

Применение информационной модели для выбора параметров алгоритмов сегментации изображений

Д.М. Мурашов¹

¹Федеральный исследовательский центр «Информатика и управление» РАН, Вавилова 42, Москва, Россия, 119333

Аннотация. Работа посвящена исследованию теоретико-информационного подхода к обеспечению качества сегментации цифровых изображений. Рассматривается система, включающая алгоритм сегментации с параметром, от которого зависит количество выделенных сегментов, и процедура выбора значения параметра, обеспечивающего минимум меры информационной избыточности. Для исследования свойств системы предложена новая упрощенная математическая модель. Показано, что для предложенной модели существует минимум меры избыточности. Адекватность модели подтверждена вычислительным экспериментом. Эксперимент, проведенный на изображениях из базы Berkeley Segmentation Dataset, показал, что сегментированное изображение, соответствующее минимуму меры избыточности, дает наибольшее информационное сходство с эталонами, имеющимися в базе BSDS500.

1. Введение

При сегментации изображений возникает проблема выбора параметров применяемых алгоритмов. Параметры выбираются, исходя из наилучшего качества сегментации. При этом для разных задач анализа изображений выбираются различные критерии качества. Это может быть визуальная оценка эксперта или какой-либо количественный показатель.

В исследованиях по сегментации результат обычно сравнивается с изображением, сегментированным экспертом и принятым в качестве эталона [1]. Возможно наличие нескольких эталонов, принятых разными экспертами. Если операция сегментации рассматривается как процесс кластеризации пикселей, то применяются теоретико-множественные, статистические и теоретико-информационные меры [2], используемые для сравнения результатов кластеризации данных. Наиболее часто применяются меры хи-квадрат; индекс Рэнда (Rand Index) [3] и его варианты; мера Жаккара (Jaccard index) [2]; мера Фаулкса-Мэллоуза (Fowlkes-Mallows) [4]; взаимная информация и варианты нормализованной взаимной информации [5]; вариация информации (variation of information) [6]. Эти меры позволяют сравнить разные варианты разбиения изображений на непересекающиеся сегменты.

В работе [7] при выборе параметров алгоритма оценивалось сходство результата сегментации с исходным изображением. В качестве меры сходства исходного и сегментированного изображений использован "взвешенный показатель недостоверности" (weighted uncertainty index), вычисляемый через значения нормализованной взаимной информации соответствующих цветовых каналов входного и сегментированного изображений [8]. По экспертным оценкам результатов сегментации фрагментов серии изображений, выполненной при различных значениях параметра, на координатной плоскости с помощью

классификатора, выделены области «недосегментации», «пересегментации» и оптимальной сегментации. При сегментации изображений параметр алгоритма для каждой точки обрабатываемого изображения выбирается с помощью итерационной процедуры на основе graph-cut алгоритма [9]. К недостаткам такого подхода следует отнести субъективность экспертных оценок, а также то, что обученный алгоритм будет давать приемлемые результаты только для тех классов изображений, которые использовались при обучении.

В работах [10], исходя из принципа минимизации избыточности информации [11], для определения наилучшего варианта сегментации изображений (и определения соответствующего параметра алгоритма) в качестве критерия предложено рассматривать минимум меры избыточности информации. Представляемая работа посвящена исследованию указанного критерия качества сегментации цифровых изображений. Для исследования свойств критерия предложена упрощенная математическая модель процесса сегментации. Показано, что для предложенной модели существует минимум меры избыточности. Адекватность модели подтверждена вычислительным экспериментом на изображениях из базы Berkeley Segmentation Dataset [1]. Сегментированное изображение, соответствующее минимуму меры избыточности, дает минимум меры различия при сравнении с эталонными изображениями.

2. Модель системы сегментации

2.1. Задача сегментации

Операция сегментации может быть описана следующей моделью [10]:

$$V = F(U, t), \quad (1)$$

где U - входное изображение; V - сегментированное изображение; F - оператор, описывающий алгоритм сегментации; t - параметр. Задача сегментации может быть сформулирована следующим образом. Пусть из заданного входного изображения U с помощью алгоритма сегментации (1) получен набор Q изображений $V = \{V_1, V_2, \dots, V_q, \dots, V_Q\}$ для различных значений параметра t . Требуется выбрать изображение V_q , обеспечивающее минимум критерию $M(U, V_q)$. В следующем разделе будет представлен информационный критерий качества сегментации, предложенный ранее в работе [10].

2.2. Информационный критерий качества сегментации

Критерий минимума избыточности информации был положен в основу теоретико-информационной модели зрительной системы человека в работе [11]. В работе [10] было предложено применять критерий минимума избыточности информации в алгоритме сегментации изображений. Для применения теоретико-информационного подхода необходима вероятностная модель связи между исходным и сегментированным изображениями.

Для простоты будем рассматривать процесс сегментации полутоновых изображений. Пусть исходное и сегментированное изображения являются входом и выходом стохастической информационной системы. Значения уровней серого изображений описываются дискретными случайными переменными U и V со значениями u и v . Переменная U имеет L , а V может иметь $1 \leq l \leq L$ градаций серого.

Операция сегментации будет представлена моделью информационного канала:

$$V = F(U, \eta), \quad (2)$$

где U – сигнал на входе канала, V - выход канала; F – функция преобразования; η – шум канала. Предполагается, что шум η является гауссовой случайной переменной с нулевым средним и дисперсией σ_η^2 ; переменные V и η независимы.

В качестве критерия качества сегментации изображений предлагается рассматривать минимум меры избыточности канала, определяемой в виде [11]:

$$R = 1 - \frac{I(U;V)}{C(V)}, \quad (3)$$

где $I(U;V)$ - взаимная информация, $C(V)$ - пропускная способность канала. Положим $C(V) = H(V)$, где $H(V)$ - энтропия выхода. Тогда, учитывая, что $I(U;V) = H(V) - H(V|U)$, выражение (7) примет вид:

$$R = \frac{H(V|U)}{H(V)}, \quad (4)$$

где $H(V|U)$ - условная энтропия выхода канала при условии, что вход равен U .

В следующем разделе будет предложена упрощенная качественная модель операции сегментации полутоновых изображений и показано, что для такой модели существует минимум критерия избыточности.

2.3. Информационная модель операции сегментации

Для исследования качественных свойств меры необходимо построить модель системы в виде совместного двумерного дискретного распределения уровней серого входа и выхода системы сегментации. Такая модель позволит изучить динамику информационных мер (энтропий), характеризующих процесс сегментации изображений.

Для построения качественной модели необходимо рассмотреть, как меняется совместное двумерное дискретное распределение уровней серого входа и выхода системы сегментации в зависимости от количества выделенных сегментов K на изображении V . Пусть на входе алгоритма сегментации предъявлено полутоновое изображение U , показанное на рисунке 1(а). На рисунке 1 (б)-(г) показаны двумерные дискретные распределения для трех значений K при числе градаций $L = 16$ серого в изображении U .

На рисунке 1 проиллюстрирована качественная картина динамики двумерного распределения при формировании сегментов. Из рисунка виден характер распределения полутонов в каждом сегменте. В каждом сегменте имеется значительное количество пикселей доминирующего уровня серого тона, образующих пики распределения, и некоторое количество пикселей с другими уровнями. При большом числе сегментов распределение уровней внутри сегментов характеризуется острыми пиками. При укрупнении сегментов и уменьшении их числа пики сглаживаются (см. рисунок 1).

Предположим, что совместное распределение уровня серого тона изображений U и V может быть представлено K компонентами, соответствующими сегментам изображения V . Для простоты положим, что все компоненты однотипные и состоят из L составляющих, которые соответствуют частоте появления пикселей уровня серого l в k -ом сегменте V . Каждая из компонент имеет пик, соответствующий доминирующему уровню серого.

Предположим, что компоненты имеют различные доминирующие уровни серого. Пусть $P(u_l, v_k) = P$, если $l = k$. Соотношение между значениями вероятностей уровней серого L в компонентах определяется коэффициентом α , $0 < \alpha \leq 1$. Например, в сегменте изображения V , закодированном уровнем v_1 $P(u_2, v_1) = P(u_3, v_1) = \dots = P(u_L, v_1) = \alpha P(u_1, v_1) = \alpha P$, где $\alpha = \alpha(K)$ зависит от количества сегментов на изображении V . Модель описанного двумерного дискретного распределения показана на рисунке 2.

Для такой модели имеют место следующие соотношения: $P(v_k) = (L-1)\alpha P + P$, $K[\alpha(L-1)+1]P = 1$, откуда следует, что $P = 1/K[\alpha(L-1)+1]$. Далее, с учетом сформированной модели совместного дискретного распределения уровней серого тона входа и выхода системы сегментации найдем выражения для энтропий, входящих в формулу (4):

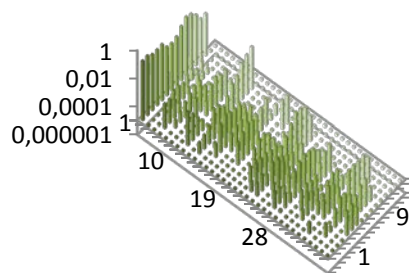
$$H(U, V) = - \sum_{l=1}^L \sum_{k=1}^K p(u_l, v_k) \log P(u_l, v_k) = \log K + \log[\alpha(L-1)+1] - \frac{(L-1)\alpha \log \alpha}{\alpha(L-1)+1} \quad (5)$$

$$H(U) = -\sum_{l=1}^L \left\{ \left[\sum_{k=1}^K P(u_l, v_k) \right] \log \left[\sum_{k=1}^K P(u_l, v_k) \right] \right\} = -LP_0 [(K_0 - 1)\alpha_0 + 1] \log [(K_0 - 1)\alpha_0 P_0 + P_0]; \quad (6)$$

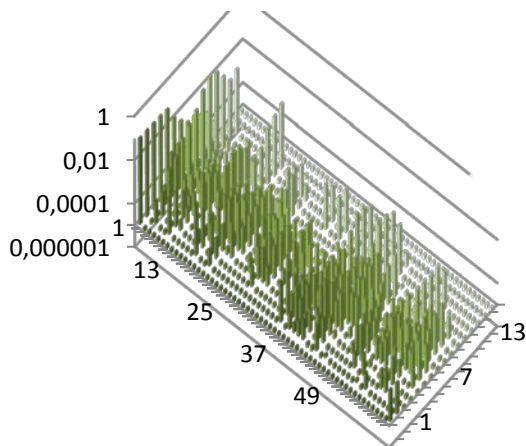
$$H(V) = -\sum_{k=1}^K \left\{ \left[\sum_{l=1}^L P(u_l, v_k) \right] \log \left[\sum_{l=1}^L P(u_l, v_k) \right] \right\} = \log K \quad (7)$$



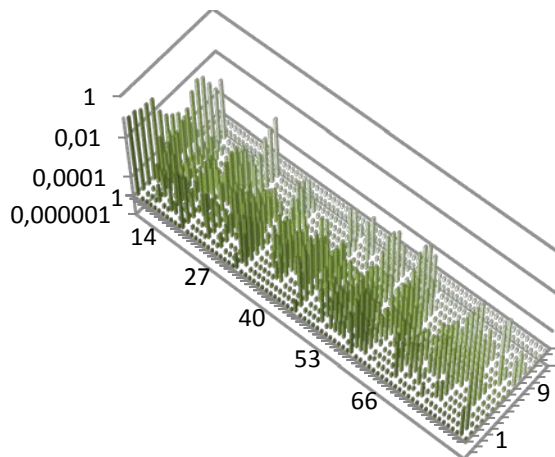
а)



б)



в)



г)

Рисунок 1. Совместное дискретное распределения уровней серого входа и выхода системы сегментации при трех значениях количества выделенных сегментов K : (а) входное изображение; (б) совместное дискретное распределение уровней серого U и V при $K = 32$; (в) при $K = 57$; (г) при $K = 75$.

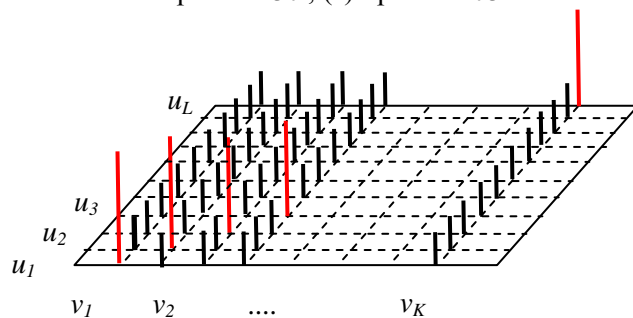


Рисунок 2. Модель совместного дискретного распределения уровней серого изображений U и V .

$$H(V|U) = H(U, V) - H(U) = \log K + \log[\alpha(L-1)+1] - \frac{(L-1)\alpha \log \alpha}{\alpha(L-1)+1} + LP_0[(K_0-1)\alpha_0+1] \log[(K_0-1)\alpha_0 P_0 + P_0], \quad (8)$$

где P_0 , K_0 и α_0 - величины, соответствующие изображению U .

Подставляя выражения (5)-(8) в (4) и преобразуя, получим:

$$R = 1 + \frac{\log[\alpha(L-1)+1] - \frac{(L-1)\alpha \log \alpha}{\alpha(L-1)+1} + LP_0[(K_0-1)\alpha_0+1] \log[(K_0-1)\alpha_0 P_0 + P_0]}{\log K}. \quad (9)$$

Замечено, что при уменьшении числа сегментов K изображения V характер распределения (см. рисунок 1) меняется. В модели, представленной на рисунке 2, соответствующая трансформация соответствует росту коэффициента α . Такая зависимость может быть представлена монотонной функцией, например:

$$\alpha(K) = \frac{a}{1 + e^{c(K-b)}} + d, \quad (10)$$

где a , b , c и d - параметры, $d=1-a$. Характер зависимости $\alpha(K)$ при различных значениях параметра c показан на рисунке 3.

Покажем, что критерий R (7-9) будет иметь минимум. Переменная K принимает значения из множества целых чисел. Для анализа функции R , описываемой выражением (9), сделаем замену целочисленной переменной K на действительную переменную z . Положим, что $\log()$ в выражении (9) - натуральный логарифм. Тогда производная dR/dz имеет вид:

$$R'(z) = \frac{\alpha(L-1)\log \alpha}{z[\alpha(L-1)+1]\log^2 z} - \frac{\log[\alpha(L-1)+1]}{z\log^2 z} - \frac{ace^{c(z-b)}(L-1)^2 \alpha \log \alpha}{[\alpha(L-1)+1]^2 [e^{c(z-b)}+1]^2 \log z} + \frac{ace^{c(z-b)}(L-1)\log \alpha}{[\alpha(L-1)+1][e^{c(z-b)}+1]^2 \log z} + \frac{H(U)}{z\log^2 z}. \quad (11)$$

Функция $\alpha(z)$, описываемая выражением (10), убывает с ростом z (см. рисунок 3). Для функции $\alpha(z)$ вида (10) показано, что $R'(z)$ при относительно малых z отрицательна, а при больших z положительна. Следовательно, существует хотя бы одна точка z , в которой $R'(z)=0$, и при этом вторая производная в этой точке положительна $R''(z)>0$. Это означает, что функция $R(z)$ имеет минимум. Графики функции $R(K)$, определяемой выражением (9), при различных значениях параметра c зависимости $\alpha(K)$, заданной выражением (10), показаны на рисунке 4(а). Из рисунка 4(а) видно, что функция $R(K)$ имеет минимум.

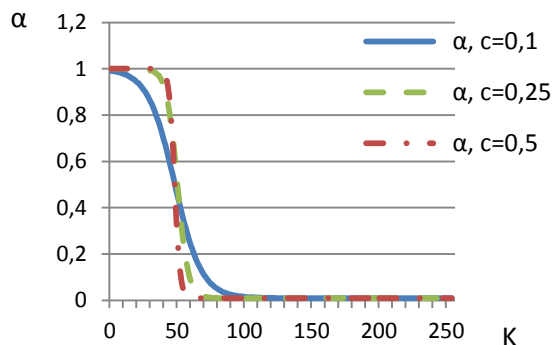


Рисунок 3. График функции $\alpha(K)$ при различных значениях параметра c .

Графики изменения энтропий (5) – (8), определяющих величину информационной избыточности (9) в зависимости от числа выделенных на входном изображении сегментов, показаны на рисунке 4(б).

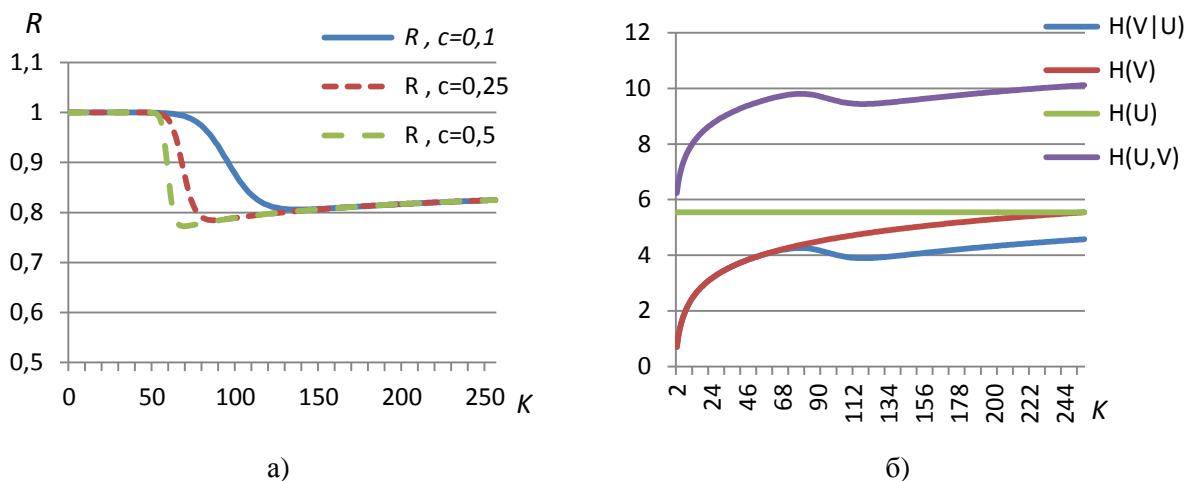


Рисунок 4. (а) График функции $R(K)$ при различных значениях параметра c ; (б) график изменения энтропий, определяющих величину информационной избыточности R , в зависимости от числа выделенных сегментов K при значении параметра $c = 0,1$.

На рисунках 5(а) и 5(б) представлены графики изменения величины информационной избыточности и энтропий, определяющих информационную избыточность (4), в зависимости от числа выделенных сегментов на изображении, приведенном на рисунке 1(а). Рисунки 4(а) и 5(а), 4(б) и 5(б) демонстрируют качественное сходство динамики величин информационной избыточности и энтропий в зависимости от числа выделенных сегментов при разбиении гипотетического и реального изображений. Эффективность критерия информационной избыточности при сегментации изображений подтверждается вычислительным экспериментом, описываемом в следующем разделе.

3. Вычислительный эксперимент

В данной работе для иллюстрации адекватности предложенной модели процесса сегментации использовались изображения из базы BSDS500 университета Беркли [1]. Каждое из исследуемых изображений сегментировалось с помощью модифицированного алгоритма SLIC с различными значениями параметра Δ_1 условия процедуры постобработки [10]. В результате сегментации изображения U получено множество из Q изображений $V = \{V_1, V_2, \dots, V_Q\}$. Далее сегментированные изображения $V_q, q = 1, 2, \dots, Q$ сравнивались с исходным изображением U и с вариантами эталонной сегментации $V_s^{GT}, t = 1, 2, \dots, S$ из базы BSDS500. Для сравнения изображений применялась вариация информации (variation of information) - информационная мера, предложенная в работе [6]. В нашем случае вариация информации характеризует различие (расстояние) между двумя версиями сегментации (или между исходным и сегментированным изображением).

Для оценки качества сегментации изображений использовалась относительная разность

$$\Delta K_{rel} = \frac{K_{min} - K_{min}^{GT}}{K_{max}},$$

где K_{min} - количество сегментов на изображении, которое обеспечивает минимум критерия избыточности; K_{min}^{GT} - количество сегментов на изображении, V_q , которое обеспечивает

наименьшую вариацию информации при сравнении с эталонной сегментацией V_s^{GT} ; K_{max} - наибольшее количество сегментов на изображениях V_1, V_2, \dots, V_Q . Гистограмма значений ΔK_{rel} , построенная по результатам сегментации 54 изображений из базы BSDS500 и сравнении с 270 эталонными сегментациями, показана на рисунке 6. Из рисунка 6 следует, что достаточно большая группа изображений, для которых сегментация по критерию минимума информационной избыточности обеспечивает высокое информационное сходство с эталонными сегментациями.

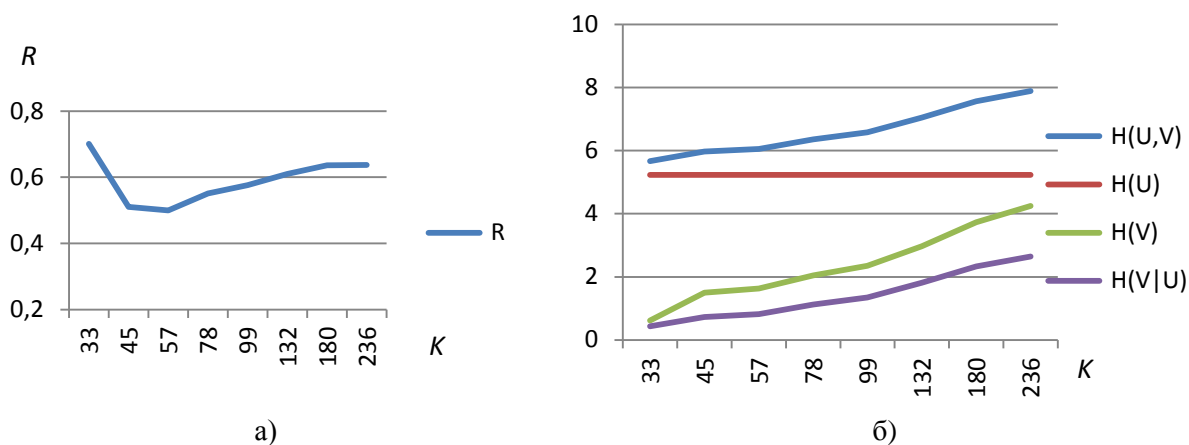


Рисунок 5. Зависимости информационных характеристик, полученные при сегментации для изображения, показанного на рисунке 1(а): (а) график функции $R(K)$; (б) графики изменения энтропий, определяющих $R(K)$.

4. Заключение

В статье предложена упрощенная математическая модель для исследования алгоритмов сегментации цифровых изображений с параметром. В качестве критерия качества используется минимум меры информационной избыточности. Показано, что предложенная модель обеспечивает минимум меры избыточности, что позволяет получить наилучший с точки зрения визуального восприятия вариант сегментированного изображения. Соответствие модели реальной системе подтверждено вычислительным экспериментом на изображениях из базы Berkeley Segmentation Dataset. Сегментированное изображение, соответствующее минимуму меры избыточности, дает минимум меры различия при сравнении с эталонными изображениями. Дальнейшие исследования будут направлены на более глубокое изучение модели и свойств меры избыточности.

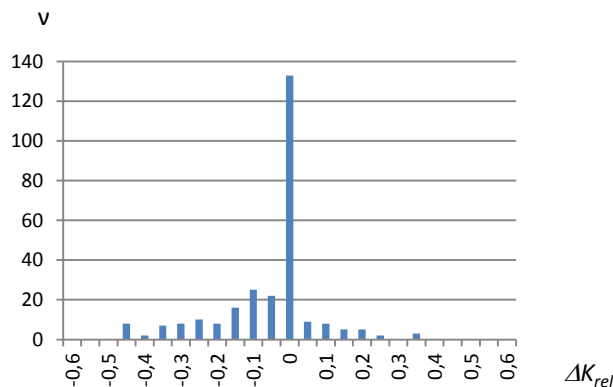


Рисунок 6. Гистограмма величины ΔK_{rel} , вычисленной при сегментации изображений из базы BSDS500, ν - частота появления определенного значения ΔK_{rel} .

5. Благодарности

Работа выполнена при частичной финансовой поддержке РФФИ, проекты № 18-07-01385 и № 18-07-01231.

6. Литература

- [1] Arbelaez, P. Contour detection and hierarchical image segmentation / P. Arbelaez, M. Maire, C. Fowlkes, J. Malik // *IEEE transactions on pattern analysis and machine intelligence*. – 2011. – Vol. 33(5) – P. 898-916.
- [2] Zhang, H. Image segmentation evaluation: A survey of unsupervised methods / H. Zhang, J.E. Fritts, S.A. Goldman // *Computer vision and image understanding*. – 2008. – Vol. 110(2). – P. 260-280.
- [3] Rand, W.M. Objective criteria for the evaluation of clustering methods / W.M. Rand // *Journal of the American Statistical association*. – 1971. – Vol. 66(336). – P. 846-850.
- [4] Fowlkes, E.B. A method for comparing two hierarchical clusterings / E.B. Fowlkes, C.L. Mallows // *Journal of the American Statistical association*. – 1983. – Vol. 78(383). – P. 553-569.
- [5] Strehl, A. Cluster Ensembles – A Knowledge Reuse Framework for Combining Multiple Partitions / A. Strehl, J. Ghosh // *Journal of Machine Learning Research*. – 2002. – Vol. 3. – P. 583-617.
- [6] Meilă, M. Comparing clusterings: an axiomatic view / M. Meilă // *Proceedings of the 22nd international conference on Machine learning / ACM*, 2005. – P. 577-584.
- [7] Frosio, I. Adaptive segmentation based on a learned quality metric / I. Frosio, E.R. Ratner // *VISAPP*. – 2015. – Vol. 1. – P. 283-292.
- [8] Ana, L.N.F. Robust data clustering / L.N.F. Ana, A.K. Jain // *Computer Vision and Pattern Recognition*. – 2003. – Vol. 2. – P. 128-133.
- [9] Felzenszwalb, P.F. Efficient graph-based image segmentation / P.F. Felzenszwalb, D.P. Huttenlocher // *Int. J. Comput. Vision*. – 2004. – Vol. 59(2). – P. 167-181. DOI: 10.1023/B:VISI.0000022288.19776.77.
- [10] Murashov, D. Theoretical-information quality model for image segmentation / D. Murashov // *Procedia engineering*. – 2017. – Vol. 201. – P. 239-248.
- [11] Atick, J.J. Towards a theory of early visual processing / J.J. Atick, A.N. Redlich // *Neural Computation*. – 1990. – Vol. 2(3). – P. 308-320.

Application of an information model for selecting parameters of image segmentation algorithms

D.M. Murashov¹

¹Federal Research Center “Computer Science and Control” of Russian Academy of Sciences, Vavilov st. 40, Moscow, Russia, 119333

Abstract. The work is devoted to the study of the information-theoretical approach to the problem of image segmentation quality. We consider a system that includes a segmentation algorithm with a parameter on which the number of selected segments depends, and a procedure for selecting parameter value that provides a minimum of segmentation quality criterion. As a quality criterion, a measure of information redundancy is used. To study the properties of the system, a new simplified mathematical model is proposed. It is shown that for the proposed model, the redundancy measure has a minimum. The validity of the model is confirmed by a computational experiment. An experiment conducted on images from the Berkeley Segmentation Dataset database showed that a segmented image corresponding to a minimum of redundancy measure gives the highest informational similarity to the ground truth images available in the BSDS500 database.