

Разработка метода терагерцового интеллектуального видеонаблюдения на основе семантического слияния терагерцового и трёхмерного видеоизображений

А.А. Морозов¹, О.С. Сушкова¹, А.Ф. Пьлупанов¹

¹Институт радиотехники и электроники им. В.А. Котельникова РАН, Моховая 11-7, Москва, Россия, 125009

Аннотация. Терагерцовое видеонаблюдение открывает новые возможности в сфере обеспечения безопасности в общественных местах, поскольку позволяет обнаруживать скрытое оружие и другие опасные предметы и тем самым предотвращать их использование. Хотя первое поколение терагерцовых систем видеонаблюдения уже создано и доступно на рынке систем обеспечения безопасности, оно до сих пор не нашло широкого применения. Главной причиной этого является то, что существующие методы анализа терагерцовых изображений не способны обеспечить скрытое и полностью автоматическое распознавание оружия и опасных предметов и могут использоваться только под управлением специально обученного оператора. В результате, терагерцовое видеонаблюдение оказывается более затратным и менее эффективным, чем традиционные методы организации периметра безопасности и ручного досмотра посетителей. В статье рассматривается проблема создания метода автоматического анализа терагерцовых видеоизображений. В качестве основы для такого метода предложено использовать принцип семантического слияния видеоизображений, полученных с помощью разных физических принципов, идея которого заключается в том, что смысловое содержание одного видеоизображения используется для анализа другого видеоизображения. Например, информация о положении человека в пространстве, координатах его рук, ног и туловища может быть использована для анализа и интерпретации цветных областей, наблюдаемых на терагерцовом изображении. Для реализации принципа семантического слияния видеоизображений созданы средства объектно-ориентированного логического программирования, включающие специальные классы для сбора, обработки и анализа видеоизображений в видимом, инфракрасном и терагерцовом диапазонах, а также трёхмерных видеоизображений, полученных с помощью времяпролётных камер.

1. Введение

В последние годы терагерцовый диапазон электромагнитных волн привлекает пристальное внимание разработчиков систем обеспечения безопасности [1–6], что связано с некоторыми особыми свойствами терагерцового излучения.

Например, для терагерцового излучения прозрачны диэлектрические материалы, такие как пластик, дерево и керамика. Оно без-опасно для человека и может использоваться для облучения в течение длительного времени (в отличие, например, от рентгеновского излучения). В терагерцовом диапазоне частот расположены вращательные и колебательные резонансные частоты сложных органических молекул, что теоретически позволяет использовать терагерцовую спектроскопию для дистанционного обнаружения взрывчатки, наркотиков и других опасных веществ.

Обычно считается, что терагерцовый диапазон частот электромагнитного излучения находится между микроволновым и инфракрасным диапазонами (см. рисунок 1) при частотах примерно равных 3 ТГц – 300 ГГц, что соответствует длинам волн примерно от 0.1 до 1 миллиметра. Границы диапазона являются условными, и их определение несколько отличается в разных публикациях

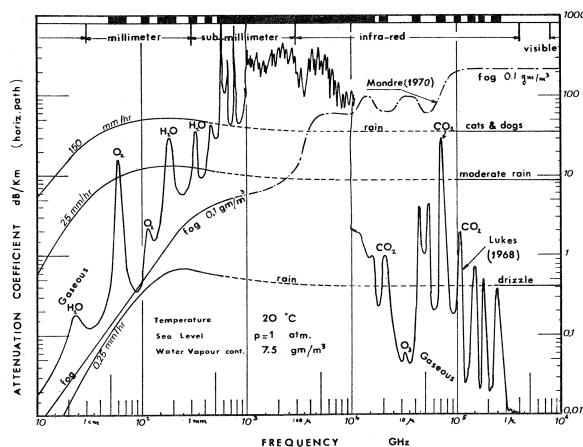


Рисунок 1. График зависимости величины затухания электромагнитных волн в атмосфере в зависимости от длины волны [7]. По оси абсцисс – частота [ГГц], по оси ординат – коэффициент затухания электромагнитных волн в атмосфере [дБ/км].

Свойства терагерцового излучения и принципы его использования различаются для разных поддиапазонов терагерцовых волн. В частности, для решения задач терагерцовой спектроскопии (распознавания опасных веществ) обычно используется излучение с частотой 0.5–3 ТГц и активная терагерцовая подсветка [8]. Для получения изображений предметов, скрытых под одеждой, применяют излучение с частотой менее 1 ТГц (субтерагерцовое излучение), соответствующее окнам прозрачности одежды. При этом могут использоваться активные, пассивные и смешанные методы исследования.

Изображения скрытых под одеждой предметов, полученные с помощью активных и пассивных методов терагерцового видеонаблюдения, существенно отличаются.

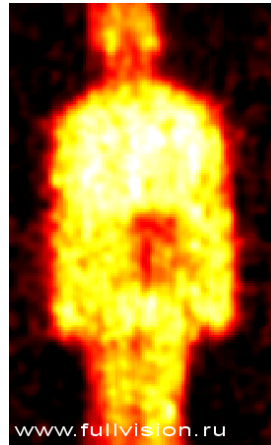


Рисунок 2. Изображение человека в терагерцовом диапазоне собственного излучения, полученное с помощью промышленной системы пассивного терагерцового видеонаблюдения THERZ-7A (Astrohn Technology Ltd [9]), диапазон 0.23–0.27 ТГц. Под одеждой за спиной человека спрятан пистолет ТТ.

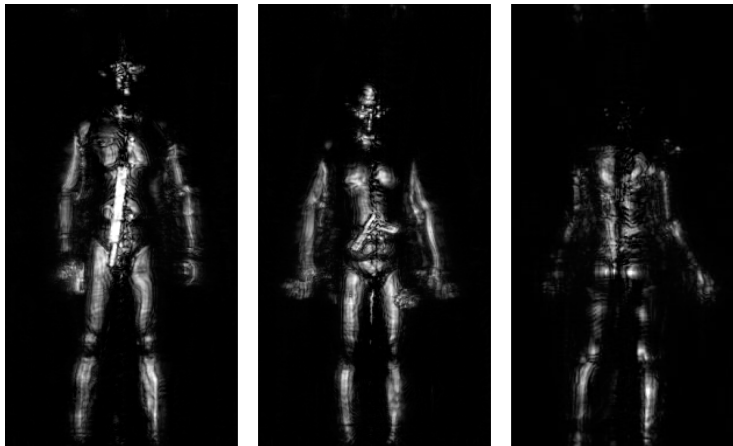


Рисунок 3. Терагерцовые изображения людей, полученные с помощью системы активного терагерцового видеонаблюдения [10]. Под одеждой людей спрятаны образцы холодного и огнестрельного оружия.

Соответственно отличаются и проблемы, возникающие при распознавании объектов на таких изображениях. Пассивное терагерцовое видеонаблюдение основано на регистрации собственного терагерцового излучения человеческого тела, при этом инородные предметы видны в качестве тёмных пятен на фоне собственного излучения тела (см. пример на рисунке 2). Трудности, возникающие при распознавании объектов на таких изображениях, связаны, прежде всего, с тем что:

- (i) Изображения объектов нечётки и характеризуются плохим разрешением и низким соотношением сигнал / шум.
- (ii) Фон терагерцового изображения более тёмный, чем изображение тела, так же как и тени от скрытых под одеждой предметов. Поэтому любые ошибки в разделении переднего плана и фона изображения могут приводить к ошибочному детектированию скрытых объектов.

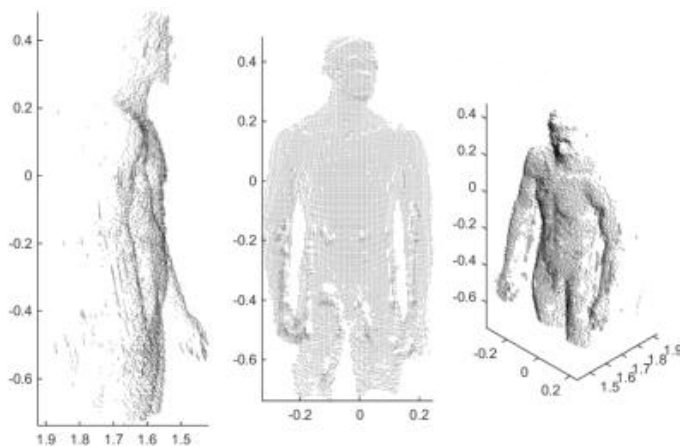


Рисунок 4. Трёхмерное терагерцовое изображение манекена, полученное с помощью терагерцового радара Pathfinder 200 GHz [11].

Активное терагерцовое видеонаблюдение предполагает использование терагерцовой подсветки, при этом регистрируется отражённое от тела человека излучение (см. пример на рисунке 3). Проблемы при анализе видеоизображений, полученных таким способом, связаны, в основном, с тем, что отражение терагерцового излучения от тела человека приводит к появлению бликов, в том числе, вытянутой формы, которые могут быть ошибочно распознаны как холодное оружие и опасные предметы, скрытые под одеждой [10].

Основными направлениями развития терагерцового видеонаблюдения считаются совмещение активного и пассивного методов сбора данных, а также реализация трёхмерного видеонаблюдения в терагерцовом диапазоне (см. пример на рисунке 4). Такие задачи ставились, в частности, в рамках европейского проекта CONSORTIS (Concealed Object Stand-Off Real-Time Imaging for Security) [11, 12], однако в настоящее время отсутствуют какие-либо сведения о разработке методов распознавания скрытых объектов достаточно надёжных для того, чтобы хотя бы приблизиться к решению проблемы полностью автоматического анализа терагерцовых видеоизображений. Все образцы экспериментальных систем терагерцового видеонаблюдения, описанные в литературе, представляют собой всего лишь новое, более совершенное поколение полуавтоматических систем досмотра.

2. Семантическое слияние многоканальных видеоизображений

Для решения задачи полностью автоматического распознавания объектов на терагерцовых видеоизображениях необходима разработка принципиально новых подходов, предусматривающих получение и использование информации о смысловом содержании (в том числе, о контексте) наблюдаемых терагерцовых видеоизображений (например, информации о положении определённого человека и положении в пространстве отдельных частей его тела), а также накопление и слияние видеоинформации, получаемой в различных спектральных диапазонах (видимом, инфракрасном и пр.) в течение определённых промежутков времени, при условии свободного перемещения людей в пространстве, по аналогии с тем, как это де-

лает человек-оператор системы терагерцового видеонаблюдения. Далее мы будем называть системы скрытого полностью автоматического терагерцового видеонаблюдения системами терагерцового интеллектуального видеонаблюдения, по аналогии с системами интеллектуального видеонаблюдения в видимом и инфракрасном диапазонах частот.

Типичное терагерцовое видеоизображение выглядит как набор размытых пятен, цветных или монохромных, в зависимости от применяемого метода визуализации данных. Одновременно с видеоизображением в терагерцовом диапазоне частот досмотровая система выводит изображение в видимом или инфракрасном диапазоне частот. Это позволяет специально обученному оператору правильно интерпретировать терагерцовое видеоизображение и выявлять наличие предметов, скрытых на теле человека. Действия оператора, сопоставляющего видеоизображения в разных диапазонах частот и наблюдающего за изменением формы цветных пятен при перемещении человека в кадре, в терминах автоматического анализа данных, являются примером семантического слияния видеоизображений, полученных с помощью разных физических принципов. Идея семантического слияния видеоизображений заключается в том, что несколько изображений объединяются в одно, и при этом смысловое содержание одного из объединяемых изображений используется для управления обработкой (анализом) объединяемого с ним другого изображения.

По мнению авторов, одним из наиболее важных источников информации для распознавания объектов на терагерцовом изображении является информация о положении человека и отдельных частей его тела в пространстве относительно наблюдаемого терагерцового изображения. Для получения такой информации целесообразно использовать облако точек и изображения скелетов, формируемые с помощью времяпролётной камеры. Для реализации этой идеи были разработаны специальные встроенные классы логического языка Акторный Пролог [13, 14]: *Astrohn*, *KinectBuffer* и др.

Встроенный класс *Astrohn* предназначен для получения терагерцовых и RGB-видеоданных из устройства THERZ-7A [9]. Встроенный класс *Astrohn* поддерживает ввод данных непосредственно с внешнего устройства, а также запись видеоданных в файл и чтение видеоданных из файла. Класс *Astrohn* поддерживает функции преобразования терагерцовых видеоданных в цветные изображения. В частности, в виде цветных изображений могут быть представлены непосредственно терагерцовые данные (с использованием псевдоцветов), RGB-данные, переданные терагерцовым комплексом, а также RGB-данные, совмещённые с текущим терагерцовым изображением. Кроме того, класс *Astrohn* поддерживает простейший режим синхронизации видеок кадров, в котором совмещаются терагерцовые и RGB-изображения, полученные максимально близко по времени. В настоящее время класс *Astrohn* поддерживает более 25 цветовых шкал высокого разрешения, включая общепринятые в тепловидении цветовые шкалы Aqua, Blackhot, Blaze, BlueRed, Gray, Hot, HSV, Iron, Red (Jet), Medical, Parula, Purple, Reptiloid, Green (Rainbow).

Встроенный класс *KinectBuffer* предназначен для получения данных с времяпролётной камеры устройства Kinect 2, а также для записи и чтения созданного 3D-видео. Подробное описание встроенного класса *KinectBuffer* можно найти в [15, 16]. Одной из функций встроенного класса

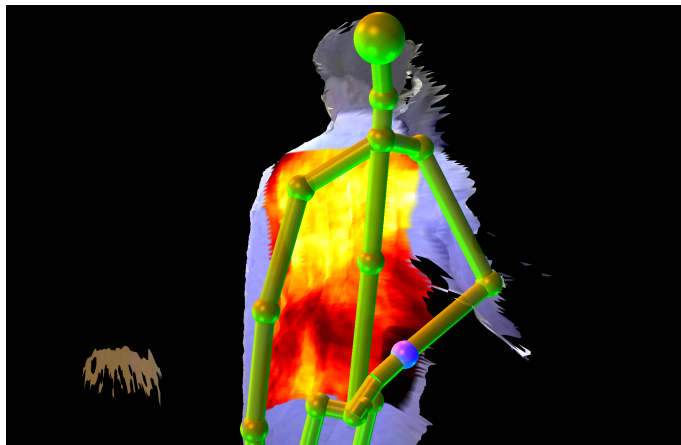


Рисунок 5. Слияние трёхмерного и терагерцового видеоизображений с помощью встроенного класса *KinectBuffer* Акторного Пролога [13, 15].

KinectBuffer является построение трёхмерной поверхности на основе облака точек и проецирование на эту поверхность заданной текстуры с помощью подготовленной заранее трёхмерной таблицы соответствий. Мы использовали эти средства для слияния трёхмерного и терагерцового видеоизображений в наших экспериментах. Для синхронизации терагерцовых изображений с трёхмерными видеоданными в классе *Astrohn* реализован метод опережающего чтения терагерцовых видеоданных. На рисунке 5 приведён пример трёхмерного изображения, построенного с помощью слияния облака точек, полученного с времяпролётной камеры, и терагерцового изображения. Терагерцовое изображение совмещено с изображением скелета человека, полученным средствами стандартного SDK устройства Кинект 2. Применение трёхмерной таблицы соответствий позволяет проецировать терагерцовое изображение на трёхмерную поверхность в реальном времени. В частности, трёхмерное видео, показанное на рисунке, может быть повернуто, отмасштабировано или сдвинуто при помощи мышки непосредственно в ходе воспроизведения. Семантическим слиянием в данном примере является использование облака точек, распознанного в качестве человеческой фигуры, для выделения на терагерцовом изображении цветowych пятен, относящихся непосредственно к предметам, скрытым на теле человека, и предотвращение ложного распознавания объектов на фоновом терагерцовом изображении.

3. Заключение

В качестве основы для интеллектуального терагерцового видеонаблюдения предложен метод семантического слияния терагерцового видео с видеоизображениями, полученными на основе иных физических принципов, в частности, с трёхмерным видеоизображением, полученным с помощью времяпролётной камеры. Для экспериментов с терагерцовым видеонаблюдением созданы средства объектно-ориентированного логического программирования, а именно, встроенные классы языка Акторный Пролог для ввода, записи, чтения и синхронизации видеопотоков в терагерцовом, тепловизионном и видимом диапазонах частот,

а также трёхмерного видео. Экспериментальная проверка продемонстрировала, что разработанные средства позволяют в реальном времени проецировать терагерцовое видеоизображение на трёхмерную поверхность человека, получаемую с помощью времяпролётной камеры. Эта информация может быть использована алгоритмами автоматического анализа для определения положения человека и отдельных частей его тела относительно терагерцового изображения, что позволит улучшить качество распознавания скрытых под одеждой предметов.

4. Литература

- [1] Federici, J.F. THz imaging and sensing for security applications – explosives, weapons and drugs / J.F. Federici, B. Schulkin, F. Huang // *Semiconductor Science and Technology*. — 2005. — Vol. 20(7). — P. S266.
- [2] Chan, W.L. Imaging with terahertz radiation / W.L. Chan, J. Deibel, D.M. Mittleman // *Reports on Progress in Physics*. — 2007. — Vol. 70(8). — P. 1325.
- [3] Sanders-Reed, J.N. Applications and challenges for MMW and THz sensors // *Micro- and Nanotechnology Sensors, Systems, and Applications VII*. International Society for Optics and Photonics. — 2015. — Vol. 9467. — P. 94672E.
- [4] Sizov, F. Infrared and terahertz in biomedicine // *Semiconductor Physics, Quantum Electronics & Optoelectronics*. — 2017. — Vol. 20(3). — P. 273-283.
- [5] Appleby, R. Millimeter wave imaging: a historical review / R. Appleby, D.A. Robertson, D. Wikner // *Passive and Active Millimeter-Wave Imaging XX*. International Society for Optics and Photonics. — 2017. — Vol. 10189. — P. 1018902.
- [6] Dhillon, S.S. The 2017 terahertz science and technology roadmap / S.S. Dhillon, M.S. Vitiello, E.H. Linfield // *Journal of Physics D: Applied Physics*. — 2017. — Vol. 50(4). — P. 043001.
- [7] Zufferey, C.H. A study of rain effects on electromagnetic waves in the 1-600 GHz range // *The MIMICAD Research Center, University of Colorado, Boulder, 1972*.
- [8] Baker, C. Detection of concealed explosives at a distance using terahertz technology / C. Baker, T. Lo, W. Tribe // *Proceedings of the IEEE*. — 2007. — Vol. 95(8). — P. 1559-1565.
- [9] Astrohn technology ltd. [Electronic resource]. – Access mode: <http://astrohn.com>.
- [10] Zhang, J. Terahertz image detection with the improved faster region-based convolutional neural network / J. Zhang, W. Xing, M. Xing, G. Sun // *Sensors*. — 2018. — Vol. 18(7). — P. 2327.
- [11] CONSORTIS. Final Publishable Summary Report. — Teknologian Tutkimuskeskus VTT, 2018.
- [12] Robertson, D.A. 220 GHz wideband 3D imaging radar for concealed object detection technology development and phenomenology studies / D.A. Robertson, D.G. Macfarlane, T. Bryllert // *Passive and Active Millimeter-Wave Imaging XIX*; ed. by D.A. Wikner, A.R. Luukanen. — 2016. — Vol. 9830. — P. 983009.
- [13] Morozov, A.A. Object-oriented logic programming of intelligent visual surveillance for human anomalous behavior detection / A.A. Morozov, O.S. Sushkova, A.F. Polupanov // *Optoelectronics in Machine Vision-Based Theories and Applications*; ed. by M. Rivas-Lopez, O. Sergiyenko, W. Flores-Fuentes, J. C. Rodriguez-Quinonez. — 2019. — P. 134-187.
- [14] Morozov, A.A. The intelligent visual surveillance logic programming / A.A. Morozov, O.S. Sushkova [Electronic resource]. – Access mode: <http://www.fullvision.ru>.

[15] Morozov, A.A. Development of agent logic programming means for multichannel intelligent video surveillance / A.A. Morozov, O.S. Sushkova, N.G. Petrova // Radioelektronika. Nanosistemy. Informacionnye Tehnologii (RENSIT). — 2018. — Vol. 10(1). — P. 101-116.

[16] Морозов, А.А. Семантическое слияние и совместный анализ терагерцового и трёхмерного видеоизображений средствами объектно-ориентированного логического программирования / А.А. Морозов, О.С. Сушкова, А.Ф. Полупанов // РЭНСИТ: Радиоэлектроника. Наносистемы. Информационные технологии. — 2018. — Т. 10, № 2. — С. 329-340.

Благодарности

Авторы выражают благодарность И. А. Кершнеру и Р. А. Толмачёвой за участие в проводимых экспериментах, а также компаниям ЗАО «Астрон» и ООО «АСофт» за предоставленное оборудование для регистрации терагерцового излучения THERZ-7A. Исследование выполнено при поддержке РФФИ (проект № 16-29-09626-офи-м).

Development of a Method of Terahertz Intelligent Video Surveillance Based on the Semantic Fusion of Terahertz and 3D Video Images

A.A. Morozov¹, O.S. Sushkova¹, A.F. Polupanov¹

¹Kotel'nikov Institute of Radio Engineering and Electronics of RAS, Mokhovaya 11-7, Moscow, Russia, 125009

Abstract. The terahertz video surveillance opens up new unique opportunities in the field of security in public places, as it allows to detect and thus to prevent usage of hidden weapons and other dangerous items. Although the first generation of terahertz video surveillance systems has already been created and is available on the security systems market, it has not yet found wide application. The main reason for this is in that the existing methods for analyzing terahertz images are not capable of providing hidden and fully automatic recognition of weapons and dangerous objects and can only be used under the control of a specially trained operator. As a result, the terahertz video surveillance appears to be more expensive and less efficient in comparison with the standard approach based on the organizing security perimeters and manual inspection of the visitors. In the paper, the problem of the development of a method of automatic analysis of the terahertz video images is considered. As a basis for this method it is proposed to use the semantic fusion of video images obtained using different physical principles, the idea of which is in that the semantic content of one video image is used to control the processing and analysis of another video image. For example, the information about 3D coordinates of a body, arms, and legs of a person can be used for analysis and proper interpretation of color areas observed on a terahertz video image. Special means of the object-oriented logic programming are developed for the implementation of the semantic fusion of the video data, including special built-in classes of the Actor Prolog logic language for acquisition, processing, and analysis of video data in the visible, infrared, and terahertz ranges as well as 3D video data collected using a time-of-flight camera.