

УДК 621.391.266

МОДИФИЦИРОВАННОЕ ОБОБЩЁННОЕ ПРЕОБРАЗОВАНИЕ ХАФА ДЛЯ ОБРАБОТКИ ТРЁХМЕРНЫХ ИЗОБРАЖЕНИЙ С НЕИЗВЕСТНЫМИ ПАРАМЕТРАМИ ВРАЩЕНИЯ И МАСШТАБИРОВАНИЯ*

А. А. Роженцов, К. В. Морозовский, А. А. Баев

*Поволжский государственный технологический университет,
424000, г. Йошкар-Ола, пл. Ленина, 3
E-mail: krtmbs@volgatech.net*

Предлагается подход к обработке трёхмерных изображений на базе модифицированного обобщённого преобразования Хафа. Показана возможность обработки изображений с неизвестными параметрами вращения и масштабирования, а также изображений, представленных отдельными фрагментами исходного образа. Исследована эффективность предложенных алгоритмов при решении задачи распознавания трёхмерных изображений.

Ключевые слова: распознавание трёхмерных изображений, обобщённое трёхмерное преобразование Хафа, инвариантность к вращению и масштабированию.

Введение. При обработке различных изображений (как плоских, так и объёмных) нередко приходится сталкиваться с ситуацией, когда в поле зрения находится неполное изображение объекта, соответствующего тому или иному эталону. Если на наблюдаемое изображение дополнительно воздействуют различные помеховые факторы и неизвестны параметры преобразований переноса, вращения, масштабирования, то принятие решения о принадлежности наблюдаемого изображения к тому или иному классу становится зачастую невозможным. В подобных ситуациях алгоритмы, базирующиеся на согласованной фильтрации, построении контурных моделей, гармоническом сферическом анализе и т. п., могут оказаться неэффективными и необходим переход к другим методам обработки, работоспособным в этих условиях. Одним из таких методов является преобразование Хафа (Hough) [1–6].

Известен ряд разновидностей этого преобразования, применяемого для обнаружения в изображениях прямых линий, параметрических кривых, кривых произвольной формы. В [1–5] предложено использовать преобразование Хафа для обработки объёмных изображений. Однако в отмеченных источниках не рассмотрен вопрос о возможности его применения при неизвестных параметрах вращения и масштабирования. Как правило, в таких случаях предлагается увеличивать размерность аккумуляторного массива в соответствии с количеством неизвестных параметров, но для описания вращений в трёхмерном пространстве необходимо знание координат оси вращения и угла поворота вокруг неё. В результате размерность аккумуляторного массива значительно возрастает, что существенно повышает требования к быстродействию и объёму памяти вычислительной системы.

Цель данной работы — модифицировать обобщённое преобразование Хафа для обработки трёхмерных изображений при неизвестных параметрах вращения и масштабирования и обеспечения умеренной трудоёмкости и возможности применения параллельных вычислений на базе графических сопроцессоров.

*Работа выполнена в рамках Аналитической ведомственной целевой программы «Развитие научного потенциала высшей школы» (проект № 8.1013.2011).

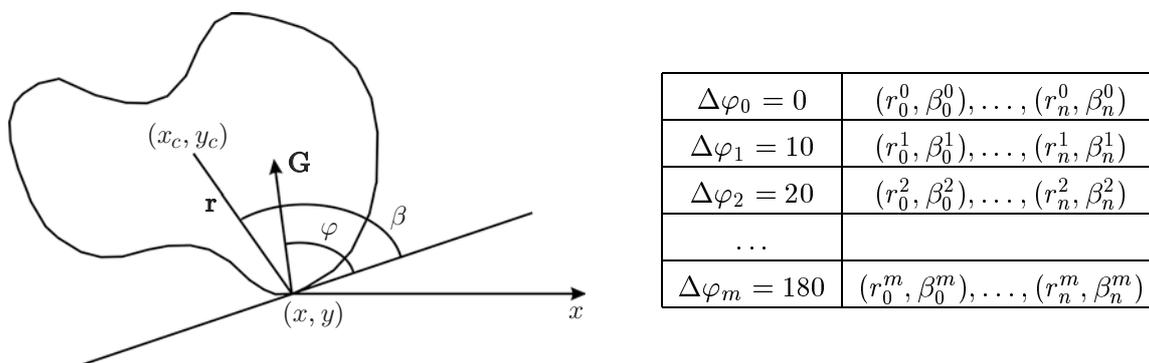


Рис. 1. Построение R-таблицы для плоского изображения

Модифицированное обобщённое преобразование Хафа. В обобщённом преобразовании Хафа [6] для плоских изображений индексация при построении R-таблицы выполняется по величине угла между нормалью \mathbf{G} к поверхности в данной точке и направлением на центр формы \mathbf{r} (рис. 1). При этом индекс таблицы является одномерным.

В случае объёмных изображений [1] индексация R-таблицы выполняется по двум углам φ и ψ (рис. 2), задающим ориентацию в пространстве нормали к данной точке поверхности. В качестве значений в R-таблицу заносятся величины углов α , β и расстояния \mathbf{r} , определяющие взаимное положение данной точки поверхности и центра формы.

При обработке трёхмерного изображения для каждой его точки определяется направление нормали к поверхности. По данным R-таблицы строится множество векторов для этого направления нормали, и в аккумуляторном массиве инкрементируются элементы, соответствующие концам этих векторов. При совпадении форм наблюдаемого объекта и эталона, а также при отсутствии вращений максимум аккумуляторного массива будет формироваться в центре формы.

При обработке изображений параметры φ и ψ могут меняться в широких пределах из-за изменения параметров вращения наблюдаемого изображения. В результате значения углов $\alpha^{(c)}$, $\beta^{(c)}$ и расстояния $\mathbf{r}^{(c)}$ для него при данных φ и ψ будут отличаться от аналогичных параметров эталонного изображения (рис. 2, c). При этом накопление отсчётов в аккумуляторном массиве производится неверно и возможно возникновение ситуаций, когда максимум получается для эталона, не соответствующего наблюдаемому объекту. Аналогичная ситуация отмечается и при изменении масштаба изображения (рис. 2, d). Использование здесь обобщённого преобразования Хафа требует увеличения размерности аккумуляторного массива в соответствии с количеством неизвестных параметров преобразований. Так, в условиях неопределённости относительно параметров вращения и масштабирования аккумуляторный массив должен иметь размерность равную семи. В вычислительном плане формирование и хранение такого массива представляет собой достаточно трудоёмкую задачу.

Устранение этого недостатка обобщённого преобразования Хафа возможно с помощью подходов, базирующихся, во-первых, на изменении принципа индексации R-таблицы, во-вторых, на занесении в R-таблицу данных, которые не зависят от параметров вращения объекта, в-третьих, на модификации принципа накопления отсчётов в аккумуляторном массиве. Сам принцип преобразования при этом аналогичен преобразованию Хафа, и далее такое преобразование будем рассматривать как модифицированное преобразование Хафа для трёхмерных изображений.

Для индексации R-таблицы воспользуемся подходом, основанным на использовании в качестве аргументов номеров отсчётов эталонного изображения. В R-таблицу предлагается заносить значения расстояния от данной точки до центра формы \mathbf{r} и угла w между

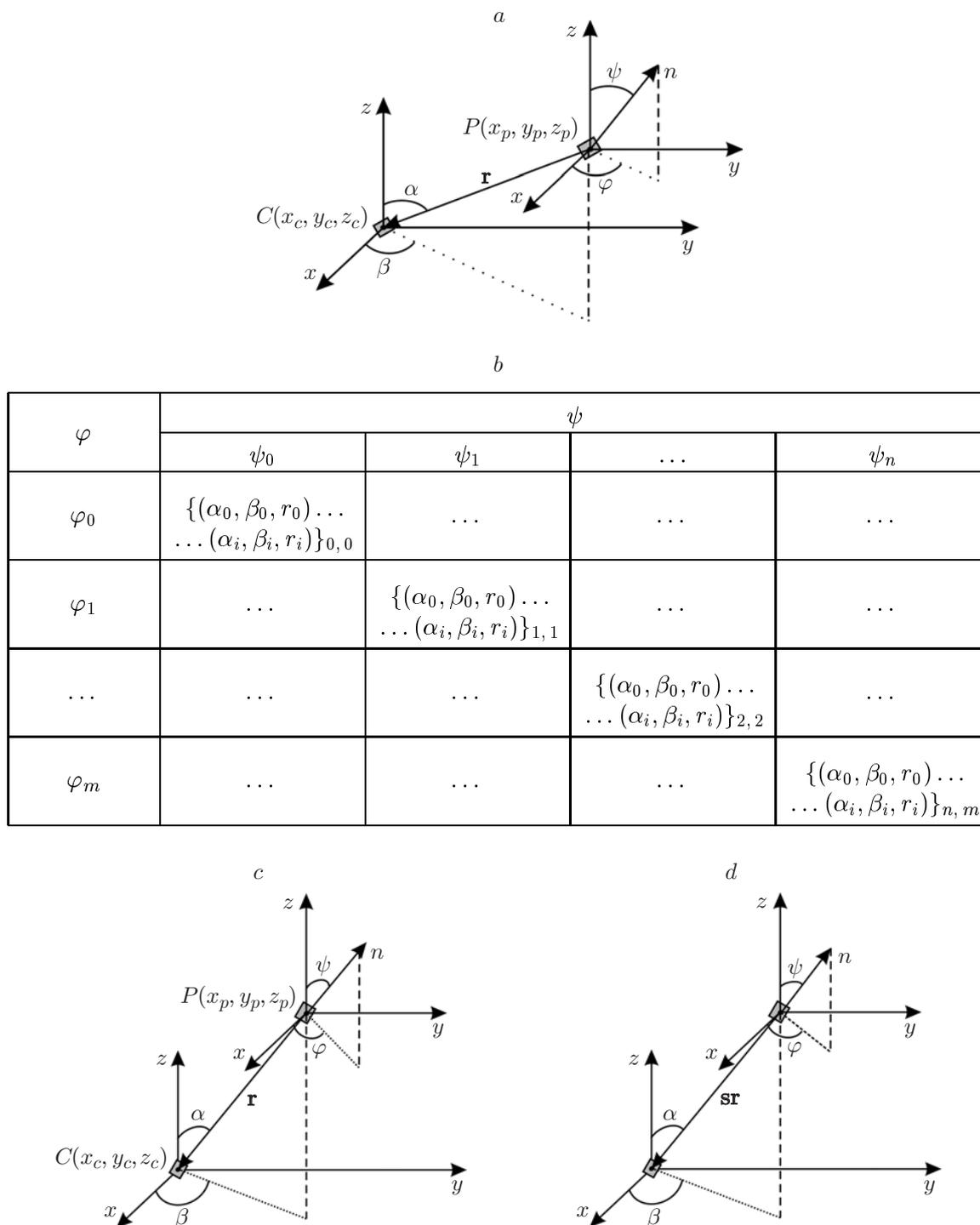


Рис. 2. Построение R-таблицы для трёхмерного изображения: *a* — задание параметров положения центра формы для формирования R-таблицы; *b* — структура R-таблицы; *c*, *d* — влияние параметров масштабирования и вращения на параметры положения центра формы

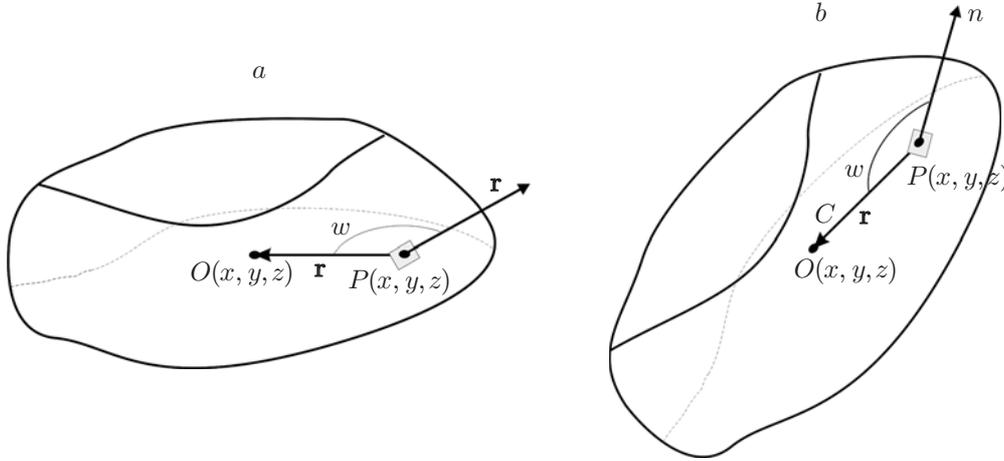


Рис. 3. Инвариантность угла w к вращениям трёхмерного объекта

нормалью и направлением на центр формы, который не зависит от параметров вращения (рис. 3). Поскольку при сканировании трёхмерных объектов количество и порядок следования отсчётов могут изменяться, при распознавании необходимо выполнять полный перебор взаимных комбинаций отсчётов эталонного и наблюдаемого изображений. При размерностях эталона $N_{\text{э}}$ и сигнального изображения $N_{\text{с}}$ общее количество комбинаций составит $N_{\text{п}} = N_{\text{э}}N_{\text{с}}$.

В случае, когда параметры вращения неизвестны, положение центра формы относительно данной точки может быть определено с точностью до расстояния и угла поворота вокруг нормали к этой точке поверхности. Таким образом, для данной точки поверхности в пространстве может быть построена окружность, точки которой образуют область, где предположительно находится центр формы (рис. 4, а). Точка, в окрестности которой проходит наибольшее количество окружностей, и будет центром формы (рис. 4, б, с).

Для аналитического описания процесса заполнения аккумуляторного массива воспользуемся алгеброй кватернионов при описании вращений [7]. Положение вектора $O_{n,m}$ в трёхмерном пространстве при заданной нормали η_m и угле w_n можно представить в виде

$$O_{n,m} = r_n b_z(\varphi_m) b_y(\psi_m) b_z(w_n) i b_z^*(w_n) b_y^*(\psi_m) b_z^*(\varphi_m), \quad (1)$$

где $b_z(\varphi_m) = \cos \varphi_m + k \sin \varphi_m$; $b_y(\psi_m) = \cos \psi_m + j \sin \psi_m$; $b_z(w_n) = \cos w_n + k \sin w_n$; i, j, k — мнимые единицы.

Координаты точек на окружности в основании конуса получим путём вращения вектора $O_{n,m}$ вокруг нормали η_m на угол φ , $\varphi \in [0, 2\pi)$:

$$O_{n,m}(\varphi) = b_m(\varphi) O_{n,m} b_m^*(\varphi), \quad (2)$$

где $b_m(\varphi) = \cos \varphi + \eta_m \sin \varphi$.

Значение индекса $C_{m,n}(\varphi)$ аккумуляторного массива, в котором производится накопление, определяется суммированием вектора $O_{n,m}(\varphi)$ со значением координаты m -й точки поверхности:

$$C_{m,n}(\varphi) = ix_m + jy_m + kz_m + O_{n,m}(\varphi),$$

$$n = 0, 1, \dots, N_{\text{э}} - 1, \quad m = 0, 1, \dots, N_{\text{с}} - 1, \quad \varphi \in [0, 2\pi). \quad (3)$$

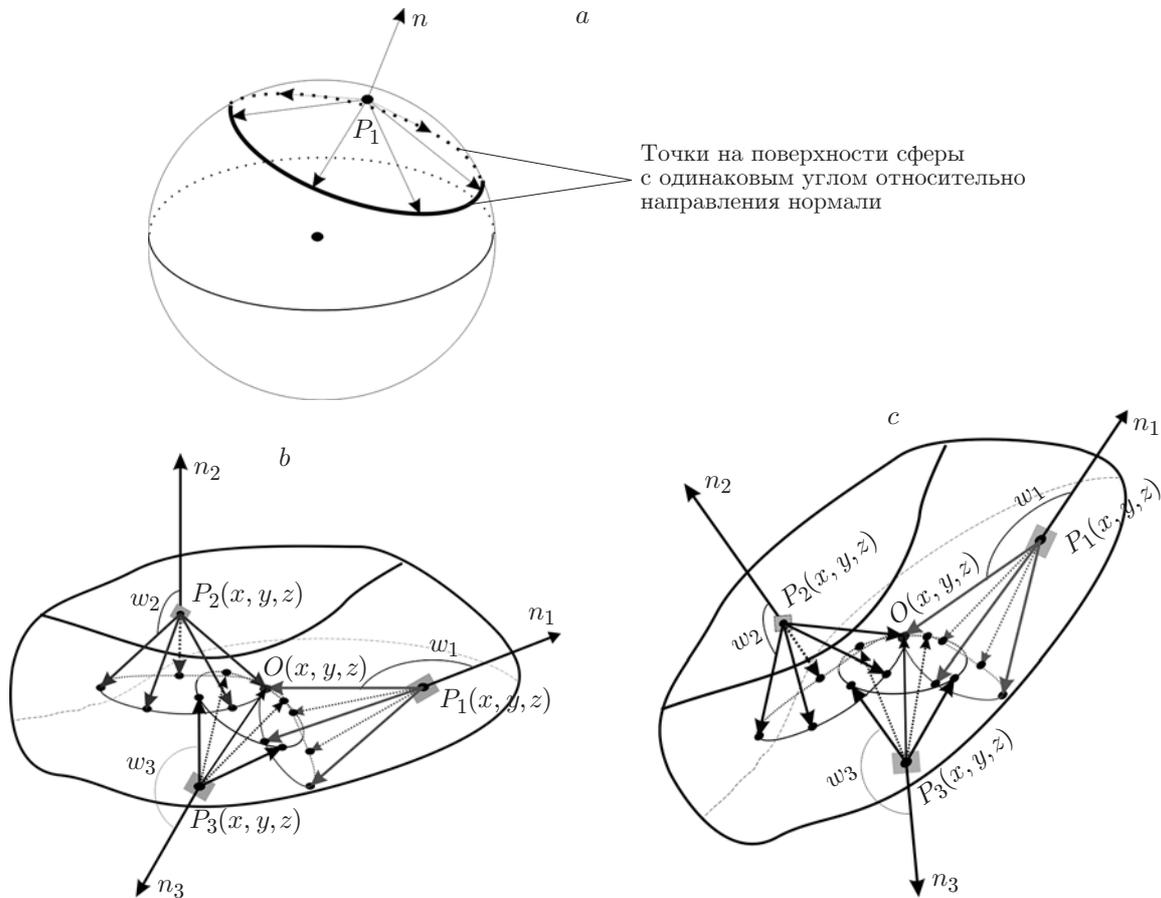


Рис. 4. Принцип голосования в аккумуляторном массиве: формирование голосующих точек в аккумуляторном массиве (а); определение центра формы при известных (b) и неизвестных (c) параметрах вращения

В простейшем случае в точках с координатами $C_{m,n}(\varphi)$ выполняется инкремент значений аккумуляторного массива:

$$A(C_{m,n}(\varphi)) = A(C_{m,n}(\varphi)) + 1. \quad (4)$$

При необходимости накопление в аккумуляторном массиве может производиться с учётом интенсивностей перепадов яркости в каждой точке поверхности, т. е.

$$A(C_{m,n}(\varphi)) = A(C_{m,n}(\varphi)) + J(ix_m + jy_m + kz_m). \quad (5)$$

При вращении обрабатываемого изображения происходит аналогичное вращение аккумуляторного массива, однако величина максимального отсчёта при этом не меняется, что позволяет выполнять распознавание трёхмерного изображения независимо от его ориентации. Поскольку при обработке происходит сопоставление каждой точки эталонного изображения с каждой точкой наблюдаемого объекта, результат обработки не будет зависеть от нумерации отсчётов, а также возможна обработка и принятие решений по отдельным фрагментам наблюдаемого изображения.

Если неизвестен масштаб изображения, то при заполнении аккумуляторного массива необходимо выполнять инкремент элементов массива вдоль образующей конуса, описываемой вектором, проходящим под углом w относительно нормали (рис. 5). Длина образующей

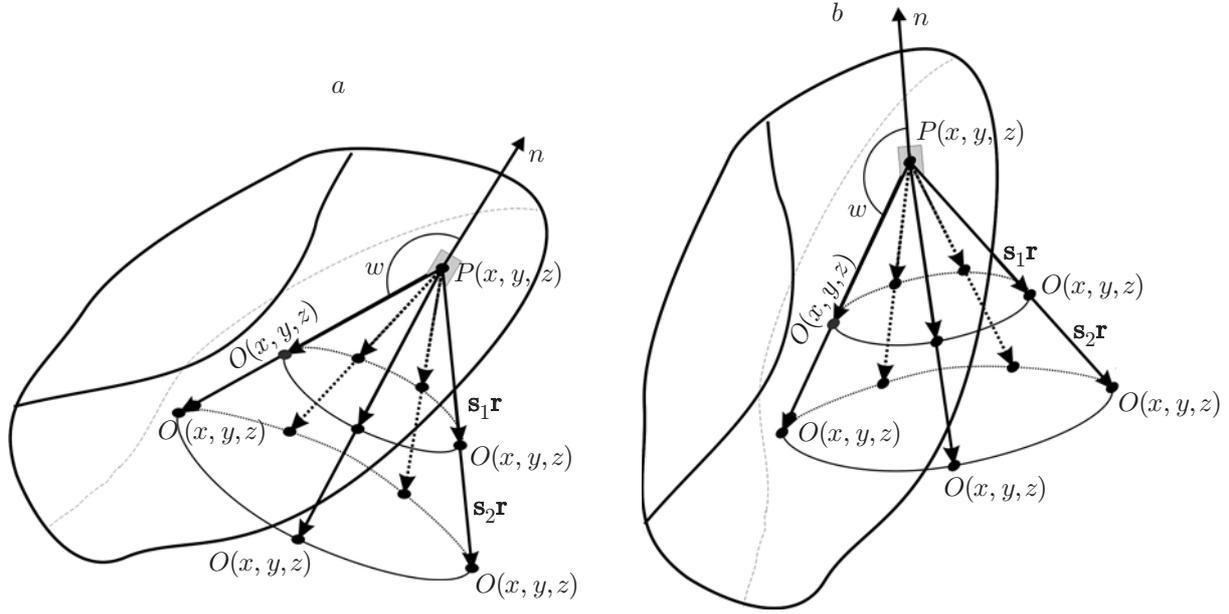


Рис. 5. Принцип голосования в аккумуляторном массиве при неизвестных параметрах вращения и масштабирования изображения

определяется возможным диапазоном изменения масштаба обрабатываемого изображения. Координаты точек, в которых инкрементируется содержимое аккумуляторного массива, вычисляются из соотношения

$$C_{m,n}(\varphi) = ix_m + jy_m + kz_m + \mu O_{n,m}(\varphi),$$

$$n = 0, 1, \dots, N_s - 1, \quad m = 0, 1, \dots, N_c - 1, \quad \mu \in [\mu_{\min}, \mu_{\max}], \quad \varphi \in [0, 2\pi), \quad (6)$$

где μ_{\min} и μ_{\max} — минимальное и максимальное значения масштаба изображения.

При наличии шума на изображении параметры центра формы находятся с некоторыми погрешностями, что вызывает «размывание» пика в аккумуляторном массиве. Дополнительные погрешности возникают из-за дискретности массива. Одним из подходов к устранению влияния этих факторов является сглаживание аккумуляторного массива за счёт его низкочастотной фильтрации, например квазилинейной. В простейшем случае может использоваться фильтр скользящего квазисреднего [8], работа которого описывается соотношением

$$\bar{A}_{i,j,k} = \left(\left(\sum_{z=k-N/2}^{k+N/2} \sum_{y=j-N/2}^{y+N/2} \sum_{x=i-N/2}^{i+N/2} A_{x,y,z} \right) / N^3 \right) A_{i,j,k}, \quad (7)$$

где N — размер окна фильтра; $A_{x,y,z}$ — текущее значение отсчёта в аккумуляторном массиве; $\bar{A}_{i,j,k}$ — значение отсчёта после сглаживания; $A_{i,j,k}$ — значение отсчёта до сглаживания.

А л г о р и т м модифицированного обобщённого преобразования Хафа состоит в следующем:

1. Для эталонов формируются R-таблицы, индексированные по номеру отсчёта с параметрами (\mathbf{r}, w) .
2. Выполняется инициализация трёхмерного аккумуляторного массива, размерность которого соответствует ожидаемому масштабу и положению центра изображения.

3. Для каждого отсчёта изображения $P_i(x, y, z)$ вычисляется вектор нормали n_i , $i = 0, \dots, N - 1$, где N — количество отсчётов в обрабатываемом изображении.
4. Для каждой j -й записи в R-таблице считаются параметры w_j и r_j , $j = 0, \dots, N_э - 1$, где $N_э$ — количество отсчётов в эталоне.
5. В соответствии с выражениями (1)–(6) производится инкремент ячеек аккумуляторного массива.
6. После перебора всех возможных комбинаций индексов i и j производится сглаживание аккумуляторного массива согласно (7).
7. В аккумуляторном массиве выполняется поиск максимума.
8. Пункты 2–7 повторяются для всех эталонов. Эталон, для которого получено наибольшее значение максимума в аккумуляторном массиве, принимается в качестве результата распознавания.

Формирование поля нормалей. При вычислении преобразования Хафа одной из промежуточных операций является нахождение нормали к данной точке поверхности. От качества выделения нормали во многом зависит правильность накопления отсчётов в аккумуляторном массиве и в конечном счёте качество обработки изображения. В компьютерной графике выделение нормалей, как правило, базируется на результатах предварительной триангуляции изображения [9], однако этот подход, с одной стороны, достаточно трудоёмкий, с другой — чувствителен к воздействию координатных шумов. Поэтому он применяется преимущественно для обработки изображений, получаемых путём компьютерного моделирования, а не сканированием реальных объектов. Здесь предлагается вычисление нормалей производить на основе анализа множества точек, попадающих в некоторый строб вокруг данной точки. Вычисление параметров нормали будет базироваться на формировании описания плоскости, проходящей вблизи точек в пределах выбранного сто́ба, методом наименьших квадратов. Рассмотрим данную процедуру подробнее.

Уравнение плоскости в общем случае может быть записано в виде

$$Ax + By + Cz + D = 0.$$

Если имеется набор из N точек, попавших в строб, то в соответствии с методом наименьших квадратов необходимо минимизировать сумму вида

$$\varepsilon = \sum_{n=0}^{N-1} (Ax_n + By_n + Cz_n + D)^2 \rightarrow \min.$$

Запишем частные производные по коэффициентам A, B, C, D и приравняем их к нулю:

$$\frac{\partial \varepsilon}{\partial A} = 2 \sum_{n=0}^{N-1} (Ax_n + By_n + Cz_n + D)x_n = 2 \sum_{n=0}^{N-1} (Ax_n^2 + By_n x_n + Cz_n x_n + Dx_n) = 0,$$

$$\frac{\partial \varepsilon}{\partial B} = 2 \sum_{n=0}^{N-1} (Ax_n + By_n + Cz_n + D)y_n = 2 \sum_{n=0}^{N-1} (Ax_n y_n + By_n^2 + Cz_n y_n + Dy_n) = 0,$$

$$\frac{\partial \varepsilon}{\partial C} = 2 \sum_{n=0}^{N-1} (Ax_n + By_n + Cz_n + D)z_n = 2 \sum_{n=0}^{N-1} (Ax_n z_n + By_n z_n + Cz_n^2 + Dz_n) = 0,$$

$$\frac{\partial \varepsilon}{\partial D} = 2 \sum_{n=0}^{N-1} (Ax_n + By_n + Cz_n + D) = 0.$$

Набор частных производных образует систему из четырёх линейных однородных уравнений, которую можно записать в матричном виде:

$$\begin{bmatrix} \sum_{n=0}^{N-1} x_n^2 & \sum_{n=0}^{N-1} y_n x_n & \sum_{n=0}^{N-1} z_n x_n & \sum_{n=0}^{N-1} x_n \\ \sum_{n=0}^{N-1} x_n y_n & \sum_{n=0}^{N-1} y_n^2 & \sum_{n=0}^{N-1} z_n y_n & \sum_{n=0}^{N-1} y_n \\ \sum_{n=0}^{N-1} x_n z_n & \sum_{n=0}^{N-1} y_n z_n & \sum_{n=0}^{N-1} z_n^2 & \sum_{n=0}^{N-1} z_n \\ \sum_{n=0}^{N-1} x_n & \sum_{n=0}^{N-1} y_n & \sum_{n=0}^{N-1} z_n & N \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

В результате решения системы [10] определяется набор коэффициентов A, B, C , нормированное значение которых задаёт направление нормали к плоскости, проходящей через точки, попавшие в сформированный строб.

Экспериментальное исследование эффективности модифицированного обобщённого преобразования Хафа. Для экспериментальной оценки эффективности предложенных алгоритмов обработки трёхмерных изображений на основе модифицированного преобразования Хафа было проведено распознавание зашумлённых изображений. В качестве эталонов использовались модели объектов из базы данных эталонного теста "Princeton Shape Benchmark" [11] (рис. 6). Одинаковые условия распознавания для всех моделей обеспечивали точечные описания их поверхностей одинаковой размерности.

Эксперимент состоит из следующих этапов:

1. Формируется алфавит из M эталонных объектов размерностью N_c отсчётов каждый.

2. Формируется набор из M R-таблиц.

3. Формируется изображение сигнального объекта путём вращения и масштабирования объекта, выбранного в качестве эталона. Производится зашумление его отсчётов аддитивным нормальным шумом с нулевым математическим ожиданием и дисперсией

$$\sigma^2 = E/qN_c, \text{ где } E = \sum_{n=0}^{N_c-1} |r_n|^2 \text{ — энергия сигнала, } q \text{ — отношение сигнал/шум:}$$

$$x^{(\xi)}(n) = x(n) + \xi_x(n), \quad y^{(\xi)}(n) = y(n) + \xi_y(n), \quad z^{(\xi)}(n) = z(n) + \xi_z(n).$$

4. В соответствии с вышеприведённым алгоритмом производится распознавание зашумлённого изображения.

5. Пункты 3 и 4 повторяются $N_{\text{оп}}$ раз, определяется общее количество правильных распознаваний $N_{\text{пр}}$ и вычисляется оценка вероятности правильного распознавания $\hat{p}_{\text{пр}} = N_{\text{пр}}/N_{\text{оп}}$.

6. Серии опытов по определению оценки вероятности правильного распознавания повторяются для различных значений отношения сигнал/шум. По их результатам строится график зависимости вероятности правильного распознавания от отношения сигнал/шум.

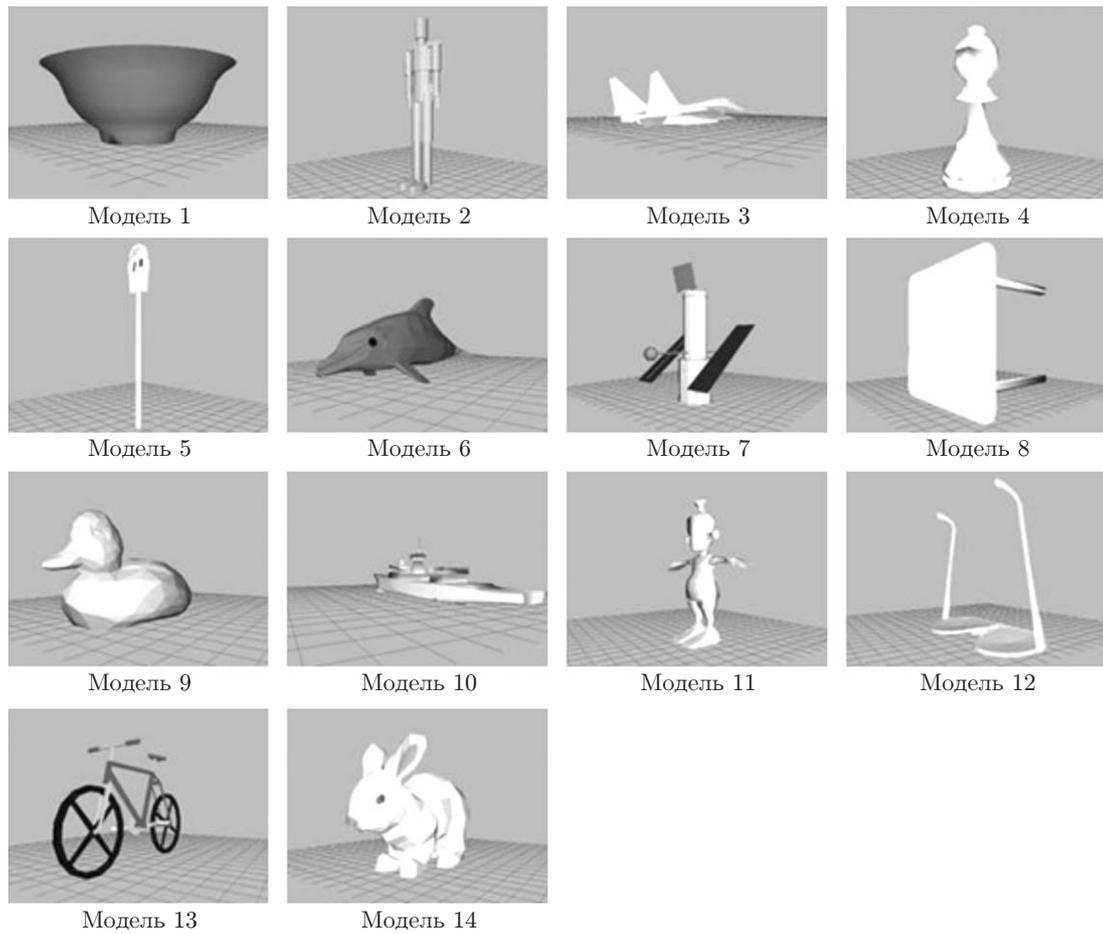


Рис. 6. Модели трёхмерных объектов, используемых в эксперименте

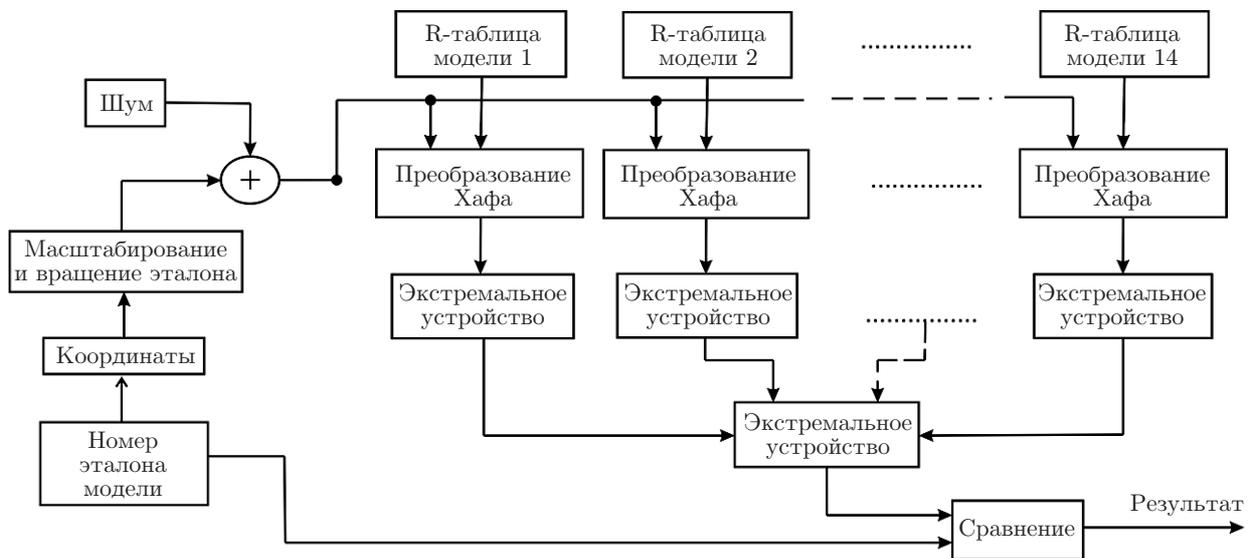


Рис. 7. Схема проведения эксперимента

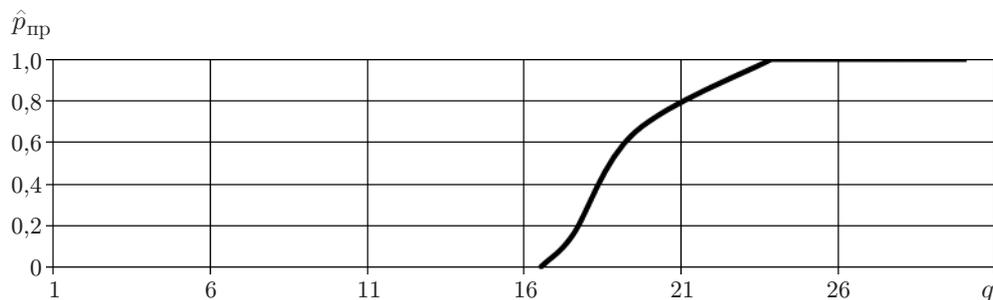


Рис. 8. Характеристика распознавания объекта с известными параметрами преобразований

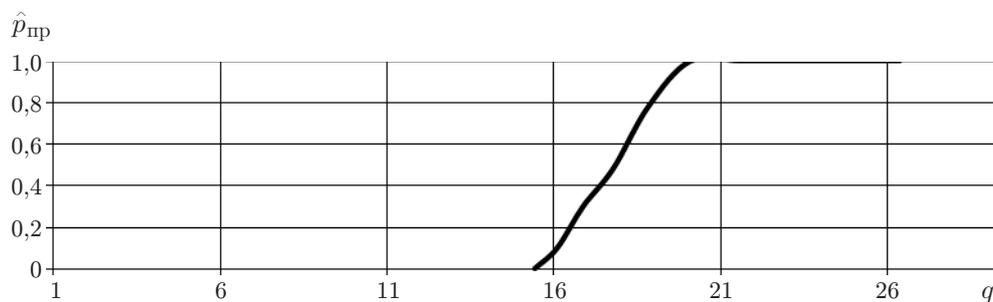


Рис. 9. Характеристика распознавания объекта с неизвестным масштабом

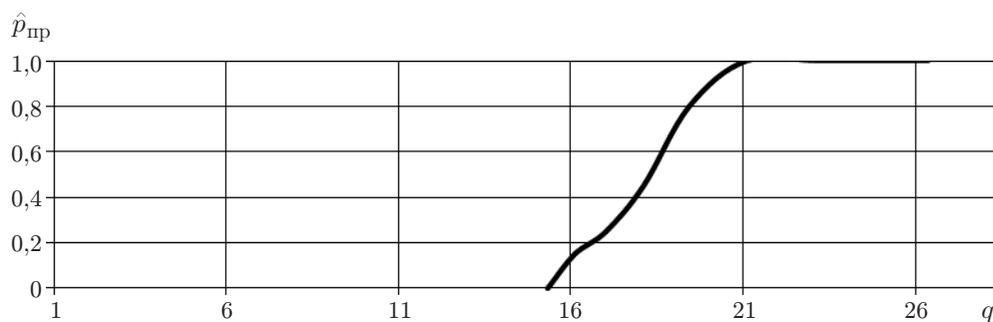


Рис. 10. Характеристика распознавания объекта, повернутого на угол 90° относительно оси $\{0,4; 0,6; 0,69\}$

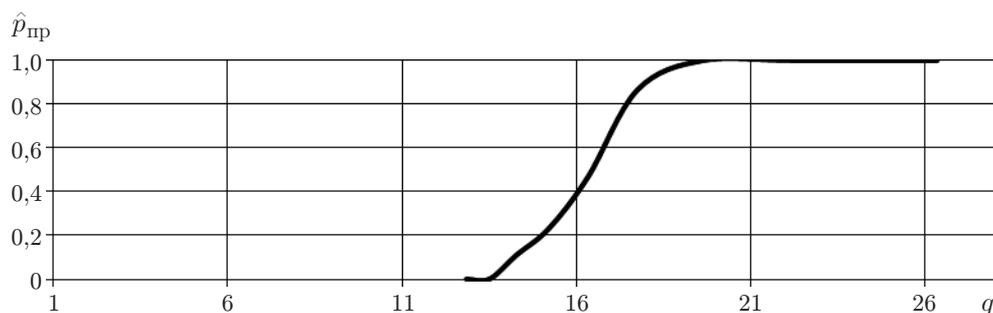


Рис. 11. Характеристика распознавания объекта, повернутого на угол 162° относительно оси $\{0,4; 0,6; 0,69\}$ с масштабом 1,3

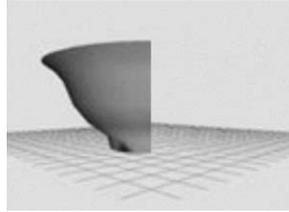


Рис. 12. Точечный фрагмент части модели 1, используемой при построении характеристики распознавания фрагмента объекта

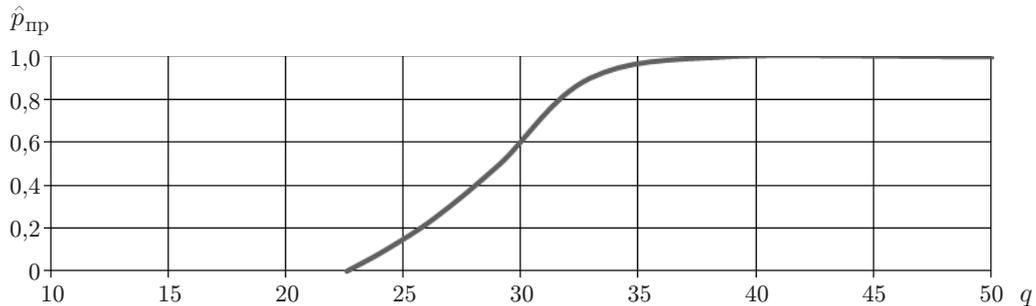


Рис. 13. Характеристика распознавания фрагмента объекта с известными параметрами преобразований

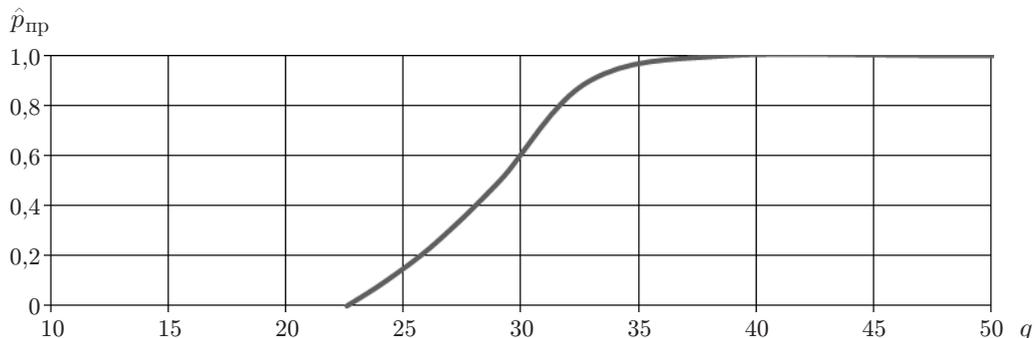


Рис. 14. Характеристика распознавания фрагмента объекта, повернутого на угол 144° относительно оси $\{0,4; 0,6; 0,69\}$ с масштабом 1,2

При проведении экспериментов (рис. 7) в качестве эталона была выбрана первая модель. На рис. 8–11 приведены результаты экспериментов для различных начальных условий. Видно, что вероятность правильного распознавания $p_{\text{пр}} > 0,9$ достигается при отношениях сигнал/шум на уровне 20–30.

При распознавании фрагмента объекта (рис. 12) вероятность $p_{\text{пр}} > 0,9$ достигается при отношениях сигнал/шум на уровне 35–40 (рис. 13, 14). Слабая зависимость графиков от параметров вращения и масштабирования свидетельствует об инвариантности разработанного алгоритма к этим преобразованиям.

Заключение. В данной работе предложено модифицированное обобщённое преобразование Хафа, обеспечивающее инвариантность результатов обработки к масштабированию и вращению трёхмерных изображений. Указанный результат достигается не за счёт значительного увеличения размерности аккумуляторного массива, а вследствие изменения принципа индексирования R-таблицы. Общая трудоёмкость алгоритма при этом квадратично зависит от количества отсчётов, задающих объект.

Благодаря своей структуре алгоритм обладает хорошими возможностями для распараллеливания, что делает его достаточно эффективным и в вычислительном плане, в том числе при реализации на графическом сопроцессоре.

Анализ эффективности алгоритма показал, что он обеспечивает вероятность правильного распознавания, близкую к характеристикам пространственно-корреляционных алгоритмов.

СПИСОК ЛИТЕРАТУРЫ

1. **Khoshelham K.** Extending generalized Hough transform to detect 3D objects in laser range data // ISPRS Workshop on Laser Scanning and SilviLaser. Espoo, Finland, 12–14 Sept., 2007. Vol. XXXVI, Pt. 3. P. 206–210.
2. **Bevilacqua V., Casorio P., Mastronardi G.** Extending Hough transform to a points' cloud for 3D-face nose-tip detection // Adv. Intell. Comput. Theor. and Appl. With Aspects of Artificial Intelligence. 2008. **5227**. P. 1200–1209.
3. **Woodford O. J.** Demisting the Hough transform for 3D shape recognition and registration // Proc. of the British Machine Vision Conference. Dundee, UK: BMVA Press, 2011. P. 32.1–32.11.
4. **Pham M.-T., Woodford O. J., Perbet F. et al.** A new distance for scale-invariant 3D shape recognition and registration // IEEE Intern. Conf. on Computer Vision (ICCV). 6–13 Nov., 2011. P. 145–152.
5. **Tombari F., Di Stefano L.** Object recognition in 3D scenes with occlusions and clutter by Hough voting // Proc. of the Fourth Pacific-Rim Symposium on Image and Video Technology (PSIVT'10). Singapore, 2010. Vol. 1. P. 349–355.
6. **Ballard D.** Generalizing the Hough transform to detect arbitrary shapes // Pattern. Recogn. 1981. **13**, N 2. P. 111–122.
7. **Кантор И. Л., Солодовников А. С.** Гиперкомплексные числа. М.: Наука, 1973. 144 с.
8. **Мироновский Л. А., Слаев В. А.** Алгоритмы оценивания результата трёх измерений. С.-Пб.: Професионал, 2010. 192 с.
9. **Gopi M., Krishnan S., Silva C. T.** Surface reconstruction based on lower dimensional localized Delaunay triangulation // Comput. Graphics Forum. 2000. **19**, N 3. P. 467–478.
10. **Ильин В. А., Позняк Э. Г.** Линейная алгебра. М.: Физматлит, 2004. 280 с.
11. **Shilane P., Min P., Kazhdan M., Funkhouser T.** The Princeton Shape Benchmark // Intern. Conf. on Shape Modeling and Applications. 2004. Vol. 1. P. 345–352.

Поступила в редакцию 30 октября 2012 г.
