

УДК 621.397

АНАЛИЗ ЭФФЕКТИВНОСТИ ПРИМЕНЕНИЯ ОРТОГОНАЛЬНЫХ ПРЕОБРАЗОВАНИЙ ДЛЯ КОДИРОВАНИЯ ДИСКРЕТНЫХ СИГНАЛОВ С КОРРЕЛИРОВАННЫМИ ОТСЧЕТАМИ

Умняшкин С.В.

Введение

Пусть $\mathbf{X} = (X_0, X_1, \dots, X_{N-1})^T$ - некоторый случайный вектор (дискретный сигнал). Требуется найти такой способ эффективного цифрового кодирования вектора \mathbf{X} , чтобы при заданной длине L (бит) двоичного кода полученный после декодирования вектор $\hat{\mathbf{X}} = (\hat{X}_0, \hat{X}_1, \dots, \hat{X}_{N-1})^T$ представлял исходный с минимальной ошибкой $\varepsilon^2 = M(\|\mathbf{X} - \hat{\mathbf{X}}\|^2)$. (Здесь и далее в качестве векторной нормы используется евклидова норма.)

Вследствие наличия межкомпонентных связей независимое покомпонентное статистическое кодирование отсчетов дискретного сигнала, следующее за их скалярным квантованием, порождает избыточные, неэффективные коды. Общий подход, который используется для повышения эффективности кодирования дискретных сигналов, состоит в предварительной обработке исходных данных с помощью обратимого дискретного преобразования с матрицей \mathbf{W} , переводящего вектор \mathbf{X} в некоторый вектор $\mathbf{Y} = \mathbf{W}\mathbf{X}$, в котором зависимость между компонентами ослаблена. Тогда независимое покомпонентное кодирование вектора \mathbf{Y} , а не вектора \mathbf{X} , становится обоснованным. Вектор \mathbf{Y} будем называть вектором трансформант, или обобщенным дискретным спектром.

Ограничим рассмотрение статистических зависимостей между отсчетами вектора \mathbf{X} рамками корреляционной модели, считая, что для исходного дискретного сигнала \mathbf{X} заданы вектор математических ожиданий компонент $\mathbf{m}_X = M(\mathbf{X}) = (m_{X_0}, \dots, m_{X_{N-1}})^T$ и ковариационная матрица $\mathbf{K}_X = (\text{cov}(X_k, X_j))_{k,j=0}^{N-1}$. Тогда смысл использования дискретного преобразования для кодирования сигналов состоит в том, чтобы получить вектор трансформант $\mathbf{Y} = \mathbf{W}\mathbf{X}$ с некоррелированными компонентами. Этого можно добиться, если матрица $\mathbf{W} = \mathbf{W}_{\text{opt}}$ составлена из собственных векторов матрицы \mathbf{K}_X - соответствующее дискретное ортогональное преобразование называется преобразованием Карунена-Лозва (или Хоттелинга). Ковариационная матрица \mathbf{K}_Y вектора \mathbf{Y} имеет в этом случае диагональный вид. Однако отсутствие быстрых алгоритмов вычисления и зависимость параметров оптимального преобразования Карунена-Лозва от структуры матрицы \mathbf{K}_X существенно ограничивают практические возможности его применения и вынуждают использовать другие преобразования. Возникает вопрос: насколько эффективно применение того или иного преобразования $\mathbf{Y} = \mathbf{W}\mathbf{X}$ по

Статья посвящена проблемам оценки эффективности использования различных ортогональных преобразований статических и динамических изображений. Предложен оригинальный критерий этой оценки.

сравнению с оптимальным преобразованием Карунена-Лозва? Как численно оценить эту эффективность для того, чтобы выбрать лучший из возможных вариантов обработки?

Энтропия случайной величины

Если источник сообщений ξ имеет N возможных состояний $\{\xi_0, \xi_1, \dots, \xi_{N-1}\}$ и принимает их вне зависимости от предыдущего состояния с вероятностями $\{p_0, p_1, \dots, p_{N-1}\}$, то двоичной энтропией источника согласно теоремы Шеннона о кодировании для канала без помех [1] является величина:

$$H(\xi) = - \sum_{k=0}^{N-1} p_k \log_2 p_k \quad (\text{бит}). \quad (1)$$

Энтропия (1) представляет собой минимально возможные средние битовые затраты R , которые необходимы для представления одного символа сообщения и является точной нижней оценкой для средней длины кода одного символа: $R \geq H$.

Следуя [1], рассмотрим возможность применения понятия энтропии к случайной величине непрерывного типа X (соответствующей непрерывному источнику информации), которая задана функцией плотности вероятности $f(x)$. Для упрощения выкладок положим, что случайная величина X может принимать значения из конечного диапазона, соответственно, функция плотности распределения вероятностей $f(x)$ отлична от нуля на конечном интервале. Положим, что $f(x) \neq 0$ непрерывно при $x \in [0, 1]$, и

$$\int_0^1 f(x) dx = 1.$$

Равномерно разобьем интервал $[0, 1]$ на N участков $\Delta_j = [ju, (j+1)u)$ длины $u = 1/N$, $j = 0, 1, \dots, N-1$. Случайной величине X , попавшей в интервал Δ_j , поставим в соответствие значение $x_j = u(j + \frac{1}{2})$. Получаем из непрерывной величины X равномерно проквантованную с шагом u дискретную случайную величину \hat{X} , энтропия (1) которой равна:

$$H(\hat{X}) = - \sum_{j=0}^{N-1} \int_{\Delta_j} f(x) dx \log_2 \int_{\Delta_j} f(x) dx = - \sum_{j=0}^{N-1} f(\theta_j) \Delta_j \cdot \log_2 (f(\theta_j) \Delta_j) =$$

$$\begin{aligned}
&= -\sum_{j=0}^{N-1} f(\theta_j) u \log_2 f(\theta_j) - \sum_{j=0}^{N-1} \int_{\Delta_j} f(x) dx \log_2 u = \\
&= -\sum_{j=0}^{N-1} f(\theta_j) \log_2 f(\theta_j) \cdot u - \log_2 u', \quad (2)
\end{aligned}$$

где некоторая точка $\theta_j \in \Delta_j$. При этом математическое ожидание квадрата ошибки квантования $\varepsilon^2 = M((X - \hat{X})^2)$:

$$\begin{aligned}
\varepsilon^2 &= \sum_{j=0}^{N-1} \int_{\Delta_j} (x - (j + \frac{1}{2})u)^2 f(x) dx = \\
&= \sum_{j=0}^{N-1} f(\xi_j) \int_{\Delta_j} (x - (j + \frac{1}{2})u)^2 dx = \frac{u^2}{12} \sum_{j=0}^{N-1} f(\xi_j) u, \quad (3)
\end{aligned}$$

где некоторая точка $\xi_j \in \Delta_j$.

Увеличивая число разбиений N , в пределе получим для энтропии (2):

$$\lim_{N \rightarrow \infty} H(\hat{X}) = \lim_{u \rightarrow 0} H(\hat{X}) = - \int_a^b f(x) \log_2 f(x) dx - \lim_{u \rightarrow 0} \log_2 u \rightarrow \infty \quad (4)$$

(в предположении существования стоящего в выражении интеграла). При этом для ошибки (3):

$$\lim_{u \rightarrow 0} \frac{\varepsilon^2}{u^2/12} = \lim_{u \rightarrow 0} \sum_{j=0}^{N-1} f(\gamma_j) u = \int_0^1 f(x) dx = 1.$$

Таким образом, при стремлении шага (дискрета) и равномерного квантования к нулю ошибка квантования (3) есть бесконечно малая величина, $\varepsilon^2 \sim u^2/12$ для любой функции $f(x)$, т.е. независимо от закона распределения.

Число бит, необходимое для кодирования дискретной величины \hat{X} , с уменьшением шага квантования $u \rightarrow 0$ непрерывной величины X бесконечно возрастает: $H(\hat{X}) \sim -\log_2 u$. Однако, если имеются две случайных величины X и Y , определяемые на интервале $t \in [0,1)$ функциями плотности вероятности $f(t)$ и $g(t)$ соответственно, то при бесконечно малом шаге разбиения интервала $[0,1)$ существует предел разности для значений энтропии (4) дискретных величин \hat{X} и \hat{Y} (в случае существования соответствующих интегралов):

$$\lim_{u \rightarrow 0} (H(\hat{X}) - H(\hat{Y})) = - \int_0^1 f(t) \log_2 f(t) dt + \int_0^1 g(t) \log_2 g(t) dt \quad (5)$$

при этом $\varepsilon^2(\hat{X}) \sim \varepsilon^2(\hat{Y}) \sim u^2/12$.

Для непрерывной случайной величины X с функцией плотности вероятности $f(x)$ (заданной в общем случае на всей числовой оси) величина

$$H(X) = - \int_{-\infty}^{+\infty} f(x) \log_2 f(x) dx \quad (6)$$

является дифференциальной энтропией [1]. Данное название подчеркивает тот факт, что реальный смысл несут именно разности значений (6): если $H(X) < H(Y)$, то при достаточно малом шаге квантования u ошибки квантования эквивалентны, $\varepsilon^2(\hat{X}) \approx \varepsilon^2(\hat{Y}) \approx u^2/12$, но, как следует из (5), для кодирования величины \hat{Y} требуется примерно на $H(Y) - H(X)$ большее число бит, чем необходимо для кодирования \hat{X} .

Критерий средней избыточной энтропии [4]

Рассмотрим вектор $\mathbf{Y} = \mathbf{W}\mathbf{X}$, полученный в результате дискретного ортогонального преобразования, для

которого матрица обратного преобразования $\mathbf{W}^{-1} = \mathbf{W}^T$. Пусть $f_k(m_k, \sigma_k, x)$ - функция плотности распределения вероятностей для Y_k k -ой компоненты вектора \mathbf{Y} , где m_k - математическое ожидание, σ_k - среднеквадратичное отклонение. Обозначив

$f_k^0(x) = f_k(0, 1, x)$ и учитывая, что

$$f_k(m_k, \sigma_k, x) = \frac{1}{\sigma_k} f_k^0\left(\frac{x - m_k}{\sigma_k}\right), \quad \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx = 1,$$

среднюю дифференциальную энтропию одной компоненты преобразованного вектора можно записать следующим образом:

$$\begin{aligned}
H_{cp} &= -\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) \log f_k(m_k, \sigma_k, x) dx = \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k \int_{-\infty}^{+\infty} f_k(m_k, \sigma_k, x) dx - \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0\left(\frac{x - m_k}{\sigma_k}\right) \log f_k^0\left(\frac{x - m_k}{\sigma_k}\right) d\left(\frac{x}{\sigma_k}\right) = \\
&= \frac{1}{N} \sum_{k=0}^{N-1} \log \sigma_k - \frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx. \quad (7)
\end{aligned}$$

Здесь и далее опущенное в обозначениях основание логарифма означает, что выбор его конкретного значения для построения изложения непринципиален.

Чем меньше средняя энтропия (7), тем эффективнее будет последующее независимое кодирование компонент преобразованного вектора.

Поскольку в рамках корреляционной модели законы распределения компонент вектора \mathbf{X} неизвестны, то и точное определение вида функций плотности распределения $f_k^0(x)$ для компонент вектора \mathbf{Y} также невозможно. Однако компоненты вектора \mathbf{Y} представляют собой определенные взвешенные суммы из компонент вектора \mathbf{X} :

$$y_k = \sum_{j=0}^{N-1} w_{k,j} x_j,$$

где весовые коэффициенты представляют собой элементы матрицы преобразования

$$\mathbf{W} = \{w_{k,j}\}_{k,j=0}^{N-1},$$

поэтому с некоторым приближением можно допустить возможность применения центральной предельной теоремы (ЦПТ) и положить, что распределение каждой из компонент вектора \mathbf{Y} подчиняется нормальному закону (отличия имеют место только в параметрах распределения). Хотя, строго говоря, для корректного использования ЦПТ необходима прежде всего статистическая независимость входящих в сумму случайных величин, именно гауссова плотность часто используется для описания законов распределения коэффициентов унитарных преобразований, в частности, при цифровой обработке изображений [2]. Используется также распределение Лапласа. Более точное описание законов распределения основано на обобщенном нормальном законе (см., например, [3]). Допущение об общем виде законов распределения компонент вектора трансформант \mathbf{Y} для различных преобразований не является строгим, однако

часто применяется на практике, поэтому можно считать, что $\forall k=0,1,\dots,N-1$:

$$-\int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx C.$$

Используем менее жесткое ограничение, а именно:

$$-\frac{1}{N} \sum_{k=0}^{N-1} \int_{-\infty}^{+\infty} f_k^0(x) \log f_k^0(x) dx \approx C,$$

где C - некоторая константа, не зависящая от вида используемого преобразования и его размерности. Среднюю энтропию (7) одной компоненты вектора трансформант можем записать тогда в следующем виде:

$$H_{cp} = \frac{1}{2N} \sum_{k=0}^{N-1} \log \sigma_k^2 + C. \quad (8)$$

Необходимые в выражении (8) значения σ_k^2 дисперсий трансформант Y_k являются диагональными элементами матрицы $\mathbf{K}_Y = \mathbf{W} \mathbf{K}_X \mathbf{W}^T$ и в общем случае могут быть найдены по формуле:

$$\sigma_k^2 = \sum_{m=0}^{N-1} w_{k,m} \sum_{j=0}^{N-1} w_{k,j} \text{cov}(x_m, x_j). \quad (9)$$

Для оптимального преобразования Карунена-Лозва расчет по формуле (9) упрощается, т.к. матрица \mathbf{K}_Y имеет диагональный вид и

$$\prod_{k=0}^{N-1} \sigma_k^2 = \det \mathbf{K}_Y = \det \mathbf{W} \cdot \det \mathbf{K}_X \cdot \det \mathbf{W}^{-1} = \det \mathbf{K}_X.$$

Поэтому
$$\frac{1}{2N} \sum_{k=0}^{N-1} \log \sigma_k^2 = \frac{1}{2N} \log \prod_{k=0}^{N-1} \sigma_k^2 = \frac{1}{2N} \log \det \mathbf{K}_X. \quad (10)$$

Рассматривая среднюю безусловную энтропию (8) как характеристику декоррелирующих свойств ортогональных преобразований, естественно в качестве точки отсчёта принять значение (8) для оптимального преобразования Карунена-Лозва и ввести в рассмотрение следующий параметр:

$$\Delta H(\mathbf{W}, \mathbf{K}_X) = H_{cp}(\mathbf{W}, \mathbf{K}_X) - H_{cp}(\mathbf{W}_{opt}, \mathbf{K}_X),$$

или, с учетом (8), (10):

$$\Delta H(\mathbf{W}, \mathbf{K}_X) = \frac{1}{2N} \log \left(\prod_{k=0}^{N-1} \sigma_k^2 / \det \mathbf{K}_X \right). \quad (11)$$

Величина (11) характеризует избыточность данных в векторе \mathbf{Y} , связанную с неполной декорреляцией компонент (чем больше значение (11), тем меньше эффективность декоррелирующего преобразования с матрицей \mathbf{W}). Назовём данную величину *средней избыточной энтропией* [4].

Комментарии к определению (11)

1. **Физический смысл величины.** Решая задачу оптимального (в смысле минимизации суммарной квадратичной ошибки) распределения заданного количества бит для кодирования компонент случайного вектора, имеющего N -мерное нормальное распределение, Хуанг и Шульцхайс [5] установили, что при определенных условиях вносимая в результате оптимального скалярного квантования по Максуму суммарная квадратичная ошибка пропорциональна произведению дисперсий компонент, а именно:

$$M \left(\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 \right) \sim N \left(\prod_{k=0}^{N-1} \sigma_k^2 \right)^{1/N}, \quad (12)$$

где σ_k^2 - дисперсия компоненты Y_k , а \mathbf{Y} и $\hat{\mathbf{Y}}$ - оригинальный и скалярно проквантованный векторы, соответственно. В свою очередь, Дэвис и Носратиниа [6] показали, что для ортогональных преобразований при фиксированном значении размерности N величина, стоящая в правой части соотношения (12), принимает минимальное значение для оптимального преобразования Карунена-Лозва. Поэтому для данного преобразования значение средней энтропии (8) минимально (см. также (10)), и значение величины (11) для любых ортогональных преобразований неотрицательно. Таким образом, чем меньше величина средней избыточной энтропии, тем ближе свойства используемого преобразования к оптимальным. Минимум величины (11), равный нулю, достигается для оптимального преобразования Карунена-Лозва. Если в формуле (11) взять двоичное основание для логарифма, то физический смысл величины (11) определяется как избыточные (по сравнению с преобразованием Карунена-Лозва) битовые затраты, приходящиеся в среднем на одну компоненту вектора \mathbf{Y} при независимом эффективном кодировании компонент после равномерного квантования с достаточно малым шагом u .

2. **О требовании к ортогональности матрицы \mathbf{W} .** Отметим, что для анализа с использованием формулы (11) требование ортогональности исследуемого преобразования является принципиальным. Действительно, анализируя среднюю неопределенность компоненты вектора $\mathbf{Y} = \mathbf{W} \mathbf{X}$ по величине средней дифференциальной энтропии (7)-(8), мы фактически предполагали, что компоненты вектора \mathbf{Y} равномерно проквантованы с шагом квантования $u \rightarrow 0$, при этом квадратичная ошибка (3), вносимая при квантовании в каждую компоненту вектора \mathbf{Y} : $\varepsilon^2 \sim u^2/12$, тогда суммарная ошибка от квантования компонент $\mathbf{Y} \xrightarrow{\text{КВАНТОВАНИЕ}} \hat{\mathbf{Y}}$ для любого преобразования:

$$\|\mathbf{Y} - \hat{\mathbf{Y}}\|^2 \sim Nu^2/12.$$

Эта ошибка всегда будет равна ошибке, внесенной в исходный вектор \mathbf{X} , $\|\mathbf{Y} - \hat{\mathbf{Y}}\| = \|\mathbf{X} - \hat{\mathbf{X}}\|$, только для ортогональных преобразований. Если анализируемые преобразования не являются ортогональными, то говорить об эквивалентности ошибки кодирования $\|\mathbf{X} - \hat{\mathbf{X}}\|^2 \sim Nu^2/12$ для разных преобразований нельзя, и сравнение декоррелирующей эффективности преобразования по параметру (11) является некорректным.

3. **Эффективность декорреляции как степень концентрации энергии в векторе трансформант.** Из формулы (11) следует, что самый «худший» вариант используемого преобразования при фиксированной размерности N дает максимальное значение параметра

$$P = \prod_{k=0}^{N-1} \sigma_k^2. \quad (13)$$

Известно, что для всех ортогональных преобразований след ковариационной матрицы инвариантен, т.е.

$$\text{trace } \mathbf{K}_Y = \sum_{k=0}^{N-1} \sigma_k^2 = \text{trace } \mathbf{K}_X = \sum_{k=0}^{N-1} \sigma_{X_k}^2.$$

Тогда наложив ограничение

$$\sum_{k=0}^{N-1} \sigma_k^2 = const'$$

методом неопределенных множителей Лагранжа можно убедиться в том, что максимум выражения (13), а также и (11), имеет место, когда $\sigma_0^2 = \sigma_1^2 = \dots = \sigma_{N-1}^2$. Оптимальность же применения преобразования следует понимать как способность последнего к концентрации основной «энергии» (определяемой следом ковариационной матрицы) исходного дискретного сигнала (вектора \mathbf{X}) в малом количестве коэффициентов-трансформант. В этом смысле задача выбора преобразования принимает вид задачи дискретной L_2 -аппроксимации: найти такое преобразование, чтобы для заданного набора Ω отбрасываемых коэффициентов преобразования их суммарная дисперсия (фактически, энергия) $\sum_{k \in \Omega} \sigma_k^2$ была минимальной. Использование такого подхода менее универсально, вполне может дать разные результаты сравнения преобразований при различном количестве отбрасываемых коэффициентов. Тем не менее, подобное изучение свойств преобразований описано в литературе [7].

Анализ эффективности применения преобразований для дискретного марковского процесса первого порядка

Для примера приведем расчетные значения параметра (11) при использовании ряда дискретных ортогональных преобразований для обработки вектора \mathbf{X} , который представляет собой реализацию стационарного марковского процесса первого порядка и имеет ковариационную матрицу

$$\mathbf{K}_X(\rho) = \begin{pmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \dots & \dots & \dots & \dots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{pmatrix}, \quad (14)$$

$0 \leq \rho \leq 1$. Данная модель имеет широкое применение и представляет большой интерес для практики, параметр ρ есть коэффициент корреляции соседних отсчетов дискретного случайного процесса [2,7].

Результаты расчета средней избыточной энтропии приведены на рис. 1, где рассмотрены дискретное косинусное преобразование (ДКП), слэнт-преобразование (СП), дискретное преобразование Хартли (ДПХ2), а также дискретные преобразования Адамара (ДПА) и Хаара (ДПХ1) (определения преобразований см. [2,7]). Выводы, которые следуют из анализа графиков рис. 1, подтверждают известные результаты оценки декоррелирующих свойств преобразований, полученные при использовании других методов, например по Пэрл [8]. В частности, критерий (11) подтверждает тот факт, что среди дискретных преобразований, имеющих быстрые алгоритмы вычислений, ДКП дает наиболее близкие к оптимальному преобразованию характеристики декорреляции для модели данных, которые описываются ковариационной матрицей (14).

Подробное сравнение результатов анализа декоррелирующей эффективности преобразований по предложенному критерию (11) и по критерию Пэрл [8] для других видов ковариационных матриц \mathbf{K}_X и других размерностей преобразований, а также более подробное

рассмотрение вопросов численного нахождения величины (11) для ряда особых случаев можно найти в работах [4,9,10].

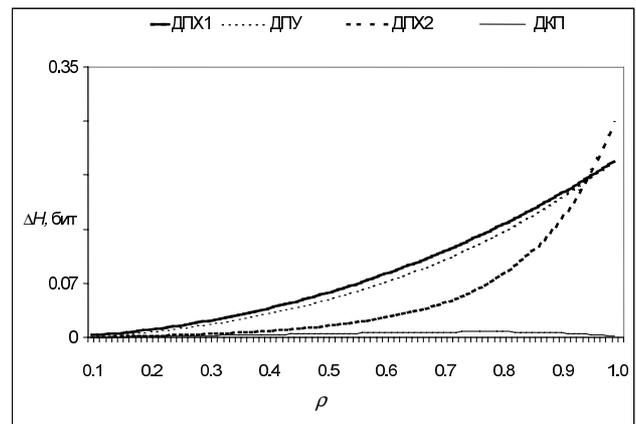


Рис. 1. Значения средней избыточной энтропии (11) при использовании некоторых ортогональных преобразований для обработки 16-ти компонентного вектора данных, имеющего ковариационную матрицу (14)

Применение предложенного критерия (11) во всех рассмотренных случаях дало оценки, которые хорошо согласуются с другими известными теоретическими и экспериментальными результатами.

Литература

1. Дмитриев В.И. Прикладная теория информации. - М.: Высш. шк., 1989. - 320 с.
2. Прэйт У. Цифровая обработка изображений: Пер. с англ. - М.: Мир, 1982. - Кн.1 и 2. - 312 и 480 с.
3. Antonini M., Barlaud M., Mathieu P., and Daubechies I., Image coding using wavelet transform// IEEE Trans. Image Proc. - Vol. 1. - №2, 1992. -P. 205-220.
4. Умняшкин С.В., Кочетков М.Е. Анализ эффективности использования дискретных ортогональных преобразований для цифрового кодирования коррелированных данных // Известия вузов. Электроника. - №6. - 1998. - С. 79-84.
5. Huang J.Y., Schultheiss P. M. Block quantization of correlated Gaussian random variables// IEEE Trans. Communications. -1963. -V. CS-11 (Sept.) -P. 289-296.
6. Davis G., Nosratinia A. Wavelet-based Image Coding: An Overview // Applied and Computational Control, Signals and Circuits. -1998. -V.1. -№1. -P. 205-269.
7. Elliott D.F., Rao K.R. Fast transforms: algorithms, analyses, applications. - London: Academic Press inc., 1982. - 488 p.
8. Pearl J. On coding and filtering stationary signals by discrete Fourier transforms // IEEE Trans. Inf. Theory. - 1973. - Vol. IT-19. - P. 229-232.
9. Кочетков М.Е., Умняшкин С.В. О сравнении критериев для оценки эффективности декоррелирующих преобразований / М.: МГИЭТ (ТУ), 1998. - 34 с. - Деп. в ВИНТИ 13.04.98, № 1069-В98.
10. Умняшкин С.В. Эффективность применения ортогональных преобразований для кодирования дискретных сигналов с точки зрения корреляционной теории // Интеллектуальные системы в производстве. - №1 - 2003. - С. 100-123.