

ISSN 1561-8323 (Print)
ISSN 2524-2431 (Online)

ИНФОРМАТИКА INFORMATICS

УДК 004
<https://doi.org/10.29235/1561-8323-2021-65-3-269-274>

Поступило в редакцию 07.05.2021
Received 07.05.2021

Б. А. Залесский

*Объединенный институт проблем информатики Национальной академии наук Беларуси,
Минск, Республика Беларусь*

МНОГОУРОВНЕВЫЙ АЛГОРИТМ ЦВЕТОВОЙ КЛАСТЕРИЗАЦИИ ИЗОБРАЖЕНИЙ

(Представлено членом-корреспондентом А. В. Тузиковым)

Аннотация. Представлен многоуровневый алгоритм цветовой кластеризации МАСС (Multilevel Algorithm for Color Clustering), предназначенный для быстрой кластеризации изображений. В настоящее время для цветовой кластеризации изображений активно используется несколько хорошо известных алгоритмов, в том числе k -средних (который является одним из наиболее часто используемых при обработке данных) и его нечеткие версии, водораздела, наращивания областей и целая серия новых более сложных нейросетевых и других алгоритмов. Однако их невозможно применять для кластеризации больших цветных изображений в режиме реального времени. Быстрая кластеризации бывает необходима, например, при обработке кадров видеопотока, создаваемого различными видеокамерами или при работе с большими базами данных изображений. Разработанный алгоритм МАСС позволяет выполнить на персональном компьютере кластеризацию больших изображений, например размера FullHD, по цвету со средним отклонением от исходных значений цвета около пяти единиц менее, чем за 20 мс, в то время как параллельная версия классического алгоритма k -средних выполняет кластеризацию этих же изображений со средней ошибкой более 12 единиц за время, превышающее 2 с. Предложенный алгоритм многоуровневой кластеризации изображений по цвету достаточно прост в реализации. Он был протестирован на большом количестве цветных изображений.

Ключевые слова: кластеризация, цветные изображения, многоуровневый алгоритм

Для цитирования. Залесский, Б. А. Многоуровневый алгоритм цветовой кластеризации изображений / Б. А. Залесский // Докл. Нац. акад. наук Беларуси. – 2021. – Т. 65, № 3. – С. 269–274. <https://doi.org/10.29235/1561-8323-2021-65-3-269-274>

Boris A. Zalesky

United Institute of Informatics Problems of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus

MULTILEVEL ALGORITHM FOR COLOR CLUSTERING OF IMAGES

(Communicated by Corresponding Member Alexander V. Tuzikov)

Abstract. The fast multilevel algorithm to cluster color images (MASS – Multilevel Algorithm for Color Clustering) is presented. Currently, several well-known algorithms of image clustering, including the k -means algorithm (which is one of the most commonly used in data mining) and its fuzzy versions, watershed, region growing ones, as well as a number of new more complex neural network and other algorithms are actively used for image processing. However, they cannot be applied for clustering large color images in real time. Fast clustering is required, for example, to process frames of video streams shot by various video cameras or when working with large image databases. The developed algorithm MASS allows the clustering of large images, for example, FullHD size, on a personal computer with an average deviation from the original color values of about five units in less than 20 milliseconds, while a parallel version of the classical k -means algorithm performs the clustering of the same images with an average error of more than 12 units for a time exceeding 2 seconds. The proposed algorithm of multilevel color clustering of images is quite simple to implement. It has been extensively tested on a large number of color images.

For citation. Zalesky B. A. Multilevel algorithm for color clustering of images. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2021, vol. 65, no. 3, pp. 269–274 (in Russian). <https://doi.org/10.29235/1561-8323-2021-65-3-269-274>

Введение. Методы и алгоритмы кластеризации изображений являются инструментом, хорошо известным в компьютерном зрении, который применяется при решении многих задач, таких, например, как обнаружение, распознавание, отслеживание объектов, признаковое описание сцен и т. д.

Среди наиболее часто применяемых в компьютерном зрении методов и алгоритмов кластеризации – алгоритм k -средних [1; 2] и его многочисленные нечеткие [3] и нейросетевые версии [4–6], метод водораздела [7] и т. д. Несколько версий алгоритма нечетких k -средних со штрафной функцией, позволяющей учитывать гладкость изображения и границы однородных областей, предложены в [8]. С актуальными статьями, посвященными кластеризации изображений и данных, можно ознакомиться в [9].

Тем не менее, программные реализации известных алгоритмов слишком медленны для решения ряда задач реального времени, например, кластеризации кадров видеопотока с разрешением HD и выше или работы с большими базами данных изображений.

Разработанный многоуровневый алгоритм МАСС предназначен для быстрой кластеризации изображений по цвету. Программная реализация алгоритма позволяет выполнить на CPU персонального компьютера (без использования графического ускорителя) кластеризацию больших цветных изображений размера FullHD по цвету со средним отклонением от исходных значений цвета около пяти единиц менее чем за 20 мс, в то время как параллельная версия классического алгоритма k -средних выполняет кластеризацию этих же изображений со средней ошибкой более 12 единиц за время, превышающее 2 с.

Алгоритм достаточно просто реализовать программно, однако он имеет особенность (которая, впрочем, присуща и другим алгоритмам кластеризации, например, алгоритмам k -средних и водораздела) – цвета полученного кластерного представления могут быть близкими или даже вообще совпадать. Для устранения этой особенности достаточно выполнить несложную операцию слияния кластеров, которая не требует большого объема вычислений и занимает мало времени. Описание операции приведено ниже.

Тестирование алгоритма проводилось на видеопоследовательностях, снятых различными видеокамерами, а также на общеизвестных наборах данных UAV123 и VOT2015. Алгоритм МАСС прошел проверку в составе разработанного автором алгоритма АСТ отслеживания объектов, наблюдаемых видеокамерой [10].

Алгоритм МАСС предназначен для быстрой кластеризации полутоновых и цветных изображений по цвету. Приведем его описание для трехканальных RGB-изображений, в которых цветовое пространство представлено $\mathbb{C} = \{0, \dots, 255\}^3$ кубом.

Обозначим через $S = \{(x, y)\}$, $x = 0, \dots, w-1$, $y = 0, \dots, h-1$, множество пикселей \mathbf{p} RGB-изображения $\mathbf{I} = \{I(\mathbf{p})\}$, $\mathbf{p} \in S$, размера $w \times h$ с векторами цвета $I(\mathbf{p}) = (I_R(\mathbf{p}), I_G(\mathbf{p}), I_B(\mathbf{p}))$. Разобьем множества значений каждого цветового канала $\{0, \dots, 255\}$ на $L = 2^k$ равных интервала длиной $h = 256 / L = 2^{8-k}$ (параметр L может принимать одно из девяти значений 1, 2, ..., 256). Разбиение цветовых каналов на равные интервалы образует разбиение цветового пространства \mathbb{C} на кубы $P_{j(1)}(1)$, содержащие h^3 векторов цвета, которые будем нумеровать мультииндексами $\mathbf{j}(1) = (j_1(1), j_2(1), j_3(1))$, $j_i(1) \in \{0, \dots, L-1\}$, такие что $P_{j(1)}(1) \cap P_{j'(1)}(1) = \emptyset$, если $\mathbf{j}(1) \neq \mathbf{j}'(1)$ и $\bigcup P_{j(1)}(1) = \mathbb{C}$ (несколько тяжеловесная нумерация используется для построения многоуровневой пирамиды кубов).

Определим традиционную многоуровневую пирамиду кубов следующим образом. Для каждой тройки четных индексов $j_i(1)$, $i = 1 \div 3$, объединим каждые восемь кубов $P_{j(1)}(1)$ с индексами либо равными, либо на единицу большими $j_i(1)$, в кубы $P_{j(2)}(2)$, содержащие $(2h)^3$ векторов цвета так, что для мультииндексов $\boldsymbol{\mu} = (\mu_1, \mu_2, \mu_3)$ с координатами μ_i , принадлежащими двухэлементным множествам $\mu_i \in \{j_i, j_i + 1\}$,

$$P_{j(2)}(2) = \bigcup_{\boldsymbol{\mu}} P_{\boldsymbol{\mu}}(1),$$

а мультииндексы связаны соотношением $j(2) = 0,5j(1)$ и, следовательно, $j_i(2) = \left\{0, \dots, \frac{L}{2} - 1\right\}$, т. е. кубов второго уровня в восемь раз меньше, чем кубов первого уровня.

Аналогично образуем кубы более высоких уровней таким образом, что

$$P_{j(m+1)}(m+1) = \bigcup_{\mu} P_{\mu}(m), m \in \{1, \dots, k\}.$$

Число кубов на m -м уровне пирамиды равно $2^{3(k-m+1)}$ и верхний $(k+1)$ -й уровень пирамиды содержит всего один куб $P_{j(k+1)}(k+1) = \mathbb{C}$.

Приведем краткое описание шагов выполнения алгоритма МАСС.

Первый шаг алгоритма состоит в задании его параметров:

– числа N , ограничивающего сверху количество кластеризованных цветов изображения, так как заранее невозможно установить точное значение числа кластеров. Например, если число цветов K входного изображения равно единице или просто $K < N$, число цветов кластеризованного по цвету изображения может быть только меньше или равно K ;

– числа интервалов значений цвета каждого канала $L = 2^k$, которое должно быть степенью двойки, задающего длину ребра наименьшего куба многоуровневого представления RGB-цветового пространства, равную $256 / L$. Параметры N и L должны удовлетворять условию $L \geq 0,5\sqrt[3]{7N+1}$, вытекающему из требования того, чтобы число кубов $P_{j(m)}(m)$ в пирамиде было не меньше числа кластеров N . Данное условие не является обременительным для использования алгоритма, так как число арифметических операций, необходимых для вычислений, есть величина $O(wh + L^3)$. Например, в случае $N = 500$ число интервалов должно удовлетворять условию $L \geq 8$. При реальном применении алгоритма число интервалов L выбирается равным 16, 32 или 64;

– числа $\varepsilon > 0$, задающего условие слияния близких по цвету кластеров, в случае выполнения необязательного шага 7 алгоритма.

На *втором шаге* алгоритма нужно вычислить средние арифметические $C_{j(1)}(1)$ векторов цвета изображения \mathbf{I} , попадающих в каждый непустой куб $P_{j(1)}(1)$ первого уровня:

$$C_{j(1)}(1) = (N_{j(1)}(1))^{-1} \sum_{I(\mathbf{p}) \in P_{j(1)}(1)} I(\mathbf{p}),$$

где $N_{j(1)}(1) = |\{I(\mathbf{p}) : I(\mathbf{p}) \in P_{j(1)}(1)\}|$ число векторов цвета $I(\mathbf{p})$ в кубе $P_{j(1)}(1)$. Для пустых кубов достаточно сохранить в памяти значение $N_{j(1)}(1) = 0$, которое используется на последующих шагах алгоритма.

Третий шаг алгоритма заключается в рекуррентном вычислении числа векторов цвета в кубах

$$N_{j(m+1)}(m+1) = \sum^* N_{j(m)}(m).$$

В каждой сумме \sum^* восемь слагаемых. Суммирование ведется по индексам $j(m)$, удовлетворяющим условию

$$P_{j(m)}(m) \in P_{j(m+1)}(m+1).$$

На *четвертом шаге* вычисляются средние арифметические векторов цвета изображения \mathbf{I} , попадающих в каждый непустой куб $P_{j(m+1)}(m+1)$, $m \in \{1, \dots, k\}$, (в этом случае $N_{j(m+1)}(m+1) > 0$), по формуле

$$C_{j(m+1)}(m+1) = (N_{j(m+1)}(m+1))^{-1} \sum^* N_{j(m)}(m) C_{j(m)}(m).$$

После завершения четвертого шага оказывается сформированным множество векторов средних цветов

$$\mathbf{Q} = \{C_{(0,0,0)}(1), \dots, C_{j(m)}(m), \dots, C_{(0,0,0)}(k+1)\},$$

содержащее $7^{-1}(8L^3 - 1)$ элементов, среди которых могут быть нулевые векторы $C_{j(m)}(m)$, соответствующие пустым кубам.

Пятый шаг МАСС состоит в удалении из множества \mathbf{Q} нулевых векторов $C_{j(m)}(m)$, соответствующих пустым кубам, путем проверки условия $N_{j(m)}(m) = 0$. После выполнения данного шага в множестве \mathbf{Q} остаются векторы реально осредненных цветов.

На *шестом шаге* алгоритма множество \mathbf{Q} сортируется так, чтобы в отсортированном $\mathbf{Q}(\text{sort})$ на первом месте оказался $C_{j(m)}(m)$, соответствующий наибольшему значению $N_{j(m)}(m)$, на втором – соответствующий второму по величине значению $N_{j(\ell)}(\ell)$ и так далее в порядке убывания чисел $N_{j(i)}(i)$.

Полученное множество отсортированных осредненных цветов $\mathbf{Q}(\text{sort})$ может быть гораздо больше требуемого числа кластеров N . Уменьшить его размерность можно путем последовательного «слияния» близких значений $C_{j(m)}(m)$, начиная с тех, которым соответствуют наибольшие значения $N_{j(m)}(m)$.

Для этого можно использовать *седьмой шаг* алгоритма (необязательный), который заключается в просмотре множества $\mathbf{Q}(\text{sort})$, начиная с проверки условия близости его первого элемента $C_{(0,0,0)}(1)$ к другим последовательно выбранным элементам $C_{j(m)}(m)$

$$\|C_{(0,0,0)}(1) - C_{j(m)}(m)\| \leq \varepsilon \quad (1)$$

для наперед заданного $\varepsilon > 0$. В случае первого выполнения условия (1) два элемента в множестве $\mathbf{Q}(\text{sort})$ заменяются одним по формуле

$$C'_{(0,0,0)}(1) = (N_{(0,0,0)} + N_{j(m)}(m))^{-1} (N_{(0,0,0)}(1)C_{(0,0,0)}(1) + N_{j(m)}(m)C_{j(m)}(m))$$

и действия повторяются пока не будет достигнут последний элемент преобразованного множества $\mathbf{Q}(\text{sort})$.

На *восьмом шаге* алгоритма из множества $\mathbf{Q}(\text{sort})$ (преобразованного на седьмом шаге) выбираются N элементов, соответствующих самым большим числам $N_{j(m)}(m)$.

Результаты экспериментов. Для исследования характеристик построенного алгоритма МАСС было проведено его тестирование на изображениях разных размеров, начиная от 320×240 до 1920×1080 (Full HD), снятых различными камерами. Помимо этого, проводилось тестирование на общеизвестном наборе тестовых видеопоследовательностей UAV123 и VOT2015. Алгоритм МАСС также был успешно использован при программной реализации разработанного ранее алгоритма АСТ [10] отслеживания объектов, наблюдаемых видеокамерой, вместо классического алгоритма k -средних.

В таблице приведены результаты тестирования МАСС на 200 изображениях, большая часть которых была взята из тестового набора VOT2015, а остальные сняты камерой дрона DJI Phantom 3 Pro. Для сравнения в этой же таблице приведены результаты использования алгоритма k -средних на этих же изображениях.

Результаты тестирования МАСС

Test results МАСС

Алгоритм Algorithm	Среднее время, мс Average time, ms	Среднее количество итераций Average number of iterations	Средняя ошибка, $N_{\text{clust}} = 20$ Average error, $N_{\text{clust}} = 20$	Средняя ошибка, $N_{\text{clust}} = 40$ Average error, $N_{\text{clust}} = 40$	Средняя ошибка, $N_{\text{clust}} = 2000$ Average error, $N_{\text{clust}} = 2000$
k -средних, RAND	2492,34	200	13,81	–	–
k -средних, FOUND	1067,80	90	11,28	–	–
МАСС	4,36	–	17,02	11,68	4,80

В качестве критериев были выбраны средняя погрешность приближения цветов исходного изображения \mathbf{I} его кластерным представлением $\mathbf{I}_{\text{clust}}$, которая вычислялась по формуле $\delta = (wh)^{-1} \|\mathbf{I} - \mathbf{I}_{\text{clust}}\|$ относительно евклидовой нормы $\|\cdot\|$ (напомним, что w и h – длина и высота изображения). Для сравнения были взяты кластерные представления, полученные алгоритмом МАСС и параллельной версией алгоритма k -средних, реализованной на CPU с параметрами, рекомендованными в библиотеке компьютерного зрения OpenCV.

Были выбраны следующие значения алгоритма МАСС: размерность множества выбранных центров кластеров $N = 2000$, количество интервалов разбиения каждого цветового канала $L = 16$. Слияние кластеров не производилось, поэтому параметр ε не использовался.

Алгоритм k -средних был использован со значениями:

количество повторений алгоритма 3;

максимальное число итераций 200;

число кластеров 20.

Из таблицы видно, что даже с выбранными значениями время применения алгоритма k -средних достаточно велико. Увеличение значений приводит к вычислительным затратам, недопустимым для решения многих задач компьютерного зрения.

Для более детального сравнения алгоритм k -средних использовался в режиме, условно названном RAND, со случайно задаваемыми начальными значениями центров кластеров (один из стандартных способов его применения), и в режиме FOUND, когда в качестве 20 начальных значений кластеров выбирались 20 первых элементов множества $\mathbf{Q}(\text{sort})$, построенного алгоритмом МАСС.

Напомним, что в таблице приведены результаты, полученные при кластеризации 200 RGB-изображений. Число кластеров N_{clust} для МАСС означает, что после первоначального построения множества $\mathbf{Q}(\text{sort})$ размера N из него были выбраны N_{clust} первых (самых больших) кластеров. Время вычисления приведено в миллисекундах.

На основании результатов тестов, приведенных в таблице, можно сделать следующие выводы: 1) на тестовых изображениях предложенный алгоритм МАСС оказался намного быстрее параллельной версии классического алгоритма k -средних; 2) время работы МАСС зависит от предварительно выбранного ограничения на количество кластеров N и практически не зависит от окончательного выбора количества кластеров N_{clust} ; 3) при использовании одинакового количества кластеров, предложенный алгоритм на 20 % уступает по точности алгоритму k -средних, но при этом он во много раз быстрее; 4) выбор из множества $\mathbf{Q}(\text{sort})$, построенного алгоритмом, большего числа кластеров N_{clust} позволяет получить более точное приближение исходного изображения, по сравнению с построенным алгоритмом k -средних, без увеличения времени выполнения МАСС; 5) использование значений центров кластеров, найденных алгоритмом МАСС, в качестве начальных значений для алгоритма k -средних позволило в 100 % случаев повысить точность последнего.

Заключение. Проведенные эксперименты с изображениями, снятыми различными камерами, а также его использование в ранее разработанном алгоритме АСТ [10] вместо алгоритма k -средних продемонстрировали применимость алгоритма для цветовой кластеризации RGB-изображений. Положительными особенностями алгоритма являются простота его реализации и быстрое действие, достаточное для его применения в приложениях реального времени.

Для уменьшения количества центров кластеров, используемых для кластеризации изображений, имеет смысл применять процедуру слияния близких по цвету кластеров (которая не является вычислительно затратной) при использовании МАСС, как в прочем, и алгоритма k -средних.

Список использованных источников

1. Steinhaus, H. Sur la division des corps materiels en parties / H. Steinhaus // Bull. Acad. Polon. – 1956. – Vol. 4, N 12. – P. 801–804.
2. Lloyd, S. Least squares quantization in PCM / S. Lloyd // IEEE Transactions on Information Theory. – 1982. – Vol. 28, N 2. – P. 129–137. <https://doi.org/10.1109/tit.1982.1056489>
3. Bezdek, J. C. Pattern Recognition with Fuzzy Objective Function Algorithms / J. C. Bezdek. – MA, USA, 1981. – 256 p. <https://doi.org/10.1007/978-1-4757-0450-1>
4. Towards K-means-friendly spaces: simultaneous deep learning and clustering / Bo Yang [et al.] // ICML'17: Proc. of the 34th International Conference on Machine Learning. – 2017. – Vol. 70. – P. 3861–3870.
5. Fully convolution neural network combined with K-means clustering algorithm for image segmentation / Bing He [et al.] // Tenth International Conference on Digital Image Processing (ICDIP 2018). – 2018. – Vol. 10806. <https://doi.org/10.1117/12.2502814>

6. Deep k -Means: Re-Training and Parameter Sharing with Harder Cluster Assignments for Compressing Deep Convolutions / Junru Wu [et al.] // Proc. of the 35th International Conference on Machine Learning, PMLR. – 2018. – Vol. 80. – P. 5363–5372.
7. Гонсалес, Р. Цифровая обработка изображений / Р. Гонсалес, Р. Вудс. – М., 2005. – 1075 с.
8. Залесский, Б. А. Метод нечеткой кластеризации k -средних со сглаживающей штрафной функцией / Б. А. Залесский // Информатика. – 2014. – № 3. – С. 14–20.
9. Image Clustering. 2018 [Electronic resource]. – Mode of access: <https://paperswithcode.com/task/image-clustering>. – Date of access: 12.04.21.
10. Залесский, Б. А. Алгоритм отслеживания объектов движущейся видеокамерой / Б. А. Залесский // Докл. Нац. акад. наук Беларуси. – 2020. – Т. 64, № 2. – С. 144–149. <https://doi.org/10.29235/1561-8323-2020-64-2-144-149>

References

1. Steinhaus H. Sur la division des corps materiels en parties. *Bulletin L'Académie Polonaise des Science*, 1956, vol. 4, no. 12, pp. 801–804.
2. Lloyd S. Least squares quantization in PCM. *IEEE Transactions on Information Theory*, 1982, vol. 28, no. 2, pp. 129–137. <https://doi.org/10.1109/tit.1982.1056489>
3. Bezdek J. C. *Pattern Recognition with Fuzzy Objective Function Algorithms*. MA, USA, 1981. 256 p. <https://doi.org/10.1007/978-1-4757-0450-1>
4. Bo Yang, Xiao Fu, Sidiropoulos N. D., Mingyi Hong. Towards K -means-friendly spaces: simultaneous deep learning and clustering. *ICML'17: Proceedings of the 34th International Conference on Machine Learning*, 2017, vol. 70, pp. 3861–3870.
5. Bing He, FengXiang Qiao, Weijun Chen, Ying Wen. Fully convolution neural network combined with K -means clustering algorithm for image segmentation. Tenth International Conference on Digital Image Processing (ICDIP 2018), 2018, vol. 10806. <https://doi.org/10.1117/12.2502814>
6. Junru Wu, Yue Wang, Zhenyu Wu, Zhangyang Wang, Ashok Veeraraghavan, Yingyan Lin. Deep k -Means: Re-Training and Parameter Sharing with Harder Cluster Assignments for Compressing Deep Convolutions. *Proceedings of the 35th International Conference on Machine Learning, PMLR*, 2018, vol. 80, pp. 5363–5372.
7. Gonzales R. C., Woods R. E. *Digital Image Processing*. Forth Edition. Pearson/Prentice-Hall, 2018. 1192 p.
8. Zalesky B. A. Method fuzzy clustering k -means with smoothing penalty function. *Informatika = Informatics*, 2014, no. 3, pp. 14–20 (in Russian).
9. *Image Clustering*. 2018. Available at: <https://paperswithcode.com/task/image-clustering> (accessed 12 April 2021).
10. Zalesky B. A. Object tracking algorithm by moving camera. *Doklady Natsional'noi akademii nauk Belarusi = Doklady of the National Academy of Sciences of Belarus*, 2020, vol. 64, no. 2, pp. 144–149 (in Russian). <https://doi.org/10.29235/1561-8323-2020-64-2-144-149>

Информация об авторе

Залесский Борис Андреевич – д-р физ.-мат. наук, заведующий лабораторией. Объединенный институт проблем информатики НАН Беларуси (ул. Сурганова, 6, 220012, Минск, Республика Беларусь). E-mail: zalesky@newman.bas-net.by.

Information about the author

Zalesky Boris A. – D. Sc. (Physics and Mathematics), Head of the Laboratory. United Institute of Informatics Problems of the National Academy of Sciences of Belarus (6, Surganov Str., 220012, Minsk, Republic of Belarus). E-mail: zalesky@newman.bas-net.by.