

Л. А. ВОРОНЦОВА, Г. П. ЧЕИДО

(Новосибирск)

**АЛГОРИТМ АВТОМАТИЧЕСКОГО ОТОЖДЕСТВЛЕНИЯ
ЗВЕЗД СНИМКА С КАТАЛОГОМ**

В задачу обработки астронегативов в самом общем плане входит определение положений естественных и искусственных небесных объектов в какой-либо системе небесных координат. Это достигается измерением на астронегативе прямоугольных координат изображений объектов и опорных звезд. Высокоточные значения небесных координат опорных звезд приведены в звездных каталогах. Для получения небесных координат объектов необходимо осуществить переход от измеренных (прямоугольных) координат к системе координат каталога. Параметры связи этих двух систем находятся из сопоставления измеренных и каталожных координат некоторого количества опорных звезд, для чего нужно отождествить эти звезды на снимке с каталогом. Отождествление сводится к нахождению положения на пластинке всех или части звезд каталога, или, что равносильно, к нахождению каталожных номеров опорных звезд снимка. Обычно эта операция производится опытным специалистом путем одновременного просмотра астронегатива и звездного атласа для сфотографированной области неба, нанесения на астронегативе метки у каждой идентифицированной звезды и составления таблицы соответствия меток и каталожных номеров. Недостатки такого решения задачи отождествления очевидны: низкая оперативность, высокие требования к квалификации обработчика, zagrożение снимка посторонними знаками, ухудшающими его информационные качества. Особенно архаично выглядит такой метод отождествления, если принять во внимание существующий [1] и перспективный [2] уровень автоматизации измерения координат изображений.

В существующих системах автоматизации обработки астронегативов координаты изображений звезд и других небесных тел измеряются полуавтоматическими или автоматическими координатно-измерительными машинами (типа "Ascorecord", GALAXY) и вводятся в ЭВМ для последующей обработки. На таком уровне автоматизации измерений уже возможно произвести автоматическое отождествление опорных звезд.

Содержание задачи сводится к следующему. В звездном каталоге выделяется множество K звезд сфотографированной области неба. По известным из каталога их экваториальным координатам и координатам оптического центра кадра можно вычислить тангенциальные координаты звезд ξ, η . Осями тангенциальной системы координат служат проекции на плоскость снимка круга склонений и круга прямых восхождений, проходящих через точку небесной сферы, соответствующую оптическому центру кадра. Тангенциальная система ξ, η «почти изометрична» прямоугольной системе координат снимка. Различие этих систем сводится к следующему. Начала их несколько сдвинуты относительно друг друга, сдвиг обусловлен погрешностью координат оптического центра, величина сдвига неизвестна, но не превышает 1—2 мм. Оси тангенциальной системы повернуты на некоторый угол, величину которого мы в дальнейшем считаем неизвестной. Вообще тангенциальные координаты являются геометрическим построением, а координатам снимка x, y присущ ряд искажений из-за влияния атмосферы, аберраций объектива и погрешностей измерительного устройства. Часть этих искажений, например дисторсию и рефракцию, можно в значи-

тельной степени компенсировать. Но остаются искажения, природа которых случайна. Все они приводят к тому, что расстояния между тождественными объектами в тангенциальной и прямоугольной системах различаются на некоторую случайную величину. Максимальное значение ε этой величины удается оценить и тогда можно говорить, что системы ξ , η и x , y изометричны с точностью до ε . Отождествление звезд, таким образом, можно осуществить путем сопоставления конфигураций, образуемых изображениями звезд в обеих системах. Описание конфигураций может быть построено заданием их геометрических элементов, таких как расстояния, углы, параметры фигур, состоящих из фиксированного числа точек и т. д. Однако выделение и анализ конфигураций при обширном их описании сложны и требуют много машинного времени. Проще всего отождествление выполняется посредством сопоставления только взаимных расстояний между объектами и углов, образованных соединяющими их векторами. Известный алгоритм [3] основан на таком анализе. Мы также следуем этому принципу.

Задача отождествления заключается в нахождении отображения множества K звезд каталога на некоторое подмножество S' множества S звезд снимка. Отображение осуществляется неизвестным преобразованием координат, сохраняющим изометрию с точностью до ε , так что элементу $k_j (j=1, \dots, n)$ из K (прообразу) соответствует единственный элемент $s_i (i=1, \dots, m)$ в S' (образ). Выделим два элемента s_p, s_q множества S' и определим расстояние d_{pq} между ними. Очевидно, что их прообразы k_p, k_s в множестве K должны удовлетворять условию

$$|d_{rs} - d_{pq}| \leq \varepsilon. \quad (1)$$

Следовательно, выделяя подмножество пар индексов элементов из удовлетворяющих (1)

$$K_{pq} = \bigcup_{\substack{\{r, s\} \\ |d_{rs} - d_{pq}| < \varepsilon \\ r, s \in K}} \{r, s\}, \quad (2)$$

мы перечислим все элементы множества K , являющиеся «претендентами» на соответствие элементам s_p, s_q . При малом ε и низкой плотности распределения звезд по полю снимка может оказаться, что K_{pq} состоит всего из двух индексов и для окончания отождествления элементов s_p, s_q остается только найти, какой из индексов r и s соответствует, например, индексу p . Однако на практике снимок может содержать десятки тысяч звезд и плотность распределения звезд на снимке оказывается столь высокой, что при данном ε K_{pq} содержит много индексов и, тем самым, отождествление по двум элементам оказывается невозможным. Приходится брать более обширное подмножество $Q = \{s_{q_1}, s_{q_2}, \dots, s_{q_l}\}$ элементов из S' . Каждый элемент подмножества Q образует с остальными семейство из $l-1$ расстояний. Соответственно с этим для каждой пары $s_{q_i} s_{q_j} (i, j = 1, \dots, l; i \neq j)$ построим множества $K_{q_i q_j}$. Всего получим C_l^2 таких множеств, причем для каждого элемента $s_{q_i} (i = 1, \dots, l)$ существует $l-1$ множество

$$K_{q_1 q_i}, K_{q_2 q_i}, \dots, K_{q_{i-1} q_i}, K_{q_{i+1} q_i}, \dots, K_{q_l q_i}, \quad (3)$$

в каждом из которых содержатся индексы претендентов на соответствие элементу s_{q_i} . Очевидно, что истинный индекс элемента, являющегося прообразом s_{q_i} , должен содержаться во всех множествах системы (3). Для его определения нужно построить пересечение множеств этой системы:

$$K_{q_i} = \bigcap_{\substack{j=1, \dots, l \\ j \neq i}} K_{q_i q_j}; \quad i = 1, \dots, l. \quad (4)$$

Наличие в каком-либо из K_{q_i} единственного индекса означает, что он является индексом прообраза элемента s_{q_i} и, тем самым, отождествление s_{q_i} закончено.

Если после выполнения операции (4) множества претендентов K_{q_i} ($i = 1, \dots, l$) содержат более чем по одному индексу, мы можем повторить операцию (2), но уже не со всем множеством K , а с элементами, входящими в K_{q_i} :

$$K_{q_i q_j}^{(2)} = \left\{ \begin{array}{l} \{r, s\} \\ |d_{rs} - d_{q_i q_j}| \leq \varepsilon \\ r \in K_{q_i}^{(1)}; s \in K_{q_j}^{(1)} \end{array} \right. \quad (5)$$

Мы получили более узкие множества $K_{q_i q_j}^{(2)}$ (из более узких множеств отбираются пары r, s , сопоставляемые по расстоянию с элементами s_{q_i}, s_{q_j}). Применяя к полученным множествам снова операцию (4), можно добиться дальнейшего сокращения множества претендентов $K_{q_i}^{(2)}$ и т. д. Этот итерационный процесс будем продолжать до тех пор, пока он приводит к сокращению количества претендентов. Если по окончании итераций отождествление оказалось неоконченным, следует расширить множество Q включением в него ряда новых элементов и повторить итерации с расширенным множеством. Вообще говоря, вся необходимая информация содержится в множествах (3), полученных уже на первом шаге, и анализ ее может производиться различными способами. Аналогично [3] мы применяем итерационный процесс, так как он представляется наиболее удобным способом такого анализа.

Операции (4) и (5) имеют смысл отыскания в множестве K точек, с семейством расстояний с точностью до ε таким же, какое имеет элемент $s_{q_i} \in Q$. Однако не исключено, что такое же семейство кроме прообраза образует ряд других элементов K , так что после первого шага итерационного процесса для некоторых элементов s_{q_i} ($i = 1, \dots, l$) остается $n_i > 1$ претендентов. Обозначим M мощность множества претендентов после первого шага: $M = \sum_{i=1}^l n_i$. Последующие итерации приводят к постепенному сокращению количества претендентов, причем как скорость сокращения, так и вероятность окончания отождествления при заданном Q зависят от M . Для среднего значения величины M можно дать оценку сверху. Она связана с вероятностью $P(A)$ события A , заключающегося в том, что произвольная из n случайно распределенных звезд множества K среди $n-1$ расстояний до других звезд содержит хотя бы одно семейство из $l-1$ расстояний ρ_i , удовлетворяющих условиям $|\rho_i - d_i| \leq \varepsilon$ ($i = 1, \dots, l-1$). Для вычисления $P(A)$ нам потребуется распределение расстояний между случайно расположенными точками.

Если в прямоугольнике $0 \leq x \leq a$; $0 \leq y \leq b$; $a, b > 0$; $a < b$ расположены точки $x_1 y_1$ и $x_2 y_2$, представляющие собой выборку из равномерного распределения, то плотность распределения величины $z_1 = x_2 - x_1$

$$\omega_{z_1}(z) = \begin{cases} 0; & |z| > a; \\ \frac{1}{a} \left(1 + \frac{z}{a}\right); & -a \leq z \leq 0; \\ \frac{1}{a} \left(1 - \frac{z}{a}\right); & 0 \leq z \leq a. \end{cases}$$

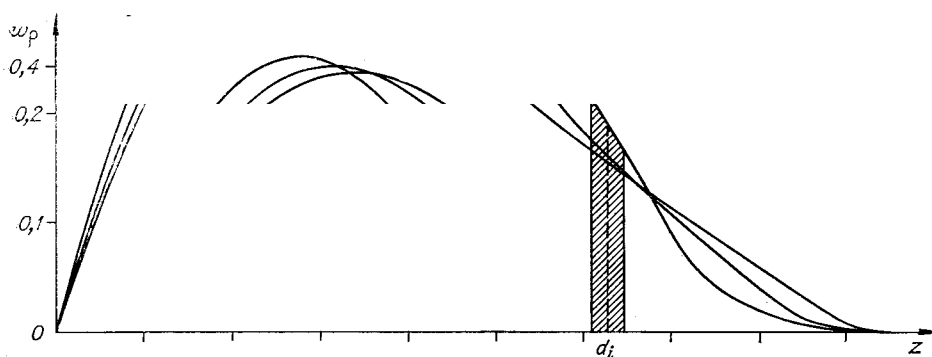


Рис. 1.

Распределение разности $z_2 = y_2 - y_1$ независимо от распределения z_1 и имеет такой же вид (с заменой параметра a на b). Исходя из этого, получаем плотность вероятности распределения расстояния

$\rho = \sqrt{z_1^2 + z_2^2}$ [4]:

$$\omega_\rho(z) = \begin{cases} \frac{2z}{ab} \left[\pi - 2z \left(\frac{1}{a} + \frac{1}{b} \right) + \frac{z^2}{ab} \right]; & 0 \leq z \leq a; \\ \frac{2z}{ab} \left[\pi - \arcsin \sqrt{\frac{z^2 - a^2}{z^2}} + \frac{2}{a} (\sqrt{z^2 - a^2} - z) \right]; & a \leq z \leq b; \\ \frac{2z}{ab} \left[2 \arcsin \frac{b}{z} - 2 \arcsin \sqrt{\frac{z^2 - a^2}{z^2}} + \frac{2}{a} \sqrt{z^2 - a^2} - 2 \frac{b}{a} - \right. \\ \quad \left. - \frac{a}{b} + 2 \sqrt{\frac{z^2 - b^2}{b}} - \frac{z^2 - b^2}{ab} \right]; & b \leq z \leq \sqrt{a^2 + b^2}. \end{cases} \quad (6)$$

Графики этой плотности при различных отношениях сторон прямоугольника приведены на рис. 1. Используя (6), можно для любого заданного d_i найти вероятность p_i попадания случайного расстояния ρ между произвольными точками в интервал $d_i \pm \varepsilon$:

$$p_i = P\{|\rho - d_i| \leq \varepsilon\},$$

что соответствует площади заштрихованной области на рис. 1. Практически при размерах астронегатива 300×300 мм величина ε заведомо меньше 1 мм, т. е. $1/300 \sqrt{2} = 0,0024$ от интервала распределения (6). Так как максимальная плотность распределения $\sim 0,04$, наибольшее значение вероятностей p_i не превышает 0,002. Надлежащим выбором d_i эти вероятности можно уменьшить еще минимум на порядок.

Обозначим k_i количество расстояний, попавших в интервал $d_i \pm \varepsilon$ ($i=1, \dots, l-1$) при предъявлении $n-1$ случайных расстояний из распределения (6). Если интервалы не перекрываются, вероятности попадания в каждый из $L=l-1$ интервалов по k_i расстояний подчиняются полиномиальному распределению

$$P(k_1, \dots, k_{l-1}, k_l) = \frac{(n-1)!}{k_1! \dots k_{l-1}! k_l!} p_1^{k_1} \dots p_{l-1}^{k_{l-1}} p_l^{k_l}.$$

Здесь $k_l = n - 1 - \sum_{i=1}^L k_i$ — количество расстояний, не попавших ни в один

Таблица 1

N	l									
	2	3	4	5	6	7	8	9	10	20
100	0,0326	0,0058	0,0010	0,0002						
200	0,1085	0,0356	0,0116	0,0038	0,0012	0,0004	0,0001			
300	0,2035	0,0916	0,0411	0,0184	0,0082	0,0037	0,0016	0,0007	0,0003	
500	0,3998	0,2525	0,1594	0,1005	0,0634	0,0399	0,0251	0,0158	0,0099	0,0001
700	0,5600	0,4278	0,3222	0,2426	0,1826	0,1374	0,1033	0,0777	0,0584	0,0031
1000	0,7480	0,6469	0,5593	0,4836	0,4181	0,3614	0,3124	0,2699	0,2330	0,0179

из интервалов;
$$p_l = 1 - \sum_{i=1}^L p_i.$$

Нас интересует вероятность события A , заключающегося в том, что ни один из интервалов $d_i \pm \varepsilon$ ($i=1, \dots, L$) не пуст. Вероятность попадания в объединение L интервалов m расстояний подчиняется биномиальному распределению

$$P(m) = \frac{(n-1)!}{m!(n-m-1)!} \left(\sum_{i=1}^L p_i \right)^m \left(1 - \sum_{i=1}^L p_i \right)^{n-m-1}.$$

Дадим оценку величины $P(A)$ сверху, для чего положим все p_i одинаковыми:

$$p_i = p, \quad i=1, \dots, L. \quad (7)$$

Тогда вероятность такого распределения m расстояний по L интервалам, при котором нет пустых интервалов [5]

$$P(0/m) = \sum_{v=0}^L (-1)^v C_L^v \left(1 - \frac{v}{L} \right)^m.$$

Отсюда граница сверху для $P(A)$

$$\hat{P}(A) = \sum_{m=L}^{n-1} \frac{(n-1)!}{m!(n-m-1)!} (Lp)^m (1-Lp)^{n-m-1} \sum_{v=0}^L (-1)^v C_L^v \left(1 - \frac{v}{L} \right)^m. \quad (8)$$

В табл. 1 приведены значения $\hat{P}(A)$ для некоторых $N=n-1$ и L при $p=0,002$. По величине $\hat{P}(A)$ определяем границу сверху для среднего значения мощности M после первого шага итераций: $M_{ср} = n\hat{P}(A) \leq \leq n\hat{P}(A)$. $M_{ср}$ так же, как и $P(A)$, убывает с расширением множества Q . Вообще говоря, не обязательно добиваться отождествления всех звезд из Q . Для полного решения задачи достаточно идентифицировать несколько звезд, а затем, воспользовавшись какой-либо формой связи координат, например, через принятые в астрометрии коэффициенты Тернера [6]

$$\begin{aligned} x &= a + b\xi + c\eta; \\ y &= d + e\xi + f\eta, \end{aligned} \quad (9)$$

найти приближенные значения коэффициентов a, \dots, f . Далее, воспользовавшись (1), для заданных координат ξ, η любой звезды можно найти координаты x, y той же звезды на снимке, тем самым устанавливая соответствие звезд снимка с каталогом.

Трудности реализации описанной процедуры имеют алгоритмический, но не теоретический характер и обусловлены перебором элементов при поиске заданных расстояний, что приводит к большим затратам машинного времени. Например, если множество K звезд каталога содержит 1000 звезд, на первом этапе нужно осуществить перебор $S_{1000}^2 \approx 5 \cdot 10^5$ расстояний и каждое сравнить с $l(l-1)/2$ расстояниями, присущими элементам множества Q . Необходимо всеми средствами сокращать количество элементов из K , по которым производится отождествление. Кроме сокращения перебора это уменьшит в соответствии с (8) вероятность $P(A)$ или при той же вероятности позволит взять более узкое множество Q , что также уменьшает время работы алгоритма. Можно предложить три способа сокращения множества K .

1. Выделение кольцевых зон в множестве K , основанное на близости центров систем координат x, y и ξ, η . Расстояние Δ между центрами обеих систем при размерах негатива 300×300 мм, как правило, не превышает 2 мм. Очевидно, что звезда s_j , находящаяся в системе x, y на расстоянии R от центра, наверняка попадает в кольцо, образованное окружностями радиусов $R - (\Delta + \epsilon)$ и $R + (\Delta + \epsilon)$ в системе координат ξ, η . Мы можем все звезды из K , попавшие в данное кольцо, считать претендентами на соответствие s_j . Максимальная площадь кольца (при $R = a/2$) в $\frac{a^2}{4\pi R(\Delta + \epsilon)}$ раз меньше площади всего снимка. В среднем во столько же раз сокращается перебор претендентов для s_j .

2. Выделение наиболее ярких звезд сфотографированной области неба. Это выделение можно произвести, используя приведенные в каталогах звездные величины. Особенно эффективен этот способ при использовании автоматической измерительной системы, описанной в [2], где оператор может визуально выделить наиболее яркие звезды снимка и построить из них подмножество Q . Для отождествления в множестве K достаточно выделить такое подмножество K' , в которое наверняка входят прообразы всех элементов Q . Разумеется, объем K' зависит от точности визуального выделения наиболее ярких звезд. Во всяком случае подмножество K' можно взять с многократным запасом, так как Q содержит не более десятка звезд.

3. Использование преобладания доли малых расстояний перед большими. Интуитивно ясно, что точки, расположенные на ограниченной площадке, образуют больше малых расстояний, чем больших. Это подтверждает распределение (6) (см. рис. 1). Плотность распределения $w_0(z)$ имеет единственный максимум при

$$z = \frac{4(a+b) - \sqrt{16(a+b)^2 - 12\pi}}{6}.$$

При $a=b$ максимум расположен в точке $z \approx 0,477a$, т. е. доля расстояний, равных половине стороны квадрата, максимальна. Отсюда очевидно нецелесообразность привлечения звезд с такими взаимными расстояниями в множество Q . Следует отдать предпочтение звездам с максимально возможными для данного снимка расстояниями.

Использование первого и второго способов позволяет во много раз сократить множество K ; даже при очень высокой плотности расположения звезд на снимке объем K может быть сокращен до сотни звезд.

Описанный выше алгоритм запрограммирован в языке «Макрокод» ЭВМ «Минск-22М». Блок-схема программы приведена на рис. 2. Блок «Претенд» осуществляет одну итерацию по сокращению множества претендентов, т. е. выполняет операции (5) и (4). После каждой итерации производится проверка окончания отождествления. Если количество $N_{от}$ отождествленных звезд равно l , производится вычисление коэффициентов Тернера и отождествление оставшихся звезд снимка

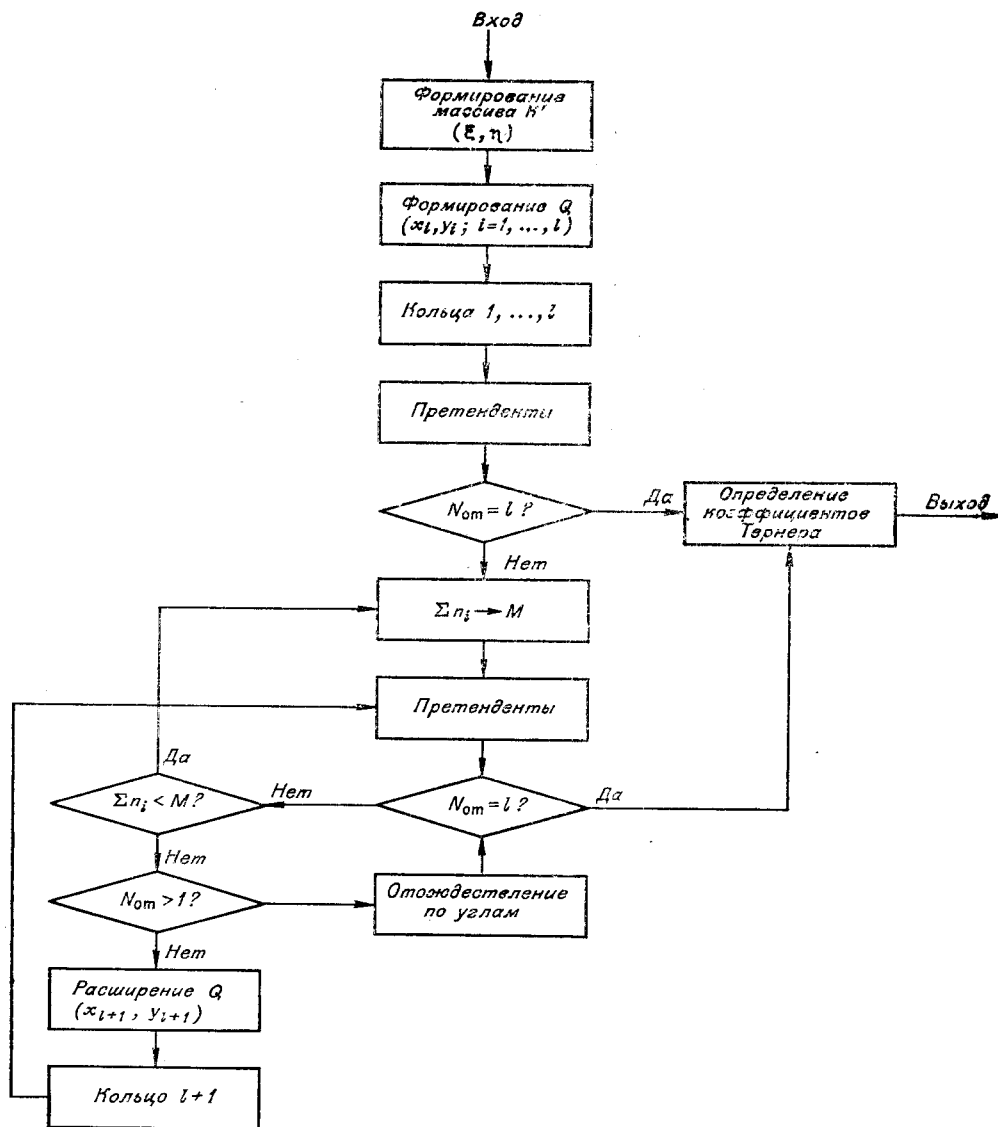


Рис. 2.

при помощи (9). В противном случае итерации продолжают, пока они приводят к сокращению количества претендентов, обозначенного $\sum n_i$, где n_i — число претендентов в K_{q_i} ($i = 1, \dots, l$). По окончании итераций, если идентифицировано более одной звезды, производится отождествление по углам, образованным векторами, соединяющими любые две звезды. Понятно, что для такого отождествления нужно иметь минимум две идентифицированные звезды. Если их не нашлось, производится расширение множества Q посредством измерения координат x, y еще одной звезды, выделяется кольцевая зона для нее и процесс продолжается.

Работа алгоритма проверялась на моделях случайного распределения звезд. Одна из моделей, содержащая 100 звезд, приведена на рис. 3. Размеры кадра 300×300 мм.

Величина сдвига Δ начала тангенциальной системы относительно системы x, y принята равной 10 мм; $\varepsilon = 7$ мм. О динамике работы алгоритма можно судить по данным табл. 2 и 3, где приводятся номера

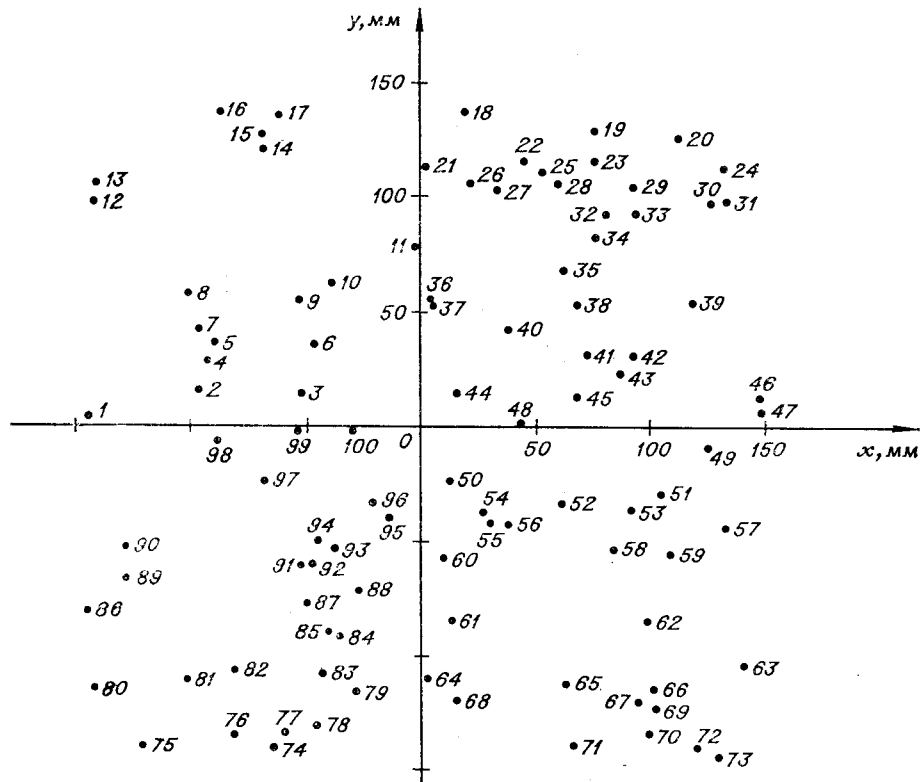


Рис. 3.

звезд множества Q и список претендентов для каждой звезды после различных этапов работы алгоритма. Из табл. 2 следует, что неоптимальный выбор звезд для множества Q не позволяет провести отождествление по трем звездам (расстояния между звездами 50, 95, 96 малы и для них находится слишком много претендентов). Возникает необходимость расширения множества Q , причем одной дополнительной звезды 100 оказывается недостаточно — она тоже имеет малые расстояния до всех звезд из Q . Только вводом пятой звезды можно достигнуть успеха. В то же время табл. 3 показывает, что звезд с большими взаимными расстояниями достаточно взять всего три и отождествление будет закончено за одну итерацию. Время работы программы в этом случае составляет несколько секунд.

Таблица 2

Q	S_i после работы блока „Кольцо“	S_i в конце итераций	Q	S_i в конце итераций	Q	S_i в конце итераций	S_i после отождествления по углам
50	44, 100, 50	44, 100, 50	50	44, 100, 50	50	50	50
95	55, 54, 48, 96, 95	48, 96, 95	95	48, 96, 95	95	96, 95	95
96	54, 48, 96, 95, 100	48, 96, 95	96	49, 96, 95	96	96	96
			100	100, 50	100	100	100
					48	48	48

Таблица 3

Q	S_i после работы блока «Кольцо»	S_i после итераций
73	72, 75, 73	73
80	16, 72, 12, 80, 75, 13, 63, 24	80
20	16, 31, 70, 20, 12, 86, 30, 63, 69, 24	20

ЛИТЕРАТУРА

1. H. M. Pratt. The Performance of CALAXY.— Publ. Roy. Observ. Edinburgh, 1971, v. 8, N 11.
2. Л. В. Бурый, В. П. Коронкевич, Ю. Е. Нестерихин, А. А. Нестеров, Б. М. Пушной, С. Е. Ткач, А. М. Щербаченко. Прецизионный фотограмметрический автомат.— Автометрия, 1974, № 4.
3. Я. К. Балодис. Об автоматическом отождествлении звезд.— Ученые записки Латв. ГУ, т. 169. Рига, 1972.
4. E. Vogel. Principes et Formules Classiques du Calcul des Probabilites. Traite du Calcul des Probabilites et de ses Applications. Paris, Gauthier — Villars, 1925.
5. В. Феллер. Введение в теорию вероятностей и ее приложения. М., «Мир», 1967.
6. В. В. Подобед. Фундаментальная астрометрия. М., «Наука», 1968.

Поступила в редакцию 28 января 1974 г.

УДК 681.3.06

А. М. ОСТАПЕНКО, Э. А. ТАЛНЫКИН

(Новосибирск)

ОБ ОДНОМ АЛГОРИТМЕ ОБРАБОТКИ ИЗОБРАЖЕНИИ

Введение. При автоматической обработке результатов эксперимента, зарегистрированных на фотопленке, возникает необходимость в устройстве, способном измерять плотность изображения в точке с заданными координатами и передавать результат в ЭВМ. Цифровой сканирующий автомат (ЦСА) [1] отвечает этим требованиям. Сложность обработки изображения таким инструментом обусловлена в основном последовательным восприятием информации вычислительной машиной. Попытка распознать в такой ситуации, что же все-таки изображено, эквивалентна попытке рассмотреть картину через черный лист бумаги с «окном» диаметра намного меньшего, чем характерный размер детали изображения.

Автоматическое распознавание изображения представляется непомерно сложной задачей, если не известна примерная конфигурация изображения. Например, если известно, что на снимке присутствует прямая линия, то сравнительно просто построить алгоритм ее нахождения, т. е. определения углового коэффициента или других параметров. В этой связи представляется эффективной начальная оценка изображения человеком для определения некоторых глобальных характеристик, таких как «типичные» точки, начало поиска и т. д. При наличии экранного нуля или другого средства активного взаимодействия такой подход может дать положительный эффект, исключая огромные затраты на автоматическую реализацию «оценочных» функций.