

УДК 621.865.8

МЕТОДЫ КОРРЕКЦИИ ПРОСТРАНСТВЕННЫХ ИСКАЖЕНИЙ ИЗОБРАЖЕНИЙ ПЛОСКИХ ОБЪЕКТОВ В УСЛОВИЯХ ДЕЙСТВИЯ ПОЛНОЙ АФФИННОЙ ГРУППЫ ПРЕОБРАЗОВАНИЙ

Н. В. Соловьев,

старший преподаватель

Санкт-Петербургский государственный университет аэрокосмического приборостроения (ГУАП)

Рассмотрены применяемые в системах распознавания робототехнических комплексов методы коррекции пространственных искажений изображений, описываемых полной аффинной группой преобразований. Основное внимание уделено методам, позволяющим распознавать объекты в процессе компенсации искажений. Предложен метод определения параметров преобразования, основанный на последовательном переборе ограниченного числа характерных точек предварительно центрированного изображения объекта.

The methods of a correction of spatial aliasings of images are considered. The distortions are circumscribed by full affine group of transformations. The methods allow to recognize objects during indemnification of distortions. Area of application algorithm – robot vision.

Введение

Задача распознавания объектов и определения параметров их положения в пространстве по изображениям возникает при разработке адаптивных промышленных роботов и робототехнических комплексов различного назначения. Подобные устройства состоят из системы распознавания, системы управления, приводов и исполнительных механизмов. Система распознавания робота после получения, обработки и анализа сенсорной информации поставляет системе управления данные об окружающей среде, на основании анализа которых последняя вырабатывает управляющие сигналы для приводов исполнительной системы. Можно сказать, что именно система распознавания, являясь непременной составляющей адаптивного робота (робототехнического комплекса), превращает программно управляемое устройство в робота, совершающего целесообразные действия в условиях изменяющейся внешней среды.

Следует отметить, что первые попытки решения проблемы машинного зрения, создания систем технического зрения и компьютерного анализа изображений сцен относятся к концу 60-х годов XX века, что связано с появлением возможности ввода в ЭВМ изображений. За прошедшее время появилось огромное количество литературы по проблемам технического зрения вообще и роботов в частности, как фундаментального характера, так и посвященных решению отдельных задач. Выпущено множество тематических сборников, обзоров, материалов конференций, посвященных различным аспектам технического зрения и смежным областям. Однако, несмотря на достигнутые успехи в решении отдельных проблем, системы технического зрения по своей универсальности все еще далеки от аналогичных систем в живой природе [1].

Основные трудности в создании систем распознавания объектов по их изображениям связаны с тем, что разные изображения одного и того же объекта в подавляющем большинстве случаев существенно отличаются друг от друга. Данное обстоятельство практически не позволяет подобрать для проведения классификации распознаваемого образа признаки, инвариантные к этим отличиям. Причины, вызывающие указанные отличия (или искажения, если считать одно из изображений объекта эталонным), можно разделить на три группы:

1) помехи, обусловленные несовершенством аппаратуры для получения, передачи, хранения и обработки изображений;

2) изменение с течением времени освещенности объектов сцены, изображение которой анализируется;

3) изменение взаимного положения объектов сцены относительно друг друга и устройства получения изображения (камеры).

Методы компенсации технических помех хорошо разработаны. Применение тестовых изображений и методов статистической обработки позволяет в большинстве случаев добиться вполне приемлемых результатов.

Компенсация искажений, вызванных второй причиной, более трудна, так как изменение освещенности приводит не только к глобальным изменениям яркости изображения, но и к изменению яркости изображений отдельных объектов, появлению и смещению бликов и теней. Однако применительно к промышленным роботам трудности компенсации данной причины можно существенно снизить, подобрав соответствующее освещение сцены, например, используя бестеневые матовые светильники для освещения объектов сцены.

По признанию большинства авторов [2, 3], основные сложности вызывают именно пространственные искажения. Действительно, даже при отсутствии технических помех и постоянном освещении сцены любое изменение взаимного расположения объектов и камеры приводит к изменению изображения.

Известно [2], что если в результате перемещения объектов или камеры не происходит перекрытие их изображений при изменении изображения сцены, то это изменение может быть описано групповым преобразованием. При центральном проецировании, когда расстояние от объектов до камеры сравнимо с фокусным расстоянием объектива камеры, преобразование описывается проективной группой. При параллельном проецировании, когда объекты находятся на значительном удалении от камеры, преобразование описывается аффинной группой, являющейся подгруппой проективной группы преобразований.

Постановка задачи

Пусть в пространстве U действует транзитивная группа преобразований G , каждый элемент $g \in G$ которой может быть представлен в виде произведения некоторых элементов ее подгрупп G_1, G_2, \dots, G_n , т. е. $g = g_1 * g_2 * \dots * g_n; g_i \in G_i$.

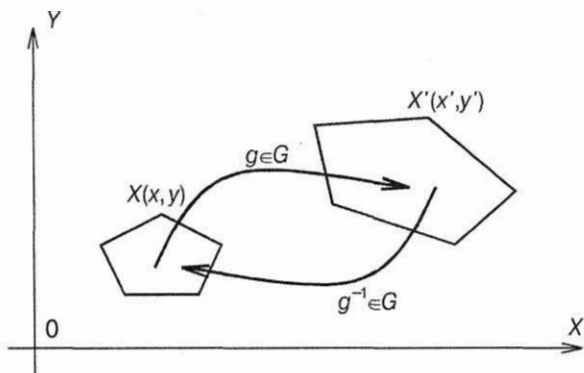
Тогда каждая точка X с координатами (x, y) исходного изображения (эталонного образа) преобразуется элементом g :

$$X' = g * X = g_1 * g_2 * \dots * g_n * X,$$

где $X':(x', y')$ – координаты соответствующей точки преобразованного или искаженного изображения. Таким образом, при наличии только пространственных искажений яркость точки X' искаженного изображения полагается равной яркости точки X исходного изображения (рис. 1).

Если из каких-либо соображений будет найден элемент g_i (или g_i^{-1}), тогда, подвергнув искаженное изображение преобразованию элементом g_i^{-1} , мы обеспечим «сближение» искаженного и эталонного изображений. Продолжая процесс последовательной нейтрализации подгрупп можно либо полностью устранить искажения (распознать все объекты и определить их параметры положения), либо существенно упростить задачу и свести ее к известному решению.

Аффинная группа имеет вид:



■ Рис. 1. Представление пространственного искажения групповым преобразованием

$$\begin{aligned} x' &= a_1x + a_2y + b_1, \\ y' &= a_3x + a_4y + b_2, \end{aligned} \tag{1}$$

где x, y, x', y' – исходные и преобразованные аффинной группой координаты точки изображения; $a_1, \dots, a_4, b_1, b_2$ – параметры аффинной группы. Очевидно, что, зная параметры, можно легко осуществить обратное преобразование $X = g^{-1} * X'$, тем самым приведя изображение к эталонному виду. Дальнейшее распознавание не представляет сложностей и может быть выполнено, например, методом маски.

Из уравнения (1) видно, что для нахождения шести параметров аффинной группы необходимо знать как минимум координаты трех не лежащих на одной прямой точек на исходном и искаженном изображениях и соответствие между ними. Тогда, решив легко получаемую из (1) систему шести уравнений с параметрами $a_1, \dots, a_4, b_1, b_2$ в качестве неизвестных и координатами соответствующих точек изображений, можно получить искомое преобразование g^{-1} . Рассмотренные далее методы в той или иной степени используют данный подход.

Метод последовательной компенсации подгрупп аффинной группы

Разработанный И. Л. Ерошем [4] метод позволяет совместить процесс компенсации пространственных искажений с собственно распознаванием, т. е. отнесением объекта к одному из известных эталонов.

Элемент g аффинной группы может быть представлен в виде:

$$g = g_1 * g_2, \tag{2}$$

где $g_1 \in G_1$ – группа сдвигов вдоль координатных осей

вида $\begin{cases} x' = x + b_1 \\ y' = y + b_2 \end{cases}; g_2 \in G_2$ – аффинная группа без сдвигов

с матрицей преобразования вида $\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$.

Если изображения эталонов распознаваемых объектов централизованы (центр формы совпадает с началом координат, что всегда можно сделать на этапе обучения системы автоматического распознавания), то, вычислив центр формы искаженного изображения, получим параметры b_1 и b_2 . После переноса начала координат в этот центр будут нейтрализованы параметры подгруппы G_1 . Следовательно, частично преобразованное, распознаваемое изображение будет связано с неизвестным пока эталоном преобразованием элементом группы G_2 . Этот элемент g_2 может быть представлен в виде

$$g_2 = g_2' * g_2'',$$

где g_2' – элемент группы масштабного преобразования с некоторой матрицей, например, вида $\begin{pmatrix} r & 0 \\ 0 & 1 \end{pmatrix}$;

g_2'' – элемент аффинной унимодулярной группы с матрицей вида $\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix}$,

$$c_1c_4 - c_2c_3 = 1. \quad (3)$$

Значения параметров r , а также c_1, c_2, c_3, c_4 могут быть найдены через параметры аффинной группы без сдвига:

$$rc_1 = a_1, \quad rc_2 = a_2, \quad c_3 = a_3, \quad c_4 = a_4, \quad \frac{a_1a_4 - a_2a_3}{r} = 1,$$

откуда $r = a_1a_4 - a_2a_3$.

Обратная матрица, которая устраняет масштабную

часть искажений, имеет вид $\begin{pmatrix} 1 & 0 \\ a_1a_4 - a_2a_3 & 1 \end{pmatrix}$. Значение r неизвестно, однако его можно определить из

сравнения площадей распознаваемого изображения и изображения эталонов. Для этого все эталонные изображения необходимо отмасштабировать так, чтобы их площади были равны между собой и имели бы величину S_0 . Вычислив величину S распознаваемого изображения, можно найти коэффициент масштаба $r = S/S_0$.

После преобразования распознаваемого изображения элементом вида

$$(g_2)^{-1} = \begin{pmatrix} 1/(S/S_0) & 0 \\ 0 & 1 \end{pmatrix}$$

оно оказывается связанным с одним из эталонов (по-прежнему неизвестным) преобразованием вида (3) или

$$\begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c_3/c_1 & 1 \end{pmatrix} \begin{pmatrix} c_1 & c_2 \\ 0 & 1/c_1 \end{pmatrix}.$$

Отношение c_3/c_1 можно определить следующим образом. Предположим, что известна одна характерная точка на каждом эталоне, и можно определить ее на исследуемом изображении. Под характерной точкой будем понимать такую точку, положение которой определяется однозначно независимо от заданных преобразований (в данном случае – аффинных). К сожалению, общий метод для выбора характерных точек на произвольных изображениях неизвестен. Для многоградационных растровых изображений в качестве характерной точки можно выбрать центр масс или центр яркости изображения, который вычисляется следующим образом:

$$x_c = \frac{1}{N} \sum_{i=1}^N x_i n_i, \quad y_c = \frac{1}{N} \sum_{i=1}^N y_i n_i,$$

где n_i – числовое значение (вес), приписанное пикселу i , например, код яркости пиксела; N – число пикселов изображения. Для таких изображений при переносе начала координат в центр формы центр яркости оказывается «сдвинутым», т. е. может не совпадать с началом координат.

Для изображений каждого из эталонов поместим на этапе обучения центр яркости на ось X (путем поворота вокруг начала координат). Тогда, вычислив центр яркости распознаваемого изображения, получим

$$\begin{pmatrix} x_c \\ y_c \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ c_3 & c_4 \end{pmatrix} \begin{pmatrix} x_c \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 x_c \\ c_3 x_c \end{pmatrix}.$$

Из полученного выражения находим:

$$\frac{c_3}{c_1} = \frac{y_c}{x_c}.$$

Тогда можно вычислить матрицу

$$\begin{pmatrix} 1 & 0 \\ c_3/c_1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -c_1/c_3 & 1 \end{pmatrix}.$$

После преобразования распознаваемого изображения этой матрицей получим, что некоторый неизвестный эталон связан с распознаваемым изображением групповым преобразованием с матрицей

$$\begin{pmatrix} c_1 & c_2 \\ 0 & 1/c_1 \end{pmatrix}. \quad (4)$$

Таким образом, от полной аффинной группы с шестью неизвестными параметрами мы перешли к группе с двумя неизвестными параметрами путем нейтрализации некоторых групповых параметров.

Теперь предположим, что нам известен эталон с номером p , которому соответствует распознаваемое изображение. Его характерная точка имеет координаты $\{x_c^p, 0\}$. После преобразования матрицей (4) координаты этой точки будут иметь значения:

$$\begin{pmatrix} x_c \\ y_c \end{pmatrix} = \begin{pmatrix} c_1 & c_2 \\ 0 & 1/c_1 \end{pmatrix} \begin{pmatrix} x_c^p \\ 0 \end{pmatrix} = \begin{pmatrix} c_1 x_c^p \\ 0 \end{pmatrix},$$

откуда

$$c_1 = \frac{x_c}{x_c^p}. \quad (5)$$

Представим матрицу (4) в следующем виде:

$$\begin{pmatrix} c_1 & c_2 \\ 0 & 1/c_1 \end{pmatrix} = \begin{pmatrix} 1 & c_1 c_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1 & 0 \\ 0 & 1/c_1 \end{pmatrix}.$$

Если известно, какому эталону соответствует распознаваемое изображение, то, преобразуя изображение эталона матрицей вида $\begin{pmatrix} c_1 & 0 \\ 0 & 1/c_1 \end{pmatrix}$, получим изображение, связанное с распознаваемым матрицей вида $\begin{pmatrix} 1 & c_1 c_2 \\ 0 & 1 \end{pmatrix}$, в которой только один параметр c_2 неизвестен.

Для каждого эталона, используя соотношение (5), найдем значения c_1^p ; $p = 1, 2, \dots, P$, где P – число эталонов. Каждый p -й эталон преобразуем с помощью матрицы

$\begin{pmatrix} c_1^p & 0 \\ 0 & 1/c_1^p \end{pmatrix}$, тогда получим, по-прежнему, P искаженных эталонов, один из которых (пока неизвестный) связан с распознаваемым изображением матрицей вида

$$\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}, \quad (6)$$

где $a = c_1^p c_2$.

Преобразование матрицей (6) не меняет координаты у любой точки, а координата x меняется линейно $x' = x + ay$. При таком преобразовании расстояние между любыми двумя точками, имеющими одинаковые значения y , не меняется. Из этих соображений легко находятся значения ρ и c_2 .

Таким образом, можно узнать эталон, которому соответствует искаженное изображение, и определить все параметры неизвестного преобразования. Тогда

$$g_2 = \begin{pmatrix} S/S_0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ c_1/c_3 & 1 \end{pmatrix} \begin{pmatrix} 1 & c_1^p c_2 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} c_1^p & 0 \\ 0 & 1/c_1^p \end{pmatrix}$$

Приведенный алгоритм позволяет распознать объект по его искаженному изображению и определить все шесть параметров аффинного преобразования. Основное ограничение области применения алгоритма – необходимость несовпадения центра формы с центром яркости на изображении каждого эталона. На бинарных изображениях они всегда совпадают. Для многоградационных изображений необходимо практически точное совпадение яркостей всех точек изображения эталона и распознаваемого объекта, что на практике наблюдается крайне редко. Кроме этого, желательно, чтобы координаты центров яркости эталонов существенно отличались друг от друга.

Метод последовательного перебора характерных точек изображения

Аналогично рассмотренному методу на первом этапе предполагается провести центрирование изображения по координатам центра формы, тем самым компенсировав подгруппу g_1 из (2).

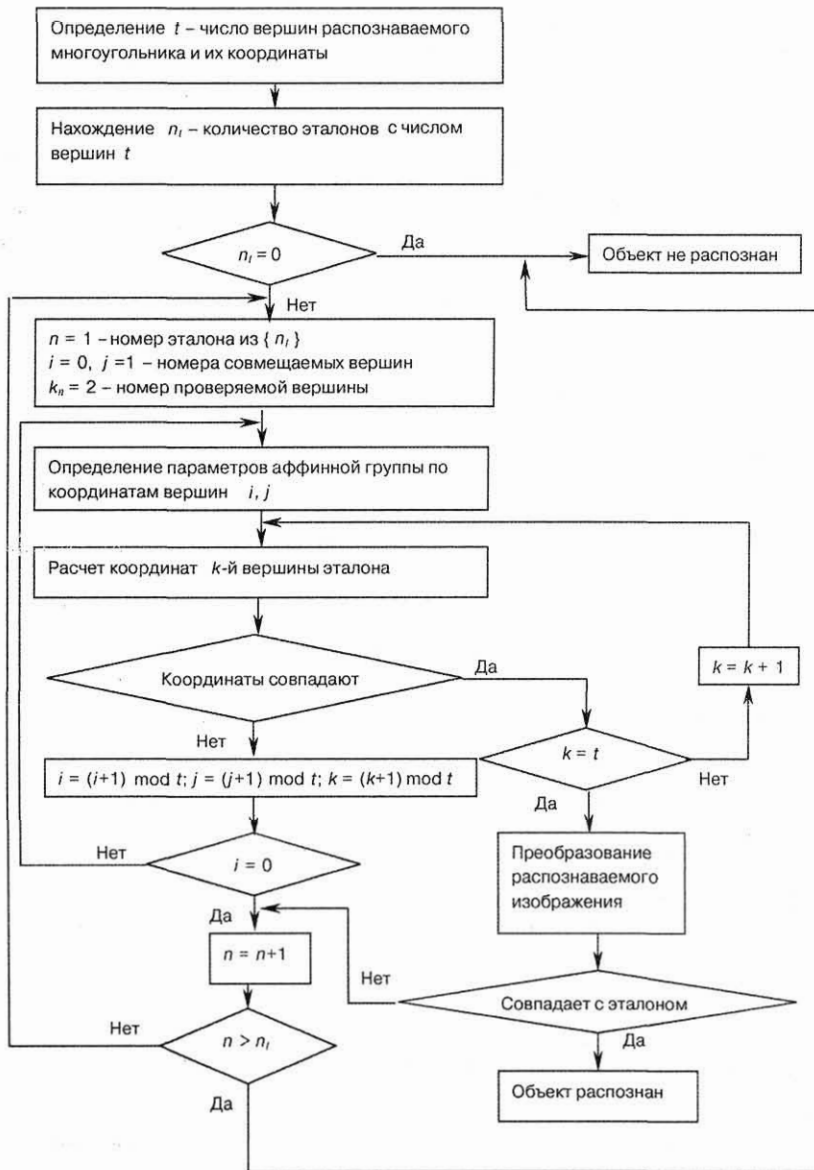
Для нахождения параметров аффинной группы без сдвига g_2 :

$\begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix}$ достаточно выбрать две

характерные точки на изображении эталона, лежащие на прямой, не проходящей через начало координат, и найти соответствующие им точки на изображении распознаваемого объекта. Если представить контур изображения распознаваемого объекта и каждого эталона в виде замкнутой ломаной линии, то в качестве характерных точек можно использовать ее вершины. Пусть число таких вершин на изображении

ρ -го эталона $s_p \in S$; $S = \{s_1, \dots, s_p\}$, где S – множество вершин эталонов, P – число эталонов. Представим контур изображения распознаваемого объекта также в виде ломаной линии. Как известно [2], аффинное преобразование переводит прямую в прямую, следовательно, число вершин на изображении распознаваемого объекта $s_x \in S$ и последовательность вершин сохраняется. Если в S нет одинаковых элементов, то, очевидно, проблема распознавания решена. Остается найти параметры элемента группы g_2 . Дальнейший алгоритм, представленный на рис. 2, предполагает перебор вершин контура изображения распознаваемого объекта при конечном числе возможных сочетаний.

Возьмем две любые последовательно расположенные вершины $\{i, j\}$ на контуре изображения эталона. В предположении, что им соответствуют две последовательно рас-



■ Рис. 2. Алгоритм распознавания перебором вершин контура

положенные вершины на контуре изображения распознаваемого объекта $\{l, k\}$, составим систему уравнений

$$\begin{cases} x'_j = a_1 x_j + a_2 y_j; \\ y'_j = a_3 x_j + a_4 y_j; \\ x'_k = a_1 x_k + a_2 y_k; \\ y'_k = a_3 x_k + a_4 y_k; \end{cases} \quad (7)$$

где $x_j, y_j, x'_j, y'_j, x_k, y_k, x'_k, y'_k$ – координаты соответствующих точек. Решив систему (7), определим параметры a_1, \dots, a_4 и преобразуем эталонное изображение группой g_2 . Если после сравнения преобразованного и распознаваемого изображений совпадение будет отсутствовать, то следует повторить описанные действия для следующей пары вершин контура изображения эталона. Рано или поздно будет получено совпадение изображений, так как число вершин контура конечно.

Если в S есть одинаковые элементы, т. е. среди эталонов есть совпадающие по числу вершин, то следует разбить S на непересекающиеся подмножества эталонов с одинаковым числом вершин контура и выполнять описанные действия по перебору вершин контура последовательно для всех эталонов подмножества до тех пор, пока не будет получено совпадение изображений. Естественно предполагается, что распознаваемый объект соответствует одному из имеющихся эталонов.

Практически при переборе пар вершин эталона не требуется подвергать преобразованию g_2 все эталонное изображение (см. рис. 2). Достаточно преобразовывать только координаты вершин контура, прерывая процесс при первом несовпадении и переходя к следующей паре вершин или эталону, если вершины данного эталона закончились. Преобразовывать все изображение следует только в случае полного совпадения вершин. Последнее связано с тем, что эталоны могут иметь одинаковый контур, но отличаться внутренней структурой, например, иметь различное число и форму отверстий.

Основное ограничение рассмотренного выше алгоритма – изображения эталонов не должны преобразовываться друг в друга аффинной группой. На практике это означает, что треугольные эталоны произвольной формы и размеров должны отличаться

внутренней структурой, а четырехугольные эталоны должны иметь хотя бы две непараллельные стороны. Другое ограничение – эталоны с плавным криволинейным контуром практически невозможно представлять в виде соответствующих многоугольников на изображениях как эталонов, так и распознаваемого объекта.

Естественно, при компьютерной реализации рассмотренного алгоритма условие точного равенства координат и совпадения изображений необходимо заменить некоторой мерой близости с экспериментально подобранным порогом, определяющим наличие совпадения. Данное требование вызвано дискретным (пиксельным) принципом хранения и обработки изображения в компьютере, что приводит к погрешностям при округлении координат.

При компьютерной реализации алгоритма возникает еще одна проблема – нахождение координат вершин многоугольника. Существующие алгоритмы поиска вершин [5] наиболее эффективны для углов линий контура в искомой вершине, близких к прямым углам. При приближении угла контура при вершине к 180° или к 0° точность определения координат вершины резко падает, так как в таких случаях трудно решить, что найдено – вершина линии контура или случайное отклонение от прямой, обусловленное помехами предварительной обработки изображения. Компьютерное моделирование данного алгоритма показало, что с учетом этих ограничений алгоритм достаточно эффективен для многоугольников, угол контура которых при вершине находится в пределах $60^\circ - 120^\circ$.

Литература

1. **Марр Д.** Зрение: Информационный подход к изучению представления и обработки зрительных образов/Пер. с англ. – М.: Радио и связь, 1987. – 400 с.
2. **Путятин Е. П., Аверин С. И.** Обработка изображений в робототехнике. – М.: Машиностроение, 1990. – 320 с.
3. **Хорн Б. К. П.** Зрение роботов/Пер. с англ. – М.: Мир, 1989. – 487 с.
4. **Ерош И. Л., Игнатъев М. Б., Москалев Э. С.** Адаптивные робототехнические системы: Методы анализа и системы обработки изображений. – Л.: Изд-во ЛИАП, 1985. – 144 с.
5. Техническое зрение роботов / В. И. Мошкин, А. А. Петров, В. С. Титов и др.; Под ред. Ю. Г. Якушенко. – М.: Машиностроение, 1990. – 272 с.