

НАУЧНЫЙ ЖУРНАЛ

**СИСТЕМЫ синхронизации,
формирования и обработки
сигналов**

№4-2023 год

Главный редактор

Пестряков Александр Валентинович,

*д.т.н., профессор, зав. кафедрой Радиооборудование и Схемотехника,
Московский технический университет связи и информатики, Москва, Россия*

Редколлегия:

Дмитриев Александр Сергеевич,

*д.ф.-м.н., профессор, Институт радиотехники и электроники им. В.А. Котельникова РАН,
Москва, Россия*

Казakov Леонид Николаевич,

*д.т.н., профессор, зав. кафедрой Радиотехнических систем, Ярославский государственный
университет им. П.Г. Демидова, Ярославль, Россия*

Карякин Владимир Леонидович,

*д.т.н., профессор, Поволжский государственный университет телекоммуникаций
и информатики, Самара, Россия*

Рыжков Анатолий Васильевич,

*д.т.н., главный научный сотрудник, профессор, Московский технический университет
связи и информатики, Москва, Россия*

Строганова Елена Петровна,

*д.т.н., профессор, Начальник Испытательной лаборатории средств связи и вещания,
Московский технический университет связи и информатики, Москва, Россия*

Учредитель:

ООО «ИД Медиа Паблшер»

Номер подписан в печать 30.05.2023 г.

СОДЕРЖАНИЕ

Гольденберг В.Л., Луговенко Д.В., Репинский В.Н. СИНТЕЗ И РЕАЛИЗАЦИЯ ПОЛИНОМИАЛЬНОГО ЗАГРАЖДАЮЩЕГО ФИЛЬТРА ПО ВЧ-ПРОТОТИПУ	4
Дегтярев В.Ф. О РЕЗОНАНСНЫХ УРОВНЯХ ПРИ НАДБАРЬЕРНОМ ДВИЖЕНИИ МИКРОЧАСТИЦЫ В ЛИНЕЙНОЙ КВАНТОВО-РАЗМЕРНОЙ СТРУКТУРЕ	11
Ким Е.Д., Митрофанов И.А., Еремин И.А., Носенко Ф.А., Егоров Д.А. МЕТОДОЛОГИЯ ПОИСКА ИДЕАЛЬНОГО РАЗМЕРА ДИСПЛЕЯ В ПОМЕЩЕНИИ	19
Комаров С.Н., Фуников И.Д. КОМПРЕССОРЫ СИГНАЛА МОДУЛЯЦИИ ДЛЯ ЭФИРНОГО АМ ПРОЦЕССОРА НА ОТЕЧЕСТВЕННОЙ ЭЛЕМЕНТНОЙ БАЗЕ	26
Пеев Д.Д., Панков К.Н., Власов А.В. ЗАЩИТА КАНАЛОВ УПРАВЛЕНИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ КРИПТОГРАФИЧЕСКИМИ МЕТОДАМИ	33
Саломатина Е.В., Сарьян В.К. КОНТЕНТНЫЙ ВЗГЛЯД НА СЕТИ	44
Силантьева А.С. КЛАССИФИКАЦИЯ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ ОБЪЕКТОВ	52

СИНТЕЗ И РЕАЛИЗАЦИЯ ПОЛИНОМИАЛЬНОГО ЗАГРАЖДАЮЩЕГО ФИЛЬТРА ПО ВЧ-ПРОТОТИПУ

Гольденберг Валерий Лазаревич,

Московский Технический Университет Связи и Информатики, старший инженер НИО-48 НИЧ МТУСИ, Москва, Россия

jewodan@rambler.ru

Луговенко Денис Владимирович,

Московский Технический Университет Связи и Информатики, магистрант 1-го года обучения, Москва, Россия,

den.lugovenko@mail.ru

Репинский Владимир Николаевич,

Московский Технический Университет Связи и Информатики, д.к. ИСВиА, ведущий научный сотрудник НИО-48 НИЧ МТУСИ, Москва, Россия,

repinski@rambler.ru

Аннотация

Описана модифицированная методика синтеза заграждающих фильтров по рабочим параметрам, приведен синтез конкретного фильтра и технологические решения по его реализации, описан выбор способа его электронной перестройки.

Ключевые слова: *фаза-частотная характеристика, сигнал, прототип, помеха, синтез, фильтр, двухполосник, четырехполосник, полоса пропускания, неравномерность.*

Введение

Частотно-избирательные цепи играют важнейшую роль в радиотехнике. Их задача - отделять желаемый сигнал из множества сигналов и помех или, наоборот, подавлять нежелательные колебания.

Применение современных методов синтеза электрических цепей позволяет создавать цепи с частотными характеристиками любой сложности, которые не требуют перестройки [1,2]. Однако, при решении задачи создания перестраиваемых фильтров, возникают дополнительные требования к фильтру и технологиям перестройки [3], особенно если речь идет о создании систем в области частот, пограничной между частотами, где реализации в виде систем с сосредоточенными и распределенными параметрами неудобна в обоих случаях. Настоящая статья посвящена синтезу заграждающего фильтра именно в таких условиях.

Результаты исследований – синтез и реализация заграждающего фильтра высокого порядка в условиях проблемной реализации.

При синтезе перестраиваемого фильтра необходимо задавать его ключевые параметры [4,5]:

- Допустимые изменения ширины полосы пропускания при перестройке с нижней частоты диапазона перестройки на верхнюю;
- Допустимые изменения уровня искажений при перестройке;
- Допустимые уровни потерь в полосе пропускания и гарантированного затухания в полосе задерживания для всего диапазона перестройки;
- Обеспечение требуемого закона изменения частоты настройки в заданном диапазоне при воздействии на исполнительное устройство, перестраивающее фильтр;
- Обеспечение оптимальной формы, габаритов, массы устройства перестройки [6,9]
- Устойчивость устройства перестройки к механическим и климатическим воздействиям

Задачей синтеза перестраиваемых фильтров является получение либо постоянной абсолютной, либо постоянной относительной ширины пропускания. Первое требование характерно для радиоприемных устройств, где нужно обеспечить минимально допустимую полосу при перестройке, второе – для

радиопередающих устройств, где требуется минимальные потери в полосе пропускания и постоянный коэффициент фильтрации гармоник [8]

Синтез фильтра по рабочим параметрам применяется при создании широкополосных фильтров, которые, в отличие от узкополосных, не удастся согласовать по входу и выходу, нагрузив характеристическими сопротивлениями, ввиду их выраженной зависимости от частоты [7]. Обычно, фильтр имеет активную нагрузку или комплексную с малой реактивной составляющей, и согласование возможно только в узкой полосе частот, где выходное сопротивление фильтра слабо зависит от частоты. То же можно сказать и про входное сопротивление фильтра и внутреннее сопротивление источника входного колебания. Теория и практика расчета фильтров, нагруженных на характеристическое сопротивление представлена фильтрами тип «к» и типа «м», про которые известны также то, что в широкой полосе частот их теоретические характеристики никогда не совпадают с экспериментальными (хотя эффект фильтрации имеет место, однако не может быть оценен заранее) [13]. Принципиальное отличие фильтров, созданных с учетом рабочих параметров от фильтров, рассчитанных по характеристическим параметрам в том, что в первом случае имеет место синтез оптимальной схемы устройства, характеристики которого соответствуют техническому заданию, а во втором подбор из базы готовых схем без каких-либо гарантий оптимальности выбранной схемы для данной технической задачи. Синтез узкополосных фильтров по рабочим параметрам часто приводит к схемам с нереализуемыми на сосредоточенных элементах значениям емкостей и индуктивностей и по этой причине не получил широкого распространения. Тем не менее, такие расчеты известны, в отличие от расчетов заграждающих фильтров, поэтому стоит уделить внимание именно такому фильтру и случаю, когда классическая реализация невозможна на сосредоточенных элементах, а диапазон частот не допускает использование элементов с распределенными параметрами.

Замещение реактивного двухполюсника резонансной цепью, построенной на магнитно-связанных индуктивных катушках

На частоте резонанса (29.93 МГц), реактивное сопротивление катушки L_1 равно 3.5 Ома. Частотная зависимость модуля сопротивления контура L_1C_1 приведена на рисунке 1.

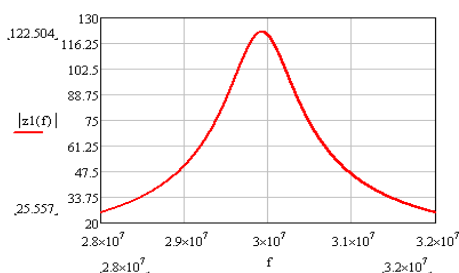


Рис. 1. Частотная зависимость модуля сопротивления контура $L_1C_1L_1C_1$

Соответственно, фазо-частотная характеристика контура $L_1C_1L_1C_1$ приведена на рисунке 2. Отметим, что фазо-частотная характеристика идеального двухполюсника без потерь представляет собой разрывную функцию и для сравнения двухполюсников малоприсгодна (рис. 2,А), поэтому при построении графика рис.2,Б мы ввели потери, эквивалентные значению добротности 200, что считается достаточным для синтеза фильтра без учета потерь [11].

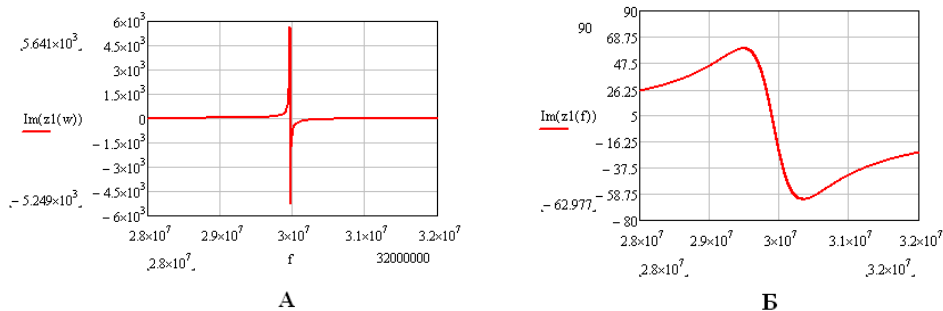


Рис. 2. Фазо-частотная характеристика контура $L_1C_1L_1C_1$

Таким образом, для реализации заграждающего фильтра с заданными характеристиками, необходимо синтезировать реализуемый на практике двухполюсник и АЧХ и ФЧХ совпадающими с соответствующими АЧХ и ФЧХ теоретического прототипа в рабочем диапазоне частот – от 25 до 35 МГц.

Простейший способ реализации избирательного звена с индуктивным элементом очень малой индуктивностью приведен на рисунке 3.

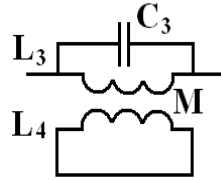


Рис. 3. Избирательное звено с индуктивным элементом

Эквивалентная индуктивность контура в этом случае определяется вносимым сопротивлением, которое носит емкостный характер и, следовательно, общая индуктивность уменьшается [12]. Подбором величин индуктивностей L_1 , L_2 и взаимной индуктивности M добиваемся требуемого значения общей индуктивности цепи:

$$L = L_3 - \frac{M^2}{L_4} \quad L = L_3 - M^2/L_4 \quad (1)$$

Реализуя расчетные индуктивности вышеизложенным образом, получим схему фильтра, по своим характеристикам полностью идентичного, полученного синтезом:

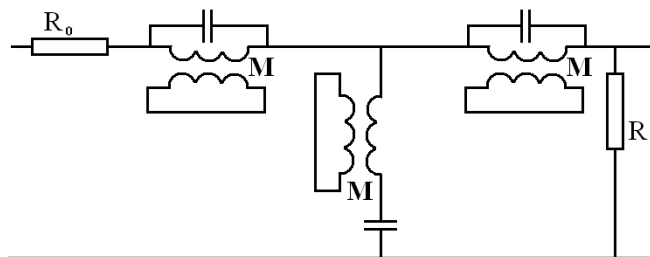


Рис. 4. Схема фильтра, по своим характеристикам полностью идентичного, полученного синтезом

Подобная реализация решает вопрос о малой индуктивности, но ничем не поможет в случае, если синтез дал очень малые значения емкости соответствующих элементов.

Универсальным является метод замены колебательных систем в фильтре, аналогичными, но реализуемыми на практике системами [12], по критерию одинаковости (в определенных пределах частотного диапазона) комплексного входного сопротивления. Одним из возможных двухполюсников такого типа является схема, изображенная на рисунке 5.

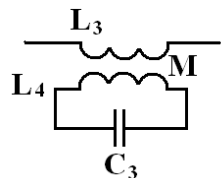


Рис.5. Схема двухполюсника

Входное сопротивление цепи, изображенной на рисунке 4, равно:

$$Z_{вх} = j\omega L_3 + \frac{\omega^2 M^2}{j\omega L_4 + \frac{1}{j\omega C_3}} \quad (2)$$

Добавим к индуктивным сопротивлениям небольшие потери, чтобы график сопротивления был более удобным для анализа, получим следующее выражение:

$$Z_{\text{вх}} = j\omega L_3 + R_{03} \frac{\omega^2 M^2}{j\omega L_4 + R_{04} + \frac{1}{j\omega C_3}} \quad (3)$$

На рисунке 6 показаны частотная зависимость реактивного сопротивления, полученного из выражения (3) – синяя кривая. На том же рисунке красным цветом показана кривая частотной зависимости реактивного сопротивления классического параллельного колебательного контура с потерями из схемы синтезированного фильтра.

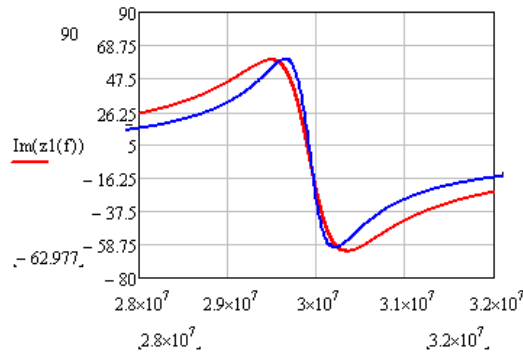


Рис. 6. Частотная зависимость реактивного сопротивления

Как видно из рисунка 6, при тех же потерях в индуктивных элементах, цепь, схема которой изображена на рисунке 4 обладает несколько более высокой избирательностью, но в целом, характеристики можно считать близкими, то есть замена правомерна. Более заметны отличия в частотных зависимостях модулей сопротивлений схемы (рис. 5) и классического параллельного контура (рис. 7).

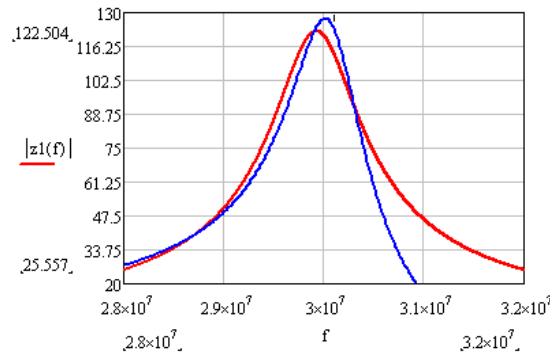


Рис. 7. Частотные зависимости модуля сопротивления системы и классического параллельного контура

На рисунке 7 бросается в глаза асимметрия АЧХ входного сопротивления колебательной системы с взаимной индуктивностью и емкостным элементом в добавочном контуре. Такое отличие характеристик может привести даже к лучшей фильтрации, однако характеристика уже не будет аналитической функцией заданного типа (Баттерворта, Чебышёва, Кауэра-Золотарева) [10]. Это обстоятельство позволяет сделать вывод о допустимости использования замены классического избирательного звена на звено типа рис. 5, однако потребуются некоторая дополнительная настройка фильтра для выравнивания характеристики, подавления нежелательных резонансов. Несмотря на перечисленные затруднения, указанный метод является практически единственным способом создания полосового избирательного устройства на сосредоточенных элементах в высшей части КВ диапазона.

Схема фильтра на элементах избирательности с индуктивной связью и емкостным элементом в дополнительном контуре, приведена на рисунке 8.

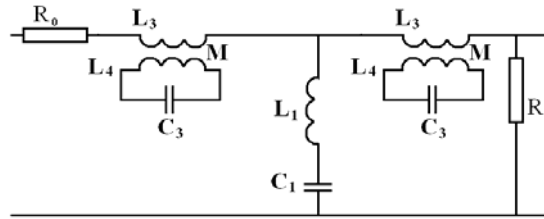


Рис. 8. Схема фильтра на элементах избирательности с индуктивной связью и емкостным элементом в дополнительном контуре

Заметим, что в параллельной ветви применение избирательного звена с взаимной индуктивностью не понадобилось, значения емкости и индуктивности получились реализуемыми. Выражение для передаточной функции цепи рис.8 имеет вид:

$$H(j\omega) = \frac{1}{\frac{1}{j\omega L_1 + \frac{1}{j\omega C_1}} + \frac{1}{R_0 + j\omega L_3 + \frac{\omega^2 M^2}{j\omega L_4 + \frac{1}{j\omega C_3}}}} \cdot \frac{1}{R_0 + j\omega L_3 + \frac{\omega^2 M^2}{j\omega L_4 + \frac{1}{j\omega C_3}} + \frac{1}{\frac{1}{j\omega L_1 + \frac{1}{j\omega C_1}} + \frac{1}{R_0 + j\omega L_3 + \frac{\omega^2 M^2}{j\omega L_4 + \frac{1}{j\omega C_3}}}}} \cdot \frac{R_0}{R_0 + j\omega L_3 + \frac{\omega^2 M^2}{j\omega L_4 + \frac{1}{j\omega C_3}}} \quad (4)$$

Ввести в это выражение потери достаточно легко, необходимо прибавить к каждому индуктивному сопротивлению резистивное сопротивление потерь [14]. На рисунке 9 показана, рассчитанная по (4) частотная зависимость передаточной функции фильтра (рис. 8):

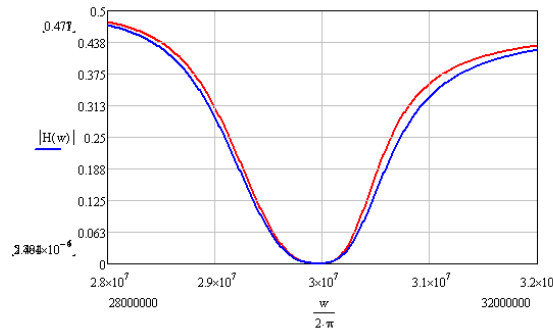


Рис. 9. Частотная зависимость передаточной функции фильтра

Введение активных потерь приводит к изменению характеристики (синяя кривая), полоса задерживания несколько расширяется, но характер кривой, ее асимметричность остаются прежними. Для выравнивания характеристики в процессе настройки был добавлен один емкостный элемент (рис. 10), в системе появился еще один резонанс и характеристику подбором емкости этого конденсатора удалось симметризовать (рис. 11).

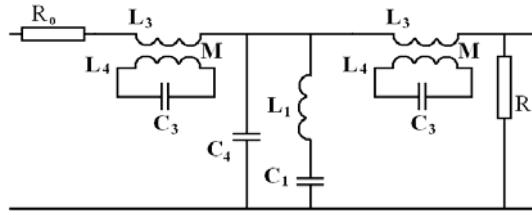


Рис. 10. Добавление емкостного элемента в схему фильтра

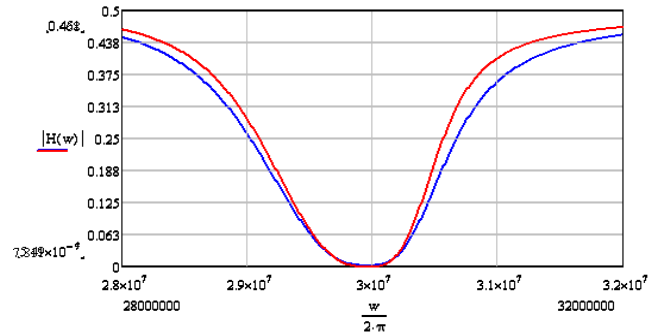


Рис. 11. Частотная характеристика фильтра после добавления емкостного элемента

При синтезе электрических цепей, в том числе и электрических фильтров, применяется параметр рабочего ослабления. Это отношение мощности в нагрузке четырехполюсника к мощности в нагрузке, согласованной с данным источником, выраженное в дБ. Рабочее ослабление связано с коэффициентом передачи четырехполюсника выражением:

$$A = -10 \log \left(\frac{1}{4|H(j\omega)|^2} \right) \quad (5)$$

На рисунке 12 приведены кривые частотных зависимостей рабочего ослабления классического синтезированного фильтра с характеристикой Баттерворта (синяя кривая) и фильтра, реализованного на избирательных элементах со связанными катушками и емкостным элементом во вторичном контуре (красная кривая).

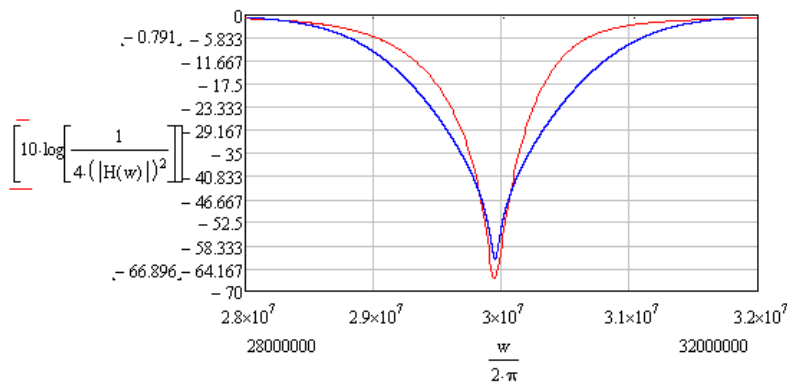


Рис. 12. Частотные зависимости рабочего ослабления классического синтезированного фильтра с характеристикой Баттерворта (синяя кривая) и фильтра, реализованного на избирательных элементах со связанными катушками и емкостным элементом во вторичном контуре (красная кривая)

Как и следовало ожидать, частотная зависимость рабочего ослабления не сохранила аналитичность описания, изменилась крутизна характеристики в переходной области, однако в целом, можно сказать, что подобная реализация эффективна и может рассматриваться даже как средство улучшения характеристики фильтра (полученные зависимости рабочего ослабления что-то среднее между кривыми Баттерворта и Чебышева).

Перестройка фильтра осуществляется путем электронного переключения матрицы емкостных элементов.

Заключение

В результате исследования синтеза заграждающего фильтра можно сделать вывод, что синтез заграждающего фильтра удобно производить по предложенной выше методике, используя в качестве прототипа не ФНЧ, как обычно, а ФВЧ, что значительно экономит время и упрощает вычислительные процедуры при расчетах.

Использование в качестве дополнительного малого емкостного элемента замкнутой системы индуктивно-связанных катушек позволяет реализовать характеристику фильтра на частотах, не позволяющих использование сосредоточенных элементов, но достаточно невысоких, чтобы реализация могла быть осуществлена с помощью систем с распределенными параметрами.

Литература

1. *Попов П.А.* Применение частотных преобразований в теории цепей. М. Энергоатомиздат, 1986, 136 с.
2. *Зелевич А.Ф.* Методические указания к курсовой работе по ТЛЭЦ. Изд. Новосибирского электротехнического института связи, 1975. 35 с.
3. *Зелевич А.Ф.* Расчет фильтров по рабочим параметрам / Попов П.А. Хазанов Г.Л. Изд. Новосибирского электротехнического института связи, 1975. 34 с.
4. *Ханзел Г.* Справочник по расчету фильтров. М. Сов. Радио, 1974.
5. *Ретинский В.Н., Смирнов Н.И., Фриск В.В.* Синтез фильтров нижних частот. М.: Информсвязьиздат, 1992.
6. *Белоус А.И., Солодуха В.А., Шведов С.В.* Космическая электроника. М.: Техносфера, 2015. 488 с.
7. *Абдуллаев Г.Б.* О реактивных свойствах кремниевых диффузионных р-п переходов // Радиотехника и электроника. Т.10, 2007. С. 789.
8. *Авеллино,* Эффективная генерация нечетных гармоник при помощи противоположно-включенных варакторных диодов // Труды Института инженеров по электронике и радиотехнике. 2009. Т. 52. 928 с..
9. *Аришинов С.С.* Температурная стабильность частоты ламповых генераторов, Госэнергоиздат, 2011.
10. *Багаев В.С.* Емкость и последовательное сопротивление германиевых диодов // Радиотехника и электроника. Т. 6. № 12, 2040 с.
11. *Воронова Л.И., Толмачев. Р.В., Усачев В.А.* Resource Development to Prevent Riots at Mass Events // 2018 Systems of signals generating and processing in the field of on board communications. 2018. IEEE Conference #43917.
12. *Трунов А С., Воронова Л.И., Воронов В.И., Сухачев Д.И., Стрельников В.Г.* Legacy Applications Model Integration to Support Scientific Experiment // 2018 Systems of signals generating and processing in the field of on board communications. 2018. IEEE Conference #43917.
13. *Воронова Л.И.* Применение компьютерного моделирования в курсе физики. Учебное пособие. Курган: изд-во КГУ, 2001, 100 с.
14. *Аджемов С.С., Гольденберг В.Л., Ретинский В.Н., Окунев Н.В.* Режекторный перестраиваемый фильтр. Патент на изобретение RU 2671042 С1, 29.10.2018/ Заявка №2017117773 от 22.05.2017.

О РЕЗОНАНСНЫХ УРОВНЯХ ПРИ НАДБАРЬЕРНОМ ДВИЖЕНИИ МИКРОЧАСТИЦЫ В ЛИНЕЙНОЙ КВАНТОВО-РАЗМЕРНОЙ СТРУКТУРЕ

Дегтярев Владимир Фомич,

Московский технический университет связи и информатики, доцент, к. ф.-м. н, Москва, Россия
vfsteel2008@gmail.com

Аннотация

В квантово-размерной цепочке изучены особенности энергетического спектра микрочастиц, обладающих энергией, превосходящей высоту потенциального барьера. Установлено, что и в этом случае образуются резонансные уровни. Для частиц с энергией, равной энергии этих уровней, прозрачность структуры равна единице. Установлено, что с увеличением числа звеньев происходит расщепление этих уровней на близкие подуровни и изменение волновой функции. Определены энергии этих подуровней и волновые функции. Предложен механизм перестройки уровней в цепочке. В основе этого механизма лежат представления о точках смены фаз колебаний осцилляторов при образовании цепи. Установлено, что параметры этих подуровней (энергия, волновая функция) зависят от параметров барьеров, ям и числа ячеек в цепочке. Предложенная модель, позволяющая определить характеристики этих подуровней, в частности их энергию и волновые функции.

Ключевые слова: Квантовая механика, квантовый барьер, волновая функция, прозрачность, наноэлектроника, резонансные уровни.

Введение

Резонансные явления в квантово-размерных структурах оказывают существенное влияние на их физические свойства и представляют собой новое быстро развивающееся направление исследований в наноэлектронике. Особенно ярко эти эффекты проявляются в сверхрешетках, содержащих большое число элементарных ячеек. Описанию свойств сверхрешеток посвящен ряд обзоров, например [1-3]. В этих работах достаточно подробно рассмотрена их энергетическая структура, оптические и электрические свойства.

В современных наноэлектронных приборах (резонансных туннельных диодах и транзисторах и др.) активная область имеет размеры порядка нескольких нанометров. Эта величина сравнима с длиной волны электрона и межатомным расстоянием типичных полупроводников, что приводит к существенному влиянию квантово-размерных эффектов на их работу. Среди этих эффектов значительную роль играют резонансные явления, связанные с прохождением микрочастиц над системой потенциальных барьеров, образующих сверхрешетку. При этом возникают новые резонансные уровни, энергетическая структура которых существенно изменяется при образовании сверхрешетки, приводя к значительному изменению их оптических и электрических свойств. В данной работе рассмотрено образование и перестройка резонансных

уровней в процессе образования квантовой сверхрешетки, предложен возможный механизм трансформации уровней. Рассмотрен случай, когда энергия микрочастицы превосходит высоту потенциальных барьеров, образующих структуру.

Модели структур, изучаемые в работе, и методика моделирования.

В настоящей работе нахождение волновых функций и коэффициентов прозрачности системы барьеров проводилось путем непосредственного решения уравнения Шредингера для заданного потенциала с соответствующими граничными условиями в системе компьютерной алгебры MAPLE. Системы компьютерной алгебры помимо численного решения поставленной задачи позволяют получить и аналитическое решение, что делает расчеты более наглядными, и позволяет не только построить теоретическую модель физического процесса, но и визуализировать полученные результаты. В работе исследована перестройка резонансных уровней, обладающих энергией, превышающей энер-

гию потенциальных барьеров. Изучена зависимость спектров прозрачности и волновых функций сверхрешетки от энергии частиц и числа ячеек в цепочке. Рассмотрено влияние высоты и толщины барьера между ячейками на спектр прозрачности и волновую функцию.

Параметры барьеров в цепочке принимались следующими: ширина барьеров (а) и ям (b) - 10 Å, высота барьера (U1) - 2 эВ. Амплитуда падающей волны принималась равной единице ($\Psi_{in}(x) = 1 \cdot e^{ikx}$). Энергия микрочастиц изменялась в пределах от 2 до 10 эВ.

Параметры резонансных уровней в одиночном звене

В рассматриваемой структуре элементарным звеном является двухбарьерная квантовая структура (ДБКС), показанная на рисунке 1.

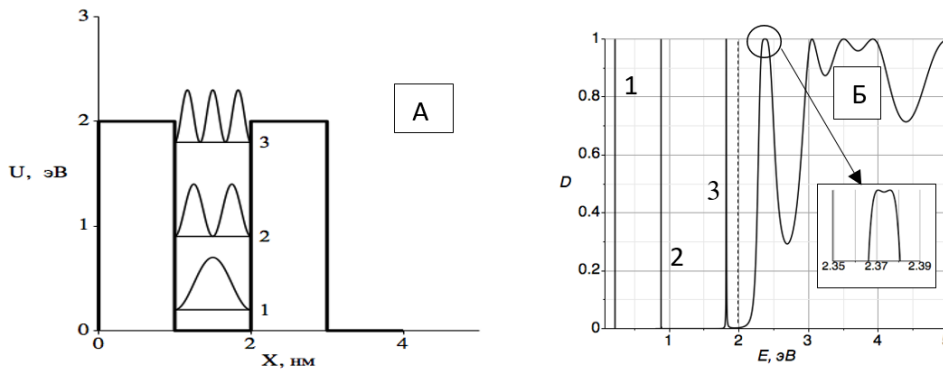


Рис. 1. А. Схема элементарного звена цепочки и волновые функции в случае: 1 – на ширине ямы укладывается одна полуволна; 2 – две полуволны; 3 – три полуволны. Б. Спектр прозрачности элементарного звена. Пунктир – высота барьера (2 эВ).

Резонансные уровни и волновые функции для первого случая ($E < U$) показаны на рисунке 1А.

1. В элементарном звене, также как и в потенциальной яме конечной глубины (толщина барьера при этом считается бесконечно большой) при $E < U$ образуется система резонансно-туннельных уровней, положение которых определяется условием $b = n \frac{\lambda}{2}$, где b – ширина ямы, λ -длина волны

для частицы с энергией E , $n = 1, 2, \dots$ – число полуволн, укладываемых на ширине ямы (рис.1А кривые 1, 2, 3). Прозрачность структуры для частиц с такой энергией равна единице. В элементарном звене рассматриваемой структуры число таких уровней равно 3 (рис. 1А). Как показывают расчеты, их энергетическое положение следующее: $E_{01} = 0.228949$ эВ, $E_{02} = 0.887202$ эВ и $E_{03} = 1.818140$ эВ. Энергия этих пиков соответствует положению энергетических уровней в яме такой же ширины и глубины, что неудивительно при малой прозрачности барьеров, разделяющих ячейки. Это легко понять, если учесть, что оба типа пиков соответствуют условию, что на ширине ямы умещается одинаковое число полуволн.

Как известно [4], резонансное туннелирование наблюдается при выполнении условия

$$\frac{1}{2} \left(\eta - \frac{1}{\eta} \right) = \text{cth}(\beta a) \text{ctg}(Kb),$$

где

$$\eta = \frac{K}{\beta}, \quad K = \sqrt{\frac{2m}{\hbar^2} E}, \quad \beta = \sqrt{\frac{2m}{\hbar^2} (U - E)} \quad (1)$$

a - ширина барьера, b – ширина ямы, E – энергия частицы (эВ), U – высота барьера (эВ).

Поскольку речь идет о резонансных уровнях их энергию при моделировании необходимо определять с высокой точностью. На кривых прозрачности (рис.1А) этим уровням соответствуют узкие резонансные линии 1, 2, 3.

2. Рассмотрим случай $E > U$. Кривая прозрачности для этого случая показана на рис. 1Б (область $E > 2$ эВ).

При этом вместо отдельных пиков наблюдаются широкие полосы. Ограничим рассмотрение только уровнем с энергией 2.37 эВ. Этот уровень является наиболее близким к барьеру и взаимодействует с ним более сильно. Эта полоса (см. вставку к рис. 1Б) представляет собой суперпозицию двух уровней с энергиями 2.371 эВ и 2.376 эВ. Волновые функции, соответствующие этим уровням, показаны на рисунке 2 А и Б. При увеличении энергии частицы уровни расширяются и превращаются в достаточно широкие разрешенные полосы.

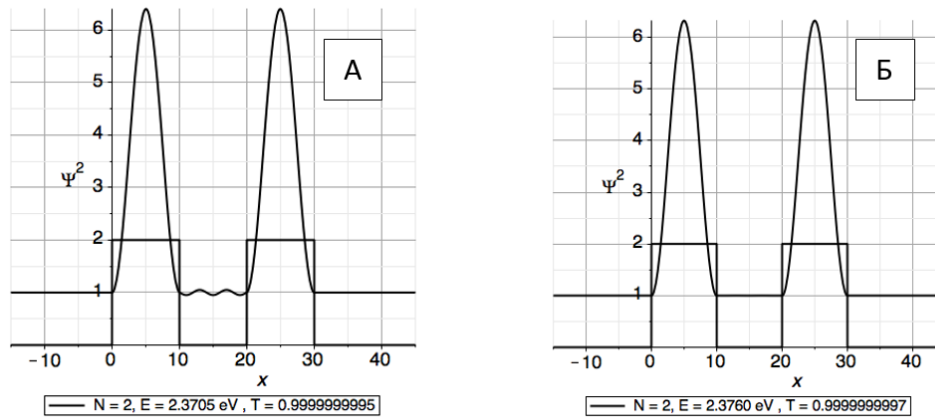


Рис. 2. Волновые функции, соответствующие уровням $E=2.371$ эВ (А) и $E=2.376$ эВ (Б)

Из рисунка 2А видно, что при энергии частицы 2.371 эВ между двумя барьерами образуется стоячая волна. Вне этой системы частица движется свободно. Это говорит о достаточно сильном взаимодействии частицы с барьером. Поэтому при рассмотрении цепочки будет рассмотрено расщепление именно этого уровня.

В случае, когда $E=2.376$ эВ, между барьерами стоячая волна не образуется и в этом промежутке частица движется свободно. Такое движение можно рассматривать как движение над системой невзаимодействующих изолированных барьеров. Это подтверждается также сходством волновых функций и равенством энергии прозрачности этого уровня и одиночного барьера.

Расчеты прозрачности для симметричной двухбарьерной структуры для случая $E > U$ приведены, например, в [4]. Прозрачность при этом описывается соотношением

$$D = \left\{ 1 + \frac{(K^2 - K_2^2)^2 \sin^2(K_2 a)}{4K^2 K_2^2} R \right\}^{-1},$$

где

$$K = \sqrt{\frac{2m}{\hbar^2} E}, \quad K_2 = \sqrt{\frac{2m}{\hbar^2} (E - U)}, \quad (2)$$

$$R = \left[\frac{(K^2 + K_2^2) \sin(K_2 a) \sin(Kb)}{K^2 K_2^2} - \frac{2KK_2 \cos(K_2 a) \cos(Kb)}{K^2 K_2^2} \right]^2 \quad (3)$$

В соответствии с этими соотношениями при $\sin(K_2 a) = 0$ или $R = 0$ коэффициент прозрачности будет равен единице. Заметим, что при $R=1$ выражение (1) совпадает с коэффициентом прозрачности для одиночного барьера.

При $\sin(K_2 a) = 0$ коэффициент прозрачности равен единице и не зависит от параметров структуры. В этом случае не возникает отраженных волн ни от первого, ни от второго барьера. Эта ситуация проиллюстрирована на рис. 2Б. При этом между барьерами концентрация частиц равна концентрации частиц с данной энергией, испускаемых источником. Следовательно, частицы в этом случае накапливаются только в области барьеров и поэтому прохождение частиц через систему барьеров подобно прохождению через систему одиночных невзаимодействующих между собой звеньев.

При $R=0$ коэффициент прохождения над каждым барьером по отдельности не равен единице. В этом случае происходит частичное отражение потока от барьеров и между барьерами образуется стоячая волна (рис. 2А). При этом за счет накопления частиц в области между барьерами полный поток частиц, прошедших второй барьер будет равен потоку частиц, испускаемых источником, и будет подавлено отражение частиц, налетающих на первый барьер. Коэффициент прозрачности системы из двух барьеров будет равным единице. Однако за счет образования стоячей волны барьеры будут взаимодействовать между собой.

Таким образом, систему уровней прозрачности, образующихся при надбарьерном движении частицы (при $E > U$) можно разбить на два типа. В первом случае происходит достаточно сильное взаимодействие частиц с барьерами. При этом происходит расщепление уровней на подуровни. Во втором случае движение частицы подобно движению над системой отдельных изолированных друг от друга барьеров. Трансформацию этих уровней при образовании цепочки можно рассматривать по отдельности независимо друг от друга.

Расщепление резонансных уровней на подуровни при увеличении числа звеньев

С увеличением числа звеньев в цепи резонансные уровни расщепляются на подуровни, число которых равно числу звеньев в системе (числу степеней свободы). Такое поведение представляется естественным, если считать систему до образования цепочки N -кратно вырожденной. Образование цепочки приводит к снятию вырождения и возникновению подуровней.

Некоторые особенности расщепления уровней в ДКБС рассмотрены в работах [4-6]. В [4,6] определены значения энергии этих подуровней и соответствующие им волновые функции. Было установлено, что прозрачность цепочки для этих значений энергии равна единице.

В настоящем разделе рассмотрен процесс расщепления исходных резонансных уровней при увеличении числа звеньев в цепи. Схема расщепления уровней приведена на рисунке 3.

При образовании цепочки происходит взаимодействие осцилляторов в отдельных звеньях. При этом возникающая волновая функция является суперпозицией волновых функций в ячейках [4,8]. Заметим, что здесь речь идет именно о самой волновой функции, а не о квадрате ее модуля. Схема расщепления уровней в случае $E < UE < U$ рассмотрена в [7,8]. Рассмотрим для определенности пер-

вый энергетический уровень (уровень $E=0.228949$ эВ на рисунке 1А). В этом случае на ширине ямы при резонансе укладывается одна полуволна. Пусть теперь к данному звену присоединили второе. В результате взаимодействия осцилляторов в звеньях образуется новая волновая функция, которая описывает колебания, происходящие в первой и второй ячейках. Если в результате взаимодействия осцилляторов колебания в обеих полуволнах происходят в одной фазе, то энергия колебаний уменьшается. Если колебания происходят в противоположных фазах, то энергия колебаний увеличивается [4,9-11].

Таким колебаниям соответствует симметричная и антисимметричная волновая функция [4,9]. В результате исходный уровень расщепляется на два. Если при этом воспользоваться аналогией с механическими колебаниями грузиков на пружинках, то в точках смены направлений колебаний пружины растянуты сильнее, что увеличивает энергию колебаний и, следовательно, энергию подуровня. При этом уменьшение энергии первого подуровня компенсируется увеличением энергии второго, что обеспечивает выполнение закона сохранения энергии. Аналогичным образом изменяется частота колебаний и в связанных электрических контурах [10, 11].

При увеличении числа звеньев взаимодействие осцилляторов в соседних ячейках происходит аналогичным образом. При этом число точек смены фаз колебаний (число узлов волновой функции, число мод колебаний) может изменяться от 0 до максимального значения $N-1$, где N – число звеньев в цепи. Число подуровней будет равно числу звеньев в цепочке. Подуровень, который соответствует колебаниям, происходящим синфазно, обладает наименьшей энергией. С появлением точек смены фаз энергия подуровня возрастает пропорционально числу этих точек. Следует отметить, что положение точек смены фаз не влияет на энергию подуровней. С увеличением длины цепочки энергетическое расстояние между подуровнями уменьшается, что связано с ослаблением взаимодействия между звеньями.

Пусть теперь $E > U$. В этом случае в элементарном звене, содержащем два барьера, образуются два подуровня с энергиями 2.371 эВ и 2.376 эВ. Второй из этих уровней (рис. 2Б) имеет волновую функцию, характерную для невзаимодействующих барьеров, и наблюдается в цепочке с любым их числом (рис. 3). Энергия этих уровней и вероятность нахождения частицы в этих состояниях не зависят от числа звеньев в цепи. Первый же уровень (с энергией 2.371 эВ) образуется только в системах с четным числом барьеров. Это обусловлено тем обстоятельством, что для образования подобных комплексов необходимы два барьера.

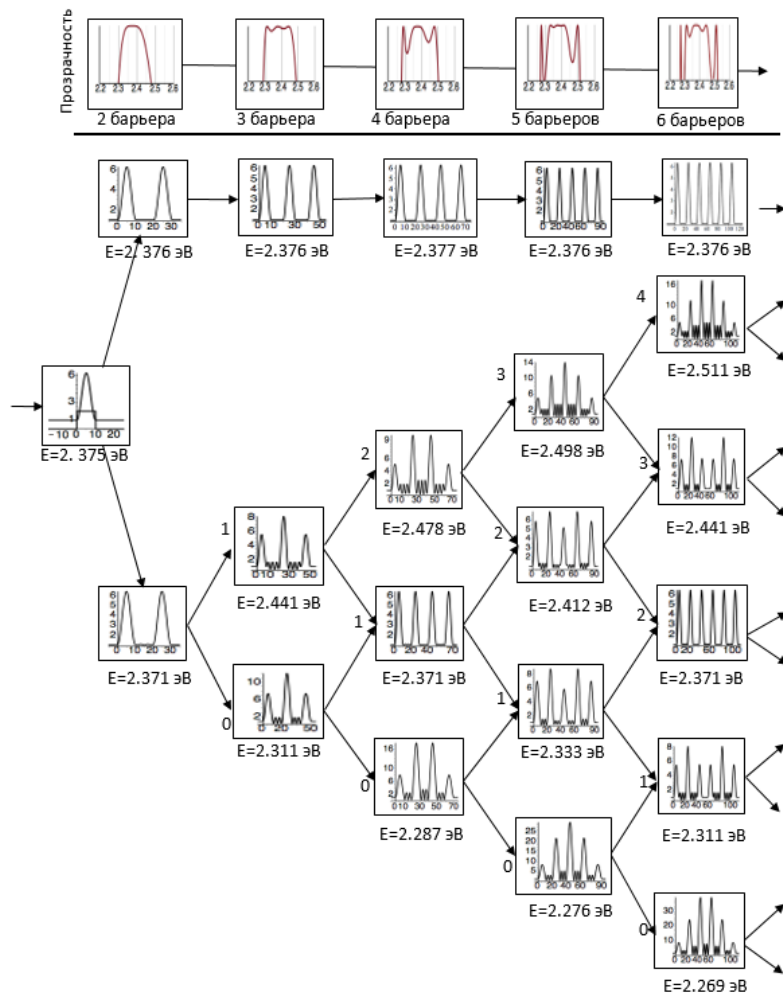


Рис. 3. Схема расщепления уровней при $E > U$ в процессе образования цепочки. Цифрами возле стрелочек обозначено число точек смены фаз колебаний.

Для уровня 2.376 эВ в этом нет необходимости. При увеличении числа звеньев происходит расщепление уровней. При этом заданному числу барьеров соответствует такое же число уровней и число главных максимумов. Между главными максимумами располагаются побочные максимумы, обусловленные наличием стоячей волны между барьерами.

Взаимодействие между ними и приводит к образованию подуровней. При этом представления о смене фаз колебаний при образовании цепочки оказываются полезными для определения характера расщепления и вида волновых функций.

Описание расщепления уровней в случае $E < U$ проведено в [8]. Проведенный расчет дает следующие выражения для энергии подуровней

$$E_{mj} = E_{0m} - 2A_m \cos\left(\frac{\pi j}{N+1}\right), \quad (4)$$

где $j=1, 2, \dots, N$.

Проведенное рассмотрение приводит к зонному характеру распределения состояний по энергии и позволяет найти энергетический спектр частиц в зависимости от количества звеньев в цепи [8]. Для квадрата модуля волновой функции получается выражение:

$$\Psi^2 = a \cdot \sin^2(kx_n) = a \sin^2\left(\frac{\pi j n}{N+1}\right) \quad (5)$$

Здесь учтено, что $x_n = nd$ $x_n = nd$, n - номер ячейки ($n=1 \dots N$), d - период структуры, j -номер моды ($j=1 \dots N$). Из этого соотношения видно, что при $j_1 + j_2 = N+1$ $\Psi_1^2 = \Psi_2^2$, а также при $n_1 + n_2 = N+1$ $\Psi_1^2 = \Psi_2^2$. Т.е. волновая функция имеет симметричный вид (рис. 4).

В случае $E > U$ расчет выполняется аналогично, но необходимо учесть, что частицы локализируются в области барьеров, и число барьеров на единицу больше, чем число ям. Соответствующий расчет приводит к соотношению

$$E_{mj} = E_{0m} - 2A_m \cos\left(\frac{\pi j}{M}\right), \quad (6)$$

где $j=1, 2, \dots, M-1$ и M – число барьеров.

При расчете числа подуровней следует помнить, что имеется еще один подуровень (в нашем случае это $E=2.376$ эВ), который не расщепляется на подуровни. Таким образом в сверхрешетке, содержащей M барьеров образуется M подуровней.

Перестройка волновых функций при образовании цепочки

При образовании слоистой квантово-размерной структуры существенные изменения испытывают не только энергетические уровни, но и волновые функции. Волновые функции для $E > U$ приведены на рисунке 4. Вид волновой функции и характер её перестройки достаточно хорошо описывается выражением (5).

В случае $E > U$ перестройка волновых функций носит достаточно сложный характер. В этом случае частицы локализируются, главным образом, в области барьеров (рис. 4), где их кинетическая энергия мала. Понятие о точках смены фаз колебаний (об узлах волновой функции) оказывается полезным и в этом случае. Здесь также с увеличением номера моды растет энергия подуровней.

Имеется, однако, и ряд особенностей по сравнению со случаем $E < U$.

1) Число пиков волновой функции равно числу барьеров. Их количество определяется только числом барьеров и не зависит от номера моды.

2) В случае совпадения узла волновой функции с барьером соответствующий пик не исчезает. Уменьшается лишь его высота.

3) Из всей совокупности подуровней можно выделить один (в данном случае это уровень $E=2.376$ эВ), который не расщепляется на подуровни. Соответствующая волновая функция представ-

ляет собой систему волновых функций одиночного прямоугольного барьера (рис. 2Б). Число пиков при этом равно числу барьеров (рис. 4).

Обращает на себя внимание симметричный вид волновых функций относительно середины цепочки. Именно такой вид волновой функции обеспечивает, по нашему мнению, равенство потоков частиц слева направо и справа налево, что и приводит к высокой прозрачности цепочки. Следует отметить также, что волновые функции m -ой и $(N+1-m)$ -ой мод имеют совершенно одинаковый вид.

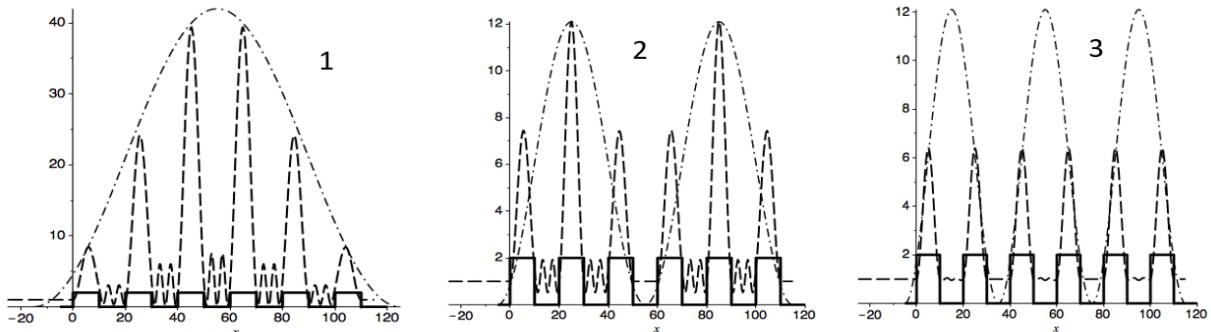


Рис.4. Квадраты модуля волновых функций частиц в цепочке, содержащей шесть барьеров:
1 - первая мода; 2 – вторая мода; 3 – третья мода

Таким образом, рассмотренная модель позволяет объяснить основные особенности перестройки волновых функций при образовании слоистой структуры.

Заключение

В работе рассмотрен процесс перестройки резонансных уровней при образовании слоистой квантово-размерной структуры.

Показано, что существенным изменениям при этом подвергается и волновая функция.

1. Установлено, что при образовании цепочки резонансные уровни расщепляются на систему подуровней, число которых равно числу звеньев в цепочке.

2. Определены значения энергии этих подуровней и соответствующие им волновые функции в зависимости от числа звеньев. Прозрачность цепочки для этих значений энергии равна единице. Предложена методика, позволяющая оценить эти энергии и построить соответствующие волновые функции.

3. Для объяснения механизма перестройки уровней развиты представления о точках смены фаз колебаний. Это такие точки, в которых фазы колебаний в соседних звеньях изменяются на противоположные. Число таких точек изменяется в пределах от нуля до N . Соответственно число подуровней, на которые расщепляется исходное состояние, равно числу звеньев в цепи. Чем больше таких точек, тем выше энергия колебаний.

4. Рассмотренные представления справедливы как для $E < U$, так и для $E > U$. Установлено, что для $E > U$ имеется энергия, при которой цепочку можно рассматривать как совокупность независимых барьеров такой же высоты.

Автор выражает искреннюю благодарность профессору Жилинскому А. П. за постоянный интерес к работе и полезные замечания.

Литература

1. Демиховский В.Я., Вугальтер Г. А. Физика квантовых низкоразмерных структур. М.: Логос, 2000. 248 с.
2. Херман М. Полупроводниковые сверхрешетки. М., Мир, 1989. 240 с.
3. Силин А.П. Полупроводниковые сверхрешетки. УФН, т. 47, вып. 3, с. 7, 485-516.
4. Драгунов В.П., Неизвестный И.Г., Гридчин В.А. Нанозлектроника, часть 1. М.: Юрайт, 2019. 285 с.
5. Аладышкин А.Ю. Туннельные явления в нанофизике. Н. Новгород: Нижегород. гос. ун-т., 2011. 32 с.
6. Давыдов А.С. Квантовая механика. М.: Наука, 1973. 702 с.
7. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике, т. 9, Квантовая механика (II). М.: Мир, 1967. 259 с.

8. *Дегтярев В.Ф., Жилинский А.П.* Трансформация резонансных туннельных уровней при образовании слоистой квантово-размерной структуры. Наноструктуры. Математическая физика и моделирование. 2020. Т.21. № 2. С. 33-48.
9. *Киттель Ч.* Введение в физику твердого тела. М.: Наука, 1978. 791 с.
10. *Стрелков С.П.* Введение в теорию колебаний. СПб: Лань, 2005. 440 с.
11. *Крауфорд Ф.* Берклеевский курс физики. Т.3, Волны. М.: Наука, 1984. 521 с.

МЕТОДОЛОГИЯ ПОИСКА ИДЕАЛЬНОГО РАЗМЕРА ДИСПЛЕЯ В ПОМЕЩЕНИИ

Ким Евгений Дмитриевич,

Московский Технический Университет Связи и Информатики, Москва, Россия
blackmail@internet.ru

Митрофанов Илья Алексеевич,

Московский Технический Университет Связи и Информатики, Москва, Россия
picfox2003@gmail.com

Еремин Иван Александрович,

Московский Технический Университет Связи и Информатики, Москва, Россия
shadowdagger@mail.ru

Носенко Фадей Александрович,

Московский Технический Университет Связи и Информатики, Москва, Россия
fadey1nosenko@gmail.com

Егоров Дмитрий Аркадьевич,

*Московский Технический Университет Связи и Информатики, старший преподаватель,
Москва, Россия*
d.a.egorov@mtuci.ru

Аннотация

Значительной частью пропускной способности Интернета является видео контент. Вопрос восприятия видеопоследовательностей является важным элементом исследований. Зрительная система человека имеет общее горизонтальное поле зрения около 200 градусов, часть этого составляет периферийное зрение, однако не весь контент предназначен для заполнения всего поля обзора человека. Проблема воздействия размера и разрешения экрана на улучшение воспринимаемого визуального качества, все еще остается нерешенной. Решения относительно размера дисплея должны требовать понимания физических, перцептивных и когнитивных возможностей каждого отдельного пользователя, а также физических ограничений рабочей среды. В представленном исследовании была разработана методология определения идеального размера экрана.

Ключевые слова: *Зрительная система человека, разрешение экрана, поле зрения человека, угол обзора, размер дисплея.*

Введение

Потоковое видео занимает большую часть пропускной способности Интернета, в 2022 году на видео пришлось 82% Интернет-трафика [1]. Вопрос восприятия мозгом видеопоследовательностей является важным элементом исследований. Большое количество информации представляется в виде статических или динамических изображений. Поскольку далее эти изображения рассматриваются и анализируются человеком, то важно знать механизмы зрительного восприятия. Эти знания являются мощным инструментом при разработке различных систем обработки изображений.

Зрительная система по-разному реагирует на излучения, которые равны по мощности, но излучаются из различных диапазонов спектра. Такая чувствительность называется спектральной. Способность глаза воспринимать минимальные различия яркости смежных областей изображения характеризуется контрастной чувствительностью. Также зрительную систему человека можно охарактеризовать различной чувствительностью к цветовому тону и насыщенности цвета [2, 15].

Поле зрения человеческих глаз соответствует не плоскости, а кривой поверхности с выпуклостью, обращенной вперед; поле зрения не ограничено центральным сектором, однако включает в себя боковые области, за восприятие которых отвечает периферийное зрение [3]. Экран изогнутого монитора

может искривляться по-разному. Параметр кривизны – обозначается буквой R и измеряется в миллиметрах. Кривизна человеческого поля зрения равна примерно 1000R. При использовании плоских мониторов, снижение четкости изображения на уровне визуального восприятия возникает в 4 раза чаще, чем при использовании изогнутых мониторов. Следовательно, на сегодняшний день перед практически каждым любителем телевизионного контента стоит проблема поиска «идеального экрана», который способен учесть, как можно больше факторов, приносящих пользователю дискомфорт.

Существующие сегодня разрешения – практически все 4k (Ultra HD). Следовательно, требуется телевизор большого размера, который необходимо смотреть с наиболее близкого расстояния, чтобы увидеть артефакты, связанные с разрешением. Благодаря этому, существует возможность находиться на более близком расстоянии к телевизору, чем при более низком разрешении, и погружаться в просмотр. По аналогии с кинотеатром: чем больше телевизор заполняет поле зрения, тем более захватывающим будет просмотр. Зрительная система человека имеет общее горизонтальное поле зрения около 200 градусов, хотя часть этого составляет периферийное зрение, не весь контент предназначен для заполнения всего поля обзора человека [4]. Проблема воздействия размера и разрешения экрана на улучшение визуально воспринимаемого качества, все еще остается нерешенной [5,16].

Любые методы, предполагающие решение этой проблемы, должны требовать понимания физических и когнитивных возможностей каждого отдельного пользователя, а также физических ограничений рабочей или операционной среды. Важнейшую роль в данном вопросе играют понятия «угол обзора» и «угловое разрешение» [1].

Угловое разрешение – это точка, в которой глаз может отличать разные детали на изображении. Угол обзора – это максимальный угол, с которого зритель способен различить четкое изображение на экране ЖК-монитора. Например, если максимальный горизонтальный угол обзора TFT-монитора составляет 160-170 градусов, то пользователь будет видеть четкое изображение, а также при просмотре с точки, смещенной относительно центральной оси на 70 градусов.

Цель данной работы заключается в создании методологии оценки качества видеопоследовательностей от физического размера отображения и ее дальнейшее использование в определении идеального размера экрана, ликвидирующего большую часть проблем среднестатистического зрителя.

Остальная часть данной работы организована следующим образом. В разделе «Анализ критериев выбора идеального дисплея» мы проведем анализ статистики выбора идеального дисплея, которую в открытом доступе предоставила компания Samsung, в разделе «Обзор исследований по смежным темам» мы коснемся смежных с нашей темой исследований, проведенных ранее. Раздел «Методология поиска идеального размера дисплея и описание лабораторных условий» представляет собой описание новой методологии оценки качества видеопоследовательностей от физического размера отображения, а также описание лабораторных условий для проведения субъективных экспериментов.

Анализ критериев выбора идеального дисплея

Перед определением идеального размера экрана, требуется отследить, как человек реагирует на искажения в видеопоследовательности, в зависимости от размера отображения. Используемая база данных будет основываться на том же методе сбора, как и в [1]. Когда задаются вопросом о выборе оптимального расстояния от экрана до субъекта, учитываются три основных фактора: размер экрана, его разрешение, а также поле зрения [6]. Данные факторы взаимосвязаны между собой. Если зритель приближается к экрану, то каждый, отдельно взятый, пиксель будет казаться ему больше. При близком расстоянии изображение на экране будет выглядеть, как старая 8-битная игра. Но если пользователь находится на удаленном расстоянии, то отображение на экране будет казаться меньше и сам экран станет занимать меньше пользовательского поля обзора. Помимо этого, размер отображения сигнала будет влиять на то, сколько визуальной площади должно занимать устройство отображения относительно нахождения пользователя мультимедийного контента. Находится идеальный баланс, когда пользователь смотрит в дисплей на таком расстоянии, когда экран занимает главную часть поля зрения, но не настолько близко, чтобы качество изображения было испорчено. По мере приближения, разрешение и поле зрения становятся больше предела комфорта [1].

Большинство современных LCD-мониторов и экранов ноутбуков позволяют нам редактировать некоторые настройки. Снижение яркости дисплея уменьшает напряжение глаз, как и регулировка внешнего освещения. Однако можно выделить некоторые нюансы: очень тусклый монитор увеличит

напряжение глаз. Чтобы отыскать удовлетворительную яркость, необходимо постепенная регулировка до тех пор, пока глаза не перестанут шуриться и напрягаться при чтении текста.

Главный критерий выбора надёжного и комфортного в использовании ноутбука — это тип дисплея, т.е. матовый дисплей или глянцевый. Человеческому глазу приятнее для восприятия матовый тип. Глянцевые экраны имеют лучшую цветопередачу и яркость, однако обладают важным недостатком - они отражают объекты. Если учитывать этот недостаток, при условии малой яркости, глаз человека будет быстро уставать и перенапрягаться из-за непрекращающийся фокусировки на тексте, ведь отражённые объекты будут мешать. Именно из-за этого недостатка, были изобретены матовые экраны - их удобно использовать при любом освещении (но при этом стоит помнить о правильной настройке самой яркости экрана). Однако у матовых экранов, в сравнении с глянцевыми, хуже яркость и цветопередача [7]. Однако для обычного потребителя параметров матовых экранов достаточно, ведь в повседневной жизни здоровье глаз важнее качества графики.

Недостаточный размер диагонали – главная причина покупки нового телекоммуникационного устройства [9]. Следующим по важности параметром, на который обращают внимание пользователи, является качество изображения. Основной причиной покупки нового телевизора является желание обновить свое старое устройство (37%). Помимо этого, пользователи выделяют маленький экран и отсутствие функций смарт ТВ. Необходимость в дополнительном телевизоре (31%) и возможность приобрести телевизор с большей диагональю (28%) – следующие популярные ответы на вопрос «Что побудило задуматься о покупке нового ТВ?». Желание сделать себе подарок выбрали, в качестве повода для покупки, 30% респондентов премиального сегмента, (рис. 1).

Рекомендуемый размер экрана телевизора зависит от размера диагонали экрана. Другими словами, это расстояние от одного угла экрана телевизора или монитора до противоположного. Важными факторами для выбора телевизора являются ширина и глубина, которые следует учитывать при покупке. Ширина – это значение, которое определяет сколько места займет телевизор. Глубина влияет на то, насколько легко установить телевизор на стене. Рекламируемый размер экрана не всегда точен, ведь рамка вокруг экрана отличается по ширине. Чем шире рамка, тем больше экран телевизора. Верный размер телевизора будет зависеть от многих факторов, таких как размер вашей комнаты, расстояние между человеком и телевизором.

Таким образом, телевизоры можно разделить на три категории: маленькие (до 32 дюймов), средние (33-43 дюйма) и большие (44 дюйма и более). В небольшой комнате маленький телевизор, вероятно, лучший вариант. Средние телевизоры также являются хорошим выбором, если планируется сидеть близко к экрану, так как будет оказываться меньшая нагрузка на поле зрения человека. Телевизоры среднего размера являются хорошим балансом между размером и ценой, также размер дисплея достаточно большой, чтобы вы могли наслаждаться изображением на расстоянии. Большие телевизоры идеально подходят для больших комнат или для просмотра фильмов. Но из-за своих преимуществ, это также самый дорогой вариант.

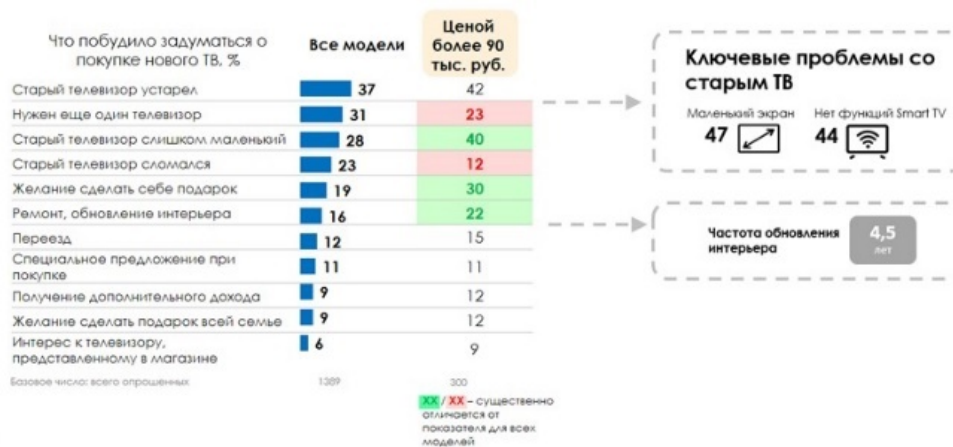


Рис. 1. Статистика покупок новых телевизоров от Samsung

Обзор исследований по смежным темам

Общество инженеров кино и телевидения рекомендует сидеть на таком расстоянии, чтобы экран занимал минимум 30 градусов вашего поля зрения для получения хорошего впечатления, рис. 2 [9].

Данное руководство полезно, но людям, которые обычно используют свои телевизоры для просмотра фильмов, может быть удобнее сесть немного ближе, с целью получения лучших впечатлений от просмотра. Минимальный угол обзора подходит для большинства случаев использования, и большинству людей, должно быть, удобно сидеть на расстоянии, на котором экран занимает 30 градусов горизонтального поля зрения. Также стоит отметить, что этот угол предполагает, что человек смотрит в телевизор на уровне глаз. Нужно учитывать также проблему сжатия. Даже при просмотре HD-контента, транслируемого в высоком разрешении, будут некоторые артефакты из-за алгоритма сжатия. Они могут проявляться в разных формах, например, шум, размытие или пиксельное изображение. Некоторые артефакты могут быть видны издали, поэтому нужно иметь в виду, что цифры показывают минимальное расстояние, на котором теряется преимущество разрешения [9].

Эксперты предполагают, что человек со зрением 20/20 (или 6/6 в метрической системе) может различать детали, находящиеся на расстоянии 1/60 градуса друг от друга [1]. Но человеческая способность различать детали характеризуется не только качеством зрительной системы, но и расстоянием до объекта. Если сидеть слишком далеко, глаза не будут различать отдельные и мелкие детали отображения на экране устройства, но, оказавшись ближе, изображение может оказаться недостаточно хорошим. Иными словами, присутствует некий дискомфорт при статическом положении телекоммуникационного экрана: для более низких разрешений придется сидеть немного дальше, чтобы не различать артефакты видеопоследовательности.

Роберт Бернхем провел исследования, в которых артефакты сравнивались на основе их углового и физического размера и окружения [10]. Был замечен небольшой эффект, со значительными различиями между участниками. Аналогичное исследование, проведенное Питером Бурхом и Гвиллимом Гриндли, не выявило нужного эффекта для размера или продолжительности презентации [11]. Габор Кутас и др. провели субъективный эксперимент, в котором участники сравнивали большую ЖК-панель в зеркальной кабине, которая заполняла FOV участников и отображала их FOV, на которой отображался плоский рисунок, с небольшой картинкой, отображаемой на ЭЛТ-мониторе [12]. Они пришли к выводу, что более крупные стимулы будут казаться светлее.

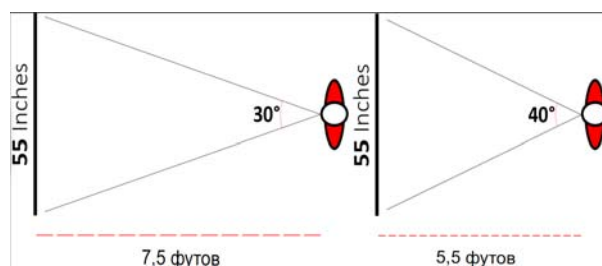


Рис. 2. Изменение угла обзора в зависимости от расстояния человека до дисплея

На сегодняшний день можно воспользоваться уже имеющимся калькулятором по подбору оптимального размера дисплея с рекомендуемым углом обзора – 30 градусов, но ввиду того, что качество транслируемого изображения во многом зависит от контента и условий просмотра, то данный калькулятор является лишь рекомендацией к пользованию [9]. Поэтому вопрос об идеальном размере экрана все еще остается в подвешенном состоянии, а об актуальности проблемы свидетельствует большое количество уже опубликованных работ на данную тематику и проведение различных исследований со стороны крупных компаний, таких как Samsung, Apple и т.д.

Методология поиска идеального размера дисплея и описание лабораторных условий

При помощи субъективных экспериментов необходимо создать набор данных, который будет использован в определении местоположения человека относительно комнаты и размера дисплея. Эксперимент проводится в специально оборудованной комнате, (рис. 5), где в качестве источника света выступает только проектор. Перед первичной оценкой руководитель замеряет яркость с помощью люксметра и фиксирует её (162 люкса) поляризационным фильтром, установленным на линзу проек-

тора. Фильтр представляет собой самодельный прибор, изготовленный из пластика и решетки от обычного жидкокристаллического монитора, позволяющий переменнно настраивать яркость.

Во время проведения эксперимента в комнате присутствуют только 1 участник и 1 руководитель. Дисплеем является проекционный экран, на котором участнику эксперимента транслируют одновременно 8 видеопоследовательностей с разным качеством отображения. Воспроизведение видео начинается в самом низком битрейте. У участника есть устройство для оценки качества восприятия видео контента, который служит манипулятором улучшения битрейта; манипулятор представляет из себя электронную педаль газа автомобиля Lada Priora, рис. 4 [13].

При просмотре видеоролика, участник эксперимента обязан следить за отсутствием искажений/артефактов видеопоследовательности на экране по своему усмотрению. Для эксперимента необходим один монитор и пять различных дистанций для того, чтобы добиться максимально схожих изображений. При смене мониторов, а не расстояний, характеристики каждого из мониторов могут сильно отличаться, что увеличивает погрешность измерений. Информация фиксируется в файл посредством автоматического сбора сведений, (рис. 5).

Участниками эксперимента являются люди в возрастном диапазоне от 20 до 40 лет с нормальным зрением. Люди, принимающие участие в эксперименте, не должны использовать средства коррекции зрения в повседневной жизни. Для правоты эксперимента, участники не должны иметь опыта работы с восприятием визуальной информации человека [14]. Требуется держать баланс между тремя важными параметрами: физической зрелостью глаз, ежедневным использованием всеми привычными пользовательским контентом, а именно просмотром видеорядов и изображений в Интернете, а также отсутствием опыта работы в сфере функционала зрительного восприятия информации, так как его наличие приводит к улучшению нахождения артефактов и обнаружения порога восприятия. Информированное согласие получается от всех участников. После теста первой позиции, проектор меняет свое местоположение, а испытуемый остается неподвижным. После чего данная процедура повторяется несколько раз.

После сбора данных производится связь с приложением «линейка» для смартфонов. Пользователь вписывает размеры комнаты, в которой ему необходимо поместить телевизионный дисплей, затем при помощи камеры смартфона определяется первоначальная точка (будущее место дисплея) и конечная (место для просмотра и времяпровождения за экраном). Ориентируясь на субъективные эксперименты, будет определяться эффективный и идеальный размер дисплея. Также, необходимо учитывать высоту размещения дисплея. Расположение экрана ниже или выше установленной точки, в первую очередь, влечет за собой угрозу для здоровья человека (его осанки) и усиливает блики, поскольку человек не сидит перпендикулярно к экрану. Также необходимо полномасштабные эксперименты, для улучшения программного обеспечения. Будут добавлены учет солнечной стороны комнаты, чтобы сохранить отображение от засветов; а также возможность предоставления типа экрана и его покрытия приложением (OLED, AMOLED, матовый, глянцевый).

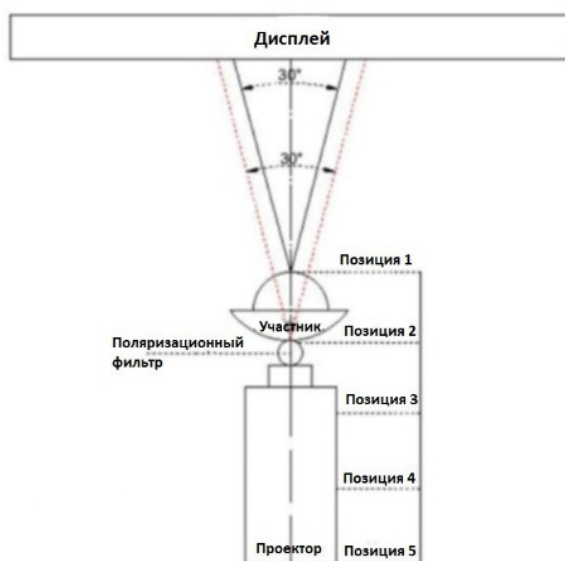


Рис. 3. Общий вид комнаты (вид сверху)



Рис. 4. Манипулятор [13]

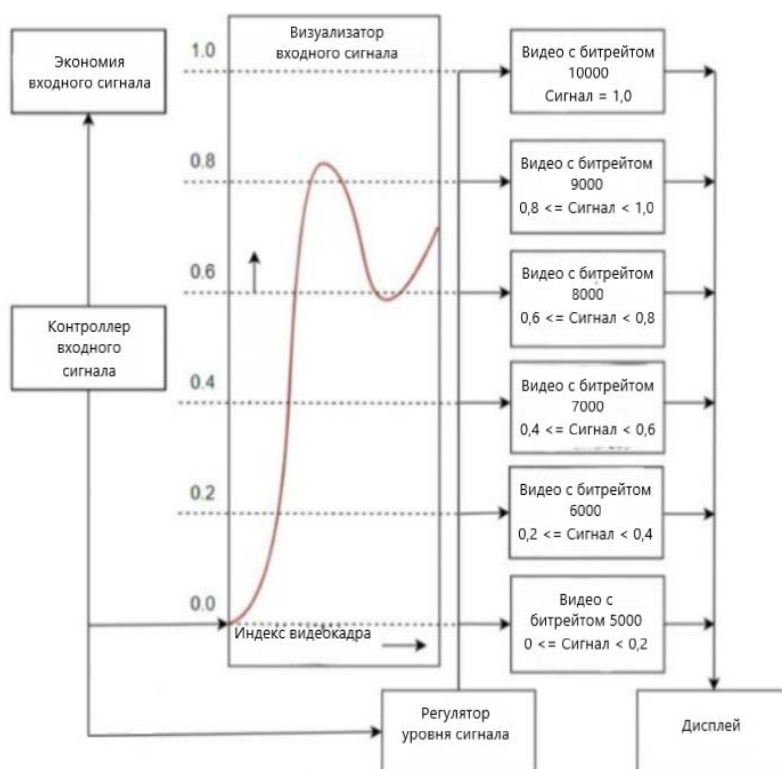


Рис. 5. Блок-схема алгоритма обработки сигнала

Заключение

Текущее исследование описывает методы оценки качества видеопоследовательностей от физического размера дисплея. Для оценки восприятия проанализировано влияние размера разрешения отображения сигнала, достигнутого путем изменения позиции проектора. Разработанный нами метод измерения характеристик зрительной системы человека позволяет проводить ранее не существующие комплексные исследования с новыми подходами к получению данных о зависимости воспринимаемого качества видео контента от расстояния до отображения сигнала. Новая методология измерения может быть использована в научных целях, касающихся практических применений для технологий визуализации, приема и обработки сигналов в устройствах информационных технологий.

Литература

1. Egorov D.A., Davydova A.A., Selivanov V.A., Kapustin P.A., Kim E.D., Taramov A.R. Method for studying the dependence of the visibility of artifacts in an image on its physical dimensions // T-Comm, vol. 16, no.6, pp. 38-44, 2022, doi: 10.36724/2072-8735-2022-16-6-38-44
2. Поташиников А.М., Можжаева А.И. Исследование пороговой разрешающей способности зрительной системы человека для основных цветов системы RGB // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 1. С. 3-10. EDN YFHTXP
3. Можжаева А.И., Власюк И.В., Поташиников А.М., Стример Л. Эталонная объективная метрика оценки качества видео совместимая с PSNR учитывающая частотные и периферическую характеристики зрения человека, // DSPA: Вопросы применения цифровой обработки сигналов. 2021. Т. 11. № 2. С. 44-54. EDN TQJSHP
4. Hubel D.H. Eye, brain, vision. M.: Mir, 1990.
5. Mozhaeva A., Mazin V., Cree M.J., Streeter L. Video Quality Assessment Considering the Features of the Human Visual System. In: Yan, W.Q., Nguyen, M., Stommel, M. (eds) Image and Vision Computing. IVCNZ 2022. Lecture Notes in Computer Science, vol. 13836. Springer, Cham. 2023. https://doi.org/10.1007/978-3-031-25825-1_21
6. Mozhaeva A., Streeter L., Vlasuyk I., Potashnikov A. Full Reference Video Quality Assessment Metric on Base Human Visual System Consistent with PSNR // 2021 28th Conference of Open Innovations Association (FRUCT), 2021, pp. 309-315, doi: 10.23919/FRUCT50888.2021.9347604.
7. Sam Medley, Anton Avdyushkin. Eye health in the age of high technology // Notebookcheck, 2019, online: <https://www.notebookcheck-ru.com/Zabota-o-zdorove-glaz-v-ehpokhu-vysokikh-tekhnologii.400976.0.html>
8. Экран превыше всего: как россияне выбирают телевизоры. <https://news.samsung.com/ru/kak-rossiyane-vybiraut-tv>
9. Demers C., Wiszniewska A. TV Size to Distance Calculator and Science. 2021. https://translated.turbopages.org/proxy_u/en-ru.ru.9dfbc03c-63e287c1-b3fb8701-74722d776562/https/www.rtings.com/tv/reviews/by-size/size-to-distance-relationship
10. Robert W. Burnham. The dependence of color upon area. The American Journal of Psychology, 1951, 64, 4 (1951), pp. 521-533.
11. Burgh Peter, Grindley G.C. Size of test patch and simultaneous contrast // Quarterly Journal of Experimental Psychology», 1962. 14, 2, pp. 89-93.
12. Kutas G., Bodrogi P. Color appearance of a large homogenous visual field. Color Research & Application, 2008, 33, 1, pp. 45-54.
13. Mozhaeva A., Potashnikov A., Vlasuyk I., Streeter L. Constant Subjective Quality Database: The Research and Device of Generating Video Sequences of Constant Quality // 2021 International Conference on Engineering Management of Communication and Technology (EMCTECH), 2021, pp. 1-5, doi: 10.1109/EMCTECH53459.2021.9618977.
14. Mozhaeva A.I., Vlasuyk I.V., Potashnikov A.M., Cree M.J., Streeter L. The Method and Devices for Research the Parameters of the Human Visual System to Video Quality Assessment // 2021 Systems of Signals Generating and Processing in the Field of on Board Communications, 2021, pp. 1-5, doi: 10.1109/IEEECONF51389.2021.9416081.
15. Иванчев В.В., Власюк И.В. Метод распределения выборки цветовых оттенков в цветовом пространстве CIE LCH // Телекоммуникации и информационные технологии. 2022. Т. 9. № 2. С. 108-119.
16. Власюк И.В., Киселева А.С. Анализ эффективности безреференсных метрик применительно к оценке качества видео при потоковой передаче // Телекоммуникации и информационные технологии. 2022. Т. 9. № 2. С. 65-74.

КОМПРЕССОРЫ СИГНАЛА МОДУЛЯЦИИ ДЛЯ ЭФИРНОГО АМ ПРОЦЕССОРА НА ОТЕЧЕСТВЕННОЙ ЭЛЕМЕНТНОЙ БАЗЕ

Комаров Сергей Николаевич,

*Заместитель генерального директора по научно-техническим разработкам ООО «Комета»,
Москва, Россия*

komarov@radiostation.ru

Фуников Илья Дмитриевич,

Московский технический Университет связи и информатики, магистрант, Москва, Россия

Аннотация

На сегодняшний день наша страна идет по курсу импортозамещения, в связи с этим было принято решение отказаться от работы с импортными интегральными микросхемами и приступить к собственной разработке компрессоров, фильтров и, в перспективе, приборов спектродинамической обработки радиовещательного сигнала полностью на отечественной элементной базе. В данной статье приведены два варианта исполнения компрессоров на основе отечественного операционного усилителя 140УД6.

Ключевые слова: *Компрессор, лимитер, динамическая обработка, послестудийная обработка, оптимизатор модуляции, эфирный процессор.*

Компрессор – устройство, выполняющее сжатие динамического диапазона звукового сигнала; иными словами, компрессор уменьшает соотношение между самым тихим и самым громким звуками. Компрессор, как и любое устройство динамической обработки, содержит звуковой тракт и цепь управления. По принципу действия он представляет собой усилитель с Автоматической Регулировкой Усиления (АРУ). Как правило, используется обратная связь по амплитуде сигнала с регулированием назад. В звуковой цепи таких приборов могут использоваться либо специализированные микросхемы, где в качестве регуляторов используются четырёхквadrантные перемножители, либо регулируемые делители напряжения на дискретных элементах с резисторной оптопарой или с полевым транзистором в качестве управляемого переменного резистора.

В любом случае компрессор по сути своей является логарифмирующим усилителем. И зависимость его выходного напряжения от входного можно записать выражением:

$$U_{\text{вых}} = A_k \ln (1 + K_u U_{\text{вх}}).$$

Где:

$U_{\text{вых}}$ и $U_{\text{вх}}$ – выходное и входное напряжения.

K_u – коэффициент усиления в тракте регулирования (в петле обратной связи).

A_k – коэффициент передачи в тракте сигнала (масштабный коэффициент усиления).

Есть еще один вариант компрессирования - лимитирование. Он отличается прежде всего степенью компрессии на большом сигнале. Основное предназначение лимитера — защита последующих узлов тракта от любых, даже самых малых, перегрузок, поскольку независимо от приращения входного сигнала уровень сигнала на его выходе увеличиваться не будет. Это немаловажная способность необходимая компрессору сигнала модуляции.

Лимитирующий радиовещательный компрессор подтягивает тихие сигналы к среднему уровню, а громким сигналам не позволяет превышать некоторое установленное значение, чтобы избежать перемодуляции в АМ передатчике. При этом лимитирование происходит за счёт плавного уменьшения усиления громких звуков.

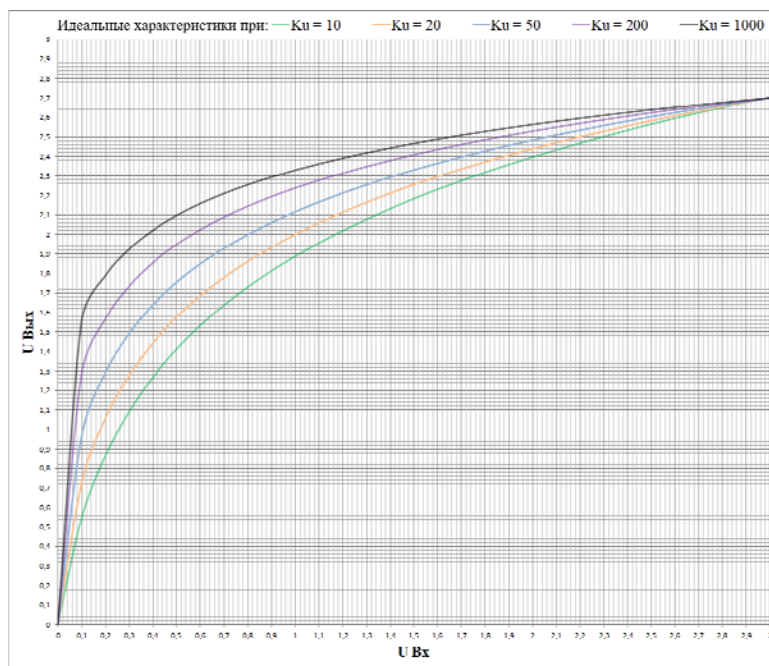


Рис. 1. Идеальные амплитудные характеристики компрессора при разных K_u .

Первый практический вариант компрессора представлен на схеме (рис. 2). Компрессор состоит из трёх основных функциональных элементов: регулируемый усилитель на интегральном операционном усилителе DA1; детектор выходного сигнала – диоды D1 и D2; регулирующий элемент на полевом транзисторе T1. Резисторы R1, R3, R5 и блокировочный конденсатор C3 создают среднюю точку для питания ОУ. В обратной связи ОУ в качестве регулятора АРУ применён полевой транзистор 2П302Б с р-п переходом и встроенным каналом, используемый в режиме обеднения. Резисторы R4 и R7 – антипаразитные. Помимо этого, резистор R4 задаёт максимальное усиление в цепи звукового тракта при отсутствии сигнала. Конденсаторы C1, C2, C7, C8 – разделительные. C4 – корректирующий в цепи обратной связи ОУ. Цепочка R8, C6 является нагрузкой детектора, и её постоянная времени определяет реакцию компрессора на импульсные воздействия.

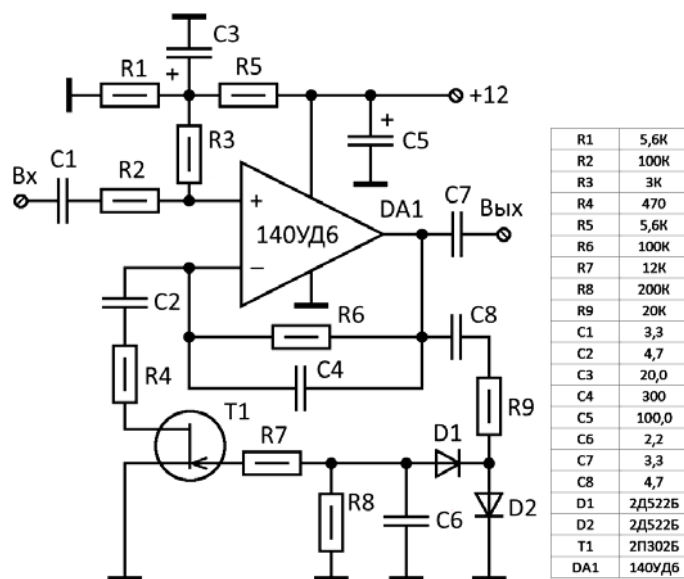


Рис. 2. Схема компрессора полевым транзистором.

Напряжение звуковой частоты с выхода ОУ детектируется диодами D1 и D2 в отрицательное напряжение, которое закрывает р-п переход транзистора. При увеличении уровня выходного сигнала

увеличивается запирающее напряжение на затворе и, соответственно, увеличивается сопротивление промежутка сток-исток, что приводит к уменьшению коэффициента усиления ОУ.

Далее было проведено снятие амплитудной характеристики на частоте 1000 Гц (рис. 3).

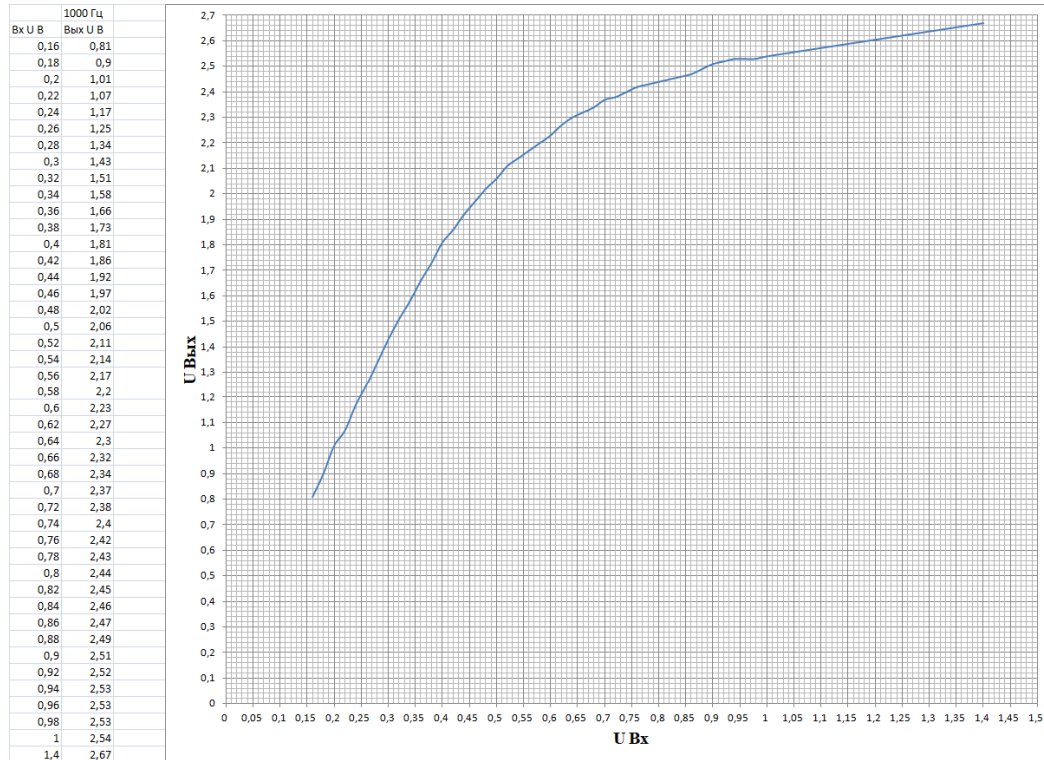


Рис. 3. График амплитудной характеристики компрессора с регулятором на полевом транзисторе.

Как видно из снятого графика, экспериментальная характеристика почти точно повторяет теоретическую зависимость (рис. 1) при небольшом коэффициенте усиления в цепи обратной связи. Такой компрессор сложно назвать лимитирующим. Однако он вполне хорош для мягкого компрессирования сигнала при обработке музыкальных радиопрограмм в студиях звукозаписи. К сожалению, при увеличении входного напряжения более 1,4 вольта эффективного значения, появляются нелинейные искажения. Поэтому в радиовещании, этот компрессор применяться не может.

Второй вариант реализации компрессора (рис. 4) состоит из пяти основных функциональных элементов: регулируемый усилитель (на таком же интегральном операционном усилителе DA1); детектор выходного сигнала – диод D1; пороговое устройство на диоде D2 и резисторе R6; усилитель управляющего сигнала на транзисторе T1; регулирующий элемент – оптрон DA2.

В качестве управляющего элемента применён резисторный оптоинвертор ЗОР124А. Он представляет собой оптоэлектронный полупроводниковый прибор, состоящий из инфракрасного излучающего светодиода и фоторезистора, между которыми имеется оптическая связь, обеспечивающая электрическую изоляцию между входом и выходом.

Выходной сигнал с ОУ 140УД6 детектируется параллельным детектором на диоде 2Д521А и при достижении порога открывания транзистора, заданного диодом D2 и резистором R6, светодиод загорается и освещает фоторезистор. В результате, сопротивление фоторезистора, включенного в цепь обратной связи ОУ, уменьшается, что приводит к уменьшению коэффициента усиления. Резистор R8 ограничивает ток через светодиод при полном открывании транзистора T1. Номиналы резистора R8 и конденсатора С7 определяют постоянную времени восстановления компрессора.

Вторая схема, аналогично предыдущей, была собрана на макете и исследована тем же способом. Экспериментальные значения (зависимость выходного сигнала от входного) были сняты на частоте 1000 Гц. Экспериментальный график изображен на рисунке 5. Этот вариант схемы уже гораздо больше похож на лимитирующий компрессор. Если его сравнивать с теоретическими кривыми на рисунке 1, то он проходит близко с кривой со значением K_u около 800.

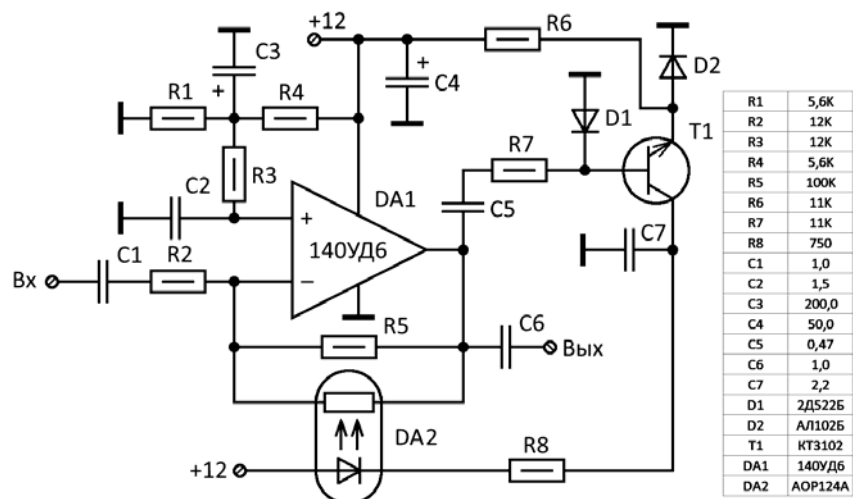


Рис. 4. Схема компрессора с оптронам.

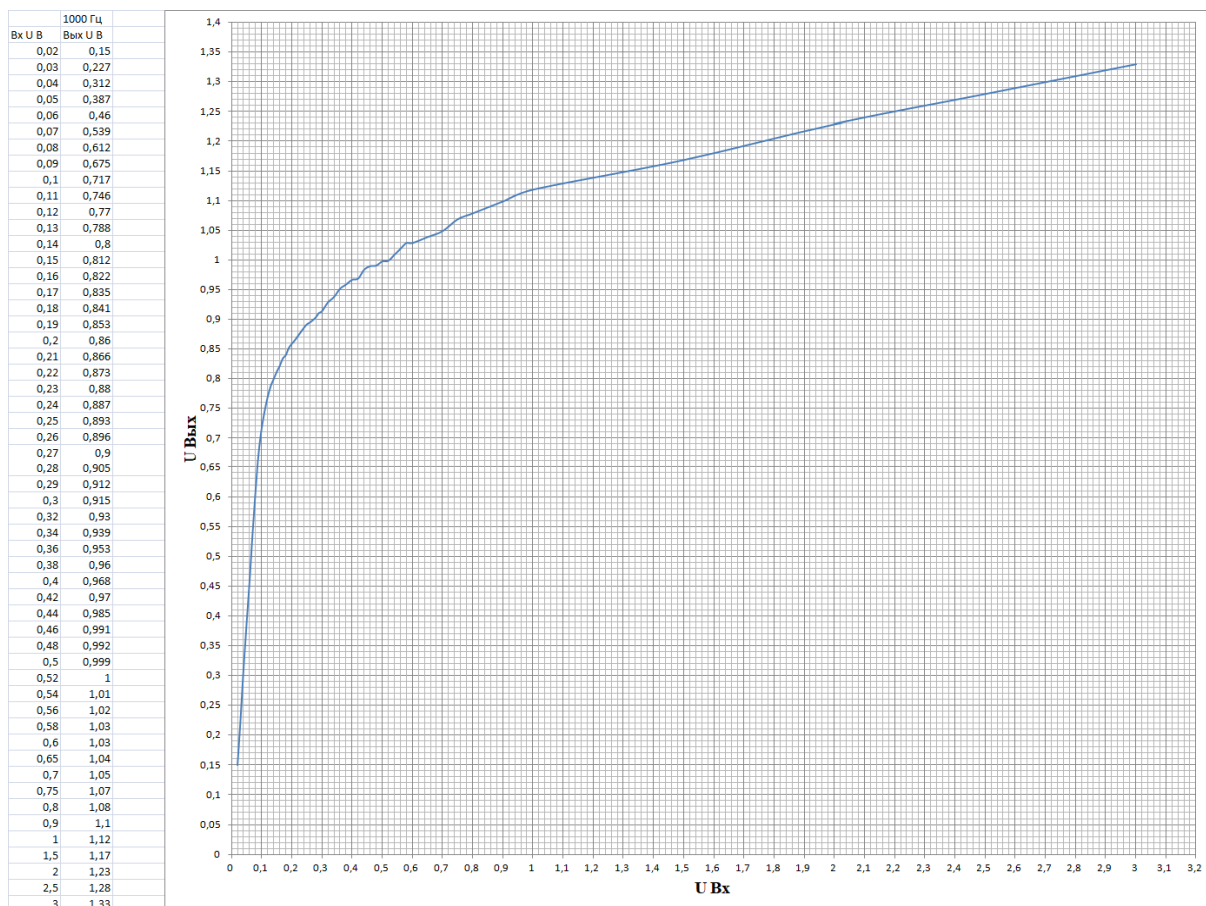


Рис. 5. График амплитудной характеристики компрессора на оптроне.

Соотношение резисторов R5 и R2 определяют коэффициент усиления на малом сигнале, когда выходное напряжение с ОУ ещё не достигает порога. В данной схеме (рис. 4) соотношение $100/12 = 8,3$. Таким образом, крутизну графика на малом сигнале определяет номинал резистора R2. Впрочем, это наглядно видно при сравнении графиков (рис. 3 и рис. 5). В обеих схемах резисторы в цепи обратной связи одинаковы, 100 килоом. Но в схеме (рис. 2) номинал входного резистора также 100 ком (как и в схеме на рис. 4), что задаёт начальное усиление на малом сигнале около единицы.

В схеме на рисунке 2 пороговое напряжение задаётся напряжением открывания кремниевых диодов в детекторе, и составляет около 0,7 вольта.

В схеме на рисунке 4 усиление в петле обратной связи, определяющее пологость амплитудной характеристики на большом сигнале (эффективность лимитирования), может быть увеличено снижением номинала резистора R7, что может также привести к перегрузке операционного усилителя и появлению нелинейных искажений. Кардинальным решением вопроса будет применение вместо одиночного транзистора T1 схемы с составным транзистором (рис. 6), когда ток эмиттера первого становится током базы второго.

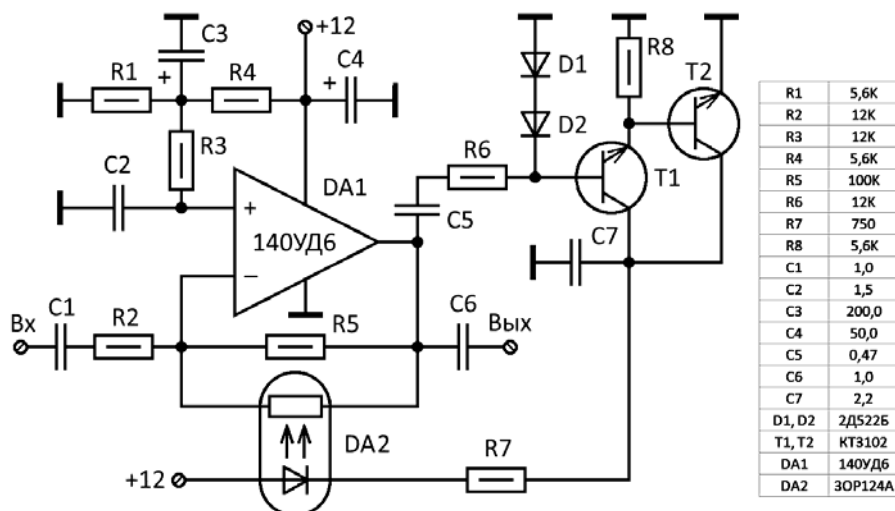


Рис. 6. Схема компрессора с оптроном и усилителем АРУ на составном транзисторе

В такой схеме возможно исключение порогового устройства на прямосмещённом диоде в цепи эмиттера, и вполне будет достаточно порога в 1,35 вольта, составленного из прямого напряжения двух последовательно включённых эмиттерных переходов составного транзистора. Для такого включения в дешёвом варианте можно рекомендовать транзисторы серии КТ315 или аналогичные, для аппаратуры с повышенной надёжностью подойдут транзисторы серии 2Т3102. Однако, для более предметных рекомендаций потребуются дальнейшие исследования.

Лимитирующий компрессор, помимо большей степени компрессии, имеет и принципиально иные динамические характеристики. Он должен очень быстро скомпенсировать сигнал перегрузки и столь же быстро вернуться к исходному состоянию. В хорошем лимитере можно установить время срабатывания вплоть до единиц миллисекунд. Изменяя постоянную времени фильтрующей цепи детектора и его схемотехнику, возможно регулировать время реакции компрессора на всплески входного сигнала, а также время спада. Реализация лимитирующих компрессоров с постоянным выходным уровнем при больших входных уровнях сигнала возможно лишь в схеме с регулировкой вперед. Однако, это приводит к усложнению схемы и появлению в ней второго регулятора. Поэтому представляется оптимальным использование составного транзистора в схеме усиления сигнала управления в петле обратной связи АРУ.

Импульсные характеристики были сняты при помощи осциллографа С1-83. Осциллограммы, демонстрирующие постоянную времени первого компрессора изображены на рисунках 7 и 8.

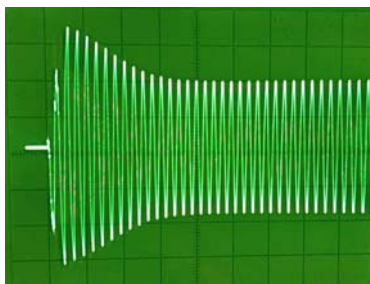


Рис. 7. Компрессор с полевым транзистором (рис. 2). Осциллограмма импульсной характеристики (2 ms/клетка, 2 В/клетка) с постоянной времени 4 ms при C6 = 0,047 мкФ. F = 2 кГц.

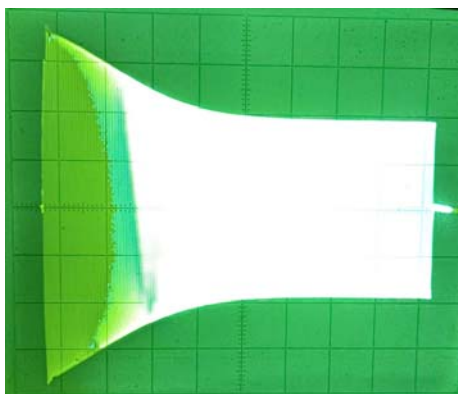


Рис. 8. Компрессор с полевым транзистором (рис. 2). Осциллограмма импульсной характеристики (20 ms/клетка, 2 В/клетка) с постоянной времени 50 ms при $C6 = 2,5$ мкФ.

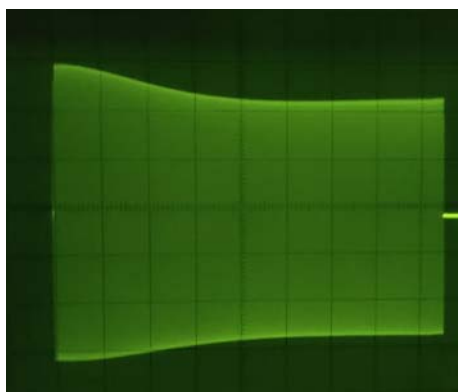


Рис. 9. Компрессор с оптроном (рис. 4). Осциллограмма импульсной характеристики (20 ms/клетка, 50 мВ/клетка) с постоянной времени 50 ms при $C7 = 4.7$ мкФ.



Рис. 10. Компрессор с оптроном (рис. 4). Осциллограмма импульсной характеристики (5 ms/клетка, 50 мВ/клетка) с постоянной времени 2 ms при $C7 = 0.47$ мкФ.

Как видно на рисунках 7 и 8, всплеск выходного сигнала весьма велик и выходное напряжение после импульсного воздействия почти в два раза превышает среднее значение. Такая реакция компрессора обусловлена малым значением коэффициента усиления в петле АРУ. Для модуляционных компрессоров такое неприемлемо. Однако, при динамической обработке музыкальных студийных записей, позволяет сохранить «живость» звучания.

Импульсная характеристика второго компрессора, благодаря наличию усиления в петле АРУ уже имеет меньший всплеск и при доработке схемы (рис. 4) до варианта (рис. 6), уже может применяться в радиовещательных проборах предэфирной обработки сигнала модуляции.

Импульсные характеристики снимались при помощи вспомогательной измерительной схемы, выполняющей роль прерывателя синусоидального сигнала подаваемого на вход компрессора. Устрой-

ство было собрано по схеме мультивибратора с механическим электромагнитным реле, в качестве которого был применён поляризованный релейный переключатель РПС32А (рис. 11).

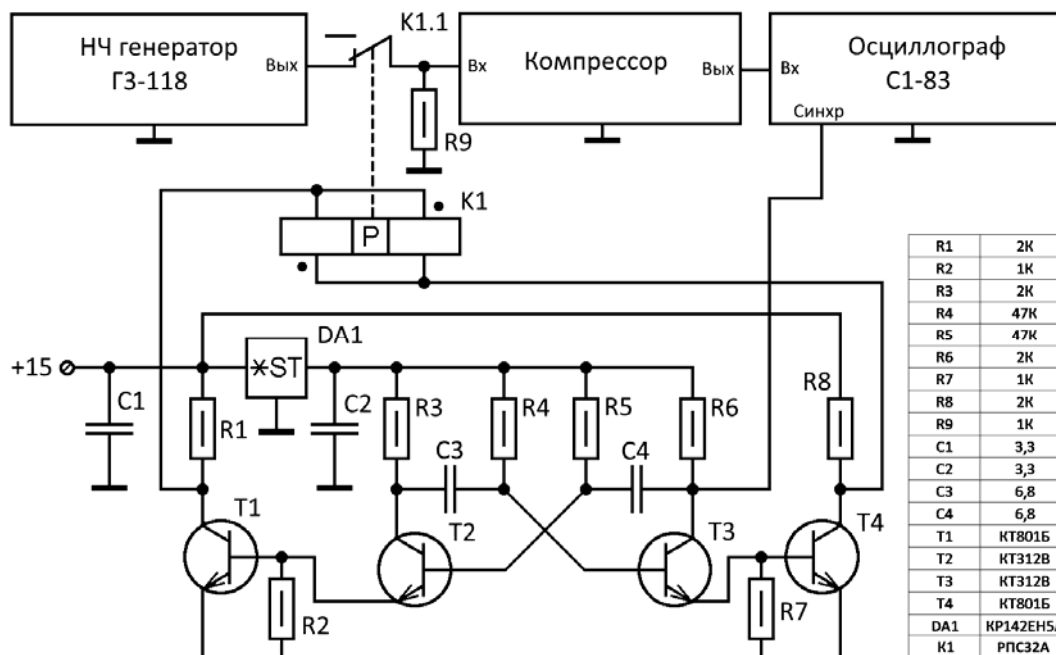


Рис. 11. Схема измерения импульсных характеристик компрессоров.

Проведённые исследования показывают, что для применения в составе приборов предэфирной спектродинамической обработки следует использовать компрессоры с большим усилением в петле обратной связи (АРУ), при значениях K_u , вплоть до 1000 (рис. 1). В этом качестве хорошо себя зарекомендовала схема компрессора, состоящая из пяти функциональных звеньев: регулируемый усилитель; детектор выходного сигнала; пороговое устройство; усилитель управляющего напряжения; регулируемый элемент. Как вариант, возможно применение схемы с оптроном.

Снятые импульсные характеристики подтверждают простое схемотехническое получение постоянных времени от первых единиц до сотен миллисекунд. Исследованные схемы компрессоров можно рекомендовать для обработки сигналов звуковых частот от 30 Гц до 15 кГц.

Литература

1. Шило В.Л. Линейные интегральные схемы в радиоэлектронной аппаратуре. М.: Советское Радио. 1979.
2. Тутце У., Шенк К. Полупроводниковая схемотехника. Под редакцией д.т.н. А.Г. Алексенко. М.: МИР, 1982.
3. Комаров С.Н., Фуников И.Д. Аналоговый эфирный процессор. Принцип работы и исследование // REDS: телекоммуникационные устройства и системы. № 4. 2022.
4. Комаров С.Н. Индивидуальное радиовещание. 2. Самодельные радиопередатчики и многофазные синтезаторы частот диапазона 200 метров. Монография. Практическое руководство для руководителей и членов конструкторских радиокружков и радиостанций Индивидуального радиовещания. М.: ИД Медиа Паблишер, 2018. 214 с.
5. Комаров С.Н. Индивидуальное радиовещание. 7. Цели, задачи и общая стратегия. Организация радиотехнических кружков. Вещательная концепция станций ИРВ. Практическое руководство для руководителей и членов конструкторских радиокружков и радиостанций Индивидуального радиовещания. М.: ИД Медиа Паблишер, 2020. 200 с.

ЗАЩИТА КАНАЛОВ УПРАВЛЕНИЯ БЕСПИЛОТНЫХ ЛЕТАТЕЛЬНЫХ АППАРАТОВ КРИПТОГРАФИЧЕСКИМИ МЕТОДАМИ

Пеев Даниил Дмитриевич,

Московский Технический Университет Связи и Информатики, Москва, Россия

peev.daniil@mail.ru

Панков Константин Николаевич,

Московский Технический Университет Связи и Информатики,

доцент кафедры «Информационная безопасность», к.ф.-м.н., Москва, Россия

pankov_kn@mtuci.ru

Власов Андрей Викторович,

Московский Технический Университет Связи и Информатики, доцент кафедры «Теория вероятностей и прикладная математика», к.ф.-м.н, доцент, Москва, Россия

pankov_kn@mtuci.ru

Аннотация

Работа посвящена актуальной теме – обеспечению информационной безопасности беспилотных летательных аппаратов (далее – БПЛА). Приводится возможный способ решения проблемы с несанкционированным доступом к каналам управления БПЛА в случае его физического захвата злоумышленником. Приведены наиболее эффективные криптографические алгоритмы, которые могут использоваться для защиты каналов управления БПЛА. Описывается работа разработанного программного обеспечения, обеспечивающего защиту канала управления.

Ключевые слова: *Информационная безопасность, беспилотные летательные аппараты, БПЛА, угрозы БПЛА, программное обеспечение для БПЛА, электронная подпись, криптография, блочные алгоритмы шифрования, низкоресурсная криптография*

Введение

В настоящее время использование беспилотных летательных аппаратов (далее – БПЛА) частными лицами, государственными и коммерческими организациями стало общепринятой практикой [1]. Существует, к примеру, множество БПЛА, которые могут в полете снимать необходимую визуальную информацию на поверхности земли (производить аэрофотосъемку) для создания карт с эффектом 3D. Так же возможна обыкновенная видеосъемка в реальном времени, которая помогает достигать невыполнимые задачи для обыкновенных людей, а также для других летательных аппаратов, в том числе, в рамках сетцентрического подхода [2]. Возможна лидарная съемка – эта технология является системой зондирования в дистанционном формате, которая при помощи лазерного импульса собирает необходимые измерения. Ещё БПЛА может являться носителем метрическомагнитного комплекса, с помощью которого становятся доступной информация о магнитном поле земли, что позволяет обнаружить аномалии магнитного поля. Такая съемка помогает в решении археологических и инженерно-геологических задач.

Исходя из анализа основных возможностей БПЛА возникает вопрос обеспечения безопасного и неприкосновенного доступа к управлению беспилотника. Ведь к настоящему моменту известно много случаев, когда злоумышленник получал доступ к управлению БПЛА, на котором передача информации была не защищена. Такой вид несанкционированного доступа обычно приводит к необратимым последствиям.

Целью данной работы является исследование некоторых уязвимостей систем БПЛА и разработка метода обеспечения неприкосновенности и защиты от злоумышленников каналов управления БПЛА с помощью криптографических методов.

Защита каналов управления беспилотных воздушных судов

Для того чтобы БПЛА совершал полеты необходимо создать канал связи. Для его создания требуется микроконтроллер, который будет обрабатывать полученную информацию от оператора. Системы связи с беспилотным воздушным судном могут осуществляться, к примеру, через спутниковые каналы связи, такой вид передачи команд характерен для коммерческого владения летательным аппаратом. При этом способе передачи воздушное судно может летать на огромные расстояния. При использовании стандартных каналов радиосвязи полет может осуществляться на небольшие расстояния, примерно до 70 километров, при этом нельзя забывать, что другие БПЛА, которые находятся в воздухе, и иные воздушные объекты могут как приводить к появлению помех, так и являться ретрансляторами.

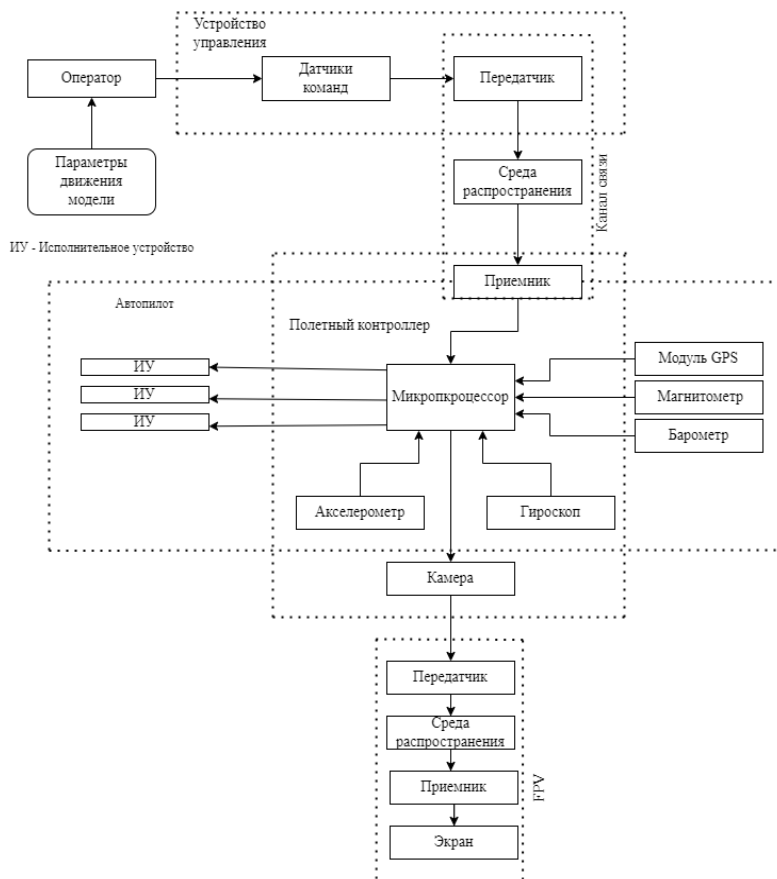


Рис. 1. Алгоритм работы заводских программных решений

Большинство БПЛА гражданского назначения (частного и коммерческого) оснащены обыкновенным контроллером. На рисунке 1 из [3] представлен алгоритм работы подобных контроллеров.

Производители подобных контроллеров, в основном, не задумываются о перехвате устройств, на которых установлено такое аппаратное решение. Поэтому, актуальна задача разработки метода, который поможет обезопасить беспилотник и его владельца (оператора) от несанкционированного доступа злоумышленника.

Для безопасности канала управления одним из авторов работы был разработан метод защиты передаваемой информации от пульта управления к беспилотнику. Для обеспечения безопасного полета в пульт управления должен быть встроен слот под USB-накопитель и шифратор. В слот будет подключаться носитель специального ключа, выполненный в роли USB-накопителя, на котором будет храниться информация с ключами для определенной криптосистемы. Шифратор будет зашифровывать показатели, выставленные оператором на пульте, взаимодействуя с USB-ключом, после чего информация будет передаваться по среде распространения в зашифрованном формате. Беспилотный летательный аппарат тоже должен иметь на своем борту слот под USB-накопитель и дешифратор (произ-

водящий расшифрование информации). На рисунке 2 продемонстрирован модернизированный алгоритм блока управления БПЛА с внедрением всех необходимых систем безопасности.

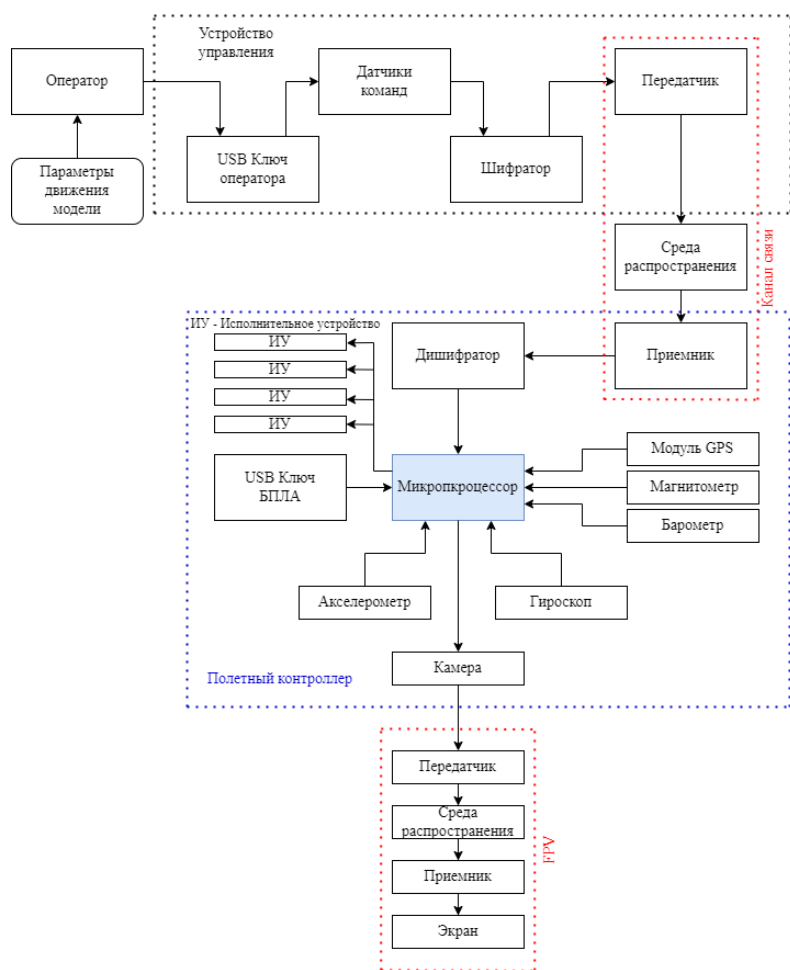


Рис. 2. Модернизированный алгоритм блока управления БПЛА с внедрением всех необходимых систем безопасности

Оператору поступает команда выполнить задачу. Ему выдают два USB-ключа (электронной подписи или ЭП), которые он должен подключить один в наземный пульт управления (далее – НПУ), а второй в сам беспилотный летательный аппарат. После подключения можно устанавливать связь. Система проверяет наличие ключа доступа (ЭП) на борту, а также в пульте и начинает давать определенные команды. В НПУ поставленные сигналы зашифровываются и передаются на борт БПЛА. Получив команду, система БПЛА понимает, что данный сигнал необходимо расшифровать, выполняет процесс расшифрования, и только после этого дает необходимую информацию на контроллер, установленный в беспилотном летательном аппарате.

Немного об значении применения ЭП. Модуль с ЭП крайне важен для организации воздушного движения БПЛА с НПУ, так же этот модуль выступает в роли пускового ключа, дающего команду запуска, как НПУ, так и БПЛА. Данный элемент выполнен в виде USB-накопителя. В нем содержится информация, в которой имеется, в том числе, секретный ключ ЭП и симметричной системы шифрования. Носитель ключа ЭП представлен на рисунке 3.



Рис. 3. Модуль ключа электронной подписи в виде USB-накопителя.

После подключения носителя ключа ЭП к пульту управления программное обеспечение опрашивает блок проверки ключей подтверждения, который имеет перечень привилегированных пользователей, тем самым находит подходящий ключ, после чего система разрешает оператору пользование наземным пультом (именуемым криптопультом) управления. Алгоритм отправки информации с НПУ представлен на рисунке 4.

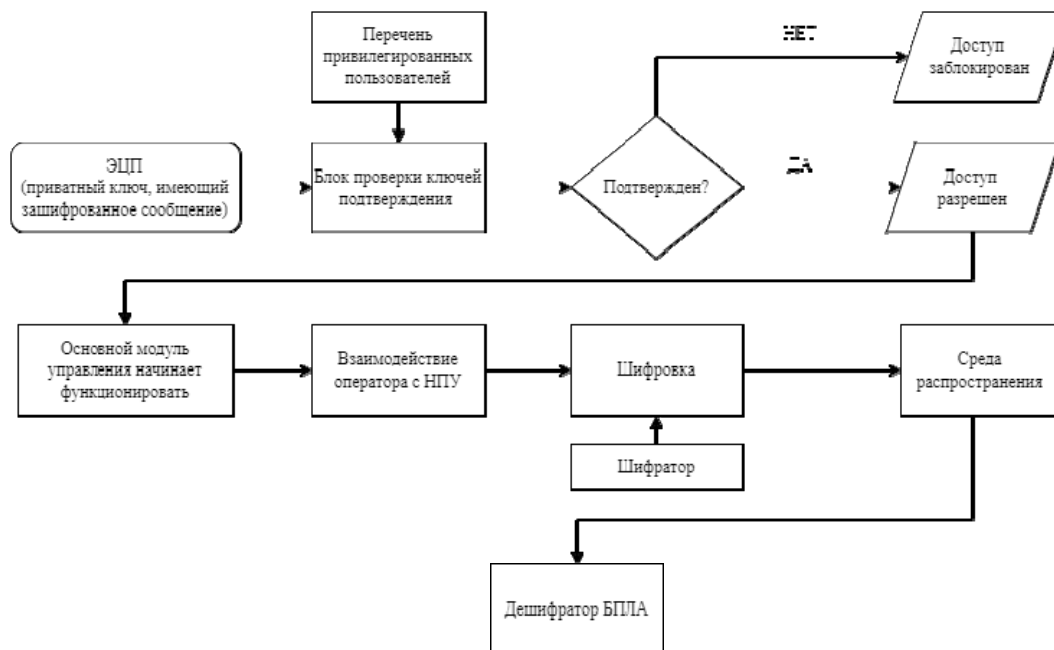


Рис. 4. Алгоритм отправки информации с НПУ

Суть метода заключается в том, что модуль носителя ключа ЭП содержит ключ для основной криптосистемы, который зашифровывает показатели, выставленные на НПУ оператором. В таком случае, данное средство не должно попадать в руки злоумышленника.

БПЛА гражданского назначения уязвимы к перехвату в воздухе. Например, с помощью специальных сетей, которые выстреливаются специальной установкой (данные сетки в полете раскрываются и движутся на встречу беспилотнику) БПЛА можно перехватить практически без каких-либо повреждений и впоследствии эксплуатировать его в своих нуждах [4]. Такому физическому методу, к сожалению, в настоящее время нельзя ничего противопоставить. При этом установленное криптографическое оборудование остается на борту, и чтобы злоумышленники не могли исследовать вышеупомянутые системы, необходимо разработать систему самоуничтожения этих модулей.

В основе модуля носителя ключевой информации ЭП лежит физический USB-носитель, для уничтожения данных, на которых уже существуют соответствующие устройства. К примеру, в 2021 году специалисты холдинга "Технодинамика" (входит в госкорпорацию "Ростех") разработали флеш-накопитель, ликвидирующий информацию нажатием кнопки [5]. При необходимости уничтожить информацию с флеш-накопителя пользователь нажимает на кнопку на торце устройства без взаимодействия с компьютером. Встроенный источник питания активирует электродетонатор, который прожигает печатную плату кумулятивной струей. Целостность корпуса при этом не нарушается, что гарантирует безопасность пользователя. В случае БПЛА такая технология не подойдет. Поэтому, необходимо использовать модуль ключа ЭП с дистанционным уничтожением.

В модуле должно быть предусмотрено два варианта ликвидации, при первом варианте используется возможность подачи команды на уничтожение оператором специально предусмотренного тумблера на наземном пульте управления беспилотным летательным аппаратом. Во втором - система самоуничтожения на БПЛА. Беспилотник самостоятельно анализирует происходящее вокруг него на основе телеметрии и показателей бортовых систем и в случае, если полет проходит не штатно (замечает перехват «сеткой»), то система подаёт сигнал на прожигание модуля ЭЦП. В таком случае, злоумышленник не сможет скомпрометировать реализованную криптографическую систему защиты информации.

Актуальным также является вопрос выбора криптографических алгоритмов для реализации системы защиты информации. БПЛА требуется быстрый отклик от пульта управления. Для обеспечения необходимого быстродействия системы управления не все алгоритмы подходят для реализации. Основным средством защиты каналов связи в системах, подобных БПЛА, на данный момент являются алгоритмы низкоресурсной криптографии (в соответствии с терминологией [6]). Особенно активно и наиболее продуктивно, в соответствии с [8], в области низкоресурсных криптографических алгоритмов в настоящее время исследованы алгоритмы блочного шифрования [7]. За последнее десятилетие было предложено достаточное количество подобных решений. Данное направление криптографических исследований развивалось двумя основными способами [9]. В первом случае разработки велись в виде эффективной программно-аппаратной реализации распространённых блочных алгоритмов, а во втором – в виде создания новых блочных шифров, изначально предназначенных для эффективной реализации на уровне микропрограмм, а также на аппаратном уровне. Результаты ряда исследований показывают, что одним из наиболее компактных в реализации и криптографически стойких является алгоритм Магма, ранее называемый ГОСТ 28147-89. Поэтому для создания системы шифрования для сигналов, передаваемых от оператора до БПЛА, с учетом отечественного законодательства ее использование наиболее оптимально.

Для систем ЭП подобных стандартизированных в Российской Федерации низкоресурсных алгоритмов в настоящее время нет. Для реализации ЭП в следующем разделе данной работы один из авторов выбрал метод RSA, но если беспилотник будет использоваться в государственных структурах или в коммерческих структурах, которые задействованы, к примеру, в государственном заказе [10], то для реализации криптосистемы следует использовать другой алгоритм. В Российской Федерации для ЭП имеется государственный стандарт – ГОСТ 34.10-2018.

Заметим, что в случае реализации квантовой угрозы в системах защиты информации в каналах управления БПЛА возникнет необходимость замены классических асимметричных алгоритмов на квантово-устойчивые, выбор которых может осуществляться с учетом информации из работ [11], [12, 16, 17].

Обеспечение безопасного канала связи

В качестве примера зашифровываемого показателя бортовой системы БПЛА выберем показатель для управления первым двигателем ENGINE_1. Для передачи команд для него необходимо выполнить ряд вычислений, первое из которых – нахождение кода целостности [13] сообщения, которое необходимо зашифровать. За основу берем алфавит, в котором каждая буква определена своим порядковым номером. В качестве примера сообщения выбираем слово «ПУСК», так как после проверки ЭП пульт и БПЛА получают команду запуска всех систем. Так же подбираем в качестве секретного ключа в системе RSA p , q и начальное значение вспомогательной последовательности H_0 . Дополнительные данные приведены в таблицах 1, 2, 3.

Таблица 1

Алфавит русского языка с порядковыми номерами букв

А	Б	В	Г	Д	Е	Ё	Ж	З	И	Й
1	2	3	4	5	6	7	8	9	10	11
К	Л	М	Н	О	П	Р	С	Т	У	Ф
12	13	14	15	16	17	18	19	20	21	22
Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
23	24	25	26	27	28	29	30	31	32	33

Таблица 2

Значения для вычисления ЭП (p , q и H_0)

Постоянные для режима шифрования RSA		Начальное значение кода целостности
p	q	H_0
67	53	3

Таблица 3

Значения для ЭЦП (ПУСК)

Сообщение	M_i	M_1	M_2	M_3	M_4
	Буква	П	У	С	К
ПУСК	Номер буквы в алфавите	17	21	19	12

1) Определение значение n:

$$n = p * q = 67 * 53 = 3551$$

2) Получение кода целостности сообщения «ПУСК». Используем хеш-функцию

$$H_i = (H_{i-1} + M_i)^2 \% n$$

В Российской Федерации для функции хеширования имеется государственный стандарт – ГОСТ 34.11-2018.

$$H_1 = (H_0 + M_1)^2 \% n = (3 + 17)^2 \% 3551 = 400$$

$$H_2 = (H_1 + M_2)^2 \% n = (400 + 21)^2 \% 3551 = 3242$$

$$H_3 = (H_2 + M_3)^2 \% n = (3242 + 19)^2 \% 3551 = 2427$$

$$H_4 = (H_3 + M_4)^2 \% n = (2427 + 12)^2 \% 3551 = 796$$

Полученный код целостности: 796.

При нахождении значения ЭП необходимо учесть, что исходное значение кода целостности и код целостности, найденный в результате проверки, должны совпадать, тогда подпись будет подлинной. При этом модуль при значениях p и q не должны быть меньше значения исходного кода целостности. Из этого есть вероятность того, что подпись может не сойтись при проверке. Поэтому, для использования ЭП нужно брать такие числа p и q, чтобы их модуль был больше или равен модулю при вычислении кода, к примеру, как приведенные в таблице 4.

Таблица 4

Константы для ЭЦП

Полученный код целостности:	Постоянные для ЭЦП RSA	
	p	q
796	43	19

3) Находим mod.

$$mod = p * q = 43 * 19 = 817$$

4) Значение функции Эйлера (формула 4):

$$\varphi = (p - 1) * (q - 1) =$$

$$(43 - 1) * (19 - 1) = 756$$

5) Подберем открытую экспоненту e:

$$(d * e) \bmod \varphi = 1 ; (5 * 605) \bmod 756 = 1$$

Таблица 5

Ключи шифрования

Генерация секретного ключа шифрования
d – Число обратное e по модулю ϕ . d взаимно простое с ϕ ; $d < \phi$; $d > 1$ Секретный ключ: $\{d, \text{mod}\} = \{5, 817\}$
Генерация открытого ключа шифрования
Открытый ключ: $\{e, \text{mod}\} = \{605, 817\}$

б) Вычисление ЭП.

Электронная цифровая подпись вычисляется по правилу:

$$s = n^d \% \text{mod}, \text{ где } n - \text{код целостности.}$$

$$s = 796^5 \% 817 = 82$$

Для проверки ЭП, используем открытый ключ (605, 756), находим Н:

$$H = 82^{605} \% 817 = 796$$

Наблюдаем, что n и H сошлись, значит, ЭП является подлинной и может использоваться в дальнейших алгоритмах шифрования для обеспечения защиты каналов управления беспилотных летательных аппаратов.

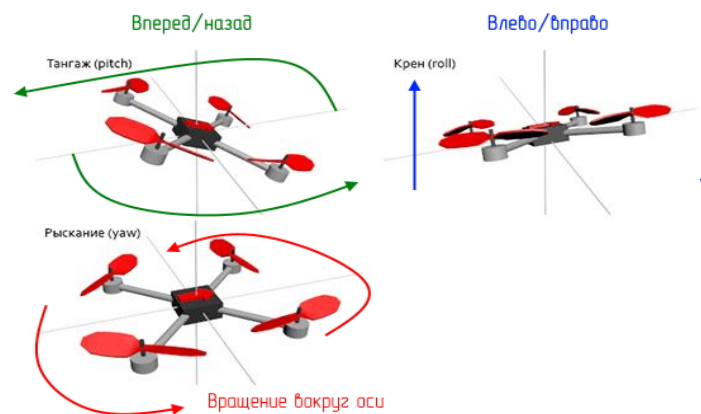


Рис. 5. Параметры полета БПЛА.

Перед тем, как провести расчет основных показателей, необходимо понять команды, которые будет выполнять беспилотник. На рисунке 5 показаны перемещения беспилотника по осям.

В таблицах 6 и 7 приведены команды для управления двигателями и рулем высоты БПЛА.

Таблица 6

Исходные данные для шифрования каждой системы БПЛА

Двигатель	Элероны	
	Влево	Вправо
ENGINE_1	ELRLFT_1	ELRRHT_1
ENGINE_3	ELRLFT_2	ELRRHT_2
ENGINE_2	ELRLFT_3	ELRRHT_3
ENGINE_4	ELRLFT_4	ELRRHT_4
ENGINE_5	ELRLFT_5	ELRRHT_5
ENGINE_6	ELRLFT_6	ELRRHT_6
ENGINE_7	ELRLFT_7	ELRRHT_7
ENGINE_8	ELRLFT_8	ELRRHT_8

Таблица 7

Дополнительные исходные данные для шифрования каждой системы БПЛА

Руль высоты	
Вверх	Вниз
ALTDUP_1	ALDOWN_1
ALTDUP_2	ALDOWN_2
ALTDUP_3	ALDOWN_3
ALTDUP_4	ALDOWN_4
ALTDUP_5	ALDOWN_5
ALTDUP_6	ALDOWN_6
ALTDUP_7	ALDOWN_7
ALTDUP_8	ALDOWN_8

Начинаем передачу зашифрованных вышеперечисленных показателей. Для этого потребуется выбранный ранее алгоритм шифрования Магма и таблица с символами английского алфавита (ASCII). За основу берем ASCII, в которой каждая буква определена своим двоичным значением. Для ключа в алгоритме Магма используем всё тоже зашифрованное сообщение «ПУСК», но с некоторыми доработками. Код целостности 796 дополнительно подвергнем операции хеширования. Сделано это для того, чтобы в случае скомпрометированной ключевой системы ЭП злоумышленник не смог получить ключ от основной криптосистемы, используемой для шифрования связи БПЛА.

Таким образом, получим: $796 * 6 = 4776$ (пусть 4776 – число W), следовательно, $W = H * j$, где j – случайное число, используемое в пульте управления и в БПЛА.

Теперь данное число преобразуем в буквы, для этого используем фрагмент английского алфавита с порядковыми номерами букв. Имеется 9 разных английских букв, которые имеют порядок, близкий в некотором смысле к случайному в соответствии с приведенной ниже таблицей 8. Процесс подготовки ключа представлен в таблице 9, также приведённой ниже

Таблица 8

Фрагмент англ. алфавита с порядковыми номерами букв

A	Z	K	D	H	R	F	P	Z
1	2	3	4	5	6	7	8	9

Таблица 9

Подготовка ключа для ГОСТ 28147-89.

M_i'	M_1'	M_2'	M_3'	M_4'
Номер буквы в алфавите	4	7	7	6
Буква (Ключ)	D	F	F	R

Для ключа X используем: DFFR

Для перевода используем кодировку алфавита и дополнительных символов. В качестве демонстрации шифрования и дешифрования в рамках данного примера выберем один показатель определенной системы, для этого, как было указано выше, выберем ENGINE_1, поскольку данный показатель является основным для работы двигателя, и в тоже время является холостым ходом двигателя БПЛА. Переводим исходный текст (ENGINE_1) и ключ в двоичную последовательность (рисунок 6).

E	N	G	I
0 1 0 0 0 1 0 1	0 1 0 0 1 1 1 0	0 1 0 0 0 1 1 1	0 1 0 0 1 0 0 1
N	E	-	1
0 1 0 0 1 1 1 0	0 1 0 0 0 1 0 1	0 0 1 0 0 0 0 0	0 0 1 1 0 0 0 1

Рис. 6. Значение системы ENGINE_1 в двоичной последовательности

L: 01000101010011100100011101001001

R: 01001110010001010101111100110001

Ключ: DFFR -> Kx: 01000100 01000110 01000110 0101 0010

После перевода холостого положения двигателя (ENGINE 1) в двоичную последовательность, необходимо выполнить операцию побитового сложения по mod 2^{32} R и Kx.

1. Выполнение операции побитового сложения по mod 2^{32} R и Kx
2. Выполняем 8 операций подстановки с помощью приведенных в таблице 10 S-блоков.

Таблица 10
S-блоки

	8	7	6	5	4	3	2	1
0	1	13	4	6	7	5	14	4
1	15	11	11	12	13	8	11	10
2	13	4	10	7	10	1	4	9
3	0	1	0	1	1	13	12	2
4	5	3	7	5	0	10	6	13
5	7	15	2	15	8	3	13	8
6	10	5	1	13	9	4	15	0
7	4	9	13	8	15	2	10	14
8	9	0	3	4	14	14	2	6
9	2	10	6	10	4	15	3	11
10	3	14	8	9	6	12	8	1
11	14	7	5	14	12	7	1	12
12	6	6	9	0	11	6	0	7
13	11	8	12	3	2	0	7	15
14	8	2	15	11	5	9	5	5
15	12	12	14	2	3	11	9	3

S8	S7	S6	S5	S4	S3	S2	S1
1001	0010	1000	1011	1010	0101	1000	0011
9	2	8	11	10	5	8	3
↓	↓	↓	↓	↓	↓	↓	↓
S8	S7	S6	S5	S4	S3	S2	S1
2	4	3	14	6	3	2	2
0010	0100	0011	1110	0110	0011	0010	0010

Рис. 7. Перестановка

После перестановки, представленной на рисунке 6, получим R: 00100100001
111100110001100100010

3. Следующий шаг - циклический сдвиг на 11 бит влево.

R: 00100100001111100110001100100010 – Значение после перестановки.

11110011000110010001000100100001 – Значение после циклического сдвига.

4. Выполнение XOR операции L и R

L: 01000101010011100100011101001001

R: 11110011000110010001000100100001
10110110010101110101011001101000

5. R присваиваем значение полученное действием выше:

R: 10110110010101110101011001101000

6. Присвоение L исходного значения R.

L: 01001110010001010101 111100110001

7. Результат первого цикла алгоритма ГОСТ 28147-89.

0100111001000101010111110011000110110110010101110101011001101000

Подобным образом проделываем до 32 цикла алгоритма шифрования.

Результат тридцать второго цикла алгоритма ГОСТ 28147-89:

1111000010011111001010011100110000000011011101100010010010000011 – зашифрованное сообщение ENGINE_1 в наземном пульте управления БПЛА.

Зашифрованное сообщение передается передатчиком по среде распространения и попадает на приёмник на БПЛА. Приняв данный зашифрованный показатель, приёмник передает информацию в дешифратор, после чего происходит расшифрование в исходный переданный показатель с наземного пульта управления, в данном случае это ENGINE_1. Производим расшифрование аналитическим путем, чтобы в дальнейшем сравнить все результаты циклов, и удостоверится в корректной работе программного обеспечения.

Дешифрование полученного сигнала на борту БПЛА проводится в обратном порядке. В системе имеется ключ Кх: 01000100 01000110 01000110 0101 0010 и результат 32 цикла – 11110000100111110010100111001100 00000011011101100010010010000011.

Имеется левая и правая часть 32 цикла. Левая часть – это исходные данные правой части начальных данных цикла. Правую часть получаем из операции XOR, в которой нам известны L и R. Так как R нам известно из полученного сообщения, то для получения исходных данных цикла необходимо найти значение L. Чтобы найти L, необходимо провести операции побитового сложения, выполнить операцию подстановки с помощью S-блоков и произвести циклический сдвиг на 11 бит влево. После чего появится возможность найти L, имея R и результат операции XOR. По факту, алгоритм шифрования и дешифрования схожи, единственное отличие в операции XOR.

Таким образом, проведя все манипуляции с шифрованием получаем защищенный канал связи, к которому не сможет подключиться злоумышленник, не обладающий большими вычислительными возможностями. Залогом этого является использование надежного симметрического алгоритма шифрования, прошедшего многочисленные тематические исследования, к примеру [14], [15].

Заключение

В результате проведенных исследований был выявлен ряд уязвимостей у беспилотников гражданского назначения, связанные с возможностью их физического захвата и вмешательства злоумышленника в незащищенные каналы управления. В связи с этим, был разработан метод для защиты каналов управления и метод защиты ключевой информации от физического доступа. Используя комбинацию данных методов, БПЛА будут максимально защищены от несанкционированного доступа злоумышленников. В отдельном разделе приведен пример моделирования работы программного обеспечения, предназначенного для защиты каналов управления.

Одним из положительных моментов данного метода защиты, является гибкость к применению различных криптографических систем и механизмов. В случае экспорта программного обеспечения за границу, в перспективе имеется возможность реализовать необходимый зарубежный алгоритм шифрования, учитывающий требования сертифицирующих органов соответствующего государства, что позволит обезопасить от действий злоумышленников БПЛА гражданского назначения других стран в соответствии с их законодательством.

Литература

1. БПЛА в гражданских областях хозяйства [Электронный ресурс] // IXBT – Режим доступа: <https://dfnc.ru/aviazcia/bespilotniki-i-terroristicheskaya-ugroza/> (дата обращения: 30.01.2023).
2. Будаев Д.С., Вошук Г.Ю., Гусев Н.А., Майоров И.В., Мочалкин А.Н. Особенности построения программно-аппаратного комплекса для согласованного выполнения задач группой беспилотных аппаратов на базе мультиагентных технологий и сетецентрического подхода // Труды 3-й Всероссийской научнотехнической конференции «РТИ Системы ВКО - 2015», 28 мая 2015 г., Москва. С. 486-505.
3. Журавлев Д.О., Зау Х.Н. Эволюция систем управления беспилотных летательных аппаратов: от появления до наших дней // Достижения и перспективы современной науки : Материалы Международной (заочной) научно-практической конференции, Астана, Казахстан, 07 февраля 2017 года / Научно-издательский центр «Мир науки» (г. Нефтекамск, Республика Башкортостан, Российская Федерация), Баспасы «Академия» (г. Астана, Казахстан); под общей редакцией А.И. Вострецова. С. 57-87.
4. Рябов К. Клин клином: беспилотники-перехватчики и их возможности. 01.08.2019 [Электронный ресурс] // Военное обозрение. Режим доступа: <https://topwar.ru/160756-klin-klinom-bespilotniki-perehvatchiki-i-ih-vozmozhnosti.html> (дата обращения: 30.01.2023).
5. Ростех создал самоуничтожающийся флеш-накопитель. 25.11.2021 [Электронный ресурс] // ТАСС – Режим доступа: <https://tass.ru/ekonomika/13023135> (дата обращения: 30.01.2023).

6. Предварительный национальный стандарт Российской Федерации. Информационные технологии. Криптографическая защита информации. Термины и определения. М.: Российский институт стандартизации, 2022. 50 с.

7. *Панков К.Н.* Основные блочные алгоритмы шифрования, предназначенные для обеспечения информационной безопасности в системе интернет-вещей // Технологии информационного общества : Материалы XIII Международной отраслевой научно-технической конференции, Москва, 20-21 марта 2019 года. Том 1. С. 458-460.

8. *Жуков А.Е.* Легковесная криптография. Часть 2 // Вопросы кибербезопасности. 2015. № 2(10). С. 2-10.

9. *Жуков А.Е.* Легковесная криптография. Часть 1 // Вопросы кибербезопасности. 2015. № 1(9). С. 26-43.

10. *Панков К.Н., Эйман А.Д.* Сертификация систем распределенного реестра как инструмент обеспечения информационной безопасности // REDS: Телекоммуникационные устройства и системы. 2021. Т. 11. № 2. С. 37-49.

11. *Pankov K.N., Glukhov M.M.* Using Error-Correcting Codes to Ensure Information Security of Unmanned Vehicles and IoT Systems // 2022 Systems of Signal Synchronization, Generating and Processing in Telecommunications (SYNCHROINFO), Arkhangelsk, Russian Federation, 2022, pp. 1-8.

12. *Панков К.Н., Миронов Ю.Б.* Использование постквантовых алгоритмов в задачах защиты информации в телекоммуникационных системах. М.: Горячая линия – Телеком, 2023, 236 с.

13. *Лось А.Б., Нестеренко А.Ю., Рожков М.И.* Криптографические методы защиты информации: учебник для академического бакалавриата. 2-е изд., испр. М.: Юрайт, 2016. 473 с.

14. *Алексеев Д.М.* Алгоритм шифрования Магма: оценка криптостойкости шифра с использованием линейного и слайдового методов анализа // Международный студенческий научный вестник. 2017. № 4-4. С. 454-459.

15. *Артемов В.Р., Фионов А.Н.* Статистический анализ шифра «Магма» // Интеллектуальный потенциал Сибири : 30-я Региональная научная студенческая конференция. Материалы конференции: в 4 частях, Новосибирск, 23-27 мая 2022 г. С. 8-11.

16. *Панков К.Н.* Оценки мощности классов отображений, применяемых в протоколах квантового распределения ключей // Научные технологии в космических исследованиях Земли. 2022. Т. 14. № 4. С. 4-18.

17. *Миронов Ю.Б., Казанцев С.Ю., Шаховой Р.А., Колесников О.В., Машковцева Л.С., Зайцев А.И., Коробов А.В.* Анализ перспектив развития источников одиночных фотонов в системах квантового распределения ключей // Научные технологии в космических исследованиях Земли. 2021. Т. 13. № 6. С. 22-33.

КОНТЕНТНЫЙ ВЗГЛЯД НА СЕТИ

Саломатина Елена Васильевна,

ПГУ им. Т. Г. Шевченко, доцент, к.т.н., Тирасполь, Приднестровье
salolew@spsu.ru

Сарьян Вильям Карпович,

ФГУП НИИР, академик НАН РА, профессор, д.т.н., Москва, Россия
sarian@niir.ru

Аннотация

В настоящее время технологии развиваются, ориентируясь на требование гарантированной доставки разнообразного контента за минимальное время и с высокой скоростью. Существующие сетевые решения и рабочие сценарии могут не справиться с лавинообразным увеличением объема трафика. В статье рассмотрены контент-ориентированные сетевые парадигмы, которые нацелены на эффективное распространение информации и устранение многих сетевых проблем, и представлены возможные направления будущих исследований.

Ключевые слова: *Контент, информационное взаимодействие, облачные вычисления, контент-ориентированные сети, сеть доставки контента, сеть как услуга, NaaS, сетевые парадигмы*

Введение

За последние несколько лет сети претерпели важные изменения, вызванные появлением и быстрым распространением новых услуг. Вместо традиционных одноадресных потоков все большее распространение получают модели произвольной (Anycast, один-к-одному-из-многих) и многоадресной (Multicast, один ко многим) передачи. Узкоспециализированные центры обработки данных, распределенные по сети, используются в качестве репозитория для вычислений и хранения данных. Наиболее популярным и объемным контентом, предлагаемым в современных сетях, является видео, которое может быть эффективно доставлено конечным пользователям с помощью многоадресной потоковой передачи. Некоторые тенденции в области ИТ, — аутсорсинг ИТ-услуг; популярность и масштабируемость современного Интернета; развертывание центров обработки данных такими компаниями, как Google, Amazon и Microsoft; технологии виртуализации и растущая потребность в обработке огромных объемов данных, обусловили популярность парадигмы облачных вычислений. Растущая популярность услуг облачных вычислений и сохраняющаяся популярность услуг распределенных вычислений привели к серьезным изменениям в интернет-трафике.

В настоящее время технологии развиваются, ориентируясь на требование доставки разнообразного контента за минимальное время и с высокой скоростью. Эффективная доставка контента по сети стала важным требованием повышения производительности сети как для стационарных, так и для мобильных пользователей. Большое внимание стало уделяться концепции контент-ориентированных сетей, где акцент делается на том, что передается, а не на том, где передается информация.

Сети в значительной степени воспринимаются в основном с точки зрения контента, которым в них обмениваются, а не сетевых элементов, ее составляющих. Основная цель разработки контент-ориентированных сетевых решений заключается в создании сетевой инфраструктуры, которая направлена на распространение контента и более устойчива к сбоям и отказам. Развертывание и реализация новых идей требует множества новых технологических решений и радикальных изменений в существующих сетях.

Модель сетевой архитектуры

Растущий объем интернет-трафика генерируется разнообразными контент-ориентированными приложениями и услугами, включая различные типы электронного контента, такие как потоковое видео и аудио, веб-объекты (текст, графика, музыка), дистрибуция программного обеспечения, потоковая передача мультимедиа по запросу и обмен файлами. Нынешняя архитектура Интернет была основана на хост-ориентированной модели, которая удовлетворяла коммуникационные потребности того времени, когда сеть была необходима для совместного использования ограниченных ресурсов

(периферийные устройства, мейнфреймы и каналы дальней связи). С развитием Интернета возникло большое количество инноваций как в приложениях и службах, работающих поверх него, так и в технологиях ниже (меж)сетевого уровня. Сейчас большинство пользователей в основном заинтересованы в доступе к огромному количеству информации, независимо от ее физического местоположения.

Общие проблемы сети, такие как асимметричная скорость полосы пропускания, масштабируемость и переменная задержка, зависящая от перегрузок и расстояния передачи, тормозят развертывание новых услуг. Если буферизация и пикселизация влияют на качество видеопотока, то задержка может оказаться критичной для поддержки принятия сложных решений или решения сложных проблем. Эффективное реагирование на кризисную ситуацию требует актуальной информации и обращения к новым, внешним источникам данных. Однако новые данные приводят к новым проблемам: их объем, достоверность, разнообразие или скорость не могут управляться только людьми, особенно в режиме реального времени и в условиях высокого стресса. Архитектура, управляемая событиями (event-driven architecture, EDA) [1] гарантирует, что при возникновении события информация об этом событии будет отправлена всем системам и людям, которые в ней нуждаются. Чем быстрее можно получать информацию о событиях там, где они должны быть, тем эффективнее можно реагировать на возможности и перераспределять ресурсы.

Историческая модель Интернет «клиент-сервер» претерпевает постепенный переход к сети, управляемой контентом [2]. Облачные провайдеры, провайдеры OTT (Over the Top) и CDN (Content Delivery Network) способствуют повсеместному развертыванию крупномасштабных центров обработки данных (Massive Scale Data Centers, MSDC), в которых могут размещаться сотни тысяч физических серверов, которые, в свою очередь, могут развертывать огромное количество виртуальных машин и контейнеры, широко используемые для разработки облачных приложений и сервисов. Базовая функциональность MSDC включает в себя вычисление, хранение и репликацию данных с использованием обмена сообщениями между серверами по поддерживающей коммуникационной инфраструктуре. Стоит отметить, что трафик между серверами составляет до 85% от общего объема трафика [3].

Передача пакетов в таких инфраструктурах требует высокого уровня масштабируемости и пропускной способности, намного превышающей исторический трафик глобальной сети Интернет. Эти требования привели к изменению моделей сетевого трафика и способов предоставления различных сетевых услуг конечным пользователям. Гибкость современных сетей обеспечивает возможность быстрого развертывания новых услуг. Фактически, различные сетевые уровни могут использоваться для реализации услуг с учетом технических, социальных, управленческих и экономических аспектов [4]. Предлагаемая в [5] архитектура учитывает наиболее популярные сценарии, обычно используемые для реализации облачных вычислений и контент-ориентированных сервисов. Большинство пользователей для этих целей применяется сетевой уровень (рис. 1).

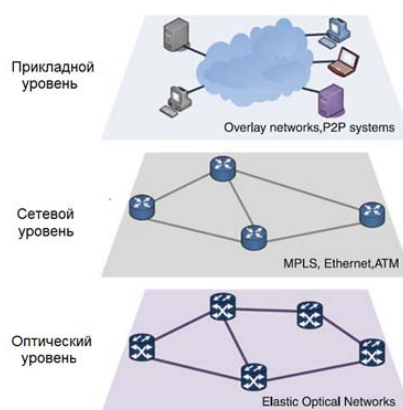


Рис. 1. Модель трехуровневой сетевой архитектуры

На вершине этой архитектуры находится прикладной уровень, который предоставляет облачные вычисления и контент-сервисы с использованием концепции оверлейных сетей. Средний уровень включает основные сетевые транспортные технологии и протоколы, например, MPLS (многопроходная коммутация по меткам), ATM (асинхронный режим передачи) и Ethernet с установлением соединения. Нижний уровень относится к оптической магистральной сети, где реализована эластич-

ная оптическая сеть (EON), используемая для эффективного и гибкого использования сетевых ресурсов [6].

Наиболее эффективным способом предоставления услуг, характеризующихся большими объемами сетевого трафика, связанного с ЦОД и передачей по сети с битрейтом выше 100 Гбит/с, является прямое использование оптического уровня, поскольку более высокие уровни либо не в состоянии обеспечить такие передачи, либо не могут обеспечить их экономически эффективным способом. Центры обработки данных, предоставляющие почти все услуги облачных вычислений и контент-ориентированных услуг, подключены к Интернету через оптические соединения, и оптимизация сетевых ресурсов напрямую на оптическом уровне обеспечивает существенные преимущества, включая сокращение накладных расходов на администрирование.

С другой стороны, для некоторых облачных вычислений и контент-ориентированных услуг, требуются дополнительные свойства и функции, недоступные на сетевом или оптическом уровнях, например, дополнительные требования безопасности, новые протоколы маршрутизации и специальные параметры QoS. Более того, использование сетевого или оптического уровня ограничивает пользователя существующими предложениями телекоммуникационных компаний и интернет-провайдеров, не желающих внедрять новые решения в свои сети или когда стоимость этих новых решений не по карману пользователям. В результате, сетевым службам, доступным на сетевом и оптическом уровнях, может не хватать ожидаемой гибкости и масштабируемости.

Для улучшения функциональности сети с требуемым качеством обслуживания идеальным решением является концепция оверлейных сетей, реализованная на прикладном уровне. Концепцию оверлейных сетей можно классифицировать как метод сетевой виртуализации, который разделяет роли традиционных интернет-провайдеров на две независимые – поставщики инфраструктуры, отвечающие за физическую инфраструктуру, и поставщики услуг, создающие виртуальные сети. Оверлейные сети преодолевают различные ограничения нижних уровней относительно простым и экономичным способом, т.к. оверлейные узлы соединены друг с другом через логические каналы, охватывающих множество физических каналов базовых сетей, с использованием различных технологий и протоколов. Основное преимущество оверлейных сетей заключается в том, что они позволяют разрабатывать и эксплуатировать новые крупномасштабные распределённые сервисы без внесения каких-либо изменений в основные протоколы сети.

Сеть как услуга (NaaS)

Облачные вычисления изменяют способы предоставления и потребления ИТ-услуг, в конечном итоге делая эти услуги более гибкими и экономичными. Облачные услуги обычно предлагаются потребителям в виде трех услуг в соответствии с NIST SP 800-146 [7], а именно:

- Инфраструктура как услуга (IaaS): вычислительные ресурсы предоставляются пользователю, например, оборудование, хранилище, серверы, виртуальные машины, сетевые компоненты, с целью разрабатывать программное обеспечение, в котором он может иметь столько приложений, сколько пожелает. Поставщик является владельцем оборудования и несет ответственность за размещение, эксплуатацию и техническое обслуживание. Пользователи вообще не контролируют физические ресурсы. Основная цель использования модели IaaS — предложить пользователям высокий уровень контроля над ее конфигурацией и использованием.

- Платформа как услуга PaaS: пользователь может использовать для разработки, тестирования и запуска приложений облачную инфраструктуру, которая будет расположена либо в общедоступном облаке, либо в частном облаке, которое поддерживается провайдером, и получать выгоду от использования операционных систем и промежуточного программного обеспечения. Пользователь имеет права на функциональную часть приложения, а не на используемые вычислительные ресурсы. Потенциальным недостатком концепции PaaS является более низкий уровень контроля над базовыми ИТ-ресурсами по сравнению с моделью IaaS.

- Программное обеспечение как услуга (SaaS): приложения размещаются у поставщика и предоставляются пользователям через сеть, обычно через Интернет. Пользователь имеет возможность использовать приложения, которые находятся в облаке, и выполнять в них определенные настройки. У него нет прав на настройку инфраструктуры, в которой находится приложение, такой как серверы и сетевые ресурсы. С одной стороны, обслуживание ИТ-ресурсов значительно упрощается. Однако в то

же время пользователи имеют очень ограниченный административный контроль над внедрением SaaS.

Термин «сеть как услуга» (NaaS) упоминается в [8], как подкатегория вышеперечисленных (рис. 2).

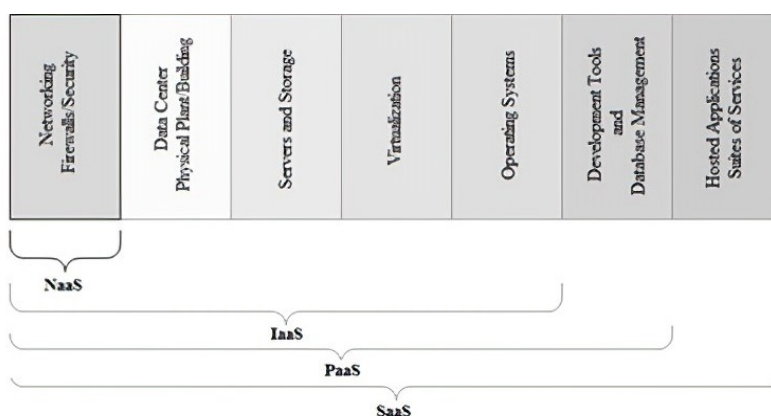


Рис. 2. NaaS как подкатегория

Согласно ITU-TY.3500 [9] NAAS является облачной услугой, в которой возможность, предоставляемая потребителю облачной услуги, относится к возможностям транспортного соединения и связанным с ним сетевым возможностям. NaaS может обеспечивать любые из трех типов облачных возможностей (IaaS, PaaS, SaaS).

NaaS объединяет сетевые ресурсы, услуги и приложения в виде продуктов, которые можно приобрести для ряда пользователей, как правило, на договорной период времени в виде подписки или с оплатой по факту использования.

NaaS включает в себя оптимизацию распределения ресурсов, рассматривая сетевые и вычислительные ресурсы как единое целое. Данная модель объединяет текущие предложения облачных вычислений с прямым, но безопасным доступом пользователя к сетевой инфраструктуре. Используя NaaS, пользователи могут легко развертывать собственные протоколы маршрутизации и многоадресной рассылки. Кроме того, изменяя содержимое пакетов в пути, они могут эффективно реализовывать расширенные сетевые службы, такие как агрегация данных в сети, устранение избыточности и интеллектуальное кэширование.

Сеть как услуга (NaaS) предоставляется крупными поставщиками облачных услуг, включая Amazon и Rackspace, а также глобальными поставщиками услуг, такими как AT&T, Telefonica и Verizon [10]. Совсем недавно в таких областях, как программно-определяемая глобальная сеть (SDWAN), появились специализированные поставщики NaaS, в число которых входят такие игроки, как Aryaka Networks, Cloudflare, Masergy Communications. В нее также входят специализированные сетевые провайдеры, такие как Akamai, у которой есть собственная сеть доставки контента (CDN) для доставки цифровых мультимедиа, а также корпоративные услуги ускорения и безопасности SaaS.

Понятие высокого уровня NaaS с использованием многоуровневой структуры согласно [ITU-T Y.3502] представлено графически на рисунке 3.

- Приложения NaaS: это виртуальные сетевые приложения или функции, предоставляемые поставщиком NaaS. Примеры приложений NaaS включают виртуальный маршрутизатор, виртуальную сеть доставки контента (vCDN), виртуальное расширенное пакетное ядро (vEPC) и виртуальный брандмауэр (vFW). В этой категории провайдер предлагает набор интерфейсов приложений для сетевых операций.

- Платформа NaaS: тип услуги, при которой пользователь может использовать сетевую платформу, предоставляемую поставщиком облачных услуг. Платформа NaaS предлагает одну или несколько сред для разработки программного обеспечения и один или несколько языков программирования для разработки, управления и запуска сетевых приложений, созданных заказчиком. Группы приложений и сетевые функции могут формировать законченное сетевое решение.

• Связность NaaS: тип услуги, при которой поставщик предоставляет возможности инфраструктуры или предлагает сетевые подключения между двумя или более конечными точками, которые включают дополнительные сетевые функции. Это включает, например, гибкую и обширную виртуальную частную сеть (VPN), пропускную способность по запросу (BoD) и многое другое. NaaS может предоставлять базовые сетевые функции, такие как подключение, используя любые физические, логические или виртуальные сетевые возможности, которые предлагает провайдер.

Сетевые возможности могут быть предоставлены клиентам с помощью любой комбинации из трех вышеперечисленных комбинаций облачных возможностей [11].



Рис. 3. Понятие высокого уровня NaaS

Несмотря на то, что модели NaaS находятся в зачаточном состоянии, они открывают огромные перспективы с точки зрения использования, управления и модернизации сетей. Платформы NaaS позволяют разрабатывать приложения, которые распространяются на различные сетевые функции, включая виртуальные маршрутизаторы, дифференциацию услуг и виртуальные брандмауэры. В отличие от других решений, решения NaaS снижают сложность сети и затраты на управление.

Контент-ориентированные сети

В настоящее время концепция контент-ориентированных сетей еще не определена в деталях и дискуссии ведутся по многим техническим аспектам, включая общую модель обслуживания, архитектуру и протоколы, которые будут стандартизированы [12-14]. Поскольку в мире сетей предпочтение отдается эволюционному прогрессу, популярным решением является концепция сетей доставки контента (CDN), рассматриваемая как шаг к полной реализации контент-ориентированной сети. Как показано на рисунке 4, значительная часть трафика в настоящее время доставляется CDN, и по прогнозу Cisco эта доля будет только расти [15].

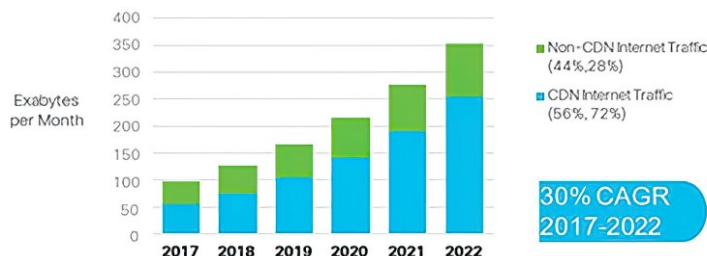


Рис. 4. Ежегодный рост интернет-трафика

Сеть доставки контента (CDN) – это группа географически рассредоточенных серверов, взаимодействующая для обеспечения своевременного и эффективного распространения информации [16] и обеспечивающая максимально возможный уровень разнообразных параметров QoS, включая задержку

ку, пропускную способность и доступность.

Хотя на рынке существует множество провайдеров сетей доставки контента, существует два основных типа CDN в зависимости от того, как они извлекают контент и предоставляют его пользователю: push CDN и pull CDN. В первом случае поставщик контента несет ответственность за контент, который CDN будет доставлять клиентам. Исходный сервер автоматически или вручную отправляет контент непосредственно на серверы CDN и ссылается на него. Этот контент кэшируется на сервере, пока не будет удален или обновлен. Основным преимуществом этого подхода является гибкость, поскольку контент-провайдер решает, какое содержимое будет отправлено на сервер CDN, когда истечет срок действия и когда его необходимо обновить. Если исходный сервер отключится по какой-либо причине, пользователи все равно смогут получить доступ к кэшированным данным. Контент загружается на сервер CDN только тогда, когда происходят какие-либо изменения, что снижает трафик.

Во втором случае контент-провайдер просто переписывает URL-адреса контента в CDN, а затем CDN берет на себя ответственность за загрузку контента в сеть и его кэширование до истечения срока его действия. Основное преимущество подхода заключается в минимальных административных затратах, так как CDN полностью управляет процессом. Кроме того, поскольку он извлекает контент только после того, как сделан первоначальный запрос, это сводит к минимуму потребление памяти.

В [17] представлена гибридная двухтактная модель, которая динамически адаптируется к изменяющимся запросам конечных пользователей, проактивно продвигая одно содержимое и реагируя на другое.

Однако независимо от модели доставки схема трафика CDN одинакова. По сути, каждая необходимая часть информации отправляется один раз на каждый сетевой узел, на котором размещен сервер CDN. Затем пользователи запрашивают получение контента с одного из серверов CDN. Если первая фаза обычно не вносит большого вклада в общий сетевой трафик, то вторая фаза – в зависимости от популярности контента – может генерировать очень большие объемы трафика.

То, что один и тот же контент, запрошенный конечными пользователями, реплицируется в различных местах, географически разбросанных по сети, дает возможность использовать передачу Anycast как лучший способ обеспечить рентабельность и высокую отказоустойчивость доставки контента. Более того, если контент должен быть доставлен нескольким пользователям одновременно (например, потоковая передача видео/аудио в реальном времени), наиболее предпочтительным подходом к передаче является Multicast передача, поскольку она сводит к минимуму использование пропускной способности сети и обеспечивает высокую масштабируемость.

Применение CDN дает такие важные преимущества как снижение перегрузки сети, более быстрая доставка запрошенного контента и оптимизация взаимодействия с пользователем, но лучшего результата можно добиться, используя сотрудничество CDN с различными сетевыми парадигмами, например, туманные вычисления, Интернет вещей, программно-определяемая сеть. Каждая из этих технологий направлена на совершенствование сетевой архитектуры в соответствии с растущими требованиями пользователей.

Заключение

Рассмотренные концепции являются одними из наиболее важных технологий, которые в последнее время вызывают широкий интерес. Данные Synergy Research Group показывают, что на основных рынках общедоступных облачных услуг и инфраструктуры доходы операторов и поставщиков за первый квартал 2022 года достигли 126 миллиардов долларов, увеличившись на 26% по сравнению с первым кварталом 2021 года. В других основных сегментах услуг, управляемых частными облачными сервисами, корпоративных SaaS и CDN, доходы от услуг увеличились на 54 миллиарда долларов, увеличившись в среднем на 21% по сравнению с прошлым годом [18]. На рисунке 5 представлена динамика ранга облачных экосистем и лидеры данного сектора.

Сохраняется тенденция, когда крупные поставщики веб-контента, как Facebook, Google, Alibaba, Microsoft, Amazon создают свои собственные CDN, предлагая своим клиентам высококачественный доступ к контенту. И их доля на рынке услуг только растет. С начала 2017 года совокупная рыночная доля европейских облачных игроков, включая SAP, Deutsche Telekom, OVHcloud, Telecom Italia, Orange, упала с 27% до 13% на их домашней территории. Напротив, на Amazon, Google и Microsoft сейчас приходится 72% регионального рынка [19, 20]. На всех рынках услуг и инфраструктуры по-

давяющее большинство ведущих игроков составляют американские компании (44% всех доходов от облачных услуг и 51% мощностей гипермасштабируемых центров обработки данных), а лидеры среди оставшихся — китайские компании (8% и 15% соответственно).

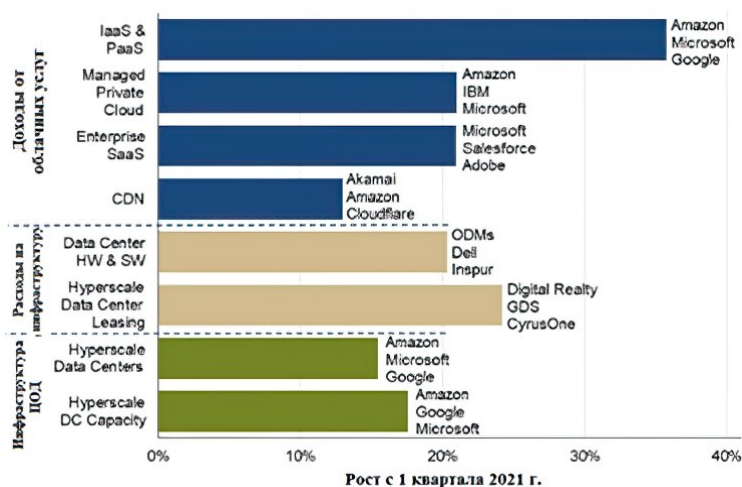


Рис. 5. Рост рынка публичных облачных экосистем и лидеры — первый квартал 2022 г.

Появление облачных вычислений и растущая динамика создания и распространения информации через Интернет привели к представлению Интернет как источника информации и услуг, которые предоставляются им по сети. Будущий Интернет сталкивается с проблемой информационной перегрузки, и существующие сетевые решения могут не справиться с быстрым увеличением объема доступной информации. Экспоненциальный рост сетевых технологий привел к значительному расширению масштабов подключенной вычислительной среды.

Внедрение Интернета вещей связано с вводом в единую инфотелекоммуникационную среду громадного количества взаимодействующих инертных и живых объектов и сыграло важную роль в координации физических и киберсистем. В принципе все эти объекты могут вступать в глобальное информационное взаимодействие. Киберфизические системы и повсеместное распространение датчиков, устройств и встроенных систем привели не только к объединению взаимодействующих в реальном времени объектов в сеть, а к переходу к объединению информации в сеть. Это требует перехода к сетевой парадигме на основе контента и открывает дверь к целому ряду инновационных возможностей.

Литература

1. Michelson B.M. Event-driven architecture overview // Patricia Seybold Group. 2006. Т. 2. №. 12. С. 10-1571.
2. Khriji S. et al. Design and implementation of a cloud-based event-driven architecture for real-time data processing in wireless sensor networks // The Journal of Supercomputing. 2022. С. 1-28.
3. Саломатина Е.В. Разработка моделей телекоммуникационных информационно-управленческих сетей и методов их эффективного использования: дис. ... канд. техн. наук: 05.12.13. - Владим. гос. университет, Владимир, 2019 - 192 с.
4. Cisco. Cisco Global Cloud Index: Forecast and Methodology, 2016–2021; White Paper; 2018. [Электронный ресурс] URL: https://virtualization.network/Resources/Whitepapers/0b75cf2e-0c53-4891-918e-b542a5d364c5_whitepaper-c11-738085.pdf (дата обращения 28.01.2023)
5. Сарьян В.К., Левашов В.К., Мецераков Р.В. Технические, социальные и управленческие аспекты становления гиперсвязанного мира // Коэволюция техники и общества в контексте цифровой эпохи: Сборник докладов, Москва, 17-18 декабря 2020 года / Под общей редакцией А.Л. Андреева, З.К. Селивановой, В.И. Герасимова. М.: Издательский дом МЭИ, 2020. С. 12-14.
6. Walkowiak K. Modeling and optimization of cloud-ready and content-oriented networks. Berlin: Springer, 2016. Т. 56.
7. Белянков Д.А. Возможности искусственного интеллекта и машинного обучения в телекоммуникационных сетях // Телекоммуникации: сети и технологии, алгебраическое кодирование и безопасность данных = Telecommunications: Networks and Technologies, Algebraic Coding and Data Security : материалы международного

научно-технического семинара, Минск, ноябрь-декабрь 2019 г. Белорусский государственный университет информатики и радиоэлектроники; ред. Бобов М. Н. [и др.]. Минск: БГУИР, 2019. С. 118-124.

8. Cloud Computing Synopsis and Recommendations [Электронный ресурс] URL: <https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-146.pdf> (дата обращения 28.01.2023)

9. *Barry D.K., Dick D.* Web services, service-oriented architectures, and cloud computing: the savvy manager's guide. 2nd ed. Waltham: Elsevier: Morgan Kaufmann, 2013. 226 p.

10. Recommendation Y.3500 «Information technology – Cloud computing – Overview and vocabulary» ITU-T, Geneva. August 2014.

11. Network As A Service (Naas) Market: Profiling Key Players to Witness Massive Revenue Growth Over Forecast Period of 2022 to 2028 [Электронный ресурс] URL: <https://www.digitaljournal.com/pr/network-as-a-service-naas-market-profiling-key-players-to-witness-massive-revenue-growth-over-forecast-period-of-2022-to-2028-att-verizon-telefonica-vodafone-bt-group-tata-communications-lu> (дата обращения 28.01.2023)

12. Recommendation Y.3512 «Cloud computing - Functional requirements of Network as a Service» ITU-T, Geneva. August 2014.

13. *Аджемов А.С., Ефимушкин В.А.* Инновационные подходы мировых операторов связи к развитию контент-ориентированных услуг // Электросвязь. 2011. № 9. С. 20-23.

14. *Perino D., Varvello M.* A reality check for content centric networking // Proceedings of the ACM SIGCOMM workshop on Information-centric networking. 2011. С. 44-49.

15. *Li Y., Wang X.* Content-oriented Framework for Future Internet // Wireless Personal Communications. 2021. Т. 120. С. 869-886.

16. Cisco VNI Global IP Traffic Forecast, 2017-2022 [Электронный ресурс] URL: <https://twiki.cern.ch/twiki/pub/HEPIX/TechwatchNetwork/HtwNetworkDocuments/white-paper-c11-741490.pdf> (дата обращения 28.01.2023)

17. *Held G.* A practical guide to content delivery networks. Auerbach Publications, 2005.

18. *Guan X., Choi B. Y.* Push or pull? toward optimal content delivery using cloud storage // Journal of Network and Computer Applications. 2014. Т. 40. С. 234-243.

19. Public Cloud Ecosystem Quarterly Revenues Leap 26% to \$126 Billion in Q1 [Электронный ресурс] URL: <https://www.srgresearch.com/articles/public-cloud-ecosystem-quarterly-revenues-leap-26-to-126-billion-in-q1> (дата обращения 28.01.2023)

20. European Cloud Providers Continue to Grow but Still Lose Market Share [Электронный ресурс] URL: <https://www.srgresearch.com/articles/european-cloud-providers-continue-to-grow-but-still-lose-market-share> (дата обращения 28.01.2023)

КЛАССИФИКАЦИЯ И СРАВНИТЕЛЬНЫЙ АНАЛИЗ МЕТОДОВ ТРЕХМЕРНОЙ РЕКОНСТРУКЦИИ ОБЪЕКТОВ

Силантьева Анна Сергеевна,

Московский Технический Университет Связи и Информатики, аспирант, Москва, Россия

ann-august@yandex.ru

Аннотация

Высококачественная трехмерная реконструкция сцены или объекта является фундаментальной и сложной задачей в разных областях использования этой современной технологии. 3D-реконструкция имеет множество применений, таких как 3D-рендеринг местности (Wu et al., 2015), виртуальная реальность, навигация роботов, задачи дополненной реальности, игры, анимация и кинофильмы. В области медицины он используется в минимально инвазивных хирургических методах, 3D-визуализации анатомии пациентов, операциях с компьютерным управлением, предоперационном планировании и так далее. Цель бесконтактного восстановления 3D-сцены или предмета достижима путем сканирования или получения из набора изображений. Собранный классификации бесконтактных методов на сегодняшний день не опубликовано и целесообразно заключить сформированные методы в единой статье с описанием их основных принципов и особенностей. Задача формирования трехмерного пространства или объекта с двумерных (2D) видов/изображений называется реконструкцией из 2D в 3D, в то время как получение посредством специальных устройств – трехмерным сканированием. В литературе доступно большое разнообразие подходов к 3D-реконструкции. Эти подходы призваны обеспечить хороший компромисс между скоростью вычислений, сложностью и точностью наряду с осуществимостью реализации. Различные методы используют разные представления сцен и обеспечивают разные выходные данные (Seitz et al., 2006). Результатом может быть либо облако точек, либо поверхность, представляющая реконструированный объект. В статье представлены классификация и обзор методов 3D-реконструкции, а приведены сравнения этих методов для различных задач и нюансов как самих объектов реконструкции, так и условий применения подходов. Также оцениваются методы реконструкции одного класса для лучшего понимания технических тонкостей их использования. В документе представлена информация о широко используемых способах при трехмерной реконструкции, что поможет исследователям понять современное состояние в этой области. Остальная часть статьи организована следующим образом. В следующем разделе представлена классификация методов 3D-реконструкции, кратко описаны методы классов. Эти подходы подробно обсуждаются далее в собственных подразделах. Каждый подраздел проливает свет на плюсы и минусы соответствующего способа. Стоит отметить, что в случае сканирования приведены сравнения непосредственно основных характеристик уже существующих устройств – 3D-сканеров с разными методами измерения. В этом разделе представлена сравнительная таблица и описаны параметры сравнения. В подразделах другого класса внимание уделяется основным принципам формирования трехмерного результата и те сложности, с которыми могут столкнуться исследователи. В последнем разделе представлено обсуждение данной работы и заключение обзора.

Ключевые слова: Трехмерная реконструкция объектов, бесконтактные методы, фотограмметрия, 3D-сканирование, восстановление поверхности

Введение

В представленном обзоре методы будут разделены на классы и подклассы. Трехмерные методы реконструкции объектов можно условно определить двумя типами: контактные и бесконтактные.

К контактными методам в основном относятся: Триггерное, непрерывное измерения, координатно-измерительные машины и робототехнические руки относятся к контактными методам. Перечисленные способы могут получать напрямую трехмерные данные о сцене/объекте с высокой точностью по средствам определенных инструментов. Следовательно, их применение обуславливается возможность прямо контактировать с объектом/сценой для проведения измерений. Например, робототехнические средства сталкиваются с тем, что робот двигается «на ощупь» и исследует перед собой проходимость поверхности, что дает высокую точность оценки его исследования, но оценка удаленной от него поверхности невозможна, так же как планирование и поиск пути.

Бесконтактные методы базируются на анализе одного или нескольких изображений в разных диапазонах, поэтому лишены указанного выше минуса за счет чего считаются более универсальными, несмотря на существенно более низкую точность восстановления.

Однако, именно из-за универсальности такие методы широко применимы, к примеру, в строительстве, в медицине, в аэрофотосъемке и т.д., поэтому большая часть данной работы будет уделена им.

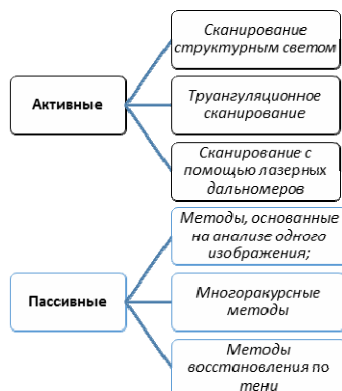


Рис. 1. Классификация бесконтактных методов трехмерной реконструкции

В свою очередь класс бесконтактных методов делят на активные и пассивные, что зависит от способа получения информации о глубине объекта исследования.

Взаимодействие радиометрическими приемами с объектом трехмерного реконструкции для получения информации о глубине называют активными методами. Пассивные же имеют влияния на объект/сцену исследования и применяют в качестве вводных данных изображение или видео.

Активные методы

Такие методы трехмерной реконструкции объектов принимают данные о положении точек в пространстве на поверхности исследуемого объекта, что достижимо с помощью сканирования, которое выполняется звуковыми или световыми волн. Применяющие свет, способы считают очень быстрыми, надежными и точными, но особенности поведения света нуждается в ограничениях при работе с полупрозрачными, прозрачными, рассеивающими, сильно отражающими, остроугольными сегментами или с глубокими вырезами. Ведь свет будет отражен, преломлен, рассеян или поглощен, что не позволит получить корректные данные об объекте.

Методы активной трехмерной визуализации серьезно затруднены из-за ухудшающих эффектов (рассеяния или поглощения), ведь они снижают уровни контрастности изображения и интенсивности света. Во избежание этих нюансов, активные методы трехмерной визуализации могут использовать 4 аппаратных решения: поляризационное, пространственное, временное и спектральное различие [17]. При помощи снятых поляризатором пары изображений с разных ракурсов, можно улучшить контраст изображения, или, увеличив расстояние от датчика до источника света, сократим обратное рассеивание [19]. Более сложные лазерные подходы применяют тонкие лучи для минимизации деграционного эффекта, который образуется рассеиванием, появляющимся в перекрывающемся потоке FOV устройства, а лазерные линейные методы на базе триангуляции уже включают оптическое сканирование FOV. Подход с временной дискриминации отбрасывает элемент шума благодаря стробируемому приемнику. Также, стоит отметить, что сократить поглощение света возможно выбором длин волны зеленого или синего цвета для источника света. Методы 3D-съемки на основе триангуляции или времени пролета применяют локальный способ сканирования объекта/сцены с помощью плоского света или узкого луча.

Системы структурированного света

В соответствии с различными режимами структурированного света, метод 3D-реконструкции структурированного света можно разделить на точечный структурированный свет, линейный структурированный свет и метод поверхностного структурированного света [14]. Метод точечного структурированного света заключается в проецировании точек света на поверхность целевого объекта и создании яркого пятна на поверхности объекта. Положение точек проекционного света определяется по извлеченному изображению.

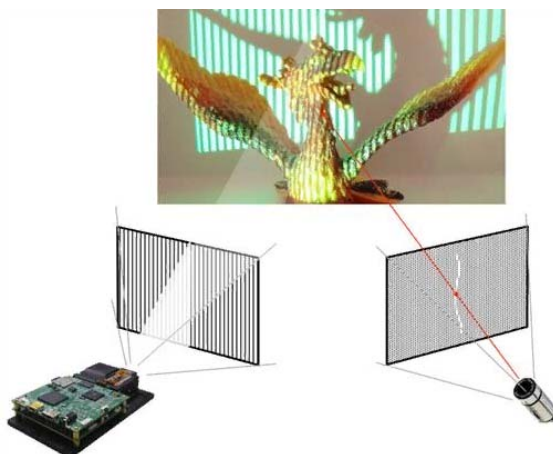


Рис. 2. Схема системы сканирования структурным светом с использованием DLP

Псевдослучайное кодированное изображение определяется заранее, и кодированное изображение может быть повторено. Он выражается в виде двоичного псевдослучайного массива, который может быть сгенерирован сдвиговым регистром, а затем может быть свернут в псевдослучайный массив определенным методом. Свернутый псевдослучайный массив имеет характеристики окна, то есть, когда окно определенного размера перемещается в массиве, содержимое окна уникально.

Для метода 3D-реконструкции со структурированным освещением основными причинами, влияющими на скорость сопоставления, являются несоответствие и окклюзия. Для поверхности нерегулярных объектов со сложными текстурами после проецирования псевдослучайного шаблона кодирования полученные закодированные изображения будут не совпадать из-за разных углов съемки камеры, что повлияет на результаты реконструкции.

Трудности также встречаются при объектах с полупрозрачными материалами, мешающих рассеиванию на поверхности, к примеру, мрамор или растения.

Технология трехмерного лазерного сканирования на основе триангуляции

Лазерная визуализация с триангуляцией получается путем проекции лазерной точки или линии на исследуемый объект, регистрируя получаемые отражения датчиками, которые определены расстоянием от лазерного источника. В ходе чего угол отражения можно описывать как данные о рельефе.

Луч лазера проецируют так, чтобы он был в поле захвата камеры, а также направляют безупречно параллельно для оптической оси камеры. Лазерная точка будет захвачена с остальными элементами объекта/сцены. Для этого на изображении находят пятно меньшей яркости, поскольку точка лазера хорошо выражена на темном куске объекта. Далее считают расстояние до объекта/сцены (D) по соотношению треугольников, применяя известные параметры.

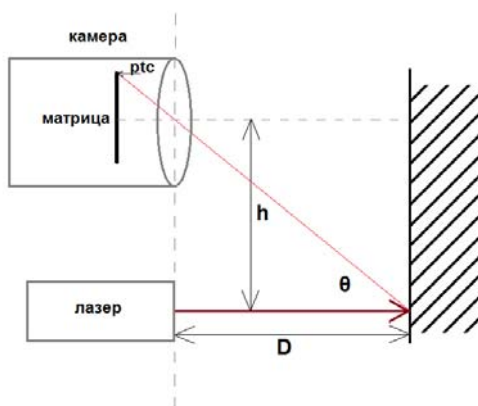


Рис. 3. Схема триангуляционного сканирования

Триангуляция получила свое название образующегося «треугольника» лазерным излучателем, камерой и точкой лазера на поверхности объекта (Рис.3). В приведенной системе известны сторона треугольника h (расстояние от лазерного излучателя до камеры), угол лазерного излучателя, а угол камеры получаем из положения лазерной точки в поле захвата камеры. Три этих параметра достаточно для формирования размера треугольника и определяют расположение угла точки лазера..

Преимущество лазерной триангуляции – это высокие показатели разрешения и точности результата.

Трехмерная визуализация лазерными дальномерами

Иначе говоря, речь о 3D-сканерах. Такое устройство включает в себя лазерный дальномер и сканер лазерного луча. Как сканер применяют полигональное зеркало для вертикального отклонения луча при вращении и сервопривод для отклонения горизонтально. Лазерный дальномер измеряет расстояния с помощью луча лазера. В лазерном 3D-сканере пользуются методами измерения расстояния импульса и фазы.

Импульсный метод базируется на измерении времени прохождения сигнала между приемопередатчиком и объектом (туда и обратно). Расстояние определяется благодаря скорости распространения электромагнитных волн. Фазовый метод заключается в определении разности фаз модулированных сигналов между отправкой и приемом.

Проблемы с лазерным сканированием:

- Данные об освещении получают через дополнительную камеру;
- Продолжительность процесса.

Трехмерная визуализация пассивными методами

Оптические подходы базируются на неоднократном захвате сцены, съемка которой производится с разных точек. Здесь искусственные источники применимы необходимости освещения сцены и заперщены при триангуляции трехмерных точек, основанной на подобных точках последовательности изображений из стереоалгоритмов сопоставлений. Стереосистемы используют две цифровые калиброванные камеры для захвата сцены. Использование трех синхронизированных камер [14] представляет собой хороший компромисс между точностью и габаритами устройства. Методы из движений применимы к сценам из потока перекрывающихся изображений, которые получают движением камеры. Такой подход заключается в извлечении точек автоматически (разреженный набор функций, таких как углы) и отслеживании такого набора функций в потоке изображений, а также оценивании их трехмерного положения благодаря нескольким ракурсам.

Помимо этого, существуют подходы извлечения карты глубины и по одному изображению, но для большинства задач они производятся с меньшей точностью.

Методы на базе анализа одного изображения; монокулярный способ

Вводной информации является одно изображения, поэтому определение карты глубины требует интерактивного подхода. Данные об объеме объекта/сцены получают анализом контуров, фактуры, затенений, освещения или с помощью проективной геометрии. Соответственно, главным плюсом данного подхода является легкость калибровки поскольку применяется одна камера, а минус состоит в том, что недостаточная информация карты глубины влечет низкую точность.

Фотограмметрия, много ракурсный способ

Фотограмметрия является процессом трехмерной визуализации объекта по нескольких изображениях с разных ракурсов/углов обзора, что повторяет бинокулярный принцип зрения людей. Здесь на наборе изображений определяют особые точки и выделяют для них особые точки соседних изображений, далее триангуляционно устанавливают трехмерные координаты. Метод имеет достаточно точное формирование карты глубины, но и трудности в работе с узорами и окклюзиями.

Метод восстановления формы по тени

Метод SFS [7] (Shape From Shading) также является более распространенным методом. Учитывая, что теневая граница изображения содержит информацию об особенностях контура изображения, яркость, темнота и тень изображения при различных условиях освещения могут использоваться для вычисления информации о глубине поверхности объекта, а модель освещения отражения используется для трехмерной реконструкции. Следует отметить, что яркость пикселя ограничена включением ин-

дикаторов источника света, параметров камеры и материала целевой поверхности. Метод формы восстановления теней имеет широкий спектр применений и может восстанавливать трехмерные модели различных объектов, кроме зеркал. Недостатком является то, что в основном это математические расчеты, а результаты реконструкции недостаточно точны, кроме того, метод SFS требует точных параметров источника света, включая информацию о положении и ориентации. Это делает его непригодным для ситуаций со сложным освещением, таких как сцены на открытом воздухе.

О алгоритмах

Монокулярная реконструкция с одного ракурса

Для этой задачи мы представляем алгоритм, принадлежащий к определенному классу методов, которые следуют тому же основному принципу, а именно представлению сетки вместе с моделью линейной деформации. [17, 18, 24, 28]. Этот алгоритм немного отличается, он состоит из двух нелинейных ограничений и способен выполнять такую реконструкцию, что форма любой изометрически деформированной поверхности оценивается с помощью только обычной камеры.

Изометрическое ограничение: это ограничение представляет собой разницу между наблюдаемой и прогнозируемой длиной ребра. Нелинейное ограничение формулируется как:

$$e_{mra} = \sum_{i=1}^{n_e} (L_i - \|s_1^{[i]} - s_2^{[i]}\|)^2 \quad (1)$$

где L_i – длина ребра i , вычисленная на шаблоне; $s_1^{[i]}$ и $s_2^{[i]}$ – значения сетки, которые учитывают конечные вершины ребра i .

Ошибка редизайна: Кроме того, бывают и ошибки редизайна, то есть ошибки в положении характерных точек на изображении. Таким образом, мы должны учесть ошибку редизайна, добавив элемент в оптимизированную функцию. Объединение уравнений 1 и 2:

$$P_i = \left[\sum_{j=13} a_{ij} s_j^{[i]} \right] \quad (2)$$

где $s_0^{[i]}$ и $S_j^{[i]}$ — подвектор s_0 и подматрица S (соответственно), соответствующие вершине j треугольника, в котором расположена характеристическая точка i . Член, соответствующий ошибке редизайна, может быть получен, как указано ниже.

$$e_{mra} = \sum_{i=1}^n \left\| \lambda_i \begin{bmatrix} q_i \\ 1 \end{bmatrix} - K_{rgb} \circ P_i \right\| \quad (3)$$

где N — количество характерных точек. λ_i добавляет к задаче оптимизации дополнительные неизвестные. Поэтому целесообразно переформулировать приведенные выше уравнения так, чтобы можно было исключить λ_i . Рассмотрим приведенное ниже уравнение:

$$\lambda_i \begin{bmatrix} q_i \\ 1 \end{bmatrix} = K_{rgb} \circ \left[\sum_{j=13} a_{ij} s_j^{[i]} \right] \quad (4)$$

После некоторых простых алгебраических манипуляций:

$$\begin{bmatrix} a_{i1} A_i & a_{i2} A_i & a_{i3} A_i \end{bmatrix}_{2 \times 9} \begin{bmatrix} s_1^{[i]} \\ s_2^{[i]} \\ s_3^{[i]} \end{bmatrix}_{9 \times 1} = \begin{bmatrix} e_{1,i} & e_{2,i} \end{bmatrix}_{2 \times 1} = 0 \quad (5)$$

где $a_i = K_{rgb}^{(1,2)} - q_i K_{rgb}^{(3)}$

Это уравнение обеспечивает 2 линейных ограничения в виде: $e_{1,i} = 0$ и $e_{2,i} = 0$. Таким образом, модифицированный $e_{ге}$ принимает форму, в которой λ_i не существует, следующим образом:

$$e_{mre} = \sum_{i=1}^N \left((e_{1,i})^2 + (e_{2,i})^2 \right) \quad (6)$$

где e_{mre} обозначает модифицированный e_{re} .

Целевая функция: теперь два ограничения описываются как два отдельных нелинейных выражения. Однако оба ограничения должны быть объединены в единую целевую функцию, чтобы они учитывались одновременно при оценке всех параметров. Для этого их взвешенное суммирование минимизируется таким образом, что члену ошибки перепроектирования присваивается вес m , учитывающий его относительное влияние в объединенной целевой функции.

$$\min_w e_{tot} = (e_{iso} + m \cdot e_{mre}) \quad (7)$$

Приведенная выше схема оптимизации представляет собой нелинейную задачу минимизации методом наименьших квадратов, обычно решаемую с использованием итеративного алгоритма, такого как Левенберг-Марквардт.

Бинокулярная реконструкция

Функциональная схема бинокулярного способа представлена на рис. 4. Буквы b и f являются расстоянием от оптических осей камеры до фокусного расстояния объективов, в которых индексы s и w ничто иное как системы координат камеры и мировая соответственно.

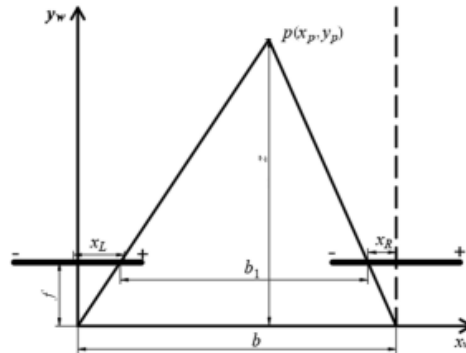


Рис. 4. Схема бинокулярной системы зрения.

Взяв левое и правое изображения и применив их величины координат L_x и R_x , определяется расстояние и координаты точке P в системе координат камеры [6]. Из подобия треугольников:

$$\begin{cases} \frac{z}{b} = \frac{z-f}{b_1} \\ b_1 = b - x_L - x_R \end{cases} ;$$

Учитывая эффект параллакса точки $d = x_L - x_R$ расстояние до точки P : $z = bf/d$.

Точка отсчета системы координат на рис. 2 совпадает с оптической осью левой камеры и имеет смещение на $b/2$ относительно центра бинокулярной камеры. Из теоремы подобия:

$$\frac{x_L}{f} = \frac{b_2}{y}$$

где $b_2 = -\frac{b}{2} - x$.

В результате получаем выражения для определения координат точки в системе координат бинокулярной камеры [7]:

$$\begin{cases} y_c = \frac{bf}{d}; \\ x_c = \frac{-b}{2} - \frac{x_L y}{f}. \end{cases}$$

Сравнительный анализ методов трёхмерной реконструкции

В таблице 1 собраны результаты сравнительного анализа методов трёхмерной реконструкции по критериям, представленным ниже и имеющих актуальность на сегодняшний день:

- Сложность вычислений.

Говоря о сложности вычислений, стоит обратить внимание на то, что параметрические методы являются менее трудоемкими в связи с уже существующей моделью восстановления, которая позволяет избавиться от лишних «особых» точек, что в следствии позволяет не сопоставлять и не ориентировать их в формируемом пространстве, что значительно упрощает требуемые расчёты. В статье Hui Chen и Wen Cui (2019) проведен эксперимент со сравнением активного структурированного света и многоракурсной трехмерной реконструкции объекта, таким образом определено, что для второго метода требовалось большое время выполнения. Если получено более точное облако точек, требуется больше входных изображений, но номера изображений не определены. Поэтому для случая, когда система проста, а облако точек реконструкции допускает некоторые отверстия, может быть выбран метод SFM.

- Простота реализации.

Здесь будет подразумеваться скорее не сама расстановка оборудования или внедрение метода, а те условия, которые необходимо обеспечить для наилучшей работы или работы вообще. Так, например, сканирование с использованием метода времени пролета является невозможным в ярком солнечном свете, как и при хорошей освещенности метод структурного света, в то время как, напротив, для методов на базе изображений не подходящим является затенение объекта.

- Пригодность к крупномасштабным сценам.

Крупномасштабные трёхмерные реконструкции объектов на сегодняшний день являются выполняемой задачей для практически всех методов, кроме триангуляционного подхода. Триангуляционные лазерные сканеры позволяют выполнять измерения с высочайшей точностью до десятых и даже сотых долей миллиметра, но на очень небольшой дистанции. Поскольку дальность действия производимых на сегодня, триангуляционных лазерных сканеров составляет от десятков сантиметров до 25 м, ошибки в измеряемых углах, вызванные влиянием атмосферы (рефракцией и затуханием электромагнитных колебаний), практически не оказывают влияния на результаты измерений. Точность получения пространственных координат точек объекта триангуляционными сканерами составляет от 50 мкм до 0,3 мм и в основном зависит от величины измеряемого расстояния, инструментальных ошибок прибора, методических ошибок, связанных с точностью математических решений, положенных в основу управляющего программного продукта сканера, а также от материала, текстуры и формы сканируемого объекта.

- Наружное применение

Для наружного применения сложно применять методы, которые не выполняют крупномасштабную 3D-реконструкцию и сложны в реализации.

- Восстановление сложных объектов

Развитие данного направления технологий настолько продвинулось вперёд, что восстановление сложных по форме или нагруженных деталями объектов стало возможных любым из методов, но результаты от пассивных методов оставляет желать лучшего, тк вводные данные низкого качества могут отрицательно повлиять на качество результата, а при теневом методе задача детализации становится весьма трудоемкой.

- Восстановление полупрозрачных объектов

Восстановление 3D-формы прозрачных объектов является отдельно изучаемой проблемой. Сложные световые пути, вызванные преломлением и отражением, не позволили как активным, так и описанными ранее пассивными методами решить эту проблему. Вариант решения был рассмотрен Чжэнцинь Ли, Ю-Ин Ёе, Манмохан Чандракер [Through the Looking Glass: Neural 3D Reconstruction

of Transparent Shapes\ Зазеркалье: нейронная 3D-реконструкция прозрачных фигур, 2020], но не получил широкого распространения из-за непригодности к основным практическим задачам.

- 3D-реконструкция в реальном времени.

Импульсные и фазовые методы сканирования позволяют получать трехмерное отображение динамической сцены, как можно увидеть в работе у многих авторов, исследующих решение задачи с помощью устройства Kinect, где нужно отметить работы с алгоритмом KinectFusion. Говоря о пассивных методах, это тоже возможно, но возникают задачи по улучшению локализации объекта. В настоящее время эти работы широко применяются в мобильных устройствах, таких как Apple Live Photos, Instagram Boomerang и проект “Cinemagraphs” (доступно: <http://cinemagraphs.com/>). Большинство из них требуют съемки более одной фотографии. Однако существуют исследования, в которых для восстановления движения объекта на фотографии требуется только одно неподвижное изображение. Было проведено исследование [16] для создания движения облака, что намного сложнее, чем анимация жестких объектов. С другой стороны, для восстановления 3D-сцены крайне важна локализация объекта, даже если система использует камеру RGB-D. Существуют исследования, посвященные изучению более эффективного способа отслеживания объекта.

Таблица 1

Сравнительный анализ методов трёхмерной реконструкции

Метод	Повышенная сложность вычислений	Простота реализации	Пригодность к крупномасштабным сценам	Применение	Восстановление сложных объектов	Восстановление полупрозрачных объектов	Работа в реальном времени
Активные	Сканирование структурным светом	+	-	+	+	-	+
	Лазерное сканирование	+	+	+	+	-	-
	Триангуляционное сканирование	-	-	-	+	-	-
Пассивные	Монокулярный метод	-	+	-	-	-	+
	Фотограмметрия	+	+	+	+	+	-
	Метод по тени	-	+	-	+	+	+

Заключение

В этой статье представлен обзор и сравнение основных методов 3D-реконструкции, хотя классифицировать эти методы очень сложно, они сгруппированы на основе их общих черт. Методы реконструирующего сканирования сравниваются на основе их производительности и ключевых параметров. Также обсуждаются результаты пассивных методов 3D-реконструкции, поможет исследователям понять современное состояние в этой области.

Решение проблемы соответствия само по себе является одной из сложных задач. С другой стороны, методы реконструкции с несколькими видами, которые работают в 3D-пространстве, такие как методы, основанные на вокселях, не нуждаются в решении сложной проблемы соответствия. Следовательно, эти подходы можно комбинировать, чтобы извлечь выгоду из обоих.

Развитие методов 3D-реконструкции в последние годы переживает огромный рост, тем не менее, эти методы нуждаются в улучшении во многих аспектах, таких как снижение сложности вычислений, простота реализации, пригодность для крупномасштабных и динамичных сцен, пригодность для приложений реального времени и так далее. Литература показывает, что многие методы реконструкции в основном предназначены для небольших базовых установок и применения внутри помещений. Акцент может быть сделан на том, чтобы сделать эти методы пригодными для наружного применения. Большинство методов моделируют только статическую сцену.

Методы реконструкции в основном сталкиваются с трудностями при работе со сложными сценами и объектами с бестекстурной поверхностью. С дополнительными подсказками / приоритетами методы могут быть изменены для работы с такими объектами. Прозрачные объекты, такие как огонь, или объекты с отражающей поверхностью, такие как зеркала, трудно реконструировать. В этом направлении необходимы дополнительные усилия. Существует возможность повысить точность пространственных методов изображения за счет повышения точности карт глубины. Активные методы позволяют получить стабильные результаты в виде облаков точек даже при неподходящих условиях реализации, что происходит благодаря применению закодированным шаблонам для расчета точек в трехмерном пространстве. Однако имеет влияние и достаточный шум рассеяния, требующий очистки облака точек вручную.

Для разрешения этого вопроса авторами Gianfranco Bianco, Алессандро Галло, Фабио Бруно и Maurizio Muzzuparra (2013) проведено исследование с использованием одного цветового канала для трехмерных визуализаций и выявлено, что процесс в зеленом канале дает хорошие показатели как уменьшение шума и увеличение точности. Пассивные способы формируют как результат наиболее чистые плотные облака точек в требуемых условиях, однако на итоговые показатели влияют текстуры. По опыту, область низкой текстуры имеет потерю контраста, что образует дыры и отсутствующие части. Кроме того, авторы предлагают интеграцию трехмерных облаков точек, которые получаемы благодаря двум способам, нацелившись на их преимущества: у пассивного метода - наилучшую плотность облака точек и наилучшее покрытие поверхности – у активного метода. Трехмерное облако точек, полученное в результате интеграции, позволяет повысить точность и плотность.

Литература

1. Aanaes H., Jensen R.R., Vogiatzis G., Tola E., Dahl A.B. Large-scale data for multiple-view stereopsis. IJCV 2016, 120(2), pp. 153-168.
2. Campbell N.D., Vogiatzis G., Hernandez C., Cipolla R. Using multiple hypotheses to improve depth-maps for multi-view stereo. In: ECCV. 2008.
3. Chen R., Han S., Xu J., Su H. Point-based multi-view stereo network. In: ICCV. 2019.
4. Chen R., Han S., Xu J., Su H. Point-based multi-view stereo network. arXiv preprint arXiv:1908.04422. 2019.
5. Cheng S., Xu Z., Zhu S., Li Z., Li L.E., Ramamoorthi R., Su H. Deep stereo using adaptive thin volume representation with uncertainty awareness. In: CVPR. 2020.
6. Cremers D., Kolev K. Multiview stereo and silhouette consistency via convex functionals over convex domains. PAMI 33(6), pp. 1161–1174. 2010.
7. Dosovitskiy A., Fischer P., Ilg E., Hausser P., Hazirbas C., Golkov V., Van Der Smagt P., Cremers D., Brox T. FlowNet: Learning optical flow with convolutional networks. In: ICCV. 2015.
8. Flynn J., Neulander I., Philbin J., Snavely N. Deepstereo: Learning to predict new views from the world's imagery. In: CVPR. 2016.
9. Furukawa Y., Ponce J. Accurate, dense, and robust multiview stereopsis. PAM 32(8), pp. 1362-1376. 2009.
10. Galliani S., Lasinger K., Schindler K. Massively parallel multiview stereopsis by surface normal diffusion. In: ICCV. 2015.
11. Goesele M., Snavely N., Curless B., Hoppe H., Seitz S.M. Multi-view stereo for community photo collections. In: ICCV. 2007.
12. Gu X., Fan Z., Zhu S., Dai Z., Tan F., Tan P. Cascade cost volume for highresolution multi-view stereo and stereo matching. In: CVPR. 2020.
13. Hartmann W., Galliani S., Havlena M., Van Gool L., Schindler K. Learned multi-patch similarity. In: ICCV. 2017.
14. He K., Zhang X., Ren S., Sun J. Deep residual learning for image recognition. In: CVPR. 2016.
15. Hiep V.H., Keriven R., Labatut P., Pons J.P. Towards high-resolution large-scale multi-view stereo. In: CVPR. 2009.
16. Honari S., Molchanov P., Tyree S., Vincent P., Pal C., Kautz J. Improving landmark localization with semi-supervised learning. In: CVPR. 2018.
17. Huang P.H., Matzen K., Kopf J., Ahuja N., Huang J.B. Deepmvs: Learning multi-view stereopsis. In: CVPR. 2018.
18. Im S., Jeon H.G., Lin S., Kweon I.S. Dpsnet: End-to-end deep plane sweep stereo. In: ICLR. 2019.
19. Ji M., Gall J., Zheng H., Liu Y., Fang L. Surfnet: An end-to-end 3d neural network for multiview stereopsis. In: ICCV. 2017.
20. Kendall A., Martirosyan H., Dasgupta S., Henry P., Kennedy R., Bachrach A., Bry A. End-to-end learning of geometry and context for deep stereo regression. In: ICCV. 2017.
21. Kingma D.P., Ba J. Adam: A method for stochastic optimization. In: ICLR. 2014.