

Исследование свойств модуляционных параметров речевых сигналов

Козлачков С.Б.,

Бонч-Бруевич А.М.,

Дворянкин С.В.

В статье приведены результаты исследований модуляционных свойств речевых сигналов. С помощью программы Matlab проведено моделирование помехоустойчивости модуляционных параметров тестовых речеподобных сигналов при воздействии на них аддитивных полосовых шумов. Выполнено сопоставление полученных результатов с данными психофизиологических исследований механизма слухового восприятия.

• разборчивость речи • речевой сигнал • коэффициент модуляции

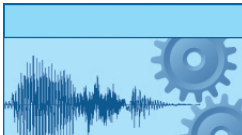
In the article there are the results of investigations of modular properties for speech signals. Through the program Matlab it is made the modeling of noise immunity modular parameters text speech-like signals through the influence on them the addictive band noise. It is made the correlation of results with psycho and physiological data for study acoustic impression mechanism.

• speech intelligibility • speech signal • modulation coefficient

Введение

Современную модель речевого сигнала (РС) можно представить как композицию вокализованных и невокализованных (шумовых) составляющих (сегментов). При этом информативная часть РС формируется путём модуляции его вокализованных и невокализованных сегментов. Уменьшение глубины модуляции (коэффициента амплитудной модуляции — m) сигнала приводит к снижению вероятности правильного восприятия и понижению его разборчивости [1].

На основании этого эффекта были разработаны соответствующие модуляционные методы (*STI*, *RASTI*, *STITEL* и *STIPA*) оценки качества РС [2], с помощью которых определяется его устойчивость к воздействию аддитивных маскирующих шумов. Принято считать, что основное информативное содержание несут в себе невокализованные (шумоподобные) сегменты РС. В соответствии с этим положением исследуемые шумовые сигналы разбиваются на октавные полосы, которые затем модулируются низкочастотными (силлабическими) тонами.



В настоящем исследовании определялся характер зависимости коэффициентов (глубины) модуляции (m) модулированных вокализованных и невокализованных сегментов РС при воздействии на него аддитивного белого шума. А также границы применимости модуляционных методов для оценки защищённости РС.

Описание эксперимента

В качестве оценочного инструмента был выбран метод *RASTI*, особенность которого заключается в определении значений m_k , получаемого для соответствующих значений частот модуляции (F_i) лишь для двух октавных полос со среднегеометрическими частотами 500 Гц и 2000 Гц (i — номер опыта, соответствующей частоты модуляции, k — номер октавной полосы частот). Такой выбор октавных полос обусловлен тем, что октавная полоса 500 Гц соответствует энергетическому максимуму усреднённого спектра речи, а в полосе 2000 Гц содержится максимальное относительное количество (до 30%) формант (т.е. информации РС).

Моделирование в программе *Matlab* выполнялось в соответствии с методикой, изложенной в работе украинских учёных [3]. В качестве тестовых использовались сигналы: вида белый шум (моделирование согласных невокализованных сегментов РС), а также вокализованный речеподобный сигнал (ВРПС — моделирование вокализованных сегментов РС, см. *рис. 1*), графическая сонограмма одного из таких ВРПС представлена на *рис. 3* [4].

Здесь, на *рис. 1*, и далее, на других рисунках спектрограмм, по оси абсцисс отложено время, по оси ординат — частота. В уровнях серого цвета указана мощность РС на данной частоте в данный момент времени. Наибольшие значения показаны чёрным цветом, наименьшие — белым.

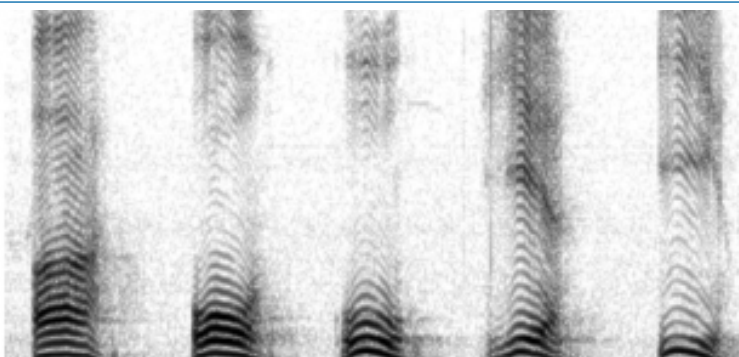


Рис. 1. Графическая сонограмма звуков «а», «о», «у», «е», «и».

Значения частот модуляции ВРПС для соответствующих октавных полос приведены в *таблице 1*.

Таблица 1

Значения частот модуляции (F_i) в октавных полосах

Среднегеометрическая частота октавной полосы, Гц	500	2000
Частоты модуляции, Гц	1; 2; 4; 8	1,4; 2,8; 5,6; 11,2

Время, в течение которого проводилось усреднение результатов, составило около 30 — 60 сек.

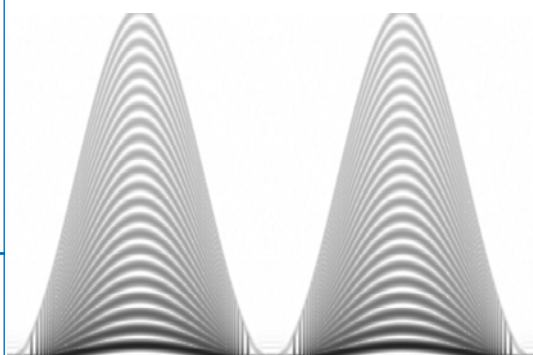


Рис. 2. Графическая сонограмма одного из тестовых вокализованных речеподобных сигналов (ВРПС)

Рис. 3. Зависимость коэффициента модуляции (m_N) сигнала вида белый шум в октавной полосе со средне-геометрической частотой 500 Гц (F_{500}) от SNR:

1 — верхняя граница значений коэффициента модуляции m_N ;
 2 — усреднённые значения коэффициента модуляции m_N ;
 3 — нижняя граница значений коэффициента модуляции m_N

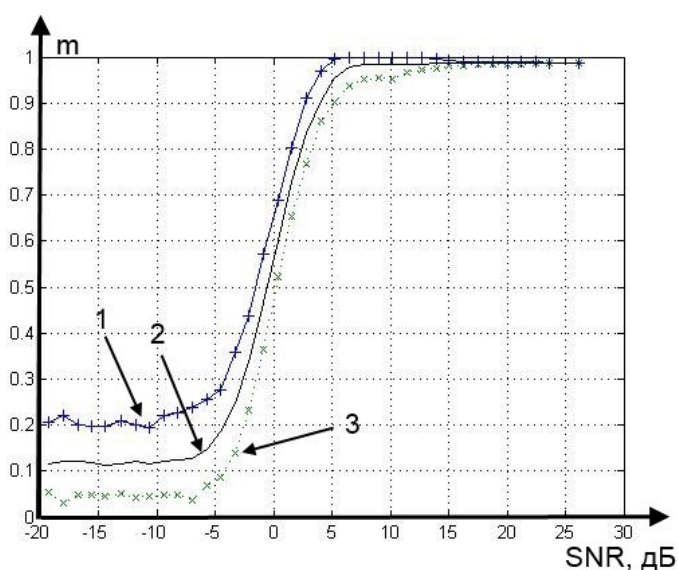
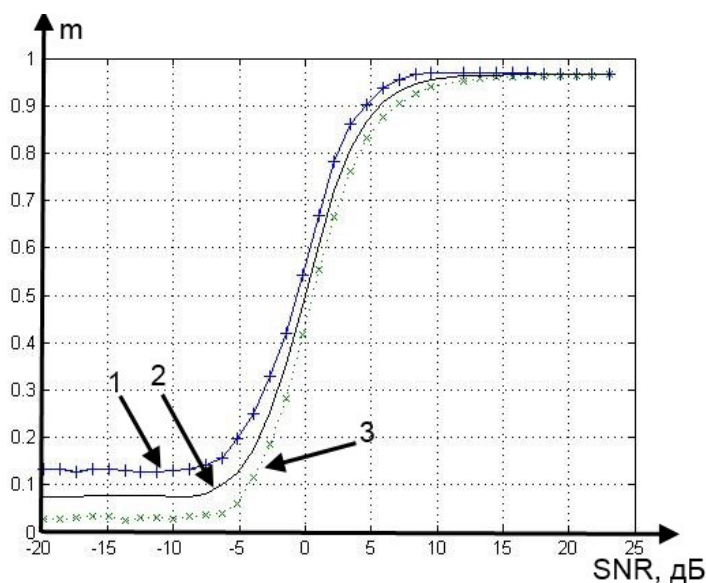


Рис. 4. Зависимость коэффициента модуляции (m_N) сигнала вида белый шум в октавной полосе со среднегеометрической частотой 2000 Гц (F_{2000}) от SNR:

1 — верхняя граница значений коэффициента модуляции m_N ;
 2 — усреднённые значения коэффициента модуляции m_N ;
 3 — нижняя граница значений коэффициента модуляции m_N



Результаты эксперимента

Полученные зависимости коэффициента модуляции (m_N — для сигналов вида белый шум и m_T — для ВРПС) от соотношения сигнал-шум (SNR) приведены на рисунках 3–6.

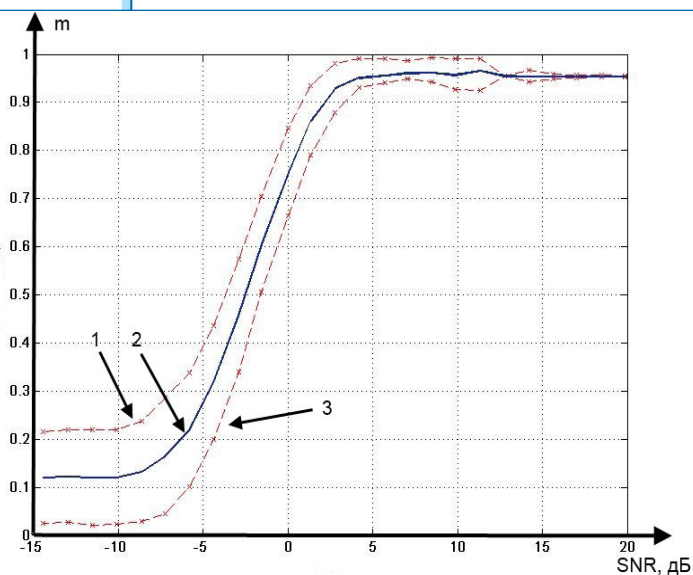


Рис. 5. Зависимость коэффициента модуляции (m_T) ВРПС в октавной полосе со среднегеометрической частотой 500 Гц (F_{500}) от SNR:

1 — верхняя граница значений коэффициента модуляции m_T ,
2 — усреднённые значения коэффициента модуляции m_T ,
3 — нижняя граница значений коэффициента модуляции m_T

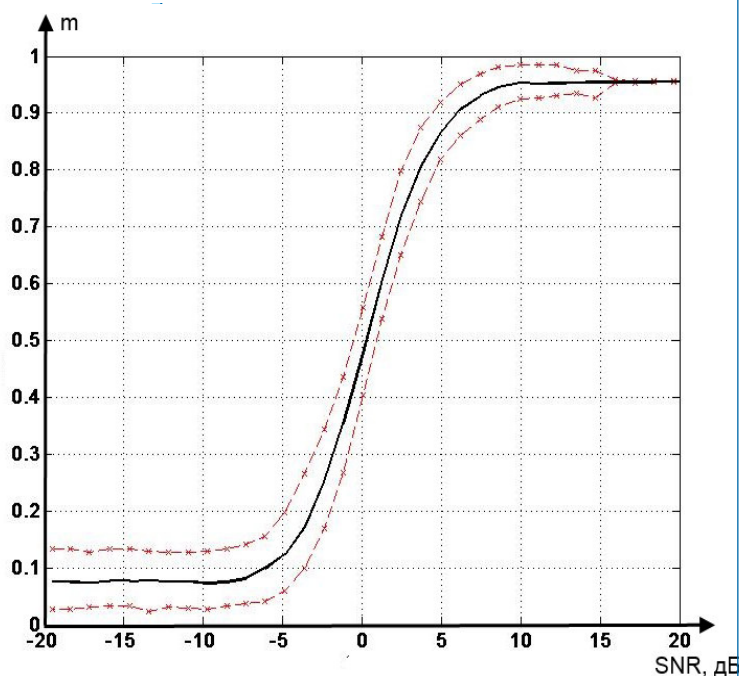


Рис. 6. Зависимость коэффициента модуляции (m_T) ВРПС в октавной полосе со среднегеометрической частотой 2000 Гц (F_{2000}) от SNR:

1 — верхняя граница значений коэффициента модуляции m_T ,
2 — усреднённые значения коэффициента модуляции m_T ,
3 — нижняя граница значений коэффициента модуляции m_T

Далее для каждого полученного значения m_k определялись значения SNR с допустимой погрешностью, отражённой в виде доверительного интервала. В общем случае задача нахождения в данных областях значений SNR и m_k с заданными погрешностями является математически некорректной [5]. Область, в которой с заданной погрешностью можно определить m_k и соот-

ветствующее ему значение SNR (линейная часть графика рис. 7), сверху и снизу ограничена областями, в которых определить эти параметры невозможно.

Для оценки защищённости необходимо определить SNR, соответствующее измеренному m_k . Погрешность определения отношения «сигнал — шум» (δ_{SNR}) при известной погрешности оценки коэффициента модуляции δ_m может быть определена как:

$$\delta_{SNR} = \frac{\delta_m}{(dm/dSNR)} \quad (1)$$

Для того чтобы найти производную от функции (см. рис.8), построенной по точкам, требуется её аппроксимация. В качестве аппроксимирующей в программе *Matlab* взята Гауссовская функция вида:

$$f(x) = a_1 * \exp\left(-\frac{(x-b_1)^2}{c_1}\right) + a_2 * \exp\left(-\frac{(x-b_2)^2}{c_2}\right) + \dots + a_7 * \exp\left(-\frac{(x-b_7)^2}{c_7}\right) \quad (2)$$

От функции берётся производная, и определяется значение дифференциала в каждой точке функции. Полученные значения соответствуют приращению этой функции. Отметим, что при $SNR < A_1$ погрешность оценки SNR становится недопустимо большой, так как производная зависимости $m(SNR)$ в этих случаях стремится к нулю, и, согласно формуле (1), погрешность

Рис. 7. Зависимость $m(SNR)$ для сигнала вида белый шум в октавной полосе со среднегеометрической частотой 500 Гц (F_{500}): точка A_1 — погрешность 3 дБ; точка A — погрешность 1 дБ.

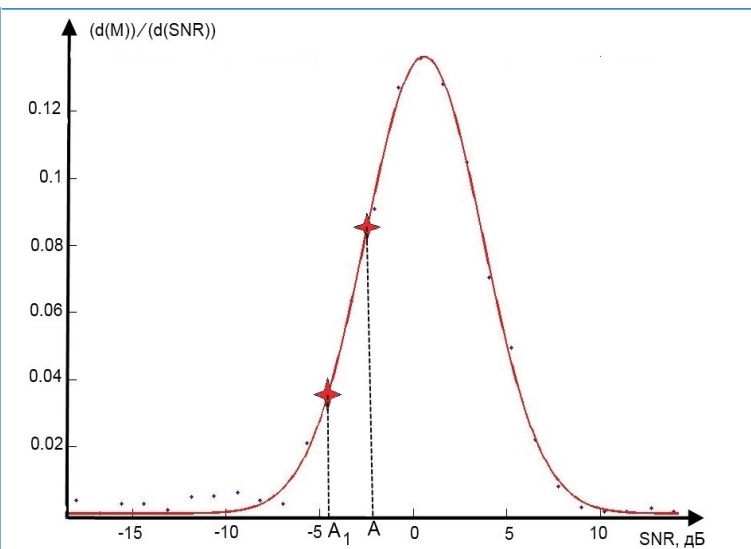
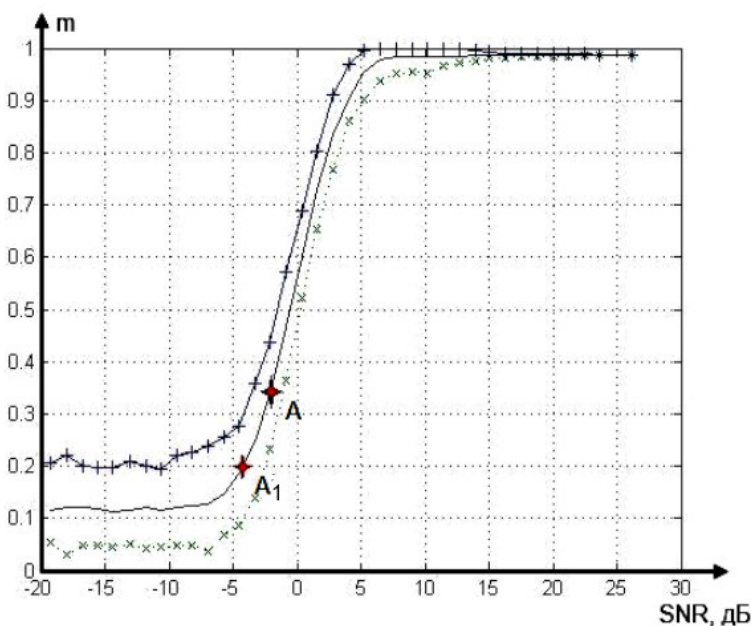


Рис. 8. График производной от функции $m(SNR)$

оценки SNR будет стремиться к бесконечности. Допустимые пределы оценки значения SNR по значению m будут находиться в области $SNR > A$. Таким образом можно определить границу применимости метода RASTI в области отрицательных значений SNR. При заданной погрешности SNR в 1 дБ, соответствующей метрологическим требованиям, получим значение $A = -2,079$ дБ. При заданной погрешности SNR в 3 дБ получим значение предела $A_1 = -4,534$ дБ.

Результаты определения границы применимости метода RASTI (значения SNR) для соответствующих видов тестового сигнала приведены в [таблице 2](#).

Таблица 2

Граничные значения (SNR) применимости метода RASTI

Величина допустимой погрешности, дБ	Значения точки граничных значений SNR (дБ) для тестовых сигналов со среднегеометрической частотой октавной полосы (Гц)			
	Для тестового сигнала вида — белый шум		Для тестового сигнала вида — ВРПС	
	500 Гц	2000 Гц	500 Гц	2000 Гц
1 дБ	-2.079 дБ	-2.656 дБ	-5.262 дБ	-5.634 дБ
3 дБ	-4.534 дБ	-5.131 дБ	-8.065 дБ	-8.432 дБ

Анализ зависимостей $m(SNR)$, приведённых на [рис. 3 — 6](#), показал, что в области $SNR < -15$ дБ разброс значений m_k определяется характеристиками маскира, а именно его пик-фактором (P_F), который, в свою очередь, зависит от ширины полосы маскира. При этом с уменьшением полосы шум по своим характеристикам приближается к свойствам гармонического сигнала, хаотически модулированного (по амплитуде) низкочастотным негармоническим колебанием, вследствие чего увеличивается его P_F .

Вышеприведённые факты полностью согласуются с результатами психоакустических исследований [1], в которых установлено, что уменьшение полосы частот шума, модулированного низкочастотным колебанием, приводит к усилению ощущений «шероховатости» звука, обусловленных собственными неоднородностями шума. При высоком фоне собственных неоднородностей шума слуховому аппарату труднее выделить низкочастотную модуляцию малого уровня. По этой причине при сокращении ширины полосы шума коэффициент минимально ощущаемой на слух модуляции (пороговые значения m_k) возрастает.

Кроме этого, на основании полученных данных можно сделать вывод, что тестовые сигналы вида белый шум, по динамическим характеристикам, на 4 — 6 дБ уступают ВРПС. Это также согласуется с результатами психоакустических исследований [1], в том, что модуляция тональных (вокализованных) сигналов воспринимается лучше, чем шумоподобных (невокализованных).

В области $SNR > 5$ дБ максимальные значения m_k определяются характеристиками тестового сигнала (P_F), который зависит от ширины полосы и степени стационарности этого сигнала. Используемый ВРПС может быть представлен набором гармонических составляющих, модулированный суммой гармонических сигналов, и является квазистационарным. При уменьшении полосы такого сигнала он трансформируется в амплитудно-модулированный тональный, с $P_F = 0$. Соответственно, уменьшение полосы таких ВРПС снижает величину его P_F . Полученные значения P_F для тестовых сигналов вида белый шум и ВРПС приведены в [таблице 3](#).

Таблица 3

Значения P_F для тестовых сигналов

Характер модуляции сигнала	Значения P_F (дБ) в полосах фильтрации сигнала (Гц)					
	Сигнал вида белый шум			Сигнал вида ВРПС		
	0...4000 Гц	1400...2800 Гц	355...710 Гц	0...4000 Гц	1400...2800 Гц	355...710 Гц
Немодулированный	6.5 дБ	11.6 дБ	12.31 дБ	10.77 дБ	10.6 дБ	7 дБ
Модулированный частотой 4 Гц	8.2 дБ	12 дБ	14.22 дБ	12.3 дБ	12 дБ	9.4 дБ

Заключение

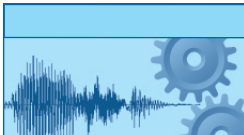
Сравнение результатов, полученных с использованием тестовых ВРПС и шумоподобных сигналов, позволило сделать следующие выводы.

Использование модуляционных методов для оценки защищённости РС ограничено значениями $SNR > -5$ — -8 дБ в соответствующих октавных полосах. При этом ВРПС имеют явное превосходство перед тестовыми шумоподобными сигналами.

Многотональные виды модуляции тестовых сигналов уступают однотоновым как по параметру помехоустойчивости, так и по восприятию в отсутствие аддитивного маскирующего шума.

Вокализованные сегменты РС имеют более высокую помехоустойчивость по сравнению с его невокализованными сегментами, что обусловлено различными модуляционными свойствами этих сигналов, т.е. их физической природой. Этот факт согласуется с результатами психофизиологических исследований слухового восприятия: дифференциальные пороги модуляционных характеристик тональных (вокализованных) сигналов в несколько раз ниже, чем для шумоподобных (невокализованных) [1]. До сих пор было принято считать, что помехоустойчивость вокализованных сегментов РС обеспечивается только за счёт их более высокого энергетического уровня по сравнению с невокализованными сегментами.

В отсутствие аддитивного маскирующего шума для оценки качества РС и параметров его модуляции решающее значение имеет его пик-фактор (P_F). При этом в условиях анализа в узких полосах, что более характерно для низкочастотного диапазона (где ширина критических полос слуха не превышает 100 Гц), преимущество на стороне вокализованных сегментов РС. В области частот свыше 2000 Гц (где ширина критических полос слуха составляет



300 — 700 Гц) модуляционные свойства вокализованных и невокализованных сегментов примерно равны. В области частот свыше 4000 Гц (где ширина полосы анализа составляет 1000 — 3500 Гц) преимущество на стороне невокализованных сегментов РС. Следует отметить, что именно в соответствии с этой закономерностью, обусловленной физическим дуализмом сигналов, распределены энергии реальных (не тестовых) РС.

Литература

1. Цвикер Э., Фельдкеллер Р. Ухо как приёмник информации / Перевод с немецкого под редакцией Б.Г. Белкина. М.: «Связь», 1971.
2. Рашевский Я.И., Каргашин В.Л. Обзор зарубежных методов определения разборчивости речи // Специальная техника. 2002. № 4. С. 37–46.
3. Дронжевская Л.Б., Климов В.А., Продеус А.Н., Шагитова Д.А. Формантный и формантно-модуляционный методы оценки разборчивости речи. Часть 1. Унификация алгоритмов // Электроника и связь. 2010. № 6 (59). Часть 2. С. 117–124.
4. Дворянкин С.В. и др. Моделирование и генерация сигналов речевых вокализов в приложениях оценки защищённости речевой информации // Проектирование и технология электронных средств. 2007. № 4. С. 33–37.
5. Тихонов А.Н., Арсенин В.Я. Методы решения некорректных задач. М.: Наука, 1979.

Сведения об авторах

Козлачков Сергей Борисович,

кандидат технических наук, старший преподаватель кафедры «Защита информации» МГТУ им. Н.Э. Баумана. Область научных интересов: речевые технологии, акустика, защита информации. E-mail: ksb.perovo@mail.ru

Бонч-Бруевич Андрей Михайлович,

кандидат технических наук, доцент кафедры «Защита информации» МГТУ им. Н.Э. Баумана. Область научных интересов: цифровая обработка сигналов, компьютерное моделирование. E-mail: 123andryb@mail.ru

Дворянкин Сергей Владимирович,

профессор, доктор технических наук, академик Российской Академии естественных наук, декан факультета кибернетики и информационной безопасности Национального исследовательского ядерного университета Московского инженерно-физического института (МИФИ), проректор по информатизации Российского нового университета, член докторского диссертационного совета по проблемам информационной безопасности и защиты информации МИФИ. Область научных интересов: распределённые защищённые системы мониторинга и управления, информационная безопасность и защита информации, информационно-телекоммуникационные системы и процессы, цифровая обработка сигналов и изображений, информационно-аналитические системы, безопасность объектов туриндустрии.