условия морфогенеза (1) по эфферентам вершин. Аналогично это положение доказывается с использованием эфферентной окрестности пары вершин A_1^m и A_2^m .

Из этой теоремы следует, что для самоподобной сети существует взаимно-однозначное соответствие между номером вершины в пределах слоя и индексами пары пересекающихся разнотипных доменов. Для регулярной сети соответствие может быть задано кортежем, представляющим собой поразрядное представление числа в многоосновной системе счисления:

$$z^m = \langle j_0 j_1 \dots j_{m-1} i_{m+1} i_{m+2} \dots i_{n-1} \rangle. \tag{2}$$

В этом кортеже допустима любая перестановка индексов.

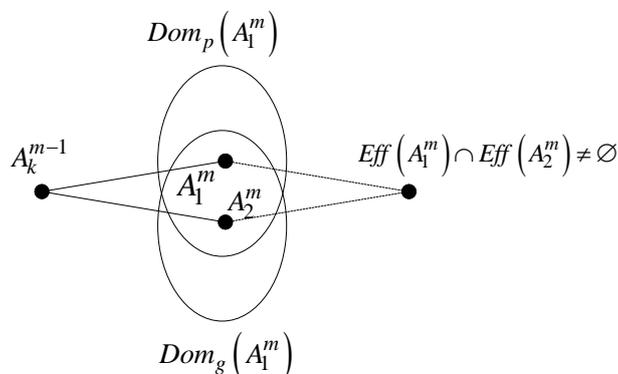


Рис. 1. Теорема о морфологии слабосвязанной сети.

3. Граф самоподобной сети. Из выражения (2), для слоя $m-1$ получим:

$$z^{m-1} = \langle j_0 j_1 \dots j_{m-2} i_m i_{m+1} \dots i_{n-1} \rangle.$$

Вершины смежных слоёв m и $m-1$ связаны дугой, если их афференты пересекаются. Афферент вершины z^{m-1} определяется разрядными числами кортежа $\langle i_m i_{m+1} \dots i_{n-1} \rangle$, а вершины z^m разрядными числами кортежа $\langle i_{m+1} i_{m+2} \dots i_{n-1} \rangle$, пересечение афферентов возможно только в том случае, когда одноименные разряды в приведённых кортежах будут совпадать. Такой же вывод можно сделать, используя условия пресечения эфферентов смежных слоёв m и $m+1$: вершины будут связаны дугой, если в кортежах $\langle j_0 j_1 \dots j_{m-1} \rangle$ и $\langle j_0 j_1 \dots j_{m-1} j_m \rangle$ одноименные разряды будут совпадать. Полученные правила позволяют достаточно просто выполнить построение графического образа самоподобной модульной сети. На рис. 2 показан пример построения четырёхслойной самоподобной сети для варианта, когда для всех m индексы принимают значения $i_m = \{0,1\}$ и $j_m = \{0,1\}$.

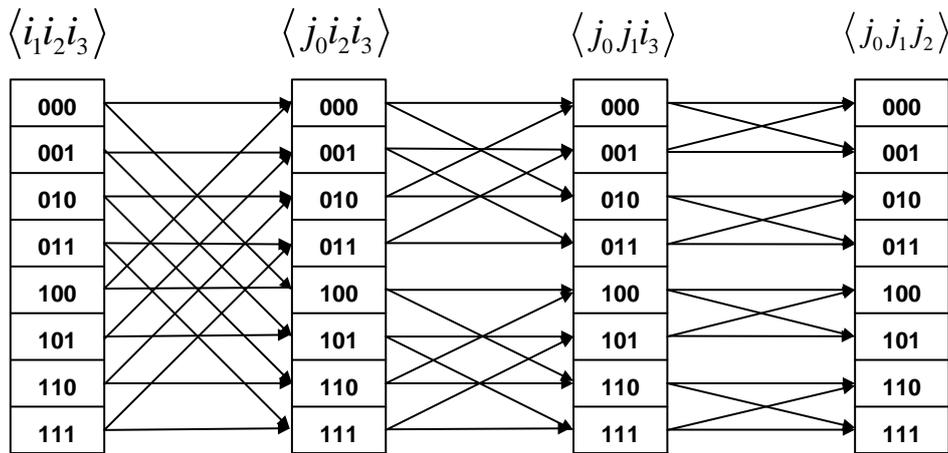


Рис. 2. Четырёхслойная самоподобная модульная сеть

Построенный граф полностью описывается лингвистическим предложением:

$$[\langle i_1 i_2 i_3 \rangle \langle j_0 i_2 i_3 \rangle \langle j_0 j_1 i_3 \rangle \langle j_0 j_1 j_2 \rangle].$$

Здесь каждое слово предложения представляет кортеж z^m номера вершины сети. В словах предложения допустима любая перестановка индексов.

4. Структурная модель модульной нейронной сети. Структурными характеристиками модульной нейронной сети являются размерности модулей по рецепторным и аксоновым полям и ранги межмодульных связей. Ограничим наше рассмотрение структурно-регулярной сетью, когда все модули в пределах слоя имеют одинаковые структурные характеристики, а все межмодульные связи имеют единичный ранг. Можно показать, что структурная регулярность влечёт за собой и морфологическую регулярность. Структурный синтез нерегулярных самоподобных сетей рассмотрен автором в работе [16].

Целесообразно упростить модель нейронной сети, полагая, что вся обработка сосредоточена в нейронных модулях, а связи осуществляют только передачу данных без внутренней обработки и искажений, т.е. осуществляют тождественное отображение с коэффициентом передачи равным единице. Такой случай характерен для алгоритмов быстрых преобразований типа БПФ.

Обозначим через p_m размерность рецепторных полей, а через g_m – размерности аксоновых полей нейронных модулей в слое m . Пример структурной модели модульной самоподобной нейронной сети с одноранговыми связями показан на рис. 3. Модули входного слоя имеют размерность $(3,2)$, а остальных слоёв – $(2,2)$. Веса всех дуг равны единицы и на рисунке не показаны. Структурная модель отличается от морфологической наличием весов вершин и дуг на графе модели.

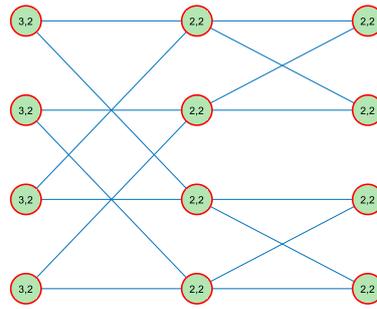


Рис. 3. Структурная модель регулярной самоподобной модульной нейронной сети с одноранговыми связями

5. Топологическая модель модульной нейронной сети. Входы и выходы структурной модели на рис. 3 являются трёх- и двух-координатными векторами без привязки к номеру отсчёта. Структурная модель непосредственно не переключается на алгоритм обработки в сети, её назначение – исследовать общие свойства модульной сети, связанные с вычислительной эффективностью и пластичностью. Для описания алгоритма необходимо построить топологическую модель. В топологической модели элементами рассмотрения являются физические контакты нейронных модулей; это либо входные рецепторы, либо выходные аксоны модуля.

Обозначим через u_m и v_m локальные номера рецепторов и аксонов для нейронного модуля в слое m . Диапазоны изменения локальных номеров определяются выражениями:

$$u_m = [0, 1, \dots, (p_m - 1)], \quad v_m = [0, 1, \dots, (g_m - 1)].$$

Позиционный номер рецептора в пределах нейронного слоя обозначим через U^m , а позиционный номер аксона через V^m . Совокупность взаимно-однозначных отображений вида:

$$\{u_m, z^m\} \rightarrow U^m, \quad \{v_m, z^m\} \rightarrow V^m$$

назовём *топологической моделью* нейронного слоя, здесь, как и прежде, z^m - номер модуля слоя m . Для регулярных нейронных сетей топологические отображения слоя могут быть выражены кортежами:

$$U^m = \langle \langle z^m \rangle \oplus u_m \rangle, \quad V^m = \langle \langle z^m \rangle \oplus v_m \rangle.$$

Символ \oplus в данном случае подчёркивает, что место размещения дополнительных разрядов u_m и v_m в кортеже $z^m = \langle j_0 j_1 \dots j_{m-1} i_{m+1} i_{m+2} \dots i_{n-1} \rangle$ может быть произвольным.

При одноранговых связях каждой дуге графа структурной модели однозначно ставится в соответствие дуга топологической модели, например, можно выбрать следующие взаимно-однозначные соответствия: $i_m \leftrightarrow u_m, \quad j_m \leftrightarrow v_m$, тогда возможен следующий вариант топологических отображений:

$$\begin{aligned} U^m &= \langle v_0 v_1 \dots v_{m-1} u_m u_{m+1} \dots u_{n-1} \rangle, \\ V^m &= \langle v_0 v_1 \dots v_{m-1} v_m u_{m+1} \dots u_{n-1} \rangle, \\ z^m &= \langle v_0 v_1 \dots v_{m-1} u_{m+1} \dots u_{n-1} \rangle. \end{aligned} \tag{3}$$

В особой ситуации находятся терминальные поля сети, поскольку для разрядных переменных u_0 и v_{n-1} нет соответствующих разрядов в кортеже z^m , но это не мешает построению кортежей топологических отображений слоя. Например, для структурной модели $[\langle i_1 i_2 i_3 \rangle \langle j_0 i_2 i_3 \rangle \langle j_0 j_1 i_3 \rangle \langle j_0 j_1 j_2 \rangle]$ можно построить следующие лингвистические предложения топологической модели:

Для рецепторных полей:

$$Rp = [\langle u_0 u_1 u_2 u_3 \rangle \langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle].$$

Для аксоновых полей:

$$Ax = [\langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle \langle v_0 v_1 v_2 v_3 \rangle].$$

Удобно эту модель представить одним лингвистическим предложением:

$$[\langle u_0 u_1 u_2 u_3 \rangle \langle v_0 u_1 u_2 u_3 \rangle \langle v_0 v_1 u_2 u_3 \rangle \langle v_0 v_1 v_2 u_3 \rangle \langle v_0 v_1 v_2 v_3 \rangle].$$

Первое слово в этом предложении соответствует рецепторному полю сети, а последнее – аксоновому. Граф топологической модели строится по тем же правилам, что и граф морфологической модели: дугами соединяются вершины, имеющие одинаковые значения разрядных переменных в смежных слоях. Граф топологической модели для данного примера показан на рис. 4.

Нетрудно убедиться, что данная модель соответствует графу быстрого преобразования Фурье (БПФ) в топологии Кули-Тьюки «с прореживанием по частоте» [17]. Другой вариант топологических отображений может быть задан в виде:

$$\begin{aligned} U^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} u_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle, \\ V^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle, \\ z^m &= \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle. \end{aligned} \quad (4)$$

Эта модель соответствует графу БПФ в топологии Кули-Тьюки «с прореживанием по времени». Граф данной топологической модели показан на рис. 5.

Из полученных результатов можно сделать вывод, что алгоритм быстрого преобразования Фурье является топологической реализацией самоподобной модульной сети, где модулями являются базовые операции типа «бабочка». Для быстрых нейронных сетей вместо базовой операции используется термин *нейронное ядро*.

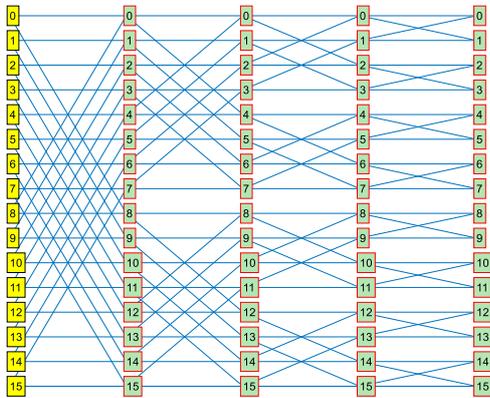


Рис. 4. Топологическая модель алгоритма БПФ «с прореживанием по частоте»

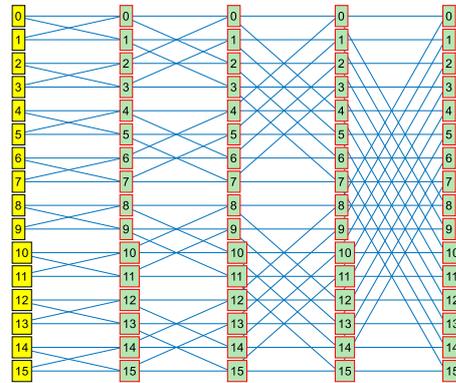


Рис. 5. Граф топологической модели алгоритма БПФ «с прореживанием по времени»

Нетрудно проверить, что для обеих выше приведённых топологий выполняется $V^{m-1} = U^m$, Из топологической модели (3) для терминальных слоёв получим:

$$U^0 = \langle u_{n-1} u_{n-2} \cdots u_1 u_0 \rangle, \quad V^{n-1} = \langle v_{n-1} v_{n-2} \cdots v_1 v_0 \rangle.$$

Если обозначить через N – размерность рецепторного поля, а через M – размерность аксонового поля сети, то из последних выражений непосредственно следует:

$$N = p_{n-1} \cdots p_1 p_0, \quad M = g_{n-1} \cdots g_1 g_0.$$

Таким образом, размерности терминальных полей сети определяются произведением размерностей нейронных модулей. На рис. 6 показан граф топологической модели (4) для структурной модели, показанной на рис. 3. Для наглядности на графе выделены базовые операции.

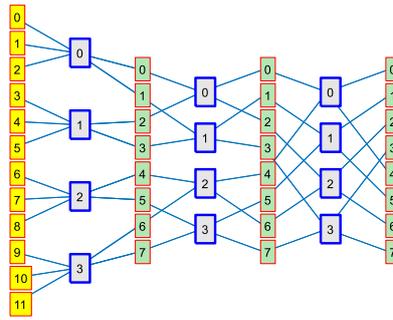


Рис. 6. Граф топологической модели с выделенными базовыми операциями

Рассмотрим ещё несколько типов регулярных топологий. На рис. 7 показан граф преобразования с топологией Гуда [18]. На графе явным образом выделены базовые операции. Работа Гуда по быстрому преобразованию Фурье появилась на семь лет раньше хорошо известной работ Кули и Тьюки [17], но прошла не замеченной.

Топологическая схема описывается моделью

$$U^m = \langle u_m u_{m+1} \dots u_{n-1} v_0 v_1 \dots v_{m-1} \rangle,$$

$$V^m = \langle u_{m+1} u_{m+2} \dots u_{n-1} v_0 v_1 \dots v_m \rangle,$$

$$i^m = \langle u_{m+1} u_{m+2} \dots u_{n-1} v_0 v_1 \dots v_{m-1} \rangle.$$

Особенность схемы Гуда состоит в том, что все слои графа имеют одинаковый топологический образ. Однако выходные операнды на каждой базовой операции не могут замещать входные отсчёты, поэтому требуется дополнительная память для хранения выходного вектора слоя. Рассмотренные выше топологические схемы Кули-Тьюки с «прореживанием по времени» и по «частоте» являются двойственными друг другу и могут быть получены зеркальным отражением относительно вертикальной оси. Алгоритмически переход к двойственной схеме выполняется заменой переменных в топологической модели по правилам: $U^m \leftrightarrow V^{n-1-m}$, $u_m \leftrightarrow v_{n-1-m}$. Двойственная топология Гуда описывается моделью:

$$U^m = \langle v_{m-1} v_{m-2} \dots v_0 u_{n-1} u_{n-2} \dots u_m \rangle,$$

$$V^m = \langle v_m v_{m-1} \dots v_0 u_{n-1} u_{n-2} \dots u_{m+1} \rangle,$$

$$i^m = \langle v_{m-1} \dots v_0 u_{n-1} u_{n-2} \dots u_{m+1} \rangle.$$

На рис. 8 показан граф быстрого преобразования с двойственной топологией Гуда.

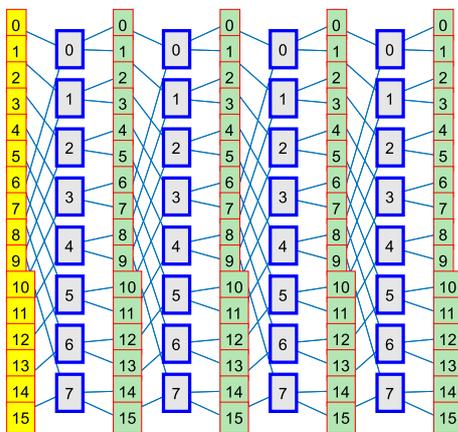


Рис. 7. Граф быстрого преобразования с топологией Гуда

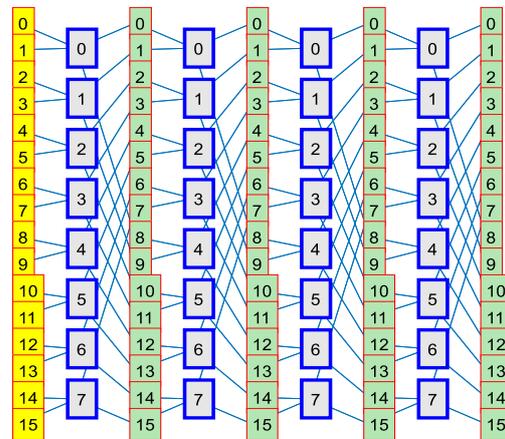


Рис. 8. Граф быстрого преобразования с двойственной топологией Гуда

Алгоритм БПФ для схем Кули-Тьюки по основанию 2 на выходе имеет спектр, гармоники которого имеют двоично-инверсную упорядоченность по частоте. Следующая топологическая схема позволяет получить на выходе спектр с естественным упорядочением по частоте при произвольных основаниях.

$$U^m = \langle u_m u_{m+1} \cdots u_{n-2} u_{n-1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle,$$

$$V^m = \langle u_{m+1} u_{m+2} \cdots u_{n-1} v_m v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle,$$

$$z^m = \langle u_{m+1} \cdots u_{n-2} u_{n-1} v_{m-1} v_{m-2} \cdots v_1 v_0 \rangle.$$

На рис. 9 показан граф быстрого преобразования, построенный по данной топологии. Двойственная топологическая модель описывается выражениями:

$$U^m = \langle v_{m-1} v_{m-2} \cdots v_1 v_0 u_m u_{m+1} \cdots u_{n-2} u_{n-1} \rangle,$$

$$V^m = \langle v_m v_{m-1} v_{m-2} \cdots v_1 v_0 u_{m+1} \cdots u_{n-2} u_{n-1} \rangle,$$

$$z^m = \langle v_{m-1} v_{m-2} \cdots v_1 v_0 u_{m+1} \cdots u_{n-2} u_{n-1} \rangle.$$

Граф алгоритма БПФ с естественным упорядочением спектра по частотам и двойственной топологией показан на рис. 10.

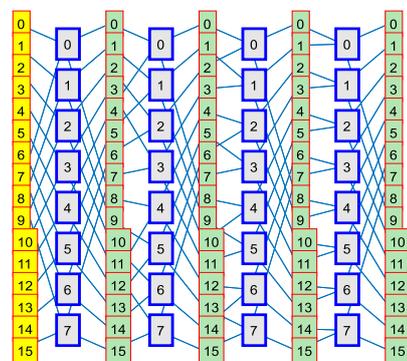
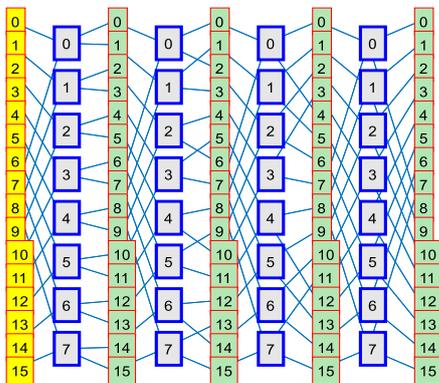


Рис. 9. Граф алгоритма БПФ с естественным упорядочением спектра по частотам следования

Рис. 10. Граф алгоритма БПФ с естественным упорядочением спектра по частотам следования и с двойственной топологией

Для быстрого преобразования Хаара используется следующая топологическая схема [3]:

$$U^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} u_m v_0 v_1 \cdots v_{m-2} v_{m-1} \rangle,$$

$$V^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_0 v_1 \cdots v_{m-2} v_{m-1} v_m \rangle,$$

$$z^m = \langle u_{n-1} u_{n-2} \cdots u_{m+1} v_0 v_1 \cdots v_{m-2} v_{m-1} \rangle.$$

Граф топологической схемы показан на рис. 11.

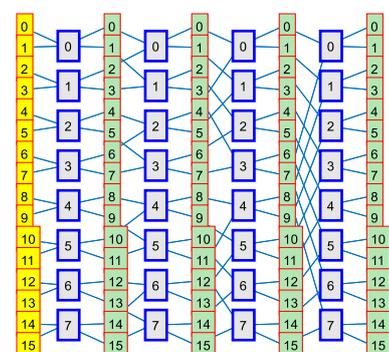
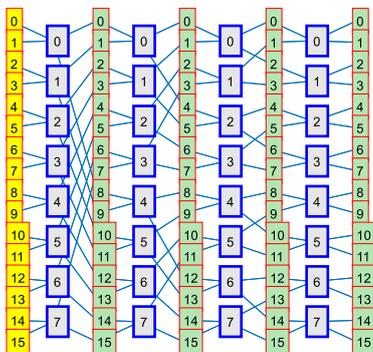


Рис. 11. Топологический граф преобразования быстрого преобразования Хаара

Рис. 12. Граф топологической схемы быстрого преобразования Хаара с двойственной топологией

Двойственна топологическая модель описывается выражением:

$$U^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} u_{n-1} u_{n-2} \cdots u_{m+1} u_m \rangle,$$

$$V^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} v_m u_{n-1} u_{n-2} \cdots u_{m+1} \rangle,$$

$$z^m = \langle v_0 v_1 \cdots v_{m-2} v_{m-1} u_{n-1} u_{n-2} \cdots u_{m+1} \rangle.$$

Топологический граф двойственной модели показан на рис. 12.

6. Параметрические модели быстрых преобразований. В графе быстрого алгоритма модулями являются базовые операции (нейронные ядра), представимые в виде матриц небольшой размерности. Для алгоритма быстрого преобразования в базовой операции z^m слоя m выполняется линейная обработка компонентов входного вектора слоя:

$$y_{z^m}^m(v_m) = \sum_{u_m} x_{z^m}^m(u_m) w_{z^m}^m(u_m, v_m),$$

где $x_{z^m}^m$ и $y_{z^m}^m$ – координаты входного и выходного векторов базовой операции; $w_{z^m}^m$ – матрица весов базовой операции. Для построения алгоритма преобразования необходимо перейти от локальных переменных ядра к глобальным переменным слоя. Этот переход реализуется на основе топологической модели. Параметрическое описание базовых операций вместе с топологической моделью образует параметрическую модель быстрого преобразования.

Заключение. В статье рассмотрен принцип построения четырехуровневой стратифицированной модели для самоподобных многослойных нейронных сетей с регулярной топологией. Стратификация позволяет исследовать систему на разных стадиях познания и описывать каждый уровень адекватными средствами. Стратифицирование можно рассматривать и как средство последовательного углубления представления о системе: при спуске вниз по иерархии страт система раскрывается в деталях; при подъёме на более абстрактные уровни яснее становится смысл и значение всей системы. В статье показано, что быстрые алгоритмы спектральных преобразований относятся к категории самоподобных нейронных сетей и могут быть представлены четырёхуровневой стратифицированной моделью.

Теоретическая основа быстрых алгоритмов долгое время базировалась на всевозможных теоремах факторизации, которые доказывали возможность разложения матрицы спектрального преобразования в произведение слабозаполненных матриц, где каждая матрица соответствует одному слою быстрого алгоритма [19, 20]. Это породило всплеск работ по теоремам факторизации. На этом пути исследователи столкнулись с тем обстоятельством, что существует большое множество различных матричных разложений для одного и того же спектрального преобразования. Когда число всевозможных теорем факторизации превысило десятки, стало понятным, что этот путь является тупиковым. Тем не менее, поток теорем не закончился и до сих пор. Проблема была чисто методическая и заключалась в смешивании понятий морфологии, структуры и топологии быстрого преобразования. Морфология является устойчивым системным инвариантом, свойственным всему классу быстрых алгоритмов, а топология – это не более чем допустимая реализация системного инварианта в связях между базовыми операциями. Каждая теорема факторизации соответствует одной из допустимых форм топологической реализации, а их число быстро растёт с ростом размерности преобразования. Автором было показано [21], что в основу теории быстрых преобразований следует положить именно системные инварианты, это устраняет необходимость изобретения новых теорем факторизации и позволяет предложить общий метод построения различных топологических реализаций быстрого преобразования при неизменной структуре. Предложенный метод существенно упрощает аналитику при построении быстрых преобразований с произвольными составными размерностями. Более того, выделение структурного и топологического уровней

даёт возможность решать новые задачи, связанные с оценкой параметрической и топологической пластичности перестраиваемых преобразований.

Список источников

1. Солодовников А.И. Основы теории и методы спектральной обработки информации / А.И. Солодовников, А.М. Спиваковский. – Л., 1986. – 272 с.
2. Дорогов А.Ю. Быстрые нейронные сети / А.Ю. Дорогов. – СПб.: Изд-во С.Петербург. ун-та, 2002. – 80 с.
3. Дорогов А.Ю. Теория и проектирование быстрых перестраиваемых преобразований и слабосвязанных нейронных сетей / А.Ю. Дорогов. – СПб.: «Политехника», 2014. – 328с.
4. Keivan Alizadeh Vahid, Anish Prabhu, Ali Farhadi, Mohammad Rastegari University of Washington butterfly transform: an efficient FFT based neural architecture design. CVPR-2020, Conference on computer vision and pattern recognition June 14-19, 2020, pp. 12024-12033.
5. Yingzhou Li, Xiuyuan Cheng, Jianfeng Lu Butterfly-Net: Optimal function representation based on convolutional neural networks, 2019, DOI: 10.48550/arXiv.1805.074512.
6. Mathieu M., Henaff M., LeCun Y., Fast training of convolutional networks through FFTs, 2013, DOI: 10.48550/arXiv.1312.5851.
7. Varsha Nair, Moitrayee Chatterjee, Neda Tavakoli, Akbar Siami Namin, Craig Snoeyink Fast fourier transformation for optimizing convolutional neural networks in object recognition, 2020, DOI: 10.48550/arXiv.2010.04257.
8. Shaohua Li, Kaiping Xue1, Bin Zhu, Chenkai Ding, Xindi Gao, David Wei, TaoWan. FALCON: A Fourier transform based approach for fast and secure convolutional Neural Network predictions. CVPR-2020, Conference on Computer Vision and Pattern Recognition June 14-19, 2020.
9. Chenlin Meng, Linqi Zhou, Kristy Choi, Tri Dao, Stefano Ermon. Butterfly Flow: building invertible layers with butterfly matrices. Proceedings of the 39-th International conference on machine learning, Baltimore, Maryland, USA, PMLR 162, 2022.
10. Lorente de Nó R. Cerebral Cortex: Architecture, intracortical connections, motor projections. Physiology of the nervous system. Oxford University Press, New York, 1949, pp. 288-312.
11. Эделмен Дж. Разумный мозг / Дж. Эделмен, В. Маунткасл. – М.: Мир, 1981. – 135.
12. Vernon B Mountcastle Sensory hand: neural mechanisms of somatic sensation. Cambridge, Harvard university press, 2005.
13. Jones E.G. Microcolumns in the cerebral cortex [Free PMC Article (full text)]. Proceedings of the National Academy of Sciences of the United States of America: journal, 2000, vol. 97, no. 10, pp. 5019-5021, PMID 10805761.
14. Волкова В.Н. Основы теории систем и системного анализа: Учеб. для студентов вузов / В.Н. Волкова, А.А. Денисов. – СПб.: Изд-во СПбГТУ, 1999. – 512 с.
15. Дорогов А.Ю. Быстрые нейронные сети: Проектирование, настройка, приложения. // Лекции по нейроинформатике Ч.1. В тр. школы-семинара «Современные проблемы нейроинформатики», науч.-техн. конф. «Нейроинформатика-2004» 28–30 января 2004 г. Москва. – М.: МИФИ, 2004, с. 69-135.
16. Дорогов А.Ю. Структурный синтез модульных слабосвязанных нейронных сетей. Часть 2 Ядерные нейронные сети / А.Ю. Дорогов // Кибернетика и системный анализ, 2001. – № 4. – с.13-20.
17. Good I.J. The Interaction algorithm and practical fourier analysis. Journal of Royal Statistical Soseity. Ser.B.- 1958, vol. 20, no.2, pp. 361–372.
18. Cooley J., Tukey J. An algorithm for the machine calculation of complex Fourier series. Math. Comput, 1965, vol. 19, pp. 297-301.
19. Рабинер Л. Теория и применение цифровой обработки сигналов / Л. Рабинер, Б. Гоулд. – М.: Мир, 1978. – 848 с.
20. Дагман Э.Е. Быстрые дискретные ортогональные преобразования / Э.Е. Дагман, Г.А. Кухарев // Новосибирск: Наука. – 1983. – 228 с.
21. Дорогов А.Ю. Структурные и топологические инварианты быстрых перестраиваемых преобразований Сб. тр. VIII Науч.-техн. Всерос. научн. конф. «Нейроинформатика-2006» 24-27 января 2006г. – М.: МИФИ, 2006. – Ч.1. – С.39-50.

Дорогов Александр Юрьевич. Доктор технических наук, доцент, профессор кафедры «Автоматики и процессов управления» Санкт-Петербургского государственного электротехнического университета (СПбГЭТУ) «ЛЭТИ», главный научный сотрудник ПАО «Информационные телекоммуникационные технологии». Основные направления исследований автора связаны с интеллектуальным анализом данных, цифровой обработкой сигналов, проектированием быстрых преобразований и нейронных сетей быстрого обучения, разработкой аналитических платформ, моделированием радио-телекоммуникационных систем. AuthorID: 17611, SPIN 8645-5873, ORCID 0000-0002-7596-6761, vaksa2006@yandex.ru, г. Санкт-Петербург, ул. Попова, 5.

UDC 004.032.26

DOI: 10.25729/ESI.2023.32.4.001

Fast transformations and self-similar deep learning neural networks Part 1. Stratified models of self-similar neural networks and fast transformations

Alexander Yu. Dorogov

PJSC "Information Telecommunication Technologies" ("Inteltech"),

St. Petersburg State Electrotechnical University, Russia, St. Petersburg, *vaksa2006@yandex.ru*

Abstract. The paper shows that the construction of fast transformations (similar to FFT) is based on self-similar structures that can equally be used to build fast neural networks (FNN). It is shown that the class of fast transformations is determined by system invariants of the morphological level and can be described as the morphogenesis of terminal projections of neural modules. Linguistic models are proposed to describe the morphology, structure and topology of regular self-made neural networks. The models are easily generalized to multidimensional variants of neural networks of this class. Due to their structure, FNN have special learning algorithms that are fundamentally different from the classic ErrorBackPropagation by the absence of a mechanism for error back propagation. The learning algorithms are based on the methods of multiplicative factorization of images and fast transformations proposed in the work. The developed algorithms are completed in a finite number of steps with guaranteed convergence. The consistent development of the concept of self-similarity leads to the development of methods for creating fast neural networks with a deep degree of learning. Self-similar neural networks have a unique opportunity to learn to new data without losing previously acquired knowledge. It is shown that FNN can be used to create high-speed random access image memory and complex combinational logic devices. The paper presents the results of the author's research on the following issues: biological prerequisites for self-similarity of neural networks; self-similar multilayer structures, morphogenesis, stratification of model representations; algorithms for fast transformations, fast neural networks, tuning methods; training of FNN to reference functions; plasticity of FNN; pyramidal neural networks of deep learning; multi-channel correlators; implementation of memory and combinational logic on pyramidal structures. The research results will be presented in three parts of the article.

Keywords: self-similar structures, fast transformations, morphogenesis, fast neural networks, pyramidal structures, image memory, combinational logic

References

1. Solodovnikov A.I., Spivakovskij A.M. Osnovy teorii i metody spektral'noj obrabotki informacii [Fundamentals of the theory and methods of spectral information processing]. Leningrad, 1986, 272 p.
2. Dorogov A.Yu. Bystryye neyronnyye seti [Fast neural networks]. St. Petersburg Publishing House. Univ., 2002, 80 p.
3. Dorogov A.Yu. Teorija i proektirovanie bystryh perestraivaemyh preobrazovanij i slabosvjazannyh neyronnyh setej [Theory and design of fast tunable transformations and loosely coupled neural networks]. St. Petersburg, Politehnika [Politehnika], 2014, 328 p.
4. Keivan Alizadeh Vahid, Anish Prabhu, Ali Farhadi, Mohammad Rastegari University of Washington butterfly transform: an efficient FFT based neural architecture design. CVPR-2020, Conference on computer vision and pattern recognition June 14-19 2020, pp. 12024-12033.
5. Yingzhou Li, Xiuyuan Cheng, Jianfeng Lu Butterfly-Net: Optimal function representation based on convolutional neural networks, 2019, DOI: 10.48550/arXiv.1805.074512.
6. Mathieu M., Henaff M., LeCun Y., Fast training of convolutional networks through FFTs, 2013, DOI: 10.48550/arXiv.1312.5851.
7. Varsha Nair, Moitrayee Chatterjee, Neda Tavakoli, Akbar Siami Namin, Craig Snoeyink Fast fourier transformation for optimizing convolutional neural networks in object recognition, 2020, DOI: 10.48550/arXiv.2010.04257.
8. Shaohua Li, Kaiping Xue, Bin Zhu, Chenkai Ding, Xindi Gao, David Wei, TaoWan. FALCON: A Fourier transform based approach for fast and secure convolutional Neural Network predictions. CVPR-2020, Conference on Computer Vision and Pattern Recognition June 14-19 2020.
9. Chenlin Meng, Linqi Zhou, Kristy Choi, TriDao, Stefano Ermon. Butterfly Flow: building invertible layers with butterfly matrices. Proceedings of the 39-th International conference on machine learning, Baltimore, Maryland, USA, PMLR 162, 2022.

10. Lorente de Nó R. Cerebral Cortex: Architecture, intracortical connections, motor projections. Physiology of the nervous system. Oxford University Press, New York, 1949, pp. 288-312.
11. Gerald M. Edelman, Vernon B. Mountcastle Razumnyy mozg [Intelligent brain]. M.: Mir, 1981, 135p.
12. Vernon B Mountcastle Sensory hand: neural mechanisms of somatic sensation. Cambridge, Harvard university press, 2005.
13. Jones E.G. Microcolumns in the cerebral cortex [Free PMC Article (full text)]. Proceedings of the National Academy of Sciences of the United States of America: journal, 2000, vol. 97, no. 10, pp. 5019-5021, PMID 10805761.
14. Volkova V.N., Denisov A.A Osnovy teorii sistem i sistemnogo analiza: Ucheb. dlya studentov vuzov [Fundamentals of systems theory and system analysis: Textbook. for university students]. St. Petersburg, Publishing house of St. Petersburg State Technical University, 1999. 512 p.
15. Dorogov A.Yu. Bystryye neyronnyye seti: Proyektirovaniye, nastroyka, prilozheniya [Fast neural networks: Design, configuration, applications]. Lekcii po nejroinformatike Ch.1. V tr. shkoly-seminara "Sovremennyye problemy nejroinformatiki", nauch.-tehn. konf. "Nejroinformatika" [Proceedings of the school-seminar "Modern problems of neuroinformatics", scientific and technical conf. "Neuroinformatics"]. Moscow, MIFI, 2004, pp. 69-135.
16. Dorogov A.Y. Strukturnyy sintez modul'nykh slabosvyazannykh neyronnykh setey. Chast' 2 Yadernyye neyronnyye seti [Structural synthesis of modular weakly connected neural networks. II Nuclear neural networks]. Kibernetika i sistemnyy analiz [Cybernetics and system analysis], 2001, vol. 37, no. 4, pp. 470-477.
17. Cooley J., Tukey J. An algorithm for the machine calculation of complex Fourier series. Math. Comput, 1965, vol. 19, pp. 297-301.
18. Good I.J. The Interaction algorithm and practical fourier analysis. Journal of Royal Statistical Soseity. Ser.B.- 1958, vol. 20, no.2, pp. 361-372.
19. Lawrence R. Rabiner, Bernard Gold Teoriya i primeneniye tsifrovoy obrabotki signalov [Theory and application of digital signal processing]. M., Mir, 1978, 848 c.
20. Dagman Je. E., Kuharev G.A. Bystrye diskretnyye ortogonal'nye preobrazovaniya [Fast discrete orthogonal transformations] Publ., Novosibirsk: Nauk :, 1983. 228 p.
21. Dorogov A.Yu. Strukturnyye i topologicheskiye invarianty bystrykh perestraivayemykh preobrazovaniy [Structural and topological invariants of fast tunable transformations]. Sb. tr. VIII Nauch.-tekhn. Vseros. nauchn. konf. "Neyroinformatika-2006" 24-27 yanvarya 2006g [Sb. tr. VIII Scientific and technical. All-Russian scientific conf. "Neuroinformatics-2006" January 24-27, 2006], M, MEPhI, 2006, Part 1, pp.39-50.

Dorogov Alexander Yurievich. Doctor of technical sciences, associate professor, professor of the department of Automation and Control Processes of St. Petersburg state electrotechnical university (SPbETU) "LETI", Chief researcher of PJSC "Information telecommunication technologies. The main directions of the author's research are related to data mining, digital signal processing, the design of fast transformations and neural networks of fast learning, the development of analytical platforms, modeling of radio and telecommunications systems. AuthorID: 17611, SPIN 8645-5873, ORCID 0000-0002-7596-6761, vaksa2006@yandex.ru , St. Petersburg, st. Popova, 5.

Статья поступила в редакцию 13.06.2023; одобрена после рецензирования 16.10.2023; принята к публикации 13.12.2023.

The article was submitted 06/13/2023; approved after reviewing 10/16/2023; accepted for publication 12/13/2023.