
Информационные технологии

УДК 004.89:004.932

Распознавание на основе инвариантных моментов

Н. С. Абрамов*, В. М. Хачумов†

* *Федеральное государственное бюджетное учреждение науки
Институт программных систем им. А.К. Айламазяна РАН
ул. Петра I, д. 4а, с. Веськово, Переславский район, Ярославская область, 152021*

† *Федеральное государственное бюджетное учреждение науки
Институт системного анализа Российской академии наук
пр-т 60-летия Октября, д. 9, Москва, Россия, 117312*

Исследуются свойства инвариантных моментов бинарных изображений, что необходимо для формирования их набора в задаче распознавания графических образов. Для распознавания образов и измерения расстояний используются инвариантные моменты Ху. Показано, что инварианты имеют разную чувствительность к изменениям входных данных, что определяет стратегию их выбора. Проведены эксперименты по распознаванию изображений текстовых символов, изображений летательного аппарата и посадочной площадке в форме креста. Рассматриваемый алгоритм распознавания характеризуется работой в режиме реального времени, использованием только одной видеокамеры, инвариантностью к повороту, сдвигу и масштабу объекта в кадре. Точность и полнота распознавания составили порядка 92% на выборке из тысячи образцов каждого типа. Приведены результаты экспериментального определения качества распознавания различных объектов по их контурным изображениям, а также результаты сравнения распознавания с использованием разных наборов инвариантных моментов. Показано, что учёт менее чувствительных инвариантных моментов сокращает время вычислений и снижает погрешность вычислений, которая возникает при флуктуации параметров объекта или сцены в кадре. Предложено объединить метод инвариантных моментов с вероятностной нейронной сетью, что позволит улучшить качество распознавания по скорости, точности и полноте.

Ключевые слова: распознавание, инварианты, нейронная сеть.

Введение

Задачи распознавания целевых объектов и определения расстояний с применением систем технического зрения возникают при необходимости осуществлять слежение за целями, построении альтернативных источников навигации, идентификации динамических объектов в охранных системах и решении других важнейших вопросов. Была исследована принципиальная возможность применения для указанных целей фото и видеокамеры с соответствующим математическим и программным обеспечением, построенным на основе алгоритмов машинной графики [1] и инвариантных моментов [2, 3]. Несмотря на большое количество работ, посвящённых отдельным вопросам применения технического зрения, в настоящее время отсутствуют универсальные системы, способные одновременно решать задачи распознавания и определения расстояния на единых методологических принципах. Подобная перспектива открывается, на наш взгляд, сочетанием ряда известных и модифицированных методов, которые будут рассмотрены в данной работе. В частности, будут рассмотрены методы и алгоритмы распознавания образов на изображении, включая метод, основанный на инвариантных моментах

Статья поступила в редакцию 3 февраля 2014 г.

Работа выполнена в рамках проекта по Программе фундаментальных исследований РАН, а также в рамках проектов РФФИ № 13-07-00025 А «Исследование методов анализа интегрированной текстовой, графической и речевой информации в системах интеллектуального управления динамическими объектами» и № 14-07-31020 мол_а «Разработка и исследование математического и программного обеспечения многофункциональных бортовых систем машинного зрения летательных аппаратов на основе технологий анализа видеоданных и методов интеллектуального управления».

Ху, а также гибридный метод распознавания на основе инвариантов и искусственной нейронной сети.

1. Распознавание образов и измерение расстояний на основе инвариантных моментов

Возможность извлекать инвариантные дескрипторы важна для распознавания динамических объектов независимо от их позиции, масштаба и ориентации на плоскости. Математическое обоснование метода основывается на теории алгебраических инвариантов Ху [4]. Пусть задан контур объекта перечислением своих N точек (x_i, y_i) . Алгоритм получения инвариантов содержит следующие этапы:

1) определение центральных моментов порядка не выше третьего:

$$m_{pq} = \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^p (y_i - \bar{y})^q, \quad p + q \leq 3;$$

\bar{x} и \bar{y} – средние значения,

$$\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i; \quad \bar{y} = \frac{1}{N} \sum_{i=1}^N y_i;$$

2) получение моментов, инвариантных к операциям поворота, переноса и зеркального отображения:

$$M_1 = m_{20} + m_{02}, \quad M_2 = (m_{20} - m_{02})^2 + 4m_{11}^2,$$

$$M_3 = (m_{30} - 3m_{12})^2 + (3m_{21} - m_{03})^2, \quad M_4 = (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2,$$

$$M_5 = (m_{30} - 3m_{12})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] +$$

$$+ (3m_{21} - m_{03})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2],$$

$$M_6 = (m_{20} - m_{02})[(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] +$$

$$+ 4m_{11}(m_{30} + m_{12})(m_{21} + m_{03}),$$

$$M_7 = (3m_{21} - m_{03})(m_{30} + m_{12})[(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] -$$

$$- (m_{30} - 3m_{12})(m_{21} + m_{03})[3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2];$$

3) получение моментов, инвариантных к полной группе аффинных преобразований:

$$M'_1 = r \cdot h, \quad M'_2 = \frac{M_2}{r^4}, \quad M'_3 = \frac{M_3}{r^6}, \quad M'_4 = \frac{M_4}{r^6},$$

$$M'_5 = \frac{M_5}{r^{12}}, \quad M'_6 = \frac{M_6}{r^8}, \quad M'_7 = \frac{M_7}{r^{12}}.$$

Величина $r = \sqrt{(\mu_{20} + \mu_{02})}$ характеризует размер изображения проецируемого на плоскость наблюдения. Если величина h есть расстояние до объекта, то из геометрических свойств проективных преобразований следует:

$$M'_1 = r \cdot h = \text{const.}$$

Таким образом, имея заранее вычисленную «псевдоконстанту» M'_1 для каждого объекта, можно вычислять в каждый момент времени величину r и расстояние h [5].

Для распознавания объектов можно использовать любые инварианты кроме первого. Инвариант M'_1 является константой для используемой камеры и наблюдаемого объекта, поэтому с его помощью по известному значению r можно определить расстояние до объекта.

Анализируемая база снимков содержала по тысяче контуров объектов различных классов. Образцы были получены путём фотографирования объектов под разными ракурсами и с разных расстояний с помощью web-камеры экспериментального стенда с последующим выделением внешних контуров.

Общий алгоритм распознавания:

1. Вход: база эталонов с вычисленными инвариантами.
2. Получение очередного кадра с видеокамеры.
3. Предобработка кадра: фильтрация шумов, повышение контрастности, нахождение контуров в кадре.
4. Вычисление инвариантов для каждого контура.
5. Вычисление расстояния между объектами A (эталонный объект) и B (объект в кадре):

$$I(A, B) = \sum_{i=2,3,4,6} \left| \frac{m_i(A) - m_i(B)}{m_i(A)} \right|,$$

где $m_i(X) = \text{sign}(M'_i(X)) \cdot \ln(|M'_i(X)|)$, $M'_i(X)$ — соответствующий инвариант изображения X .

6. Распознавание объекта по заданному критерию (шаг 5).
7. Возврат к шагу 2.

В табл. 1 показан пример определения инвариантных моментов для контурных изображений некоторых самолётов [6].

Таблица 1

Инвариантные моменты самолётов

| Момент | Миг-29 | Миг-31 | Су-27 | Су-37 |
|--------|----------|----------|-----------------------|-------------------------|
| M_1 | 44, 117 | 106, 886 | 45, 200 | 44, 98 |
| M_2 | 0, 1573 | 0, 20717 | 0, 3574 | 0, 35327 |
| M_3 | 0, 7588 | 0, 3819 | 0, 7403 | 0, 40862 |
| M_4 | 0, 2332 | 0, 13167 | 0, 44153 | 0, 16889 |
| M_5 | 0, 09715 | 0, 02767 | 0, 2524 | 0, 04436 |
| M_6 | 0, 08701 | 0, 05982 | 0, 2639 | 0, 1003 |
| M_7 | 0, 01343 | 0, 01031 | $-1, 2 \cdot 10^{-3}$ | $-4, 146 \cdot 10^{-5}$ |

Использование алгоритмов, анализирующих контуры с помощью инвариантных моментов, позволяет быстро и качественно распознавать объекты разных классов, кроме того, такие алгоритмы устойчивы к изменениям освещения. Недостатком служит их чувствительность к помехам, связанных с дрожанием изображения видеокамеры.

2. Проверка устойчивости моментов к искажениям данных

Положим, что точность измерения координат точек объекта равна некоторой величине δ , $0 < \delta \leq 1$. Причём «1» соответствует наивысшей точности. Положим для простоты, что $\delta x_i = \delta y_i$ — возможные изменения (флуктуации) измеряемых

переменных x_i, y_i , вызванные ошибками измерений. Тогда, центральный момент μ_{11} с учётом этих ошибок будет вычислен как:

$$\frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})(\delta y_i - \delta \bar{y}) = \delta^2 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) = \delta^2 \mu_{11},$$

следовательно, чувствительность μ_{11} к изменениям или точности данных пропорциональна или равна δ^2 . Аналогично получим для других моментов:

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})^2 &= \delta^2 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2 = \delta^2 \mu_{20}, \\ \frac{1}{N} \sum_{i=1}^N (\delta y_i - \delta \bar{y})^2 &= \delta^2 \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^2 = \delta^2 \mu_{02}, \\ \frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})(\delta y_i - \delta \bar{y})^2 &= \delta^3 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})^2 = \delta^3 \mu_{12}, \\ \frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})^2(\delta y_i - \delta \bar{y}) &= \delta^3 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^2(y_i - \bar{y}) = \delta^3 \mu_{21}, \\ \frac{1}{N} \sum_{i=1}^N (\delta x_i - \delta \bar{x})^3 &= \delta^3 \frac{1}{N} \sum_{i=1}^N (x_i - \bar{x})^3 = \delta^3 \mu_{30}, \\ \frac{1}{N} \sum_{i=1}^N (\delta y_i - \delta \bar{y})^3 &= \delta^3 \frac{1}{N} \sum_{i=1}^N (y_i - \bar{y})^3 = \delta^3 \mu_{03}. \end{aligned}$$

Чувствительность моментов инвариантных к операциям поворота и сдвига:

$$\begin{aligned} \delta^2 \mu_{20} + \delta^2 \mu_{02} &= \delta^2 (\mu_{20} + \mu_{02}) = \delta^2 M_1, \\ (\delta^2 \mu_{20} - \delta^2 \mu_{02})^2 + 4(\delta^2 \mu_{11})^2 &= \delta^4 ((\mu_{20} - \mu_{02})^2 + 4\mu_{11}^2) = \delta^4 M_2, \\ (\delta^3 \mu_{30} - 3\delta^3 \mu_{12})^2 + (3\delta^3 \mu_{21} - \delta^3 \mu_{03})^2 &= \delta^6 M_3, \\ (\delta^3 \mu_{30} - \delta^3 \mu_{12})^2 + (\delta^3 \mu_{21} + \delta^3 \mu_{03})^2 &= \delta^6 M_4, \\ (\delta^3 \mu_{30} - 3\delta^3 \mu_{12})(\delta^3 \mu_{30} + \delta^3 \mu_{12}) \times [(\delta^3 \mu_{30} + \delta^3 \mu_{12})^2 - \\ &- 3(\delta^3 \mu_{21} + \delta^3 \mu_{03})^2] + (3\delta^3 \mu_{21} - \delta^3 \mu_{03})(\delta^3 \mu_{21} + \delta^3 \mu_{03}) \times \\ &\times [3(\delta^3 \mu_{30} + \delta^3 \mu_{12})^2 - (\delta^3 \mu_{21} + \delta^3 \mu_{03})^2] = \delta^{12} M_5, \\ (\delta^2 \mu_{20} - \delta^2 \mu_{02})[(\delta^2 \mu_{30} + \delta^2 \mu_{12})^2 - (\delta^2 \mu_{21} + \delta^2 \mu_{03})^2] + \\ &+ \delta^2 \mu_{11}(\delta^2 \mu_{30} + \delta^2 \mu_{12})(\delta^2 \mu_{21} + \delta^2 \mu_{03}) = \delta^6 M_6, \\ (3\delta^3 \mu_{21} - \delta^3 \mu_{03})(\delta^3 \mu_{30} + \delta^3 \mu_{12})[(\delta^3 \mu_{30} + \delta^3 \mu_{12})^2 - \\ &- (3\delta^3 \mu_{21} + \delta^3 \mu_{03})^2] - (\delta^3 \mu_{30} - 3\delta^3 \mu_{12})(\delta^3 \mu_{21} + \delta^3 \mu_{03}) \times \\ &\times [3(\delta^3 \mu_{30} + \delta^3 \mu_{12})^2 - (\delta^3 \mu_{21} + \delta^3 \mu_{03})^2] = \delta^{12} M_7. \end{aligned}$$

Результаты исследований чувствительности моментов к искажениям представлены в табл. 2.

Таблица 2

Инвариантные моменты самолётов

| Моменты | M_1 | M_2 | M_3 | M_4 | M_5 | M_6 | M_7 |
|------------------|------------|------------|------------|------------|---------------|------------|---------------|
| Чувствительность | δ^2 | δ^4 | δ^6 | δ^6 | δ^{12} | δ^6 | δ^{12} |

Полученная оценка свидетельствует о том, что разные моменты имеют различную устойчивость к искажениям входных данных (флуктуациям), причём для распознавания (классификации) объектов целесообразно использовать моменты $M_1 - M_4, M_6$ как наиболее устойчивые к помехам и погрешностям измерений. В то же время при классификации близких по форме объектов лучше применять наиболее чувствительные моменты M_5, M_7 .

3. Эксперименты по распознаванию объектов

В эксперименте были использованы контуры изображений текстовых символов, изображений летательного аппарата и посадочная площадка в форме креста (по тысяче образцов каждого вида). Применялась бюджетная веб-камера Microsoft LifeCam VX-3000 и библиотека для работы с изображениями OpenCV. На рис. 1 слева показаны результаты отдельных этапов работы программы, реализующей алгоритм распознавания посадочного креста в задаче посадки вертолётки по видеоизображению камеры. Кроме этого, проводился эксперимент, связанный с поиском в режиме реального времени самолётов определённого типа на снимке, сделанном с космического аппарата (рис. 1, справа). Найденные на видеокадре объекты обведены окружностью.

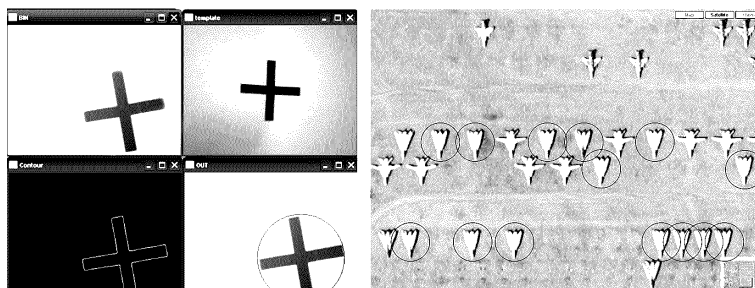


Рис. 1. Примеры работы алгоритма распознавания

Некоторые числовые показатели эксперимента представлены в табл. 3.

Таблица 3

Точность и полнота распознавания объектов

| Тип объекта | Точность, % | Полнота, % |
|---------------------|-------------|------------|
| Элементы текста | 92,2 | 93,7 |
| Летательный аппарат | 97,5 | 95,7 |
| Посадочная площадка | 98,2 | 94,3 |

Из табл. 3 видно, что метод инвариантов обеспечивает приемлемую точность и полноту распознавания различных объектов. Алгоритм успешно распознал посадочный крест, который является композицией сдвига, поворота и масштабирования относительно эталонного изображения, и самолёты заданного типа. Достоинства алгоритма: использование только одной камеры, инвариантность к повороту,

сдвигу и масштабу объекта в кадре, высокая скорость работы, отсутствие этапа обучения (возможность в процессе работы задавать разные объекты для поиска).

Наличие разной чувствительности обуславливает возможность оптимизации выбора состава моментов для решения задачи. В табл. 4 представлены результаты экспериментов по распознаванию летательных аппаратов в зависимости от набора инвариантных моментов. При этом определяется расстояние (мера сходства) между эталонным объектом и исследуемым образцом $I(A,B)$, а также величина максимального отклонения вычисленного расстояния I от его среднего значения на множестве кадров для разных конфигураций сцены (расположение объекта относительно видеокамеры, масштаб).

Таблица 4

Проверка качества распознавания

| Набор моментов | Расстояние $I(A,B)$ | Отклонение |
|------------------|---------------------|------------|
| 2, 3, 4, 6 | $0,04 \pm 0,01$ | 25% |
| 2, 3, 4, 5, 6, 7 | 3 ± 1 | 33% |

Из таблицы следует, что если использовать не все, а выборочно моменты (2, 3, 4, 6), то результаты распознавания исследуемых в эксперименте объектов становятся лучше.

Таким образом, предлагаемый метод учёта менее чувствительных инвариантных моментов позволяет, во-первых, сократить время вычисления расстояния $I(A,B)$ (что, в свою очередь, ускоряет работу всего алгоритма распознавания), и, во-вторых, снизить погрешность вычислений, которая возникает при изменении параметров объекта/сцены в кадре.

4. Объединение метода инвариантов и искусственной нейронной сети

Для обработки изображений зачастую используют искусственные нейронные сети (ИНС). Недостатки алгоритма распознавания на основе метода инвариантных моментов с предложенной мерой сходства: качество распознавания объектов уступает алгоритму с ИНС; алгоритм на основе инвариантных моментов неудовлетворительно справляется с сильным зашумлением входного изображения и наличием большого количества объектов на сцене. Преимущества алгоритма на основе инвариантных моментов перед алгоритмом, использующим ИНС: отсутствие процесса обучения и необходимости переобучения сети; низкие требования к вычислительной мощности, малое количество анализируемых признаков (инвариантов).

С целью улучшения качества распознавания по комплексу показателей (скорость, точность, полнота) в работе [7] предлагается интеграция метода инвариантных моментов с вероятностной нейронной сетью, в основу построения которой положен классификатор Байеса. Пример сети с четырьмя входами представлен на рис. 2.

Вероятностные сети учитывают то обстоятельство, что параметры эталона (элементы входного вектора) сами носят, как правило, вероятностный характер. Функция активации j -го нейрона для слоя образцов:

$$Q_j = \exp \left(\frac{- \sum_{i=1}^4 (w_{ij} - x_i)^2}{\sigma_j^2} \right),$$

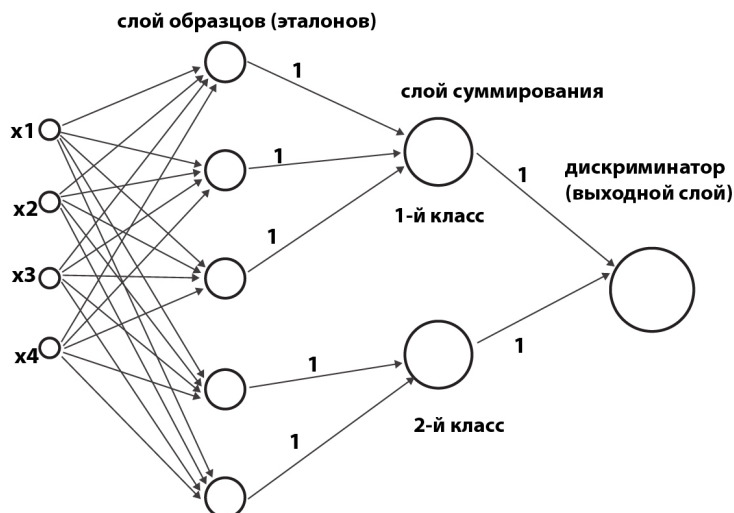


Рис. 2. Пример схемы для двух классов и четырёх входных признаков

где σ_j — среднее квадратическое отклонение j -го эталона; w_{ij} — значение весового коэффициента, соответствующее математическому ожиданию i -го входа j -го эталона, x_i — текущее значение i -го входа. Нулевой (фиктивный) слой сети служит для распределения входных данных между нейронами первого слоя нейронов. При объединении с методом инвариантов на вход системы подаются инварианты неизвестного входного образца. Функцией активации на слое суммирования является «линейный скачок». Последний слой содержит один нейрон (дискриминатор), который сравнивает выходы нейронов второго слоя и выбирает выход с максимальным сигналом.

Объединённый подход «инварианты—нейронная сеть» позволяет существенно сократить размерность пространства признаков за счёт метода инвариантов при сохранении преимуществ нейронной сети.

Заключение

В статье исследованы свойства инвариантных моментов, в частности, определена их теоретическая чувствительность к флуктуациям входных данных. Неравнозначность инвариантов по этому параметру позволяет отбрасывать часть инвариантов при общем усилении распознающих свойств в задаче анализа графических образов. Приведённые результаты экспериментального определения качества распознавания показывают целесообразность учёта разной чувствительности моментов. Кроме того, предложено объединить метод инвариантных моментов с вероятностной нейронной сетью, что позволяет существенно сократить объём входных данных и одновременно с этим построить метод распознавания инвариантный к аффинным преобразованиям.

Литература

1. Роджерс Д., Адамс Д. Математические основы машинной графики. — М.: Машиностроение, 1980. — 240 с. [Rogers D., Adams J. Mathematical Foundations of Computer Graphics. — Moscow: Engineering, 1980. — 240 p. — (in russian).]
2. Dudani S. A., Breeding K. J., McChee R. B. Aircraft Identification by Moment Invariants // IEEE Transaction on Computers. — 1977. — No 1. — Pp. 39–45.
3. Wong R. Y. Scene Matching with Invariant Moments // Computer Graphics and Image Processing. — 1978. — No 8. — Pp. 16–24.

4. *Hu M. K.* Visual Pattern Recognition by Moment Invariants // IRE Transactions on Information Theory. — 1962. — No 8. — Pp. 179–187.
5. *Абрамов Н. С., Фраленко В. П.* Определение расстояний на основе системы технического зрения и метода инвариантных моментов // Информационные технологии и вычислительные системы. — 2012. — № 4. — С. 32–39. [Abramov N. S., Fralenko V. P. Determination of Distances Based on the Computer Vision System and the Method of Invariant Moments // Information Technologies and Computing. — 2012. — No 4. — Pp. 32–39. — (in russian).]
6. *Пережигин А. А., Хачумов В. М.* Обнаружение и автоматическое определение параметров летательного объекта на видео потоке // Информационные технологии и вычислительные системы. — 2005. — № 1. — С. 38–48. [Perezhigin A. A., Khachumov V. M. Detection and Automatic Determination of the Parameters of the Aircraft on the Video Stream // Information Technologies and Computer Systems. — 2005. — No 1. — Pp. 38–48. — (in russian).]
7. *Хачумов В. М.* О расширении функциональных возможностей искусственных нейронных сетей // Авиакосмическое приборостроение. — 2008. — № 5. — С. 53–59. [Khachumov V. M. The Extension of the Functionality of Artificial Neural Networks // Aerospace Instrument. — 2008. — No 5. — P. 53–59. — (in russian).]

UDC 004.89:004.932

Object Recognition Based on Invariant Moments

N. S. Abramov*, V. M. Khachumov[†]

* *Ailamazyan Program Systems Institute of the Russian Academy of Sciences
4a, Peter I, Veskovo, Pereslavl-Zalessky Yaroslavl Region, Russia, 152021*

[†] *Institute for Systems Analysis of the Russian Academy of Sciences
9, pr. 60-letiya Otkryabrya, Moscow, Russia, 117312*

We investigate the properties of invariant moments of binary images, it is necessary for the formation of their set in order to recognize graphic images. Pattern recognition and distance measurement used Hu invariant moments. It is shown that the invariants have different sensitivity to changes in input data that defines the strategy of their choice. Experiments on pattern recognition of text characters, images of the aircrafts and the landing site in the form of a cross were carried out. The considered recognition algorithm works in real time, use only one camera, is invariant to rotation, shift and scale the object in the frame. The accuracy and completeness of recognition amounted to about 92% on a set of thousands of samples of each type. The results of the experimental determination of the quality of recognition of various objects based on their contour images, as well as the results of comparing recognition using a different set of invariant moments are presented. It is shown that the inclusion of the less sensitive invariant moments reduces the computation time, and lowers the computational error that occurs when fluctuations in the parameters of an object or scene in the frame take place. It is proposed to combine the method of invariant moments with probabilistic neural network, which will improve the quality of recognition, making it more fast, accurate and complete.

Key words and phrases: invariants, recognition, neural network.