

СТАТИСТИЧЕСКИЙ МЕТОД ВОССТАНОВЛЕНИЯ ПРОФИЛЯ ПО ДАННЫМ ЛАЗЕРНОГО СКАНИРОВАНИЯ

Броневиц А.Г., д.ф.-м.н., главный научный сотрудник Центра перспективных фундаментальных и прикладных исследований Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), e-mail: brone@mail.ru .

Каркищенко А.Н., профессор, руководитель Центра перспективных фундаментальных и прикладных исследований Научно-исследовательского и проектно-конструкторского института информатизации, автоматизации и связи на железнодорожном транспорте (ОАО «НИИАС»), e-mail: karkishalex@gmail.com

Уманский В.И., кандидат технических наук, Генеральный директор ЗАО «ИнтехГеоТранс», e-mail: umanvi@yandex.ru

Ключевые слова: профиль земляного полотна, лазерное сканирование, случайная последовательность, функция правдоподобия, дискриминантная функция, разладки случайного процесса.

Введение

В настоящее время развиваются методы автоматического мониторинга профиля земляного полотна и верхнего строения железнодорожного пути на протяженных участках дороги. Под воздействием динамических поездных нагрузок и влиянием природно-климатических факторов происходит деформация полотна, так что со временем оно может не удовлетворять требованиям нормативной документации и условиям дальнейшей эксплуатации. В этом случае предпринимаются меры по поддержке земляного полотна. Постоянная диагностика земляного полотна, учитывая протяженность существующих железнодорожных линий, является трудоемкой, требующей постоянного внимания процедурой [1].

Эффективность диагностики во многом определяется такими свойствами как: объективность, достоверность, простота проведения, оперативность, дешевизна. При этом требуется мониторинг целых участков и направлений, применяющийся для деформаций массового характера на большом протяжении. Особую ценность приобретают методы мониторинга, главными свойствами которых являются быстрота проведения и дешевизна. Наиболее простыми, но важными являются наблюдения за соответствием сложившихся профилей земляного полотна с типовыми нормативными профилями для заданных грунтов, слагающих насыпи или выемки. На основании таких наблюдений и их сравнительного анализа требуется фиксировать такие параметры как:

- ненормативная ширина плеча балластной призмы;
- ненормативная ширина обочины земляного полотна;
- места с завышенной крутизной откосов;
- нарушение структурных линий полотна, представляющих собой продольные линии излома поверхности земли в полосе отвода и др.

В связи с этим возникает необходимость разработки

Статья посвящена восстановлению поперечного профиля земляного полотна по зашумленным данным лазерного сканирования. Приведена формальная постановка задачи. Показано, что задача может быть сведена к анализу случайной последовательности измерений. Дано решение задачи с помощью дискриминантной функции Фишера. Показано, как с помощью построенной дискриминантной функции локализуются точки переключения в кусочно-параметризованной случайной последовательности, и предложен вычислительно эффективный способ приближенного отыскания таких точек. Приведен пример реализации метода на моделируемых данных.

методов автоматического сравнения поперечных профилей на основе использования специальных технических средств получения информации о поверхностном рельефе земли. Для решения этой задачи перспективным является использование современных лазерных сканеров, которые, по сути, представляют собой лазерные дальномеры. Технически это сводится к тому, что посредством геометрических преобразований оказывается возможным восстановить в некоторой координатной системе положение большого количества точек (облака точек) поверхности. Анализ полученных данных позволяет сделать выводы о поверхностном строении просматриваемого участка. Современные лазерные сканеры импульсного типа способны сканировать поверхность объекта со скоростью до 10000 точек в секунду, а сканеры фазового типа - до 500000 точек.

Стандартный способ получения информации о профиле состоит в создании технических комплексов наземного лазерного сканирования и обработки информации непосредственно на перемещающемся поезде [2]. Технологически данная задача решается путем установки двух лазеров на задней стенке последнего вагона. В процессе движения вагона каждый лазер сканирует свою часть полупространства – соответственно, слева и справа относительно движения, формируя совокупность замеров (рис. 1).

В данной работе приведена формальная постановка задачи восстановления поперечного профиля земляного полотна по зашумленным данным лазерного сканирования, описан подход к решению задачи путем сведения ее к анализу случайной последовательности измерений. Дано решение данной задачи с помощью дискрими-

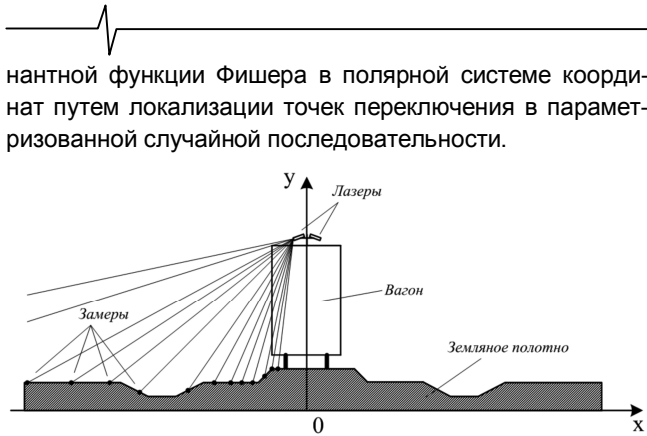


Рис. 1. Схема измерений земляного полотна

Описание и общая постановка задачи

При организации лазерного сканирования следует принимать во внимание ряд особенностей:

- совокупность измерений (замеров) представляет собой дискретное множество точек, количество которых определяется частотой излучения лазера и может варьироваться в зависимости от его технических возможностей;

- за один проход лазеров получается совокупность замеров по общему углу охвата в 360 градусов (по 180 градусов на каждый лазер);

- распределение замеров зависит от высоты расположения лазера над поверхностью земли;

- так как сканирующий луч перемещается дискретно во времени, а значит, и в пространстве, через равный угловой шаг, то замеры расположены относительно равномерно по угловой мере;

- замеры располагаются неравномерно в линейном измерении, поскольку происходит отражение от поверхностей, расположенных под разным углом по отношению к лучу лазера и находящихся на разном расстоянии от источника излучения (при этом замеры сгущаются при приближении луча к балластной призме);

- данные реальных измерений оказываются сильно зашумленными; ошибки измерений возникают из-за погрешностей собственно лазерного дальномера, а также в большей степени ввиду естественной неровности отражающей поверхности земляного полотна, наличия на нем посторонних предметов, растительности и пр.

Ниже рассматривается задача восстановления по лазерным замерам формы земляного полотна и балластной призмы на достаточно протяженном участке дороги. Теоретически решение данной задачи позволит провести автоматическое сравнение реального рельефа полотна с рельефом, соответствующим требованиям нормативных документов.

Рассмотрим один цикл съема информации о земляном полотне с помощью лазерного датчика. Под циклом будем понимать период лазерного сканирования, соответствующий обзору лазера сверху-вниз, т.е. на 180 градусов. Далее будем рассматривать информацию, получаемую только при сканировании земляного полотна, и не будем принимать во внимание точки, возникающие при отражении луча сканера от объектов, находящихся

выше полотна. В результате сканирования получается совокупность замеров в поперечном сечении, т.е. в плоскости, перпендикулярной направлению пути¹. Если точки соединить последовательно отрезками прямых линий, то можно получить представление о форме реального профиля. Полученные замеры будут лишь приближенно описывать земляное полотно. Даже в том случае, если оно соответствует требованиям действующих нормативов, измерения будут содержать случайную составляющую, появление которой вызвано различными причинами: неровностями поверхности земли; различными предметами, отражающими лазерный луч; растительным покровом; снегом; собственно погрешностями лазерного дальномера; продольными и поперечными колебаниями вагона и др.

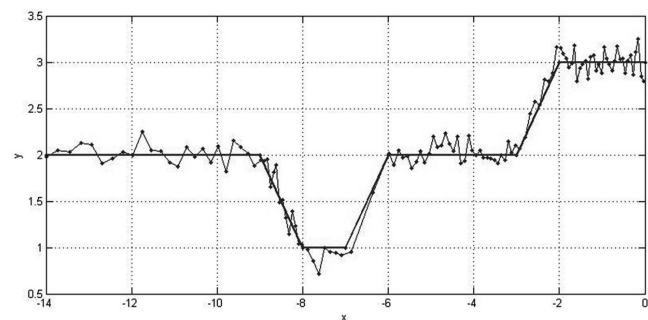


Рис. 2. Пример последовательности замеров

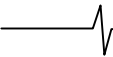
Будем считать, что выполняются следующие предположения:

1. Ошибка каждого измерения в одном цикле является суммой большого количества независимых или слабо зависимых факторов, поэтому она может быть достаточно точно описана случайной величиной, подчиненной нормальному закону распределения;

2. Случайные величины имеют короткий «промежуток зависимости». Иными словами, возмущения слабо связаны с их предысторией, и поэтому статистическая зависимость двух измерений быстро уменьшается с ростом расстояния между ними. Это условие гарантирует, что получаемые на основе таких измерений статистические оценки являются асимптотически гауссовскими в смысле центральной предельной теоремы [3].

В результате одного цикла измерений получается случайная последовательность замеров, которая примерно отражает рельеф земляного полотна слева и справа от железнодорожного пути. Пример такой последовательности, полученной в результате моделирования с гауссовским зашумлением слева от дальномера, приведен на рис. 2. Отметим, что точки показаны в декартовой системе координат, в которой ось абсцисс расположена перпендикулярно направлению перемещения дальномеров, как показано на рис. 1. На рисунке видно, что точки сгущаются по мере приближения к началу координат.

¹ На самом деле из-за движения поезда точки замеров будут лежать на пересечении винтовой линии, описываемой лучом лазера, и поверхности земли, пересекающейся с этой винтовой линией. Будем считать в дальнейшем, что замеры на участке одного цикла путем геометрических преобразований приведены к рассматриваемой перпендикулярной плоскости.



Задача обработки полученных данных состоит в том, чтобы по последовательности измерений восстановить исходную форму профиля полотна (показанную на рис. 2 более толстой линией – графиком кусочно-линейной функции).

Формальное описание статистического анализа профиля полотна

Последовательность замеров можно рассматривать как реализацию некоторого случайного процесса, который является нестационарным, т.е. меняет статистические характеристики на различных участках сканирования. При этом могут быть участки, на которых эти характеристики примерно постоянны, но могут быть и такие, на которых они меняются по неизвестному неслучайному закону. Выявление этого закона или функции, описывающей эти изменения, а также определение точек, в которых вид этой функции меняется, позволяют провести анализ реального профиля. Задачи в подобной или похожей формулировке рассматриваются во многих областях, связанных с обработкой статистических данных (случайные процессы и последовательности, временные ряды, конкретные приложения). Многообразие методов решения указанной задачи определяется тем, какие априорные предположения допускаются в каждом конкретном случае.

Рассмотрим вначале широко используемую и удобную модель кусочно-стационарного случайного процесса, в котором изменение свойств происходит скачком. Такой процесс можно представить в виде последовательности стационарных участков, на которых его свойства одинаковы. Обнаружение стационарных участков можно рассматривать как классическую задачу математической статистики – оценивания точек, в которых происходит изменение свойств кусочно-стационарного случайного процесса. Подобные случайные процессы возникают в разных прикладных задачах – в системах управления в пространстве состояний, при мониторинге окружающей среды, анализе сейсмической активности и др. С этой точки зрения стационарные участки представляют собой устойчивые и возможно периодически меняющиеся в процессе наблюдения состояния порождающего их объекта. Мы будем рассматривать случайные процессы, стационарные в широком смысле [4, 5].

В практических задачах информация об исследуемом случайном процессе обычно представляет собой последовательность измерений (y_1, y_2, \dots, y_N) в некоторых точках (x_1, x_2, \dots, x_N) , а вероятностной моделью описания данных измерений является последовательность $(\xi_1, \xi_2, \dots, \xi_N)$ случайных величин. Таким образом, считается, что измерения (y_1, y_2, \dots, y_N) являются реализациями данных случайных величин. При описании последовательностей $(\xi_1, \xi_2, \dots, \xi_N)$ можно использовать цепи Маркова, различного рода авторегрессионные модели, стохастические разностные уравнения [6].

В силу сделанного выше предположения о коротком «промежутке зависимости» ограничимся в дальнейшем наиболее простой моделью такого типа, а именно, будем считать, что последовательность состоит из независимых случайных величин, т.е. принимаем предположение, что случайные величины ξ_i попарно независимы.

Рассматриваемая ниже модель последовательности независимых случайных величин ориентирована на использование

1) при небольших по объему обучающих выборках (y_1, y_2, \dots, y_N) , которые не могут гарантировать получение оптимальных статистических выводов при использовании стандартных вероятностных методов;

2) когда вычислительные ресурсы не позволяют реализовать применение более трудоемких и сложных вероятностных моделей.

Известно, что даже при использовании модели последовательности независимых случайных величин в задаче обнаружения многократных изменений свойств (или разладок) случайного процесса возникают труднопреодолимые проблемы вычислительного характера при решении ее классическими методами [7].

Предположим, что последовательность измерений (y_1, y_2, \dots, y_N) в точках (x_1, x_2, \dots, x_N) можно описать с помощью последовательности одномерных независимых случайных величин $(\xi_1, \xi_2, \dots, \xi_N)$. Будем вначале считать для простоты, что в этой последовательности можно выделить некоторое число n стационарных участков $[\chi_1, \chi_2 - 1], [\chi_2, \chi_3 - 1], \dots, [\chi_n, \chi_{n+1} - 1]$, $\chi_1 = x_1$, $\chi_{n+1} - 1 = x_N$. Внутри каждого стационарного участка $[\chi_m, \chi_{m+1} - 1]$ случайные величины ζ_i и ζ_j , $x_i, x_j \in [\chi_m, \chi_{m+1} - 1]$, должны обладать постоянными математическими ожиданиями и дисперсиями, т.е. $E(\zeta_i) = E(\zeta_j)$, $D(\zeta_i) = D(\zeta_j)$.

Рассмотрим оценивание точек нарушения стационарности χ_i , $i = 1, 2, \dots, n$, по методу максимального правдоподобия. Будем считать, что на каждом стационарном участке $[\chi_m, \chi_{m+1} - 1]$ функцию $f_m(y)$ плотности распределения случайных величин можно задать параметрически вектором параметров a_m , т.е. $f_m(y) = f(y, a_m)$. В этом случае логарифм функции правдоподобия Фишера имеет вид:

$$\Phi(a_1, \dots, a_n, \chi_1, \dots, \chi_n) = \sum_{m=1}^n \sum_{i=\chi_m}^{\chi_{m+1}-1} \ln[f(y_i, a_m)].$$

Таким образом, для оценивания точек χ_1, \dots, χ_n необходимо найти максимум логарифмической функции правдоподобия относительно аргументов a_1, \dots, a_n , χ_1, \dots, χ_n .

В более общем случае в последовательности одномерных независимых случайных величин на каждом из n участков может наблюдаться тренд математического ожидания, описываемый неслучайной функцией $g(x)$, при постоянной дисперсии, т.е. для ξ_i и ξ_j , таких, что $x_i, x_j \in [\chi_m, \chi_{m+1} - 1]$ имеет место $E(\xi_i) = g_m(x_i)$, $D(\xi_i) = D(\xi_j)$. Тогда для оценки точек переключения χ_i , $i = 1, 2, \dots, n$, по методу максимального правдоподобия следует считать, что на каждом участке $[\chi_m, \chi_{m+1} - 1]$ функция $f_m(y)$ плотности распределения случайной величины ξ_i имеет вид $f_m(y) = f(y, g_m(x_i), a_m)$. В частности, если функции, описывающие тренд, принадлежат одному классу (например, классу линейных или квадратичных функций), то $f_m(y) = f(y, g(x_i, b_m), a_m)$,

где b_m - вектор параметров, описывающих конкретную функцию из данного класса. В дальнейшем такой случайный процесс будем называть кусочно-параметризованным процессом. Заметим, что при $g_m(x) = c_m = const$ данный случай сводится к предыдущему случаю стационарных в широком смысле участков. Логарифм функции правдоподобия Фишера имеет вид:

$$\Phi(a_1, \dots, a_n, b_1, \dots, b_n, \chi_1, \dots, \chi_n) = \sum_{m=1}^n \sum_{i=\chi_m}^{\chi_{m+1}-1} \ln[f(y_i, g(x_i, b_m), a_m)]$$

Следовательно, для оценивания точек χ_1, \dots, χ_N необходимо найти максимум логарифмической функции правдоподобия относительно большего, чем в предыдущем случае, числа аргументов, а именно, $a_1, \dots, a_N, b_1, \dots, b_N, \chi_1, \dots, \chi_N$.

Как правило, построить рекуррентный алгоритм максимизации функционала Φ не удастся, поэтому остается единственная возможность – проводить полный перебор всех возможных положений точек переключения χ_m . Эта задача для случая многократных разладок является недопустимо трудоемкой – требуется оценивать функцию правдоподобия Фишера порядка N^n раз. Эти выводы согласуются с результатами, полученными в [8], где показано, что применение методов максимального правдоподобия позволяет построить эффективные решающие правила для оценки моментов разладки, но приводят к алгоритмам очень высокой сложности, трудно реализуемым на практике. Для решения этой проблемы в [9] предлагается непараметрический метод обнаружения многократных нарушений стационарности, основанный на идее многократного разрезания и проверке однородности отдельных частей общей выборки. Этот метод использует достаточно сложную двумерную статистику, что затрудняет его применение. В работе [10] предложен алгоритм апостериорного обнаружения многократных разладок случайной последовательности по изменению моментной функции первого порядка. В нем используется одномерная статистика, имеющая смысл разности оцениваемых математических ожиданий случайного процесса на отрезках различной длины.

Идея описываемого ниже метода состоит в предварительной локализации участков, содержащих «нестационарные» точки, т.е. точки, в которых происходит изменение характеристик случайной последовательности. Кратко суть метода состоит в следующем. На множестве независимых наблюдений вводится дискриминантная функция, позволяющая судить о стационарности случайной последовательности в произвольной точке x . По априорным сведениям о возможной длине и отличительных свойствах параметризованного участка делаются правдоподобные предположения о поведении этой функции и вводится понятие нестационарных точек. После этого определение таких участков сводится к задаче кластеризации нестационарных точек и определению точек разладки.

Дискриминантная функция для случая кусочно-параметризованной последовательности

Как отмечалось выше, распределение замеров в условиях лазерного сканирования характеризуется нерав-

номерностью (см. рис. 2). Это усложняет использование статистических методов оценивания параметров случайной последовательности. Вместе с тем, в данном случае можно избежать указанной неравномерности, если воспользоваться равномерностью распределения точек по угловой мере. Для этого целесообразно перейти к полярной системе координат, связанной с точкой подвеса лазерного дальномера. На рис. 3 показано соотношение декартовой и полярной системы координат.

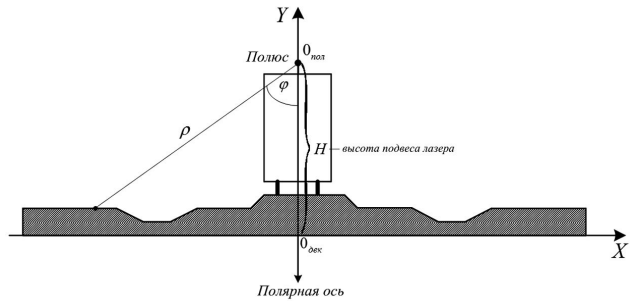


Рис. 3. Полярная система координат, связанная с лазерным дальномером

Из рисунка видно, что формулы, связывающие координаты (x, y) некоторой точки в декартовой и (p, φ) в полярной системах координат, имеют вид: $x = p \sin \varphi$, $y = H - p \cos \varphi$, где H – высота подвеса лазера, т.е. расстояние между полюсом полярной и началом декартовой системы координат.

На рис. 4 видно, как кусочно-линейная функция, задающая некоторый идеальный профиль и представленная в декартовой системе координат (слева), выглядит в полярной системе (справа). На правом графике по горизонтальной оси отложены значения полярного угла, а по вертикальной – расстояние от точки до полюса. Далее будем предполагать, что все статистические данные представлены именно в полярной системе координат.

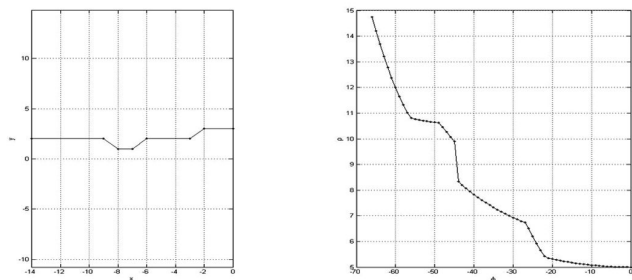


Рис. 4. Кусочно-линейная функция, описывающая «идеальный» профиль, в декартовой (слева) и полярной (справа) системах координат

Поскольку в соответствии с существующими нормативами все регламентированные профили земляного полотна являются непрерывными кусочно-линейными функциями, то при переходе к полярной системе координат они будут представлять собой непрерывную кусочно-параметризованную кривую, состоящую из кусков прямых линий, заданных в полярной системе:

$$\rho = \frac{r_m}{\cos(\varphi - \psi_m)}, \quad m = 1, 2, \dots, n, \quad \text{где } r_m \text{ – расстояние от}$$

прямой, описывающей m -й участок, до полюса, а ψ_m – угол, который образует нормаль к прямой с полярной осью. Поэтому наблюдаемый профиль в поляр-

ной системе координат будет реализацией (p_1, p_2, \dots, p_N) кусочно-параметризованной последовательности случайных величин $(\eta_1, \eta_2, \dots, \eta_N)$, которая на m -м участке описывается трендом вида $E(\eta_i) = \frac{r_m}{\cos(\varphi_i - \psi_m)}$. Построим дискриминантную функцию для анализа таких случайных последовательностей в полярной системе координат.

Основное требование, которое предъявляется к дискриминантной функции, состоит в том, что она должна принимать максимальные значения в областях изменения свойств последовательности независимых случайных величин η_i и минимальные – там, где эти свойства относительно неизменны или подчинены регулярной зависимости, т.е. внутри параметризованных участков. Для построения дискриминантной функции потребуется условие минимально возможной «длины» произвольного параметризованного участка. Будем предполагать, что количество точек-замеров на каждом параметризованном участке не меньше некоторого значения $2S$, где S – натуральное число, т.е. для любого $i \in \{1, 2, \dots, n\}$ выполняется неравенство $|\{\varphi_i | \chi_i \leq \varphi_i < \chi_{i+1}\}| \geq 2S$. Это условие представляется менее жестким, чем априорно известное число стационарных участков, правдоподобные предположения о котором в реальных задачах сделать достаточно трудно.

Рассмотрим последовательность независимых случайных величин $(\eta_1, \eta_2, \dots, \eta_N)$ с функциями плотности $f_i(\rho)$ и математическими ожиданиями, подчиняющимися закону $E(\eta_i) = g(\varphi_i, r_m, \psi_m) = \frac{r_m}{\cos(\varphi_i - \psi_m)}$ при $\chi_i \leq \varphi_i \leq \chi_{i+1}$ с неизвестными и подлежащими определению параметрами ρ_m и ψ_m . Пусть в точке φ_i происходит изменение статистических свойств этой последовательности, т.е.

$$f_i(\rho) = \begin{cases} f(\rho, g(\varphi, r_1, \psi_1), a_1), & i > t, \\ f(\rho, g(\varphi, r_2, \psi_2), a_2), & i > t. \end{cases}$$

Локализуем момент разладки по реализации случайной последовательности $(\rho_1, \rho_2, \dots, \rho_N)$ с помощью дискриминантной функции, строящейся по логарифмической функции правдоподобия:

$$L(k) = \max_{r_1, \psi_1, a_1} \sum_{i=k-s+1}^k \ln f(\rho_i, g(\varphi_i, r_1, \psi_1), a_1) + \max_{r_2, \psi_2, a_2} \sum_{i=k+1}^{k+s} \ln f(\rho_i, g(\varphi_i, r_2, \psi_2), a_2) + \max_{r_3, \psi_3, a_3} \sum_{i=k-s+1}^{k+s} \ln f(\rho_i, g(\varphi_i, r_3, \psi_3), a_3).$$

Данную функцию можно интерпретировать как логарифм отношения правдоподобия двух конкурирующих гипотез: H_1 – в точке φ_k произошла разладка случайного процесса (первое и второе слагаемое), против гипотезы H_2 – отрезок $[\varphi_{k-s+1}, \varphi_{k+s}]$ принадлежит кусочно-параметризованному участку (третье слагаемое). При этом параметры $r_1, \psi_1, a_1, r_2, \psi_2, a_2, r_3, \psi_3, a_3$ слу-

чайной последовательности на отрезках $[k-s+1, k], [k+1, k+s], [k-s, k+s]$ оцениваются по методу максимального правдоподобия. Можно предположить, что построенная таким образом дискриминантная функция будет обладать достаточно хорошими статистическими свойствами, например, можно ожидать, что момент разладки вероятнее всего произошел в тех точках случайной последовательности, в которых функция $L(k)$ принимает относительно большие значения, и наоборот – значения этой функции будут сравнительно невысоки внутри стационарных участков.

Для нормально распределенных случайных величин $\eta_1, \eta_2, \dots, \eta_N$ дискриминантная функция принимает вид:

$$L(k) = \max_{r_1, \psi_1, \sigma_1} \sum_{i=k-s+1}^k \ln \left\{ \frac{1}{\sqrt{2\pi}\sigma_1} \exp \left[-\frac{1}{2\sigma_1^2} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)} \right)^2 \right] \right\} + \max_{r_2, \psi_2, \sigma_2} \sum_{i=k+1}^{k+s} \ln \left\{ \frac{1}{\sqrt{2\pi}\sigma_2} \exp \left[-\frac{1}{2\sigma_2^2} \left(\rho_i - \frac{r_2}{\cos(\varphi_i - \psi_2)} \right)^2 \right] \right\} - \max_{r_3, \psi_3, \sigma_3} \sum_{i=k-s+1}^{k+s} \ln \left\{ \frac{1}{\sqrt{2\pi}\sigma_3} \exp \left[-\frac{1}{2\sigma_3^2} \left(\rho_i - \frac{r_3}{\cos(\varphi_i - \psi_3)} \right)^2 \right] \right\}.$$

Для построения дискриминантной функции необходимо, таким образом, решить три экстремальные задачи относительно неизвестных параметров $r_j, \psi_j, \sigma_j, j = 1, 2, 3$, соответственно. Рассмотрим решение первой из них. Приравнивая частные производные по переменным r_1, ψ_1, σ_1 к нулю, после эквивалентных преобразований получаем

$$\begin{cases} \sum_{i=k-s+1}^k \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)} \right) \frac{1}{\cos(\varphi_i - \psi_1)} = 0, \\ \sum_{i=k-s+1}^k \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)} \right) \frac{\sin(\varphi_i - \psi_1)}{\cos^2(\varphi_i - \psi_1)} = 0, \\ \sum_{i=k-s+1}^k \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)} \right)^2 = s\sigma_1^2, \end{cases}$$

откуда следуют выражения:

$$\sigma_1^2 = \frac{1}{s} \sum_{i=k-s+1}^{k+s} \left(\rho_i - \frac{r_1}{\cos(\varphi_i - \psi_1)} \right)^2, \\ r_1 = \frac{\sum_{i=k-s+1}^k \frac{\rho_i}{\cos(\varphi_i - \psi_1)}}{\sum_{i=k-s+1}^k \frac{1}{\cos^2(\varphi_i - \psi_1)}}, \\ \sum_{i=k-s+1}^k \sum_{j=k-s+1}^k \frac{\rho_i \sin(\varphi_i - \varphi_j)}{\cos^2(\varphi_i - \psi_1) \cos^3(\varphi_j - \psi_1)} = 0.$$

Из данных равенств находим $r_1^*, \psi_1^*, \sigma_1^*$. Последнее уравнение можно решить относительно ψ_1 каким-либо итерационным методом². Можно показать, что на отрез-

² Например, данное уравнение можно решить в MatLab с помощью процедуры fzero. Для значения $s = 10$ время отыскания корня ψ_1^* с точностью 10^{-4} равно примерно 0,009 с.

ке $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ хорошим начальным приближением будет

$$\psi_1^{(0)} = \arcsin \frac{\operatorname{sgn} b}{\sqrt{1+W^2}}, \text{ где}$$

$$W = \frac{\sum_{i=k-s+1}^k \rho_i \cos \varphi_i \left(s \rho_i \sin \varphi_i - \sum_{i=k-s+1}^k \rho_i \sin \varphi_i \right)}{\sum_{i=k-s+1}^k \rho_i \cos \varphi_i \left(s \rho_i \cos \varphi_i - \sum_{i=k-s+1}^k \rho_i \cos \varphi_i \right)},$$

$$b = \sum_{i=k-s+1}^k \rho_i (\sin \varphi_i - W \cos \varphi_i).$$

Аналогично находятся решения $r_2, \psi_2, \sigma_2, r_3, \psi_3, \sigma_3$ второй и третьей задач:

$$\sigma_2^2 = \frac{1}{s} \sum_{i=k+1}^{k+s} \left(\rho_i - \frac{r_2}{\cos(\varphi_i - \psi_2)} \right)^2,$$

$$r_2 = \frac{\sum_{i=k+1}^{k+s} \frac{\rho_i}{\cos(\varphi_i - \psi_2)}}{\sum_{i=k+1}^{k+s} \frac{1}{\cos^2(\varphi_i - \psi_2)}},$$

$$\sum_{i,j=k+1}^{k+s} \frac{\rho_i \sin(\varphi_i - \varphi_j)}{\cos^2(\varphi_i - \psi_2) \cos^3(\varphi_j - \psi_2)} = 0,$$

$$\sigma_3^2 = \frac{1}{2s} \sum_{i=k-s+1}^{k+s} \left(\rho_i - \frac{r_3}{\cos(\varphi_i - \psi_3)} \right)^2$$

$$r_3 = \frac{\sum_{i=k-s+1}^{k+s} \frac{\rho_i}{\cos(\varphi_i - \psi_3)}}{\sum_{i=k-s+1}^{k+s} \frac{1}{\cos^2(\varphi_i - \psi_3)}}$$

$$\sum_{i,j=k-s+1}^{k+s} \frac{\rho_i \sin(\varphi_i - \varphi_j)}{\cos^2(\varphi_i - \psi_3) \cos^3(\varphi_j - \psi_3)} = 0$$

После подстановки найденных параметров в выражение для дискриминантной функции и соответствующих преобразований получим окончательную формулу для вычисления $L(k)$:

$$L(k) = s \ln \frac{\sigma_3^2}{\sigma_1 \sigma_2}.$$

Можно ожидать, что вблизи точек переключения значения функции $L(k)$ будут большими, чем внутри параметризованного участка. Выберем величину порога h так, чтобы можно было отслеживать поведение дискриминантной функции при изменении параметра k по истинности неравенства $L(k) \geq h$. При этом должно наблюдаться скопление точек $NS = \{\varphi_k \mid L(k) \geq h\}$ вблизи точек φ_i , в которых происходит переход с одного параметризованного участка на другой, и напротив – разреженность таких точек внутри параметризованного участка. Будем для простоты называть такие точки нестационарными. Поэтому возникает проблема кластеризации нестационарных точек на нестационарные участки $LN_j, j = 1, 2, \dots, n$, в которых по определению наблюдается лишь один переход с одного параметризованного участка на другой.

Определение параметризованных участков с помощью дискриминантной функции

Нестационарные участки $LN_j, j = 1, 2, \dots, n$, задают разбиение множества NS . Обозначим через θ отношение эквивалентности, связанное с этим разбиением. Определим условие принадлежности двух произвольных точек $\varphi_{i1}, \varphi_{i2} \in NS, \varphi_{i1} < \varphi_{i2}$, одному нестационарному участку. Для этого рассмотрим функцию $E[\varphi_{i1}, \varphi_{i2}] = \frac{1}{i_2 - i_1 - 1} \sum_{j=i_1}^{i_2} L(j)$, которая, очевидно, дает среднее значение функции L на отрезке $[\varphi_{i1}, \varphi_{i2}]$.

Внутри нестационарного участка наблюдается сгущение нестационарных точек, поэтому следует ожидать, что значение $E[\varphi_{i1}, \varphi_{i2}]$, при условии $(\varphi_{i1}, \varphi_{i2}) \in \theta$, будет больше выбранного порога. Если же точки φ_{i1} и φ_{i2} принадлежат различным нестационарным участкам, т.е. $(\varphi_{i1}, \varphi_{i2}) \notin \theta$, то отрезок $[i_1, i_2]$, вероятно, включает интервал, в точках которого значения функции L меньше h . Следуя этим рассуждениям, получаем следующее условие принадлежности произвольных $\varphi_{i1}, \varphi_{i2} \in NS$ одному нестационарному участку: $(\varphi_{i1}, \varphi_{i2}) \in \theta$, когда $E[\varphi_{i1}, \varphi_{i2}] \geq h$. Это условие, очевидно, определяет толерантное, но не транзитивное отношение $\tilde{\theta} \subseteq \theta$.

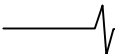
Введем дискретную функцию $\pi(\varphi_i)$, характеризующую плотность распределения нестационарных точек в окрестности точки φ_i , необходимую при описании нестационарных участков. Пусть $[\varphi_{i1}, \varphi_{i2}]$ – произвольный отрезок на $[\varphi_1, \varphi_N]$. Тогда

$$\pi(\varphi_i) = \max_{\substack{[\varphi_{i1}, \varphi_{i2}] \\ \varphi_i \in [\varphi_{i1}, \varphi_{i2}]}} E[\varphi_{i1}, \varphi_{i2}],$$

где максимум берется по всем отрезкам $[\varphi_{i1}, \varphi_{i2}]$, содержащим точку φ_i . В основе конструктивного метода определения нестационарных участков лежит следующая теорема, доказанная в [11].

Теорема. Пусть $\varphi_i, \varphi_m \in NS, \varphi_m > \varphi_i$. Тогда $(\varphi_i, \varphi_m) \in \theta$ в том и только том случае, если для любого $\varphi_k \in [\varphi_i, \varphi_m]$ имеет место неравенство $\pi(\varphi_k) \geq h$.

Покажем, как функцию $\pi(\varphi_j)$ можно использовать для определения нестационарных участков. Пусть нестационарные точки φ_i, φ_m принадлежат некоторому нестационарному участку LN_j , т.е. $(\varphi_i, \varphi_m) \in \theta$, тогда он включает отрезок, т.е. $[\varphi_i, \varphi_m] \subseteq LN_j$. По теореме функция π не меньше пороговой величины h в каждой точке этого отрезка, т.е. для любого $\varphi_k \in [\varphi_i, \varphi_m]$ $\pi(\varphi_k) \geq h$. Поэтому будем считать, что точка $\varphi_k \in [\varphi_i, \varphi_m]$ принадлежит некоторому нестационарному участку тогда и только тогда, когда $\pi(\varphi_k) \geq h$; точки $\varphi_i, \varphi_m \in [\varphi_1, \varphi_N]$ при $\pi(\varphi_i) \geq h, \pi(\varphi_m) \geq h, \varphi_i < \varphi_m$, принадлежат одному нестационарному участку в том и только том случае, если для любого $\varphi_k \in [\varphi_i, \varphi_m]$ $\pi(\varphi_k) \geq h$. Это определение нестационарного участка удовлетворяет всем необходимым условиям. Легко проверяется, что каждая нестационарная точка принадле-



жит некоторому нестационарному участку, причем каждый такой участок имеет непустое пересечение с множеством нестационарных точек, т.е. для любого $j \in [2, n]$ $LN_j \cap NS \neq \emptyset$, и определенное таким образом множество $\{LN_j\}$ индуцирует отношение эквивалентности θ на множестве нестационарных точек NS .

Оценка моментов нарушения стационарности

Оценка положения точек переключения случайной последовательности с одного параметризованного участка на другой существенно упрощается, так как эти моменты локализованы внутри нестационарных участков, и можно перейти к однопараметрическим задачам определения точек переключения внутри сегментов, каждый из которых включает лишь один нестационарный участок. При этом могут оказаться полезными предварительные оценки положения точек переключения, учитывающие поведение дискриминантной функции L . В частности, для точки переключения $\hat{\varphi}_k$ можно использовать первоначальное приближение вида

$$\hat{\varphi}_k = \frac{\sum_{\varphi_i \in LN_k} \varphi_i L(i)}{\sum_{\varphi_i \in LN_k} L(i)}$$

В качестве $\hat{\varphi}_k$ в последней формуле берется ближайшее целое значение.

Данный метод обнаружения точек, в которых меняются свойства случайных величин, был апробирован на данных моделирования зашумленного профиля. При этом оказалось, что последняя оценка дает достаточно точное приближение и не возникает необходимости ее улучшения каким-либо другим способом.

Пример численного моделирования

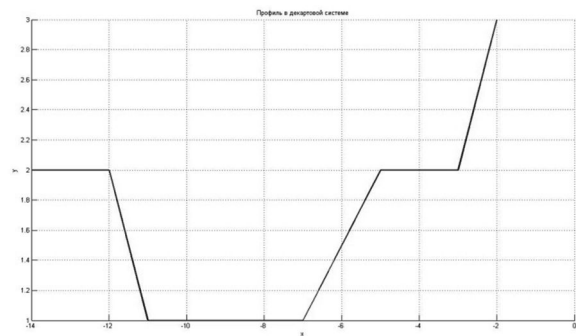
Рассмотрим численный пример обнаружения непараметрических участков в статистической последовательности. На рис. 5, а) показана форма идеального профиля, на рис. 5, б) - результаты численного моделирования случайной последовательности в полярной системе координат, на каждом параметризованном участке которой случайная последовательность описывается нормальным вероятностным распределением.

По результатам численного моделирования была построена дискриминантная функция и функция плотности нестационарных точек (рис. 5, в) и рис. 5, г)). Обратим внимание на то, что локальный метод обнаружения непараметрических участков в данном случае позволил осуществить правильную локализацию точек переключения случайной последовательности, если использовать порог $h = 22$. Восстановление геометрической формы профиля производилось по методу наименьших квадратов в декартовой системе координат. Восстановленный профиль изображен на рис. 5, д).

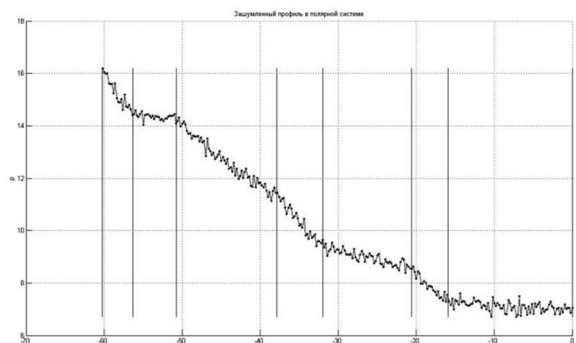
Заключение

В работе показано, что автоматизация процедуры мониторинга поперечного профиля земляного полотна может быть сведена к задачам анализа случайной последовательности в различных постановках. Полезные результаты при этом открываются при использовании подходов, основанных на применении функций правдо-

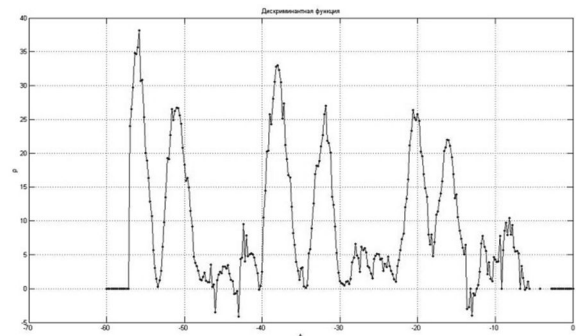
подобия. В частности, на этом пути оказывается возможным решение задачи восстановления структурных линий, о чем упоминалось в начале статьи. Использование статистических методов накладывает и ряд ограничений, за рамками которых трудно рассчитывать на приемлемое по качеству решение. К числу таких ограничений следует отнести, прежде всего, требование достаточно представительной статистической выборки замеров. Методы вряд ли будут хорошо работать в условиях разрушений или сильных деформаций земляного полотна, когда не представляется возможным сделать правдоподобные предположения о форме профиля.



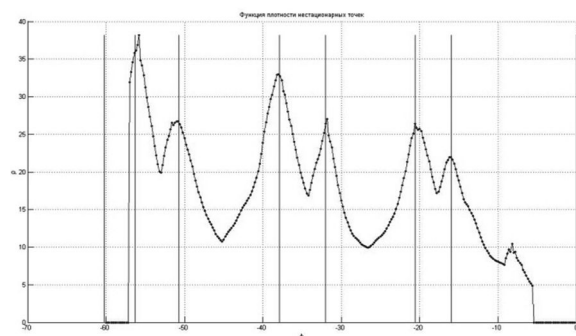
а)



б)



в)



г)

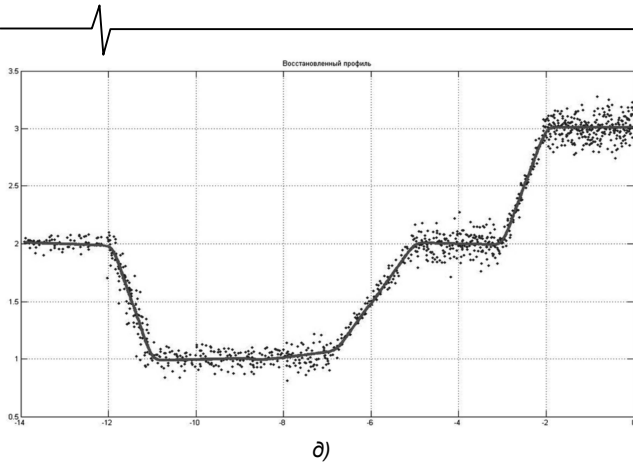


Рис. 5. Результаты моделирования:

- а) исходный идеальный профиль в декартовой системе;
 б) зашумленный профиль в полярной системе координат;
 в) дискриминантная функция;
 г) функция плотности нестационарных точек;
 д) восстановленный профиль

Остается открытым вопрос об эффекте лазерных «теней», когда на профиле имеются места с крутизной, не позволяющей осуществить лазерное сканирование в принципе. Но это проблема самой технологии съема информации. По-видимому, в этом случае для восстановления профиля в «затененных» местах следует разрабатывать методы, основанные не только на чисто математических подходах, что могло бы стать темой отдельного исследования.

Литература

1. Сазонов В., Кармалин Б., Лебедев А. и др. Современные направления диагностики и мониторинга земляного полотна. Путь и путевое хозяйство. 2009, 6, с.34-37.
2. Уманский В.И. Технология построения трехмерных моделей железнодорожного полотна в высокоточном координатном пространстве. Сборник докладов 6-й Международной научно-практической конференции «Геопространственные технологии и сферы их применения». М., 2010.
3. Бриллинджер Д. Временные ряды. Обработка данных и теория. М.: Мир, 1980.
4. Гихман И.И., Скороход А.В. Введение в теорию случайных процессов. М.: Наука, 1965.
5. Королюк В.С., Портенко Н.И., Скороход А.В., Тур-

бин А.Ф. Справочник по теории вероятностей и математической статистике. М.: Наука, 1985.

6. Клигене Н., Телькснис Л. Методы обнаружения моментов изменения свойств случайных процессов (обзор). Автоматика и телемеханика. 1983, № 10.

7. Торговицкий И.Ш. Методы определения момента изменения вероятностных характеристик случайных величин. Зарубежная радиоэлектроника. 1976, №1, с. 3-52.

8. Харин Ю.С. Классификация случайных серий неизвестной длины. Проблемы передачи информации. 1985, т. XXI, вып. 4, с. 64-65.

9. Дарховский Б.С. Непараметрический метод оценивания интервалов однородности случайной последовательности. Теория вероятностей и ее применение. 1985, т. 30, №4, с. 795-799.

10. Бродский Б.Е., Дарховский Б.С. Алгоритм апостериорного обнаружения разладок случайной последовательности. Автоматика и телемеханика. 1993, №1, с. 62-67.

11. Броневиц А.Г., Каркищенко А.Н. Вероятностные и возможные методы классификации случайных последовательностей. Таганрог: ТРТУ, 1996.

STATISTICAL METHOD OF RESTORING SUBGRADE FROM DATA OF LASER SCANNING

*Bronevich A.G., Karkishchenko A.N.,
Umanskiy V.I.*

The article is devoted to restoring the cross-section of subgrade from noisy data of laser scanning. A formal statement of this problem is given. It is shown that the problem can be reduced to the analysis of a random sequence of measurements. The paper presents a solution by using the Fisher discriminant function. It is also shown how to localize switching points in a piecewise parameterized random sequence with the help of the constructed discriminant function. A computationally efficient way of estimating such points is proposed. Finally an example on simulated data is given.



10-ый Симпозиум по проектированию и тестированию

IEEE EAST-WEST DESIGN
& TEST SYMPOSIUM - EWDTIS 2012

пройдет с 14 по 17 сентября 2012 года
в Харькове (Украина)

Подробнее: <http://ewdtest.com/conf>