

Решение задачи сверхразрешения с использованием полносвязной модели нейронной сети

September 20, 2023

Аннотация

В работе приведена математическая модель цифровой антенной решетки (ЦАР) для определения положения объектов в угломерном пространстве. Поставлена задача определения положения нескольких объектов с ЦАР различной конфигурации. Сформулирована задача определения значений коэффициентов принятого сигнала на приемных устройствах решетки, позволяющая решать задачу сверхразрешения на основе задачи экстраполяции. Приведен метод решения задачи экстраполяции на основе модели полносвязной нейронной сети прямого распространения. Описаны практические результаты решения задачи сверхразрешения как с выбранной моделью, так и в сравнении с другим методом экстраполяции.

Введение

Повышение информативности угломерных систем является актуальной научной и технической проблемой. Поисковые задачи, связанные с определением положения или очерчиванием границ объектов, находят большое приложения в таких различных областях науки как сейсмика, геология, акустика, электромагнетизм и т.д. Улучшение точности угловых измерений и эффективной угловой разрешающей способности с использованием новых методов цифровой обработки сигналов является важным шагом в направлении развития таких информационных радиотехнических систем.

Радиолокационные системы (РЛС), основанные на цифровых антенных решетках (ЦАР), преобразуют сигналы в цифровую форму и обрабатывают их с помощью специальных алгоритмов. ЦАР способны анализировать и обрабатывать гораздо больше информации, содержащейся в структуре принимаемого сигнала, чем обычные антенны. В результате возможно достижение разрешающей способности, превышающей предел, определенный критерием Рэлея - сверхразрешение.

Методы обработки принятого сигнала на основе ЦАР подходят для широкого класса задач, основанных на волновых процессах, где необходимо исследовать положение объектов и границ, например в задачах гидролокации, сейсмических исследованиях, акустике. Приведенные задачи относятся к классу поисковых, в которых важен как результат различия разных объектов так и точность их поиска, что увеличивает значимость данного исследования.

Угловое разрешение, основанное на критерии Рэлея, может быть выражено следующим образом:

$$\delta\theta \approx \lambda / D \quad (1)$$

где D - линейный размер антенны, λ - длина волны, которая соответствует ширине диаграммы направленности (ДН) до половинной мощности.

Из-за важности задачи повышения углового разрешения во многих странах разрабатываются и внедряются методы цифровой обработки сигналов, которые позволяют превысить (1) и достигать сверхразрешения. [1, 2, 3, 4, 5, 6]

Постановка задачи

Принятый со стороны ЦАР сигнал $U(\alpha)$ может быть модельно представлен в виде интегрального преобразования:

$$U(\alpha) = \int_{\Omega} F(\alpha - \varphi) I(\varphi) d\varphi, \quad (2)$$

где $\Omega(\alpha)$ - угловая область (сектор сканирования) расположения источника сигнала; $I(\alpha)$ - угловое распределение амплитуды отраженного (или излучаемого сигнала), равное 0 вне Ω ; $F(\alpha)$ - ДН измерительной системы. Достижимая при проведении прямых измерений в соответствии с (2) угловая разрешающая способность совпадает с классической величиной (1).

Для ЦАР ставится задача на основе интеллектуального анализа принятого сигнала $U(\alpha)$ и известной ДН системы $F(\alpha)$ восстановить изображение источника сигналов $I(\alpha)$ с угловым разрешением, превышающим критерий Рэлея, во много более возможном масштабе.

В математическом аспекте задача сводится к приближенному решению (2), рассматриваемому как линейное интегральное уравнение (ЛИУ) Фредгольма первого рода типа свёртки относительно неизвестной функции $I(\alpha)$, с максимально достижимой точностью.

В общем случае попытки увеличить разрешение по сравнению с (1) путём решения ЛИУ (2) приводят к появлению неустойчивых решений, т.к. поставленная задача относится к классу обратных и не удовлетворяет второму и третьему требованию корректности задачи по Адамару [3, 7, 8, 9, 10].

Метод решения

Для ЦАР возможен иной подход к задаче, напрямую не сводящийся к решению ЛИУ (2), что повышает устойчивость используемых алгоритмов.

Для упрощения получаемых выражений рассмотрим большую линейную ЦАР с числом элементов $2N + 1$. Амплитуды токов на излучателях полагаем равными, область расположения источника Ω - одномерной, с размерами меньшими ширины луча ЦАР $\theta_{0.5}$. Значения амплитудных ДН отдельных излучателей ЦАР в пределах $\theta_{0.5}$ считаем не изменяющимися. Тогда, заменив $\sin(\alpha) \approx \alpha$, ДН решётки можно представить в виде:

$$F(\varphi - \alpha) = \sum_{n=-N}^N \exp(ikdn(\varphi - \alpha)) \quad (3)$$

где $k = 2\pi/\lambda$ - волновое число, d - расстояние между соседними излучателями, α - направление максимума ДН. Для ЦАР сигнал (2) при введённых условиях преобразуется в сумму:

$$U(\alpha) = \sum_{n=-N}^N \exp(-ikdn\alpha) \int_{\Omega} \exp(ikdn\varphi) I(\varphi) d\varphi = \sum_{n=-N}^N C_n \exp(-ikdn\alpha), \quad (4)$$

$$C_n = \int_{\Omega} \exp(ikdn\varphi) I(\varphi) d\varphi, n = 0, \pm 1, \dots, \pm N, \quad (5)$$

где коэффициенты C_n - сигнал, принятый n -ым элементом ЦАР.

Обобщение (4) и (5) и получаемых в дальнейшем выражений на общий случай, в том числе и на двумерные задачи, принципиальных трудностей не вызывает.

Ядро ЛИУ (2) для ДН в виде (3) является вырожденным. Несложно показать, что в этом случае решение $I(\alpha)$ выражается посредством коэффициентов C_n :

$$I(\alpha) = A \sum_{n=-N}^N \exp(-ikdn\alpha) C_n + \psi(\alpha) \quad (6)$$

где $\psi(\alpha)$ - произвольная функция, ортогональная всем собственным функциям ядра, т.е. экспонентам из (6) на отрезке $[-\lambda/d; \lambda/d]$, A - нормировочный коэффициент.

Конкретный вид функции $\psi(\alpha)$ не может быть непосредственно найден из условий задачи. Для её определения или хотя бы оценки требуется какая-либо априорная информация. Приближенное

решение, которое представлено первым слагаемым суммы (6), повторяет по форме сигнал $U(\alpha)$, и следовательно, полученное разрешение соответствует критерию Рэлея.

Рассмотрим одну из возможностей увеличения углового разрешения ЦАР. В отличие от обычных фазированных антенных решеток (ФАР), сигналы, принятые каждым излучателем ЦАР - C_n , могут быть зарегистрированы в цифровом виде и применены далее при цифровой обработке.

Из (5) следует, что различия значений C_n , $n = 1, 2, \dots$ определяются только пространственным положением отдельных излучателей ЦАР. Следовательно, проанализировав расположение достаточно большого количества излучателей и значения соответствующего им набора коэффициентов C_n , можно с определенной точностью найти зависимость принимаемого сигнала C_j от координат или положения излучателя на решетке, что позволит с определенной точностью оценить значения принимаемого сигнала виртуальными излучателями за пределами апертуры ЦАР и, с определенной точностью, продолжить эту зависимость. Далее, используя прогнозные значения C_j , $j = \pm(N+1), \pm(N+2), \dots, \pm M$ вместе с измеренными ранее значениями C_j , $j = -N, -N+1, \dots, N$, распространим суммирование в (4) до $2M+1$ членов. В результате получим ДН виртуальной ЦАР с $2M+1$ излучателями и увеличенной в M/N раз апертурой. Ширина ДН уменьшилась во столько же раз, и при идеальных условиях, в M/N раз увеличиваются точность угловых измерений и угловая разрешающая способность.

В итоге, восстановленное со сверхразрешением приближенное изображение объекта представлено в виде:

$$I(\alpha) \approx \sum_{n=-M}^M \exp(-ikdn\alpha) C_n \quad (7)$$

где M - количество приемных устройств виртуальной ЦАР в одну сторону от её центра.

Чем больше синтезированная апертура ЦАР, т.е. чем уже ДН, тем, в соответствии с (1), выше угловое точное решение и точнее найденное приближенное решение (7). В пределе при $N \rightarrow \infty$ найдем точное решение, полученное с помощью ДН в виде дельта-функции, которое представляет собой разложение $I(\alpha)$ в ряд Фурье на соответствующем интервале.

Таким образом, приближённое решение ЛИУ (2-4) сведено к определению с значений возможно большего количества комплексных коэффициентов C за пределами апертуры ЦАР. Восстановление значений коэффициентов C для виртуальных элементов ЦАР (т.е. за пределами реальной апертуры) может быть выполнено методами экстраполяции.

Для примера задачи обнаружения двух равноинтенсивных объектов представим распределение интенсивности источников излучения $I(\alpha)$, $\alpha \in \Omega$ в виде комбинации дельта-функций Дирака [22]:

$$I(\alpha) = \delta(\alpha - t_1) + \delta(\alpha - t_2), \quad t_1, t_2 \in \Omega, \quad (8)$$

где t_1, t_2 - известные положения объектов в угловом одномерном пространстве Ω . Изобразим возможную зависимость коэффициентов принятого сигнала на приемных устройствах в зависимости от его возможного положения в апертуре ЦАР с равноудаленными приемными устройствами (рис.1а).

На рисунке выше (рис.1а) можно предположить что теоретическая функция действительных значений принятого сигнала в зависимости от номера устройства является четной, то есть симметричной относительно $n = 0$. Экстраполяция действительных значений необходима лишь в одну сторону с зеркальным отображением возможных предсказанных значений в другую относительно центрального устройства. Заметим также что модельная теоретическая зависимость несимметричных целей является осциллирующей как комбинация нескольких гармонических составляющих, а также затухающей по амплитуде значений, что показывает возможную сложность в экстраполяции значений на основе классических интегральных кривых решения динамических систем.

Также приведем иллюстрацию зависимости мнимых частей значений принятого сигнала на физических и модельных виртуальных устройствах ЦАР (рис.1б) в случае, если объекты расположены несимметрично относительно центра области наблюдения $|t_1| \neq |t_2|$, то есть объекты неравноудалены от начала координат.

На рисунке (рис. 1б) заметим, что теоретическая модельная зависимость возможных мнимых значений принятого сигнала на виртуальных устройствах при конкретной реализации объектов в угломерном пространстве представлена в виде нечетной функции, то есть в виде функции

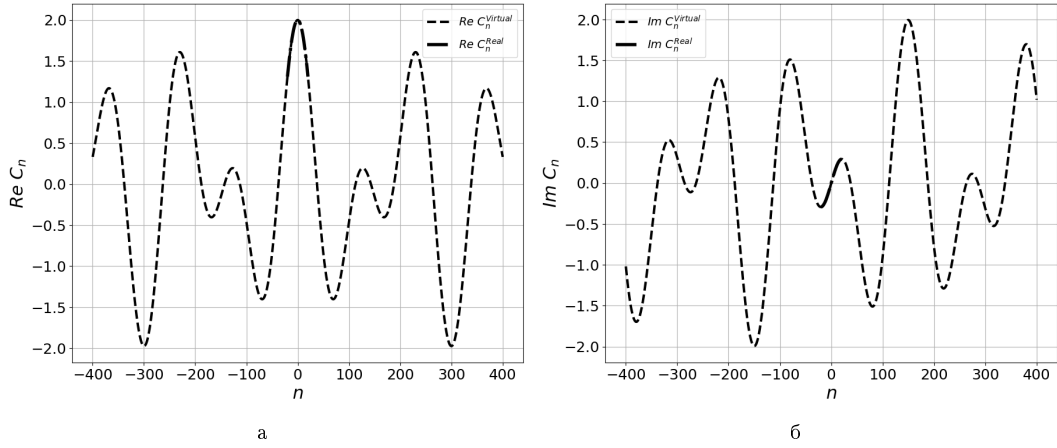


Figure 1: (а) действительная часть комплексных значений принятого сигнала на приемных устройствах в зависимости от номера позиции в ЦАР, (б) Мнимая часть комплексных значений принятого сигнала на приемных устройствах в зависимости от номера позиции в ЦАР; сплошная жирная линия - значения на физических устройствах ЦАР в соответствии с выражением (5); штриховая линия - значения на возможных виртуальных устройствах ЦАР в соответствии с выражением (5) в предположении о возможном расширении физической конфигурации ЦАР

сопряженных значений относительно центра ЦАР $n = 0$. Из графиков зависимостей (рис1) можно сделать вывод о достаточности решения задачи экстраполяции комплексных значений принятого сигнала на устройствах ЦАР при заданных K физических устройствах только в одну сторону от центра ЦАР, то есть в положительную полуось положений виртуальных устройств и значения в отрицательной полуоси представлять сопряженными к полученным значениям.

В результате решения задачи экстраполяции на основе полученных оценок \hat{C}_n , $n \in [-M, M]$ возможно восстановить распределение принятого сигнала $U(\alpha)$ на ЦАР так, как если бы в сканировании области с интенсивностью $I(\alpha)$ принимало участие $2 \cdot M + 1$ устройств. По итогам мы имеем возможность получить изображение объектов с большим разрешением, превышающим критерий Рэля для физической ЦАР заданной конфигурации (рис.2).

Существует множество видов экстраполяции (линейные, полиномиальные, сплайновые и т. д.), которые характеризуются различными показателями сложности вычислений и быстродействием. В данной работе приводится решение задачи экстраполяции значений принятого сигнала на основе полносвязной модели нейронной сети прямого распространения [18], [19].

Принцип прогнозирования зависимостей на основе алгоритмов искусственных нейронных сетей (ИНС) широко применяется в различных областях науки и техники последние 20 лет. ИНС часто используются в прикладных задачах распознавания образов, детекции объектов на изображениях и видео, генерации и обработки текстовой информации, обработки аудиосигналов, а также в различных прикладных задачах робототехники и управления.

Существует множество разновидностей и архитектур нейронных сетей прямого распространения. Для одномерной задачи, поставленной ранее, мы имеем возможность использовать как полносвязную искусственную нейронную сеть прямого распространения [11, 12, 13], так и свёрточные аналоги [14, 15]. Полносвязные нейронные сети прямого распространения применяются в задачах обработки одномерных зависимостей, ввиду простоты архитектуры нейронной сети, скорости работы и эффективности использования вычислений с одномерными зависимостями. По скорости работы и принципу действия в одномерном случае полносвязные и свёрточные нейронные сети идентичны, для ядер свертки размером равным длине входных данных (размерности вектора данных) в свёрточных сетях. Для простоты пояснения и построения системы экстраполяции зависимости будем использовать классическую реализацию полносвязной нейронной сети прямого распространения, показав по результатам возможности подобного класса алгоритмов решать поставленную задачу.

Полносвязная нейросетевая модель прямого распространения является функцией, представляемой в виде:

$$\hat{y} = FCNN(x|W), \quad (9)$$

где W - вектор матриц весов (параметров) модели, $x \in \mathbb{R}^f$ - вектор значений входного примера,

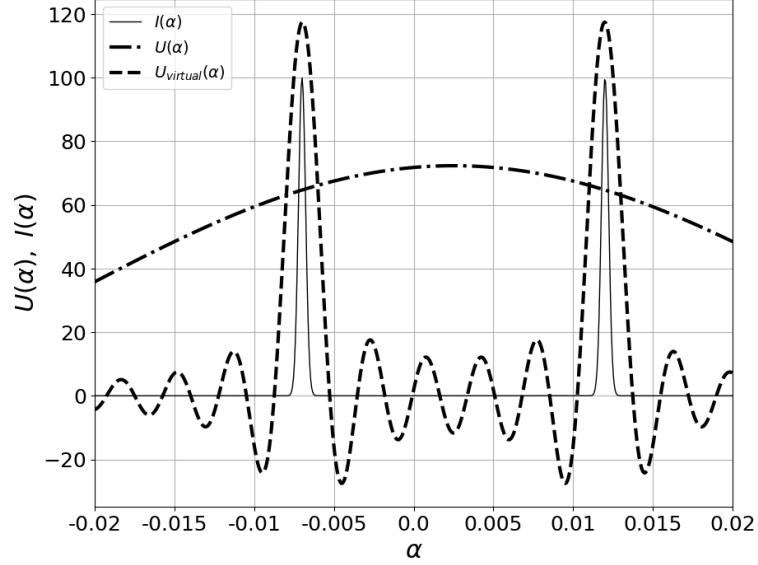


Figure 2: Значения функции распределения интенсивности источников и распределение принятого сигнала на ЦАР в угловом пространстве $\alpha \in \Omega$; сплошная тонкая линия - значения распределения интенсивности объектов $I(\alpha)$, моделируемых на основе дельта-функций Дирака [22]; штриховая жирная линия - принятый сигнал на физической ЦАР, рассчитанный на основе (4); сплошная жирная линия - принятый сигнал, , рассчитанный на основе (4), с применением моделируемых виртуальных устройств при их идеальной работе в соответствии с моделью (3)

$\hat{y} \in \mathbb{R}^d$ - вектор оценки целевых значений. Нейронная сеть отображает вектор входных данных x в выходной вектор оценки \hat{y} , посредством серии линейных преобразований с нелинейными функциями активации.

Полносвязная нейросетевая модель состоит из соединенных в скрытые слои линейные нейроны с нелинейной функцией активации. По мере прямого распространения информации от входов к выходу ко входному вектору x применяется последовательная серия линейных преобразований с матрицами весов слоя и нелинейных преобразований на основе выбранных функций активации. Входные данные часто представляются в виде матрицы векторов X :

$$X = (x^{<1>}, x^{<2>}, \dots, x^{<s>}) \in \mathbb{R}^{s \times f}, \quad x^{<i>} \in \mathbb{R}^f, \quad i \in 1, 2, \dots, s, \quad (10)$$

где s - количество примеров, f - размер вектора значений каждого примера. Значения на выходе нейронной сети представляются также в виде матрицы оценок $\hat{Y} \in \mathbb{R}^{s \times d}$, где d - размер одного вектора меток, соответствующий своему входному вектору.

Каждый слой нейронной сети с L слоями обладает матрицей весов W^l , $l \in 1, 2, \dots, L$, взвешивающей выход из предыдущего слоя или входа на каждом нейроне текущего слоя. Все матрицы весов $W = (W^1, W^2, \dots, W^L)$ нейронной сети обладают различными размерами, зависящими от размера входных векторов f , требуемого размера выходных векторов d и количества нейронов в каждом из слоев нейронной сети.

Веса нейронной сети W настраиваются в процессе обучения на тренировочной выборке примеров X_{train} и эталонных меток Y_{train} таким образом, чтобы минимизировать выбранный функционал ошибки $Loss(Y, \hat{Y})$ на пространстве параметров нейронной сети вследствие чего ставится следующая задача оптимизации:

$$Loss(Y, FCNN(x|W)) \rightarrow \min_W \quad (11)$$

Функционал ошибки между оценками \hat{Y} (9) и метками Y выбирается исходя из задачи и должен как отражать качество решения задачи нейронной сетью в виде числовой метрики. Существует множество функционалов качества для различных задач, решаемых алгоритмами искусственного интеллекта [24], [25]. Минимизация метрики качества (11) решения задачи (9) производится на основе алгоритма из семейства методов градиентного спуска с различными модификациями [26] в котором веса W итеративно обновляются на основе тестовых прогонов работы модели на

тренировочной выборке X_{train} (10) и минимизации полученных ошибок между ответами модели (9) и метками Y_{train} . Качество итогового обучения принято проверять на тестовой отложенной выборке схожих примеров X_{test} в сравнении с эталонными метками Y_{test} .

Численные эксперименты и сравнение методов экстраполяции

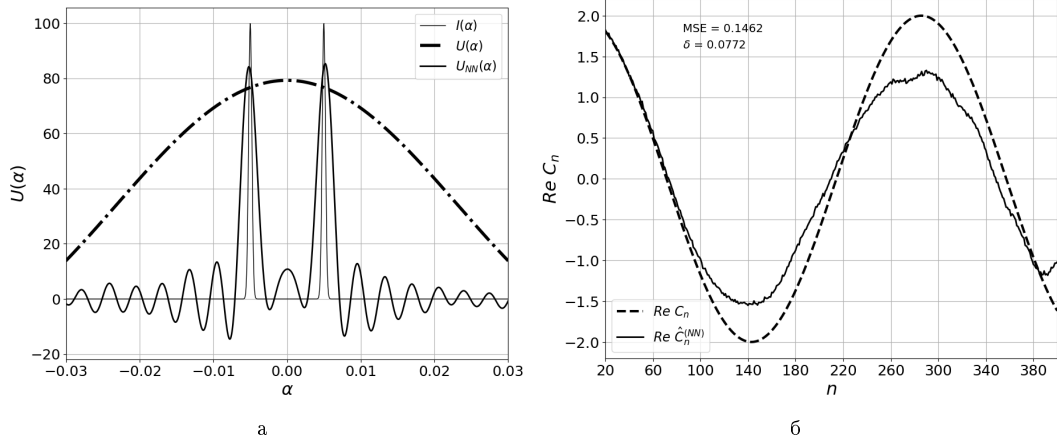


Figure 3:

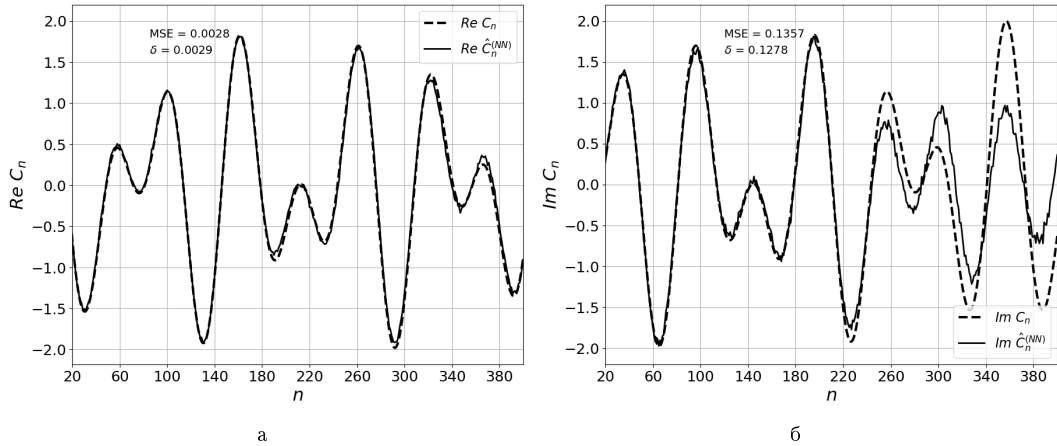


Figure 4:

В итоге сравним разрешающую способность, обеспечиваемую сопоставляемыми в данной работе методами экстраполяции. Для этого проведем серию экспериментов с сближающимися четными целями $t_1 = -t_2$, в которых будем измерять минимальное расстояние $\theta_{0,p}$ при которых объекты будут все еще различимы по критерию Рэлея (1) в зависимости от числа предсказываемых значений принятого сигнала виртуальных излучателей по различным алгоритмам экстраполяции.

По построенному графику (рис. 8) заметим, что добавление первых виртуальных устройств в содействие физической апертуре оказывает большее влияние на разрешающую способность по сравнению с последующими, что показывает снижающуюся эффективность добавления все более большого количества точек экстраполяции. Приблизительное положение излучателя в заданной апертуре ЦАР, после которого ширина ДН уменьшается значительно медленнее равно $n \approx 240$. В сравнении с экстраполяцией методом нейронной сети, алгоритм экстраполяции Берга показывает худшие результаты, что обусловлено достижением плато показателя $\theta_{0.5}$ при увеличении числа точек экстраполяции. По динамике показателя также можно объяснить результаты эксперимента на рисунке (рис.7). Алгоритм Берга при увеличении виртуальной апертуры не позволяет в данной реализации экстраполяции разрешить две цели, стоящие на

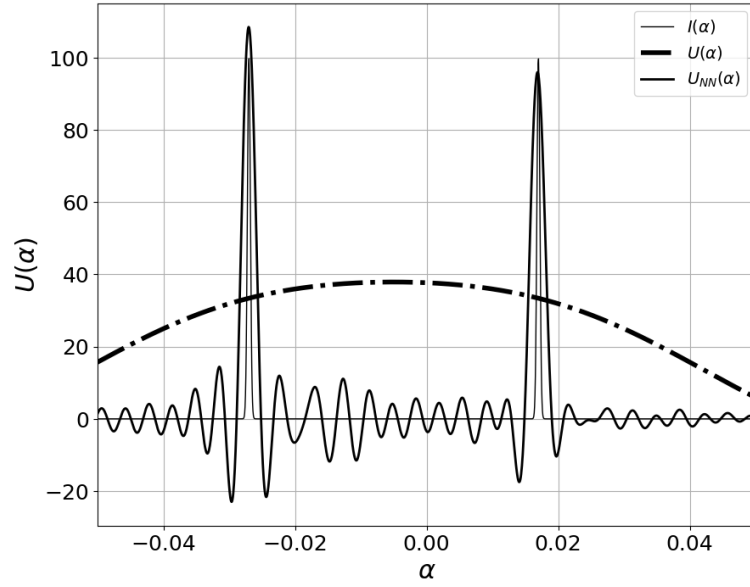


Figure 5:

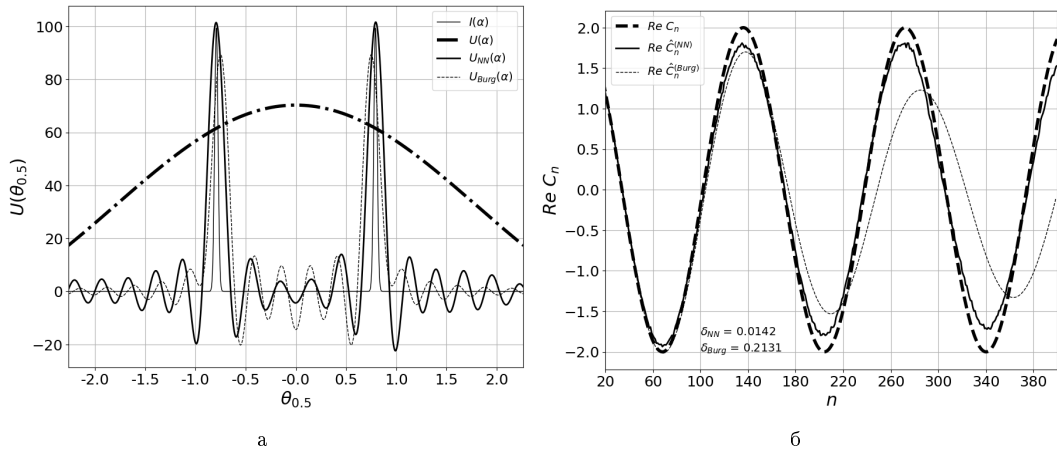


Figure 6:

расстоянии $|t_1 - t_2| \leq 0.014$. Также по графику (рис. 8) можно увидеть хорошее качество работы нейронной сети для данной задачи сверхразрешения, которая проходит близко к теоретической зависимости для виртуальных устройств, работающих по модели (3)-(5).

Заклучение

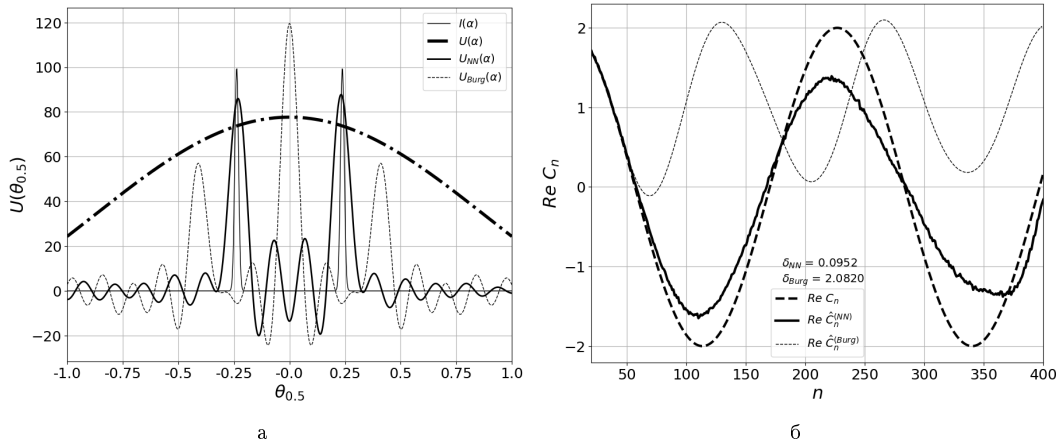


Figure 7:

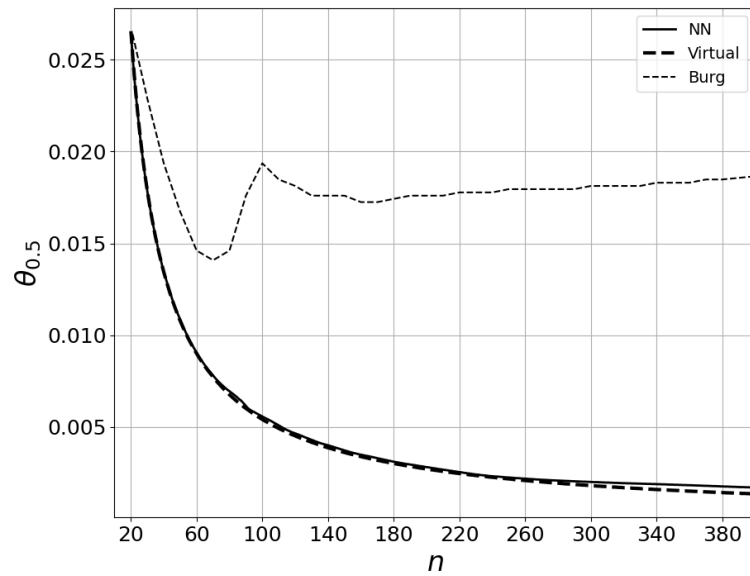


Figure 8: График динамики показателя ширины диаграммы направленности до половинной мощности в зависимости от числа предсказываемых виртуальных устройств

References

- [1] Bhaskar D., Rao B. D., Hari K.V.S. Performance analysis of ROOT – MUSIC // IEEE Trans. on Acoustics, Speech, and Signal Processing. 1989. Vol. ASSP – 37, pp. 1939-1949.
- [2] Cetin M., Karl W. Feature-enhanced synthetic aperture radar image formation based on non-quadratic regularization // IEEE Trans. Image Processing. 2001. Vol. 10, № 4, pp. 623–631.
- [3] Quinquis A., Radoi E., Totir F. Some radar imagery results using superresolution techniques // IEEE Trans. Antennas and Propagation. 2004. Vol. 52, № 5, pp. 1230-1244,
- [4] Changyin S, Zheng B. Super-resolution algorithm for instantaneous ISAR imaging // Electron. Letters. 2000. 36, № 3, pp. 253-255.
- [5] М.В. Ратынский. Адаптация и сверхразрешение абстрактнодействующих антенных решетках. Радио и связь, М 2003 г. с.199.
- [6] Herman M. A., Strohmer T. High-resolution radar via compressed sensing // IEEE Trans. Signal Processing. 2009. Vol. 57, № 6, pp. 2275–2284.

- [7] Lagovsky B.A. Image Restoration of the Objects with Superresolution on the Basis of Spline – Interpolation // Progress In Electromagnetics Research Symposium (PIERS 2012-Moscow), PIERS Proceedings 2012. pp. 989 – 992.
- [8] Лаговский Б.А., Самохин А.Б. Алгебраические методы восстановления изображения источников радиоизлучения с повышенным угловым разрешением // Электромагнитные волны и электронные системы. 2009, № 9, т.14, с. 7-18.
- [9] Лаговский Б.А. Методы повышения эффективного углового разрешения малоразмерных целей в задачах радионавигации и радиолокации. // Антенны. 2007, № 9(124), с. 50 -55.
- [10] Лаговский Б.А. Восстановление изображения групповой цели цифровыми антенными решетками // Антенны. 2011. № 2(165), с. 40 -46.
- [11] Лаговский Б.А., Рубинович Е.Я. Алгоритмы цифровой обработки данных измерений, обеспечивающие угловое сверхразрешение. Мехатроника, автоматизация, управление. 2021;22(7):349-356. <https://doi.org/10.17587/mau.22.349-356>
- [12] Тормозов, В. С. Настройка, обучение и тестирование нейронной сети долгой краткосрочной памяти для задачи распознавания образов / В. С. Тормозов, А. Л. Золкин, К. А. Василенко // Промышленные АСУ и контроллеры. – 2020. – № 3. – С. 52-57. – DOI 10.25791/asu.3.2020.1171. – EDN JJWSSG.
- [13] Минаев Е. Ю., Кутикова В. В., Никоноров А. В. Трекинг объектов в видеопотоке на основе сверточных нейронных сетей и фрактального анализа //Сборник трудов IV международной конференции и молодежной школы «Информационные технологии и нанотехнологии» (ИТНТ-2018)-Самара: Новая техника. – 2018. – С. 2792-2798.
- [14] Козар Б. А., Кугуракова В. В., Сахибгареева Г. Ф. Структуризация сущностей естественного текста с использованием нейронных сетей для генерации трехмерных сцен //Программные продукты и системы. – 2022. – Т. 35. – №. 3. – С. 329-339.
- [15] Семенюк В. В., Складчиков М. В. Разработка алгоритма распознавания эмоций человека с использованием сверточной нейронной сети на основе аудиоданных //Информатика. – 2022. – Т. 19. – №. 4.
- [16] Яшин В. И. РАСТУЩИЕ НЕЙРОННЫЕ СЕТИ В РОБОТОТЕХНИКЕ //Роботизация Вооружённых Сил Российской Федерации. – 2020. – С. 279-284.
- [17] Ахтёров А. В., Кирильченко А. А. Основы теоретической робототехники. Искусственные нейронные сети.(Обзор) //Препринты Института прикладной математики им. МВ Келдыша РАН. – 2008. – №. 0. – С. 2-20.
- [18] Сукина С. А. Нейронные сети //Новая наука: стратегии и векторы развития. – 2016. – №. 5 (82). – С. 248.
- [19] Ямашкин А. А., Ямашкин С. А. Использование нейронных сетей прямого распространения для ландшафтного картографирования на базе космических снимков //Геодезия и картография. – 2014. – №. 11. – С. 52-58.
- [20] Иванько А. Ф., Иванько М. А., Сизова Ю. А. Нейронные сети: общие технологические характеристики //Научное обозрение. Технические науки. – 2019. – №. 2. – С. 17-23.
- [21] Дычков И. Н. Сверточные нейронные сети //Тенденции развития науки и образования. – 2021. – №. 73-1. – С. 38-41.
- [22] Назаров А.Н. Обработка потоков в мониторинговом облачном кластере. Russian Technological Journal. 2019;7(6):56-67. <https://doi.org/10.32362/2500-316X-2019-7-6-56-67>
- [23] Дорожко Игорь Владимирович, Осипов Никита Алексеевич, Иванов Олег Андреевич ПРОГНОЗИРОВАНИЕ ТЕХНИЧЕСКОГО СОСТОЯНИЯ СЛОЖНЫХ ТЕХНИЧЕСКИХ СИСТЕМ С ПОМОЩЬЮ МЕТОДА БЕРГА И БАЙЕСОВСКИХ СЕТЕЙ // Труды МАИ. 2020. №113.

- [24] Виноградова Е. П., Головин Е. Н. Метрики качества алгоритмов машинного обучения в задачах классификации //Научная сессия ГУАП. – 2017. – С. 202-206.
- [25] Шишкин Ю. Е., Скатков А. В. Метрики качества для оценки и прогнозирования критических состояний //Качество и жизнь. – 2019. – №. 1. – С. 61-66.
- [26] Каширина И. Л., Демченко М. В. Исследование и сравнительный анализ методов оптимизации, используемых при обучении нейронных сетей //Вестник ВГУ. Серия: Системный анализ и информационные технологии. – 2018. – №. 4. – С. 123-132.