

УДК 004.93

АЛГОРИТМ ОПРЕДЕЛЕНИЯ ПОЛОЖЕНИЯ И ОРИЕНТАЦИИ ТРЕХМЕРНЫХ ОБЪЕКТОВ ПО ВИДЕОИЗОБРАЖЕНИЯМ НА ОСНОВЕ ВЕРОЯТНОСТНОГО ПОДХОДА

Захаров А.А., Тужилкин А.Ю., Веденин А.С.

Муромский институт (филиал), Владимирский государственный университет имени Александра Григорьевича и Николая Григорьевича Столетовых, Муром, e-mail: aa-zaharov@ya.ru

В статье представлен алгоритм для определения положения и ориентации трехмерных объектов по видеоизображениям. Показана актуальность работы. Приведены возможные сферы использования результатов исследований. Рассмотрены существующие подходы для определения положения и ориентации трехмерных объектов. На первом этапе определяются приближенные значения положения и ориентации объекта. Для сокращения пространства поиска на начальном этапе выбираются несколько основных положений объекта в пространстве с фиксированным углом поворота по координатным осям. Определение позиции и ориентации основано на нахождении максимума апостериорной вероятности по отношению к существующим шаблонам изображения наблюдаемого объекта. При анализе используется описание трехмерных объектов на основе комплексной информации о форме, взаимном положении примитивов в пространстве, пропорциях. Используя условные вероятности, можно задавать степень важности свойств для определенных объектов сцены. Приведены результаты работы алгоритма.

Ключевые слова: определение положения и ориентации, трехмерные объекты, видеоизображения, вероятностный подход

THE ALGORITHM FOR DETERMINING THE POSITION AND ORIENTATION OF THREE-DIMENSIONAL OBJECTS FROM VIDEO IMAGES ON THE BASIS OF A PROBABILISTIC APPROACH

Zakharov A.A., Tuzhilkin A.Y., Vedenin A.S.

Murom Institute (Branch) of Vladimir State University named after Alexander G. and Nicholas G. Stoletovs, Murom, e-mail: aa-zaharov@ya.ru

The algorithm for determining the position and orientation of three-dimensional objects by video images is considered in the article. Actuality of work is shown. Possible areas of the use of the research results are presented. Existing approaches for determining the position and orientation of the three-dimensional objects are considered. Approximate values of the position and orientation of the object are determined in the first phase. Several basic object positions in a space with a fixed angle of rotation are selected to reduce the search space in the initial step. Determination of the position and orientation based on finding the maximum a posteriori probability with respect to the existing template image of the observed object. Description of three-dimensional objects used on the basis integrated information on the form, the mutual position of the primitives in the space, proportions. The degree of importance of the properties for certain objects in the scene can be specified using the conditional probabilities. The results of the algorithm are given.

Keywords: determination of the position and orientation, three-dimensional objects, video images, the probability approach

Задача автоматического определения положения и ориентации трехмерных объектов с использованием систем технического зрения актуальна в таких областях, как контроль технологических процессов в промышленности, охранные системы, человеко-машинное взаимодействие, навигация транспортных средств, индустрия развлечений и т.д. Решение этой задачи в автоматическом режиме позволит обеспечить автономность функционирования систем, повысить точность контроля, увеличить производительность работы, снизить затраты на подготовительные операции и обслуживание.

Целью работы является разработка и исследование алгоритма автоматического определения положения и ориентации трех-

мерного объекта по видеоизображениям. Смысл предлагаемого подхода в том, чтобы сократить пространство поиска за счет определения на начальном этапе приближенных значений положения и ориентации объекта. Это осуществляется с использованием метода максимального правдоподобия.

Обзор методов определения положения и ориентации объектов в пространстве

В настоящее время все подходы к определению положения и ориентации трехмерных объектов в пространстве можно разделить на две группы: неоптические и оптические [7]. К неоптическим технологиям относятся электромагнитные, механические, инерционные, акустические и гибридные [7].

В электромагнитных системах используются магниты, размещенные на объекте. Система рассчитывает позиции по искажениям магнитного потока. Недостатками подобных систем являются зависимость от магнитных и электрических помех, переменчивая чувствительность сенсоров в зависимости от положения, ограниченное рабочее пространство.

В механических системах используется каркас, который крепится к объекту. Движение составных частей объекта приводит в движение элементы каркаса. В этом случае анализируются углы между скелетными элементами каркаса. Однако перемещение всего объекта в пространстве механическая система фиксировать не способна.

Инерциальные системы для сбора информации о движении используют гироскопы, расположенные на объекте. Система определяет не только положение сенсора, но также угол его наклона. Инерционные системы устойчивы к помехам различного рода, но дрейф во время движения приводит к накоплению ошибки.

Акустические системы используют датчики, которые формируют сигнал. Приемники, установленные на некотором расстоянии, измеряют время полета переданного сигнала. Акустические системы восприимчивы к помехам окружающей среды. Также имеются ограничения, связанные с количеством одновременно используемых устройств, размером рабочего пространства. Гибридные системы для повышения эффективности могут использовать в своем составе несколько технологий.

Оптические методы основаны на использовании маркерных и безмаркерных принципов. Маркерные оптические системы бывают пассивными и активными. Оптические пассивные маркерные системы используют датчики, которые отражают падающий на них свет. В оптических активных системах свет испускается самими маркерами.

Вышеописанные технологии требуют использования специальных датчиков, средств сопряжения с блоками приема и обработки данных, ограниченным рабочим пространством. Основным недостатком рассмотренных методов при использовании в специализированных системах является невозможность применения датчиков при постоянном наблюдении множества сцен, что существенно ограничивает их применение. Безмаркерные оптические методы являются наиболее перспективными с точки зрения простоты использования и универсальности. Поэтому важна разработка методов и программных средств определения положения и ориентации объ-

ектов с использованием безмаркерных оптических методов.

Вопросы распознавания на основе пассивных оптических методов в настоящее время изучены недостаточно. Использование компьютерного зрения ограничено следующими факторами наблюдения: наличием шумовых и оптических эффектов, изменяющимся фоном, эффектами резкой смены освещения, взаимным перекрытием объектов в сцене. Кроме того, при реализации систем существуют неопределенности, связанные с выделением информативных признаков объекта, так как двумерное изображение не обладает полнотой информации о свойствах трехмерного объекта. Поэтому для реализации подхода необходимо использовать заранее предопределенную информацию о структуре сцены и условиях получения изображения.

Безмаркерные оптические методы определения положения и ориентации объекта принято делить на трекары и детекторы [4]. Трекары отслеживают состояние объекта от кадра к кадру. Детекторы определяют положение объекта в текущем кадре без учета состояния в предыдущих кадрах. Применение конкретного безмаркерного оптического метода зависит от условий наблюдения сцены. Подходы, использующие контуры, лучше применять в тех случаях, когда изображение визуальной обстановки получено в градациях серого цвета, на снимках можно достаточно четко выделить края объекта, присутствуют эффекты резкой смены освещения [5, 8]. Если объекты имеют выраженную текстуру, сцена хорошо освещена, видимое пространство небольшое, то лучше использовать методы оптического потока [1]. Методы детектирования часто используют точки интереса. Однако методы на основе использования точек интереса чувствительны к изменению освещения и хорошо подходят для цветных изображений [2, 3, 6].

Алгоритм определения положения и ориентации трехмерных объектов по видеоизображениям на основе вероятностного подхода

Входными данными для описания структуры сцены являются видеоизображения. При анализе используется описание трехмерных объектов на основе комплексной информации о форме элементов, взаимном положении примитивов в пространстве, пропорциях, яркостных характеристиках фона сцены.

Для реализации алгоритма необходимо обучение системы, в результате которого будет сформирован набор изображений объекта в различных положениях. Происходит

съемка объекта из различных положений. Предполагается, что камера перемещается по окружности, в центре которой располагается объект. Съемка производится через одинаковые угловые расстояния (рис. 1).

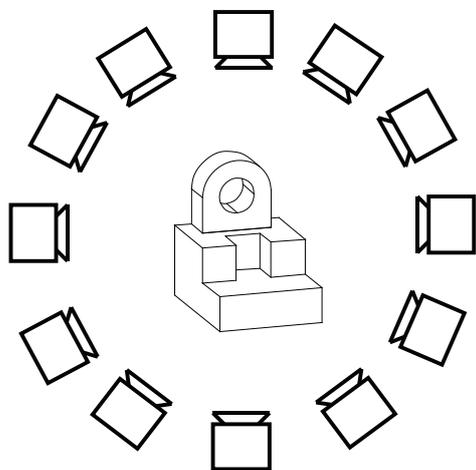


Рис. 1. Получение изображений объекта с изменением угла вращения камеры вокруг центральной оси объекта

Для распознавания объекта необходимо выделить признаки на изображении. Например, если рассматриваемый объект представляет собой параллелепипед, то на изображении он представлен видимыми гранями. Таким же образом можно провести декомпозицию изображения реального объекта. Для этой задачи используются глобальные дескрипторы изображения – гистограммы ориентированных градиентов

(HOG). Множество HOG-дескрипторов, локализованных в одном регионе, позволяет судить о конкретной части объекта. Выделим в каждом регионе центр. Совокупность выделенных таким образом точек будет характеризовать положение и ориентацию объекта (рис. 2).

При нахождении соответствий рассматриваются выделенные области на изображении-эталоне и на анализируемом изображении. Для нахождения соответствий предлагается использовать алгоритм ожидания-максимизации (Expectation-maximization, EM). Алгоритм позволяет итеративно вычислять значение максимального правдоподобия при наличии скрытых данных. Основная идея алгоритма заключается в задании модели для некоторого набора данных и многократном вычислении параметров. Вычисление максимального правдоподобия позволяет оценить параметры модели, для которых наблюдаемые данные являются наиболее вероятными. EM-алгоритм итеративно выполняется в два этапа. На первом шаге (ожидание) отсутствующая информация вычисляется с учетом наблюдаемых данных и текущих параметров модели. На втором шаге (максимизация) функция правдоподобия достигает максимума при условии, что отсутствующие данные известны. На этапе максимизации алгоритм использует текущие значения отсутствующих данных. Для обеспечения сходимости EM-алгоритма должна увеличиваться вероятность на каждой итерации.

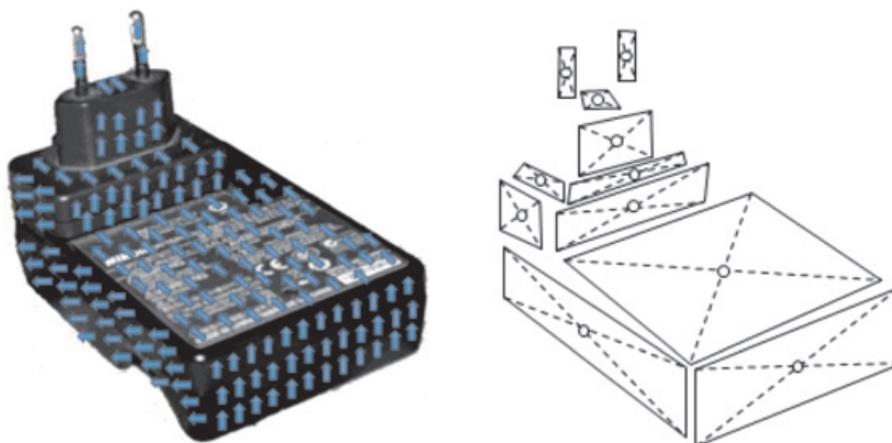


Рис. 2. Выделение характерных точек на изображении

В нашем случае изображение формы объекта представлено точечным множеством произвольного размера. Порядок точек не имеет значения. Получаем два мно-

жества точек $\mathbf{X} = (x_1, x_2, \dots, x_M)^T \in \mathbb{R}^{M \times 2}$ и $\mathbf{Y} = (y_1, y_2, \dots, y_N)^T \in \mathbb{R}^{N \times 2}$, $M \neq N$. Задача состоит в нахождении соответствий между

X и Y . Предположим, что y_n представляют собой наблюдения с моделью смеси:

$$p(y_n) = p(y_n | v_f) p(v_f) + p(y_n | v_b) p(v_b); (1)$$

$$p(y_n | v_f) = \sum_{m=1}^M p(y_n | m, v_f) p(m), (2)$$

где v_b – компонент смеси, представляющий «модель фона», которая позволяет учитывать возможность выбросов; v_f – компонент смеси, который описывает распознаваемые объекты.

Изменение положения объекта описывается параметрами: s – параметр масшта-

бирования; \mathbf{D} – матрица поворота; $\mathbf{t} \in R^2$ – вектор переноса.

Назначим равную вероятность для каждого компонента смеси $p(m) = 1/M$. Это предотвращает соответствие одного объекта изображения нескольким объектам другого изображения.

Оценки максимального правдоподобия для других параметров модели s , \mathbf{D} , \mathbf{t} и $p(v)$ находятся с помощью EM-алгоритма. Алгоритм в этом случае имеет следующий вид.

Е-шаг. Вычислить выражение, используя текущие значения параметров:

$$p(m, v_f | y_n) = \frac{p(y_n | m, v_f) p(m) p(v_f)}{p(y_n)}. (3)$$

М-шаг. Обновить параметры:

$$p(v_f) = \frac{1}{N} \sum_n p(v_f | y_n); (4)$$

$$p(v_b) = 1 - p(v_f); (5)$$

$$(s, D, t) = \arg \min_{s, R, t} \sum_{m, n} p(m, v_f | y_n) \|y_n - s D x_m - t\|^2. (6)$$

Формула (6) описывает задачу нахождения соответствий между двумя точечными множествами размерами N и M . Степень соответствия пары (x_m, y_n) задается величиной $p(m, v_f | y_n)$.

Исследование алгоритма

Исследование алгоритма проводилось на тестовых изображениях (рис. 3).

Для обучения системы была произведена съемка объекта с угловым расстоянием 10 градусов. Таким образом, было получено 36 кадров. На каждом снимке были выделены координаты опорных точек. Условием работоспособности алгоритма является успешное распознавание каждого положения объекта в интервале ± 10 градусов.



Рис. 3. Тестовые изображения объекта

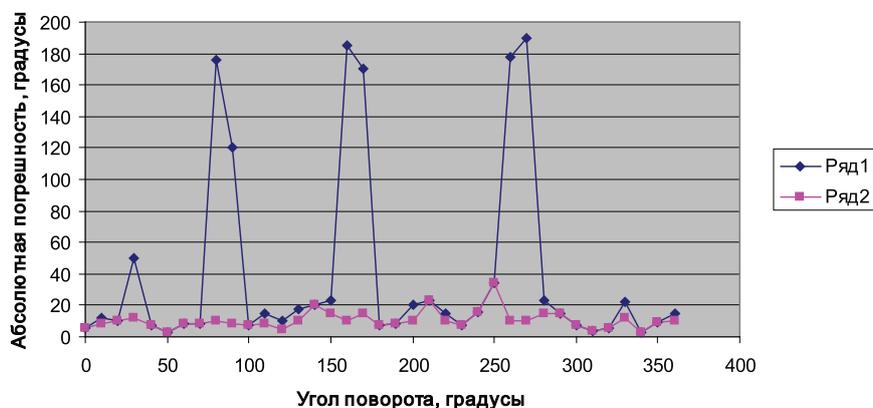


Рис. 4. Результаты определения угла поворота объекта: а – определение положения с помощью одного вида (ряд 1); б – определение положения с помощью двух видов (ряд 2)

Положение и ориентация объекта оценивались вначале на основе одиночного изображения. Исследование алгоритма проводится с объектом, который вращается вокруг одной из координатных осей на 360 градусов. В общей сложности было сделано около 100 снимков объекта. Расстояние от камеры до объекта для всех снимков одинаковое. Была измерена абсолютная погрешность определения ориентации объекта. Следует отметить, что ряд результатов имеет ошибку ориентации 180 градусов, что может быть объяснено симметричностью объекта.

Процент правильно распознанных положений объекта при использовании одиночного изображения составляет около 40% (рис. 4, а). При использовании двух видов при распознавании эта величина увеличивается до 75% (рис. 4, б).

Таким образом, при использовании двух видов в большинстве случаев абсолютная погрешность определения угла поворота объекта не превышает 10 градусов. Поэтому алгоритм можно использовать для определения приближенного угла поворота объекта.

Заключение

В работе представлен алгоритм для определения положения и ориентации объектов по изображениям на основе вероятностного подхода. Подход основан на использовании EM-алгоритма и позволяет итеративно оценивать параметры проекции трехмерной модели на плоскости изображения. Проведенные эксперименты показывают, что алгоритм способен восстанавливать угол поворота объекта в пределах ± 10 градусов. В дальнейшем предполагается усовершенствовать алгоритм для более точного определения положения и ориентации на основе вычисления параметров линейных признаков объекта на изображении.

Работа выполнена при поддержке гранта РФФИ 13-07-97523, задания № 2014/13 на выполнение государственных работ в сфере научной деятельности в рамках базовой части государственного задания Минобрнауки России.

Список литературы

1. Basu S., Essa I., Pentland A. Motion regularization for modelbased head tracking // Proceedings ICPR. – 1996. – pp. 611–616. (оптический поток).
2. Gordon I., Lowe D.G. What and where: 3D object recognition with accurate pose // Lecture Notes in Computer Science. – 2006. – Vol. 4170. – P. 67–82. (Tracking by detection – texture based)
3. Jurie F., Dhome M. A simple and efficient template matching algorithm // Proceedings ICCV. – 2001. – P. 544–549. (recursive based – texture – template).
4. Lepetit V., Fua P. Monocular Model-Based 3D Tracking of Rigid Objects: A Survey // Foundations and Trends in Computer Graphics and Vision. – 2005. – Vol. 1, № 1. – P. 1–89. (обзор).
5. Ulrich M., Wiedemann C., Steger C. CAD-based recognition of 3D objects in monocular images // ICRA'09 Proceedings of the 2009 IEEE international conference on Robotics and Automation. – 2009. – P. 2090–2097. (Tracking by detection – line based).
6. Vacchetti L., Lepetit V., Fua P. Stable real-time 3D tracking using online and offline information // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2004. – vol. 26, № 10. – P. 1385–1391. (recursive based – texture – interest point).
7. Welch G., Foxlin E. Motion Tracking: No Silver Bullet, but a Respectable Arsenal // Computer Graphics and Applications. – 2002. – Vol. 22, № 6. – P. 24–38.
8. Wuest H., Vial F., Stricker D. Adaptive line tracking with multiple hypotheses for augmented reality // Proceedings ISMAR. – 2005. – P. 62–69. (recursive tracking- ребра).

References

1. Basu S., Essa I., Pentland, A. Proceedings ICPR, 1996, pp. 611–616.
2. Gordon I., Lowe, D.G. Lecture Notes in Computer Science, 2006, Vol. 4170, pp. 67–82.
3. Jurie F., Dhome M. Proceedings ICCV, 2001, pp. 544–549.
4. Lepetit, V., Fua P. Foundations and Trends in Computer Graphics and Vision, 2005, Vol. 1, no. 1, pp. 1–89.
5. Ulrich M., Wiedemann C., Steger, C. ICRA'09 Proceedings of the 2009 IEEE international conference on Robotics and Automation, 2009, pp. 2090–2097.
6. Vacchetti L., Lepetit V., Fua, P. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, vol. 26, no. 10, pp. 1385–1391.
7. Welch G., Foxlin E. Computer Graphics and Applications, 2002, Vol. 22, no. 6, pp. 24–38.
8. Wuest H., Vial F., Stricker D. Proceedings ISMAR, 2005, pp. 62–69.

Рецензенты:

Жизняков А.Л., д.т.н., профессор, первый зам. директора, МИ (ф) ВлГУ, г. Муром;
 Орлов А.А., д.т.н., доцент, зав. кафедрой физики и прикладной математики, МИ (ф) ВлГУ, г. Муром.

Работа поступила в редакцию 06.11.2014.