

## ПЕРЕДАЧА И ОБРАБОТКА ИНФОРМАЦИИ

УДК 621.395

*С.Н. Кириллов, В.Т. Дмитриев, Я.О. Картавенко*

### АЛГОРИТМ ОБЪЕКТИВНОЙ ОЦЕНКИ КАЧЕСТВА ДЕКОДИРОВАННОГО РЕЧЕВОГО СИГНАЛА НА ОСНОВЕ ИЗМЕНЕНИЯ СПЕКТРАЛЬНОЙ ДИНАМИКИ КРИТИЧЕСКИХ ПОЛОС СПЕКТРА

*Предложен алгоритм объективной оценки качества речевого сигнала (РС) на выходе декодера на основе изменения спектральной динамики в пределах критических полос. Обоснованы параметры алгоритма и доказано, что степень корреляции субъективной и объективной оценок качества речи составляет 0,78-0,97.*

*Ключевые слова:* речевой сигнал, объективная оценка качества речи, фонетическая функция, спектральная динамика, критическая полоса.

**Введение.** Одной из важнейших проблем при проектировании и эксплуатации систем кодирования РС, наряду с надежностью, является обеспечение высокого качества передачи РС (КПРС).

Современные системы передачи обеспечивают при достаточной громкости до 89 % слоговую и до 100 % фразовую разборчивость РС, зачастую за счет потери узнаваемости. Поэтому в настоящее время к КПРС по каналам связи предъявляются требования как по разборчивости, так и по узнаваемости [1].

Разборчивость есть объективная количественная мера, характеризующая способность тракта передавать содержащуюся в РС смысловую информацию в данных конкретных акустических условиях окружающей среды. Под мерой разборчивости понимается выраженное в процентах или долях единицы отношение числа правильно принятых элементов речи (звуков, слогов, слов и фраз) к достаточно большому объему числа переданных. Разборчивость представляет собой семантическое содержание РС. Узнаваемость отражает информацию о говорящем абоненте, т.е. его индивидуальных особенностях речи и т.д. [1].

Субъективная оценка КПРС определяется в соответствии с [10,11] и требует проведения большого количества тестовых испытаний, что приводит к значительным организационным и временным затратам. Кроме того, сильное влия-

ние на результаты оказывают условия проведения испытаний, язык, настроение, возрастная группа экспертов и другие факторы. Все это приводит к низкой повторяемости результатов и высокому разбросу субъективных оценок КПРС.

Объективные оценки качества РС получают с помощью технических средств, что обеспечивает лучшую повторяемость результатов по сравнению с субъективными оценками. Однако объективные методы позволяют только предсказать ожидаемое КПРС, но не способны оценивать качество речи так, как это делает человек. При этом объективные методы оценки КПРС, позволяющие оперативно определить качество речи на выходе декодеров, в большинстве своем менее надежны, чем субъективные процедуры. Цель инструментального подхода к предсказанию КПРС - достижение высокой корреляции между предсказанными оценками качества и субъективной оценкой, полученной путем тестирования группой экспертов. Большинство инструментальных методов оценки КПРС сравнивают РС источника (исходный сигнал) и выходной РС системы передачи (декодированный сигнал) [1].

**Цель работы.** Целью работы является обоснование алгоритма объективной оценки КПРС, а также исследование предложенного алгоритма для ряда стандартных кодеров при различных видах акустических помех и искажений.

**Известные методы оценки КПРС.** Существует ряд алгоритмов объективной оценки КПРС, основывающихся на измерении динамических и статических характеристик (например, коэффициента передачи, частотной и амплитудной характеристик, динамического диапазона и др.) [2...5]:

– критерий отношения сигнал-шум (ОСШ)

$$\text{ОСШ} = 10 \lg \frac{\sum_{n=1}^{N_d} x^2(n)}{\sum_{n=1}^{N_d} (x(n) - y(n))^2}, \quad (1)$$

где  $x(n)$  и  $y(n)$  –  $n$ -е отсчеты исходного и декодированного РС,  $N_d$  – количество отсчетов дискретного РС;

– критерий сегментного отношения сигнал-шум (ОСШ<sub>сег</sub>)

$$\begin{aligned} \text{ОСШ}_{\text{сег}} &= \\ &= \frac{10}{M} \sum_{m=1}^M \lg \sum_{n=1}^{N_{\text{сег}}} \left( \frac{x^2(n, m)}{(x(n, m) - y(n, m))^2} \right), \end{aligned} \quad (2)$$

где  $x(n, m)$  и  $y(n, m)$  –  $n$ -е отсчеты на  $m$ -м сегменте исходного и декодированного РС,  $N_{\text{сег}}$  – количество отсчетов в сегменте РС,  $M$  – число сегментов РС;

– расстояние Итакура-Саито (Itakura-Saito Distance - ISD)

$$\begin{aligned} \text{ISD} &= \\ &= \frac{1}{N_d} \sum_{n=1}^{N_d} \left( \frac{|X(f_n)|}{|Y(f_n)|} - \log \frac{|X(f_n)|}{|Y(f_n)|} - 1 \right), \end{aligned} \quad (3)$$

где  $X(f_n)$  и  $Y(f_n)$  – спектры исходного и декодированного РС;

– искажение спектра барков (Bark Spectral Distortion – BSD)

$$\text{BSD} = \sum_{g=1}^G \sum_{n=1}^{N_{b,g}} (|X(f_n, g)| - |Y(f_n, g)|)^2, \quad (4)$$

где  $X(f_n, g)$  и  $Y(f_n, g)$  –  $g$ -я критическая полоса спектров исходного и декодированного РС,  $N_{b,g}$  – количество отсчетов в  $g$ -й критической полосе,  $G$  – общее количество критических полос;

– модифицированная оценка искажения спектра барков (Modified Bark Spectral Distortion – MBSD)

$$\begin{aligned} \text{MBSD} &= \\ &= \frac{1}{M} \sum_{m=1}^M \left[ P \sum_{g=1}^G \sum_{n=1}^{N_{b,g}} |X(f_n, m, g) - Y(f_n, m, g)|^2 \right], \end{aligned} \quad (5)$$

где  $X(f_n, m, g)$  и  $Y(f_n, m, g)$  –  $g$ -я критическая

полоса спектров  $m$ -го сегмента исходного и декодированного РС,  $P$  – мера восприятия искажений (в случае, когда искажения ощутимы  $P=1$ , в других случаях  $P=0$ );

– расстояние COSH (COSH Distance – COSH)

$$\begin{aligned} \text{COSH} &= \\ &= \frac{1}{N_d} \sum_{n=1}^{N_d} \left[ \frac{1}{2} \left( \frac{|Y(f_n)|}{|X(f_n)|} + \frac{|X(f_n)|}{|Y(f_n)|} - 1 \right) \right]; \end{aligned} \quad (6)$$

– искажение фазового спектра (Spectral Phase Distortion – SP)

$$\text{SP} = \frac{1}{N_d} \sum_{n=1}^{N_d} (\theta_x(f_n) - \theta_y(f_n))^2, \quad (7)$$

где  $\theta_x(f_n)$  и  $\theta_y(f_n)$  – фазовые спектры исходного и декодированного РС,  $N_d$  – количество отсчетов дискретного РС;

– корреляция возбуждения спектра (Excitation Spectral Correlation – ESC)

$$\begin{aligned} \text{ESC} &= \frac{\left( \sum_{n=1}^{N_d} |X(f_n)| \cdot |Y(f_n)| \right)^2}{\sum_{n=1}^{N_d} |X(f_n)|^2 \sum_{n=1}^{N_d} |Y(f_n)|^2}; \end{aligned} \quad (8)$$

– искажение спектра (Spectral Distortion – SD)

$$\text{SD} = \frac{10}{N_d} \sum_{n=1}^{N_d} \left[ \log_{10} |X(f_n)| - \log_{10} |Y(f_n)| \right].$$

Кроме того, для оценки качества РС может использоваться фонетическая функция [6...8], предложенная А.А. Пироговым. Физический смысл фонетической функции заключается в оценке изменения спектральной динамики речи. Согласно этой концепции фонемы отличаются друг от друга, прежде всего, изменениями спектрального распределения. Фонетическая функция (функция ощущения спектральной динамики – ФОСД) имеет вид:

$$\begin{aligned} P_H(\omega, t) &= P_2(\omega, t) - P_1(\omega, t) = \\ &= C \left[ \int_0^{\infty} \exp\left(\frac{-\tau}{T}\right) \cdot \lg \left( \frac{S_2(\omega, t)}{S_2(\omega, t-\tau)} \right) d\tau - \right. \\ &\quad \left. - \int_0^{\infty} \exp\left(\frac{-\tau}{T}\right) \cdot \lg \left( \frac{S_1(\omega, t)}{S_1(\omega, t-\tau)} \right) d\tau \right], \end{aligned} \quad (9)$$

где  $C$  – константа,  $S_1(\omega, t)$  – мгновенный спектр исходного РС,  $S_2(\omega, t)$  – мгновенный спектр декодированного РС,  $P_2(\omega, t) - P_1(\omega, t)$  – разность ФОСД декодированного и исходного РС,  $T$  – длина РС,  $\tau$  – сдвиг РС.

**Обоснование модернизации алгоритма ФОСД.** Человеческая система восприятия звука имеет ограниченное, зависящее от частоты разрешение. С точки зрения восприятия человеком, равномерное измерение частоты может быть выражено в единицах ширины критических полос [9]. В иностранной литературе критическая полоса получила название «барк» (bark) [4,5].

Согласно [2] усредненный спектр РС определяется по следующему выражению:

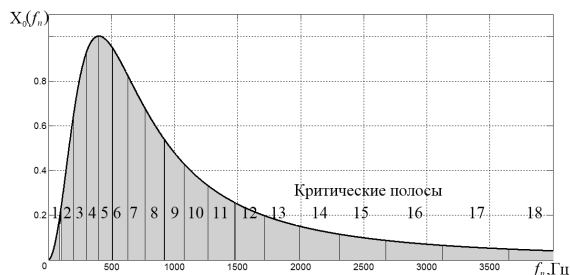
$$X_0(f_n) = 4\sigma^2 \frac{(f_n/f_0)}{\left[1 + (f_n/f_0)^2\right]^2}, \quad (10)$$

где  $\sigma$  – среднеквадратическое отклонение,  $f_0 = 400\text{Гц}$  – частота основного тона речи,  $f_n$  –  $n$ -я спектральная составляющая РС, в пределах от 0,1 до 3,4 кГц.

Введем понятие коэффициента «значимости» критической полосы  $\beta_g$ , показывающего соотношение энергии РС, сосредоточенной в критической полосе  $E_g$  по отношению к энергии  $E_{X_0}$  усредненного спектра РС, в виде:

$$\beta_g = \{E_g / E_{X_0}\}. \quad (11)$$

На рисунке 1 представлен усредненный спектр РС, полученный согласно (10), а также показаны критические полосы, попадающие в диапазон РС.



**Рисунок 1 – Усредненный спектр РС**

В соответствии с выражением (11) для каждой  $g$ -й из критических полос были получены следующие коэффициенты «значимости»  $\beta_g$  (таблица 1).

**Таблица 1**

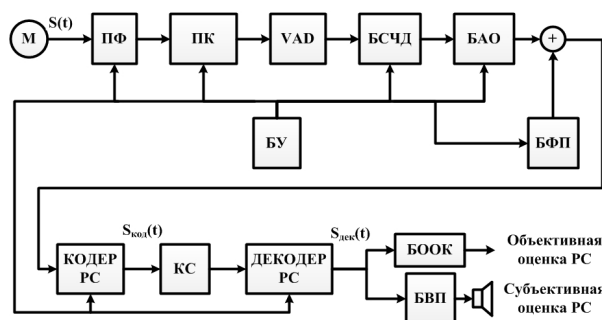
$g$	$\beta_g$	$g$	$\beta_g$	$g$	$\beta_g$
1	0.0049	8	0.0775	15	0.0317
2	0.0371	9	0.0653	16	0.0282
3	0.0685	10	0.0611	17	0.0254
4	0.0837	11	0.0521	18	0.0235
5	0.0926	12	0.0455	19	0.0212
6	0.0908	13	0.0409	20	0.0178
7	0.0890	14	0.0355	21	0.0145

На основе полученных данных была предложена модификация алгоритма ФОСД для вычисления изменения спектральной динамики критических полос спектров соседних сегментов РС:

$$MFOSD = \sum_{m=2}^M \sum_{g=1}^G \beta_g \sum_{n=1}^{N_{b,g}} P_H(f_n, g, m) = \sum_{m=2}^M \sum_{g=1}^G \beta_g \sum_{n=1}^{N_{b,g}} \left( \frac{|Y(f_n, g, m)|}{|Y(f_n, g, m-1)|} - \frac{|X(f_n, g, m)|}{|X(f_n, g, m-1)|} \right), \quad (12)$$

где  $M$  – количество сегментов РС,  $P_H(f_n, g, m)$  – ФОСД  $g$ -й критической полосы спектра сегмента  $m$  с длиной  $N_{seg}$ ,  $\beta_g$  – коэффициент «значимости» спектральных составляющих  $g$ -й критической полосы,  $X(f_n, g, m)$  и  $Y(f_n, g, m)$  –  $g$ -я критическая полоса спектра сегмента  $m$  с длиной  $N_{seg}$  исходного и декодированного РС.

**Экспериментальные исследования.** Экспериментальные исследования проводились по схеме, показанной на рисунке 2.



**Рисунок 2 – Схема получения объективной оценки качества речи**

На рисунке 2 используются следующие обозначения: М – микрофон, ПФ – полосовой фильтр, ПК – программный конвертор, VAD (Voice Activity Detector) – детектор речевой активности, БСЧД – блок сокращения частотного диапазона, БАО – блок амплитудного ограничения, БФП – блок формирования помехи, БУ – блок управления, КОДЕР РС – первичный кодек РС, КС – канал связи, БООК – блок объективной оценки качества РС, БВП – блок восстановления пауз.

Регистрируемый на микрофоне М тестовый РС с частотой дискретизации 44,1 кГц и разрядностью квантования 16 бит поступает на ПФ с полосой пропускания 0,1-3,4 кГц и далее преобразуется программным конвертором ПК с частотой дискретизации 8 кГц. Схема VAD формирует на выходе РС без пауз. В блоках БСЧД и

БАО осуществляется искусственное искажение РС вследствие сокращения его частотного диапазона и амплитудного ограничения соответственно. Сформированные в блоке формирования помехи БФП широкополосные и низкочастотные, квазистационарные и нестационарные помехи суммируются с РС с выхода БАО, и полученная аддитивная смесь РС и помехи поступает на вход блока КОДЕР РС. В блоках КОДЕР РС и ДЕКОДЕР РС происходит кодирование и декодирование РС в соответствии с принятыми по международным стандартам правилами. Декодированный РС поступает на динамик после блока восстановления пауз БВП, где осуществляется субъективная оценка КППС, а также в блок объективной оценки качества БООК, где выходными данными является оценка КППС, включающая оценки слоговой разборчивости и узнаваемости декодированного РС. В блоке БООК также осуществляется предоценочная обработка сигнала: выравнивание по уровню и времени исходного и декодированного РС, а также слуховое преобразование с использованием психоакустических моделей восприятия звуков человеком. Блок управления БУ осуществляет выбор режимов работы таких управляемых блоков, как БСЧД, БАО, БФП, КОДЕР РС и ДЕКОДЕР РС.

Исследования проводились для основных кодеков систем передачи: ММВЕ при скоростях передачи 1,2 кбит/с и 2,4 кбит/с, ICELP – 4,8 кбит/с и 6 кбит/с, G723.1 – 5,3 кбит/с, G729 – 8 кбит/с, AMR-WB – 15,85 кбит/с и 23,85 кбит/с.

Для сравнения точности представленных выше методов объективной оценки качества речи использовался их коэффициент корреляции с показателями субъективной оценки по разборчивости и узнаваемости, полученными согласно ГОСТ Р 50840-95 и ГОСТ Р 51061-97 [10, 11]. Получение субъективных оценок КППС по разборчивости и узнаваемости осуществлялось с помощью 10 дикторов и 10 аудиторов при комфортном уровне громкости, в помещении с условиями, удовлетворяющими ГОСТ 15150 [12].

Были проведены экспериментальные исследования влияния фильтрации спектра РС на выделение частотных диапазонов, отвечающих за разборчивость речи и узнаваемость диктора. Установлено, что для полосы пропускания 0,3-3,4 кГц частотный диапазон 0,3-2,2 кГц отвечает в основном за разборчивость речи, т.е. ее информационное наполнение, а частотный диапазон 1,5- $f_g$  кГц – за узнаваемость диктора.

Вычисление корреляции осуществлялось с помощью следующего выражения:

$$CORR = \frac{\sum_{d=1}^D (OB(d) - \overline{OB})(SB(d) - \overline{SB})}{\sqrt{\sum_{d=1}^D (OB(d) - \overline{OB})^2 \sum_{d=1}^D (SB(d) - \overline{SB})^2}}, \quad (13)$$

где  $OB(d)$  и  $SB(d)$  – объективный и субъективный показатели качества РС,  $\overline{OB}$  и  $\overline{SB}$  – средние значения объективного и субъективного показателей качества на множестве реализаций РС,  $d$  – номер реализации РС,  $D$  – общее количество реализаций РС,

В таблицах 2 и 3 приведены значения корреляций объективных и субъективных оценок разборчивости и узнаваемости РС (темным цветом в таблице выделены максимальные коэффициенты корреляции для различных алгоритмов оценки КППС и кодеков; разработанный алгоритм обозначен как *MFOSD*).

Таблица 2

Кодеки	ММВЕ	ММВЕ	ICELP	G723
Методы			Р	.1
Скорость, кбит/с	1,2	2,4	4,8	5,3
Разборчивость				
BSD	0,52	0,82	0,89	0,95
COSH	0,21	0,14	0,87	0,44
ESC	0,68	0,81	0,96	0,91
ISD	0,31	0,31	0,49	0,95
MBSD	0,45	0,9	0,85	0,91
PESQ	0,47	0,9	0,96	0,91
SD	0,86	0,9	0,96	0,91
CKO	0,5	0,71	0,04	0,91
SNR	0,51	0,7	0,11	0,92
SNR <sub>segm</sub>	0,22	0,56	0,56	0,22
SP	0,21	0,19	0,96	0,63
FOSD	0,65	0,62	0,44	0,4
MFOSD	0,93	0,9	0,96	0,95
Узнаваемость				
BSD	0,98	0,46	0,8	0,89
COSH	0,54	0,03	0,78	0,19
ESC	0,93	0,45	0,81	0,91
ISD	0,91	0,07	0,53	0,81
MBSD	0,98	0,77	0,76	0,88
PESQ	0,93	0,75	0,93	0,95
SD	0,84	0,79	0,87	0,94
CKO	0,98	0,64	0,1	0,88
SNR	0,98	0,64	0,16	0,9
SNR <sub>segm</sub>	0,79	0,85	0,43	0,49
SP	0,34	0,47	0,89	0,84
FOSD	0,88	0,25	0,4	0,72
MFOSD	0,98	0,53	0,95	0,92

Таблица 3

Кодеки Методы	G729	ICELP	AMR- WB	AMR- WB
Скорость, кбит/с	8	6	15,85	23,85
Разборчивость				
BSD	0,86	0,97	0,7	0,71
COSH	0,61	0,7	0,38	0,36
ESC	0,9	0,94	0,26	0,08
ISD	0,62	0,68	0,56	0,64
MBSD	0,9	0,97	0,69	0,7
PESQ	0,88	0,94	0,15	0,13
SD	0,87	0,76	0,67	0,77
CKO	0,72	0,85	0,76	0,77
SNR	0,76	0,83	0,84	0,81
SNR <sub>segm</sub>	0,36	0,58	0,12	0,11
SP	0,74	0,78	0,36	0,29
FOSD	0,76	0,71	0,38	0,68
MFOSD	0,75	0,97	0,89	0,83
Узнаваемость				
BSD	0,74	0,91	0,74	0,61
COSH	0,59	0,92	0,46	0,44
ESC	0,78	0,81	0,15	0,16
ISD	0,6	0,91	0,56	0,61
MBSD	0,8	0,81	0,73	0,59
PESQ	0,77	0,79	0,17	0,15
SD	0,77	0,95	0,71	0,85
CKO	0,6	0,94	0,79	0,69
SNR	0,64	0,96	0,82	0,78
SNR <sub>segm</sub>	0,25	0,46	0,14	0,25
SP	0,73	0,61	0,22	0,13
FOSD	0,69	0,93	0,5	0,65
MFOSD	0,79	0,92	0,81	0,84

Как следует из анализа таблиц 2 и 3, полученных на основе экспериментальных исследований, предложенный алгоритм *MFOSD* обеспечивает наибольшую степень корреляции объективных и субъективных оценок для рассмотренных кодеков РС в широком диапазоне скоростей передачи.

УДК 004.932

### **Б.А. Алпатов, П.В. Бабаян, С.А. Смирнов**

## **КОМБИНИРОВАННЫЙ АЛГОРИТМ СЛЕЖЕНИЯ ЗА ВОЗДУШНЫМИ ОБЪЕКТАМИ**

*Предложен алгоритм отслеживания воздушного объекта на основе автоматического выбора алгоритма измерения координат на основе оце-*

**Заключение.** Показано, что предложенный алгоритм объективной оценки КИРС на основе изменения спектральной динамики критических полос спектра обеспечивает степень корреляции объективных и субъективных оценок по разборчивости и узнаваемости речи не ниже 0,78 для рассмотренных алгоритмов кодирования РС.

#### **Библиографический список**

1. Цыбулин М.К., Бочаров М.О. Анализ методов оценки качества передачи речевой информации по каналам связи различной структуры // Электросвязь. - 2008. - №11. - С. 46-48.
2. Шелухин О.И., Лукьянцев Н.Ф. Цифровая обработка и передача речи. - М.: Радио и связь, 2000. - 456 с.
3. H. Ozer, I. Avciabas, B. Sankur, N. Memon, Steganalysis of audio based on audio quality metrics, in: SPIE Electronic Imaging Conf. on Security and Watermarking of Multimedia Contents // vol. V, Santa Clara, January 20-24, 2003. - PP. 55-66.
4. S. Wang, A. Skey, and A. Gersho. An objective measure for predicting subjective quality of speech coders. IEEE Journal on Selected Areas in Communications, 10(5), 1992. - PP. 74-77.
5. Wonho Yang, Enhanced modified bark spectral distortion (EMBSD): an objective speech quality measure based on audible distortion and cognition model, Ph.D dissertation, May 1999. -163 p.
6. Соболев В.Н. Информационные технологии в синтетической акустике. Монография - М.: ИРИАС, 2007. - 360 с.
7. Пирогов А.А. Синтетическая телефония. - М.: Связьиздат, 1963. - 120 с.
8. Пирогов А.А. Вокодерная телефония. - М.: Связь, 1974. - 536 с.
9. Цвикер Э., Фельдкеллер Р. Ухо как приемник информации. 2-е изд., перераб. и доп. - М.: Связь, 1971. - 255 с.
10. ГОСТ Р 50840-95. Передача речи по трактам связи. М.: Госстандарт России, 1995. - 180 с.
11. ГОСТ Р 51061-97. Системы низкоскоростной передачи речи по цифровым каналам. - М.: Госстандарт России, 1995. - 12 с.
12. ГОСТ 15150-69. Машины, приборы и другие технические изделия - М.: ИПК издательство стандартов, 1969. - 84 с.

нивания характеристик фоноцелевой обстановки. Выбор осуществляется из двух алгоритмов измерения координат объектов: алгоритма на основе байесовской сегментации и алгоритма на основе пространственной фильтрации. Представлены результаты экспериментальных исследований предложенного алгоритма с использованием обширной базы видеосюжетов.

**Ключевые слова:** слежение, измерение координат, алгоритм на основе байесовской сегментации, алгоритм на основе пространственной фильтрации, оценка работоспособности, комплексирование.

**Введение.** Системы технического зрения, предназначенные для слежения за воздушными объектами по данным видеонаблюдения, получили широкое применение как в военных, так и в гражданских областях. В таких системах достаточно широкое распространение получили два алгоритма измерения координат объектов: алгоритм на основе байесовской сегментации [1] и алгоритм на основе пространственной фильтрации [2]. Кратко охарактеризуем условия применения названных алгоритмов и укажем их основные достоинства и недостатки.

*Алгоритм на основе байесовской сегментации* предназначен для выделения движущихся и неподвижных объектов, наблюдаемых на относительно однородном фоне. Он основан на использовании априорной информации о различии статистических свойств объекта и фона. Этот алгоритм сегментации наиболее эффективен при выделении воздушных объектов на фоне неба. Для его работы необходимо начальное указание области изображения, в пределах которой следует искать объект. Вычислительная сложность алгоритма на основе байесовской сегментации невысока. Другим его преимуществом является устойчивость к геометрическим изменениям объекта. Отрицательной чертой этого алгоритма является низкая эффективность при наблюдении объектов на неоднородном фоне. Еще один недостаток данного алгоритма заключается в том, что для успешного слежения необходимо выполнение условия статистического различия точек фона и объекта. Кроме того, при малых размерах объекта полученная статистика может быть недостоверной, что неизбежно приводит к росту числа ошибок.

*Алгоритм на основе пространственной фильтрации* наиболее эффективен при выделении движущихся и неподвижных объектов на фоне ясного или облачного неба, в том числе и при малых отношениях сигнал/шум. Преимуществом этого подхода являются невысокие требования к объёму априорной информации, описывающей объекты наблюдения. Начальное целеуказание не является обязательным, необходима лишь грубая подстройка алгоритма под размеры объектов. Для успешного использова-

ния данного алгоритма требуется различие в свойствах пространственной структуры объекта и фона. Следует отметить, что вычислительная сложность алгоритма на основе пространственной фильтрации, как правило, больше, чем у алгоритма на основе байесовской сегментации.

Каждый из этих алгоритмов измерения координат объектов характеризуется определёнными требованиями к типам фоноцелевой обстановки, при которых он работоспособен. В некоторых случаях, имея априорные сведения о фоноцелевой обстановке, можно заранее выбрать алгоритм измерения координат объектов, обеспечивающий высокие качественные показатели слежения. Однако зачастую при слежении за воздушными объектами фоноцелевая обстановка значительно изменяется, в результате чего алгоритм измерения координат, выбранный в момент захвата объекта, может оказаться неработоспособным.

В настоящей статье описан комбинированный алгоритм, позволяющий повысить эффективность слежения за объектами за счет автоматического выбора наиболее работоспособного алгоритма измерения координат в зависимости от типа фоноцелевой обстановки. Отметим, что такой принцип повышения показателей качества слежения соответствует задаче *комплексирования алгоритмов обработки изображений*. Алгоритмы, между которыми осуществляется выбор, в дальнейшем будем называть *базовыми*.

**Цель работы.** Необходимо разработать алгоритм слежения за воздушными объектами, который на каждом кадре видеопоследовательности производит автоматический выбор алгоритма измерения координат, обеспечивающего наиболее устойчивое слежение за объектом с учётом условий наблюдения. Выбор должен осуществляться из двух алгоритмов измерения координат объектов: алгоритма на основе байесовской сегментации и алгоритма на основе пространственной фильтрации. Переход с одного алгоритма на другой должен производиться автоматически без необходимости повторного захвата объекта. Работа комбинированного алгоритма должна обеспечиваться при наличии начального целеуказания, содержащего координаты

наты центра и размеры объекта.

**Общий подход к решению задачи.** Подходы к комплексированию алгоритмов обработки изображений, получаемых односпектральным датчиком, можно условно разделить на две группы: *объединение* и *комбинирование* базовых алгоритмов. *Объединение* сводится к такой модификации базовых алгоритмов, при которой удается использовать сильные стороны одного алгоритма для повышения качественных характеристик другого алгоритма. При этом используемые алгоритмы могут соответствовать как одной, так и разным группам методов обработки изображений. При *комбинировании* базовых алгоритмов осуществляется слияние информации, полученной в результате работы этих алгоритмов.

В работах [3, 4] описывается подход к объединению алгоритмов, позволяющий устранить фактор, приводящий к ухудшению работоспособности алгоритма сопоставления с эталоном, с помощью алгоритмов сегментации. В работе [5] рассматривается алгоритм обнаружения динамических малоразмерных объектов, позволяющий уменьшить число ложных тревог за счет объединения информации, полученной от алгоритма выделения движения и алгоритма пространственного выделения объекта в кадре. В работах [6, 7, 8] представлены подходы, основанные на комбинировании базовых алгоритмов. В [6] рассматривается способ слежения за объектами при отсутствии априорных сведений о фоноцелевой обстановке, который заключается в выборе наиболее работоспособного алгоритма измерения координат. Для определения работо-

способности алгоритмов вводятся признаки, позволяющие оценить эффективность того или иного алгоритма. В работах [7, 8] описан подход, аналогичный тому, что был рассмотрен в работе [6]. Отличие заключается в том, что выбор осуществляется между четырьмя пороговыми алгоритмами сегментации.

В данной работе выбран подход, основанный на комбинировании базовых алгоритмов. Предлагается при поступлении очередного кадра производить измерение координат объекта с помощью двух алгоритмов одновременно. Затем производится оценка работоспособности каждого алгоритма. На основании оценок работоспособности принимается решение о выборе наиболее эффективного для данного типа фоноцелевой обстановки алгоритма измерения координат. Результаты работы наиболее работоспособного алгоритма принимаются за текущие координаты объекта, а также используются для корректировки зоны поиска второго алгоритма измерения координат на следующем кадре.

На рисунке 1 представлена структурная схема системы слежения, реализующей данный подход. Преимущество выбранного подхода перед подходами, основанными на объединении базовых алгоритмов, связано с относительной простотой построения системы слежения, базирующейся на апробированных алгоритмах измерения координат. При реализации данного подхода в аппаратуре не потребуется много затрат для разработки нового программного и аппаратного обеспечения. Вычислительная сложность данного подхода приблизительно равна сумме вычислительных затрат базовых алгоритмов.

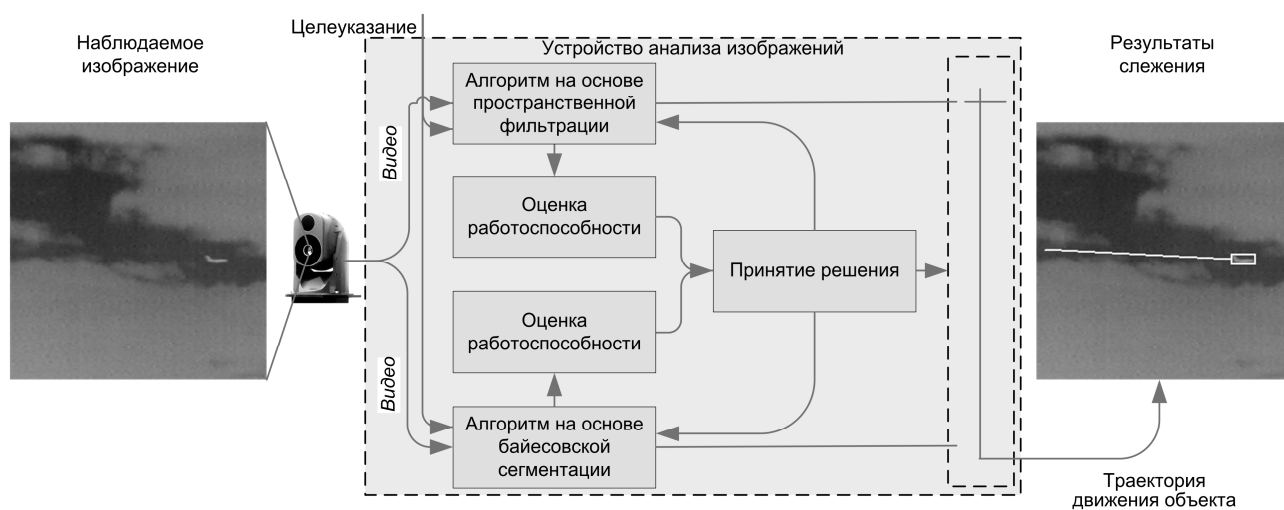


Рисунок 1 – Структурная схема системы автоматического слежения за объектами с выбором алгоритма измерения координат

**Решение задачи.** Для оценивания работоспособности алгоритмов измерения координат в работе [6] используются два типа метрик.

Первый тип метрик – количественные характеристики работоспособности (КХР), которые позволяют объективно оценить эффективность

алгоритма за счет сравнения измеренных координат с эталонными данными. Второй тип метрик, не требующий эталонных данных, – признаки работоспособности (ПР). ПР дают лишь оценочную характеристику работоспособности. Существенным недостатком ПР по сравнению с КХР является невозможность непосредственного сравнения качества измерения координат для различных алгоритмов, поскольку ПР, характеризующие разные алгоритмы, обычно не совпадают ни по размерности, ни по диапазону значений. Для использования этих характеристик предварительно необходимо установить статистическую связь между КХР и ПР. Таким образом, на основе данной связи можно по значениям ПР получить оценки КХР.

В данной работе предложены ПР, позволяющие охарактеризовать алгоритмы измерения координат на основе байесовской сегментации и на основе пространственной фильтрации. Для того чтобы производить сравнение этих ПР, осуществляется переход от значений ПР к значениям двоичной КХР. Приведем ПР для используемых базовых алгоритмов и двоичную КХР.

Признак, характеризующий алгоритм измерения координат на основе байесовской сегментации, основывается на вычислении статистического различия объекта и фона и определяется выражением:

$$K_{stat} = \left| \hat{H}(P) - \hat{H}(O) \right|, \quad (1)$$

где  $\hat{H}(P) = \sum_{x=1}^N \sum_{y=1}^M p(X|P) \log_2 p(X|P)$  – энтропия

изображения в рамке;

$\hat{H}(O) = \sum_{x=1}^N \sum_{y=1}^M p(X|O) \log_2 p(X|O)$  – энтропия

изображения в окне;  $p(x, y)$  – совместная яркостно-градиентная гистограмма;  $N, M$  – число разрядов гистограммы по яркости и по градиенту соответственно.

Алгоритм измерения координат на основе пространственной фильтрации базируется на оценке средней яркости объекта. Признак, характеризующий ожидаемую работоспособность такого алгоритма, можно определить по формуле:

$$K_{spat} = \frac{\max_{(i,j) \in H} |d_{strobe}(i, j)|}{\hat{\sigma}}, \quad (2)$$

где  $|d_{strobe}(i, j)|$  – модуль разностного изображения в точке  $(i, j)$ , принадлежащей строке;  $\hat{\sigma}$  – оценка СКО остаточного фона;  $H$  – множество точек, принадлежащих объекту, определяемое на основе целеуказания оператора или

результатов работы алгоритма на предыдущем кадре. Для вычисления разностного изображения и оценки СКО остаточного фона должны использоваться процедуры, аналогичные тем, что применяются в алгоритме [2].

Двоичная КХР позволяет описать работоспособность алгоритма измерения координат в терминах “следит”/”не следит”. КХР принимает значение “следит” в том случае, если измеренный центр объекта не выходит за границы прямоугольной области, заданной эталонными данными о расположении и размерах объекта, в противном случае критерий имеет значение “не следит”.

Переход от значений ПР к двоичной КХР должен осуществляться путём сравнения признаков работоспособности с пороговыми значениями. Процедура нахождения пороговых значений подробно описана в работе [9].

Установление статистической связи между ПР и КХР, а следовательно, и выбор пороговых значений производится экспериментально на основе базы данных видеосюжетов. Отметим, что при определении порогового значения возможны ошибки двух видов:

- ошибка первого рода – ошибочное принятие решения, что алгоритм неработоспособен;
- ошибка второго рода – ошибочное принятие решения, что алгоритм работоспособен.

В таблице 1 приведены пороговые значения для базовых алгоритмов и ошибки классификации, полученные на основе экспериментальных исследований.

**Таблица 1 – Пороговые значения признаков и ошибки классификации**

Алгоритм	Пороговое значение	Ошибка первого рода, %	Ошибка второго рода, %
на основе байесовской сегментации	0,7	11,1	3,9
на основе пространственной фильтрации	5,04	0	1,6

После перехода от значений ПР к оценке КХР можно непосредственно сравнивать работоспособность двух алгоритмов с целью принятия решения о выборе алгоритма измерения координат. Так как используемая в работе КХР может иметь только два значения “следит”/”не следит”, выбор алгоритма измерения координат очевиден только в случае, когда один из алгоритмов оказывается работоспособным. В случае, когда обе оценки КХР имеют значение “следит”, предлагается использовать алгоритм с наименьшей ошибкой второго рода, т.е. алго-



ритм на основе пространственной фильтрации. В этом случае, если оценка КХР принимает значение “следит”, можно говорить о практически безошибочном отнесении алгоритма к работоспособному. В случае, когда оценки КХР принимают значение “не следит”, целесообразно формировать признак срыва слежения.

Очевидно, что после принятия решения о выборе алгоритма измерения координат в текущем кадре необходимо скорректировать положение зоны поиска для менее работоспособного алгоритма. Для этого используются координаты, полученные в наиболее работоспособном алгоритме.

На основании описанного выше разработан комбинированный алгоритм слежения за воздушными объектами с автоматическим выбором алгоритма измерения координат. Можно выделить следующие основные этапы алгоритма.

а. Измерение координат алгоритмами на основе пространственной фильтрации и байесовской сегментации.

б. Расчет признаков по формулам (1), (2).

с. Сравнение признаков с пороговыми значениями. Если расчетные значения больше пороговых значений, то оценка КХР принимает значение “следит”, в противном случае – “не следит”.

д. Принятие решения о выборе алгоритма измерения координат.

е. Коррекция зоны поиска для менее работоспособного алгоритма.

**Экспериментальные исследования.** Экспе-

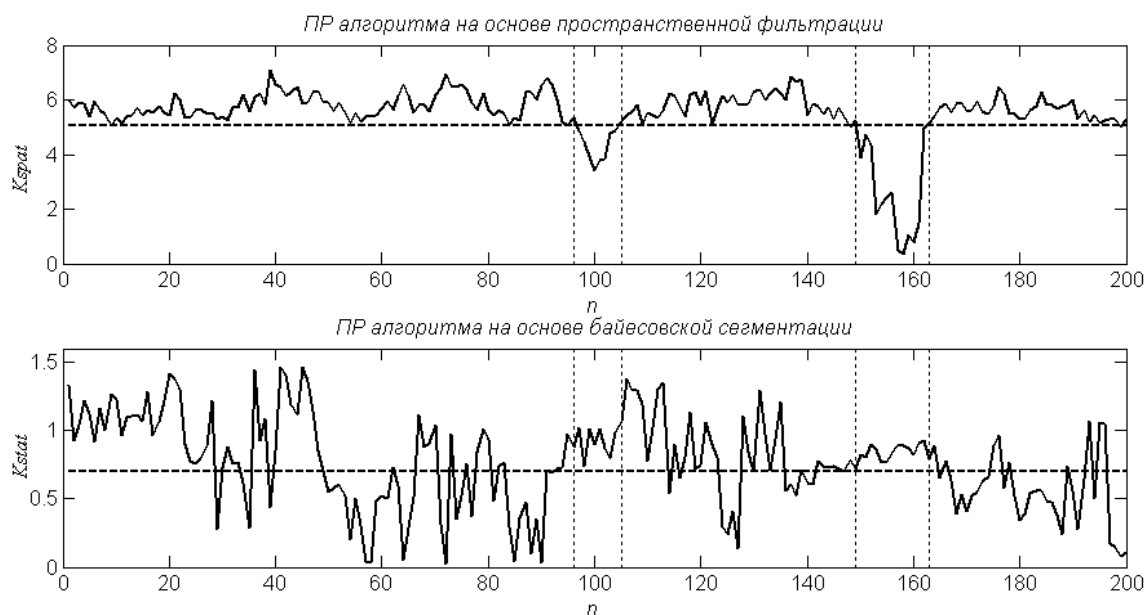
риментальная проверка комбинированного алгоритма производилась с использованием базы данных видеосюжетов, включающей шестьдесят натуральных видеопоследовательностей продолжительностью от 100 до 300 кадров. Фон на тестовых сюжетах был неподвижным или медленно движущимся с присутствием кучевых облаков, поверхности земли, домов и деревьев.

Для оценки эффективности предлагаемого алгоритма слежения подсчитывалось количество сюжетов, на которых наблюдался срыв слежения, при этом срыв слежения фиксируется, если измеренные координаты объекта выходят за границы прямоугольной области, заданной эталонными данными о расположении и размерах объекта.

В таблице 2 приведены результаты экспериментальных исследований для базовых алгоритмов и предлагаемого подхода. Анализ результатов показывает, что комбинированный подход позволяет улучшить качество слежения за воздушными объектами.

**Таблица 2 – Результаты экспериментальных исследований**

Алгоритм	Количество сюжетов, на которых происходит срыв слежения
на основе пространственной фильтрации	6
на основе байесовской сегментации	13
комбинированный	2



**Рисунок 2 – Признаки работоспособности для алгоритма на основе пространственной фильтрации (вверху) и алгоритма на основе байесовской сегментации (снизу)**

На рисунке 2 для одного из сюжетов приведены графики изменения ПР для алгоритмов на основе пространственной фильтрации и на основе байесовской сегментации при их работе в составе предлагаемого алгоритма. Отметим, что при отдельной работе этих алгоритмов срыв слежения наблюдается на 97-м и 29-м кадрах соответственно. С 1-го и по 97-й кадр измерение координат производится алгоритмом на основе пространственной фильтрации. Так как после 97-го и до 106-го кадра ПР вида (2) меньше порогового значения, происходит переключение на алгоритм на основе байесовской сегментации. С 106-го кадра измерение координат производится алгоритмом на основе пространственной фильтрации. Аналогичная ситуация наблюдается и со 149-го кадра по 163-й. На рисунке сплошной линией отмечены признаки вида (1) и (2), горизонтальной пунктирной линией – пороговые значения.

**Выводы.** Основываясь на результатах экспериментальных исследований, можно сделать вывод об эффективности комбинированного алгоритма. При этом достигается трехкратное снижение количества срывов слежения. Невысокая вычислительная сложность позволяет рекомендовать комбинированный алгоритм для использования в бортовых системах обнаружения и сопровождения объектов.

Работа проводилась при финансовой поддержке Министерства образования и науки РФ (госконтракт № 14.740.11.0284).

#### **Библиографический список**

1. Бакут П.А., Лабунец В.Г. Телевизионная

следающая система с байесовским дискриминатором цели // Зарубежная радиоэлектроника. – 1987. – № 10. – С. 81 - 93.

2. Муравьев В.С., Муравьев С.И. Алгоритм выделения и измерения координат объектов, наблюдаемых на облачных фонах // Вестник Рязанского государственного радиотехнического университета. – 2007. – № 21. – С. 20 - 24.

3. E. Hueber, L. Bigué, and P. Ambs. Active Contour Segmentation by Use of a Multichannel Incoherent Optical Correlator // Applied Optics, Vol. 42, Issue 23, 2003. – pp. 4681-4687.

4. Алтатов Б.А., Балашов О.Е., Шубин Н.Ю. Алгоритм измерения координат движущихся объектов последовательности изображений // Вестник Рязанского государственного радиотехнического университета. – 2010. – № 4. – С. 32 - 36.

5. Борисова И.В., Ерошкин Р.Э., Кузнецов А.А. и др. Обнаружение динамических малоразмерных объектов в различных спектральных диапазонах // Прикладная физика. – 2006. – № 5. – С. 100 - 103.

6. Алтатов Б.А., Бабаян П.В., Смирнов С.А. Автоматическое слежение за объектами при отсутствии априорных сведений о фоновой обстановке // Цифровая обработка сигналов. – 2009. – № 3. – С. 52 - 56.

7. Yong X., Feng D.D. and Rongchun Z. Optimal Selection of Image Segmentation Algorithms Based on Performance Prediction // Visual Information Processing, Sydney, Australia, 2003. – pp. 105-108.

8. Zhang, Y.J. and Luo, H.T. Optimal selection of segmentation algorithms based on performance evaluation // Optical Engineering. – 2000. – 39(6). – pp. 1450-1456.

9. Бабаян П.В., Смирнов С.А. Оценивание ожидаемой работоспособности алгоритмов измерения координат объектов // ГрафиКон' 2010: 20-я международная конференция. СПб. – 2010. – С. 329 - 330.

УДК 004.9: 681.518

**В.К. Злобин, А.Н. Колесенков, Б.В. Костров**

## **КОРРЕЛЯЦИОННО-ЭКСТРЕМАЛЬНЫЕ МЕТОДЫ СОВМЕЩЕНИЯ АЭРОКОСМИЧЕСКИХ ИЗОБРАЖЕНИЙ**

*Рассматривается проблема сокращения объема вычислений в корреляционно-экстремальных алгоритмах идентификации изображений. Предлагаются методы, основанные на применении вещественно-диадной свертки.*

**Ключевые слова:** корреляционно-экстремальные алгоритмы, идентификация изображений, вещественно-диадная свертка.

**Введение.** Совмещение аэрокосмических изображений (АКИ) является наиболее типичной задачей для практически всех видов применения АКИ. Цель работы – разработка методов корреляционно-экстремального совмещения изобра-

жений, основанных на применении вещественно-диадной свертки. В качестве типовых примеров использования подобных методов можно привести [1]:

- получение единого изображения по серии

синхронно формируемых снимков;

- решение задачи пространственной привязки изображений и преобразования их в заданную проекцию;

- определение местоположения летательных аппаратов в процессе их пилотирования;

- автоматическое обновление информации в ГИС мониторинга чрезвычайных ситуаций техногенного и природного характера.

Методологическую основу построения технологий решения перечисленных задач составляет цифровая реализация корреляционно-экстремальной обработки пространственно-временных сигналов, требующая больших вычислительных затрат. В связи с этим разработка простых алгоритмов, обеспечивающих сокращение вычислительных затрат при допустимом уровне точностных характеристик, является актуальной задачей. Существующие квазиоптимальные алгоритмы были в свое время разработаны интуитивным путем, что затрудняет и в некотором смысле делает невозможным аналитическое исследование и проведение их сравнительного анализа. Однако можно сформулировать некоторые принципы и определить наиболее общие этапы синтеза квазиоптимальных корреляционных алгоритмов и технологий:

- определение вида предварительной обработки изображений;

- определение типа критериальной функции, т.е. меры, с помощью которой будет выноситься решение о степени схожести изображений;

- выбор способа поиска экстремума критериальной функции.

**Теоретические исследования.** Введем следующие обозначения:  $f_{ТИ}(i, j)$  - функция распределения яркости некоторого изображения размером  $N \times K$  элементов в прямоугольных целочисленных координатах ( $i = \overline{0, N-1}$ ,  $j = \overline{0, K-1}$ ), будем считать его текущим изображением (ТИ);  $f_{ЭИ}(m, l)$  - функция распределения яркости другого изображения, связанного некоторым образом с ТИ, но не обязательно повторяющее его, размером  $M \times L$  ( $M \leq N$ ;  $L \leq K$ ). Будем называть его эталонным изображением (ЭИ).

Первый этап синтеза алгоритма основан на том, что обрабатываемые изображения содержат избыточную информацию, устранение которой не несет за собой снижения вероятности и точности их взаимной корреляционной привязки. Снижение исходной информативности изображений является одним из способов уменьшения объема вычислений при корреляционной обработке. Возможный уровень снижения инфор-

мативности изображений определяется допустимой степенью ухудшения помехоустойчивости системы.

Второй этап синтеза корреляционных алгоритмов предполагает выбор критериальной функции  $K(p, g)$ . Наиболее часто используется классический алгоритм корреляционной обработки изображений, основанный на вычислении взаимной корреляции или интеграла типа свертки с последующим поиском максимума этой функции:

$$K(p, g) = \frac{1}{NK} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} f_{ЭИ}(i, j) f_{ТИ}(i+p, j+g), \quad (1)$$

где  $0 \leq p \leq N-M-1$ ;  $0 \leq g \leq K-L-1$  - смещение фрагмента  $f_{ЭИ}$  относительно  $f_{ТИ}$ .

Принято считать, что  $p$  и  $g$  отсчитывается от левого верхнего угла  $f_{ТИ}$ , для которого  $p=0$ ,  $g=0$ .

В рамках корреляционной обработки изображений можно также использовать в качестве меры сходства коэффициент корреляции:

$$P(p, g) = \frac{1}{\sigma(f_{ЭИ})\sigma(f_{ТИ})} \times \left[ \frac{1}{NK} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} (f_{ЭИ}(i, j) - \overline{f_{ЭИ}})(f_{ТИ}(i+p, j+g) - \overline{f_{ТИ}}) \right], \quad (2)$$

где  $\sigma(f_{ЭИ})$ ,  $\sigma(f_{ТИ})$  - среднеквадратичное отклонение яркости фрагментов  $f_{ЭИ}$  и  $f_{ТИ}$ ;  $\overline{f_{ЭИ}}$  и  $\overline{f_{ТИ}}$  - среднее значение отсчетов яркости изображений.

К недостаткам классического алгоритма следует отнести большой объем вычислений и большое количество локальных экстремумов взаимнокорреляционной функции. Нахождение глобального экстремума при совмещении АКИ возможно только в случае предварительной обработки совмещаемых изображений, производимой с целью устранения с них избыточной информации. Трудоемкость выполнения классического алгоритма может быть оценена следующим образом:

$$T_{КЛ} \cong N^2(M_{ПР}^2 + M_{СЛ}^2 + m_{ПР}^2 + m_{СЛ}^2), \quad (3)$$

где  $M_{ПР}$  - количество умножений, необходимых для вычисления функции корреляции,  $M_{СЛ}$  - количество сложений, необходимых для вычисления функции корреляции,  $m_{ПР}$ ,  $m_{СЛ}$  - количество умножений и сложений, необходимых для предварительной обработки изображений.

Если считать, что операция умножения выполняется в 5 раз дольше, чем операция сложения, а трудоемкость выполнения предварительного этапа обработки примерно равна трудоемкости вычисления корреляционной функции, то

$$T_{КЛ} \cong 52N^2M_{СЛ}^2. \quad (4)$$

Чтобы в некоторой степени снять остроту этой проблемы, корреляционный алгоритм иногда реализуется в спектральном пространстве преобразования Фурье:

$$K(p, g) = F^{-1} \{G(f_i, f_j)T(f_m, f_l)\}, \quad (5)$$

где  $F^{-1}$  - символ обратного преобразования Фурье;  $G(f_i, f_j)$  и  $T(f_m, f_l)$  - преобразования Фурье изображений  $f_{ТИ}(i, j)$  и  $f_{ЭИ}(m, l)$ .

Однако даже при реализации быстрого преобразования Фурье (БПФ) в цифровой форме не всегда удается достичь желаемого результата. Если потребовать, чтобы процесс корреляции имел быстрый алгоритм, а реализация его была наиболее простой, то мы придем к понятию диадной корреляции, которая может быть вычислена с использованием преобразования Уолша, которое, как известно, может быть в десятки раз быстрее, чем преобразование Фурье [2]. Функция вещественной корреляции (1) заменяется на диадную корреляцию:

$$K_D(p, g) = \frac{1}{NK} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} f_{ЭИ}(i, j) \times \\ \times f_{ТИ}(i \oplus p, j \oplus g), \quad (6)$$

где  $p$  и  $g$  - дискретный временной сдвиг в хемминговой метрике.

Использование диадной корреляции напрямую бывает затруднительно. Лучшие результаты могут быть получены при вычислении корреляционной функции через теорему корреляции, которая в диадном пространстве приводит к понятию вещественно-диадной свертки. Докажем следующую теорему.

**Теорема.** Если  $\{X(m)\}$  и  $\{Y(m)\}$  - цифровые последовательности с периодом  $N$ , то элементы корреляционной последовательности  $\{Z(s)\}$  будет определяться следующим образом:

$$Z(s) = \sum_{s=0}^{N-M-1} \sum_{u=0}^{N-1} C_Y(u)C_X(u), \quad (7)$$

где  $C_X(u)$ ,  $C_Y(u)$  - спектральные коэффициенты преобразования Уолша последовательностей  $\{X(m)\}$ ,  $\{Y(m)\}$ ;  $s = \overline{0, N-M-1}$  - вещественный сдвиг последовательности  $\{Y(n)\}$ ;  $u$  - переменная секвентного пространства.

**Доказательство.** Спектральные коэффициенты сдвигаемой последовательности  $Y(m+s)$  можно определить следующим образом:

$$Y(m+s) = \sum_{u=0}^{N-1} C_Y(u)wal_s(u), \quad (8)$$

где  $wal_s(u)$  - базисные функции системы Уолша.

С другой стороны,

$$Z(s) = \frac{1}{N} \sum_{s=0}^{N-M-1} \left[ \sum_{m=0}^{M-1} X(m)Y(m+s) \right]. \quad (9)$$

Подставив выражение (8) формулу (9), получим:

$$Z(s) = \frac{1}{N} \sum_{s=0}^{N-M-1} \left[ \sum_{m=0}^{M-1} X(m) \sum_{u=0}^{N-1} C_Y(u)wal_s(u) \right]. \quad (10)$$

Так как  $C_X(u) = \frac{1}{N} \sum_{m=0}^{M-1} X(m)wal_s(u)$ , то окончательно получим:

$$Z(s) = \sum_{s=0}^{N-M-1} \sum_{u=0}^{M-1} C_X(u)C_Y(u), \quad (11)$$

что и требовалось доказать.

Поскольку первая сумма в выражении (11) определяет сдвиг по  $s$ , то для вычисления  $Z(s)$  она не имеет принципиального значения. Окончательно получим:

$$Z(s) = \sum_{u=0}^{M-1} C_X(u)C_Y(u). \quad (12)$$

Полученное выражение будем называть вещественно-диадной сверткой.

**Применение вещественно-диадной свертки.** Для случая вычисления взаимнокорреляционной функции двух изображений формула для вещественно-диадной свертки принимает вид:

$$K_{ВД}(p, g) = \frac{1}{ML} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} [H_w F_{ЭИ}(m, l) H_w] \times \\ \times [H_w F_{ТИ}(m+p, l+g) H_w], \quad (13)$$

где  $H_w$  - матрица Адамара,  $F_{ЭИ}(m, l)$  и  $F_{ТИ}(m+p, l+g)$  - матрицы элементов эталонного и текущего изображений.

Совмещение текущего и эталонного изображений может осуществляться с ошибками округления координат точки привязки, поскольку вычисляемая оценка корреляционной функции  $K(p, g)$  представляет собой случайную функцию, статистические характеристики которой определяют, прежде всего, надежность совмещения. Надежность совмещения может быть оценена по следующей формуле:

$$r = M[K(p_n, g_n)] / [D[K(p_B, g_B)]]^{1/2},$$

где  $(p_n, g_n)$  и  $(p_B, g_B)$  - значения аргументов

оценки корреляционной функции, соответствующие точке истинного совмещения и одной из точек области боковых пиков корреляционной функции;  $M[K(p_n, g_n)]$  - математическое ожидание оценки корреляционной функции в точке истинного совмещения изображений;  $D[K(p_B, g_B)]$  - дисперсия оценки корреляционной функции в области боковых пиков.

Чем больше  $r$ , тем ниже вероятность ложного совмещения. Однако получить аналитические зависимости для расчета вероятности ложного совмещения по известной  $r$  не удастся. Результат привязки можно считать ложным при выполнении неравенства  $r < r^*$ , где  $r^*$  - допустимое значение оценки  $r$ , при котором результат вычисления корреляционной функции считается достоверным. Более удобным с практической точки зрения можно считать способ отбраковки ложных совмещений, основанный на оценке значений корреляционной функции в окрестности её максимума и второго локального экстремума

$$r = 1 - \frac{K_{02}}{K(i_{\max}, j_{\max})} \geq r^*,$$

где  $K_{02}$  - величина второго по значению экстремума корреляционной функции,  $K(i_{\max}, j_{\max})$  - значение корреляционной функции в её экстремуме  $(i_{\max}, j_{\max})$ .

**Методы совмещения изображений**

**на основе вещественно-диадной свертки**

**1. С устранением информационной избыточности высокочастотной фильтрацией.** Метод основан на устранении из коррелируемых спектров низкочастотных составляющих. Для этого на спектры отождествляемых фрагментов накладывается фильтр верхних частот вида:

$$H_{BЧ}(i, j) = \begin{bmatrix} 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 1 & 1 \\ 0 & 0 & 0 & \dots & 1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 1 & 1 & \dots & 1 & 1 \\ 1 & 1 & 1 & \dots & 1 & 1 \end{bmatrix} \text{ или}$$

$$H_{BЧ}(i, j) = \begin{cases} 0, & \text{если } i \leq a, j \leq b \\ 1, & \text{если } M-1 \leq i \leq a; L-1 \leq j \leq b' \end{cases}$$

где  $a$  и  $b$  - номера позиций в  $H_{BЧ}(i, j)$ , до которых низкочастотные секвенты принимаются равными нулю.

Выражение (13) в этом случае будет записываться следующим образом:

$$K_{BЧ}(p, g) = \frac{1}{ML} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} [H_w F_{ЭИ}(m, l) H_w H_{BЧ}] \times [H_w F_{ТИ}(m+p, l+g) H_w H_{BЧ}]. \quad (14)$$

Сечения двумерной функции корреляции  $K_{BЧ}(p, g)$  по координатам  $p$  и  $g$ , вычисленные по формулам (10) и (11) для изображений ТИ и ЭИ (рисунок 1), представлены на рисунке 2 (а, б, в, г) соответственно.

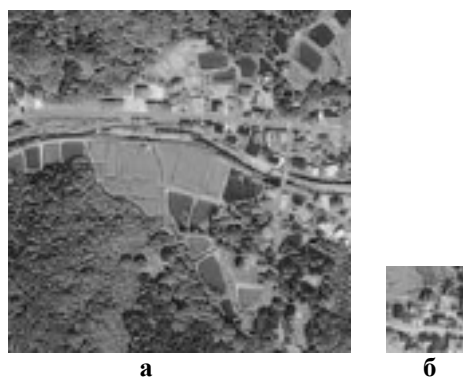


Рисунок 1 – Изображения ТИ (а) и ЭИ (б)

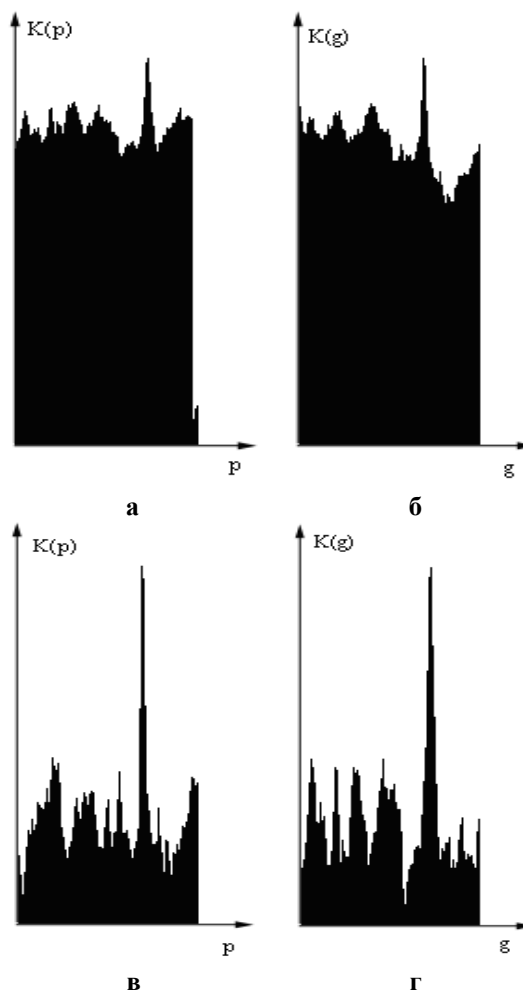


Рисунок 2 – Двумерная функция корреляции, вычисленная по формуле (13) [сечение по координате  $p$  - (а), сечение по координате  $g$  - (б)]. Вычисление по формуле (14) – сечения по координатам  $p$  и  $g$  (в, г)

Количество операций, необходимых для вычисления критериальной функции в соответствии с (14), можно оценить как

$$T_{BЧ} \cong 2N^2 M_{СЛ}^2 (\log_2 M + 13), \quad (15)$$

где принято, что  $M = L = 2^l$  и  $N = K = 2^n$ .

Если размер ЭИ равен  $32 \times 32$  пикселей, то

$$T_{BЧ} \cong 38N^2 M_{СЛ}^2, \quad (16)$$

что на 30 % меньше, чем трудоемкость, оцененная по формуле (4).

Результаты вычислений оценки  $r$  для  $a$  и  $b$ , изменяющихся в пределах  $0 \div 7$ , представлены в таблице.

**Результаты расчета оценки  $r$**

№ метода	$r$ при $a = b$							
	0	1	2	3	4	5	6	7
1	0,24	0,31	0,31	0,25	0,13	0,13	0,1	0,05
2	0,24	0,55	0,56	0,58	0,48	0,52	0,52	0,5
3	0,24	0,31	0,58	0,53	0,54	0,58	0,53	0,57

**2. С устранением информационной избыточности методом прореживания базисных функций [3].** Метод основан на устранении низкочастотных составляющих коррелируемых изображений в процессе вычисления их спектров. Для этого используется модифицированная матрица Адамара следующего вида:

$$H_W^M(i, j) = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 1 & 1 & 1 & \dots & -1 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & -1 & 1 & \dots & 1 & -1 \end{bmatrix}. \quad (17)$$

В этом случае выражение (13) можно записать следующим образом:

$$K_M(p, g) = \frac{1}{ML} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} [H_W^M F_{ЭИ}(m, l) H_W^{-M}] \times [H_W^M F_{ТИ}(m + p, l + g) H_W^{-M}], \quad (18)$$

где  $H_W^{-M}$  - модифицированная матрица Адамара, транспонированная по отношению к  $H_W^M$ .

Количество операций, необходимых для вычисления критериальной функции в соответствии с (14), можно оценить как

$$T_M \cong 36N^2 M_{СЛ}^2. \quad (19)$$

Результаты вычислений оценки  $r$  для  $a = b$ , изменяющихся в пределах  $0 \div 7$ , представлены в таблице. Выигрыш в количестве вычислений

может быть еще более очевиден, если учесть, что не все строки и столбцы в матрице Адамара участвуют в образовании критериальной функции.

**3. С использованием квазидвумерного спектрального представления коррелируемых изображений.** Операция получения секвентного спектра для применения вещественно-диадной свертки выражается следующим образом:

$$C_{ij}(u, v) = H_W F(i, j) H_W, \quad (20)$$

где  $C_{ij}$  - спектральные коэффициенты;  $F(i, j)$  - фрагмент изображения;  $H_W$  - матрица Адамара.

Выражение (20) может быть представлено в показательной форме:

$$C_{ij}(u, v) = \frac{1}{NK} \sum_{j=0}^{K-1} \sum_{i=0}^{N-1} F(i, j) (-1)^{\langle i, u \rangle + \langle j, v \rangle}, \quad (21)$$

где  $F(i, j)$  - входной массив изображения,  $C_{ij}(u, v)$  - спектральные коэффициенты преобразования;  $\langle i, u \rangle = \sum_{s=0}^{n-1} i(s)u(s)$  и  $\langle j, v \rangle = \sum_{s=0}^{n-1} j(s)v(s)$ ;  $i(s), j(s), u(s), v(s)$  - двоичные представления  $i, j, u$  и  $v$ , т.е.

$$[i]_{10} = [i(n-1), i(n-2), \dots, i(1), i(0)]_2.$$

В выражении (21) рассмотрим внутренне суммирование.

$$\frac{1}{N} \sum_{i=0}^{N-1} F(i, j) (-1)^{\langle i, u \rangle} = \frac{1}{N} \{ F(0, j) (-1)^{\langle 0, u \rangle} + F(1, j) (-1)^{\langle 1, u \rangle} + \dots + F(N-1, j) (-1)^{\langle N-1, u \rangle} \}. \quad (22)$$

В правой части полученного выражения записано  $(WH)_n$  каждой строки матрицы  $F(i, j)$ . Введем обозначение

$$\frac{1}{N} \sum_{i=0}^{N-1} F(i, j) (-1)^{\langle i, u \rangle} = C_i(u, j). \quad (23)$$

Тогда коэффициенты  $C_i(u, j)$  можно записать в матричном виде:

$$[C_i(u, j)] = \begin{bmatrix} C_0(0,0) & \dots & C_0(0, N-1) \\ C_1(1,0) & \dots & C_1(1, N-1) \\ \dots & \dots & \dots \\ C_{K-1}(K-1,0) & \dots & C_{K-1}(K-1, N-1) \end{bmatrix}. \quad (24)$$

Выражение (24) назовем квазидвумерным спектром изображения  $F(i, j)$ , поскольку оно представляет упорядоченные в виде двумерной матрицы одномерные спектральные коэффициенты строк изображения.

В соответствии с равенством Парсеваля можно записать [2]:

$$\frac{1}{NK} \sum_{i=0}^{N-1} \sum_{j=0}^{K-1} F(i, j)^2 = \sum_{u=0}^{N-1} \sum_{v=0}^{K-1} C_{ij}(u, v)^2 =$$

$$= \sum_{u=0}^{N-1} \sum_{j=0}^{K-1} C_i(u, j)^2. \quad (25)$$

Поскольку автокорреляционная функция является пространственным отображением спектральной плотности мощности, то можно сделать вывод о приемлемости применения квазидвумерных спектров изображений для нахождения критериальной функции:

$$K_{КВД}(p, g) = \frac{1}{ML} \sum_{m=0}^{M-1} \sum_{l=0}^{L-1} [H_W^M F_{ЭИ}(m, l)] \times \\ \times [H_W^M F_{ТИ}(m + p, l + g)]. \quad (26)$$

Сложность вычислений для данного метода можно оценить по формуле

$$T_{КВД} \cong 18N^2 M_{СЛ}^2. \quad (27)$$

Результаты вычислений оценки  $r$  представлены в таблице.

**Выводы.** Применение вещественно-диадной свертки позволяет строить эффективные экстремально-корреляционные системы для совмеще-

ния аэрокосмических изображений. В результате сравнения данных таблицы можно сделать вывод о надежности алгоритмов ( $r = 0,57$ ), основанных на квазидвумерных способах корреляции при получении выигрыша в объеме вычислений примерно в 3 раза по сравнению с классическим алгоритмом совмещения изображений. Приведенные результаты получены в процессе проведения исследований в рамках ФЦП «Научные и научно-педагогические кадры инновационной России» на 2009 - 2013 гг.

#### **Библиографический список**

1. Злобин В.К., Еремеев В.В. Обработка аэрокосмических изображений. – М.: ФИЗМАТЛИТ, 2006. – 288 с.
2. Ахмед Н., Рао К.Р. Ортогональные преобразования при обработке цифровых сигналов. – М.: Связь, 1980. – 248 с.
3. Колесенков А.Н., Костров Б.В. Метод прореживания базисных функций в корреляционно-экстремальных алгоритмах совмещения изображений // Вопросы радиоэлектроники. Сер. общетехническая. Выпуск 1. – 2010. С. 176 - 183.

УДК 004.93 1

**Е.М. Дондик, А.Н. Пылькин, Н.Е. Скоробогатова**

## **ИНФОРМАЦИОННАЯ СИСТЕМА РАСПОЗНАВАНИЯ ОБРАЗОВ ПЕРЕМЕННОЙ СТРУКТУРЫ**

*Рассматривается алгоритм распознавания образов, формируемых единой динамической системой с переменной структурой, в которой распознаваемые информационные признаки некоррелированы между собой и создаются комбинацией её составляющих, находящихся в непрерывном движении. Примером таких систем может быть лицо говорящего человека или рука, диктирующая речь глухих людей. Предлагается алгоритм селекции и последующего распознавания таких образов.*

**Ключевые слова:** образы переменной структуры, распознавание, вектор, контурный анализ, дактильная речь.

**Введение.** Общая теория распознавания достаточно полно разработана применительно к статическим и изменяющимся динамическим образам при возможности отслеживания объекта за счет медленного изменения его формы или поворота по известным уравнениям. Для таких динамических процессов известен фиксированный момент распознавания, и при этом обычно сохраняется определенная корреляция между последовательностями представления динамического образа.

Задача распознавания существенно усложняется, когда информационные отсчеты образа абсолютно не коррелированы между собой,

моменты распознавания не фиксированы и, в общем, неизвестны, а промежуточные переходные формы составляют большое количество комбинаций траекторий, включая даже формирование ложных образов.

Именно с полезными информационными образами должна работать система распознавания дактильной речи, которая призвана помочь людям с дефектами слуха и речи в общении с окружающими, в том числе при посещении поликлиник и других учреждений социальной сферы.

Предлагаемый алгоритм для наглядности удобно рассмотреть применительно к особен-

ностям представления знаков дактильной речи (дактилем). При этом распознающая система переменных образов должна учитывать ряд особенностей представления дактильной речи.

Во-первых, переменные образы поступают на вход системы непрерывной последовательностью, состоящей из информационных знаков и переходов между ними, что вытекает из правила плавной и слитной демонстрации дактильной речи. С учетом того, что дактилирование ведется в соответствии с нормами правописания, т.е. слова разделяются паузами, а фразы – остановкой, в распознающей системе необходимо предусмотреть также выделение этих пауз.

Во-вторых, реальная дактильная речь может демонстрироваться на различном расстоянии от видеокамеры и людьми с различным размером руки, взрослыми или детьми, и поэтому распознающая система должна работать с переменными образами изменяемого размера.

В-третьих, положение руки при демонстрации знака не фиксировано на поле кадра видеоизображения и может смещаться или наклоняться в зависимости от привычки демонстрирующего дактильную речь. Следовательно, распознающая система должна обеспечивать инвариантность по отношению к смещению или повороту переменного образа.

Все приведенные проблемы накладывают на систему распознавания образов переменной структуры достаточно сложные и отчасти противоречивые требования.

**Формулировка проблемы.** С учетом рассмотренных особенностей, а также, если подразумевается, что система должна быть единой, универсальной, и возникает проблема распознавания переменных образов определенной структуры. При этом в данном случае под переменными образами подразумевается видеоизменяемый объект, который обладает следующими особенностями:

- формирует счетное множество распознаваемых информационных образов, несущих конкретную информацию;

- все распознаваемые информационные образы объекта не коррелированы между собой и статистически независимы;

- траектория перехода от одного информационного образа к другому протекает во времени и включает определенное количество промежуточных дискретных отсчетов ложных образов, которые сохраняют взаимную коррелированность в основном между несколькими смежными отсчетами;

- траектория видеоизменения информационных образов, хотя и несет определенную

информацию, но из-за большого объема едва ли применима в данной системе и может рассматриваться как случайная.

**Цель работы.** В общем случае перед распознающей системой переменных образов встают две относительно автономных, но функционально координируемых задачи: первая – выделить информационный образ и вторая – распознать его. На первом этапе распознавания переменного образа осуществляется селекция информационного образа из видеопоследовательности изображений, в которой следует непрерывный ряд промежуточных переходных изображений, имеющих такую же структуру и фактически ничем не отличающихся от подлежащих распознаванию. Возникает достаточно неопределенная ситуация, когда необходимо из множества однотипных изображений с неизвестным (пока нераспознанным) знаком определить, где он находится в последовательности, и только потом распознать, какой это знак. Добавляет сложности при этом возможность того, что в переходной последовательности случайно может демонстрироваться один из распознаваемых знаков.

Таким образом, в общей системе, представленной на рисунке 1, входная информация подается на систему селекции образа (ССО), которая должна указать, какое из изображений последовательности представляет собой информационный дактильный знак для передачи его в систему распознавания (СПО), которая передает результат на систему воспроизведения (СВ).

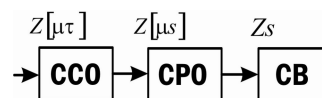


Рисунок 1 - Структурная схема системы распознавания

В качестве признака для селекции информационного образа предлагается воспользоваться априорно известной особенностью представления знака, которая позволяет его идентифицировать и отличить от промежуточной траекторной информации. Такими селектирующими признаками основного информационного образа относительно промежуточных (переходных) предлагается считать:

- скорость изменения параметров знака и его элементов;

- изменение траектории движения элементов информационного знака относительно фона;

- сравнительный анализ множества переходных траекторий от каждого информационного знака ко всем остальным информационным знакам;



– изменение направления движения элементов знака после демонстрации текущего основного распознанного образа.

Каждый из приведенных признаков характеризуется своими особенностями и непосредственно влияет не только на принцип построения, но и на структуру системы распознавания. Некоторые из таких систем требуют огромного объема вычислительных ресурсов или значительного времени распознавания. При этом время селекции информационного образа в реальном времени ограничивается минимальным интервалом демонстрации дактилем. В качестве наиболее простого и приемлемого по времени способа селекции информационного образа предлагается использовать изменение направления движения элементов знака. Следовательно, оператор селекции информационного образа призван выделить из непрерывной последовательности переходящих образов момент его демонстрации.

Входной информацией для оператора селекции служат кадры видеопоследовательности, которые подаются с определенной частотой так, чтобы иметь определенное число промежуточных образов между информационными. Число промежуточных образов зависит от логической структуры оператора и от взаимно корреляционных свойств последовательности промежуточных образов в формате изображений.

Для выделения режима молчания, когда наступает пауза между словами или предложениями, оператор селекции с целью исключения выделения ложного образа формирует дополнительную команду запрета распознавания. Система селекции фактически работает по своим алгоритмам автономно от системы распознавания и только формирует служебные команды момента распознавания и запрета на него. Выделенные из непрерывной последовательности переменных образов с помощью оператора селекции информационные образы передаются в распознающую систему в виде кадра статического изображения.

**Теоретические исследования. Алгоритм селекции.** Распознаваемые переменные образы, в реальности представляемые в аналоговой форме  $Z(t)$ , воспринимаются цифровой видеокамерой в виде цифровых изображений  $Z[t_b]$  с периодом  $t_b$ , определяемым параметрами съемки. Для работы системы такой объем информации избыточен, что может затруднить обработку. Поэтому для покадровой обработки выбирается период поступления кадров

$\tau = t_b/q$ , где  $q$  играет достаточно важную роль как для селекции, так и для распознавания образов. При выборе коэффициента деления кадров  $q$ , т.е. увеличения периода поступления переменного образа на вход системы, приходится исходить из двух противоположных требований. С одной стороны, необходимо обеспечить за средний период селекции  $s\tau$  достаточное число отсчетов, чтобы не пропустить существенные изменения образа, а с другой – большое число промежуточных отсчетов снижает чувствительность оператора селекции, т.к. изменения амплитуд сравниваемых векторов становятся соизмеримы с шумами дискретизации. Принимается, что порядковая нумерация отсчетов  $\mu$  в системе селекции начинается каждый раз с момента выделения текущего образа  $\mu = 1, 2, \dots, s$  и заканчивается в момент  $s\tau$  демонстрации следующего образа. При этом необходимо иметь в виду, что величина  $s$  непостоянна, может меняться от образа к образу и зависит от скорости представления образов или темпа демонстрации дактильной речи. Измеряемая величина  $s$  в случае превышения ею некоторого порога  $s \geq s_0$  может служить признаком окончания слова или фразы. В полной системе распознавания образов переменной структуры предлагается на этапе селекции использовать векторное веерное, с постоянным шагом по углу, представление контуров, как показано на рисунке 2.

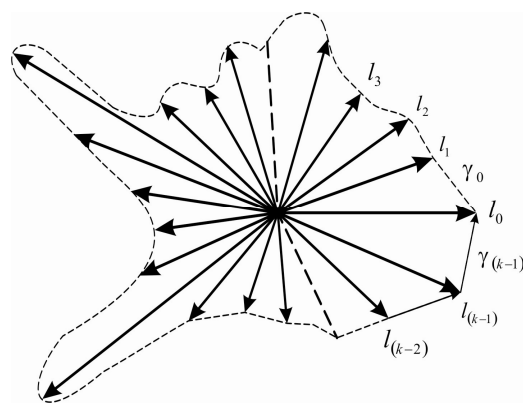


Рисунок 2 - Кодирование контура дактилемы веером векторов

В процессе функционирования в системе реализуются последовательные процедуры бинаризации и выделения замкнутых контуров дактильных знаков, приведение их к общему полюсу и определение длины каждого из составляющих  $l_n$  векторов, где  $n = 0, 1, \dots, (k-1)$ , а  $k$  определяет количество векторов в веере. Последовательность образов дактилем  $Z[\mu\tau]$  в

каждом интервале распознавания  $\mu = 0, 1, \dots, s$  будет представлено веером векторов  $l_n[\mu\tau]$  с постоянным шагом по углу. Таким образом, в комплекснозначном пространстве в момент  $\mu\tau$  дактилема полностью описывается множеством векторов:

$$Z[\mu\tau] = \{l_n[\mu\tau]\}_{0,(k-1)} = \left\{ l_n[\mu\tau] \exp\left(\frac{2\pi}{k} n\right) \right\}_{0,(k-1)}, \quad (1)$$

где  $n$  – номер текущего вектора,  $k$  – число векторов в веере.

Такое представление позволяет однозначно определять длину и координаты вектора, учитывая его действительную  $l_{n1} = \text{Re}(l_n)$  и мнимую  $l_{n2} = \text{Im}(l_n)$  составляющие.

При пересечении контура знака в нескольких точках за отсчет берется максимально удаленное пересечение, т.к. переходные образы сильнее всего изменяются на удаленных границах контура. Сам контур изображения знака принципиально может быть приближенно представлен в данном случае набором контурных векторов  $\gamma_n$ , связанных с текущими значениями двух смежных векторов в момент  $\mu\tau$  соотношением:

$$\gamma = l_{n+1} - l_n. \quad (2)$$

Контур играет важную роль в представлении и распознавании образов, однако на этапе селекции может не учитываться.

Оператор селекции признаков должен использовать не менее трех последовательных по времени отсчетов значений каждого из  $k$  векторов веера. Таким образом, для текущего вектора  $n$  оператор селекции  $G_n$  будет зависеть от величины трех значений длин вектора  $l_n$  в разные моменты времени:

$$G_n = G_n \{l_n[(\mu-1)\tau], l_n[\mu\tau], l_n[(\mu+1)\tau]\}. \quad (3)$$

Операция сравнения длин векторов осуществляется с некоторым порогом отсечения  $\eta$ , ниже которого разность не фиксируется. Величина порога выбирается с учетом дискретности представления переходных образов и в зависимости от выбранного коэффициента  $q$  деления кадров видеоканеры.

Входной информацией для селекции являются две разности вектора  $l_n$  между последовательностями отсчетов:

$$R_{n1} = l_n[(\mu-1)\tau] - l_n[\mu\tau],$$

$$R_{n2} = l_n[\mu\tau] - l_n[(\mu+1)\tau], \quad (4)$$

если они превышают порог отсечения  $\eta$ . При этом абсолютные значения величин знаков разностей  $R_{n1}$  и  $R_{n2}$  переводятся в фиксированные значения логических переменных  $L_{1r}, L_{1d}, L_{2r}, L_{2d}$  в соответствии с соотношениями:

$$L_{1r} = L_{2r} = \{1 \text{ при } R_{n1} > 0, R_{n2} > 0\}, 0 \text{ – иначе,}$$

$$L_{1d} = L_{2d} = \{1 \text{ при } R_{n1} < 0, R_{n2} < 0\}, 0 \text{ – иначе.} \quad (5)$$

Эти логические переменные в конечном счете и будут определять момент представления переменного образа, предназначенного для распознавания. Логическая структура оператора селекции информационного знака приводится на рисунке 3, где показана схема формирования команды распознавания  $f_n$  и служебной команды окончания слова с запретом распознавания  $h_n$ , только для одного вектора  $l_n$ .

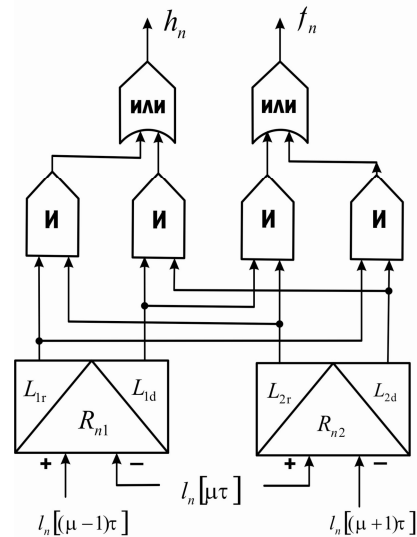


Рисунок 3 - Структура оператора селекции знака

При этом сами команды формируются при следующих соотношениях логических переменных:

$$f_n = 1, \text{ при } (L_{1r} \langle \text{«и»} L_{2d}) \text{ или } (L_{1d} \langle \text{«и»} L_{2r}),$$

$$h_n = 1, \text{ при } (L_{1r} \langle \text{«и»} L_{2r}) \text{ или } (L_{1d} \langle \text{«и»} L_{2d}). \quad (6)$$

Значения  $f_n$  и  $h_n$  используются для накопления сравнительных данных о каждом промежуточном образе, на основании которых формируется команда селекции  $F$ . Команда запрета  $H$  одновременно определяет момент движения элементов знака и поэтому может служить информацией о прерывании движения, т.е. об окончании слова или фразы.

Окончательное решение о селекции момента  $s\tau$  демонстрации информационного образа из

числа промежуточных  $\mu\tau$  принимается по максимуму  $F$  на основании анализа всего веера  $k$  векторов:

$$F = \sum_{n=0}^{k-1} f_n, \quad (7)$$

и с учетом значения комбинации логических данных команды  $H$ :

$$H = \sum_{n=0}^{k-1} h_n. \quad (8)$$

Комбинация этих двух команд позволяет также определять дактильные знаки с движущимися элементами в процессе распознавания.

**Алгоритм распознавания.** В основу русской дактильной речи положена одноручная система представления знаков алфавита, и поэтому для распознавания можно выбрать замкнутые одиночные контуры бинаризованного изображения. В рассматриваемой системе используется кодирование контура элементарными векторами, которые формируют вектор-контур каждой дактилемы [1,2,3].

Выделенный в растровом формате образ поступает на вход программы кодирования, которая выделяет контур изображения знака в бинарном представлении и осуществляет кодирование кадра. Таким образом, каждый знак представляется в закодированном виде цепочкой векторов в комплексном пространстве, позволяющем учесть направления каждого составляющего вектора. В таком же закодированном виде хранится обучающая выборка знаков, которая может в определенной мере подстраиваться под различные виды освещения и фона распознаваемых образов. Система распознавания осуществляет сравнение демонстрируемого знака со всеми эталонными, предварительно компенсируя различие в масштабе, повороте или смещении, и выдает результат на блок индикации, а также в систему речевого воспроизведения. Управление всей системой осуществляет блок синхронизации, который обеспечивает последовательное выполнение операции селекции знака, его кодирование и распознавание.

Ввод образа осуществляется веб-камерой в покадровом растровом полутоновом формате и обычно в цветном варианте. Для ускорения обработки выбирается наиболее информативный монохромный канал изображения, и оно переводится путем выбора уровня отсечения яркости пикселей в бинарный формат.

Каждый контур анализируемого образа  $\Gamma = \{\gamma(i)\}_k$  кодируется набором элементарных векторов  $\gamma(i)$   $i = 0, \dots, k-1$ , где  $k$  – составляет

число элементарных векторов в вектор-контуре. Таким образом, комплекснозначный код контура  $\Gamma = \{\gamma(i)\}_k$  будет представлен на растровой плоскости изображения последовательностью векторов как функцией дискретного комплексного переменного.

Распознавание знаков дактильной речи осуществляется сравнением их вектор-контуров в пространственной или частотной области.

В качестве решающей процедуры распознавания в пространственной области в системе используется вычисление нормированного скалярного произведения вектор-контура анализируемой дактилемы  $\Gamma$  и вектор-контура  $V$  каждого из числа хранящихся в базе данных знаков дактилем. Использование в качестве информативных признаков образов замкнутых вектор-контуров позволяет выполнить распознавание при изменении масштаба дактилемы и ее наклона. На основе скалярного произведения вводится понятие взаимно корреляционной функции этих контуров, которая является функцией переменного сдвига  $m$  между контурами. Для каждого значения  $m$  получается набор  $k$  скалярных произведений, которые и образуют взаимно корреляционную функцию:

$$T_{(m)} = \sum_{i=0}^{k-1} v(i)\gamma^*(i+m), \quad m = 0, 1, \dots, k-1.$$

Отсчеты взаимно корреляционной функции являются комплексными:

$$\begin{aligned} (V, \Gamma) &= \sum_{i=0}^{k-1} v(i)\gamma^*(i), \quad (V, \Gamma) = \\ &= \sum_{i=0}^{k-1} v^*(i)\gamma(i) = (V, \Gamma), \end{aligned} \quad (9)$$

где знак (\*) соответствует комплексносопряженному значению контура.

Практическая реализация процесса распознавания в пространственной области осуществляется с помощью сравнения анализируемого контура с пространственным контурным согласованным фильтром каждого образа. Признаком конкретного знака будет импульсная характеристика  $\lambda(m)$  фильтра размерностью  $k$ , у которой все составляющие элементарные векторы хранятся в запоминающем устройстве. Естественно, что пространственный согласованный фильтр для выделения вектор-контура  $\gamma(i)_{0,(k-1)}$  должен быть настроен на этот контур и иметь импульсную характеристику  $\gamma(m) = \gamma^*(i-m+k-1)$ , для которой выходной сигнал фильтра будет определяться соотношением:

$$u(m) = \sum_{i=0}^{k-1} v(i) \gamma(i-1+k-1). \quad (10)$$

Таким образом, фильтр, согласованный с комплекснозначным вектор-контуром, повторяет этот сигнал в обратном порядке, с задержкой на число отсчетов  $k$ , соответствующих длительности сигнала, и является комплексно-сопряженной копией этого сигнала.

Особенностью контурного спектрального анализа является возможность представить базисные функции разложения также в виде визуальных образов, а именно элементарных контуров. Это позволяет оперировать при спектральном анализе и фильтрации контуров визуально сопоставимыми образами.

Спектр  $P = \{p(m)\}_k$  вектор-контюра  $\Gamma = \{\gamma(i)\}_k$  в форме ДПФ будет:

$$p(m) = \sum_{i=0}^{k-1} \gamma(i) \exp\left\{-j \frac{2\pi}{k+1} mi\right\}, \quad (11)$$

$m=0, \dots, (k+1)$ , где элементарные векторы спектра  $p(m)$  являются комплексными числами, и при этом их упорядоченная последовательность  $\{p(0), \dots, p(k-1)\}$  также образует на плоскости спектральный векторный контур  $P$ . Таким образом,  $p(m)$  получаются суммированием всех элементарных векторов исходного контюра  $\Gamma$  с поворотом каждого из них на угол  $\left(\frac{2\pi}{k} mi\right)$ .

В спектральном виде произвольный контур  $\Gamma$  может быть представлен линейной комбинацией элементарных контуров  $\Gamma_0, \Gamma_1, \dots, \Gamma_{k-1}$ , весами в которой служат компоненты спектра. Спектр контюра на выходе согласованного фильтра определяется как:

$$p_{\text{вых}}(m) = |p(m)|^2 \exp\left\{j \frac{2\pi}{k} m\right\}, \quad (12)$$

при  $m=0, \dots, (k-1)$ , т.е. через квадрат модуля частотных компонент  $|p(m)|$ .

Частотный коэффициент передачи фильтра будет равен  $\Omega = \{\omega(m)\}_{0, k-1}$ .

Контурный пространственный сигнал на входе фильтра  $\Gamma = \{\gamma(i)\}_{0, (k-1)}$  формирует на выходе фильтра текущий сигнал  $H = \{u(s)\}_{0, (k-1)}$  с частотным коэффициентом передачи  $\Omega = \{\omega(m)\}_{0, k-1}$  в виде:

$$H = \frac{1}{k} \sum_{i=0}^{k-1} \omega(m) p_{\text{ex}}(m) \Gamma_m = \frac{1}{k} \sum_{m=0}^{k-1} p_{\text{вых}} \Gamma_m, \quad (13)$$

т.е. выходной сигнал в виде вектор-контюра  $H$  равен взвешенной сумме элементарных контуров  $\Gamma_m$ ,  $m=0, \dots, (k-1)$ , весами в которой выступают элементарные векторы  $p_{\text{вых}}(m)$  контюра спектра.

Для распознавания каждого знака используется свой контурный согласованный фильтр, параметры которого хранятся в памяти системы распознавания. Составляющая частотного коэффициента передачи определяется в виде:

$$\omega(m) = P_A^*(m) \exp\left\{j \frac{2\pi}{k} m\right\}, \quad (14)$$

где  $P_\Gamma = \{p_\Gamma(m)\}$  – спектр входного сигнала,  $P_\Gamma^*(m)$  – комплексно-сопряженное значение составляющих спектра входного сигнала.

Если спектр фильтруемого контюра согласован с фильтром:

$$p_{\text{вых}}(m) = |p(m)|^2 \exp\left\{j \frac{2\pi}{k} m\right\}, m=0, 1, \dots, k-1, \quad (15)$$

то сигнал на выходе фильтра в момент  $m=k-1$  будет вещественной величиной и максимально возможным по модулю.

Таким образом, при поступлении на вход фильтра последовательности контуров с одинаковой энергией максимальный по модулю отсчет будет только у контюра, согласованного с фильтром. При этом в фильтре реализуются только линейные преобразования, и, следовательно, контурный согласованный фильтр вырабатывает количественную меру сходства и позволяет распознать контюры дактильных образов, сохраняя свойство инвариантности к изменению масштаба, сдвигу начальных точек контуров и повороту изображений.

**Выводы.** Рассмотренный алгоритм распознавания образов переменной структуры показывает теоретическую и практическую возможность построения таких систем, которые в сочетании с уже реализованными в мобильной связи процедурами видеосъемки и воспроизведения могут предоставить дополнительные функции звукового озвучивания дактильной речи.

#### Библиографический список

1. Фурман Я.А. Согласованная фильтрация контуров изображений // Радиотехника.-1995.-№ 6.-С. 30-33
2. Furman Ya.A., Yanshin V.V. Extraction and Linear Filtering of Closed Polygonal Contours of Images// Pattern Recognition and Image Analysis.-1994.-V.4, № 2.
3. Дондик Е.М., Скоробогатова Н.Е. Система распознавания и воспроизведения знаков дактильной речи // Сб. науч. трудов «Инновационные подходы к применению информационных технологий в профессиональной деятельности», НАЧОУ ВПО СГА. – Белгород: ГИК, 2010.-С. 287-291.

УДК 621.396.96

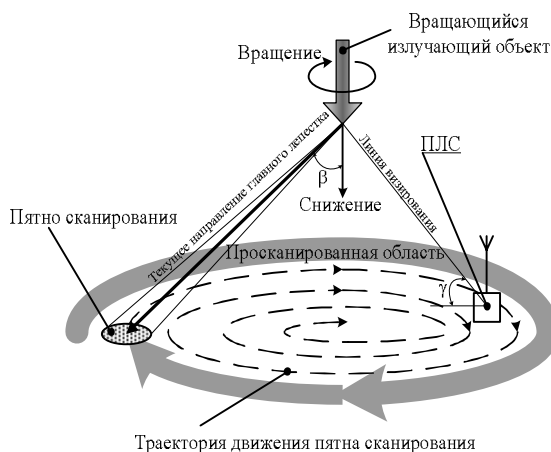
**В.Г. Андреев****МЕТОД ОБНАРУЖЕНИЯ СИГНАЛОВ  
ОТ НЕПРЕРЫВНО ИЗЛУЧАЮЩИХ ВРАЩАЮЩИХСЯ ОБЪЕКТОВ**

*Предложен и исследован метод обнаружения радиосигналов, излучаемых вращающимся непрерывно излучающим объектом при частично известных параметрах его вращения. Показано, что предлагаемый метод дает возможность уменьшить на 1,5...3 дБ пороговое отношение сигнал-шум при обнаружении излучения по сравнению с известными методами обнаружения. Данный выигрыш дает возможность увеличить в 1,3...1,4 раза дальность обнаружения излучения от вращающегося непрерывно излучающего объекта за счёт использования дополнительных процедур фильтрации и нормировки её результатов, основанных на априорных сведениях о параметрах вращения объекта.*

**Ключевые слова:** преобразование Фурье, спектр, накопление сигнала, многоканальный фильтр.

**Введение.** Обнаружение вращающихся непрерывно излучающих объектов пассивными локационными станциями (ПЛС) [1], работающими в тепловом и радиодиапазонах электромагнитных волн, представляет собой задачу, характерную для спектрального анализа. Её специфика связана с тем, что непрерывно излучающий вращающийся объект облучает неподвижную ПЛС периодически за счет циклического изменения положения диаграммы направленности антенны излучающего объекта.

Примерами таких вращающихся излучающих объектов могут служить самоприцеливающиеся боеприпасы типа XM898 «SADARM» (Sense and Destroy Armor) [2], которые, снижаясь, совершают вращательные движения с целью обеспечения спирального сканирования земной поверхности (см. рисунок 1) [3].

**Рисунок 1**

Главный лепесток диаграммы направленности самоприцеливающегося боеприпаса отклонен от вертикали для обеспечения заданной ширины зоны сканирования. Отметим, что ПЛС может обнаружить вращающийся излучающий объект не только по его излучению по главному, но и по боковым лепесткам диаграммы направленности антенны объекта.

Известен [4] метод обнаружения подобных излучений, заключающийся в том, что осуществляют аналого-цифровое преобразование (АЦП) детектированного сигнала и получают его дискретные выборки, складывают возведенные в квадрат выборки и получают их сумму. Аналогичным образом измеряют уровень шумов в дополнительном канале измерения уровня шумов, снабженном дополнительным антенным устройством. Нормируют сумму к измеренному уровню шумов и получают нормированную сумму. По превышению порога обнаружения величиной нормированной суммы принимают решение об обнаружении радиосигналов.

Недостатком подобного подхода является сложность его реализации из-за необходимости введения дополнительного канала измерения уровня шумов для стабилизации вероятности ложной тревоги путем выбора порога обнаружения. Кроме того, данный метод не использует априорную информацию о периодичности обнаруживаемого радиосигнала, что приводит к потерям в вероятности его правильного обнаружения. Наличие дополнительного канала измерения уровня шумов снижает помехоустойчивость из-за неэквивалентности основного и

дополнительного каналов за счет, например, различий в диаграммах направленности антенн основного и дополнительного каналов.

Альтернативный метод [5] обнаружения излучений заключается в том, что осуществляют АЦП детектированного сигнала и получают его дискретные выборки, складывают возведенные в квадрат выборки и получают их сумму. Сумму запоминают. Отключают приемный тракт от антенного устройства и измеряют аналогичным образом уровень шумов.

Нормируют сумму к измеренному уровню шумов путем её деления на измеренный уровень шумов и получают нормированную сумму. По превышению порога обнаружения величиной нормированной суммы принимают решение об обнаружении радиосигналов.

Недостатком описанного метода являются значительные временные затраты для измерения мощности собственных шумов приемного тракта с целью стабилизации уровня ложных тревог. В течение всего дополнительного времени приемный тракт не участвует в обнаружении, что приводит к пропускам обнаруживаемого альтернативным методом сигнала. Кроме того, информация о периодичности обнаруживаемого радиосигнала не используется. Перечисленные недостатки приводят к тому, что вероятность правильного обнаружения сигналов альтернативным методом меньше, чем методом [4], использующим два приёмных канала.

**Цель работы** — повышение эффективности обработки сигналов, излученных вращающимся непрерывно излучающим объектом, путем использования спектральных свойств сигналов и шумов при обнаружении излучения введением дополнительных процедур их фильтрации и нормировки её результатов. При этом используется свойство равномерного распределения плотности мощности шума по спектру и сосредоточения в узком частотном диапазоне спектральной плотности мощности периодического за счет вращения излучающего объекта сигнала от него.

**Описание сигнала.** На рисунке 2 представлена временная зависимость  $P(t)$  нормированной мощности  $P$  экспериментального сигнала от вращающегося излучающего объекта с характеристиками, аналогичными самоприцеливающемуся боеприпасу ХМ898: частота  $f$  вращения боеприпаса  $f=25$  Гц, скорость снижения 25 м/с, ширина  $\Delta\psi$  главного лепестка диаграммы направленности по уровню 0,5 по мощности  $\Delta\psi=5^\circ$ , форма диаграммы направленности вида  $(\sin x)/x$ , угол  $\beta$  отклонения центра главного лепестка диаграммы направленности антенны

объекта от вертикали  $\beta=30^\circ$ , скорость ветра около 3 м/с, время наблюдения 0,1 с, высота объекта — около 130 м, раскачивание объекта незначительное, угол  $\gamma$  визирования объекта  $\gamma \approx 62^\circ$  (см. рисунок 1).

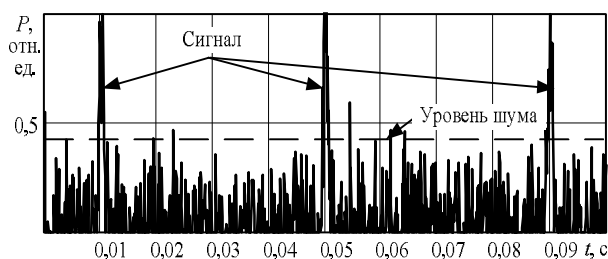


Рисунок 2

Периодический характер процесса  $P(t)$ , очевидный из анализа рисунка 2, обуславливает сосредоточение полезной компоненты сигнала в узком спектральном диапазоне, однако период вращения излучающего объекта, точные форма и положение диаграммы направленности его антенны, угол  $\gamma$  визирования априорно неизвестны, что препятствует построению согласованного с сигналом фильтра его обработки. Поэтому используются некогерентные системы обнаружения излучения от непрерывно излучающих вращающихся объектов, реализующие принцип энергетического приемника [5].

**Метод решения.** Увеличение вероятности  $D$  правильного обнаружения сигналов, излученных вращающимся излучающим объектом, а также исключение дополнительного канала измерения уровня шумов достигается тем, что (см. рисунок 3) отсчеты  $x_l$ ,  $l=0, 1, \dots, (L-1)$  выборки  $x=[x_l]$  детектированного сигнала, полученные после аналого-цифрового преобразования (АЦП) детектированного сигнала, запоминают в устройстве памяти, которое организовано как сдвиговый регистр, подвергают выборку  $x$   $L$ -точечному дискретному преобразованию Фурье (ДПФ) и формируют  $L$  спектральных комплексных отсчетов  $c_l$ ,  $l=0, 1, \dots, (L-1)$ .

Каждый из  $L$  спектральных комплексных отсчетов  $c_l$  подвергают взятию квадрата модуля и получают  $L$  спектральных амплитудных отсчетов  $s_l$ ,  $l=0, 1, \dots, (L-1)$

$$s_l = |c_l|^2. \quad (1)$$

В полученном массиве  $s_l$  выявляют максимальный спектральный амплитудный отсчет  $s_{\max}$  из всех  $L$  спектральных амплитудных отсчетов  $s_l$ :

$$s_{\max} = \max(s_0, s_1, \dots, s_{L-1}). \quad (2)$$

Вид нормированного к своему максимальному значению  $s_{\max}$  массива  $s_l/s_{\max}$  показан на рисунке 4.

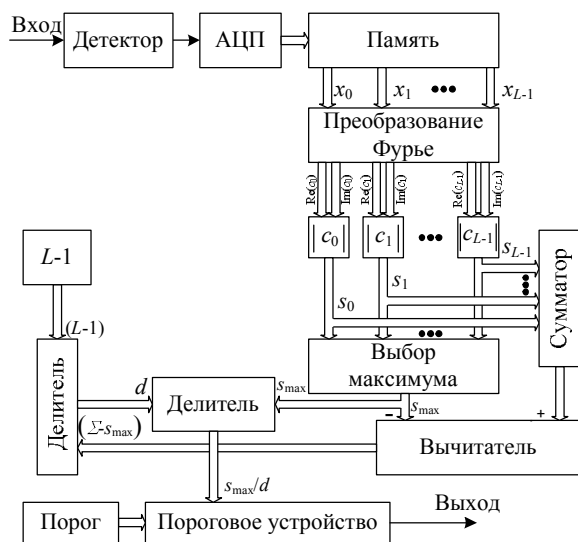


Рисунок 3

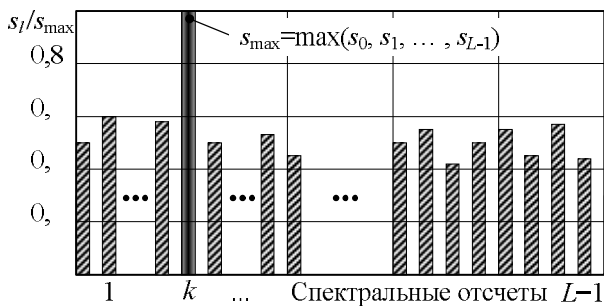


Рисунок 4

Предполагается, что  $s_{max}$  — это полезный сигнал, а остальные  $(L-1)$  спектральных амплитудных отсчетов — это шум. После суммирования всех  $L$  спектральных амплитудных отсчетов  $s_i$  получают их сумму  $\Sigma$ :

$$\Sigma = \sum_{i=0}^{L-1} s_i, \tag{3}$$

а путём вычитания из суммы  $\Sigma$  максимального спектрального амплитудного отсчета  $s_{max}$  получают разность  $(\Sigma - s_{max})$ , которая характеризует сумму амплитудных значений  $(L-1)$  отсчетов шума.

Тогда оценку  $d$  уровня шума можно получить делением разности  $(\Sigma - s_{max})$  на уменьшенное на единицу число  $(L-1)$  спектральных отсчетов:

$$d = (\Sigma - s_{max}) / (L-1). \tag{4}$$

При этом используется свойство равномерности распределения мощности белого гауссовского шума во всем спектральном диапазоне. Максимальный спектральный амплитудный отсчет  $s_{max}$  нормируется к оценке  $d$  уровня шума путем их деления. Получаемое частное  $s_{max}/d$  характеризует соотношение полезного сигнала и уровня шума:

$$s_{max}/d = \frac{(L-1)s_{max}}{\Sigma - s_{max}}. \tag{5}$$

По превышению порога обнаружения величиной частного  $s_{max}/d$  принимают решение об обнаружении радиосигналов, излученных непрерывно излучающим вращающимся объектом.

Отметим, что величина порога не зависит от уровня шума ввиду его оценки  $d$ , входящей в частное (5), величина которого сравнивается с порогом. Это дополнительно упрощает реализацию рассматриваемого метода, т.к. возможно использование фиксированного порога, величина которого зависит лишь от заданной вероятности  $F$  ложной тревоги и не требует адаптации к уровню шумов.

Вычисление ДПФ дискретной временной выборки  $x$  может осуществляться, например, путем домножения временной выборки  $x$  на векторы  $f_l$  преобразования Фурье:

$$c_l = x^T f_l,$$

где  $x = [x_l]$ ,  $f_l = [f_{n,l}] = [e^{-i 2\pi n l / L}]$ ,  $n = 0, 1, \dots, (L-1)$ ,  $l = 0, 1, \dots, (L-1)$ ,  $^T$  — знак транспонирования; компоненты  $f_{n,l}$  векторов  $f_l$  могут быть рассчитаны заранее и храниться в запоминающем устройстве.

Используя предложенный метод, удаётся достичь по сравнению с известными подходами к решению аналогичных задач следующих преимуществ:

1) исключить дополнительный канал оценки уровня шума за счет её осуществления в спектральной области по выражению (4) в едином канале обнаружения;

2) исключить необходимость адаптации порога обнаружения к уровню шума при обнаружении радиосигнала от вращающегося излучающего объекта за счет использования при нормировке (5) оценки уровня шума, полученной в едином канале обнаружения в спектральной области по выражению (4).

**Оценка эффективности.** Рассмотрим в качестве показателя эффективности предлагаемого метода и его аналогов [4, 5] характеристику обнаружения, т.е. зависимость  $D(Q)$  вероятности  $D$  правильного обнаружения от отношения  $Q$  сигнал-шум при фиксированной величине вероятности  $F$  ложной тревоги, что соответствует принятому в радиолокации критерию Неймана — Пирсона.

Для определения степени преимущества предлагаемого метода по сравнению с известным [4], проведено статистическое моделирование методом Монте-Карло при следующих параметрах:  $F = 0,01$  — вероятность ложной тревоги системы обнаружения; число повто-

рений опытов  $N=10000$ ; частота  $f$  вращения излучающего объекта  $f=25$  Гц, ширина  $\Delta\psi$  по уровню 0,5 по мощности направленного на приемное устройство системы обнаружения лепестка антенны вращающегося объекта  $\Delta\psi=5^\circ$ , время наблюдения 1 с, число  $L$  частотных отсчетов  $L=256$ .

Результаты моделирования изображены на рисунке 5, где зависимости  $D(Q)$  для предлагаемого метода показаны сплошными линиями, а для известного — пунктиром.

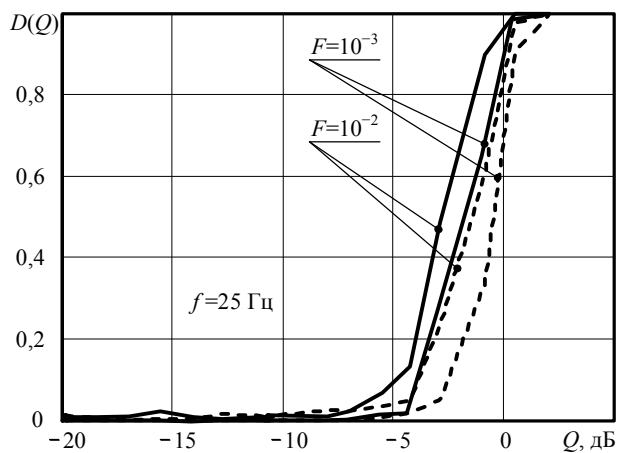


Рисунок 5

Анализ характеристик обнаружения, представленных на рисунке 5, показывает, что предлагаемый метод при обеспечении вероятности правильного обнаружения  $D=0,8$  имеет выигрыш  $\Delta q$  в пороговом отношении сигнал-шум по сравнению с известным методом  $\Delta q=2...3$  дБ, что обеспечивает увеличение дальности обнаружения вращающегося излучающего объекта в 1,3...1,4 раза.

При уменьшении вероятности ложной тревоги до  $F=0,001$  выигрыш  $\Delta q$  сохраняется и составляет около 2 дБ, что обеспечивает увеличение дальности обнаружения излучающего объекта до 1,2 раз (см. рисунок 5). Отметим, что наращивание частоты  $f$  вращения излучающего объекта до  $f=50$  Гц повышает достигаемые при  $F=0,001$  выигрыши. Так, при сохранении прочих перечисленных выше

параметров выигрыш составляет  $\Delta q=3$  дБ. Это обеспечивает увеличение дальности обнаружения излучающего объекта до 1,4 раз по сравнению с известным [4] методом.

**Выводы.** Таким образом, предложен метод повышения эффективности обнаружения радиосигналов, излучаемых вращающимся непрерывно излучающим объектом при частично известных параметрах его вращения. Показано, что предлагаемый метод дает возможность уменьшить на 1,5...3 дБ пороговое отношение  $Q$  сигнал-шум по сравнению с известными методами обнаружения. Данный выигрыш в пороговом отношении дает возможность увеличить в 1,3...1,4 раза дальность обнаружения излучения от вращающегося непрерывно излучающего объекта за счёт использования дополнительных процедур фильтрации и нормировки её результатов, основанных на априорных сведениях о параметрах вращения объекта. Кроме того, используя предложенный метод, удаётся исключить: 1) дополнительный канал оценки уровня шума за счёт её осуществления в спектральной области; 2) необходимость адаптации порога обнаружения к уровню шума путём использования при нормировке его оценки, полученной в едином канале обнаружения.

#### Библиографический список

1. Пассивная радиолокация. Методы обнаружения объектов / Р.П. Быстров, Г.К. Загорин, А.В. Соколов, Л.В. Федорова.— М.: Радиотехника, 2008.— 320 с.
2. Строев В. Кассетные боеприпасы с самоприцеливающимися боевыми элементами // Зарубежное военное обозрение.— 2000.— № 8.— С. 20-25.
3. Чубасов В.А. Высокоточные боеприпасы: учеб. пособие для вузов.— СПб.: БГТУ "ВОЕНМЕХ", 2008.— 116 с.
4. Караваев В.В., Сазонов В.В. Статистическая теория пассивной локации.— М.: Радио и связь, 1987.— 240 с.
5. Николаев А.Т., Перцов С.В. Радиотеплолокация. Пассивная радиолокация.— М.: Сов. радио, 1964.— 327 с.