

References

1. Finkel'shtein M.I. *Radar Basics*. Moscow: Radio i svjaz', 1983, 536 p. (In Russ.)
2. Varakin L.E. *Communication Systems with Noise-Like Signals*. Moscow: Radio i svjaz', 1985, 384 p. (In Russ.)
3. Arkhangel'skiy V.B. et al. Optimal algorithm for generating and processing noise-like signals in optical reflectometry. *Obrabotka signalov v sistemah svyazi: sb. nauch. tr. uchebn. zaved. svyazi*, 1996, pp. 36–39. (In Russ.)
4. Listvin A.V., Listvin V.N. *Optical Fiber Reflectometry*. Moscow: LESARart, 2005, 208 p. (In Russ.)
5. Bylina M.S. et al. *Measurement of Parameters of Fiber-Optic Line Paths*. A Tutorial. Saint Petersburg: SPbGUT, 2002, 80 p. (In Russ.)
6. Andreev V.A. et al. *Backscattering Fiber Optic Measurements*. A Tutorial. Samara: SRTTs PGATI, 2001, 121 p. (In Russ.)
7. Solonina A.I. *Fundamentals of Digital Signal Processing*. Saint Petersburg: BHV-Peterburg, 2005, 768 p. (In Russ.)
8. Ivanov A.B. *Fiber Optics. Components, Transmission Systems, Measurements*. Moscow: SYRUS SYSTEMS, 1999, 671 p. (In Russ.)
9. Gauer J. *Optical Transmission Systems*. English trans. Moscow: Radio i svjaz', 1989, 501 p. (In Russ.)
10. Anderson D.R., Johnson L.M., Bell F.G. *Troubleshooting Optical Fiber Networks: Understanding and Using Optical Time*. Amsterdam: Elsevier Academic Press, 2004, 437 p.
11. Baklanov I.G. *Testing and Diagnostics of Communication Systems*. Moscow: Eko-Trendz, 2001, 264 p. (In Russ.)

Received 16.06.2021

УДК 621.39

АНАЛОГО-ЦИФРОВОЙ НАКОПИТЕЛЬ С КОЛЬЦЕВЫМ РЕГИСТРОМ ПАМЯТИ

Архангельский В.Б.¹, Глаголев С.Ф.², Хричков В.А.²

¹ ООО «Оптические технологии», Всеволожск, РФ

² Санкт-Петербургский государственный университет телекоммуникаций

имени профессора М.А. Бонч-Бруевича, Санкт-Петербург, РФ

E-mail: v.b.arh@mail.ru, glagolevsf@yandex.ru, hrichkovv@gmail.com

В данной работе рассматривается необходимость фрагментарной регистрации рефлектограммы для увеличения динамического диапазона рефлектометра. Для обеспечения высокой точности оцифровки сигнала предлагается разбить линию связи на несколько фрагментов и регистрировать сигнал обратного рассеяния каждого фрагмента с разными коэффициентами усиления, изменяя их с помощью масштабных усилителей. Это решение позволит поддерживать амплитуду сигнала на входе аналого-цифрового преобразователя, близкую к номинальному значению, и тем самым снизить погрешности преобразования сигнала. После оцифровки сигналы каждого фрагмента записываются в свой кольцевой регистр памяти для последующего проведения их корреляционной обработки. В предложенном способе для получения полной рефлектограммы необходимо сложить автокорреляционные функции всех фрагментов. Представленные диаграммы работы рефлектометра показывают эффективность предложенного способа.

Ключевые слова: оптическое волокно, оптическая рефлектометрия, сигнал обратного рассеяния, сложный зондирующий сигнал, корреляционный рефлектометр, кольцевой регистр памяти

Введение

В корреляционных рефлектометрах, зондирующие сигналы которых представляют собой фрагменты М-последовательностей (рисунок 1, а) с изменяющейся начальной фазой, длительность импульсов τ определяет разрешающую способность рефлектометра, а длительность М-последовательности — длину исследуемого волоконно-оптического тракта.

Сигналы обратного рассеяния с выхода усилителя фототока оцифровываются аналого-цифровым преобразователем (АЦП) и накапливаются в кольцевом регистре памяти. Число ячеек памяти регистра равно числу символов периода

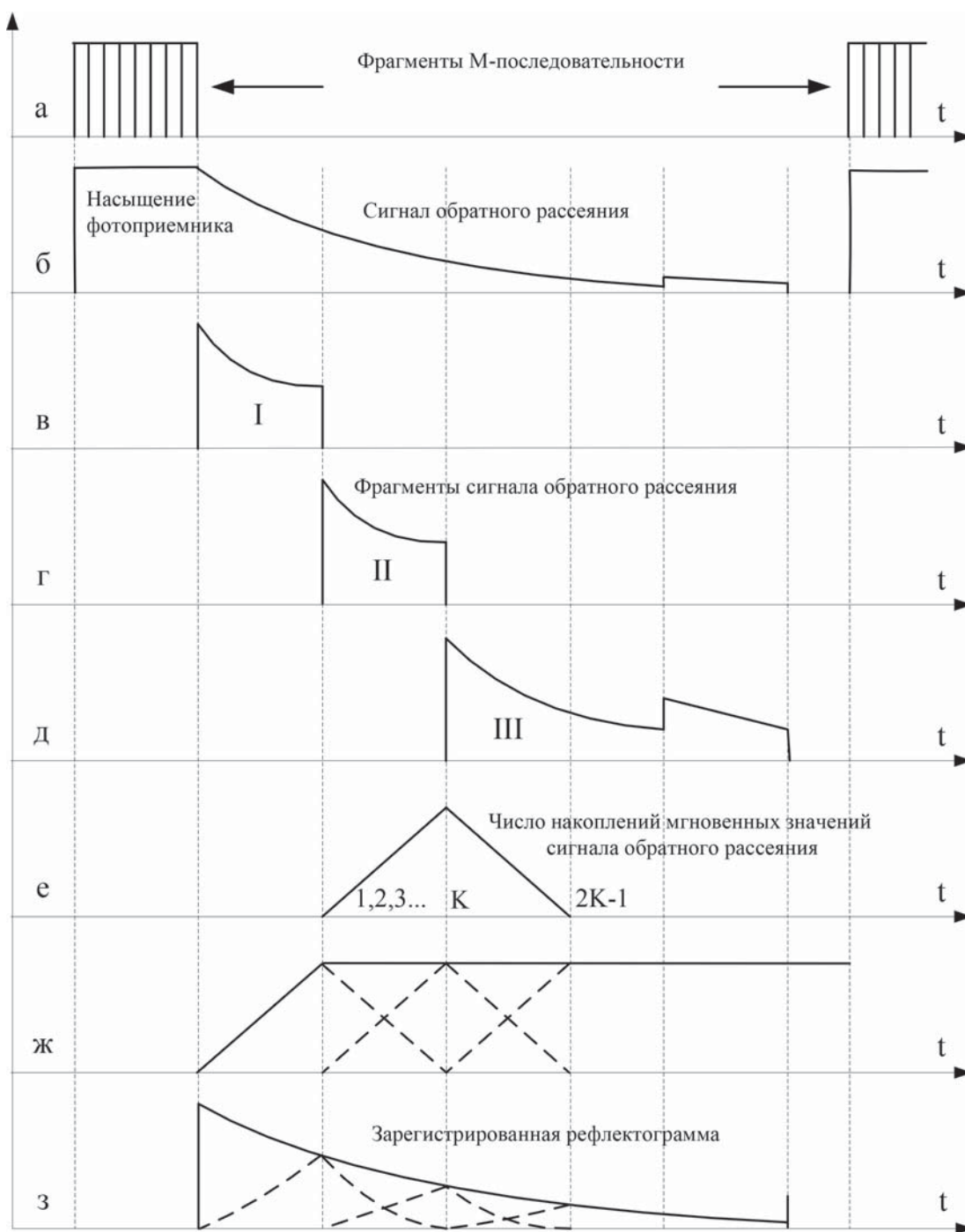


Рисунок. Диаграммы работы корреляционного рефлектометра

М-последовательности и должно быть больше суммы числа элементов регистрируемой рефлектограммы R и числа импульсов в фрагменте зондирующего сигнала K :

$$M = (2^\gamma - 1) \geq (R + K),$$

где γ – любое целое число.

Фрагментарная регистрация сигнала обратного рассеяния

Амплитуда сигнала обратного рассеяния волоконно-оптического тракта затухает во времени

по экспоненциальному закону (рисунок, б), и затухание может достигать 100 дБ. Динамический диапазон АЦП значительно меньше, порядка 30 дБ, поэтому для регистрации сигнала обратного рассеяния во всем диапазоне затуханий его необходимо усиливать. Если сигнал обратного рассеяния разделить на ряд фрагментов T_F одинаковой длительности, то затухания сигналов в каждом фрагменте будут примерно одинаковыми. Эти затухания могут быть компенсированы линейкой масштабных усилителей с одинаковыми

коэффициентами усиления, равными среднему затуханию сигнала во фрагментах (рисунок, *в-д*).

Во время первого фрагмента на АЦП поступает сигнал обратного рассеяния непосредственно с выхода усилителя фототока и оцифрованные сигналы накапливаются в первом регистре памяти. Во время второго фрагмента на АЦП поступает сигнал обратного рассеяния с выхода первого масштабного усилителя и оцифрованные сигналы накапливаются во втором регистре памяти. Во время третьего фрагмента на АЦП поступает сигнал обратного рассеяния с выхода второго масштабного усилителя и оцифрованные сигналы накапливаются в третьем регистре памяти и т. д. Таким образом, на АЦП поступают сигналы с различных масштабных усилителей приблизительно равной амплитудой, и в каждом регистре памяти накапливается только один фрагмент сигнала обратного рассеяния.

Для обеспечения максимального отношения сигнал – шум длительность фрагмента M -последовательности зондирующего сигнала выбирается равной длительности фрагмента сигнала обратного рассеяния, накапливаемого в каждом регистре. Так как на выходе последнего масштабного усилителя шум сигнала обратного рассеяния максимален и его среднее квадратичное значение соизмеримо с номинальным значением входного напряжения АЦП, то длительность фрагмента сигнала обратного рассеяния, накапливаемого в последнем регистре памяти, больше длительности остальных фрагментов и заканчивается только с началом следующего зондирующего сигнала. Число ячеек в каждом регистре памяти одинаковое, равное числу импульсов M -последовательности.

Запись значений сигнала обратного рассеяния в кольцевые регистры памяти

После генерации первого фрагмента зондирующего сигнала мгновенные значения сигнала обратного рассеяния заносятся в каждом регистре памяти в ячейки с номерами 1, 2, ..., K , а в последнем регистре памяти мгновенные значения заносятся в большее число ячеек с номерами 1, 2, ..., $(M - VK)$, то есть до начала генерации второго фрагмента зондирующего сигнала. После его генерации мгновенные значения заносятся в каждом регистре в ячейки с номерами 2, 3, ..., $(K + 1)$, а в последнем регистре – с номерами 2, 3, ..., $(M - VK + 1)$, то есть до начала генерации третьего фрагмента зондирующего сигнала.

После его генерации мгновенные значения заносятся в каждом регистре в ячейки с номерами 3, 4, ..., $(K + 2)$, а в последнем регистре – с номерами 3, 4, ..., $(M - VK + 2)$ и т. д. Последовательно генерируется M фрагментов последовательности, и мгновенные значения в каждой ячейке каждого регистра памяти суммируются с накопленными ранее, а при превышении номера ячейки значения M заносятся по кольцу в ячейки с номером 1, 2, ... и т. д.

После регистрации сигналов обратного рассеяния от всех M фрагментов зондирующего сигнала в ячейках каждого регистра памяти накапливаются мгновенные значения сигналов обратного рассеяния от ряда элементов волоконно-оптического тракта с различным числом накоплений. Так, например, во втором регистре памяти число накоплений от первого элемента фрагмента – 1, от второго – 2, от третьего – 3 и т. д., линейно увеличиваясь до K -го – K , от $(K + 1)$ -го – $(K - 1)$, от $(K + 2)$ -го – $(K - 2)$, от $(K + 3)$ -го – $(K - 3)$ и т. д., линейно уменьшаясь до $(2K - 1)$ -го – 1. На рисунке, *е* показана зависимость числа накоплений мгновенных значений сигнала обратного рассеяния от элементов волоконно-оптического тракта во втором регистре памяти. В последнем регистре памяти от K -го и последующих элементов число накоплений постоянно и равно K до начала следующего зондирующего сигнала. Если просуммировать число накоплений мгновенных значений сигнала обратного рассеяния во всех регистрах памяти (рисунок, *ж*), то число накоплений всех, за исключением первых K , элементов будет одинаково и равно K .

Корреляционная обработка сигнала обратного рассеяния

Для корреляционной обработки накопленных в регистрах памяти сигналов обратного рассеяния в качестве опорного сигнала используется исходная M -последовательность. В результате корреляционной обработки накопленных в регистрах памяти сигналов формируются фрагменты рефлектограммы, сумма которых и является полной рефлектограммой исследуемого волоконно-оптического тракта (рисунок, *з*). Для устранения искажений рефлектограммы в ее первом фрагменте используется множитель:

$$x(r) = z(r)K/r,$$

где $x(r)$ – рефлектограмма как функция от расстояния r ; $z(r)$ – зарегистрированная рефлектограмма с искажениями.

Погрешности при измерениях

При измерении мгновенных значений сигнала обратного рассеяния АЦП возникают составляющие погрешности преобразования, обусловленные шумами фотоприемника, неточностью установки коэффициента передачи масштабных усилителей и нелинейностью передаточной функции преобразователя. Составляющая погрешности из-за шумов фотоприемника уменьшается при увеличении числа импульсов в фрагменте зондирующего сигнала и увеличении времени регистрации рефлектограммы, составляющая погрешности из-за неточности установки коэффициента передачи масштабных усилителей уменьшается калибровкой рефлектометра, а составляющая погрешности от нелинейности преобразователя носит случайный характер и зависит от величины мгновенного значения оцифрованного сигнала обратного рассеяния.

Сигнал обратного рассеяния волоконно-оптического тракта затухает во времени по экспоненциальному закону:

$$y(t) = U \exp(-\alpha ct/n),$$

где U – максимальное значение сигнала обратного рассеяния в начале тракта; α – коэффициент затухания оптического волокна; c – скорость света в вакууме; n – коэффициент преломления сердцевины волокна.

Сигнал обратного рассеяния в V -фрагменте волоконного тракта:

$$y_V(t) = U_V \exp(-\alpha ct/n).$$

АЦП имеет Q положительных и Q отрицательных единиц младшего разряда Δ , поэтому сигнал обратного рассеяния измеряется в целых числах единиц этого разряда $y_V(t) = i\Delta$ с погрешностью преобразования, не превышающей половины единицы младшего разряда $\varepsilon \leq \Delta/2$. Преобразователь имеет динамический диапазон $D = 2Q/\Delta$ и последовательно формирует цифровые значения сигнала обратного рассеяния, которые изменяются в моменты времени перехода этого сигнала через границу, разделяющую младшие разряды.

Каждое цифровое значение сигнала обратного рассеяния на выходе АЦП существует некоторый интервал времени:

$$\begin{aligned} t(i) &= -\frac{n}{\alpha c} \left[\ln \left(i\Delta - \frac{\Delta}{2} \right) - \ln \left(i\Delta + \frac{\Delta}{2} \right) \right] = \\ &= \frac{n}{\alpha c} \ln \left(\frac{i+0,5}{i-0,5} \right). \end{aligned}$$

Цифровые значения сигнала обратного рассеяния через интервал времени, равный дли-

тельности импульсов зондирующего сигнала τ , заносятся и накапливаются в регистре памяти рефлектометра.

Длительность всего регистрируемого фрагмента:

$$T_F = \frac{n}{\alpha c} \sum_{i=Q_1}^{Q_2} \ln \left(\frac{i+0,5}{i-0,5} \right),$$

где Q_1 и Q_2 – минимальное и максимальное цифровое значение сигнала обратного рассеяния в V -м фрагменте.

Нормированный интервал времени каждого цифрового значения:

$$t_n(i) = \ln \left(\frac{i+0,5}{i-0,5} \right) / \sum_{i=Q_1}^{Q_2} \ln \left(\frac{i+0,5}{i-0,5} \right).$$

Так как зондирующие сигналы представляют собой фрагменты M -последовательностей, начальная фаза которых последовательно изменяется от 1 до M , а измеренные мгновенные значения сигнала обратного рассеяния заносятся в ячейки регистра памяти начиная с номера, который так же последовательно изменяется от 1 до M , то в каждой из M ячеек каждого регистра памяти накапливается сумма K – мгновенных значений сигналов обратного рассеяния, изменяющихся в пределах затухания сигнала в выбранном фрагменте:

$$\Xi_V = \frac{n}{\alpha c} \sum_{i=Q_1}^{Q_2} i \ln \left(\frac{i+0,5}{i-0,5} \right).$$

Погрешность измерения мгновенного значения сигнала обратного рассеяния $\varepsilon(i)$ не превышает половины единицы младшего разряда. Погрешность накопленной суммы мгновенных значений:

$$\xi_\Sigma = \frac{n}{\alpha c} \sum_{i=Q_1}^{Q_2} \varepsilon(i) \ln \left(\frac{i+0,5}{i-0,5} \right),$$

а среднее квадратичное значение этой погрешности:

$$\xi = \frac{\Delta n}{2\alpha c} \sqrt{\sum_{i=Q_1}^{Q_2} \ln^2 \left(\frac{i+0,5}{i-0,5} \right)},$$

где $\Delta/2$ – среднее квадратичное значение погрешности измерения мгновенного значения сигнала обратного рассеяния.

Средняя квадратичное значение погрешности измерения i -х значений сигнала обратного рассеяния:

$$\xi_i = \frac{\Delta}{2} \sqrt{\sum_{i=Q_1}^{Q_2} \ln^2 \left(\frac{i+0,5}{i-0,5} \right)} / \sum_{i=Q_1}^{Q_2} \ln \left(\frac{i+0,5}{i-0,5} \right).$$

Например, при $\alpha = 0,05$, $n = 1,5$, $Q_1 = 10$ и $Q_2 = 100$ эта погрешность составляет $\xi_i \approx$

$\approx 0,13(\Delta/2)$, то есть погрешность уменьшается приблизительно в восемь раз и во столько же раз расширяется динамический диапазон. Рассмотренное уменьшение погрешности АЦП справедливо только при малых шумах фотоприемника на его входе, при оцифровывании сигналов с выходов первых масштабных усилителей. Механизм уменьшения погрешности обусловлен ее усреднением в диапазоне изменений сигнала обратного рассеяния в каждом фрагменте.

Для обеспечения высокой точности оцифровки сигнала обратного рассеяния амплитуда выходного сигнала усилителя фототока устанавливается приближенно равной номинальному значению АЦП. Через интервал времени T_F вход АЦП переключается к выходу первого масштабного усилителя, который частично компенсирует затухание сигнала обратного рассеяния за интервал времени T_F . Через интервал времени $2T_F$ вход АЦП переключается к выходу второго масштабного усилителя, который частично компенсирует затухание сигнала обратного рассеяния за второй интервал времени T_F и т. д. Общее число масштабных усилителей ограничивается шумами на выходе последнего масштабного усилителя, которые не должны приводить к насыщению АЦП. Амплитуда сигнала обратного рассеяния должна быть меньше шума и не оказывать существенного влияния на насыщение АЦП.

На выходе последнего масштабного усилителя присутствуют значительные шумы фотоприемника, которые при накоплении уменьшают погрешность согласно:

$$\xi_i = \frac{\Delta}{2} \sqrt{\frac{\Delta}{2\sigma\sqrt{\pi}}},$$

где σ – среднее квадратичное значение шумов фотоприемника на выходе последнего масштабного усилителя.

Например, при $\sigma = 50\Delta$ эта погрешность составляет $\xi_i \approx 0,08(\Delta/2)$, то есть погрешность уменьшается приблизительно в двенадцать раз и во столько же раз расширяется динамический диапазон. Рассмотренное уменьшение погрешности АЦП справедливо только при больших шумах фотоприемника на выходе последнего масштабного усилителя. Механизм уменьшения погрешности обусловлен ее усреднением в диапазоне изменений шумового сигнала фотоприемника.

При преобразовании выходных сигналов остальных масштабных усилителей погрешность уменьшается за счет обоих усреднений.

Шумы фотоприемника последнего масштабного усилителя не зависят от параметров измеряемого волоконно-оптического тракта, а зависят только от коэффициентов усиления масштабных усилителей, их полосы пропускания и являются характеристикой рефлектометра. Амплитуда сигнала обратного рассеяния на выходе этого усилителя зависит от длительности фрагмента последовательности зондирующего сигнала, равного длительности фрагмента сигнала обратного рассеяния, и затухания волоконно-оптического тракта. Это позволяет, не изменяя уровня шумов и незначительно изменяя длительность фрагментов, изменять амплитуду сигнала обратного рассеяния на выходе последнего масштабного усилителя и устанавливать ее уровень меньше шумов фотоприемника.

Заключение

Рассмотренные выше положения показывают, что в корреляционных рефлектометрах, зондирующие сигналы которых представляют собой фрагменты М-последовательностей с изменяющейся начальной фазой, погрешности преобразования АЦП значительно меньше, чем собственная погрешность АЦП, а динамический диапазон шире.

Работа выполнена при финансовой поддержке Фонда содействия инновациям в рамках договора № 3855ГСИ/63276.

Литература

1. Финкельштейн М.И. Основы радиолокации. М.: Радио и связь, 1983. 536 с.
2. Варакин Л.Е. Системы связи с шумоподобными сигналами. М.: Радио и связь, 1985. 384 с.
3. Оптимальный алгоритм генерации и обработки шумоподобных сигналов в оптической рефлектометрии / В.Б. Архангельский [и др.] // Обработка сигналов в системах связи: сб. науч. тр. учебн. завед. связи. 1996. С. 36–39.
4. Листвин А.В., Листвин В.Н. Рефлектометрия оптических волокон. М.: ЛЕСАРпт, 2005. 208 с.
5. Измерение параметров волоконно-оптических линейных трактов: учеб. пособие / М.С. Былина [и др.]. СПб.: СПбГУТ, 2002. 80 с.
6. Измерения на ВОЛП методом обратного рассеяния: учебное пособие / В.А. Андреев [и др.]. Самара: СРТТЦ ПГАТИ, 2001. 121 с.

7. Солонина А.И. Основы цифровой обработки сигналов. СПб.: БХВ-Петербург, 2005. 768 с.
8. Иванов А.Б. Волоконная оптика. Компоненты, системы передачи, измерения. М.: SYRUS SYSTEMS, 1999. 671 с.
9. Гауэр Дж. Оптические системы передачи / пер. с англ. М.: Радио и связь, 1989. 501 с.
10. Anderson D.R., Johnson L.M., Bell F.G. Troubleshooting Optical Fiber Networks: Understanding and Using Optical Time. Amsterdam: Elsevier Academic Press, 2004. 437 p.
11. Бакланов И.Г. Тестирование и диагностика систем связи. М.: Эко-Трендз, 2001. 264 с.

Получено 16.06.2021

Архангельский Владимир Борисович, к.т.н., старший научный сотрудник ООО «Оптические технологии». 188643, Российская Федерация, г. Всеволожск, ул. Первомайская, 26а. Тел. +7 950 019-21-95. E-mail: v.b.arh@mail.ru

Глаголев Сергей Федорович, к.т.н., доцент кафедры фотоники и линий связи (ФилС) Санкт-Петербургского государственного университета телекоммуникаций имени М.А. Бонч-Бруевича (СПбГУТ). 193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, 22, корп. 1. Тел. +7 921 887-67-95. E-mail: glagolevsf@yandex.ru

Хричков Валентин Александрович, старший преподаватель кафедры ФилС СПбГУТ. 193232, Российская Федерация, г. Санкт-Петербург, пр. Большевиков, 22, корп. 1. Тел. +7 911 849-49-82. E-mail: hrichkovv@gmail.com

ANALOG-TO-DIGITAL STORAGE RING WITH MEMORY RING

Arkhangelsky V.B.¹, Glagolev S.F.², Khrichkov V.A.²

¹ *Optical Technologies LLC, Vsevolozhsk, Russian Federation*

² *St. Petersburg State University of Telecommunications*

named after Professor M.A. Bonch-Bruевич, Saint Petersburg, Russian Federation

E-mail: v.b.arh@mail.ru, glagolevsf@yandex.ru, hrichkovv@gmail.com

This paper discusses the need for fragmentary recording of the reflectogram to increase the dynamic range of the reflectometer. To ensure high accuracy of signal digitization, it is proposed to split the communication line into several fragments, and register the backscattered signal of each fragment with different gains by changing them using scale amplifiers. This solution will allow maintaining the signal amplitude at the input of the analog-to-digital converter close to the nominal value, and thereby reduce the signal conversion errors. After digitization, the signals of each fragment are written into its own circular memory register for subsequent correlation processing. In the proposed method, to obtain a complete reflectogram, it is necessary to add the autocorrelation functions of all fragments. The presented diagrams of the reflectometer operation show the effectiveness of the proposed method.

Keywords: *optical fiber, optical reflectometry, backscatter signal, complex probe signal, correlation reflectometer, ring memory register*

DOI: 10.18469/ikt.2021.19.3.06

Arkhangelsky Vladimir Borisovich, LLC «Optical technology», 26a, Pervomayskaya Street, Vsevolozhsk, 188643, Russian Federation; Senior Researcher, PhD in Technical Sciences. Tel. +7 950 019-21-95. E-mail: v.b.arh@mail.ru

Glagolev Sergey Fedorovich, The Bonch-Bruевич St. Petersburg State University of Telecommunications, 22, building 1, Bolshevikov Avenue, Saint Petersburg, 193232, Russian Federation; Associate Professor of Photonic and Telecommunication Lines Department, PhD in Technical Sciences. Tel.+7 921 887-67-95. E-mail: glagolevsf@yandex.ru

Khrichkov Valentin Alexandrovich, The Bonch-Bruевич St. Petersburg State University of Telecommunications, 22, building 1, Bolshevikov Avenue, Saint Petersburg, 193232, Russian Federation; Senior Lecturer of Photonic and Telecommunication Lines Department. Tel. +7 911 849-49-82. E-mail: hrichkovv@gmail.com

References

1. Finkel'shtein M.I. *Radar Basics*. Moscow: Radio i svjaz', 1983, 536 p. (In Russ.)
2. Varakin L.E. *Communication Systems with Noise-Like Signals*. Moscow: Radio i svjaz', 1985, 384 p. (In Russ.)
3. Arkhangel'skiy V.B. et al. Optimal algorithm for generating and processing noise-like signals in optical reflectometry. *Obrabotka signalov v sistemah svyazi: sb. nauch. tr. uchebn. zaved. svyazi*, 1996, pp. 36–39. (In Russ.)
4. Listvin A.V., Listvin V.N. *Optical Fiber Reflectometry*. Moscow: LESARart, 2005, 208 p. (In Russ.)
5. Bylina M.S. et al. *Measurement of Parameters of Fiber-Optic Line Paths*. A Tutorial. Saint Petersburg: SPbGUT, 2002, 80 p. (In Russ.)
6. Andreev V.A. et al. *Backscattering Fiber Optic Measurements*. A Tutorial. Samara: SRTTs PGATI, 2001, 121 p. (In Russ.)
7. Solonina A.I. *Fundamentals of Digital Signal Processing*. Saint Petersburg: BHV-Peterburg, 2005, 768 p. (In Russ.)
8. Ivanov A.B. *Fiber Optics. Components, Transmission Systems, Measurements*. Moscow: SYRUS SYSTEMS, 1999, 671 p. (In Russ.)
9. Gauer J. *Optical Transmission Systems*. English trans. Moscow: Radio i svjaz', 1989, 501 p. (In Russ.)

Received 16.06.2021

ТЕХНОЛОГИИ КОМПЬЮТЕРНЫХ СИСТЕМ И СЕТЕЙ

УДК 004.932

О ПРЕЦЕДЕНТНОЙ ИДЕНТИФИКАЦИИ ФРАГМЕНТОВ ИЗОБРАЖЕНИЙ СКАНИРОВАННОГО РУКОПИСНОГО ТЕКСТА

Жиляков Е.Г.¹, Заливин А.Н.², Белов С.П.², Черноморец Д.А.¹, Васильева Н.В.¹

¹ Белгородский государственный национальный исследовательский университет, Белгород, РФ

² Белгородский университет кооперации экономики и права, Белгород, РФ

E-mail: belov@bsu.edu.ru

В настоящее время накоплены большие хранилища данных, полученных при сканировании рукописных текстов. Существенное место в них занимают сканированные напечатанные документы, которые содержат рукописные подписи должностных лиц. Полученные в процессе сканирования изображения текстов часто подвергаются компьютерному анализу в связи с той или иной необходимостью. Существенный интерес представляет поиск в этих изображениях фрагментов, содержащих заданные словоформы, например в филологии при исследовании частоты использования одним и тем же автором некоторых слов. Можно также указать случаи поиска слов с позиций обеспечения безопасности социально-экономических процессов. Важным примером является обнаружение фальсификаций подписей должностных лиц и т. п. Особенностью автоматического поиска идентичных словных фрагментов на изображениях сканированных документов является возможность их идентификации с использованием только одного образца текста (прецедента), что требует создания специальной методики машинного обучения. В представленной статье разработана решающая процедура отнесения словных фрагментов изображений сканированного рукописного текста к классу идентичных заданному прецеденту. В качестве элементов признаков пространства предложено использовать проекции векторов на соответствующие ненулевым собственным числам собственные векторы субполосных матриц. Обоснован способ формирования суммарных субполосных матриц на основе введенного понятия информационных субполос в области пространственных частот. Предложена процедура обучения на основе одного прецедента. В основе этой процедуры используется разработанный метод формирования векторов, совокупность которых моделирует обучающую выборку. Сформирован алгоритм обработки изображений при поиске идентичных заданному фрагменту.

Ключевые слова: изображения сканированного рукописного текста, поиск фрагментов, идентичных заданному, субполосный анализ