

ЭНТРОПИЯ КОНТУРОВ МЕСТНОСТИ, ИЗОБРАЖЕННЫХ НА ТОПОГРАФИЧЕСКИХ КАРТАХ

Рассмотрим информационную модель контура местности с точки зрения теории информации К. Э. Шеннона [6]. Если криволинейный контур аппроксимирован ломаной линией, то в упорядоченном множестве характерных точек контура $u_0, u_1, \dots, u_{i-1}, u_i, u_{i+1}, \dots, u_n$ энтропию положения точки u_{i+1} относительно предшествующих точек u_{i-1}, u_i будем рассматривать как условную энтропию и обозначать ее $H(u_{i+1}|u_{i-1}, u_i)$.

Положение точки u_{i+1} относительно точек u_{i-1}, u_i однозначно определяется углом поворота Θ и длиной элементарного отрезка l . Последовательности случайных величин

$$l_1, l_2, \dots, l_n;$$

$$\Theta_1, \Theta_2, \dots, \Theta_n$$

независимы, причем длина элементарных отрезков имеет гамма-распределение, а углы поворота распределены по нормальному закону [2]. Вот почему можно принять

$$H(u_{i+1}|u_{i-1}, u_i) = H(l, \Theta) = H(l) + H(\Theta). \quad (1)$$

Энтропия элементарного отрезка $H(l)$ есть не что иное, как математическое ожидание функции $-\log [f(l)\Delta l]$, где $f(l)$ — плотность вероятностей; Δl — шаг квантования l ; \log — символ двоичного логарифма [1]. Поскольку

$$f(l) = \frac{\lambda^\alpha}{\Gamma(\alpha)} l^{\alpha-1} e^{-\lambda l},$$

где λ и α — параметры распределения, $\Gamma(\alpha)$ — гамма-функция, можем записать

$$H(l) = M \left\{ -\log \left[\frac{\lambda^\alpha}{\Gamma(\alpha)} l^{\alpha-1} e^{-\lambda l} \Delta l \right] \right\} =$$

$$= M \left[\alpha \log \frac{1}{\lambda} + \log \Gamma(\alpha) - (\alpha - 1) \log l + \lambda l \log e - \log \Delta l \right] =$$

$$= \alpha \log \frac{1}{\lambda} + \log \Gamma(\alpha) - (\alpha - 1) M(\log l) + \lambda \log e M(l) - \log \Delta l. \quad (2)$$

В выражении (2) M обозначает математическое ожидание. Для случайной величины l , имеющей гамма-распределение, математическое ожидание [4]

$$M(l) = \frac{\alpha}{\lambda} = \bar{l} = \frac{1}{\Lambda}, \quad (3)$$

где \bar{l} — средняя длина элементарного отрезка; $\bar{\Lambda}$ — средняя плотность характерных точек генерализованного изображения контура.

Поскольку график функции $f = \log l$ — выпуклая вверх кривая, справедливо неравенство Йенсена [5], $M(\log l) \leq \log [M(l)]$, т. е., учитывая равенство (3), принимаем

$$M(\log l) = \log \frac{\alpha}{\lambda} - \log t,$$

где t — некоторая переменная. Проведенные исследования показали, что для случайных величин, имеющих гамма-распределение, переменная t зависит от параметра α . Если $1 \leq \alpha \leq 6,6$, эту зависимость можно выразить эмпирической формулой

$$t = \frac{\alpha + 0,43}{\alpha - 0,21}.$$

Следовательно,

$$M(\log l) = \log \left[\frac{\alpha(\alpha - 0,21)}{\lambda(\alpha + 0,43)} \right]. \quad (4)$$

Подставляя значения $M(l)$ и $M(\log l)$ из равенств (3) и (4) в выражение (2), получаем

$$H(l) = \log \left[\Gamma(\alpha) \left(\frac{e}{\alpha} \right)^\alpha \left(\frac{\alpha + 0,43}{\alpha - 0,21} \right)^{\alpha-1} \frac{\bar{l}}{\Delta l} \right]. \quad (5)$$

Соответственно энтропию углов поворота найдем из уравнения

$$H(\Theta) = M\{-\log [f(\Theta) \Delta \Theta_i]\} = -M[\log f(\Theta)] - M(\log \Delta \Theta_i). \quad (6)$$

Как известно [1, 6], при нормальном распределении вероятностей

$$-M[\log f(\Theta)] = \log(\sqrt{2\pi}e\sigma), \quad (7)$$

где σ — среднее квадратическое отклонение углов поворота.

Переменный шаг квантования $\Delta \Theta_i$ выберем под условием

$$\Delta \Theta_i = \frac{\Delta l}{l_i}.$$

Принимая во внимание уравнения (3) и (4), будем иметь

$$-M(\log \Delta \Theta_i) = \log \left[\frac{\bar{l}}{\Delta l} \left(\frac{\alpha - 0,21}{\alpha + 0,43} \right) \right]. \quad (8)$$

На основании выражений (6), (7) и (8) можем записать

$$H(\Theta) = \log \left[\sqrt{2\pi}e \left(\frac{\alpha - 0,21}{\alpha + 0,43} \right) \frac{\sigma \bar{l}}{\Delta l} \right]. \quad (9)$$

И, наконец, подставляя значение $H(l)$ и $H(\Theta)$ из выражений (5) и (9) в уравнение (1), найдем условную энтропию положения характерной точки контура

$$H(u_{i+1}/u_{i-1}u_i) = \log \left\{ \sqrt{2\pi e} \Psi(\alpha) \sigma \left(\frac{\bar{l}}{\Delta l} \right)^2 \right\} \text{ бит/точку,} \quad (10)$$

где

$$\Psi(\alpha) = \Gamma(\alpha) \left(\frac{e}{\alpha} \right)^\alpha \left(\frac{\alpha + 0,43}{\alpha - 0,21} \right)^{\alpha-2} \quad (11)$$

Значения функции $\Psi(\alpha)$ для α от 1 до 6,5

α	0	2	4	6	8
1	1,50	1,64	1,73	1,80	1,83
2	1,84	1,85	1,85	1,84	1,84
3	1,83	1,82	1,80	1,79	1,77
4	1,75	1,73	1,71	1,69	1,67
5	1,66	1,64	1,62	1,60	1,59
6	1,58	1,56	1,55	1,54	—

функция, численные значения которой приведены в таблице.

Попробуем теперь определить энтропию контура конечной длины L . Энтропия последовательности характерных точек u_0, u_1, \dots, u_n обладает свойством иерархической аддитивности [5]. При фиксированной начальной точке u_0 ее можно представить в виде суммы

$$H(L) = H(u_1, u_2, \dots, u_n/u_0) = H(u_1/u_0) + H(u_2/u_0, u_1) + \dots + H(u_n/u_0, u_1, u_2, \dots, u_{n-1}). \quad (12)$$

Возможные значения азимута линии, соединяющей точки u_0 и u_1 , будем предполагать равновероятными в интервале $0-2\pi$. Следовательно, энтропия данного направления с учетом выражения (8) будет

$$H(A) = M \left[-\log \left(\frac{1}{2\pi} \Delta\Theta_i \right) \right] = \log \left(2\pi \frac{\alpha - 0,21}{\alpha + 0,43} \frac{\bar{l}}{\Delta l} \right).$$

Полагая, что длина отрезка $[u_0u_1]$ не зависит от азимута линии, и принимая во внимание выражения (5) и (11), получаем

$$H(u_1/u_0) = H(l) + H(A) = \log \left[2\pi \Psi(\alpha) \left(\frac{\bar{l}}{\Delta l} \right)^2 \right] \text{ бит/точку.} \quad (13)$$

Для генерализованного изображения контура $u_0, u_1, u_2, \dots, u_n$ — последовательность без последствия или с ограниченным последствием [1, 2]. Поэтому при $i \geq 2$ можно принять

$$H(u_{i+1}/u_0, u_1, \dots, u_{i-1}, u_i) = H(u_{i+1}/u_{i-1}, u_i). \quad (14)$$

Среднее число характерных точек контура длиной L будет $n = L\bar{\lambda}$.

Из выражений (10), (12) — (14) следует, что при достаточно большом n энтропия изображения контура, выражаемая в битах, неограниченно приближается к

$$H(L) = L\bar{\lambda} \log \left\{ \sqrt{2\pi e} \Psi(\alpha) \sigma \left(\frac{\bar{l}}{\Delta l} \right)^2 \right\}. \quad (15)$$

Соответственно удельную энтропию, выражаемую в битах на километр, в расчете на единицу длины контура запишем в виде

$$H(\bar{\Lambda}) = \frac{H(L)}{L} = \bar{\Lambda} \log \left\{ \sqrt{2\pi e} \Psi(\alpha) \sigma \left(\frac{\bar{L}}{\Delta l} \right)^2 \right\}. \quad (16)$$

Формулы (15) и (16) — среднее количество информации, необходимой для того, чтобы воспроизвести положение, форму, размеры, а для замкнутых контуров — также и площадь изображаемого контура.

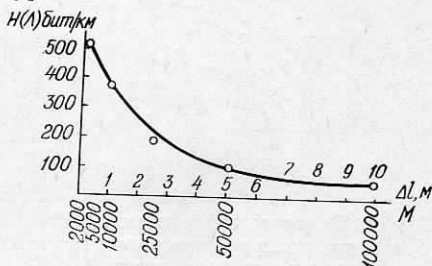


Рис. 1. Зависимость удельной энтропии контура $H(\bar{\Lambda})$ от масштаба топографической карты.

Значение энтропии в выражениях (10), (15) и (16) существенно зависит от шага квантования Δl , выбираемого в некоторых пределах произвольно. Однако для контурной нагрузки топографических карт неопределенность выбора Δl устраняется, если принять, что она равна точности масштаба соответствующей карты или плана, т. е. $\Delta l = 10^{-4} M$, где M — знаменатель численного масштаба.

По мере перехода к картам более мелких масштабов удельная энтропия $H(\bar{\Lambda})$ уменьшается как в результате увеличения Δl , так и сокращения числа характерных точек вследствие генерализации изображения. Аналогично тому, как это было сделано Ю. К. Неумывакиным [3], проследим зависимость величины $H(\bar{\Lambda})$ от масштаба карты или плана. На основании опытных данных по формуле (16) вычислена удельная энтропия для группы контуров (грунтовые и шоссейные дороги, границы лесных массивов, ручьи, реки и озеро), изображенных на топографических картах в масштабах 1 : 2000 (три контура), 1 : 10 000 (восемь контуров), 1 : 25 000 (четыре контура), 1 : 50 000 (восемь контуров) и 1 : 100 000 (три контура). Средние значения $H(\bar{\Lambda})$ показаны кружками на рис. 1. Зависимость удельной энтропии изображения контура от масштаба карты в интервале масштабного ряда 1 : 2000—1 : 100 000 хорошо аппроксимируется кривой (рис. 1), уравнение которой

$$H_i(\Delta) = \frac{A}{1 + \beta (10^{-3} M_i)^T} \frac{\text{бит}}{\text{км}}, \quad (17)$$

где A, β, γ — параметры, значение которых обусловлено характером местности. В нашем эксперименте получены следующие значения $A=540$ бит/км, $\beta=0,0272$, $\gamma=1,27$.

Как и следовало ожидать, с увеличением знаменателя масштаба изображение теряет часть информации. Относительной мерой потерь информации, а потому и мерой обобщения контура, может служить относительная избыточность изобра-

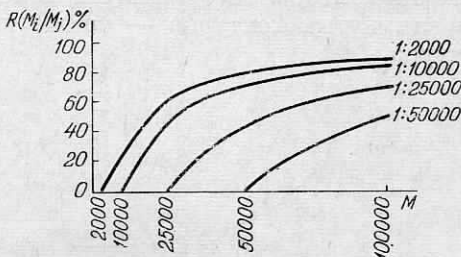


Рис. 2. Относительная избыточность изображения контура для топографических карт различных масштабов.

жения в масштабе $1/M_i$ по отношению к более мелкому масштабу $1/M_j$

$$R(M_i/M_j) = 1 - \frac{H_j(\bar{\Delta})}{H_i(\bar{\Delta})}.$$

Выразив при помощи формулы (17) $H_i(\bar{\Delta})$, $H_j(\bar{\Delta})$ через M_i , M_j и имея в виду, что $M_j > M_i$, получим

$$R(M_i/M_j) = \beta \frac{M_j^\gamma - M_i^\gamma}{10^{3\gamma} + \beta M_j^\gamma}. \quad (18)$$

Смысл величины $R(M_i/M_j)$ состоит в том, что если в масштабе $1/M_i$ изображение контура длиной L содержит, учитывая погрешности измерений, не более $H(L)$ единиц информации, то изображение того же контура в более мелком масштабе $1/M_j$ будет заключать в себе не более $[1-R(M_i/M_j)]H(L)$ тех же единиц. Таким образом, при переходе к более мелкому масштабу $1/M_j$ $100R(M_i/M_j)$ % исходной информации T_i (рис. 2) из-за сужения изобразительных возможностей оказывается излишней и теряется.

Список литературы: 1. Вентцель Е. С. Теория вероятностей. М., Физматгиз, 1962. 2. Войславский Л. К. Вероятностная модель изображения контуров местности на топографических картах. — Геодезия, картография и аэрофото съемка, 1979, вып. 29. 3. Неумывакин Ю. К. Обоснование точности топографических съемок для проектирования. М., Недра, 1976. 4. Румишский Л. З. Элементы теории вероятностей. М., Наука, 1970. 5. Стратонович Р. Л. Теория информации. М., Советское радио, 1975. 6. Шеннон К. Э. Работы по теории информации и кибернетике. М., Иностранная литература, 1963.

Работа поступила в редколлегию 22 ноября 1977 года. Рекомендована кафедрой геодезии Харьковского института инженеров коммунального строительства.