

# Определение расстояний на основе системы технического зрения и метода инвариантных моментов<sup>1</sup>

Н.С. Абрамов, В.П. Фраленко

**Аннотация.** В работе рассматривается методика вычисления расстояний и распознавания объектов на основе системы технического зрения и метода инвариантных моментов. Предложена новая нормализация центральных моментов. Проведены эксперименты и даны погрешности вычислений.

**Ключевые слова:** определение расстояний, распознавание, инвариантные моменты Ху, контур, оценка погрешности, эксперимент.

## 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. В данной работе исследуется 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 с анализом инвариантов, подобные универсальные системы, способные одновременно решать задачи распознавания и определения расстояния на единых методологических принципах, в настоящее время отсутствуют, хотя потребность в них достаточно велика.

## 1. Определение расстояния до объекта на основе инвариантов для бинарных изображений

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 (Рис. 1) перечислением своих  $N$  точек  $(x_i, y_i)$ .

Алгоритм получения инвариантов содержит следующие этапы:

1) определение центральных моментов порядка не выше третьего:

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

и  $\bar{y}$  – средние значения,

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

<sup>1</sup> Работа выполнена при поддержке РФФИ (проект № 12-07-00205-а), Программы № 18 Президиума РАН и Программы ОНИТ РАН (проект 2.10).

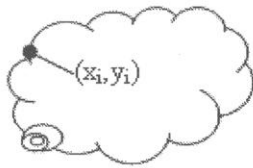


Рис. 1. Контур объекта

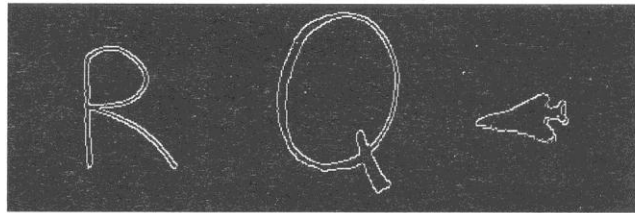


Рис. 2. Контур изображения объектов

2) получение моментов, инвариантных к операциям поворота, переноса и зеркального отображения [1]:

$$\begin{aligned}
 M_1 &= m_{20} + m_{02} \\
 M_2 &= (m_{20} - m_{02})^2 + 4m_{11}^2 \\
 M_3 &= (m_{30} - 3m_{12})^2 + (3m_{21} - m_{03})^2 \\
 M_4 &= (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2 \\
 M_5 &= (m_{30} - 3m_{12})(m_{30} + m_{12}) \times \\
 &\quad \times [(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] + \\
 &\quad + (3m_{21} - m_{03})(m_{21} + m_{03}) \times \\
 &\quad \times [3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] \\
 M_6 &= (m_{20} - m_{02})[(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] + \\
 &\quad + 4m_{11}(m_{30} + m_{12})(m_{21} + m_{03}) \\
 M_7 &= (3m_{21} - m_{03})(m_{30} + m_{12}) \times \\
 &\quad [(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] - \\
 &\quad - (m_{30} - 3m_{12})(m_{21} + m_{03}) \times \\
 &\quad \times [3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2];
 \end{aligned} \tag{1}$$

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

$$\begin{aligned}
 M'_1 &= r \cdot h, \quad M'_2 = \frac{M_2}{r^4}, \quad M'_3 = \frac{M_3}{r^6}, \quad M'_4 = \frac{M_4}{r^6}, \\
 M'_5 &= \frac{M_5}{r^{12}}, \quad M'_6 = \frac{M_6}{r^8}, \quad M'_7 = \frac{M_7}{r^{12}}.
 \end{aligned} \tag{2}$$

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 [2]:

$$M'_1 = r \cdot h = \text{const.} \tag{3}$$

Для распознавания объектов можно использовать любые инварианты (2) кроме первого. Инвариант (3) теоретически является константой для используемой камеры и наблюдаемого объекта, поэтому с его помощью по известному значению  $r$  можно определить расстояние до объекта.

Для проведения эксперимента с целью проверки практической пригодности формулы (3) применялась малобюджетная web-камера Life-Cam VX-3000 [5]. Объектами служили изображения латинских букв «R» и «Q» и контур самолета. Все три изображения были представлены на матовом листе бумаги, который перемещался относительно камеры, обеспечивая тем самым изменение масштаба. Инварианты вычислялись для контуров объектов (Рис. 2).

Эксперимент включал следующую последовательность действий:

1. Устанавливается начальное расстояние камеры до объекта ( $h_0 - 14$  см).
2. Осуществляется равномерное движение камеры к объекту. Расстояние от начального положения до конечного ( $h_1 - 7-8$  см) разбивается на  $N$  ( $N=90$ ) фиксированных фаз, которым соответствуют определенные кадры изображений.
3. Обеспечивается захват видеопотока и обработка каждого кадра: с применением библиотеки OpenCV [6] находится контур изображения (Рис. 2) и вычисляется величина  $r$ .

4. Рассчитывается текущее расстояние по формуле:

$$h_i = \frac{h_1 - h_0}{N}.$$

5. Проверяется утверждение (3).

На Рис. 3 показана экспериментальная зависимость  $r = \sqrt{M_1} = \varphi(h)$  для исследуемых объектов.

Видно, что значение функции  $r$  почти линейно уменьшается с увеличением расстояния  $h$ . Небольшие всплески функции интерпретируются как помехи на видеокдрах, возникающие из-за ошибок эксперимента.

На Рис. 4 показаны значения  $M'_1 = \sqrt{M_1}h$ , которые теоретически должны быть константами для каждого наблюдаемого объекта.

Видно, что экспериментальные значения величины  $M_1'$  довольно стабильны на разных кадрах. В каждом тесте отклонение не превышало 9%, причем максимальные отклонения пришлось на начало и конец, а наименьшие – на середину эксперимента. Это объясняется тем, что тестирование выполнялось вручную одним человеком, для которого характерны дрожание рук, запаздывание нажатия клавиш (команд) и другие недостатки. Погрешность измерения константы объясняется следующими факторами: неточностью измерения расстояния от объектива

камеры до объекта, плохим качеством получаемого изображения из-за ограничений камеры.

Из экспериментов следует, что вычисление моментов (2) выполняется с отклонениями, которые можно нивелировать хорошей аппаратурой, автоматизацией тестирования, исключающей человеческий фактор, и постобработкой. Таким образом, имея заранее вычисленную «псевдоконстанту»  $M_1'$  для каждого объекта, можно вычислять в каждый момент времени величину  $r$  и расстояние  $h$ .

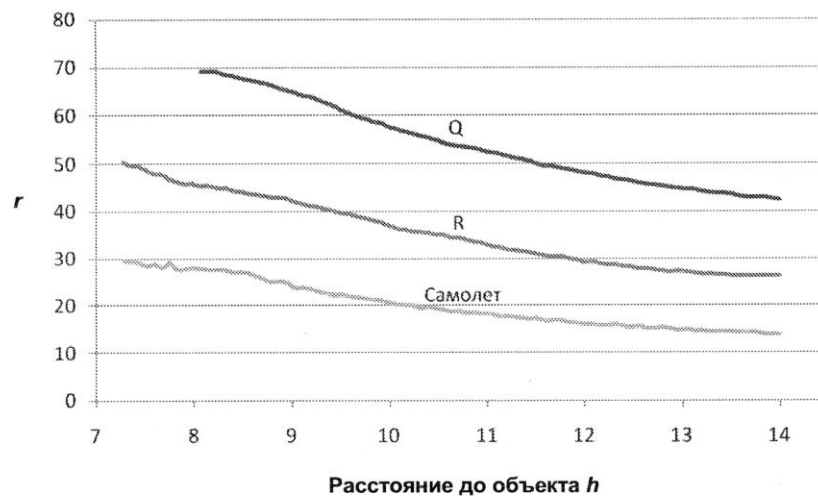


Рис. 3. Зависимость  $\sqrt{M_1} = \varphi(h)$

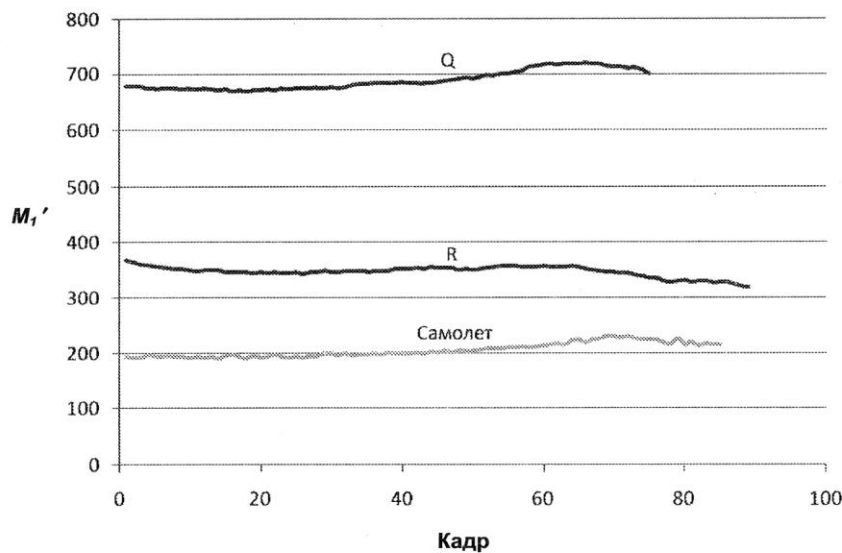


Рис. 4. Значения  $M_1' = \sqrt{M_1}h$

## 2. Оценка погрешности измерения инвариантов

Пусть имеется полутоновое изображение с функцией яркости  $f(x, y)$ , где  $(x, y)$  – координаты пикселя. Область значений функции лежит в интервале  $[0,1]$ . Тогда центральные моменты для цифрового изображения

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y),$$

где  $\bar{x}, \bar{y}$  – центр тяжести.

Центральные моменты порядка три или меньше [3]:

$$\begin{aligned} \mu_{00} &= m_{00} \\ \mu_{01} &= 0 \\ \mu_{10} &= 0 \\ \mu_{20} &= m_{20} - \bar{x} m_{10} \\ \mu_{02} &= m_{02} - \bar{y} m_{01} \\ \mu_{11} &= m_{11} - \bar{y} m_{10} \\ \mu_{30} &= m_{30} - 3\bar{x} m_{20} + 2\bar{x}^2 m_{10} \\ \mu_{12} &= m_{12} - 2\bar{y} m_{11} - \bar{x} m_{02} + 2\bar{x}\bar{y} m_{01} \\ \mu_{21} &= m_{21} - 2\bar{x} m_{11} - \bar{y} m_{20} + 2\bar{x}\bar{y} m_{01} \\ \mu_{03} &= m_{03} - 3\bar{y} m_{02} + 2\bar{y}^2 m_{01} \end{aligned} \quad (4)$$

Нормализованные центральные моменты, инвариантные к масштабу:

$$m_{pq} = \mu_{pq} / \mu_{00}^{(p+q)/2}. \quad (5)$$

Найденные центральные моменты  $m_{pq}$  позволяют найти инвариантные моменты по формулам (1).

Рассмотрим практический пример анализа достаточно простого изображения буквы «R», повернутой на разные углы и с разными масштабирующими коэффициентами. Заметим, что на реальных изображениях нормализация центральных моментов вида (5) для  $(p+q) > 1$  дает плохие результаты, что наглядно показано в работе [3]. Для высокой устойчивости к изменениям масштаба используется нормализация

$$m_{pq} = \mu_{pq} / \mu_{00}^{(p+q)/2+1}. \quad (6)$$

Моменты  $M_i$  с нормализацией (6) применяются в библиотеке [6] для анализа контуров. Однако оказывается, что с этой нормализацией  $M_i$  перестает реагировать на изменение масштаба, что не позволяет определять расстояния до целевых объектов.

В настоящей работе предлагается использовать нормализацию центрального момента (5) и следующие инварианты:

$$\begin{aligned} M'_1 &= r \cdot h, \quad M'_2 = \frac{M_2}{r^4}, \quad M'_3 = \frac{M_3}{r^4}, \quad M'_4 = \frac{M_4}{r^4}, \\ M'_5 &= \frac{M_5}{r^8}, \quad M'_6 = \frac{M_6}{r^6}, \quad M'_7 = \frac{M_7}{r^8}. \end{aligned} \quad (7)$$

В Табл. 1 приведены полученные значения инвариантов и оценки отклонений. За «идеальные» значения приняты величины инвариантов  $x_{ideal}$ , полученные для изображения с масштабирующим коэффициентом 1.0 (первая строка таблицы). Применялись два вида нормализации (2) и (7) и вычислялось относительное отклонение в процентах для каждого измерения:

$$\delta = 100 \cdot \left| \frac{x_{ideal} - x_{real}}{x_{ideal}} \right|,$$

где  $x_{real}$  – величина инварианта, полученного в эксперименте.

Из Табл.1 следует, что отклонения при аффинных преобразованиях в случае нормализаций (2) и (7) совпадают в пределах погрешности.

Проведем два эксперимента по измерению расстояния с помощью первого инварианта  $M_1$  для бинарных и полутоновых изображений соответственно (Рис. 5 и Рис. 6).

Экспериментальные данные представлены в Табл. 2 и в Табл. 3, где вычисленный масштабирующий коэффициент для  $i$ -го эксперимента получен делением  $r_i$  на  $r_{10}$  (в десятом эксперименте реальный масштабирующий коэффициент равен 1.0).

Рис. 7 и Рис. 8 отражают изменение оценки величины  $M'_1$  при разных масштабирующих коэффициентах. Оценка получается делением  $r_i$  на реальный масштабирующий коэффициент  $i$ -го эксперимента. Рисунки показывают, что

погрешность измерения  $M_1'$  существенно возрастает при масштабирующем коэффициенте менее 0.4. Однако эти изменения не оказывают

критичными для точности определения реального масштабирующего коэффициента, что подтверждается Табл. 2 и Табл. 3.

Табл. 1. Определение инвариантов для изображений буквы «R»







Изображение	№	Полученное значение инв. Ху с нормализацией (2)	Отклонение $\delta$ , %	Полученное значение инв. Ху с нормализацией (7)	Отклонение $\delta$ , %
 Масштабир. коэффициент 1.0	1	6.459095e-01	0.000	3.773190e+02	0.000
	2	8.600896e-03	0.000	8.600896e-03	0.000
	3	2.531461e-02	0.000	1.056123e-02	0.000
	4	1.314034e-02	0.000	5.482136e-03	0.000
	5	-5.362404e-05	0.000	-9.333536e-06	0.000
	6	1.218637e-03	0.000	5.084141e-04	0.000
	7	-2.335836e-04	0.000	-4.065642e-05	0.000
 Поворот на 45 градусов	1	6.459014e-01	0.001	3.773623e+02	0.011
	2	8.701333e-03	1.168	8.701333e-03	1.168
	3	2.532890e-02	0.056	1.056693e-02	0.054
	4	1.318073e-02	0.307	5.498852e-03	0.305
	5	-5.635347e-05	5.090	-9.808116e-06	5.085
	6	1.229409e-03	0.884	5.128955e-04	0.881
	7	-2.341479e-04	0.242	-4.075259e-05	0.237
 Поворот на 135 градусов	1	6.459843e-01	0.012	3.773068e+02	0.003
	2	8.614365e-03	0.157	8.614365e-03	0.157
	3	2.535300e-02	0.152	1.057970e-02	0.175
	4	1.313217e-02	0.062	5.479998e-03	0.039
	5	-5.504314e-05	2.646	-9.584978e-06	2.694
	6	1.218751e-03	0.009	5.085794e-04	0.033
	7	-2.332102e-04	0.160	-4.061023e-05	0.114
 Сдвиг	1	6.459095e-01	0.000	3.773190e+02	0.000
	2	8.600896e-03	0.000	8.600896e-03	0.000
	3	2.531461e-02	0.000	1.056123e-02	0.000
	4	1.314034e-02	0.000	5.482136e-03	0.000
	5	-5.362404e-05	0.000	-9.333536e-06	0.000
	6	1.218637e-03	0.000	5.084141e-04	0.000
	7	-2.335836e-04	0.000	-4.065642e-05	0.000
 Масштабир. коэффициент 0.5	1	6.451008e-01	0.125	3.774628e+02	0.038
	2	8.458464e-03	1.656	8.458464e-03	1.656
	3	2.526833e-02	0.183	1.051554e-02	0.433
	4	1.313848e-02	0.014	5.467646e-03	0.264
	5	-4.906593e-05	8.500	-8.497485e-06	8.957
	6	1.208232e-03	0.854	5.028119e-04	1.102
	7	-2.343078e-04	0.310	-4.057860e-05	0.191
 Масштабир. коэффициент 1.4	1	6.458462e-01	0.010	3.770741e+02	0.065
	2	8.551000e-03	0.580	8.551000e-03	0.580
	3	2.529550e-02	0.075	1.055119e-02	0.095
	4	1.311735e-02	0.174	5.471474e-03	0.194
	5	-5.348968e-05	0.251	-9.306503e-06	0.290
	6	1.212968e-03	0.465	5.059499e-04	0.485
	7	-2.328770e-04	0.303	-4.051754e-05	0.342



Рис. 5. Бинарное изображение для экспериментов (880x1000 для масштабирующего коэффициента 1.0)

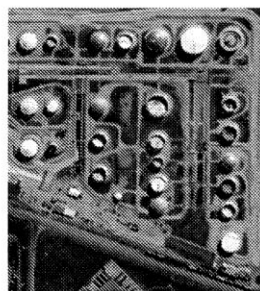


Рис. 6. Полутоновое изображение для экспериментов (880x1000 для масштабирующего коэффициента 1.0)

Табл. 2. Эксперимент с оценкой для бинарных изображений

№ эксперимента, $i$	Масштабирующий коэффициент	$r_i$	Вычисленный масштабирующий коэффициент	Оценка $M'_i$
1	0.1	3.789206e+01	0.1004244684	3.789206e+02
2	0.2	7.556593e+01	0.2002706729	3.778297e+02
3	0.3	1.131250e+02	0.2998126254	3.770833e+02
4	0.4	1.510832e+02	0.4004123832	3.77080e+02
5	0.5	1.886205e+02	0.4998966393	3.772410e+02
6	0.6	2.263941e+02	0.6000071559	3.773235e+02
7	0.7	2.643648e+02	0.7006400421	3.776640e+02
8	0.8	3.019051e+02	0.8001322490	3.773814e+02
9	0.9	3.395968e+02	0.9000257079	3.773298e+02
10	1.0	3.773190e+02	1.0	3.773190e+02
11	2.0	7.546383e+02	2.000000795	3.773192e+02
12	3.0	1.131957e+03	3.000000001	3.773190e+02
13	4.0	1.509277e+03	4.000002651	3.773193e+02
14	5.0	1.886596e+03	5.000002651	3.773192e+02
15	6.0	2.263915e+03	6.000002651	3.773192e+02
16	7.0	2.641234e+03	7.000002652	3.773191e+02
17	8.0	3.018553e+03	8.000002652	3.773191e+02
18	9.0	3.395873e+03	9.000005302	3.773192e+02
19	10.0	3.773192e+03	10.00000530	3.773192e+02

Табл. 3. Эксперимент с оценкой  $M'_i$  для полутоновых изображений

№ эксперимента, $i$	Масштабирующий коэффициент	$r_i$	Вычисленный масштабирующий коэффициент	Оценка $M'_i$
1	0.1	3.840157e+01	0.1000210192	3.840157e+02
2	0.2	7.678956e+01	0.2000066678	3.839478e+02
3	0.3	1.151949e+02	0.3000375064	3.839830e+02
4	0.4	1.535813e+02	0.4000190137	3.839533e+02
5	0.5	1.919645e+02	0.4999921862	3.839290e+02
6	0.6	2.303722e+02	0.6000291717	3.839537e+02
7	0.7	2.687581e+02	0.7000093767	3.839401e+02
8	0.8	3.071630e+02	0.8000390692	3.839538e+02
9	0.9	3.455491e+02	0.9000197951	3.839434e+02
10	1.0	3.839350e+02	1.0	3.839350e+02
11	2.0	7.678188e+02	1.999866644	3.839094e+02
12	3.0	1.151738e+03	2.999825492	3.839127e+02
13	4.0	1.535636e+03	3.999729121	3.839090e+02
14	5.0	1.919536e+03	4.999637960	3.839072e+02
15	6.0	2.303434e+03	5.999541590	3.839057e+02
16	7.0	2.687332e+03	6.999445219	3.839046e+02
17	8.0	3.071226e+03	7.999338430	3.839033e+02
18	9.0	3.455127e+03	8.999249874	3.839030e+02
19	10.0	3.839021e+03	9.999143085	3.839021e+02

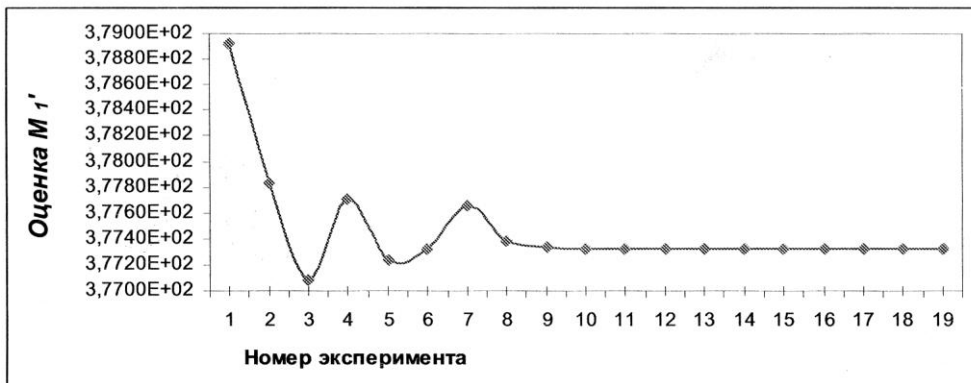


Рис. 7. Оценка  $M_i'$  на бинарных изображениях для разных масштабирующих коэффициентов



Рис. 8. Оценка  $M_i'$  на полутоновых изображениях для разных масштабирующих коэффициентов

### 3. Эксперимент по распознаванию объектов с помощью инвариантов

В качестве распознаваемых с помощью инвариантов  $X_u$  объектов в эксперименте были использованы контуры изображений латинских букв «R» и «Q» и контуры схематичного изображения самолета. База снимков содержала по тысяче контуров объектов каждого класса. Образцы были получены путем фотографирования объектов под разными ракурсами и с разных расстояний с помощью web-камеры. Расстояние между парой объектов  $A$  и  $B$  вычислялось по формуле:

$$I(A, B) = \sum_{i=1}^7 \left| \frac{m_i(A) - m_i(B)}{m_i(A)} \right|,$$

где  $m_i(A) = \text{sign}(M_i'(A)) \cdot \log(M_i'(A))$ ,  
 $m_i(B) = \text{sign}(M_i'(B)) \cdot \log(M_i'(B))$  (натуральный логарифм). Результаты эксперимента представлены в Табл. 4.

Табл. 4. Точность и полнота распознавания объектов

Тип объекта	Точность, %	Полнота, %
Буква «R»	90.0	90.0
Буква «Q»	90.9	100.0
Самолет	100.0	90.0

Можно отметить, что точность и полнота распознавания тестовых объектов достаточно низкая, однако ошибки могут быть скомпенсированы тем, что в реальной ситуации камера делает несколько десятков кадров в секунду. Видеоряд дает дополнительную информацию и правильное распознавание обеспечивается сопоставлением нескольких результатов обработки одного и того же объекта в кадрах.

### Заключение

В статье рассмотрен метод измерения расстояний до целевых объектов с помощью системы технического зрения на основе инвариантных моментов. Предложена новая нормализация центральных моментов. Даны оценки точности

измерения расстояния и распознавания объектов. Технология определения расстояний включает реализацию нескольких этапов:

1. Формирование базы знаний: установление коэффициента подобия и константы  $M_1'$  как одного из параметров камеры для наблюдаемого (известного) объекта, построение зависимости  $r = \sqrt{M_1} = \varphi(h)$ . Эта задача сопряжена с проведением соответствующих предварительных экспериментов над камерой.

2. Нахождение (распознавание) опорного объекта на снимке (видеокадре) на основе инвариантных моментов или других подходов, извлечение из базы знаний реальных (мировых) координат  $(x, y)$ , установление экранных  $(x^*, y^*)$  координат, соответствующих наблюдаемому объекту.

3. Вычисление характерного значения  $r$  по видеокадру опорного изображения.

4. Определение расстояний до объекта с использованием коэффициента пропорциональности на основе значения первого инвариантного момента. Усреднение расстояний с целью повышения точности измерений.

Описанный метод может практически применяться в системах технического зрения

беспилотных летательных аппаратов при наличии фото-видео-камеры с разрешающей способностью не менее 1280x720. Полученные результаты использованы при разработке программного комплекса автономного позиционирования беспилотного летательного аппарата на основе системы технического зрения (ПК «Навигатор-М») [7].

## Литература

1. Роджерс Д., Адамс Дж. Математические основы машинной графики. – М.: Машиностроение, 1980. – 240 с.
2. Dudani S. A., Breeding K. J., McChee R. B. Aircraft identification by moment invariants. – IEEE transaction on computers, vol. C26, No 1, 1977, pp.39-45.
3. Wong R. Y. Scene matching with invariant moments. – Computer Graphics and image processing, 1978, vol.8, pp.16-24.
4. Hu M. K. Visual pattern recognition by moment invariants. – IRE Transactions on Information Theory 8, 1962, pp. 179-187.
5. LifeCam VX-3000. URL: <http://www.microsoft.com/hardware/ru-ru/d/lifecam-vx-3000> (дата обращения 16.09.2012).
6. Библиотека OpenCV. URL: <http://opencv.willowgarage.com> (дата обращения: 16.09.2012).
7. Свидетельство о государственной регистрации программы для ЭВМ № 2012618314.

**Абрамов Николай Сергеевич.** Инженер-исследователь, аспирант Института программных систем им. А.К. Айламазяна РАН. Окончил Институт программных систем им. А.К. Айламазяна в 2010 году. Автор трех печатных работ. Область научных интересов: методы обработки изображений, системы компьютерного зрения, искусственный интеллект.  
E-mail: n-say@nsa.pereslavl.ru, nsay@joker.botik.ru

**Фраленко Виталий Петрович.** Инженер-исследователь Института программных систем им. А.К. Айламазяна РАН. Окончил Институт программных систем им. А.К. Айламазяна в 2007 году. Кандидат технических наук. Автор 41 печатной работы. Область научных интересов: искусственный интеллект, нейронные сети, шифрование, сжатие.  
E-mail: alarmod@pereslavl.ru