

Сегментация изображений на основе RGBD данных

Е.В. Медведева¹, Е. А. Варко¹

¹Вятский государственный университет, Московская 36, Киров, Россия, 610000

Аннотация. Предлагается метод сегментации изображений на основе совместного использования данных о цвете и глубине. Метод состоит из двух этапов. На первом этапе выполняется сегментация RGB изображения на основе обнаружения контуров с последующим окрашиванием замкнутых областей. А на втором этапе – совместная сегментация данных о цвете и глубине. Данные о глубине позволяют различать пиксели с похожими яркостными характеристиками для разных объектов и повысить качество сегментации изображения. Для снижения вычислительных ресурсов предлагается выделять контуры по битовым плоскостям старших разрядов цифрового изображения с использованием математического аппарата двумерных цепей Маркова. Результаты моделирования подтверждают эффективность разработанного метода.

1. Введение

Для решения ряда задач, связанных с обнаружением и распознаванием статических и динамических объектов в таких областях как видеонаблюдение, автономное вождение и др., применяют сегментацию.

Традиционные методы сегментации сосредоточены в основном на использовании цветовых или яркостных признаков. В таких методах качество сегментации изображений существенно зависит от характера сцены: плавного или резкого изменения освещения; теней, создаваемых объектами; сложного фона и т. п. Поэтому, несмотря на огромное количество исследований ни один из существующих методов сегментации не позволил получить полностью удовлетворительные результаты, основываясь только на одних цветовых данных.

С разработкой новых датчиков RGBD, таких как Microsoft Kinect, которые обеспечивают синхронизированное видео глубины и цвета, открылись новые возможности для задач связанных с обнаружением и распознаванием объектов. В отличие от RGB данных, данные о глубине более устойчивы к изменениям освещения и динамическим фоновым объектам и могут являться эффективным дополнительным признаком для сегментации изображений.

В последнее время совместное использование данных о цвете и глубине стало новой темой исследования в области компьютерного зрения. В ряде работ предлагаются различные методы сегментации RGBD данных: на основе объединения алгоритмов вычитания фона с данными глубины [1]; на использовании сверточных нейронных сетей [2]; кластеризации [3]; с использованием контура, яркости и глубины [4] и др.

Однако практически все методы сегментации на основе объединения данных глубины и цвета имеют либо недостаточную гибкость, либо требуют значительных вычислительных ресурсов. Поэтому исследование в этой области является актуальной задачей.

Целью данной работы является разработка метода сегментации изображений на основе совместного использования данных яркости и глубины, позволяющего повысить качество сегментации при снижении вычислительных ресурсов.

2. Сегментация изображений на основе RGBD данных

В цветовом пространстве RGB каждая компонента является цифровым полутоновым изображением (ЦПИ), пиксели которого представлены g -разрядными двоичными числами. Компонента D также представляет собой многоразрядное цифровое изображение (карту глубины), в котором каждому элементу соответствует информация о расстоянии от камеры до каждой точки наблюдаемой сцены.

Реализовать сегментацию RGBD данных можно двумя способами: на первом этапе выполнить сегментацию на базе цветовых признаков, а на втором – на базе признаков глубины или наоборот. На первом этапе предпочтительнее использовать цветные данные. Это связано с рядом дефектов на карте глубины – потерянными и искаженными значениями глубины, неравномерными и шумными границами объектов, неверно измеренными значениями глубины для некоторых материалов с зеркальными или мелкозернистыми поверхностями и т.п. Поэтому использование на первом этапе данных глубины приведет к существенному искажению границ объектов и к разрыву контуров объектов – на втором этапе.

В работе на первом этапе выполняется сегментация RGB изображения. Для повышения точности выделения границ объектов интереса, используется метод на основе обнаружения контуров с последующим окрашиванием пикселей в замкнутых областях. А на втором этапе – совместная сегментация данных о цвете и глубине. Данные о глубине позволяют различать пиксели с похожими яркостными или цветовыми характеристиками для разных объектов и тем самым повысить качество сегментации изображения.

Цифровые полутоновые изображения, соответствующие цветовым компонентам, можно представить набором g разрядных двоичных изображений (РДИ). Наиболее информативные (детальные) области выделены на РДИ старших разрядов ЦПИ, а РДИ младших разрядов представляют собой изображения двумерного шума. Поэтому обнаружение контуров интересующих объектов предлагается выполнять по РДИ старших разрядов ЦПИ. Это позволяет использовать для обнаружения контуров математическую модель на основе двумерных цепей Маркова с двумя равновероятными состояниями и матрицами вероятностей переходов по горизонтали ${}^1\Pi = \left\| \pi_{ij}^{(l)} \right\|_{2 \times 2}$ и вертикали ${}^2\Pi = \left\| \pi_{ij}^{(l)} \right\|_{2 \times 2}$ ($i, j = \overline{1, 2}; l = \overline{1, g}$) [5,6]. Такой подход к обнаружению контуров позволит снизить вычислительные ресурсы за счет оперирования матрицами вероятностей переходов размером 2×2 .

На рисунке 1 представлен фрагмент РДИ с принятыми обозначениями элементов изображения.

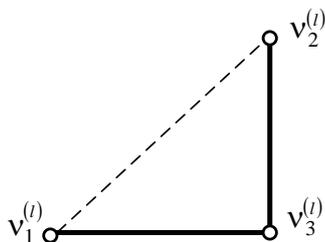


Рисунок 1. Фрагмент РДИ.

В соответствии с математической моделью двумерных цепей Маркова количество информации в элементе $v_3^{(l)}$ для различных сочетаний окрестных элементов $\Lambda_{i,j,k} = \{v_1^{(l)}, v_2^{(l)}\}$ определяется по формулам, представленным в таблице 1 [5,6].

Таблица 1. Выражения для вычисления количества информации в элементе $v_3^{(l)}$.

$v_1 v_2 \rightarrow v_3$	Выражения
00 → 0	$I(v_3^{(l)} v_1^{(l)}, v_2^{(l)}) = -\log \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}}$,
01 → 0	$I(v_3^{(l)} v_1^{(l)}, v_2^{(l)}) = -\log \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ij}^{(l)}}$,
M	M
11 → 0	$I(v_3^{(l)} v_1^{(l)}, v_2^{(l)}) = -\log \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}}$, (1)
00 → 1	$I(v_3^{(l)} v_1^{(l)}, v_2^{(l)}) = -\log \frac{{}^1\pi_{ii}^{(l)} {}^2\pi_{ii}^{(l)}}{{}^3\pi_{ii}^{(l)}}$,
M	M
11 → 1	$I(v_3^{(l)} v_1^{(l)}, v_2^{(l)}) = -\log \frac{{}^1\pi_{ij}^{(l)} {}^2\pi_{ij}^{(l)}}{{}^3\pi_{ii}^{(l)}}$,

где ${}^r\pi_{ij}^{(l)}$ ($i, j = \overline{1,2}; r = \overline{1,3}$) - элементы матриц вероятностей переходов в одномерных цепях Маркова с двумя состояниями - ${}^1\Pi^{(l)}$ (по горизонтали), ${}^2\Pi^{(l)}$ (по вертикали) и ${}^3\Pi^{(l)} = {}^1\Pi^{(l)} \times {}^2\Pi^{(l)}$.

Предполагается, что элементы матриц вероятностей переходов априорно известны и получены на основе большого числа выборок реальных изображений.

Решение о принадлежности анализируемого элемента точке контура осуществляется на основании сравнения вычисленной величины количества информации с порогом. Значение порога определяется как среднее значение между минимальным количеством информации и количеством информации, когда хотя бы один из элементов окрестности принимает другое состояние.

Для восьмиразрядного ЦПИ, представленного 256 значениями яркости, по РДИ старшего (8-го) разряда можно выделить все светлые области с яркостью от 128 до 255 на темном фоне, либо, наоборот, все темные объекты - на фоне с яркостью выше 128. Для выделения областей на менее контрастных изображениях с нечетко выраженными границами необходимо выделить контуры на следующих бинарных изображениях 7-го или 6-го разрядов ЦПИ. Контурное изображение, в этом случае, будет представлять сумму контурных изображений нескольких разрядов.

Предложенный метод выделения контуров требует небольших вычислительных ресурсов, определяемых операциями сравнения с двумя окрестными элементами. В результате предложенного алгоритма получается замкнутый контур толщиной в один пиксел. Это свойство является важным при выполнении следующей процедуры – окрашивания замкнутых областей.

Для окрашивания замкнутых областей задается диапазон значений яркости $[Y_{\min}; Y_{\max}]$, которые может принимать объект. Всем элементам внутри области объекта присваивается среднее значение яркости Y_{cp} (или метка с заданным значением). Для окрашивания областей

выбран построчный алгоритм заполнения с затравкой [7]. Он дает существенный выигрыш в объеме памяти и времени обработки за счет хранения только одного затравочного элемента для каждого заполняемого участка. В результате такой обработки изображения объект интереса может быть разбит на несколько частей или иметь неточные границы из-за неравномерной освещенности, наличия теней или бликов. Кроме того, на изображении наряду с объектами интереса могут быть выделены посторонние объекты на заднем плане сцены. Все эти факторы будут влиять на качество решения последующих задач - обнаружения, классификации и распознавания изображений.

На втором этапе задается диапазон значений данных $[X_{\min}; X_{\max}]$ на карте глубины, которые может принимать объект интереса, и формируется маска. Далее маска накладывается на результат сегментации RGB изображения и выполняется окончательный этап выделения объектов.

Данная процедура позволяет различать объекты, имеющие близкие яркостные или цветовые характеристики, но отличные характеристики по дальности, а также улучшить сегментацию объектов, при неравномерном освещении, наличии теней и т.п.

Блок-схема, поясняющая работу алгоритма, представлена на рисунке 2.

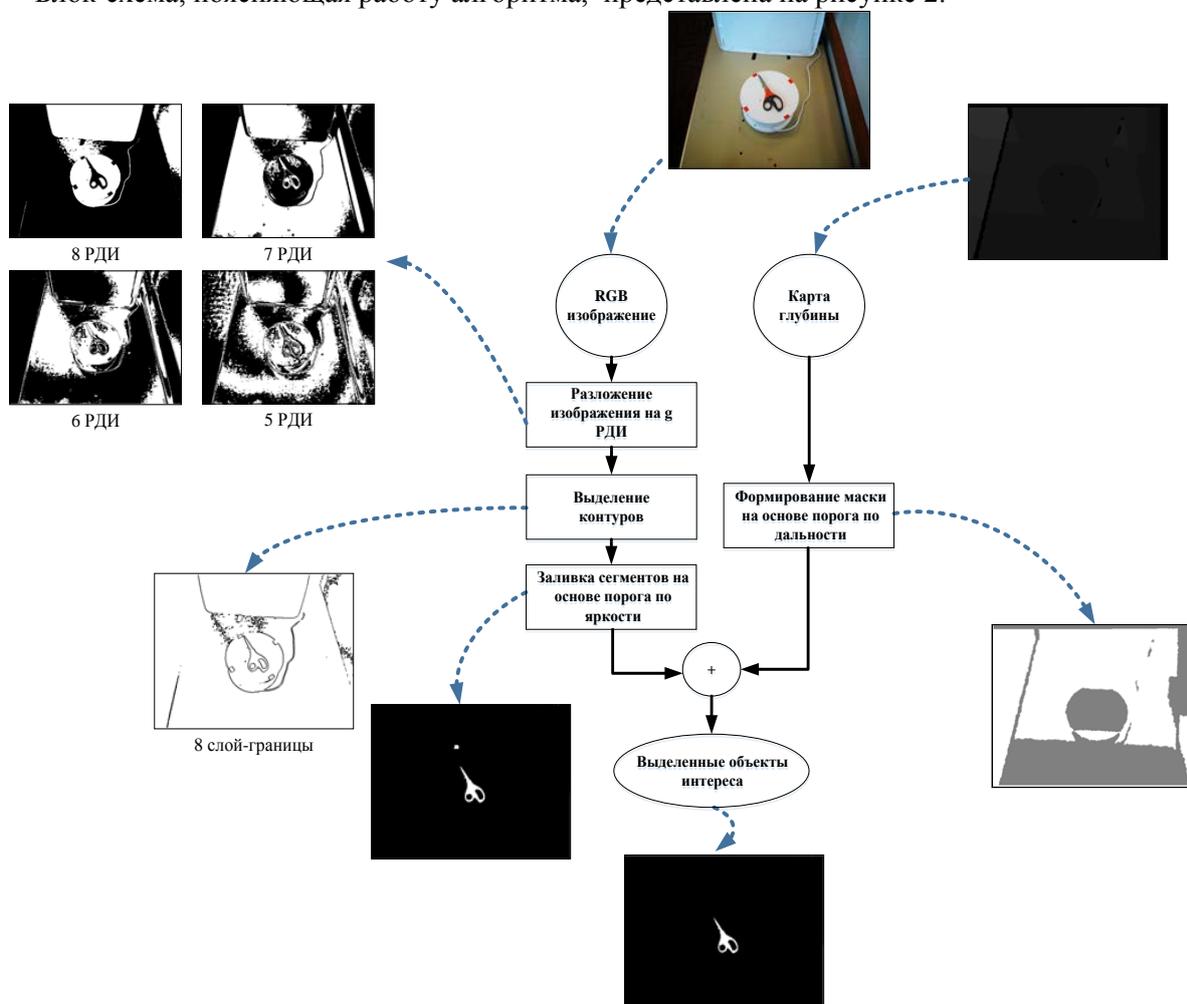


Рисунок 2. Блок-схема алгоритма совместной сегментации RGBD изображений.

3. Экспериментальные результаты

Для исследований использовался набор данных RGBD Object Dataset [8]. Набор данных RGBD содержит пары последовательностей цветных изображений и карт глубины, а также результаты сегментации на основе данных глубины и цвета, с применением алгоритма RANSAC и модели

адаптивной гауссовой смеси (АГС) [9]. Каждая видеопоследовательность состоит из 199 кадров размером 640×480 . На каждом изображении объектом интереса является только один предмет.

На рисунке 3 показаны примеры алгоритмов сегментации: (а) – исходное RGB изображение; (б) – эталонная разметка; (в) – сегментация с применением алгоритма RANSAC и АГС; (г) – результат алгоритма сегментации по данным яркости; (д) – результат совместной сегментации по данным яркости и глубины. Сегментация по яркости изображения «Apple» выполнена по R-компоненте, РДИ 8-го разряда; изображений «Banana» и «Scissors» - по В-компоненте, РДИ 8-го разряда; изображения «Coffee mug» - по G-компоненте, РДИ 8-го разряда.

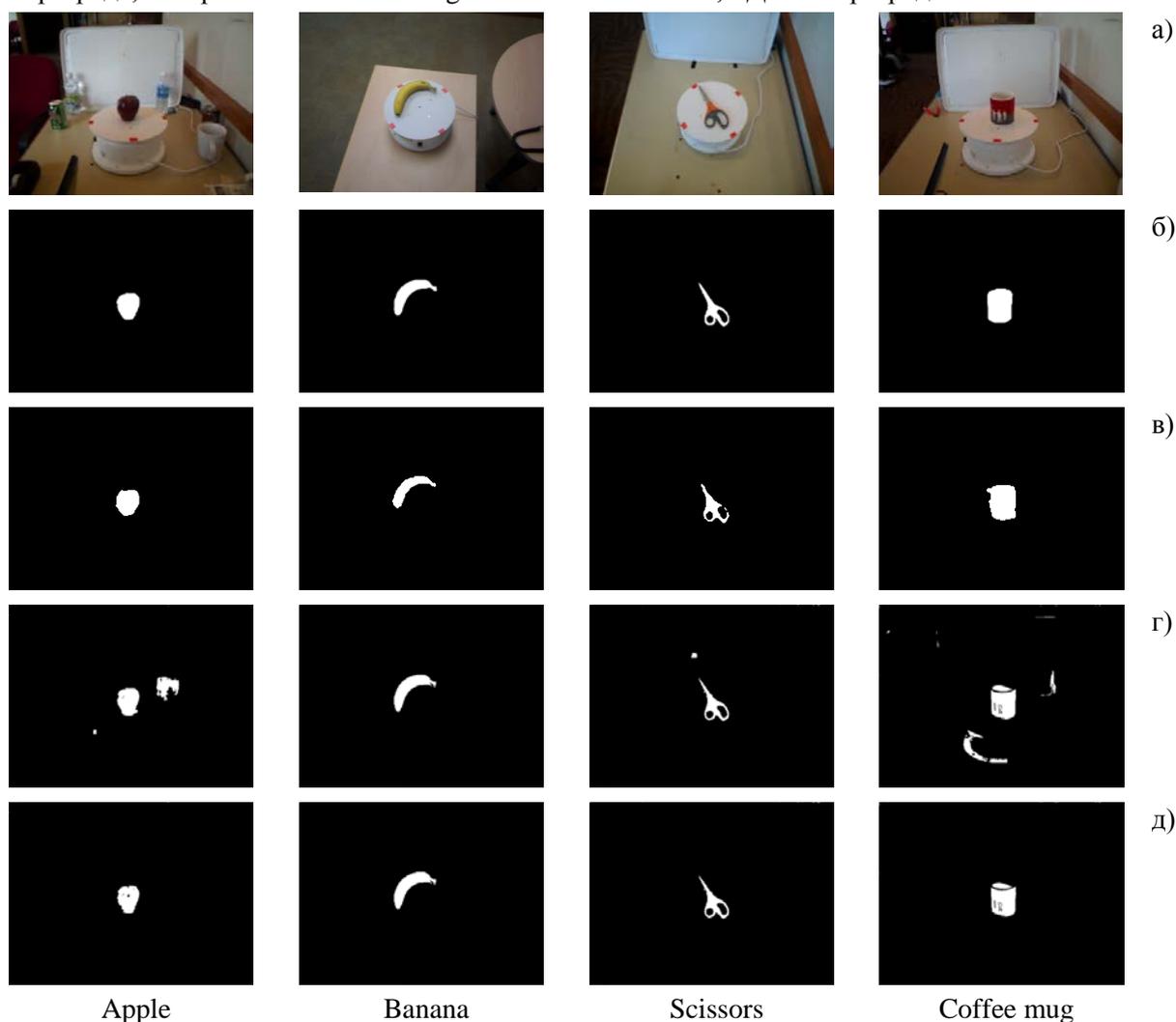


Рисунок 3. Сравнение методов сегментации RGBD данных.

Из приведенных результатов (рис.3г) видно, что алгоритм сегментации на основе обнаружения контуров точно локализует границы объектов. Дополнительное использование данных глубины (рис.3д), позволяет улучшить качество сегментации: убрать выделенные фрагменты близкие по яркости объекту интереса, избавиться от теней и т.п. Также из сравнения результатов на рис. 3г и рис.3д видно, что разработанный метод позволяет точнее выделить объекты интереса, чем метод, предложенный в работе [9].

Для типовых изображений (или последовательности видеок кадров), на которых объекты интереса имеют близкие характеристики по яркости и глубине, процесс сегментации может быть выполнен автоматически.

Для оценки качества сегментации использовались критерии точности (P), полноты (R) [10] и коэффициент ошибки (E):

$$P = \frac{TP}{TP + FP}, \quad (2)$$

$$R = \frac{TP}{TP + FN}, \quad (3)$$

$$E = \frac{FP + FN}{TP + TN + FP + FN}, \quad (4)$$

где TP — истинно-положительное решение; TN — истинно-отрицательное решение; FP — ложно-положительное решение; FN — ложно-отрицательное решение.

Показатель точности системы в пределах сегментированной области – это доля пикселей, действительно принадлежащих данной области относительно всех пикселей, которые отнесены к этой области. Показатель полноты измеряет долю всех истинно определенных пикселей, принадлежащих сегментированной области, относительно всех пикселей. Коэффициент ошибки E учитывает все ошибочные пиксели к общему числу пикселей. Для вычисления оценок точности, полноты сегментации и коэффициента ошибки использовались эталонные изображения сегментации.

В таблице 2 представлены результаты оценок качества сегментации разработанным методом и известным [9]. Оценки вычислялись по отдельным изображениям и усреднялись по всей видеопоследовательности.

Таблица 2. Оценки результатов алгоритмов сегментации.

Видео последовательности	на основе яркости			на основе яркости и глубины			На основе RANSAC и модели АГС [9]		
	P	R	E	P	R	E	P	R	E
Apple	0.93	0.58	0.0086	0.89	0.98	0.0015	0.90	0.97	0.0013
Banana	0.93	0.93	0.0018	0.93	0.93	0.0018	0.80	0.97	0.0027
Scissors	0.80	0.91	0.0026	0.82	0.97	0.0021	0.78	0.93	0.0026
Coffee mug	0.83	0.63	0.0118	0.98	0.80	0.0030	0.95	0.94	0.0014
Comb	0.89	0.40	0.0237	0.91	0.90	0.0035	0.85	0.94	0.0035

Совместная сегментация имеет близкие значения оценок точности с оценками для яркостной сегментации, но увеличивает показатель полноты (до 2.1 раз) и уменьшает ошибку сегментации (до 5.7 раз).

4. Заключение

Предложенный метод сегментации изображений на основе совместного использования данных о цвете и глубине позволяет точно выделять границы объектов интереса и эффективно различать пиксели с похожими яркостными характеристиками для разных объектов. Выделение контуров по битовым плоскостям старших разрядов цифрового изображения с использованием математического аппарата двумерных цепей Маркова позволяет снизить вычислительные ресурсы на реализацию алгоритма. Алгоритм может быть применен решения ряда задач, связанных с обнаружением и распознаванием объектов в системах видеонаблюдения, автономного вождения и т.п.

5. Литература

- [1] Trabelsi, R. Efficient and fast multi-modal foreground-background segmentation using RGBD data / R. Trabelsi, I. Jabri, F. Smach, A. Bouallegue // Pattern Recognition Letters. – 2017. – Vol. 97. – P. 13-20.

- [2] Qi, X. 3D Graph Neural Networks for RGBD Semantic Segmentation / X. Qi, R. Liao, J. Jia, S. Fidler, R. Urtasun // IEEE International Conference on Computer Vision – Venice, Italy, 2017. – P. 5199-5208.
- [3] Yalic, H. Automatic Object Segmentation on RGB-D Data using Surface Normals and Region Similarity / H. Yalic, A. Can // Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. – 2018. – Vol. 4. – P. 379-386.
- [4] Xiang, F. Image segmentation using contour, surface, and depth cues. / F. Xiang, C. Chen, C. Wang, C.-C. Jay Kuo // IEEE International Conference on Image Processing, 2017.
- [5] Petrov, E.P. Mathematical Models of Video-Sequences of Digital Half-Tone Images / E.P. Petrov, I.S. Trubin, E.V. Medvedeva, S.M. Smolskiy // Integrated models for information communication systems and net-works : design and development – GI Global, 2013. – P. 207-241.
- [6] Medvedeva, E.V. Image Segmentation Based on Two-Dimensional Markov Chains / E.V. Medvedeva, E.E. Kurbatova // Computer Vision in Control Systems-2. Innovations in Practice. Springer International Publishing Switzerland, 2015. – P. 277-295.
- [7] Павлидис, Т. Алгоритмы машинной графики и обработки изображений / Т. Павлидис – М.: Радио и связь, 1986. – 394 с.
- [8] Набор данных RGBD Object Dataset [Электронный ресурс]. – Режим доступа: <http://rgbd-dataset.cs.washington.edu/dataset.html> (23.05.2019).
- [9] Lai, K. A Large-Scale Hierarchical Multi-View RGB-D Object Dataset / K. Lai, L. Bo, X. Ren, D. Fox // IEEE International Conference on Robotics and Automation, 2011.
- [10] Рассел, С. Искусственный интеллект. Современный подход / С. Рассел, П. Норвинг – М.: Издательский дом «Вильямс», 2006. – 1408 с.

Image segmentation based on RGBD data

E.V. Medvedeva¹, E.A. Varko¹

¹Vyatka State University, Moskovskaya str. 36, Kirov, Russia, 610000

Abstract. A method for image segmentation based on the combined use of color and depth data is proposed. The method consists of two stages. At the first stage, RGB image segmentation is performed based on the detection of contours with the subsequent filling of closed areas. The second stage is the joint segmentation of color and depth data. Depth data allows you to distinguish between pixels with similar brightness characteristics for different objects and improve the quality of image segmentation. To reduce computational resources, it is proposed to detect contours in binary images of high-order bits of a digital grayscale image using the mathematical apparatus of two-dimensional Markov chains. The simulation results prove the effectiveness of the developed method.