



УДК 621.3. 01

© Ш. С. Фахми, 2010

АНАЛИТИЧЕСКАЯ МОДЕЛЬ ОЦЕНКИ СТЕПЕНИ ПРИБЛИЖЕНИЯ К ЭПСИЛОН-ЭНТРОПИИ НА ОСНОВЕ ПИРАМИДАЛЬНО- РЕКУРСИВНОГО МЕТОДА КОДИРОВАНИЯ ИЗОБРАЖЕНИЙ

Фахми Ш. С. – канд. техн. наук, доц. кафедры САПР, тел.: (812) 234-36-75, e-mail: Shakeebf@mail.ru СПбГЭТУ «ЛЭТИ»

В статье предлагается аналитическая модель для оценки степени приближения к эpsilon-энтропии источника сигналов нестационарных изображений на основе пирамидально-рекурсивного метода разбиения исходного изображения. Проведены экспериментальные исследования и определены параметры системы кодирования изображений.

The analytical model to estimate the approximation degree to epsilon-entropy of the source of non-stationary image signals is provided with the use of the pyramidal-recursive approach for the initial image partitioning. The experimental investigation has been carried out, and the parameters of the system for encoding images have been found.

Ключевые слова: аналитическая модель, энтропия источника, кодирование источника, трихотомия.

В настоящий момент, с появлением новой элементной базы в виде систем на кристалле и современных САПР на ее основе, открылись новые возможности создания специализированных систем обработки нестационарных сигналов и почти сняты многие ограничения по скорости и сложности для проектирования систем обработки видеoinформации [1]. Однако проблема не исчезла, а лишь возросло понимание ее значимости и специфических сложностей, если учесть еще и фактор наличия фотоприемника в составе вычислителя [1]. Сложившаяся ситуация вынуждает исследователей делать все больше упор на развитие новых подходов и способов в обработке изображений, чтобы найти пути преодоления разрыва между «машинным» и «человеческим» восприятием изображений [2, 3].

Известно, что для человека характерно так называемое симультанное, или синтетическое, узнавание, когда наблюдаемый объект воспринимается как единое целое (например, мгновенное узнавание лица человека) [2]. Таким

образом, узнавание идет без осознания, как условный рефлекс, следовательно, оно основано на принципах структурированного представления видеоданных [2, 3]. Этот механизм до сих пор не раскрыт и почти не изучен, следовательно, поиск оптимальных алгоритмов представления и структуризации видеоданных является актуальным и востребованным при решении задач кодирования нестационарных изображений [4, 5, 11].

В работах [6–10] предложены алгоритмы для кодирования нестационарных изображений с применением пирамидально-рекурсивного метода кодирования, в основе которого лежит принцип рекурсивного разбиения изображения с целью представления в виде структурированной совокупности семантически значимых единиц или опорных точек (ОТ) и правил их взаимосвязей (соединения).

Постановка задачи исследования. Суть пирамидально-рекурсивного метода состоит в следующем. Все изображение рассматривается как единая область (полигон) и на ней проверяется критерий однородности (наличие или отсутствие каких-либо объектов или опорных точек) (рис. 1). На этапе кодирования, если критерий выполняется, то изображение считается содержащим один лишь фон (далее обозначается как пустой полигон). В противном случае проводится разбиение этого полигона на одинаковые равные полигоны¹ и на каждом из них вновь вычисляется критерий однородности. Затем выполняется проверка, и дальнейшему разбиению подвергаются лишь те полигоны, для которых критерий не соблюдается. И так, в результате анализа всего изображения получаем полигональную двумерную сетку, содержащую как пустые полигоны, так и полигоны, содержащие ОТ. Каждая опорная точка характеризуется своими абсолютными координатами и яркостью в пределах полигона. Методы и алгоритмы разбиения, алгоритмы поиска объектов и критерии однородности в рамках данной статьи не рассматриваются, а с некоторыми из них можно познакомиться в работе [10].

Информация о местоположении ОТ состоит из описания последовательности разбиения, что определяет нахождение ОТ с точностью до покрывающего ее полигона, и координат ОТ относительно полигона. Способ описания последовательности разбиения, в частном случае, при разбиении на два, три и четыре равных прямоугольных полигона, рассмотрен в [6–10], а общий случай, при разбиении полигона на несколько равных частей, рассмотрен в [3]. При покрытии всех ОТ в соответствии с правилами разбиения могут образоваться пустые полигоны, не покрывающие ни одной ОТ.

Общий подход к кодированию по опорным точкам. Каждая опорная точка характеризуется своими абсолютными координатами и яркостью. Процесс покрытия полигонами представлен в следующем виде: ОТ последова-

¹ В экспериментах исследовано разбиение априорной области на различное число частей: $d = 2, \dots, 6$, т. е. дихотомия (деление на две части), трихотомия (деление на три части), тетрахотомия (деление на четыре части) и т. д.



тельно, в произвольном порядке наносятся на поле изображения, при этом, если в соответствующем полигоне окажутся две ОТ, то полигон разбивается на равные части до тех пор, пока каждая ОТ не окажется единственной в своем полигоне.

Введем следующие обозначения: n – число ОТ, равное числу заполненных (покрывающих ОТ) полигонов (рис. 1, а), m – максимальное число разбиений, приводящее к размеру полигона в один пиксел изображения, d – число полигонов, полученных при разбиении исходного полигона, N – общее число пикселей исходного изображения; очевидно, что $N = d^m$, B – число бит информации, необходимое для кодирования только местоположения ОТ, v – разрядность яркости ОТ.

Требуется оценить функцию $B = F(n, d, m)$ и найти оптимальное число полигонов после разбиения d , минимизирующее функцию B (т. е. $F \rightarrow \min$). В нормированных единицах искомое отношение запишется в виде

$$b = f(n_{\text{нор}}, d, m), \text{ где } b = B/N, n_{\text{нор}} = n/N.$$

Учитывая то обстоятельство, что все ОТ появляются независимо друг от друга и с одинаковой вероятностью, то число бит информации для кодирования местоположений ОТ в общем виде будет $I_{\text{max}} = n \cdot \log N$.

Коэффициент сжатия определяется выражением

$$K = \frac{(n \cdot v + n \cdot \log N)}{(n \cdot v + B)}, \quad (1)$$

т. е.

$$k = \frac{(n_{\text{нор}} \cdot v + n_{\text{нор}} \cdot \log N)}{(n_{\text{нор}} \cdot v + b)},$$

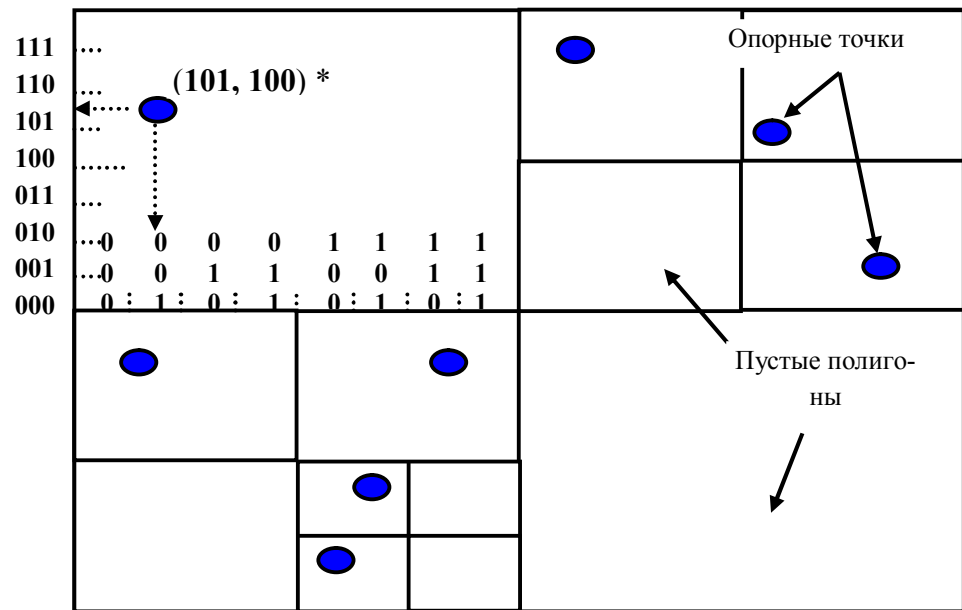
где $\log N$ – разрядность абсолютных координат ОТ, $b = B/N$, $n_{\text{нор}} = n/N$.

Способ сжатия изображений по ОТ. Для решения поставленной задачи рассмотрим алгоритм формирования сжатого описания изображения (рис. 1). После каждого разбиения в описание заносится один из следующих признаков:

Пр₁ – признак того, что полигон подлежит следующему дроблению, т. е. покрывает более одной ОТ.

Пр₂ – признак того, что полигон далее не подлежит дроблению и покрывает одну ОТ. В этом случае в сжатое описание дополнительно заносятся координаты ОТ в пределах полигона.

Пр₃ – признак того, что полигон не содержит ни одной ОТ и, следовательно, не подлежит последующему дроблению. Таким образом, сжатое описание B состоит из двух составляющих: $B = B_1 + B_2$, где B_1 – информация, необходимая для кодирования и передачи признаков Пр₁ и Пр₃; B_2 – информация, необходимая для кодирования и передачи ОТ, сопутствующая Пр₂.



а – исходное изображение

| № | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Сжатое описание | Пр ₁ | Пр ₂ | Пр ₁ | Пр ₁ | Пр ₃ | Пр ₂ | Пр ₂ | Пр ₃ | Пр ₂ |
| Код | 11 | 0* | 11 | 11 | 10 | 0* | 0* | 10 | 0* |
| 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | | |
| Пр ₂ | Пр ₂ | Пр ₃ | Пр ₁ | Пр ₂ | Пр ₂ | Пр ₂ | Пр ₃ | | |
| 0* | 0* | 10 | 11 | 0* | 0* | 0* | 10 | | |

б – сжатое описание, где * – информация о координатах опорных точек в пределах полигона

Рис. 1. Формирование сжатого описания

Видно (рис. 1, б), что общее число признаков x вычисляется по формуле

$$x = \frac{(d(n + \Delta n) - 1)}{(d - 1)},$$

где Δn – число пустых полигонов, т. е. количество Пр₃ в сжатом описании. Число Пр₂ равно числу ОТ n , число Пр₁: $x_1 = x - n - \Delta n$.

Для кодирования признаков используем коды Хаффмана, наиболее часто встречающийся признак кодируем одним битом, а остальные – двумя битами (рис. 1, б). Сравнивая величины x , n , и Δn , получим:



при $d = 2$: $x_1 > n$, Δn ; а при $d > 2$: либо $n > x_1$, Δn , если $E \leq 1$, либо $\Delta n > x_1$, n , если $E > 1$, где $E = \Delta n/n$. Следовательно, можно легко вывести и показать, что

$$B_1 = \begin{cases} x_1 + 2(n + \Delta n), & \text{при } d = 2; \\ n + 2(x_1 + \Delta n), & \text{при } d > 2 \text{ и } E \leq 1; \\ \Delta n + 2(x_1 + \Delta n), & \text{при } d > 2 \text{ и } E > 1. \end{cases}$$

После преобразования получим [6, 9]:

$$B_1 = \begin{cases} 3n_{\text{нор}} \cdot N \cdot (E + 1), & \text{при } d = 2; \\ \frac{n_{\text{нор}} \cdot N \cdot (2dE + d + 1)}{(d - 1)}, & \text{при } d > 2 \text{ и } E \leq 1; \\ \frac{n_{\text{нор}} \cdot N \cdot (2dE + (d + 1))}{(d - 1)}, & \text{при } d > 2 \text{ и } E > 1. \end{cases} \quad (2)$$

Для кодирования местоположения ОТ внутри полигона необходимо число бит, зависящее от размера полигона, т. е. уровня разбиения ℓ . Исходное изображение назовем полигоном с уровнем $\ell = 0$. После первого разбиения имеем полигоны уровня $\ell = 1$ и т. д. Пиксель будет являться неотделимым полигоном уровня m . Тогда

$$B_2 = \sum_{\ell=1}^m r_{\ell} \cdot Z_{\ell}, \quad (3)$$

где Z_{ℓ} – число заполненных (т. е. покрывающих ОТ) полигонов уровня ℓ ; r_{ℓ} – необходимая разрядность для кодирования координат ОТ.

Полигон уровня ℓ содержит N/d^{ℓ} пикселей, тогда

$$r_{\ell} = \log \frac{N}{d^{\ell}} \quad (4)$$

Подставим выражение (4) в соотношение (3) и, учитывая, что

$$\sum_{\ell=1}^m Z_{\ell} = n \quad (5)$$

получим

$$B_2 = N \cdot \log d \left(n_{\text{нор}} m - \left(\frac{1}{d^m} \right) \right) \sum_{\ell=1}^m \ell \cdot Z_{\ell}. \quad (6)$$

Окончательное нормированное выражение для определения количества информации будет:

$$B = \log d \cdot \left(n_{\text{нор}} m - \left(\frac{1}{d^m} \right) \right) \sum_{\ell=1}^m \ell \cdot Z_{\ell} \begin{cases} 3n_{\text{нор}} \cdot E \cdot (E + 1), & \text{при } d = 2; \\ \frac{n_{\text{нор}} (2dE + d + 1)}{(d - 1)}, & \text{при } d > 2. \end{cases} \quad (7)$$

Таким образом, данный способ сжатия позволяет по выражению (7) вычислить число бит информации, необходимое для определения местоположе-

ния ОТ. Однако для вычисления по выражениям (7) и (1) нужно определить следующее:

$E = f(n_{\text{нор}}, d, m)$ – долю пустых полигонов по отношению к числу ОТ;

$Z_{\ell} = f(n_{\text{нор}}, d, m, \ell)$ – число заполненных полигонов различных уровней ℓ .

Рекуррентные соотношения для вычисления числа пустых и заполненных полигонов. Для нахождения E и Z_{ℓ} рассмотрим следующую модель. Изображение представим полем площадью 1; ОТ последовательно наносятся на поле с равной вероятностью появления в любой точке, что приводит к дроблению соответствующих полигонов, при котором появляются как заполненные, так и пустые полигоны.

Обозначим: r_{ji} – суммарную площадь пустых полигонов i -го уровня после появления j -й точки; S_{ji} – суммарную площадь заполненных полигонов уровня i после появления точки j ; p_{ji} – вероятность появления ОТ $j+1$ в пустом полигоне уровня i ; q_{ji} – вероятность появления ОТ $j+1$ в заполненных полигонах уровня i ; Z_{ji} – число заполненных полигонов уровня i после появления точки j ; Y_{ji} – число пустых полигонов уровня i после появления точки j .

Площадь полигона уровня i равна $1/d$. С учетом этого напишем ряд очевидных соотношений [2, 3, 6, 9]:

$$\sum_{\ell=1}^m (r_{j\ell} + s_{j\ell}) = 1, \sum_{\ell=1}^m (p_{j\ell} + q_{j\ell}) = 1, \sum_{\ell=1}^m Z_{j\ell} = \sum_{\ell=1}^m S_{j\ell} \cdot d_{j\ell} = j \text{ и } \sum_{\ell=1}^m (r_{j\ell} + s_{j\ell}) = 1, \quad (8)$$

$Z_{j\ell} = S_{j\ell} \cdot d^{\ell}$, $Y_{j\ell} = r_{j\ell} \cdot d^{\ell}$, где $j = 1 \dots n$; $\ell = 1 \dots m$;

$r_{j\ell} = (d-1)/d$, $S_{11} = 1/d$, $p_{11} = (d-1)/d$, $q_{11} = 1/d$ и площадь полигона уровня i равна $1/d$.

Искомые величины Z_{ℓ} и E для вычисления по выражению (7) будут:

$$E(n_{\text{нор}}, d, m) = \frac{\sum_{i=1}^m Y_{ji}}{j} \text{ при } j = n,$$

$$Z_{\ell}(n_{\text{нор}}, d, m, \ell) = Z_{n,\ell}, \quad (9)$$

где $\ell = 1 \dots m$.

После появления каждой ОТ величины p_{ji} и q_{ji} вычисляются по площадям r_{ji} , S_{ji} :

$$p_{ji} = \left(\frac{r_{ji}}{\left(1 - j \left(\frac{1}{d^m}\right)\right)} \right);$$

$$q_{ji} = \left(\frac{S_{ji} - z_{ji} \left(\frac{1}{d^m}\right)}{\left(1 - j \left(\frac{1}{d^m}\right)\right)} \right); \quad (10)$$



В знаменателе выражений и числителе второго выражения учитываются те площади, где невозможны повторные появления ОТ из-за конечного числа пикселей на изображении.

Площади r_{ji} и S_{ji} вычисляются как математическое ожидание случайных величин, принимающих значение различных площадей с соответствующими вероятностями в зависимости от появления ОТ в заполненных или пустых полигонах для всех уровней.

Рассмотрим вычисление площади $r_{j+1, i}$. В зависимости от координат очередной ОТ возможны следующие ситуации (рис. 2):

1. Уменьшение площади: $r_{j+1, i} = r_{ji} - \frac{1}{d^i}$. Это соответствует появлению $j+1$

ОТ в пустом полигоне i -го уровня; вероятность этого события – p_{ji} (цифра 1 на рис. 2).

2. Увеличение площади: $r_{j+1, i} = r_{ji} + \frac{d-2}{d^i}$. Это соответствует появлению

ОТ в заполненном полигоне более низкого уровня при условии, что ОТ вызвала цепочку дроблений до уровня ℓ , т. е. появилось d новых полигонов, из них $d-2$ – пустых и два заполненных: предшествующий и текущий $j+1$ -й ОТ (цифра 2 на рис. 2). Пусть ОТ появилась в заполненном полигоне уровня $i-1$; вероятность этого события $q_{j, i-1}$. Тогда возможны две ситуации:

– с вероятностью $1/d$ потребуется более одного разбиения (цифра 2' на рис. 2);

– с вероятностью $1-1/d$ потребуется одно разбиение (цифра 2 на рис. 2).

Следовательно, интересующая нас вероятность при условии появления ОТ в полигоне уровня $i-1$ составит $q_{j, i-1} = 1 - \frac{1}{d}$. Если ОТ появилась в запол-

ненном полигоне уровня $i-2$ (вероятность – $q_{j, i-2}$), то с вероятностью $\left(\frac{1}{d}\right)\left(\frac{1}{d}\right)$

потребуется более двух разбиений, а с вероятностью $\left(\frac{1}{d}\right)\left(1 - \frac{1}{d}\right)$ – два раз-

биения, т. е. до уровня i . Следовательно, интересующая нас вероятность составит $q_{j, i-1} \cdot \frac{1}{d} \cdot \left(1 - \frac{1}{d}\right)$. В итоге, суммируя вероятности по всем уровням, по-

лучим искомую вероятность $\sum_{\ell=1}^{i-1} q_{j, i-\ell} \frac{1}{d^{\ell-1}} \left(1 - \frac{1}{d}\right)$, следует иметь в виду, что

найденная вероятность справедлива при условии, что $i = m$. Если $i = m$, то дальнейшие разбиения невозможны. При появлении ОТ в полигоне $(i-1)$ -го уровня (вероятность – q_{ji}) произойдет только одно разбиение. При появлении ОТ в полигоне $(i-2)$ -го уровня с вероятностью $1/d$ произойдет два разбиения,

т. е. до уровня i . В итоге при $i = m$ искомая вероятность составит $\sum_{\ell=1}^{i-1} q_{j, i-\ell} \frac{1}{d^{\ell-1}}$.

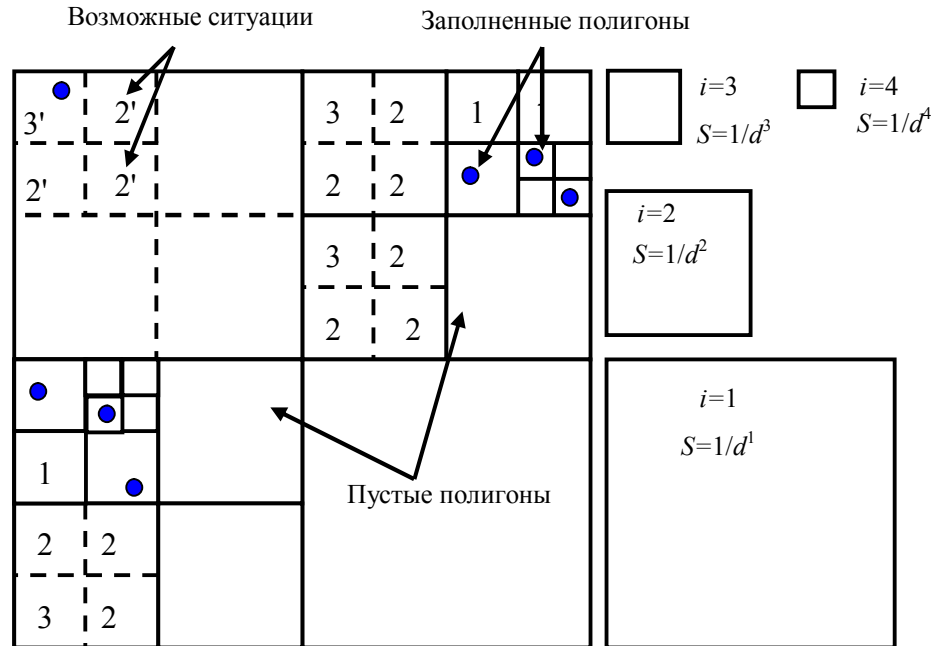


Рис. 2. Возможные варианты появления опорных точек

3. Увеличение площади: $r_{j+1,i} = r_{ji} + \frac{d-1}{d}$. Этот случай аналогичен второму, но разница состоит в том, что цепочка разбиений не заканчивается на уровне i . Очевидно, что этот случай исключается при $i = m$ (цифра 3 на рис. 3). Проводя аналогичные рассуждения, получим вероятность этого случая:

$$\sum_{\ell=1}^{i-1} q_{j,i-j} \frac{1}{d^{\ell-1}} \frac{1}{d}$$

4. Площадь сохранится: $r_{j+1,i} = r_{ji}$. Эта ситуация возникает во всех остальных случаях, ее вероятность составит $1 - p_{ji} - \sum_{\ell=1}^{\ell-1} q_{j,i-\ell} \frac{1}{d^{\ell-1}}$ при $\ell \leq m$.

Суммируя произведения площадей на вероятности соответствующих рассмотренных выше четырех возможных исходов, после элементарных преобразований получим окончательное выражение:

По аналогии проводим вычисление площади $S_{j+1,i}$. С учетом выражений (8), (10, 11) получим следующие рекуррентные выражения [6, 9]:



$$\begin{aligned}
 Y_{j+1,1} &= Y_{j,1} + \frac{\left\{ \begin{array}{l} (d-1)^2, \text{ если } \ell < m \\ d(d-2), \text{ если } \ell = m \end{array} \right\} \cdot \sum_{\ell=1}^{m-1} (1-d)^{\ell-m} Z_{j1} - Y_{j1}}{d^\ell - jd^{\ell-m}} \\
 Z_{j+1,1} &= Z_{j,1} + \frac{\left\{ \begin{array}{l} 2(d-1), \text{ если } \ell < m \\ 2d, \text{ если } \ell = m \end{array} \right\} \cdot \sum_{\ell=1}^{m-1} (1-d)^{\ell-m} Z_{j1} - Z_{j1}(1-d^{\ell-m})}{d^\ell - jd^{\ell-m}} \quad (12)
 \end{aligned}$$

Оценка степени приближения к эpsilon-энтропии. Рассмотренная аналитическая модель позволяет вычислить значение количества информации b (см. постановку задачи) по следующей схеме:

1. По выражению (12) вычисляется число заполненных и пустых полигонов (y_{ji}, z_{ji}) , при заданных ℓ, m, d, i и j .
2. По выражению (9) вычисляются E и Z_ℓ , где $\ell = 1, \dots, m$.
3. По выражению (7) вычисляется b .

Для оценки эффективности предложенного способа кодирования информации о расположении ОТ будем сравнивать результаты моделирования с теоретическим минимумом (рис. 3).

На поле изображения, содержащем N пикселей, n ОТ могут распределиться

$$C_N^n = \frac{N!}{n!(N-n)!} \text{ способами. С учетом равновероятности распределения ОТ}$$

все комбинации независимы, следовательно, количество информации, необходимое для кодирования любой комбинации, составит

$$b_{\min} = \log C_N^n = b_{\min} = -n_{\text{нор}} \log n_{\text{нор}} - (1 - n_{\text{нор}}) \log(1 - n_{\text{нор}}). \quad (13)$$

Выражение (13) позволяет вычислить эpsilon-энтропию источника b_{\min} .

Заключение. На рис. 3 приведены зависимости $E(n/N)$ при различных d, m . Анализ результатов позволяет утверждать, что E слабо зависит от m и при $N = d^m > 10^3$, этой зависимостью можно пренебречь (ошибка на уровне десятых долей процента). Из графиков следует, что $d = 3$ или 4 являются оптимальными.

Определено, что оптимальным числом полигонов после разбиения при проектировании систем кодирования на основе ПРП с равновероятным распределением ОТ на исходных изображениях является $d = 3$. Показано, что эффективным оказывается ПРП изображений с равномерным распределением ОТ, когда количество ОТ не превышает 30 % от общего количества точек исходного изображения.

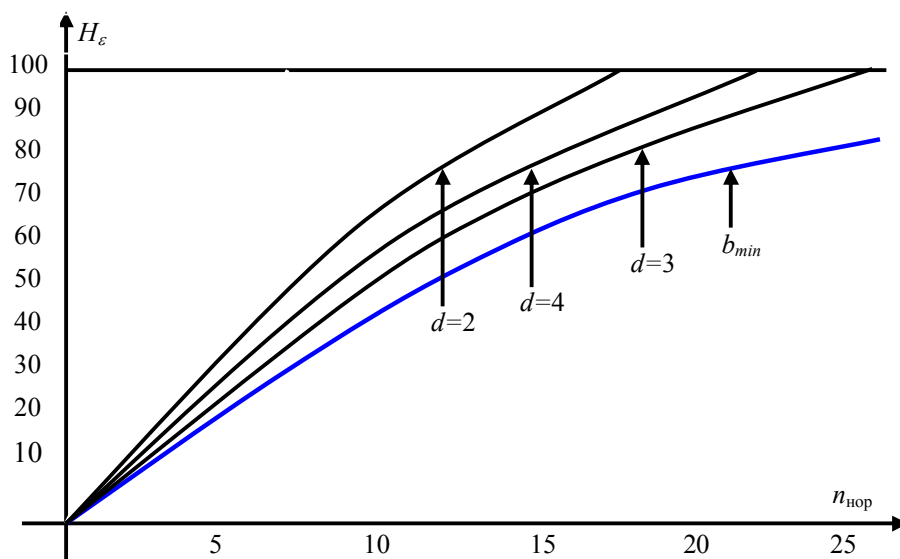


Рис. 3. Степени приближения к эpsilon-энтропии при различных значениях d

Библиографические ссылки

1. *Твердотельная революция в телевидении* / В. В. Березин, А. А. Умбиталиев, Ш. С. Фахми, А. К. Цыцулин, Н. Н. Шипилов. М., 2007.
2. *Александров В. В., Горский И. Д.* Представление и обработка изображений. Рекурсивный подход. Л., 1985.
3. *Александров В. В., Горский И. Д.* Алгоритмы и программы структурного метода обработки данных. Л., 1983.
4. *Эксперименты с машинным зрением* / В. П. Андреев, Д. А. Белов и др.; под ред. И. А. Овсевича. М., 1987.
5. *Фу К.* Структурные методы в распознавании образов. М., 1980.
6. *Фахми Ш. С.* Оценка степени сжатия изображений с использованием рекурсивного метода разбиения. СПб., 1993.
7. *Фахми Ш. С.* Аналитическая модель оценки эффективности кодирования по опорным точкам. СПб., 2006.
8. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Алгоритм сжатия и восстановления изображений с пирамидально-рекурсивной структурой данных. Деп. в ВИНТИ 11.02.93, № 350-В93.
9. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Аналитическая модель для оценки степени сжатия информации по опорным точкам // Методы и аппаратно-программные средства цифровой обработки сигналов. СПб., 1993. № 1.
10. *Фахми Ш. С., Шах В. В., Шмидт В. К.* Алгоритм сжатия и восстановления методом деления на треугольники. СПб., 1993.
11. *Бахрушин А. П.* Спектральный анализ видеок кадров на основе системы импульсных функций с целью синхронизации процессов внедрения и поиска цифровых водяных знаков // Вестник Тихоокеанского государственного университета. 2008. № 4(11).