

М. А. СТАРКОВ
 (Новосибирск)

СТАТИСТИЧЕСКАЯ МОДЕЛЬ БИНАРНЫХ ИЗОБРАЖЕНИЙ

При решении задач кодирования изображений, распознавания образов, обнаружения объектов в шумах, а также при построении оптимальных алгоритмов большую роль играет способ математического представления изображений. Здесь успех определяется тем, насколько близко математическая модель описывает реальный физический объект. Ниже предложен способ описания бинарных изображений, т. е. изображений, представленных в виде нулей и единиц на равномерной квадратной сетке. Эффективность модели показана на примере кодирования.

Рассмотрим множество бинарных изображений $\{A^i\}$, которые могут быть представлены на матрице размерностью $m \times n$. Очевидно, что их число равно 2^{mn} . Если для каждого изображения A^i определить вероятность его реализации P_i , то среднее количество битов, необходимое для кодирования, приблизительно будет равно энтропии

$$H = - \sum_{i=1}^{2^{mn}} P_i \log_2 P_i. \quad (1)$$

Вряд ли возможно реализовать эту идею именно в такой постановке [1], поэтому несколько ее изменим. Пусть на некотором множестве точек (или элементов матрицы) заданы значения изображения. Будем называть такое множество опорным множеством точек Ω . Тогда, как показывают исследования [2], вероятность реализации единицы в окрестности единицы опорного множества будет выше, чем в окрестности нуля. Теперь попытаемся вычислить вероятности реализации единицы $p_{j,k}$ на множестве точек, дополняющем Ω до всей матрицы или бесконечной сетки. Будем искать $p_{j,k}$ на $\bar{\Omega}$ в виде решения линейного разностного уравнения второго порядка. Уравнения этого типа при своей простоте дают достаточно широкий набор решений. Естественно, что такое предположение налагает ограничения на класс исследуемых изображений, которые выясняются в ходе дальнейших рассуждений.

Заметим, что разностное уравнение второго порядка дает линейную комбинацию значений $p_{r,s}$ в (j, k) -й точке и ее окрестности, причем под окрестностью (j, k) -й точки будем понимать 4 ближайшие к ней точки. Итак,

$$a_1 p_{j,k} + a_2 p_{j-1,k} + a_3 p_{j+1,k} + a_4 p_{j,k-1} + a_5 p_{j,k+1} + a_6 = 0, \quad (2)$$

где из шести неизвестных коэффициентов пять линейно-независимые. Ограничимся рассмотрением статистически изотропных изображений;

тогда точки окрестности (j, k) -й точки должны войти с равным весом, т. е. $a_2 = a_3 = a_4 = a_5$, и уравнение (1) можно записать в виде

$$\Delta p_{j, k} - \lambda(p_{j, k} - h) = 0, \quad (3)$$

где

$$\Delta p_{j, k} = p_{j-1, k} + p_{j+1, k} + p_{j, k-1} + p_{j, k+1} - 4p_{j, k} \quad (4)$$

есть оператор Лапласа на сетке.

Выясним физический смысл коэффициентов λ и h . Пусть множество Ω пусто, тогда единственным решением уравнения (3) на сетке будет $p_{j, k} = h$. Таким образом, h есть априорная вероятность получения единицы в произвольной точке (при этом все точки сетки равноправны).

Для выяснения смысла коэффициента λ решим следующую задачу. Пусть при $j=0$ $p_{0, k} = 1$. Уравнение (3) вырождается в одномерное:

$$p_{k-1} - 2p_k + p_{k+1} - \lambda(p_k - h) = 0, \quad (5)$$

которое имеет единственное положительное, не возрастающее на бесконечности решение:

$$p_k = (1 - h)(1 + \lambda/2 - \sqrt{\lambda + \lambda^2/4})^{|k|} + h \quad (6)$$

при $\lambda \geq 0$. Рассмотрим случай: $h = 0$; тогда

$$p_k = (1 + \lambda/2 - \sqrt{\lambda + \lambda^2/4})^{|k|}. \quad (7)$$

Для отрезка прямой, накрывающего точку $k=0$, определим вероятность того, что длина этого отрезка равна l при условии, что вероятность появления единицы в остальных точках определена выражением (6). Пусть P_l — вероятность реализации отрезка длиной l . Вычислим вероятность появления единицы в k -й точке p_k . Заметив, что отрезки длиной k и меньше никакого вклада в p_k не вносят (так как они не могут одновременно накрыть и k -ю точку, и начало отсчета), рассмотрим случай, когда $l > k$. Тогда отрезок может занять относительно начала отсчета одно из l равновероятных положений, при которых только в $l-k$ случаях он накроет k -ю точку. Итак, вероятность появления единицы в k -й точке при условии, что отрезок имеет длину l ,

$$p_k(l) = (l - k)/l. \quad (8)$$

Теперь нетрудно записать полную вероятность получения единицы в k -й точке:

$$p_k = \sum_{l=k+1}^{\infty} P_l (l - k)/l.$$

Учитывая, что эта же вероятность определяется выражением (7), получим следующее уравнение для P_l :

$$\sum_{l=k+1}^{\infty} P_l (l - k)/l = (1 + \lambda/2 - \sqrt{\lambda + \lambda^2/4})^{|k|}, \quad (9)$$

которое имеет решение

$$P_l = \lambda(1 + \lambda/2 - \sqrt{\lambda + \lambda^2/4})^l, \quad (10)$$

удовлетворяющее условию нормировки

$$\sum_{l=1}^{\infty} P_l = 1.$$

Вычислим среднюю длину отрезка

$$\bar{l} = \sum_{l=1}^{\infty} l P_l = \lambda \sum_{l=1}^{\infty} l^2 (1 + \lambda/2 - \sqrt{\lambda + \lambda^2/4})^l = \sqrt{4/\lambda + 1}. \quad (11)$$

откуда выразим λ через \bar{l} :

$$\lambda = 4/(\bar{l}^2 - 1). \quad (12)$$

Полученное выражение позволяет утверждать, что параметр λ в уравнении (3) зависит от средней длины отрезка, получаемого сечением компонент изображения прямой, причем λ убывает достаточно быстро с увеличением среднего размера компонент.

Рассмотрим два предельных случая: $\bar{l} \rightarrow 1$ и $\bar{l} \rightarrow \infty$. При этом λ будет стремиться к ∞ и к нулю соответственно. В первом случае уравнение (3) преобразуется к виду

$$p_{j,k} - h = 0, \quad (13)$$

и во втором —

$$\Delta p_{j,k} = 0. \quad (14)$$

Решение уравнения (13) на $\bar{\Omega}$ не зависит от задания изображения на точках опорного множества; этим решением всегда будет $p_{j,k} = h$, т. е. здесь уравнение соответствует некоррелированному или слабокоррелированному изображению, состоящему из отдельных не связанных черных точек.

Уравнение же (14) позволяет предсказать все изображение по опорному множеству, состоящему из одной точки (j, k) . Действительно, если $a_{j,k} = 1 \vee 0$ и $(j, k) \in \Omega \equiv (j, k)$, то единственным ограниченным решением уравнения (14) будет $p_{r,s} = 1 \vee 0$ соответственно. В этом случае уравнение (3) следует рассматривать для ансамбля изображений и компонента либо вообще отсутствует, либо покрывает всю сетку, или, говоря иначе, представляет собой либо белый, либо черный лист, причем вероятность реализации черного листа равна h .

Оба рассмотренных случая ($\bar{l} \rightarrow 1$ и $\bar{l} \rightarrow \infty$) для практики никакого интереса не представляют, поскольку обычно имеют дело с изображениями, состоящими из достаточно больших компонент, заданных на матрице ограниченной размерности. Для разрешимости уравнений (3) и (14) на матрице к множеству Ω должна быть присоединена ее граница; тогда предсказание статистики на множестве точек $\bar{\Omega}$ представляет собой задачу интерполяции, а в решении этих задач имеется достаточный опыт. Понятно, что предсказать поведение изображения более или менее определенно по малому количеству известных точек нельзя, поэтому будем считать, что точки Ω достаточно густо покрывают матрицу, причем под словом «достаточно» следует понимать, что расстояние между ближайшими точками Ω меньше \bar{l} . Решения уравнений (3) и (14) для этого случая практически совпадают.

Пусть теперь значения изображения заданы на множестве точек Ω и на границе матрицы и $p_{j,k}$ есть решение уравнения (3). Рассмотрим множество изображений $\{X^l\}$, таких, что $x_{j,k}^l = \omega_{j,k}$, если $(j, k) \in \Omega$. Обозначим через X_1^l множество точек изображения X^l , имеющих значение единицы, и через X_0^l — имеющих значение нуля. Каждому изображению указанного множества можно сопоставить вероятность его реализации, равную

$$P(X^l) = \prod_{(j,k) \in X_1^l} p_{j,k} \prod_{(r,s) \in X_0^l} (1 - p_{r,s}). \quad (15)$$

Заметим, что множество $\{X^l\}$ представляет собой полную систему событий, т. е.

$$\sum_l P(X^l) = 1,$$

в чем можно убедиться непосредственным суммированием (15).

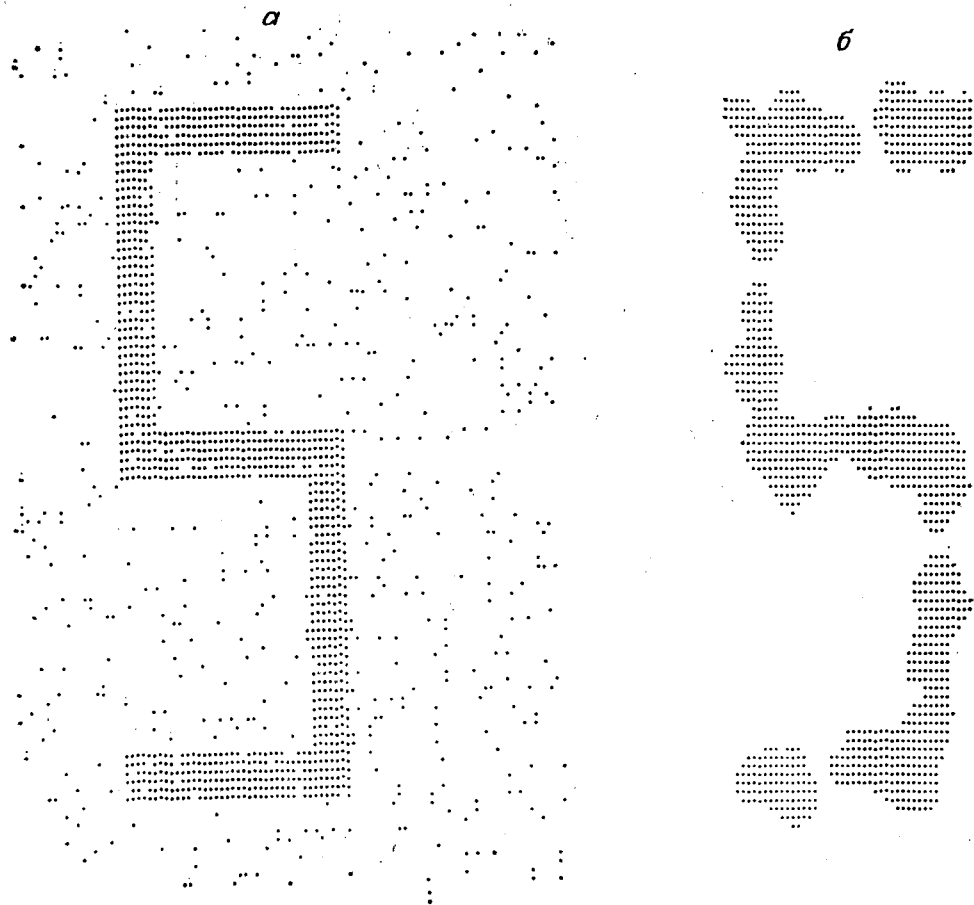


Рис. 1.

Нетрудно видеть, что вероятность (15) достигает максимума на таком X^s , что $(j, k) \in X_1^s$, если $p_{j, k} > 0,5$. Поскольку точки со значением $p_{j, k} = 0,5$ не изменяют величины $P(X^s)$ и являются подмножеством границы X^s , то X^s определена с точностью до граничных точек.

На рис. 1, а изображена цифра 5 на матрице размерностью 100×100 элементов. Множество Ω формировалось датчиком случайных чисел, и на рисунке оно показано черными точками на фоне и отсутствием точек на изображении (всего 1000 точек); к этому множеству присоединена граница матрицы со значением 0. Значение $p_{j, k}$ на Ω определялось из уравнения (3) при $h = 0,1$ и $\lambda = 0,2$. На рис. 1, б показано множество точек при $p_{j, k} > 0,5$. Эксперименты показывают, что изображение на рис. 1, б слабо зависит от значения λ при его изменении от 0 до 1, и, таким образом, для вычисления $p_{j, k}$ на Ω можно использовать уравнение (14). Точки, с $p_{j, k} = 0,5$ и близкие к ним по значению, соответствуют точкам с наибольшей энтропией, поэтому для дальнейшего уточнения исследуемого изображения к ним следует обращаться в первую очередь.

Изображение, представленное на рис. 2, исследовалось в точках со значениями j и k , кратными 5 (всего 361 точка), после чего решалось уравнение (14) при значениях $p_{j, k}$ на границе матрицы, равных нулю.

Далее определялось множество Ω' из следующего условия: если $0,5 - \delta < p_{j, k} < 0,5 + \delta$, то $(j, k) \in \Omega'$. При этом $\mu(\Omega')$ — количество точек

множества Ω' — равнялось 100. Исходное изображение исследовалось в точках Ω' , и таким образом формировалось новое множество $\Omega_{k+1} = \Omega_k \cup \Omega'$. Алгоритм заканчивал работу, когда энтропия, вычисленная по всей матрице,

$$H = - \sum_{j,k=1}^{100} [p_{j,k} \log_2 p_{j,k} + (1 - p_{j,k}) \log_2 (1 - p_{j,k})]$$

становилась равной нулю.

Изображения, соответствующие наибольшей вероятности X^s , выводились после каждого 3-го цикла и показаны последовательно на рис. 2, а—г. Во время работы алгоритма контролировались моменты первого порядка

$$m = 10^{-4} \sum_{j,k=1}^{100} p_{j,k}$$

и частота появления единицы на множествах Ω' . При этом для изображения X , представленного на рис. 2, момент первого порядка совпадал с величиной

$$m' = 10^{-4} \sum_{j,k=1}^{100} x_{j,k}$$

со среднеквадратической ошибкой, вычисленной по всем циклам и равной 10%. Частота же получения единиц равнялась 0,54, среднеквадратичная ошибка — 7%. Причем, как и следовало ожидать, наибольшие ошибки получались в первых циклах. (Напомним, что первоначально изображение было известно на количестве точек, меньшем 4%, и выбор Ω_0 никак не связан с видом X .) Приведенные примеры позволяют сделать вывод, что для стационарных изотропных изображений полученная модель вполне удовлетворительно отражает их статистические свойства.

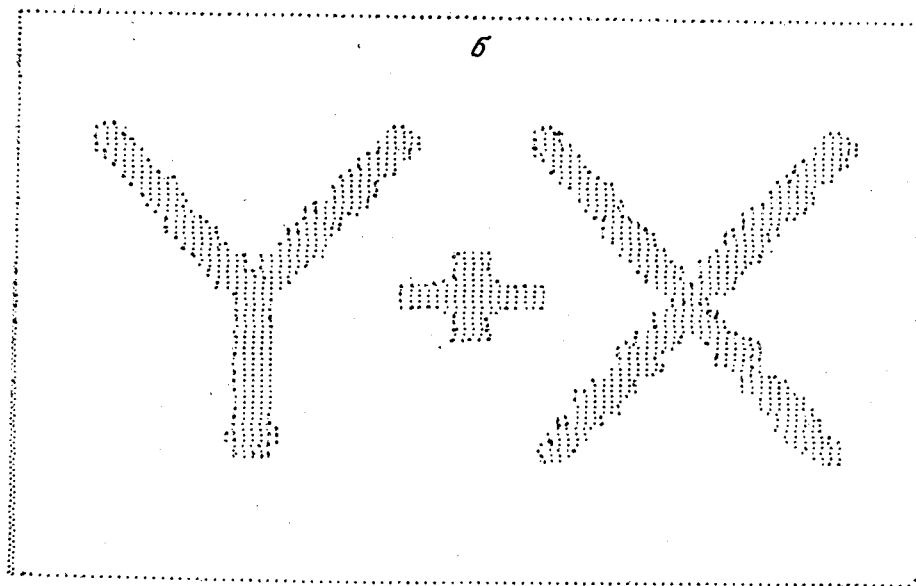
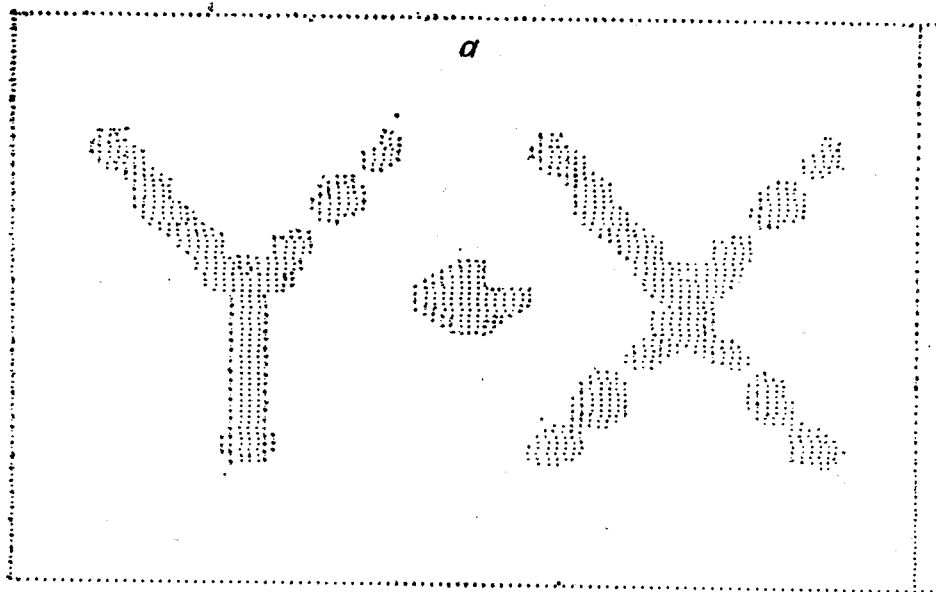
Предложенный выше алгоритм может быть применен для передачи бинарных изображений по каналу связи. Для этого на обоих концах канала должны стоять ЭВМ, вычисляющие поле вероятностей на $\bar{\Omega}$ после каждого обращения к исходному изображению. После чего в памяти ЭВМ ищется точка (a, b) , соответствующая $\min |p_{j,k} - 0,5|$, значение которой снимается с изображения, и образуется очередное множество $\Omega_{k+1} = \Omega_k \cup (a, b)$. Очевидно, что если по каналу связи будут переданы множества точек, принадлежащие границам компонент изображения $\Gamma(X_1)$ и $\Gamma(X_0)$, то при нулевых условиях на границе матрицы и при решении уравнения (10) энтропия станет равной нулю. Таким образом, множество

$$\Omega_{\min} = \Gamma(X_1) \cup \Gamma(X_0) \quad (16)$$

следует понимать как необходимое и достаточное для передачи изображения; при этом количество переданных точек составляет $\mu(\Omega_{\min})$. В реальном же процессе энтропия станет равной нулю при некотором $\Omega_{\text{кон}}$ и всегда будет выполняться неравенство $\mu(\Omega_{\text{кон}}) \geq \mu(\Omega_{\min})$. Отношение этих мер можно принять за коэффициент эффективности алгоритма

$$K_s = \mu(\Omega_{\text{кон}}) / \mu(\Omega_{\min}). \quad (17)$$

В примере, показанном на рис. 2, K_s равнялся 1,3. Заметим, что для передачи всех компонент изображения необходимо, чтобы на каждую из них (или на «дыру» в компоненте) попала хотя бы одна точка. Пусть передаваемое изображение не содержит компонент (или «дыр») с площадью, меньшей S . Перед началом работы алгоритма исследуем изображение на некотором опорном множестве точек Ω_0 , которое может



быть сформулировано аналогично примеру на рис. 2. Тогда вероятность необнаружения минимальной компоненты можно определить из выражения

$$p_{\text{но}} \approx [1 - S/(mn)]^{\mu(\Omega_0)}, \quad (18)$$

где mn — размерность всего изображения. Ограничивая $p_{\text{но}}$ допустимой величиной, после несложных преобразований (18) получим оценку для количества точек опорного множества

$$\mu(\Omega_0) \geq \ln p_{\text{но}} / \ln [1 - S/(mn)]. \quad (19)$$

Очевидно, что точки опорного множества должны равномерно располагаться на исходном изображении.

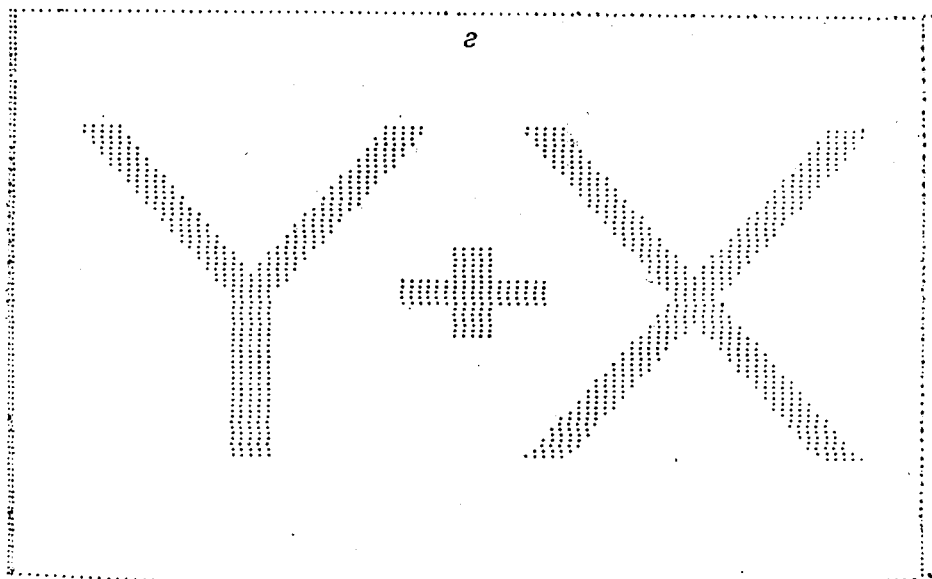
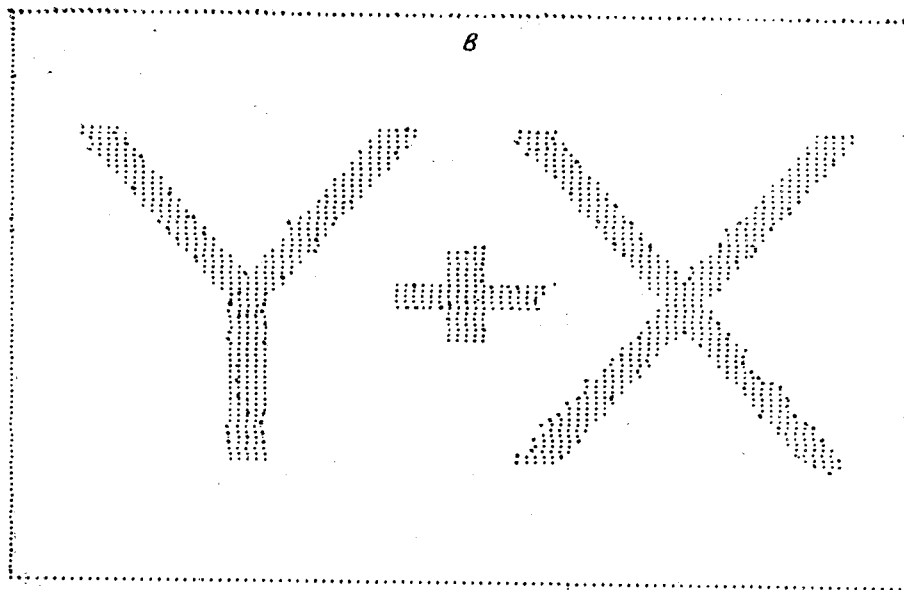


Рис. 2.

В заключение автор выражает благодарность В. Р. Кирейтову, О. Е. Трофимову и В. А. Цецохо за активное обсуждение работы и замечания.

ЛИТЕРАТУРА

1. Уинти Р. Кодирование изображений посредством преобразований.— В кн.: Обработка изображений при помощи цифровых вычислительных машин. М., «Мир», 1973.
2. Лебедев Д. С., Пийль Е. И. Экспериментальные исследования статистики телевизионных сообщений.— «Техника кино и телевидения», 1959, № 3.

Поступила в редакцию 3 января 1979 г.