

ПЕРЕДАЧА И ОБРАБОТКА ИНФОРМАЦИИ

УДК 537.86+621.396.96

А.Ю. Паршин, Ю.Н. Паршин

ВЫДЕЛЕНИЕ ГРАНИЦ ФРАКТАЛЬНОГО ОБЪЕКТА МЕТОДОМ МАКСИМАЛЬНОГО ПРАВДОПОДОБИЯ ПО НЕЗАВИСИМОЙ И ЗАВИСИМОЙ ВЫБОРКАМ

Предложен алгоритм выделения границы объекта с использованием корреляционных размерностей объекта и фона методом максимального правдоподобия. Рассмотрены различные варианты структуры кадра обработки, проведен анализ вероятностей ошибок выделения границ при различном размере кадра. Исследовано влияние геометрической зависимости выборок на качество выделения границ.

Ключевые слова: корреляционная размерность, оптимальное оценивание, максимально правдоподобные оценки, зависимые выборки, выделение границы.

Введение. Расширение областей применения беспилотных аппаратов накладывает более высокие требования к организации канала связи с наземным пунктом приема. Для повышения надежности и скорости передачи информации необходимо уменьшать объем передаваемых данных, сохраняя при этом информативность. Внедрение компьютерных методов обработки позволяет выполнять обработку радиолокационных изображений с достаточно высокой эффективностью. Вместе с тем в настоящее время возможности различения объектов на фоне подстилающей поверхности и относительно друг друга у человеческого глаза заметно выше, чем у существующих автоматических компьютерных алгоритмов [1], что свидетельствует о большом потенциале повышения качества обработки изображений.

Современные методы обработки изображений широко используют различия в свойствах отдельных участков изображения с применением текстурного анализа. Наиболее распространенным способом является оценивание градиента текстурных признаков в окрестности исследуемого элемента изображения, который реализуют алгоритмы Собеля, Лапласа, Превитта, Кирша и Кенни [2]. При обработке изображения на область в окрестности выбранного элемента накладывается маска, которая выполняет преобразование области для нахождения локального гра-

диента яркости. Основной операцией указанных алгоритмов является вычисление производной. При этом алгоритм Кенни использует более совершенную и точную процедуру вычисления производной. Данный алгоритм включает также сравнение градиента с порогом с использованием гистерезиса и шумоподавление, что позволяет характеризовать данный алгоритм выделения границ объектов как наиболее эффективный по сравнению с другими градиентными методами [3].

Вследствие значительной пространственной изменчивости отражательных характеристик основных классов земной поверхности и объектов искусственного происхождения применением градиентных методов приводит к некорректным результатам. Поэтому возникла необходимость учета контекстной информации при распознавании объектов. Наличие статистической связи между близко расположенными элементами изображения приводит к появлению дополнительного признака для классификации. Источники контекстной информации могут находиться в самом изображении в виде определенных правил пространственной организации элементов.

Среди различных текстурных признаков наибольший интерес в последнее время вызывают фрактальные свойства изображения [4]. Вместе с тем широко используемые алгоритмы анализа фрактальной размерности [5] не обеспечи-

вают предельной точности при ограниченной выборке.

Целью статьи является описание методов повышения точности выделения границ на изображении с помощью текстурного анализа путем использования фрактальных свойств анализируемого фрагмента изображения, а также оптимальных по критерию максимального правдоподобия алгоритмов выделения границы. Особое внимание уделено обоснованию независимости анализируемой выборки и анализу эффективности при частично зависимой выборке.

Использование независимых расстояний между векторами. Для оценивания корреляционной размерности объекта [6], заданного временным рядом, требуется сформировать векторы в многомерном пространстве вложения, задающие состояние динамической системы. В работе [10] предложен способ реконструкции динамической системы путем использования значений наблюдаемой одной компоненты в качестве значений ненаблюдаемых компонент $\mathbf{x}_n = \{x_n, x_{n+1}, \dots, x_{n+d_E-1}\}$, где d_E – размерность пространства вложения, $n = 1, \dots, N$ – номер вектора при заданном способе упорядочивания наблюдаемых данных. Для повышения точности восстановления необходимо, чтобы соседние отсчеты были слабо коррелированными, а в пределах одного вектора должна сохраняться зависимость между его координатами. В связи с этим число координат вектора должно быть небольшим, что позволит отразить нелинейную статистическую зависимость. Для полного восстановления динамической системы размерность вектора выбирается с учетом корреляционной размерности d восстанавливаемого динамического процесса по соотношению $d_E \geq 2 \text{int}(d) + 1$ [10]. Однако для частной задачи оценивания корреляционной размерности достаточным является соотношение $d_E \geq \text{int}(d) + 1$ [16].

Для оценивания корреляционной размерности необходимо определить расстояния между всеми парами N векторов в выборке. В дальнейшем расстояния между векторами рассчитываются в соответствии с евклидовой метрикой. Если для проведения расчета корреляционной размерности используются все расстояния, то такой набор называется V-статистикой [11], а число расстояний равно $M_V = N^2$. Для данной статистики характерна повторяемость расстояний, что приводит к неинформативности более половины элементов статистики. Кроме того, часть неповторяющихся расстояний является взаимозависимой, что не согласуется с моделью,

принятой при синтезе максимально правдоподобного алгоритма оценивания корреляционной размерности. Вследствие этого V-статистика практически не используется для получения оптимальных оценок.

Для получения более информативной U-статистики [11] выделяют неповторяющиеся расстояния между векторами $\|\mathbf{x}_i - \mathbf{x}_j\|$, $i = 2, \dots, N$, $j = 1, \dots, i-1$. Количество расстояний в статистике равно $M_U = N(N-1)/2$. В этом случае также имеется геометрическая взаимозависимость части расстояний, что не позволяет строго применять модель независимых расстояний для синтеза оптимального алгоритма.

Для уменьшения взаимозависимости расстояний используют T-статистику [11] с применением окна Тейлера [15]: статистика содержит расстояния

$$\|\mathbf{x}_i - \mathbf{x}_j\|, i = T+1, \dots, N, j = 1, \dots, i-T,$$

где T – размер окна Тейлера. При этом разность индексов устанавливается равной T , что влияет на независимость расстояний между векторами. Число расстояний в статистике равно $M_T = (N-T)(N-T+1)/2$. В частном случае $T=1$ получаем U-статистику. В некоторых работах [13] предлагается выбирать размер окна Тейлера равным $T = d_E - 1$.

Более точно независимость расстояний может быть установлена путем рассмотрения геометрических свойств множества векторов. Рассмотрим d_E -мерное пространство векторов общего положения, в котором задана метрика и определены расстояния между всеми парами векторов. Координаты векторов могут быть независимыми, например статистически независимыми случайными числами.

Для выявления свойства взаимозависимости расстояний между N векторами общего положения в евклидовом пространстве рассмотрим сначала более простой случай векторов на плоскости: $d_E = 2$, анализируя расстояния последовательно для каждого вектора. Необходимо определить возможность задать положение каждого из векторов в пространстве расстояний с точностью до сдвига и поворота. Расстояние r_{12} между 1-м и 2-м векторами может быть произвольным и задается только положением этих векторов в пространстве. Расстояние r_{13} между 1-м и 3-м векторами также может быть произвольным, а расстояние r_{23} между 2-м и 3-м векторами зависит от значений выше рассмотренных расстояний в соответствии с правилом тре-

угольника: $0 \leq r_{23} < r_{12} + r_{13}$. Вместе с тем минимальное значение расстояния зависит от заданной метрики. Например, в пространстве с евклидовой метрикой нижняя граница диапазона расстояния r_{23} равна $|r_{12} - r_{13}| \leq r_{23} \leq r_{12} + r_{13}$. Таким образом, появляется расстояние, которое зависит от других расстояний, а положение 3-го вектора определяется расстояниями r_{13} , r_{23} с точностью до отражения. Однозначность положения 3-го вектора гарантируется, если добавляется 4-й вектор и фиксируется еще одно расстояние r_{34} .

Обобщая проведенные рассуждения на произвольное число векторов, можно заключить:

1) полностью независимым является только $N - 1$ расстояние;

2) расстояние между 2-м и 3-м векторами является условно независимым от r_{12} , r_{13} и ограничивается правилом треугольника $|r_{12} - r_{13}| < r_{23} < r_{12} + r_{13}$;

3) расстояние между 3-м и 4-м векторами принимает одно из двух predetermined значений; величина этих расстояний зависит от выбранной метрики, остальные расстояния до 3-го вектора полностью определяются предыдущими расстояниями;

4) добавление каждого последующего вектора общего положения однозначно фиксирует все предыдущие векторы, что делает все условно независимые расстояния фиксированными, и дает одно независимое расстояние и несколько условно независимых расстояний.

Учитывая сложный характер взаимозависимости расстояний между векторами, затруднительно установить общее выражение для распределения вероятности этих расстояний. Более простой задачей является определение максимального числа M_1 независимых расстояний между N векторами в d_E -мерном пространстве.

Пространство конфигураций, задаваемых координатами всех векторов, имеет размерность $d_{conf} = Nd_E - d_E(d_E + 1)/2$, где вычитаемый член соответствует количеству движений d_E -мерного пространства, преобразующих конфигурации векторов в тождественные, то есть d_E сдвигов и $d_E(d_E - 1)/2$ вращений [12]. При этом количество фиксирующих расстояний между векторами не может быть меньше размерности пространства конфигураций, т.е. если количество расстояний меньше d_{conf} , то они конфигурацию однозначно задать не могут. И наоборот, если количество расстояний больше d_{conf} , то

они образуют избыточную систему координат и являются зависимыми.

Вместе с тем система, содержащая d_E выбранных точек и расстояния от каждой точки до всех остальных, имеет как раз такой размер вращений [12]. Однако пространство конфигураций отображается в соответствующее евклидово пространство неоднозначно: каждому набору расстояний соответствует дискретный набор конфигураций, расположенных по разные стороны от d_E выбранных точек. Для устранения неоднозначности необходимо добавить расстояния до еще одной точки от всех, т.е. $N - d_E - 1$ расстояний. Таким образом, получается верхняя и нижняя границы для необходимого количества расстояний:

- нижняя граница $M_{UL} = d_{conf}$,

- верхняя граница $M_{UH} = d_{conf} + N - d_E - 1$.

Например, для двумерного изображения $d_E = 2$, а число расстояний равно $M_{UL} = 2N - 3$, $M_{UH} = 3N - 6$. Можно предположить, что в приложениях более существенной является размерность системы, т.е. нижняя граница необходимого числа расстояний.

На рисунке 1 представлена зависимость отношения $R = M_{UL} / M_U$ от числа векторов N для нескольких значений d_E . При увеличении N относительное количество расстояний, необходимых для фиксирования конфигурации, уменьшается. Поэтому целесообразно провести сравнительный анализ для независимой и более полной зависимой выборки.

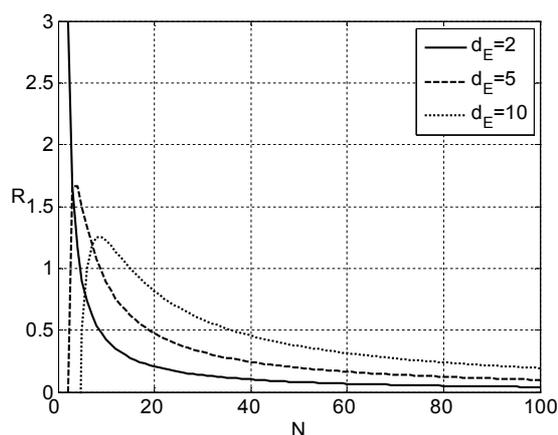


Рисунок 1 – Зависимости отношения $R = M_{UL} / M_U$ от числа векторов N

Наиболее точные оценки корреляционной размерности получаются при использовании алгоритма максимального правдоподобия [7,9]. При условии нормировки расстояний $r_m = l_m / l_{max}$, $m = 1, \dots, M$, закон распределения

вероятностей для расстояний между векторами, для которых корреляционная размерность равна d , задается степенной зависимостью $F(r) = r^d$, а функция плотности распределения вероятности имеет вид $w(r) = \frac{dF(r)}{dr} = d \times r^{d-1}$, $0 < r < 1$ [9].

При условии, что измеренные значения расстояний между векторами в псевдофазовом пространстве статистически независимы, в работе [9] получена многомерная плотность распределения вероятностей, которая также является функцией корреляционной размерности d , а значит, и функцией правдоподобия. Максимально правдоподобная оценка корреляционной размерности, полученная на основании данной статистической модели, имеет вид [9]:

$$\hat{d} = -\frac{M}{\sum_{m=1}^M \ln r_m}. \quad (1)$$

Для решения задачи выделения границы между фоном и объектом также использовался метод проверки статистических гипотез на основе достаточной статистики – отношения правдоподобия [8]:

$$\Lambda = \frac{w_1(\mathbf{x})}{w_2(\mathbf{x})} = \frac{\prod_{m=1}^M d_1 r_m^{d_1-1}}{\prod_{m=1}^M d_2 r_m^{d_2-1}}, \quad (2)$$

где d_1, d_2 – корреляционные размерности объекта и фона.

Алгоритм выделения границы по фрактальным признакам фона и объекта. Для выделения границы объекта изображение разбивается на кадры квадратной формы, состоящие из полей анализа. При формировании границы кадра из 2-х полей граница может располагаться в одном из четырех априорно заданных направлений – горизонтальном, вертикальном и по двум направлениям диагонали. Оценка границы представляется как одно из полей. Для определенности граница отображается верхним, правым и нижним диагональным полем.

При формировании границы кадра из 9 полей граница также может располагаться в одном из четырех априорно заданных направлений: горизонтально, вертикально посередине кадра, а также по двум направлениям диагонали. Оценка границы представляется как набор из 3 полей, расположенных вертикально или горизонтально в середине кадра, а также по диагонали кадра.

Граница разделяет кадр на две части, в каждой из которых формируется N векторов [14]. Статистика для задачи выделения границы равна

отношению правдоподобия для границы различного вида в предположении, что в частях кадра по обе стороны границы располагаются объекты с различным значением корреляционной размерности. При условии равновероятных расположений объекта и фона относительно границы отношение правдоподобия принимает вид:

$$\Lambda_k = \frac{\frac{1}{2} \prod_{m=1}^{M_1} d_1 r_{1km}^{d_1-1} \times \prod_{m=1}^{M_1} d_2 r_{2km}^{d_2-1} + \frac{1}{2} \prod_{m=1}^{M_2} d_1 r_m^{d_1-1} + \frac{1}{2} \prod_{m=1}^{M_2} d_2 r_m^{d_2-1}}{\frac{1}{2} \prod_{m=1}^{M_1} d_2 r_{1km}^{d_2-1} \times \prod_{m=1}^{M_1} d_1 r_{2km}^{d_1-1} + \frac{1}{2} \prod_{m=1}^{M_2} d_1 r_m^{d_1-1} + \frac{1}{2} \prod_{m=1}^{M_2} d_2 r_m^{d_2-1}}, \quad (3)$$

где $k = 1, \dots, 4$ – номер разновидности границы, r_{1km}, r_{2km} – расстояния в полях анализа 1 и 2 при k -й разновидности границы, r_m – расстояния между всеми векторами в кадре.

Определение наличия границы производится путем сравнения с порогом максимального из отношений правдоподобия (3), вычисленного для каждой из проверяемых гипотез, состоящей в наличии объектов с различными значениями корреляционной размерности в соседних полях одного и того же кадра при наличии границы определенного вида. Принимается решение о наличии границы, соответствующей кадру с номером $\hat{k} = \arg \max_k \Lambda_k$, если значение отношения правдоподобия $\Lambda_{\hat{k}}$ превышает некоторое пороговое значение.

Анализ эффективности алгоритма выделения границы проведен методом вычислительно-эксперимента. Независимые расстояния между векторами со степенным законом распределения вероятностей сформированы на основе датчика независимых равномерно распределенных в диапазоне (0;1) случайных чисел $n: r = n^{1/d}$. Зависимые расстояния генерируются в соответствии со следующим алгоритмом.

1. Генерируются $N-1$ независимых случайных чисел со степенным распределением вероятностей в диапазоне значений (0;1); данные числа задают расстояния от первого вектора до всех $N-1$ остальных векторов.

2. Генерируется $N-2$ независимых случайных чисел со степенным распределением вероятностей в диапазоне значений $(r_{\min}; r_{\max})$; данные числа задают условно независимые расстояния от второго вектора до всех $N-2$ остальных, кроме 1-го векторов. Минимальное и

максимальные значения определяются правилом треугольника:

$$r_{\min} = |r_{12} - r_{1i}|, \quad r_{\max} = r_{12} + r_{1i}, \quad i = 3, \dots, N.$$

3. Формируются координаты 1-го и 2-го векторов:

$$x_1 = 0, \quad y_1 = 0, \quad x_2 = r_{12}, \quad y_2 = 0.$$

Координаты остальных $i = 3, \dots, N$ векторов определяются из геометрии их положения с использованием теоремы косинусов и расстояний от i -го вектора до 1-го и 2-го векторов:

$$x_i = \frac{r_{12}^2 + r_{1i}^2 - r_{2i}^2}{2r_{12}}, \quad y_i = \pm \sqrt{r_{1i}^2 - x_i^2}.$$

В дальнейшем знак координаты y_i выбирается случайным и равновероятным.

4. По найденным координатам всех векторов вычисляются оставшиеся $\frac{N^2 - 5N}{2} - 3$ зависимых расстояний: $r_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}$, $i = 3, \dots, N$, $j = i + 1, \dots, N$.

Анализ эффективности предложенных алгоритмов выделения границ изображения проведен на примере объекта с границей в виде восьмиугольника, что позволяет более полно отразить влияние нелинейности границы объекта. Выделение границы проводилось методом статистического моделирования отношения правдоподобия (3).

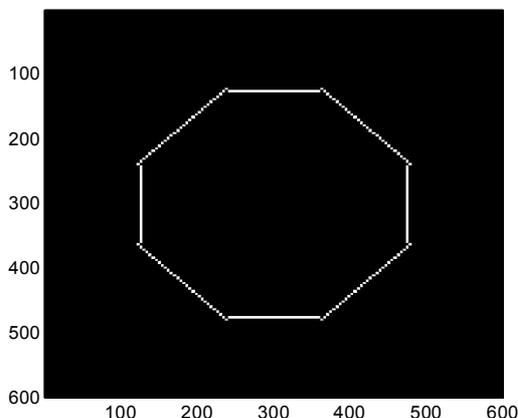


Рисунок 2 – Идеальная граница объекта для исследуемого метода выделения границы

На рисунке 2 представлена оценка границы, полученная при идеальных условиях измерений границы объекта: ошибки, вызванные конечным объемом выборки при статистическом оценивании исключены. Неровность границы на рисунке 2 объясняется особенностями формирования кадра, принятыми видами границы и конечным размером исследуемого кадра. На рисунке 3 приведены результаты выделения границы объекта с использованием алгоритма максимального

правдоподобия (3). Необходимо отметить ошибки выделения границы в виде ложных отметок и пропуска границы.

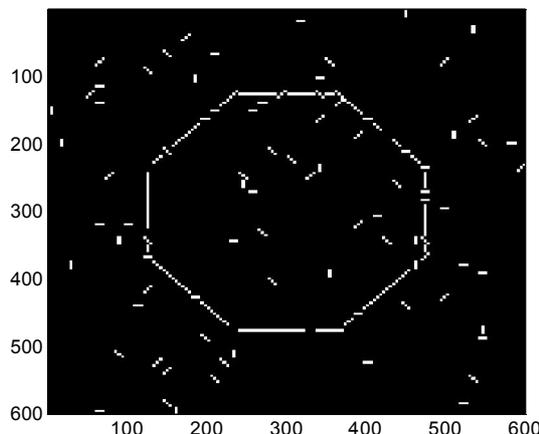


Рисунок 3 – Результат обработки алгоритмом максимального правдоподобия, $d_1 = 1,3$, $d_2 = 1,8$, граница из трех полей

Визуальные методы оценки качества обработки изображения не позволяют объективно произвести сравнение анализируемых алгоритмов и определить их эффективность. Предложено оценивать эффективность алгоритмов выделения границ по вероятностным характеристикам, в качестве которых приняты вероятность ложных полей границы на всем изображении

$$F = \frac{N_{ЛГ}}{N_{НГР}}$$

$$P_{УР} = \frac{N_{ПГ}}{N_{ГР}}, \quad \text{где } N_{ЛГ} - \text{число полей на изображении, не совпадающих с отметками эталонной границы, } N_{НГР} - \text{число полей эталонного изображения, не относящихся к эталонной границе, } N_{ПГ} - \text{число пропущенных полей на эталонной границе, } N_{ГР} - \text{общее число полей на эталонной границе.}$$

Выбирая априорную величину вероятности появления ложных полей, можно устанавливать величину порога, с которым производится сравнение статистики (3). Результаты моделирования зависимости ошибок 1-го и 2-го рода от числа независимых расстояний между векторами для кадров различного вида приведены на рисунках 4,5.

Исследование статистических характеристик проводится для числа векторов $N = 18; 72; 162$. Используются две структуры кадра – 2×1 и 3×3 поля. Количество независимых расстояний равно $M = N - 1$. Число зависимых расстояний рассчитывается как $M = \frac{N(N-1)}{2}$.

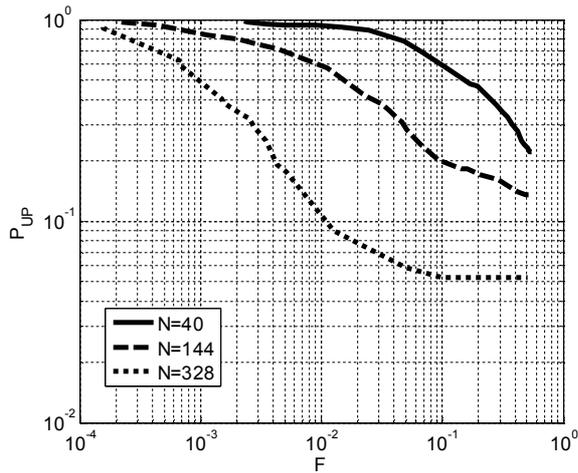


Рисунок 4 – Вероятности ошибок при использовании независимых выборок и различном числе расстояний в кадре 2x1, $d_1 = 1,3, d_2 = 1,8$

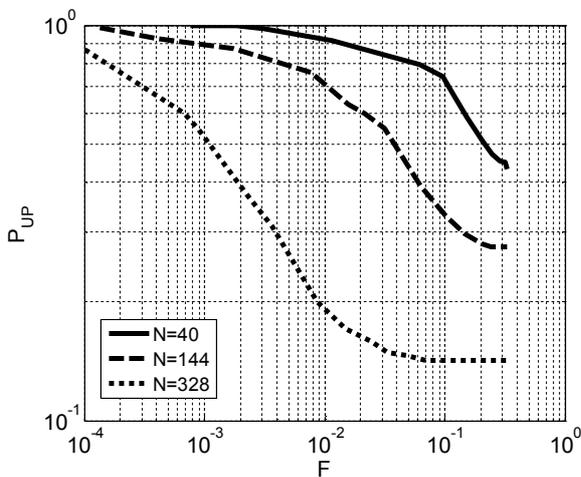


Рисунок 5 – Вероятности ошибок при использовании независимых выборок и различном числе расстояний в кадре 3x3, $d_1 = 1,3, d_2 = 1,8$

При малом значении порога обнаружение производится в каждом кадре изображения, в том числе и в части изображения, не занимаемого границей. При формировании границы в виде 3-х полей из 9 в кадре 3x3 число полей, занимаемых ошибочно обнаруженной границей, равно 3. Поэтому вероятность появления ложных полей границы имеет максимальное значение $F = 1/3$. При формировании границы в виде одного из 2-х полей в кадре со структурой 2x1 поле предельное значение вероятности появления ложных отметок равно $F = 1/2$. В этом случае обнаруживаемая граница при такой структуре занимает половину кадра, а максимальное число отмеченных полей составляет половину общего объема данных изображения.

Вероятность пропуска границы также имеет предел, который определяется числом расстояний M и разницей между d_1 и d_2 . С увеличением M уменьшается ошибка вычисления кор-

реляционной размерности и, следовательно, вероятность пропуска границы. В общем случае вероятность пропуска границы определяется вероятностью обнаружения границы и вероятностью классификации границы. При уменьшении порога обнаружения граница обнаруживается достоверно, и вероятность пропуска границы перестает зависеть от порога и принимает минимальное значение. Дальнейшее уменьшение вероятности пропуска границы можно получить при увеличении M , а также при увеличении разности между корреляционными размерностями объекта d_1 и фона d_2 .

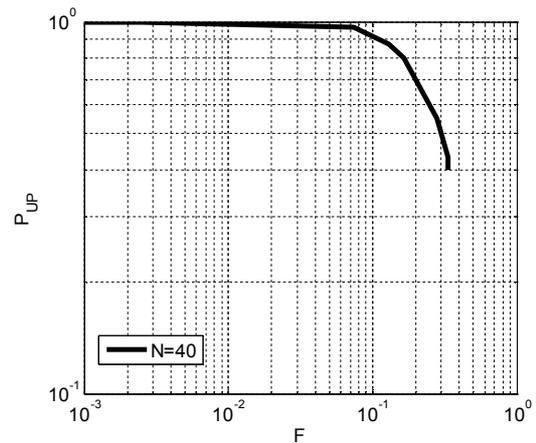


Рисунок 6 – Вероятности ошибок при использовании зависимых выборок в кадре 3x3, $d_1 = 1,3, d_2 = 1,8$

На рисунке 6 приведены вероятности ошибок, полученные для зависимых расстояний, сформированных предложенным в данной работе алгоритмом. При этом обработка осуществляется с помощью алгоритма максимального правдоподобия (3), полученного в предположении независимых расстояний. Для уменьшения геометрической зависимости расстояний предлагается аналогично работам [7,9] применить усечение, то есть учитывать расстояния, не превышающие некоторого заданного значения. Усечение зависимой выборки расстояний производится по сформированной гистограмме и осуществляется приблизительно на уровне 0,4 от максимального значения расстояний. При этом оценка корреляционной размерности получается смещенной и асимптотически эффективной. Это позволяет использовать алгоритм максимального правдоподобия, синтезированный для независимой выборки. Однако число усеченных расстояний невелико и приближенно соответствует числу введенных в данной статье независимых расстояний.

Сравнение графиков на рисунках 4, 5, полученных для независимых расстояний, с графиками

ком на рисунке 6 показывает заметное увеличение вероятностей ошибок при использовании зависимых расстояний. Полученные результаты согласуются с результатами работы [7], в которой также установлено сильное влияние зависимости расстояний на точность оценивания корреляционной размерности.

Заключение. В данной статье на основе предложенной методики проведен анализ эффективности выделения границы для 4-х вариантов модели границы, разделяющей квадратный кадр. Установлено, что при использовании наблюдаемых данных в виде независимых расстояний и оптимального по критерию максимального правдоподобия алгоритма вероятности ошибок уменьшаются при увеличении числа сформированных векторов.

При использовании совокупности независимых и зависимых расстояний вероятности ошибок получаются заметно выше вследствие неоптимальности применяемого алгоритма для этого случая. Увеличение числа векторов не позволяет значительно уменьшить вероятности ошибок, так как в этом случае увеличивается доля зависимых расстояний в наблюдаемых данных. Возможность уменьшения вероятностей ошибок связана с разработкой оптимального алгоритма для случая зависимых расстояний.

Исследование выполнено при поддержке гранта Российского научного фонда (проект 14-19-01263) в Рязанском государственном радиотехническом университете.

Библиографический список

1. Теоретические основы радиолокации и радионавигации: учеб. пособие для вузов / Сосулин Ю.Г. М.: Радио и связь, 1992.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005. 1072 с.
3. Прэнтт У. Цифровая обработка изображений. М.: Мир, 1982. 781 с.
4. Потапов А.А., Герман В.А. Фрактальный непараметрический обнаружитель радиосигналов // Ра-

диотехника. 2006. № 5. С. 30-36.

5. Сосулин Ю.Г., Русскин А.Б. Фрактальное обнаружение протяженных малоcontrastных объектов на изображениях // Радиотехника. 2009. № 12. С. 48-57.

6. Паркер Т.С., Чжуа Л. О. Введение в теорию хаотических систем для инженеров // ТИИЭР. 1987. Т. 75. № 8. С. 6-40.

7. Паршин Ю.Н., Паршин А.Ю. Максимально правдоподобное оценивание корреляционной размерности с учетом влияния смещения оценки и усечения диапазона масштабов // Вестник Рязанского государственного радиотехнического университета. 2013. № 4-1 (46). С. 18-25.

8. Паршин А.Ю., Паршин Ю.Н. Использование максимально правдоподобных оценок фрактальной размерности в негауссовских статистиках для обнаружения радиосигналов // Цифровая обработка сигналов. 2013. №1. С. 42-46.

9. Luciana De Luca, Dario Luzio, Massimo Vitale. A ML Estimator of the Correlation Dimension for Left-hand Truncated Data Samples // Pure and applied geophysics. V.159. № 11-12. 2002. P. 2789-2803.

10. Takens F. Detecting Strange Attractors in Turbulence // Lecture Notes in Mathematics. V. 898. 1981. P. 366-381.

11. Cees Diks. Nonlinear Time Series Analysis: Methods And Applications (Nonlinear Time Series and Chaos: v.4) World Scientific. 1999. 209 p.

12. Фрид Э. Элементарное введение в абстрактную алгебру / пер. с венгерского Ю.А. Данилова. М.: Мир, 1979. 260 с.

13. Provenzale A., Smith J.A., Vio R., Murate G. Distinguishing between low-dimensional dynamics and randomness in measured time series. Physica D 58. P. 31-49. 1992.

14. Parshin A.Yu., Parshin Yu.N. Synthesis and analysis of efficiency of algorithm for objects edges detection by maximum likelihood method // 2nd Mediterranean Conference on Embedded Computing (MECO). Budva. Montenegro. June 15th-21st 2013. 282 p.

15. Theiler J. Spurious dimensions from correlation algorithms applied to limited time-series data // Phys. Rev. A 34. 1986. P. 2427-2432.

16. Антипов О.И., Неганов В.А., Потапов А.А. Детерминированный хаос и фракталы в дискретно-нелинейных системах. М.: Радиотехника, 2009. 235 с.

УДК 621.395:519.2

О.Н. Титов, А.А. Афанасьев

МОДЕЛЬ СИСТЕМЫ ПРЕДОБРАБОТКИ РЕЧЕВОГО СИГНАЛА НА ОСНОВЕ ПОЛИСПЕКТРАЛЬНОГО АНАЛИЗА

Рассмотрены современные тенденции в вопросах развития систем предобработки речевого сигнала в условиях зашумления. Показаны основные недостатки существующих систем, обоснована и предложена математиче-

ская модель на основе полиспектрального анализа для решения частных задач в контексте эффективного шумоподавления в речи.

Ключевые слова: полиспектральный анализ, спектральная плотность мощности (СПМ), тройная автокорреляционная функция (ТАКФ), речевого сигнал (РС), биспектр речевого сигнала, сложная помеховая обстановка.

Введение. Ведущей тенденцией развития отечественных систем обработки речи, обусловленной ее важностью в процессе обмена информацией и возрастающими современными требованиями абонентов, является задача повышения качества предоставляемых телекоммуникационных услуг.

При непосредственной обработке зашумленной речи очевидно наличие тепловых и других шумов, а также возможно зашумление его другими потенциально нежелательными акустическими помехами.

Цель работы – описание математической модели предобработки зашумленного РС на основе полиспектрального анализа.

Дополнительно отметим, что статья носит рекомендательный характер, так как рассматриваются вопросы теоритического моделирования с целью создания новой системы предобработки зашумленного речевого сигнала.

Теоретические исследования. Среди подходов, реализующих задачу шумоподавления в обрабатываемом РС на практике и в теории, хорошо известны такие, как винеровская и инверсная фильтрации и методы, основанные на параметрическом анализе сигналов, на основе процедуры регуляризации с целью решения обратной задачи. Однако стоит отметить, что эффективность данных «классических» методов зачастую зависит от ряда допущений и ограничений как к РС, так и к помехе [1, 2].

При проведении анализа существующих на сегодняшний момент методов устранения шумов в РС можно провести определенную классификацию [1-3]:

- 1) адаптивные компенсаторы помех;
- 2) методы, основанные на использовании статистических моделей РС во временной области;
- 3) методы, основанные на обработке РС с использованием аппарата скрытых марковских моделей;
- 4) методы, основанные на использовании, отдельных характерных свойств РС;
- 5) методы, основанные на оценке спектральных характеристик шума;
- 6) метод оценивания минимальной среднеквадратической ошибки;
- 7) подавление аддитивного квазистационарного шума методом вычитания амплитудных спектров;

8) методы адаптивного трешолдинга на основе вейвлет-преобразования;

9) методы на основе нелинейной фильтрации.

Особо стоит отметить метод под № 7, получивший наибольшее теоретическое распространение и «завоевавший» нишу практического применения в последние 8 - 10 лет.

Однако на современном этапе развития технологий и способов шумоподавления все существующие методы слабо приспособлены к шумовому воздействию высокой энергетики и ориентированы, прежде всего, на «нормальный» закон распределения.

Поэтому в условиях сложной помеховой обстановки, а именно при воздействии помех с законом распределения, отличающимся от нормального, к примеру, в условиях воздействия шума смешанного вида, который может включать аддитивную гауссову компоненту и шум, отличающийся от гауссова, лучшие по отношению к линейной фильтрации результаты обеспечивают методы и алгоритмы нелинейной фильтрации [2, 4]. Данная группа методов, несмотря на характерные для них специфические динамические искажения, обеспечивает успешное решение задачи подавления смешанного (негауссова) шума в условиях полной или частичной априорной неопределенности в отношении свойств шумов за счет использования какого-либо фильтра из большого банка разнообразных сглаживающих фильтров, разработанных к настоящему времени.

Однако эффективность методов нелинейной фильтрации в задаче шумоподавления падает при обработке «импульсных» сигналов, т.е. сигналов, представляющих собой набор импульсов, длительность которых сравнима с длительностью помеховых импульсных выбросов, а также в случаях, когда для нескольких наблюдаемых реализаций обрабатываемых сигналов имеют место случайные сдвиги информационной составляющей, которые могут быть, например, обусловлены влиянием нелинейного искажения сигналов.

Повышение качества синтезированной речи при ее низкоскоростном кодировании в условиях сложной помеховой обстановки может быть достигнуто с использованием следующего подхода – это представление РС аддитивной смесью сиг-

налов речи, имеющих квазистационарную природу, а также помех, которые, в свою очередь, имеют собственный закон распределения, носящие как гауссов $n_G^{(m)}(t)$, так и негауссов характер $n_{aG}^{(m)}(t)$:

$$U^{(m)}(t) = S^{(m)}(t) + n^{(m)}(t), \quad (1)$$

$$n^{(m)}(t) = \sum n_G^{(m)}(t) + \sum n_{aG}^{(m)}(t), \quad (2)$$

где $U^{(m)}(t)$ – обрабатываемый РС, $S^{(m)}(t)$ – сигналы речи, $n^{(m)}(t)$ – шумовое воздействие.

Закон распределения РС, представленного выражением (1), в условиях, когда энергетически шум значительно превышает сигналы речи, стремится, в приближенной степени, к нормальному закону, однако имеет отличия от него. Данными отличиями часто пренебрегают, что вносит в некоторых случаях весомые искажения, возникающие при обработке зашумленной речи, особенно на конечном этапе при ее восприятии [2, 3].

На основе результатов, представленных такими учеными, как Вентцель Е.С., Малахов А.Н., Пугачев В.С., Баскаков С.И., в работах которых описаны взаимосвязи канонических разложений случайных процессов с их характеристическими функциями и функциями плотности вероятности:

$$W_\zeta(u) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Theta_\xi(v) \exp[-jv\zeta] dv, \quad (3)$$

$$\Theta_\xi(v) = \int_{-\infty}^{\infty} W_\zeta(u) \exp[jvu] du, \quad (4)$$

где $W_\zeta(u)$ – функция плотности вероятности (ФПВ) случайной величины; $\Theta_\xi(v)$ – характеристическая функция (ХФ), данную пару еще называют «парой преобразования Фурье», что позволяет говорить о ХФ как о тождественном представлении ФПВ.

Кроме того, (4) в отличие от (3) может и не обладать сходимостью, что позволяет говорить о некотором суммировании в смысле обобщенных функций [4,5].

Так как ХФ тождественно представляет ФПВ, то естественно, что все параметры функции $W_\zeta(u)$ можно «достать» из $\Theta_\xi(v)$, например моменты [9,10]:

$$a_n = j^{-n} \left(\frac{\partial^n \Theta_\xi(v)}{\partial v^n} \right)_{v=0} \equiv j^{-n} \Theta_\xi^n(0). \quad (5)$$

Таким образом, ХФ можно записать с помощью степенного ряда, членами которого яв-

ляются моменты случайной величины:

$$\Theta_\xi(v) = 1 + \sum_{k=1}^{\infty} \frac{a_k}{k!} (jv)^k. \quad (6)$$

Если же записать ХФ как некую экспоненциальную функцию от определенной функции $B(v)$:

$$\Theta_\xi(v) = \exp(B(v)), \quad (7)$$

то в случае нахождения обратной функции можно записать:

$$B(v) = \ln \Theta_\xi(v) = \sum_{k=1}^{\infty} \frac{\chi_k}{k!} (jv)^k, \quad (8)$$

вследствие чего саму ХФ можно представить через члены ряда функции $B(v)$:

$$\Theta_\xi(v) = \exp \left[\sum_{k=1}^{\infty} \frac{\chi_k}{k!} (jv)^k \right], \quad (9)$$

где χ_k – коэффициенты ряда функции $B(v)$, носящие название кумулянтов или семиинвариантов. Набор кумулянтов однозначно описывает статистические характеристики случайной величины и в совокупности служит тождественным представлением ее закона распределения. Анализируя (6) и (9), можно легко сделать вывод, что кумулянты и моменты случайной величины связаны между собой:

$$\begin{aligned} \chi_1 &= a_1 = m \\ \chi_2 &= a_2 - a_1^2 = D \\ \chi_3 &= a_3 - 3a_1 a_2 + a_1^3 = A \\ \chi_4 &= a_4 - 3a_2^2 - 4a_1 a_3 + 12a_1^2 a_2 - 6a_1^4 = E \end{aligned} \quad (10)$$

Аналогично показана и обратная связь:

$$\begin{aligned} a_1 &= \chi_1 \\ a_2 &= \chi_2 + \chi_1^2 \\ a_3 &= \chi_3 + 3\chi_1 \chi_2 + \chi_1^3 \\ a_4 &= \chi_4 + 3\chi_2^2 + 4\chi_1 \chi_3 + 6\chi_1^2 \chi_2 + \chi_1^4 \end{aligned} \quad (11)$$

Необходимо также сказать, что коэффициенты A и E в выражении (10) характеризуют асимметрию и эксцесс ФПВ, кроме того, для гауссова процесса данные коэффициенты равны нулю.

Непосредственно для практического применения, раскрываемого в исследовании, наиболее важна функция плотности вероятности, получаемая путем обратного преобразования Фурье от (9):

$$W_{aG_\xi}(u) = W_{G_\xi}(u) + \sum_{k=3}^{\infty} (-1)^k \frac{\chi_k}{k!} W_{G_\xi}^{(k)}(u). \quad (12)$$

В работах [4-6] показаны оптимальность представления и полнота описания стационарно-

го сигнала конечной мощности совокупностью центральных моментов, кумулянтов, а также, что немаловажно, набором полиспектров:

$$A^2(w), B^3(w_1, w_2), \dots, C^n(w_1, \dots, w_{n-1}), \quad (13)$$

где $A^2(w)$ – спектр мощности; $B^3(w_1, w_2)$ – биспектр; $C^n(w_1, \dots, w_{n-1})$ – мультиспектр n -го порядка.

Очевидно, что для анализа РС целесообразно использовать оценку совокупности спектра и биспектра, которая несет в себе более полную и «открытую» информацию о РС, нежели только одномерный Фурье-спектр.

В практике полиспектрального анализа более известен метод косвенного получения биспектра путем прямого преобразования Фурье от тройной автокорреляционной функции (ТАКФ) $R_U(a, b)$, раскрывающей физическое значение бикорреляции и описывающей в статистическом смысле сечение случайного процесса (СП) по кумулянту 3-го порядка – коэффициенту асимметрии (для случая эргодического СП):

$$R_U(a, b) = \left\langle \sum_{k=0}^{K-1} U[kT]U[(k+a)T]U[(k+b)T] \right\rangle_K, \quad (14)$$

$$R_U(a) = \left\langle \sum_{k=0}^{K-1} U[kT]U[(k+a)T] \right\rangle_K, \quad (15)$$

$$R_U(a) = R_U(-a). \quad (16)$$

Как и классическая функция автокорреляции (15), которая отличается свойством симметрии вида (16), ТАКФ также обладает следующим важным для практики свойством симметрии:

$$\begin{aligned} R_U(a, b) &= R_U(b, a) = R_U(b-a, -a) = \\ &= R_U(a-b, -b) = R_U(-a, b-a). \end{aligned} \quad (17)$$

Из определений, сформулированных в фундаментальных работах по полиспектральному анализу [7, 8], биспектр – это преобразование Фурье от ТАКФ:

$$\dot{B}_U(p, q) = \sum_{a=-K+1}^{K-1} \sum_{b=-K+1}^{K-1} R_U(a, b) \exp[-j2\pi(ap + bq)]. \quad (18)$$

При этом биспектр – это комплексная функция двух частотных переменных p и q :

$$\dot{B}_U(p, q) = |\dot{B}_U(p, q)| \exp[j\gamma_U(p, q)], \quad (19)$$

где $|\dot{B}_U(p, q)|$ – биамплитуда, $\gamma_U(p, q)$ – бифаза.

С учетом симметрии ТАКФ и биспектра выражение (19) преобразуем к виду:

$$\dot{B}_U(p, q) = \sum_{a=0}^{K-1} \sum_{b=0}^{K-1} R_U(a, b) \exp[-j2\pi(ap + bq)]. \quad (20)$$

В работах [7-9] приводятся основные свойства биспектра, такие как:

1) биспектр – периодическая функция с периодом, равным 2π :

$$\dot{B}_U(p, q) = \dot{B}_U(p + 2\pi, q + 2\pi), \quad (21)$$

2) биспектр – это симметричная функция на биспектральной плоскости (плоскости двух независимых частотных переменных p и q):

$$\begin{aligned} \dot{B}_U(p, q) &= \dot{B}_U(q, p) = \dot{B}_U(-q, -p) = \\ &= \dot{B}_U(-p, -q) = \dot{B}_U(-p - q, q) = \\ &= \dot{B}_U(p, -p - q) = \dot{B}_U(-p - q, p) = \\ &= \dot{B}_U(q, -p - q) = \dots \end{aligned}, \quad (22)$$

3) Биспектр обладает свойством подобия, что явно раскрывает прямую взаимосвязь между биамплитудой и амплитудным Фурье-спектром, а также взаимосвязь бифазы и фазового Фурье-спектра:

$$|\dot{B}_U(p, q)| = |\dot{A}_U(p) \dot{A}_U(q) \dot{A}_U(p+q)|, \quad (23)$$

$$\gamma_U(p, q) = \varphi_u(p) + \varphi_u(q) - \varphi_u(p+q). \quad (24)$$

Свойства, показанные в выражениях (23, 24), широко применяются в итерационных рекурсивных алгоритмах восстановления сигнала по оценке биспектра, а также непосредственно при прямом методе получения значений биспектра сигнала из Фурье-спектров.

Оценка биспектра есть совокупность двумерного преобразования Фурье от ТАКФ, полученной в результате усреднения по ансамблю временных реализаций СП, однако при рассмотрении РС как квазиэргодического процесса на участках локальной стационарности дает возможность перейти к вычислению биспектра по конкретной реализации СП:

$$\dot{B}_U(p, q) = \left\langle \dot{B}_S(p, q) \right\rangle + \left\langle \dot{B}_N(p, q) \right\rangle, \quad (25)$$

В условиях «стандартного» зашумления, т.е. в случае помехи, представляющей собой процесс вида белого шума с нулевым средним значением, биспектр искаженного РС примерно равен биспектру чистого сигнала речи, так как *статистики высоких порядков гауссовых процессов стремятся к нулю*, что дает возможность в некоторой степени ослабить шумоподобные компоненты. Однако данный факт не является основой предложенного подхода при моделировании системы предобработки, а именно, если сигнал помехи отличается либо по уровню среднего значения, либо по своему распределению от

нормального закона, то применение такого подхода укажет на факт наличия зашумления полезного сигнала негауссовым процессом. Тогда при этом получение биспектральных оценок шумового воздействия позволяет эффективно провести очистку речи от него путем вычитания биспектров. Исходя из этого, необходимо более точно описывать статистические свойства РС, дающие в приближенной степени реальное описание случайного процесса.

На сегодняшний момент алгоритмы спектрального вычитания обладают следующими недостатками [2]:

- 1) необходимость обнаружения явного факта паузы в обрабатываемом РС для оценки и расчета уровня помехи;
- 2) аппроксимация закона распределения сигнала помехи нормальным законом;
- 3) пренебрежение влиянием помехи на фазовый спектр сигнала;
- 4) появление после спектрального вычитания в реконструированном РС искажений, известных как «музыкальные тона».

Основным предположением, которое используется в современных алгоритмах с применением спектрального вычитания, является то, что оценка шумового сигнала в паузе сигнала равна аддитивному воздействию помехи на сам сигнал на активном участке:

$$A_N^{pause}(j\omega) = A_N^{aktiv}(j\omega), \quad (26)$$

$$A_U^{aktiv}(j\omega) = A_S^{aktiv}(j\omega) + A_N^{aktiv}(j\omega). \quad (27)$$

В связи с этим сильно возрастает зависимость данных алгоритмов от самого факта правильного обнаружения паузы, который наглядно можно представить согласно следующим пунктам:

- 1) поиск паузы при слабом зашумлении, основным классификационным признаком сегментации служит интервальная оценка от $-U_1^{trech}$ до U_1^{trech} пороговой амплитуды:

$$\left| U(kT) \right| \leq U_1^{trech}, \quad (28)$$

$$\rho_\Sigma = \sum_{k=0}^{K-1} \rho_k, \quad (29)$$

$$\rho_\Sigma \geq \rho_\Sigma^{trech} \quad (30)$$

при выполнении условия (28) отсчету $U(kT)$ присваивается признак паузы $\rho_k=1$, в противном случае – 0.

Тогда суммарный коэффициент (29) показывает суммарное количество отсчетов, удовлетворяющих пороговому условию (28), вследствие чего очевидно появление в оценке порогового значения ρ_Σ^{trech} , зависящего от количества отсчетов, в сегменте обрабатываемого РС, тогда при выполнении условия (30) сегмент определяется как пауза.

В работах [1, 3] показаны решения по адаптации порогового значения U_2^{trech} , получаемого путем увеличения U_1^{trech} (28);

- 2) значительное увеличение U_2^{trech} ведет к увеличению участков, потенциально определяемых как пауза, и влечет за собой ошибку второго рода или «ложную тревогу», когда активные составляющие обрабатываемого РС будут ошибочно идентифицированы как сегменты пауз, что приводит, в конечном счете, к ошибочной оценке шумового воздействия.

С учетом этих недостатков в работе [2] предложены решения поиска сегмента паузы во временной области РС путем подсчета количества знакоперемен на сегменте μ_Σ^\pm относительно амплитудной точки «0», основанного на предположении о шумовом воздействии как о гауссовом процессе:

$$\mu^{+trech} - \varepsilon \leq \mu_\Sigma^\pm \leq \mu^{+trech} + \varepsilon, \quad (31)$$

$$\mu^{+trech} \cong K/2 \quad (32)$$

и в энергетической области путем нахождения мгновенной энергии сегмента сигнала E_m :

$$E_m = \sum_{k=0}^{K-1} U^2(kT), \quad (33)$$

$$E_m \leq E_m^{trech}. \quad (34)$$

Отметим, что пороговое значение μ^{+trech} зависит от K – количества отсчетов в сегменте РС; ε – параметр, характеризующий некий доверительный интервал; E_m^{trech} – пороговое значение по мгновенной мощности, определяемое, как правило, эмпирическим путем. Тогда при выполнении условий (32, 34) сегмент определяется как пауза.

В случае шумового воздействия высокой энергетики, представляющего собой процесс негауссова вида (2), способы, представленные выше, которые, в свою очередь, ориентированы на нормальный закон распределения шумового воздействия, дают значительные ошибки, так как

сегмент паузы определяется неточно или вообще не определяется.

Вследствие этого исследование корреляции высоких порядков обрабатываемых сегментов зашумленной речи и сегментов квазистационарного акустического шума позволит провести правильную сегментацию с целью выделения паузы в обрабатываемом зашумленном РС. При этом анализ энергетических характеристик необходимо свести к минимуму, отталкиваясь, прежде всего, от параметров, получаемых путем оценки корреляции различных порядков.

Как видно из выражения (16), даже в случае шумового воздействия высокой энергетики, обрабатываемый РС достаточно легко разделяется

на биспектр сигналов чистой речи $B_S(p, q)$ и

биспектр шумового воздействия $B_N^{aktiv}(p, q)$; в случае же сложной помеховой обстановки определение в паузе обрабатываемого РС – $U(kT)$

биспектра шумового воздействия $B_U^{pause}(p, q)$, основываясь на тех же принципах, что и (26, 27), возможно проведение шумоподавления в обрабатываемом РС согласно следующим выражениям:

$$B_U^{pause}(p, q) = B_N^{aktiv}(p, q), \quad (35)$$

$$B_S(p, q) = B_U^{aktiv}(p, q) - B_N^{aktiv}(p, q). \quad (36)$$

Экспериментальные исследования. В рамках моделирования системы предобработки зашумленного РС рассматривались вопросы чувствительности и адекватности модели, проводилась оценка реакции модели на шумовое воздействие квазистационарного характера, а также на вносимую погрешность, связанную с возрастанием вычислительной сложности и размерности пространства обработки.

Непосредственно об условиях проведения эксперимента:

1) моделирование осуществляется в программной среде MATLAB;

2) для проверки адекватности модели при воздействии различных сигналов речи используется тестовая фраза – «этот дачный бугор после пожара фальшиво пахнет югославской похлебкой из щавеля»; данная фраза содержит в себе информацию обо всех фонемах русского языка, включая информацию и об аллофонах, кроме того все статистические особенности также учтены;

3) для проверки чувствительности модели используется аддитивно зашумленная тестовая фраза различными видами идеализированных и

реальных шумов квазистационарной природы (АБГШ, двигатель, город, ветер) при различных значениях отношения сигнал-шум (ОСШ);

4) запись тестовой фразы и шумовых сигналов производится от 40 дикторов по 10 записей с частотой дискретизации 8000 Гц и нормировкой, учитывающей количество бит на отсчет с общей длительностью 8 секунд (64000 отсчетов), кроме того, обработка РС ведется на сегменте квазистационарности 20 мс (160 отсчетов), дополнительно используется взаимное окно пересечения соседних сегментов, равное 10 мс, для уменьшения эффекта посегментной обработки;

5) при проверке вносимой погрешности при вычислениях используется значение среднеквадратического отклонения (СКО) исходного сигнала от сигнала, прошедшего все этапы обработки (получение биспектра, восстановление спектров Фурье по оценке биспектра, получение сигнала во временной области путем обратного преобразования Фурье), кроме того, для каждого опыта выбираются произвольным образом диктор, номер записи и шум, различный по виду и уровню.

В ходе экспериментальной проверки получены следующие важные для практического применения результаты и выводы.

1. В условиях сложной помеховой обстановки при низких отношениях сигнал-шум исходный сигнал претерпевает различные структурные изменения (в зависимости от вида шума), что явно указывает на необходимость индивидуального подхода к каждому шумовому воздействию. Однако из-за наличия различных шумов в реальных условиях подход должен быть обобщенным, что, несомненно, указывает на возможность применения аппарата полиспектрального анализа.

2. Дополнительные вычисления, связанные с возрастанием степени пространства анализа исследуемой зашумленной речи, не вносят существенных искажений во временные реализации РС и шумов на сегменте локальной стационарности. Кроме того оценка СКО по 2500 опытам показала, что математическое ожидание дисперсии, анализируемой на сегменте, стремится к нулю, ограничено степенью округления абсолютных значений отсчетов сигналов и не превышает значения 10^{-14} (конкретно для проверки математической модели использовался тип данных double).

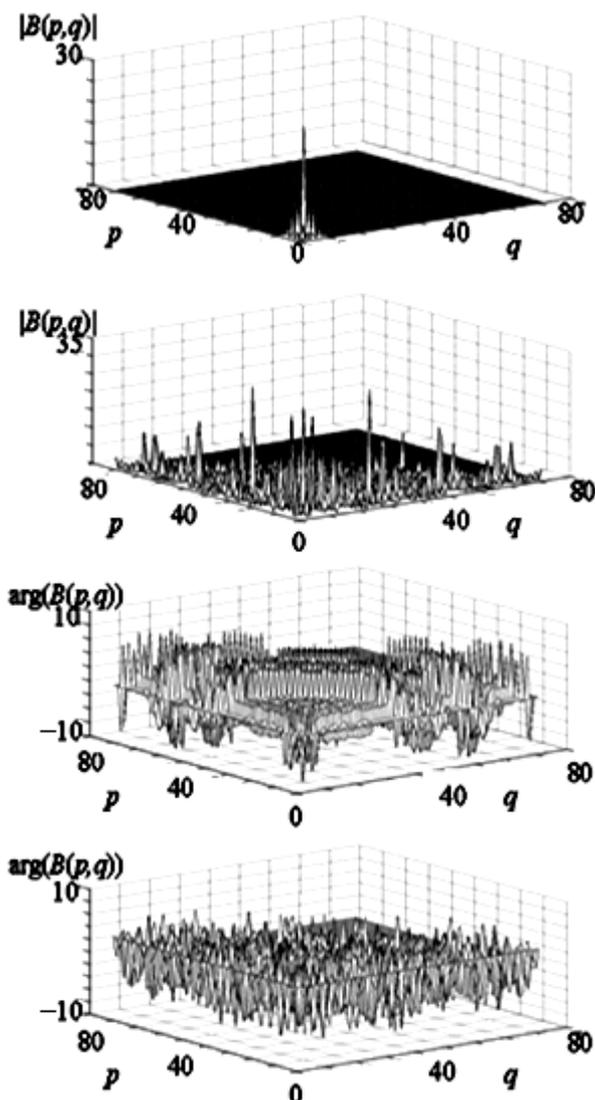
3. Из проведенного анализа очевидно, что шум высокой энергетики ведет к серьезным изменениям РС как в высокоплотностной, так и низкоплотностной областях биспектра. Это вле-

чет за собой и нелинейные искажения, а также полное разрушение исходных фазовых соотношений, которые имеют значение в условиях сложной помеховой обстановки.

Вследствие этого предложенный подход, состоящий в вычитании из биспектра зашумленной речи биспектра шума, несомненно, позволит провести очистку и фазовых соотношений от изменений, возникших из-за воздействия шума.

Факт изменения оценок биспектра исследуемых сегментов зашумленной речи предлагаем проиллюстрировать следующим рисунком, на котором представлены (сверху вниз):

- биамплитуда К-сегмента тестовой фразы при отсутствии шума;
- биамплитуда К-сегмента тестовой фразы, зашумленной АБГШ до ОСШ 0 дБ;
- бифаза К-сегмента тестовой фразы при отсутствии шума;
- бифаза К-сегмента тестовой фразы, зашумленной АБГШ до ОСШ 0 дБ.



Исследуемые сигналы в биспектральной области

Дополнительно отметим, что в статье с целью экономии информационного пространства не рассматриваются аспекты восстановления спектров Фурье по оценкам биспектра. Данные итерационные рекурсивные алгоритмы достаточно подробно представлены с учетом всех модификаций в источниках [10-11].

Заключение. В контексте представленной статьи основной упор делается на раскрытии преимуществ дополнительной оценки биспектра РС относительно «классических» методов обработки речи. При этом основой является не снижение влияния статистик высоких порядков на поведение РС, а именно использование полученной информации для оценки шумового воздействия и проведения шумоподавления согласно выражениям (35, 36). Это, по своей сути, не ограничивает возможности применения спектрального анализа для решения частных задач (взвешивания уровня шумового воздействия, восстановления фазы «квазичистого» сигнала путем оценивания спектра мощности для определения группы фонем обрабатываемого сегмента и др.) с целью эффективного шумоподавления обрабатываемого зашумленного РС в условиях сложной помеховой обстановки. Также данная статья не ориентирована на возможные алгоритмы, которые разработаны согласно предложенному теоретико-математическому подходу, а раскрывает именно момент возможности применения аппарата полиспектрального анализа для системы предобработки зашумленного РС различными видами квазистационарных шумов.

Таким образом, корреляция третьего порядка позволяет получить несколько больше информации о характере исследуемого зашумленного РС путем раскрытия вопросов взаимосвязи между компонентами Фурье-спектров. При этом дополнительный анализ биспектра в совокупности с «классическим» спектральным анализом дает более точные оценки шумовых воздействий. Стоит отметить, что информация, полученная из дополнительной оценки биспектра, носит, прежде всего, качественный, а не количественный характер. Пояснить данный факт можно согласно выражению (12), носящему также название закона Эджворта. Анализ показывает, что по сравнению с существующими методами аппроксимации закона распределения зашумленного РС в совокупной оценке мултиспектров мы имеем на $\sum_{k=3}^{\infty} (-1)^k \frac{\chi_k}{k!} W_{G_c}^{(k)}(u)$ больше качественной информации.

Однако количество, или, скажем, «цена», этой информации убывает с обратно фактори-

альной зависимостью, вследствие чего дополнительный анализ биспектра РС позволяет получить на $\frac{1}{3!} = \frac{1}{6}$ больше качественной информации об анализируемом сегменте РС, представленном квазиэргодическим процессом на участке локальной стационарности.

Потенциал полиспектрального анализа относительно биспектрального анализа и спектрального анализа по отдельности виден и неоспорим, однако стоит отметить и значительное увеличение вычислительной сложности алгоритмов и повышение требований к ресурсам как постоянной, так и оперативной памяти устройства обработки речи. Данный аспект можно легко представить в виде закономерности увеличения количества анализируемых значений в различных порядках мультиспектров, например, при работе со спектром мощности РС, полученного с частотой дискретизации, равной 8000 Гц, необходимо проанализировать ≈ 80 значений (при условии, что на сегменте квазистационарности 160 отсчетов сигнала во временной области), а для анализа биспектра необходимо уже оценить ≈ 3240 значений.

Дополнительно укажем на то, что недостатки существующих способов шумоподавления на основе спектрального вычитания в предложенной математической модели во многом устранены. Так как анализу подвергается, в первую очередь, не вся спектральная плотность мощности в предположении о нормальном законе распределения параметров шумового воздействия, а конкретно каждая биспектральная компонента. Кроме того, все фазовые связи в обрабатываемом РС сохраняются, что улучшает узнаваемость

диктора по сравнению с «классическими» методами обработки зашумленной речи.

Библиографический список

1. Рабинер Л., Шафер Р. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 97 с.
2. Чучупал В. Я. Цифровая фильтрация зашумленных речевых сигналов. – М.: Вычислительный центр РАН, 1998. – 52 с.
3. Покровский Н. Б. Расчет и измерение разборчивости речи. – М.: Связьиздат, 1962. – С. 42–64.
4. Тоцкий А. В. Восстановление сигналов по оценкам биспектров в присутствии гауссовых и негауссовых помех / Зарубежная радиоэлектроника. – 2002. – № 11. – с. 44–58.
5. Малахов А.Н. Кумулянтный анализ негауссовых случайных процессов и их преобразований. – М.: Сов. радио, 1978. – 372 с.
6. Astola J., Heinonen P., and Neuvo Y., On root structures of median and median-type filters, IEEE Trans. Acoust., Speech, Signal Processing, vol. ASSP–35, Aug. 1987, pp. 1199 – 1201.
7. Нукиас Х.Л., Рагувер М.Р. Биспектральное оценивание применительно к цифровой обработке сигналов // ТИИЭР. 1987. Т.75. – №7. – С. 5-30.
8. Важинский В.Н. Корреляционная функция третьего порядка и биспектр в задачах обработки сигналов // ОМП. 1991. – №4. – С. 4-14.
9. Zhang Ji-Wu, Zheng Chong-Xun, and Xie Au, Bispectrum Analysis Of Focal Ischemic Cerebral EEG Signal Using Third-Order Recursion Method, IEE Trans. Biomedical Engineering, Vol. 47, No. 3, March 2000, pp. 765 – 774.
10. Bartelt H., Lohman A. W., and Wirtitzer B., Phase and amplitude recovery from bispectra, Applied Optics, vol. 23, Sept. 1984, pp. 3121–3129.
11. Бочков Г.Н., Горохов К.В. Способ синтеза биспектрально-организованных сигналов // Письма в ЖТФ. 1995. Т.21. В.16. – С. 27-32.

УДК 621.396.663

М.А. Коробков

МЕТОД УСТРАНЕНИЯ ЛОЖНОГО ПЕЛЕНГА МНОЖЕСТВЕННЫХ ИСТОЧНИКОВ РАДИОИЗЛУЧЕНИЯ ПРИ ИСПОЛЬЗОВАНИИ ПЕЛЕНГАЦИОННОГО АЛГОРИТМА РЕДУКЦИИ РАНГА

Предложен метод устранения ложных координат возникающих в процессе пеленга сигналов источников радиоизлучения (ИРИ), принимаемых однородной кольцевой антенной решеткой (ОКАР) с помощью пеленгационного алгоритма редукции ранга (АРР). Приведены результаты численного моде-

лирования процесса пеленга трех ИРИ, имеющих самое неблагоприятное для работы APP взаимное угловое расположение. Показано, что с помощью предложенного метода удастся полностью устранить ложные координаты.

Ключевые слова: пеленгация, кольцевая решётка, алгоритм редукции ранга.

Введение. В настоящее время известны различные алгоритмы пеленга множественных ИРИ, основанные на обработке ковариационной матрицы (КМ) сигналов, принимаемых элементами антенных решеток (АР) [1-3]. Базовым является алгоритм многосигнальной классификации (МСК), который в зарубежной литературе обозначается аббревиатурой MUSIC [4]. К сожалению, алгоритм МСК, обладая многочисленными достоинствами, вместе с тем, характеризуется и весьма существенным недостатком, состоящим в том, что его применение приводит к необходимости выполнения двухмерного поиска координат ИРИ, что влечет за собой значительные затраты ресурсов вычислительной системы и времени поиска решения. Поэтому в литературе предложены алгоритмы, позволяющие в той или иной мере устранить этот недостаток. К их числу относится и рассматриваемый здесь алгоритм редукции ранга (APP) [5-7].

APP является обобщением вычислительно эффективной модификации алгоритма МСК, которая может быть применена для ОКАР. Используя теорию редукции ранга, удаётся преобразовать ОКАР в виртуальную линейную решётку, к которой может быть применен алгоритм К-МСК – алгоритм МСК с нахождением корней (в зарубежной литературе именуемый Root-MUSIC) [5]. Основная идея APP состоит в сведении решаемой задачи двухмерного углового поиска к одномерному поиску угловых координат ИРИ сначала по азимуту, а затем и по углу места. Однако на первом же из указанных двух этапов возникает серьезная проблема, связанная с появлением ложных координат.

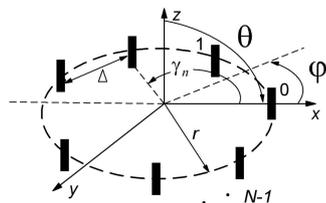


Рисунок 1 – Однородная кольцевая антенная решётка

Известны различные методы отбраковки ложных решений, возникающих при работе APP. Например, в работе [5] предлагается устранять их путем решения дополнительного полиномиального уравнения, корни которого нейтрализуют

ложные решения. Однако последующие исследования, проведенные в работах [6, 7], показали, что при определенных угловых положениях предложенный метод неверно выделяет координаты, соответствующие заданным угловым положениям. Несостоятельность данного метода проявляется в следующих случаях:

1) когда азимутальный угол одного из источников равен среднеарифметическому значению азимутальных углов двух других ИРИ, то есть $\varphi_k = (\varphi_m + \varphi_n) / 2$. При этом попытка устранения зеркальной координаты для одного из источников m или n может привести к удалению истинной координаты φ_k ;

2) когда координата одного источника отличается от координаты другого источника на π , например $\varphi_k = \varphi_m + \pi$.

В работах [6, 7] предлагаются другие подходы к решению задачи, требующие затратных в вычислительном отношении процедур.

Цель работы состоит в разработке альтернативного алгоритма, позволяющего устранить ложные координаты, получаемые в результате работы APP.

Теоретические исследования. Пусть $\mathbf{x}_{e,(N \times 1)}$ – вектор сигналов, принимаемых элементами ОКАР в некоторый момент времени. Здесь и далее индексами в скобках будем обозначать размерности векторов и матриц. В соответствии с APP преобразуем его в пространство лучей с помощью матрицы $\mathbf{W}_{m,n} = (e^{j2\pi nm/N}) / \sqrt{N}$, формируя вектор $\mathbf{x}_{b,(M' \times 1)} = \mathbf{W} \mathbf{x}_e$, где $M' = 2M + 1$ – размерность пространства лучей, $m = -M, \dots, M$, $n = 0, \dots, N - 1$, $M = \lfloor k_0 r \rfloor$, k_0 – волновое число, r – радиус ОКАР, символ $\lfloor \cdot \rfloor$ определяет целую часть числа.

В APP вначале независимо от угла места проводится вычисление азимута путём решения следующего полиномиального уравнения [5]

$$f_{RARE}(z)_{|z|=1} = \det \left\{ \mathbf{T}(1/z)^T \mathbf{E}_{N_b} \mathbf{E}_{N_b}^H \mathbf{T}(z) \right\} = 0, \quad (1)$$

где $\det \{ \cdot \}$ – определитель матрицы, $\mathbf{E}_{N_b(M' \times (M'-L))}$ – матрица, составленная из собственных векторов ковариационной матрицы $\mathbf{R}_b = E \{ \mathbf{x}_b \mathbf{x}_b^H \}$,

соответствующих подпространству шума, $E\{\cdot\}$ – означает определение математического ожидания величины, стоящей в фигурных скобках, L – число ИРИ, верхние индексы $(\cdot)^H$ и $(\cdot)^T$ означают выполнение операций эрмитова сопряжения и транспонирования матрицы, соответственно, матрица $\mathbf{T}_{(M' \times (M+1))}(z)$ имеет следующую структуру:

$$\mathbf{T}(z) = \begin{bmatrix} \mathbf{Q}(z) & \mathbf{0} \\ \mathbf{0} & 1 \\ \mathbf{PQ}(1/z) & \mathbf{0} \end{bmatrix}, \quad (2)$$

где $\mathbf{Q}_{M \times M}(z) = \text{diag}\{z^{-M}, z^{-(M-1)}, \dots, z^{-2}, z^{-1}\}$ – диагональная матрица, $z = e^{j\varphi}$, φ – азимут, $\mathbf{P}_{M \times M}$ – обменная матрица.

Таблица 1 – Этапы выполнения предлагаемого метода

Этап	Процедура
1	Перейти от вектора отсчета сигналов \mathbf{x}_e к вектору \mathbf{x}_b в пространстве лучей
2	Сформировать ковариационную матрицу \mathbf{R}_b
3	Определить матрицу \mathbf{E}_{N_b} , составленную из собственных векторов матрицы \mathbf{R}_b , соответствующих ее собственным значениям в пространстве шума
4	Найти корни полинома $f_{RARE}(z)$, решая уравнение (1).
5	Провести предварительную фильтрацию оценок с помощью соотношения $ 1 - z_i \leq \delta_{\max}$
6	Объединить корни z_i , имеющие одинаковые значения φ_i путем их усреднения
7	Для каждой из оставшихся оценок азимутального угла φ_i рассчитать зависимости функции $f_{\text{МСК}}(\varphi_i, \theta)$ от угла θ в интервале от 0° до 90° и путем одномерного поиска определить θ_i , при которых достигаются максимумы функции $f_{\text{МСК}}(\varphi_i, \theta_i)$
8	Отсортировать пары (φ_i, θ_i) в порядке убывания функции $f_{\text{МСК}}(\varphi_i, \theta_i)$ и выделить среди них первые L пар углов, при которых максимумы принимают наибольшие значения. Они и будут трактоваться как истинные оценки положений ИРИ

Исследование полинома в уравнении (1) показывает, что его корни, пример которых приведен на рисунке 2,а, обладают следующими свойствами:

- 1) общее число корней $\eta = M'^2 - 1$;
- 2) если z является корнем уравнения (1), то $1/z^*$ также является корнем уравнения (1), где $*$ – знак комплексного сопряжения. Это означает, что корню, находящемуся внутри единичной окружности, соответствует корень, расположенный с внешней стороны окружности;
- 3) если z является корнем уравнения (1), то $-z$ также является корнем уравнения (1). Это означает, что корню уравнения (1) соответствует корень, находящийся симметрично относительно начала координат. Иными словами, корню z соответствует корень $z \pm \pi$. Далее будем называть такие корни зеркальными.

Согласно свойству 1, общее число корней η больше числа ИРИ L . Следовательно, для определения координат ИРИ необходимо выполнить процедуру удаления ложных корней, соответствующих ложным координатам.

Пошаговая последовательность предлагаемого решения задачи указана в таблице 1. Первые четыре этапа процедуры являются стандартными для АРР. Полагается, что число ИРИ L либо известно, либо верно оценено. Пятый этап также выполняется в АРР. Многочисленные расчеты, проведенные на данном этапе, привели нас к следующему выводу: при различном взаимном расположении и числе ИРИ, мощности их сигналов, а также при изменении параметров, описывающих ОКАР, корни уравнения (1), соответствующие ИРИ, находятся на комплексной плоскости в кольце $|1 - |z_i|| \leq \delta_{\max}$, где $\delta_{\max} \leq 0.2$. На шестом этапе процедуры происходит объединение пары корней, удовлетворяющих свойству 2, путем их усреднения. На седьмом этапе осуществляется одномерный поиск максимумов спектральной функции МСК, которая имеет следующее представление:

$$f_{\text{МСК}}(\theta, \varphi) = \left[\mathbf{a}_b^H(\theta, \varphi) \mathbf{E}_{N_b} \mathbf{E}_{N_b}^H \mathbf{a}_b(\theta, \varphi) \right]^{-1}, \quad (3)$$

где $\mathbf{a}_b(\theta, \varphi) = \mathbf{W} \mathbf{a}_e(\theta, \varphi)$ – вектор отклика элементов ОКАР в пространстве лучей на ИРИ, приходящий с углового направления, задаваемого углами θ и φ . Для ОКАР с изотропными элементами

$$\mathbf{a}_e(\varphi, \zeta) = \left[e^{j\zeta \cos(\varphi - \gamma_0)}, \dots, e^{j\zeta \cos(\varphi - \gamma_{N-1})} \right]^T, \quad (4)$$

где $\zeta = k_0 r \sin(\theta)$, $\gamma_n = 2\pi n / N$ – угловое положение n -го элемента ОКАР.

На последнем, восьмом этапе, пары (φ_i, θ_i) сортируются в порядке убывания функции $f_{\text{МСК}}(\varphi_i, \theta_i)$ и среди них выделяются первые L пар углов, при которых максимумы спектральной функции достигают наибольших значений. Эти пары углов трактуются как верные координаты ИРИ.

Экспериментальные исследования. При проектировании линейных, планарных и кольцевых антенных решеток обычно выбирают расстояние между излучателями, равное, примерно, половине длине волны λ в вакууме. Однако в пеленгационных ОКАР, в которых при обработке сигнала используются АРР, следует выбирать шаг Δ между элементами, меньший половины длины волны [8]. Обычно его задают в интервале $0.3\lambda \leq \Delta \leq 0.5\lambda$.

Приведем пример, демонстрирующий результат применения предложенной методики. Рассмотрим ОКАР с числом элементов $N = 12$ и расстоянием между элементами $\Delta = 0.4\lambda$. Предположим, что излучают три ИРИ, взаимное угловое расположение которых соответствует наиболее неблагоприятному случаю для работы АРР, при котором $\varphi_2 = (\varphi_1 + \varphi_3) / 2$, $\varphi_3 = \varphi_1 + \pi$, а именно, $(\varphi_1 = 30^\circ, \theta_1 = 30^\circ)$, $(\varphi_2 = 120^\circ, \theta_2 = 35^\circ)$ и $(\varphi_3 = 210^\circ, \theta_3 = 40^\circ)$. Сигналы ИРИ предполагаются некоррелированными, количество выборок сигнала с каждого элемента ОКАР $K = 512$, отношение сигнал-шум (ОСШ) 15 дБ, в канале действует аддитивный белый гауссовский шум.

На рисунке 2,а треугольниками изображены положения точек, соответствующих заданным трем ИРИ, а кружочками – решения уравнения (1). Видно, что помимо корней, приближающихся к точкам, соответствующим заданным ИРИ (сигнальные корни), присутствуют также ложные (паразитные и зеркальные), соответствующие ложным пеленгам ИРИ. При реально заданных параметрах ОКАР получаем в данном случае $\eta = 80$ корней уравнения (1). После пятого этапа, при $\delta_{\text{max}} = 0.05$, число корней уменьшается до $\eta = 12$, рисунок 2, б), а после шестого этапа – $\eta = 4$, что приводит к сокращению количества корней по сравнению с исходным значением в 20 раз.

Теперь, выполняя седьмой этап процедуры, строим для всех четырёх углов $(\varphi_i = 29.8^\circ, 119.9^\circ, 209.8^\circ, 299.9^\circ)$ зависимости спектральной функции алгоритма МСК от зенитного угла θ , рисунок 3, и путем одномерного

поиска определяем их максимумы, шаг поиска 0.1° . Наконец, на последнем этапе сортируем значения этих максимумов в порядке убывания и выбираем из них те $L = 3$, которые имеют наибольшее значение. Получаем три пары углов, положение которых максимально приближено к заданным ИРИ (см. таблицу 2). Также в таблице 2 приведен пеленг, полученный алгоритмом МСК для рассматриваемого примера (шаг поиска по координатам угла места и азимута составлял 0.1°).

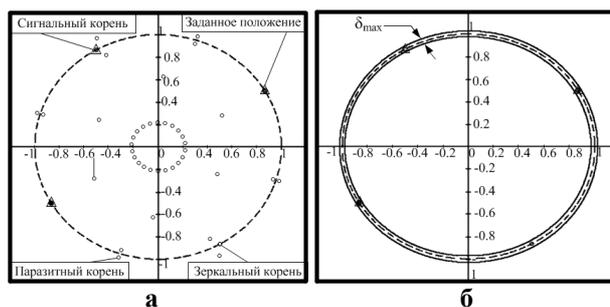


Рисунок 2: а – результат решения уравнения (1); б – результат предварительной фильтрации корней $\delta_{\text{max}} = 0.05$

Приведенный пример наглядно демонстрирует, что пеленг ИРИ с наиболее неблагоприятным для работы АРР взаимным угловым расположением, с помощью предложенного алгоритма, действительно позволил избавиться от всех паразитных корней уравнения (1).

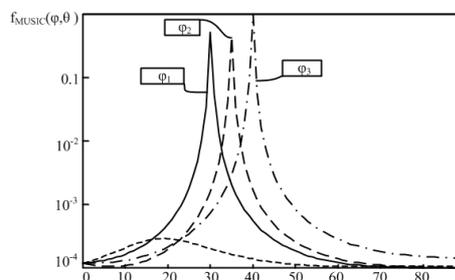


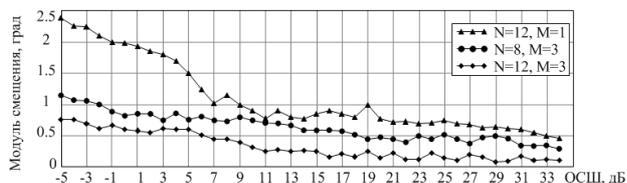
Рисунок 3 – Нормированное к единице значение функции (3), вычисленное для фиксированных значений азимута

Таблица 2 – Окончательные оценки координат ИРИ

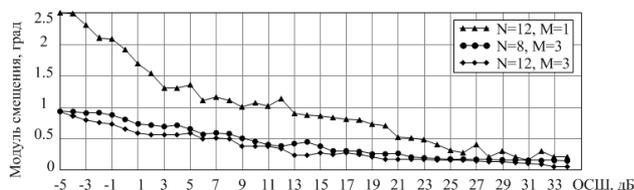
№ ИРИ	Оценки координат				Заданные координаты	
	АРР+метод		МСК		азимут	угол места
	азимут	угол места	азимут	угол места		
1	29.8°	30°	30°	30°	30°	30°
2	119.9°	35°	120°	35°	120°	35°
3	209.8°	40°	210°	40°	210°	40°

Модуль смещения пеленга при различных параметрах конфигурации ОКАР при различном уровне ОСШ приведен на рисунке 4. Для построения зависимостей проведено от 1500 до

2000 испытаний, предполагалось, что пеленгуется один ИРИ, координаты которого задавались случайным образом.



а



б

Рисунок 4 – Модуль смещения пеленга одного ИРИ при различных параметрах ОКАР в зависимости от ОСШ: а – по азимуту; б – по углу места

Заключение. Описан метод устранения ложных корней, возникающих в процессе решения задачи пеленга ИРИ с помощью АРР. Достоинство предложенного варианта выполнения алгоритма состоит в том, что он надежно функционирует даже при таком взаимном расположении ИРИ, которое традиционно считается наиболее неблагоприятным для работы АРР.

УДК 004.93

М.Ю. Волынец, К.А. Майков

КОМБИНИРОВАННЫЙ МЕТОД РАСПОЗНАВАНИЯ ЛИЦ В РЕЖИМЕ РЕАЛЬНОГО ВРЕМЕНИ

Предложен комбинированный метод распознавания и обнаружения лиц на цифровых изображениях и видеопоследовательностях в режиме реального времени на основе сравнительного анализа известных методов решения задачи распознавания и обнаружения лиц. Предложенный метод обладает большей производительностью, достоверностью решения при различном пространственном расположении объекта анализа и не зависит от условий освещения.

Ключевые слова: распознавание лиц, метод Виолы – Джонса, нейронные сети, обнаружение лиц, метод главных компонент, вейвлет-преобразование Габора.

Введение. За последние годы для решения задачи распознавания человека по изображению лица широкое практическое использование получили такие методы, как метод главных компонент [1], нейросети [2], байесовские сети [3], метод Виолы – Джонса [4] и ряд других подходов.

Результативность применения метода в основном определяется спецификой решаемой задачи, при этом ни один из известных методов не может быть признан однозначно предпочтительнее других [5].

Однако, как показано в [6], имеется возмож-

Библиографический список

1. Коробков М. А. Корреляционные методы пеленгования источников излучения // Молодой ученый. — 2014. — №13. — С. 55-58.
2. Kellogg R.L., Mack E.E., Crews C.D. Direction Finding Antennas and Systems, in J. Volakis (ed.), *Antenna Engineering Handbook, Fourth Edition*, New York, McGraw-Hill, 2007, Chapter 47. P. 1755.
3. Tuncer E., Frienlander B. Classical and Modern Direction-of-Arrival Estimation // Elsevier Inc., 2009. — P. 429.
4. Schmidt R.O. Multiple Emitter Location and Signal Parameter Estimation // IEEE Trans. Antennas Propag. 1986. Vol. AP-34. P. 276—280.
5. Pesavento M., Böhme J.F. Direction of arrival estimation in uniform circular arrays composed of directional elements // Proc. Sensor Array and Multichannel Signal Processing Workshop, 2002, P. 503–507.
6. Wang B.H., Hui H.T., Leong M.S. Decoupled 2D Direction of Arrival Estimation Using Compact Uniform Circular Arrays in the Presence of Elevation-Dependent Mutual Coupling // IEEE Trans. Antennas Propag., 2010, Vol. 58, No. 3.
7. Goossens R., Rogier H. A hybrid UCA-RARE/Root-MUSIC approach for 2-D direction of arrival estimation in uniform circular arrays in the presence of mutual coupling. // IEEE Trans. Antennas Propag. 2007, Vol. 55, P. 841–849.
8. Belloni F, Koivunen V. Beamspace transform for UCA: error analysis and bias reduction // IEEE Transactions on signal processing, Vol. 54, No. 8, August, 2006. P. 3078-3089.

ность построения некоторой комбинации методов обнаружения и распознавания лиц, обеспечивающей более качественные результаты. Выбор конкретных компонентов совокупности методов и последовательность их реализации требуют теоретического обоснования и экспериментальных исследований. *Цель работы* – построить метод распознавания лиц, позволяющий функционально объединить решение задач обнаружения и распознавания в реальном времени.

Теоретические исследования комбинированного метода обнаружения и распознавания лиц. Предлагаемое решение включает два этапа: этап обнаружения лиц и этап распознавания лиц.

Предлагаемый метод обнаружения лиц представляет комбинацию двух методов [7]: метода адаптивного усиления и метода, основанного на нейронных сетях.

В предложенном модифицированном методе обнаружения лиц шаг сканирующего окна, в отличие от стандартной реализации метода Виолы – Джонса, является переменным, что обеспечивает меньшую ресурсоемкость для обработки области анализируемого изображения, с высокой вероятностью не содержащей графический образ лица. Таким образом, скорость перемещения окна в областях, не содержащих изображения лица, выше, чем в областях, в которых содержится изображение анализируемого лица.

Метод распознавания лиц [8] представляет функциональное единство двух методов – модифицированного «Подпространство ЛДА» и метода, основанного на вейвлет-преобразованиях Габора.

В методе «Подпространство ЛДА» предлагается использовать модифицированный метод главных компонент (МГК) применительно к процедуре уменьшения размерности. Модификация заключается в генерации матрицы проекции случайным способом вместо матрицы ковариации, которая используется в стандартной реализации МГК. В отличие от стандартного МГК, где данные проецируются на сжатое пространство, полученное из тренировочного набора, модифицированный является независимым от тренировочных данных, что показывает преимущество предложенного решения в скорости вычисления по сравнению с МГК, приобретающим существенное практическое значение для случая входных данных большой размерности. Предложенный метод не требует переобучения в случае изменения обучающего набора, например, добавления новых изображений с лицами или удаления старых изображений. В МГК при обновлении обучающего набора необходимо пересчитывать вычисленные ранее векторы и повторять

процедуру проекции лиц на пространство. Предложенное решение требует меньше ресурсов для вычисления в силу того, что сложность вычисления матрицы размерности $d \cdot p$ равна $O(pd^2)$, в то время как сложность МГК равна $O(p^2M) + O(p^3)$, где p и d – размерности исходного и преобразованного пространств соответственно, M – количество изображений в обучающем наборе. В [9] показано, что случайные отображения не приносят искажений в данные и сохраняют сходство между объектами.

Экспериментальные исследования комбинированного метода обнаружения и распознавания лиц. Для сравнения реализации с постоянным и переменным шагами были проведены эксперименты на базе лиц ORL [5]. Для анализа производительности предложенного метода используются две метрики, а именно полнота и точность, определяемые следующими формулами [10]:

$$\text{полнота} = \frac{TP}{TP + FN};$$

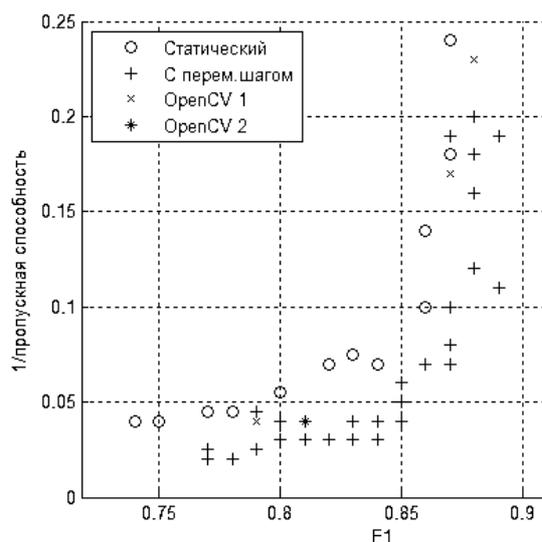
$$\text{точность} = \frac{TP}{TP + FP},$$

где TP – число правильно определенных объектов, FN – число пропущенных объектов, FP – количество ложных обнаружений.

Для систем обработки изображений в режиме реального времени первостепенное значение имеет пропускная способность – число обработанных кадров в секунду. Для визуализации Парето-минимальных решений [11] при учете пропускной способности используется метрика $F1$ [10], объединяющая в себе точность и полноту. Данную метрику можно рассматривать как средневзвешенное значение полноты и точности

$$F1 = 2 \cdot \frac{\text{точность} \cdot \text{полнота}}{\text{точность} + \text{полнота}}.$$

На рисунке представлены точки Парето, полученные для базовых и предлагаемого подходов в двумерном пространстве: «мера $F1$ -пропускная способность». Анализ результатов, приведенных на рисунке 1, показывает, что для заданной точности предложенный подход всегда обеспечивает хотя бы одну конфигурацию, которая обеспечивала бы пропускную способность такую же, как и конфигурации из базовых реализаций с той же (или лучшей) точностью, что говорит о преимущественности предложенного метода. Выигрыш в производительности метода движения сканирующего окна с переменным шагом не является постоянным, а изменяется в зависимости от начальных координат сканирующего окна.



Точки Парето для различных подходов

Для сравнения предложенного решения по модификации метода МГК с использованием базы лиц ORL проведен ряд экспериментов с применением трех различных процедур для оценки точности распознавания:

- сопоставление результатов модифицированного и стандартного МГК при различном числе эталонов в базе;
- сопоставление результатов модифицированного (усредненные значения) и стандартного МГК при различном числе эталонов в базе;
- сопоставление результатов модифицированного и стандартного МГК при поиске заданного количества эталонов лица по заданному тестовому изображению.

При проведении эксперимента по сопоставлению результатов модифицированного и стандартного МГК при различном числе эталонов в базе выполняется сравнение лица из тестового набора с лицами из тренировочного набора. В этом случае точность распознавания вычисляется как отношение корректно распознанных лиц к общему количеству лиц в наборе.

В процедуре сопоставления результатов модифицированного (усредненные значения) и стандартного МГК при различном числе эталонов в базе пять раз (число выбрано произвольно, чем больше это число, тем выше будет точность алгоритма) запускается процедура поиска для каждого лица из тестового набора и вычисляется мера близости в каждом конкретном случае. После этого выбирается наилучшее значение из полученных результатов. Точность распознавания вычисляется как отношение количества корректно распознанных лиц из тестового набора к общему количеству лиц в тестовом наборе.

В процедуре сопоставления результатов модифицированного и стандартного МГК при по-

иске заданного количества эталонов лица по заданному тестовому изображению для тестового лица извлекается объект с минимальным значением меры близости. Затем вычисляется полнота для каждого тестового лица - отношение изображений лиц, полученных из тренировочного набора, которые относятся к тому же объекту, чье лицо изображено на входном изображении, к общему количеству изображений, сохраненных в тренировочном наборе рассматриваемого объекта. Точность распознавания вычисляется усреднением вычисленной полноты для каждого тестового изображения.

В каждом эксперименте изменяется размер изображений, а также количество изображений рассматриваемого объекта в тестовом наборе.

В рассматриваемом эксперименте входное тестовое лицо проецировалось на случайное пространство пониженной размерности, полученное модифицированным методом МГК и стандартным МГК. Для каждого эксперимента изменялась размерность пространства и вычислялась точность распознавания.

В целях исследования модификации метода были выполнены эксперименты, в которых обучающий и тестовый наборы изображений были различными с частичным совпадением изображений.

В первом эксперименте обучающее множество состояло из трех изображений каждого из 40 объектов, т.е. всего 120 изображений. Для второго эксперимента обучающее множество состояло из 130 изображений - по 13 изображений на каждый отдельный объект. Изображения оставшихся 27 объектов были использованы для тестового множества. Источником тестового и тренировочного набора являлась база ORL.

Для более глубокого анализа каждого метода для каждого рассматриваемого объекта из тестового набора изменялось его количество изображений.

Анализ результатов проведенных экспериментов показывал, что в случае, когда тестирование проводится на тех же лицах, на которых было проведено обучение, стандартный метод МГК не проигрывает модифицированному алгоритму, а в случае низкой размерности вычисленного пространства показывает лучшие результаты, однако его преимущественность утрачивается по мере роста размерности.

В случае, когда для обучения и тестирования использовались разные изображения, результат применения двух рассматриваемых методов показывает практически одинаковое значение точности распознавания с некоторым преимуществом модифицированного алгоритма при боль-

ших значениях размерности. Это объясняется тем, что стандартный МГК зависит от входных данных построения собственного пространства, в то время как для модифицированного метода такие зависимости отсутствуют. Однако результаты эксперимента показывают, что результативность указанных методов зависит от количества изображений каждого объекта из тренировочного набора. Производительность алгоритмов снижается при уменьшении количества изображений в тренировочном наборе и увеличении количества изображений в тестовом наборе.

В [9] доказано, что метод случайной проекции показывает лучшую производительность, когда усредняются значения, полученные при нескольких прогонах модифицированного метода. Результаты экспериментов показывают, что при усредненных значениях производительность модифицированного метода выше, чем у стандартного.

В случае одинаковых объектов в обучающей и тестовой выборке стандартный МГК и модифицированный показывают схожий результат, однако преимущество модифицированного алгоритма становится явным при увеличении размерности пространства.

В таблице представлены результаты проверки комбинации выбранных методов обнаружения и распознавания лиц. Модифицированный метод «Подпространство ЛДА» при совместном использовании как с методом, основанном на нейронных сетях, так и с методом адаптивного усиления показывает скорость распознавания, при которой достигается пропускная способность 20 кадров в секунду. Результаты метода Габора зависят от применяемого метода обнаружения лиц: при использовании совместно с методом адаптивного усиления метод обеспечивает меньшее время распознавания сравнительно с комбинацией данного метода с методом, основанным на нейронных сетях, что объясняется большим количеством ложных обнаружений нейросетевым методом.

Из анализа результатов, приведенных в таблице, следует, что последовательное использование методов распознавания модифицированного «Подпространства ЛДА» и метода, основанного на вейвлет-преобразованиях Габора, показывает более высокие результаты верных распознаваний по сравнению с использованием этих методов по отдельности. Увеличение времени обработки не является критичным, поскольку практически важной является точность распознавания, определяющая риск принятия неправильного решения.

Таблица 1 – Результаты проверки комбинированного метода обнаружения и распознавания

Метод	Время обнаружения, с	Время распознавания, с	Общее время, с	Точность, %
Нейронные сети + мод. подпространство ЛДА	0.9	0.005	0.905	86
Нейронные сети + вейвлет-преобразования Габора	0.9	0.75	1.65	87
Модиф. адаптивного усиления + мод. подпространство ЛДА	0.06	0.005	0.065	83
Модиф. адаптивного усиления + вейвлет-преобразования Габора	0.06	0.12	0.18	84
Модиф. адаптивного усиления + мод. подпространство ЛДА \ вейвлет-преобразования Габора	0.06	0.03	0.09	89
Модиф. адаптивного усиления \ нейронные сети + мод. подпространство ЛДА \ вейвлет-преобразования Габора	0.1	0.02	0.12	94

Заключение. Предложен комбинированный метод, представляющий функциональное единство двух составных компонент, первая из которых, решающая задачу обнаружения лиц, образована модифицированным методом адаптивного усиления и методом на основе нейронных сетей. Вторая компонента, решающая задачу распознавания лиц, образована методом на основе вейвлет-преобразований Габора и модификации метода «Подпространство ЛДА». Проведенные эксперименты показали практическую пригодность и преимущественность разработанного

комбинированного метода обнаружения и распознавания лиц применительно к задачам обработки изображений в режиме реального времени.

Библиографический список

1. *Zhao W., Chellapa R., Krishnaswamy A., Swets D., Weng J.* Discriminant analysis of principle components for face recognition // 2nd International Conference on Automatic Face and Gesture Recognition. — April 1998. — pp. 336-341.
2. *Rowley H.* Neural Network-Based Face Detection // School of Computer Science. — 2009.
3. *Yang M.H., Kriegman D.J., Ahuja N.* Detecting faces in images // IEEE Transactions on Pattern Analysis and Machine Intelligence. — 2002. — vol. 24. — pp. 34-58.
4. *Viola P., Jones M.J.* Rapid object detection using a boosted cascade of simple features // In conference on computer vision and pattern. — 2001.
5. *Кухарев Г.А., Щеголева Н.Л.* Системы распознавания человека по изображению лица. СПб.: Изд-во СПбГЭТУ «ЛЭТИ», 2006.
6. *Beumer D., Poggio T.* Face recognition from one example view, Proceedings of the Fifth International Conference on Computer Vision.— 2005. — pp.500-507.
7. *Вольнец М.Ю., Майков К.А.* Метод обнаружения лиц с использованием иерархической структуры признаков // Вестник МГУП. — М., 2013. — С.139-145.
8. *Вольнец М.Ю., Майков К.А.* Комбинированный метод обнаружения и распознавания лиц в реальном режиме // Молодежный научно-технический вестник. — М., январь 2014.
9. *Fradkin D., Madigan D.* Experiments with random projection for machine learning // ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. — 2009. — pp. 571–522.
10. *Goutte C., Gaussier E.* A probabilistic interpretation of precision, recall and f-score, with implication for evaluation // ECIR. — 2010.
11. *Geilen M., Basten T., Theelen B. D., Otten R.* An algebra of pareto points // Fundam. Inform. — 2007. — vol. 78, —pp. 35–74.