Example-Based Object Detection in the Attached Image

Mikhail Kharinov [1]

[1] *St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), 39, 14th Line V.O., St. Petersburg, 199178, Russia*

Abstract

The paper solves the problem of detecting exemplified objects in a color image. A solution provides the representation of similar objects in the same colors, and different objects in different colors. This is achieved by combining images of object examples and a target image into a single joint image, which is represented in sequential number 1, 2, ..., etc. colors. The mentioned effect is demonstrated by detecting irises and pupils in a test image. It is explained by the fact that: a) the joint image is approximated by a hierarchy of approximations in sequential color numbers; b) the hierarchy of approximations is described by a convex sequence of *approximation errors* (values of the total squared error); c) due to the convexity, the approximation errors are reduced for all approximations of the joint image. In the last explanation item, it is applied the operation of combining hierarchically organized objects into a single object, which is introduced in this paper. To produce the required hierarchy of image approximations Ward's pixel clustering is used. Ward's method is generalized for image processing *by parts* (within pixel subsets) that provides generation of multiple proper approximation hierarchies and accelerates the calculations. To do so, the so-called split-and-merge pixel cluster CI-method is embedded into Ward's generalized method to provide a real-life minimization of the error for image approximation in a fixed number of colors.

Keywords

Color image, object detection, Ward's pixel clustering, split-and-merge methods, total squared error, convex sequence.

1. Introduction

Any specialist in detecting objects in a digital image knows that it is easier to detect objects within the same image than in different images. But hardly anyone is able to estimate how much easier? Obviously, this can be a practical guide to action, which is being done. The answer to the question and the theoretical understanding of the experimental results turns out to be non-trivial and leads to the pixel clustering model developed by the author for detecting objects in color images.

A pronounced improvement in object detection when images are merged into single one was noticed for stereopairs [1] and is used to recognize remote images of the same scene [2,3]. To develop an utilization of the idea, let's attach the input image to the image of object-of-interest examples and approach the joint image by a sequence of piecewise constant approximations in the color numbers from 1 to N, where N is the number of pixels in the image. Let's find adequate methods, analyze the dependencies characteristic for the image, come up with computer calculation techniques and understand how does it works. Then we'll better understand how the image and natural visual perception are arranged.

The purpose of the paper is to explain the effect of explicit classification of objects due to image merging and to indicate the exploited property of adequately ordered information inherent in the image and objects in the image.

The paper content is available to the reader who is familiar with the concept of piecewise constant approximation and elementary methods of cluster analysis [4,5], which provide a reduction in the *approximation error* E (total squared error):

GraphiCon 2022: 32nd International Conference on Computer Graphics and Vision, September 19-22, 2022,
St. Petersburg Federal Research Center of the Russian Academy of Sciences (SPC RAS), St. Petersburg, Russia
EMAIL: khar@iias.spb.su (M. Kharinov)
ORCID: 0000-0002-5166-1381 (M. Kharinov)

- Ward's clustering;
- split-and-merge methods;
- K-means method,

as well as Otsu method for a *halftone* (grayscale, gray) image. For a clear understanding of the meaning of the high-speed network computer calculations mentioned in the paper, it is also desirable to distinguish a tree from a cyclic graph.

The central concept of the listed methods is the optimal piecewise constant approximation of the image with the minimum approximation error E_g^{opt} for a given number g of pixel clusters, i.e. colors or tones, in the case of a halftone image. In the general case of arbitrary data, the exact calculation of optimal approximations for any admissible number of clusters is a practically unsolvable NP-hard problem [6]. But, according to our experience, for digital images in the form of pixel multisets with values in narrow intensity ranges, the problem is quite solvable. And the methods of practical calculation of optimal or close to optimal pixel clusters are probably of theoretical and practical interest not only for image processing, but also for cluster analysis.

In the theory of clustering, we start from the results of the work [6], which studies the features of Ward's clustering, and also poses the problem of joint application of Ward's method and K-means one. In [6], to minimize the approximation error E by K-means method, it is proposed to pre-calculate the hierarchy of Ward's approximations and memorize it using ordinary trees (dendrograms).

According to Willem J. Heiser's editorial comments, the paper [6] belongs to the top ones, and the problem posed is extremely difficult. This is indeed true, but with the caveat that, in contrast to Ward's clustering, K-means method does not accurately match to the minimized functional E [7]. Therefore, in order to perfectly minimize approximation error E and obtain optimal or closest to optimal approximations, it is important to refine Ward and K-means methods themselves and additionally consider the combined use of Ward's clustering with the method of splitting/merging image pixel clusters, which is the focus of the paper. As for dendrograms, it is known that in computer calculations, to memorize and work with the hierarchy of image approximations, it is more convenient to use Sleator-Tarjan dynamic trees [8,9] instead of dendrograms, which affects the speed of calculations.

2. Method Demonstration

Obviously, the technique of attaching an image to an image of object examples can be applied to a wide variety of objects-of-interest. For definiteness, let's test it in the problem of detecting pupils and irises.

Many works are devoted to the specific problem of iris detection, for example [10–15]. At the input, the information of the original image is taken without prior hierarchical or other ordering. The color image when detecting iris is usually converted to grayscale. A typical processing technique is to take into account the rounded shape of the pupils and iris through local or integral transformations. Replacing a color image with a grayscale one obviously reduces the amount of input data and the possibility of their recognition. Heuristic consideration of a priori assumptions about the content of the image narrows the implementation area. Such iris detection programs are likely to perform poorly when presented with, say, images of the eyes of animals whose pupils are not always rounded.

Let's check if the color in the image is interfering with us. Let's abandon the a priori geometric description of the objects-of-interest, leaving it for the next processing stage, when the objects-of-interest are already localized. Let's take images attached to each other obtained at different scales (Figure 1).

Figure 1 shows a joint image prepared for the iris detection experiment. On the left in Figure 1 there are 16 images of differently colored eyes, available at https://commons.wikimedia.org/w/index.php?curid=14527332 (scaled). The eyes in the two leftmost images in the top row look lighter than the rest ones. The rightmost "black eye" in the lower row of eye images has a broken structure, because the pupil merges with the iris. On the right is a fragment of a modern photo of "Lena" model from 319×512 pixels (https://www.reddit.com/r/pics/comments/erqtr5/in_1973_a_photo_of_a_swedish_playboy_model_named/) with a better shot of the eyes than in the famous "Lena" image. There are a total of 1114×512 pixels in the joint image that made up of 17 nested images.

Figure 1: Joint image containing sixteen eye images in the four rows on the left and a face photo attached on the right

2.1. Formulation of the Problem

The objects considered are the pupils, the irises, as well as the reflections in the pupil areas of the lighting devices with aureoles, which must be "subtracted" from the images of the pupils or the irises. Objects on the left are considered *marked up* with object type labels, i.e., pupils and irises on the left are assumed to have known sets of pixel coordinates.

Let's treat the problem of detecting pupils and irises in the image of "Lena", whose pixel coordinates are to be found.

The following simplest object detection procedures are wanted.

Let's get an approximation of image Figure 1 in several colors. As is known, a strong reduction in the number of colors is accompanied by image segmentation and a relatively small number of segments in the image approximation, which contributes to automation.

1. Then the detection of coordinates using Figure 1 is reduced to reading the average color values of the pixels of object segments in 16 eye images on the left and searching for segments marked with the same colors in "Lena" image on the right.
2. And vice versa, we read the colors of the segments in "Lena" image on the right and find the segments of which objects are marked with these colors on the left. Then we select the most probable objects to which these segments belong.

2.2. Experimental Results

The results of the experiment are presented in Figure 2.

Figure 2 in horizontal rows shows the initial approximations of the joint image of Figure 1 processed as a whole (right column) and by 17 parts (left column). The right column from top to bottom shows the initial four approximations in one, two, three and four colors for a single joint image of marked up objects and a recent photograph of "Lena" model. The left column shows 17 approximations each in the same number of colors, but calculated for every of the 17 separated nested images.

Let's look at the left column.

The total number of colors in 17 approximations is 17, 34, 51, 68, etc., at least up to and including 170 for representation in 10 colors. I.e. in approximations of each *level* (in a given number of colors), all averaged colors are different even after rounding off the actual real (more precisely, rational) values. Therefore, coloring objects with exactly the same colors is unlikely, and the solution of the problem is not trivial, as it is intended in its formulation.

Figure 2: Hierarchical structuring of 17 images individually (left) and as a single joint image (right)

A different picture emerges when all 17 images are merged into a single image. Compared to the approximations in Figure 2 in the left column color differences between objects are leveled in the right column, which contributes to their simplest detection.

Pupil pixels in the approximations of the second and third levels all have color $93,62,56$ in terms of RGB components, and in the approximation of the fourth level they are marked with color $44,28,30$.

In 16 approximations of the eyes of the third level, the irises mostly merge with the pupils, and in the approximation of the fourth level, the pixels of the irises in the majority have colors $115,77,68$ or $197,141,118$. In the two images of light eyes of the third and fourth levels, the irises consist of pixels with values of $197,141,118$ and light inclusions of $228,212,194$.

In "Lena" image, the pixels of a pair of irises are perfectly localized in the fourth level approximation, where they coincide with the prevailing color $115,77,68$ of eleven of the fourteen "dark" irises, except for the image approximation with a broken structure in the lower right corner and two central ones in penultimate column of eye images.

The iris pixels in "Lena" image continue to form connected segments in approximations up to the eighth level inclusive. At the same time, at the fifth and sixth levels, the segments have color $115,77,68$. And at the seventh and eighth levels, they are colored with color $125,75,54$, just like the pixels of the irises in the six approximations of the eye images on the left in the two lower rows, with the exception of two approximations on the diagonal that descends from left to right.

As for the reflections of the lighting devices, it manifests itself at all *non-trivial* levels from the second to the fourth level in almost all approximations of marked up objects. Aureoles appear at the fourth level, where they occupy most of the pupil in some approximations. Therefore, these aureoles must necessarily be considered, although neither the reflections of lighting devices nor aureoles are reliably detected at the approximations of "Lena" image due to insufficient resolution.

Obviously, in order to expand the possibilities of detecting objects, it is useful to distinguish pixels of a given color from each other, in addition to color, also by other features. Due to the hierarchical structure, the pixels of given cluster are characterized by the following additional attributes as the level g increases:

- a non-increasing sequence of approximation errors E_g;
- non-decreasing negative increments ΔE_g of the approximation errors E_g;
- a non-increasing *segmental* sequence of errors E_g^{seg} calculated for segments of image approximations in different color numbers g;
- segmental sequence of increments ΔE_g^{seg} of approximation errors E_g^{seg};
- the ranges of levels within which the cluster or segment does not change, as well as other attributes, for example, an attribute indicating to what, larger or smaller by approximation error nested cluster (segment), a given pixel belongs when the cluster (segment) containing this pixel is divided in two.

The study of these and other new features is an interesting topic for further experimental research.

The advantage of detecting objects in a joint image (on the right in Figure 2) in comparison with the conventional use of independent data items (on the left in Figure 2), at first glance, seems difficult to implement. Indeed, the hierarchy of approximations for separated images (on the left in Figure 2) compared to the hierarchy of approximations of the joint image (on the right in Figure 2) is calculated much faster due to the polynomial increase in computational complexity with increasing number of pixels. But this shortcoming was overcome. In fact, the desired single hierarchy for the joint image (right in Figure 2) is obtained by transforming approximation hierarchies for the image parts (left in Figure 2) by simple algorithms. In other words, on a set of hierarchically structured objects, the operation of merging objects with each other into single objects is introduced. This ensures the acceleration of calculations by dividing the image into parts, as well as the construction of various options for the hierarchy of approximations (right column in Figure 2), since the required hierarchy of approximations is not univocal.

2.3. Interpretation

Let's figure out what caused the positive effect of merging images into a single joint image using the example of processing Figure 1 by parts and as a whole (Figure 3).

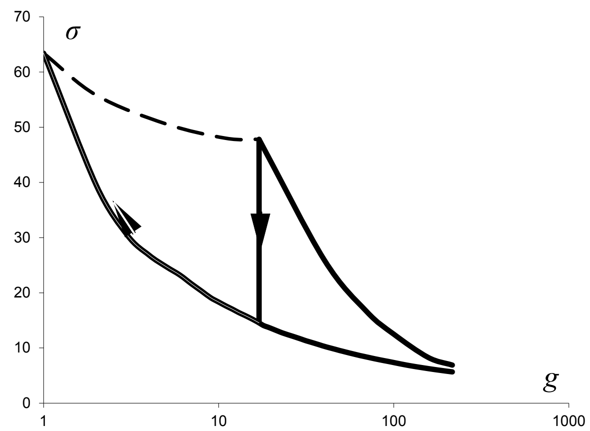

Figure 3: Reducing approximation errors when merging hierarchical image approximation sequences into an aggregate hierarchy

Figure 4 shows the dependence of the mean square deviation $\sigma = \sqrt{E/3N}$ of the image approximations as a function of the number g of colors in the approximations.

The lower curve describes the sequence of the deviations σ_g of the joint image approximations (the right column in Figure 2). When translating the values of σ_g into approximation errors $E_g = 3N\sigma_g^2$, the sequence of E_g will be convex:

$$E_g \le \frac{E_{g-1} + E_{g+1}}{2}, \quad g = 2, 3, ..., N-1. \tag{1}$$

Thanks to the convexity of the approximation error sequence, the errors E_g obviously are limited:

$$E_g \le E_1\left(1 - \frac{g-1}{N-1}\right), \quad g = 1,2, \ ..., N. \tag{2}$$

Note: in fact, the experimental values of E_g lie much closer to the optimal values, like ones of twenty-one charts in Figure 4, built for the "black eye" image of $1860 \times 1158 = 2153880$ pixels for a different numbers N of pre-enlarged pixels, where N takes values from 10000 to 100000 skipping 10000, from 100000 to 200000 skipping 50000, from 200000 to 500000 skipping 100000, and from 500000 to 2000000 skipping 250000.

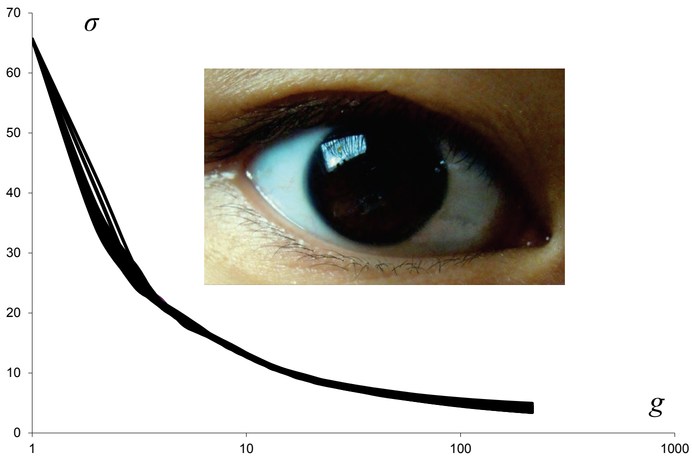

Figure 4: Experimental dependences of the approximation error E on the color number g for different values of the number N of pre-scaled pixels

When separating the image Figure 1 to 17 independent nested images, for each of them its own hierarchy of approximations is built. As it is easy to figure out, all 17 hierarchies can be combined into single hierarchy by re-ordering the splitting of pixel clusters into two nested ones as follows.

1. All 17 pixel clusters are united into one set.
2. At the first step, out of 17 pixel clusters, the one that provides the maximum drop in the approximation error E is split into two.
3. Then, such E minimization continues at each subsequent step until all clusters have disintegrated into individual pixels.

It turns out an incomplete hierarchy of approximations from $17, 18, ..., N$ pixel clusters, described in Figure 4 by the upper dangling solid curve, which is convex in the $g \times E$ coordinates. If the resulting hierarchy is completed to a full one (dashed curve), then the total curve will turn out to be piecewise convex with violation of convexity at $g = 17$ colors. Because of this single convexity violation, constraint (2) ceases to work.

Thus, the effect of detecting similar objects in the same image is explained by the improvement in the quality of the image approximation at all *levels* for each number of colors. At the same time, the improvement in quality is achieved due to the construction of a complete hierarchy of approximations, described by a convex dependence of the approximation error E on g, corresponding to the squared values of the lower curve in Figure 4.

The area between the upper and lower curves connecting the common points in Figure 3, as well as the defference between upper and lower curve at $g_0 = 17$, can serve as a numerical estimate of the hierarchical pixel clustering improvement when images are merged into one.

To quickly obtain the required hierarchy (lower curve in Figure 3), the approximation hierarchy in $17,18,...,N$ colors (upper solid unfinished curve) using the modernized Ward's method in combination with CI method is transformed into a similar hierarchical sequence of approximations in $17,18,...,N$ colors, described by a "tail" of the lower curve (down arrow in Figure 3) and is completed to the full hierarchy by Ward's original method [4,5] (arrow to the left up along the thin white line in Figure 3). The methods mentioned are detailed in the next Section.

3. Upgrading Methods for Minimizing Approximation Error *E*

In this Section, we will discuss the simplest clustering methods from the exhaustive number of methods presented in [4,5]. Unfortunately, in modern applications to image pixel clustering, the accurate understanding of even the simplest methods leaves much to be desired.

3.1. Working Equations

The classical methods considered here for the approximation error E minimizing are justified by the formula (3) for the increment ΔE_{merge} caused by the merging of two pixel clusters a,b and the pair of formulae (4), (5), which are derived from the formula (3):

$$\Delta E_{merge} = \frac{n_a n_b}{n_a + n_b} \left\| I_a - I_b \right\|^2, \tag{3}$$

$$\Delta E_{split} = -\Delta E_{merge}, \tag{4}$$

$$\Delta E_{correct} = k\left(\frac{n_b}{n_b + k} \left\| I_b - I_k \right\|^2 - \frac{n_a}{n_a - k} \left\| I_a - I_k \right\|^2 \right), \tag{5}$$

where (5) describes the increment $\Delta E_{correct}$ of the approximation error E caused by reclassifying of k pixels from cluster a to cluster b, I_k is averaged intensity of k pixels, n_a, n_b are the pixel numbers in clusters a, b and I_a, I_b are the average intensities within the clusters a and b. The trivial formula (4) is rather non-trivial in practical implementation via reversible calculations, in which an unlimited rollback is supported along with E minimizing in the reverse course of calculations.

Ward's method utilizes (3). The formulae (3) and (4) are sufficient for transparent split-and-merge methods. The formula (5) should be taken into account in relation to the commonly used K-means method, which is deduced from (5) under a fairly rough assumptions: $\sqrt{n_b/(n_b + k)} \approx \sqrt{n_a/(n_a - k)} \approx 1$.

3.2. Ward's Pixel Clustering

Ward's method is indispensable method for generating a hierarchy of pixel clusters describing by convex curve. Ward's original method is undeservedly rarely used for image pixel clustering, firstly, because of the supposedly excessive computational complexity, and secondly, because of some instability of the results when varying the image, changing the number of pre-enlarged pixels, etc. In our experience, Ward's original pixel clustering is good for detecting instances of objects in the same color in one and the same image, which makes it possible to improve object detection using the image merging technique [1–3]. However, to make full use of Ward's pixel clustering, the method needs to be upgraded in accordance with the features of the input image data.

The instability of original Ward's pixel clustering [4,5] is caused by the fact that the method generates the single convex curve from more than N existing ones. This is easy to verify if, when merging clusters, to take into account the options for incrementing the approximation error E in the vicinity of its current minimum value. If the optimal image approximation in $g = g_0$ colors is known, then to obtain an image approximation hierarchy described by a convex curve and containing a given optimal approximation in g_0 colors, it is sufficient to perform the original Ward's method within each pixel cluster of the optimal image approximation in g_0 colors, then, without modifying the pixel clusters obtained, re-sort their merging, and, finally, complete the generation of the hierarchy by iteratively merging the pixel clusters of the optimal image approximation in g_0 colors accordingly to original Ward's method. It should be noted that, using the above upgraded Ward's method, we simultaneously solve both the problem of controlled selection of a specific suboptimal hierarchy depending on the tuning parameter g_0, and the problem of reducing computational complexity. Of course, in real-life calculations, the optimal approximation of the image in g_0 colors is not known in advance. Therefore, to perform high-speed Ward's pixel clustering *by parts* (within the clusters), some another initial approximation is taken. In this case, the mentioned calculation sequence is preserved, but before the reordering of the cluster mergings, the chosen approximation is processed accompanied with a decrease in the approximation error E according to so-called CI (Clustering Improvement)-method of improving the quality of the image approximation in g_0 colors, which specifies the division of the image into parts processed as separate images.

CI-method is described in the next subsection.

The computational complexity of Ward's pixel clustering by image parts is estimated as $N^{\frac{4}{3}}$ depending on the pixel number N [16]. Moreover, when the modernized Ward's method is applied recursively, the computational complexity is drop as $N^{\frac{4}{3}} \to N^{\frac{16}{15}} \to N^{\frac{256}{255}} \to$ etc., depending on the recursion step.

So, instead of complaining about the polynomial increase in computational complexity with increasing number of processed pixel sets, we use that with a decrease in the number of pixel sets, the complexity drops to the same extent.

Thus, Ward's recursive pixel clustering by image parts provides an almost linear dependence of computational complexity on the pixel number N. Taking into account the reduction in complexity due to the reduction in the number of pre-enlarged pixels, the dependence of the clustering speed on N is ensured no worse than linear. And, at the same time, the adjustment of the recursive Ward's clustering by the image parts to the desired number g_0 of bases colors is provided.

Ward's original pixel pre-enlarged pixel clustering uses the basic operation of cluster merging (3). The remaining methods use an additional basic operation of splitting pixel clusters, which is performed as the inverse of the merge operation described by (4). As an independent operation with pixel clusters, reclassification of pixels from one cluster to another is also used in accordance with (5). Evidently, this operation is boiled down to mentioned basic ones. All three considering approximation error minimization algorithms are *greedy*, i.e. out of several available variants of the cluster transformation, accompanied by a change in the approximation error E, the one that provides the smallest error is selected. If the equal options occur, the first one that comes across is so far selected.

3.3. CI (Clustering Improvement)-method

CI-method is a split-and-merge method for hierarchically structured pixel or pre-enlarged pixel clusters, intended to minimize the approximation error E of a given image approximation in g colors. It is assumed that the hierarchy for each cluster is binary and is not limited by the conditions of coincidence of pixel clusters with some connected segments or other similar conditions. In a typical case, the current binary hierarchy within each cluster and the binary hierarchy of the clusters themselves is obtained by Ward's methods and is described by a convex sequence of approximation errors. The joint application of the methods under consideration is accompanied with violations of the convexity property, which are detected online and immediately corrected by reducing the approximation error *E*.

In CI-method, the cluster is first split into two, and then a pair of clusters merges with the maximum possible total drop (minimum negative increment) of the approximation error E, if it occurs and while it occurs. So CI-method is as follows.

1. The pixel cluster with the maximum *drop* $\left|\Delta E_{split}\right|$ (minimum negative increment) of the approximation error E according to (4) is found and split in two.
2. Let's then find a pair of clusters whose merging is accompanied by a minimum increment ΔE_{merge} of the approximation error E in accordance with (3). If the total increment $\Delta E_{total} = \Delta E_{merge} + \Delta E_{split}$ as a result of splitting the first cluster of pixels in two and merging the found two clusters is negative, then the found clusters are merged. Otherwise, with a non-negative total increment $\Delta E_{total} \geq 0$, the cluster divided in two is restored and processing ends.
3. Let's go back to point 1.

CI-method leaves unchanged any approximation obtained by the original Ward's method as a member of the binary hierarchy . CI-method supports the convexity of corresponding dependence of approximation errors for current color number g . In such a way CI-method ensures local convexity of the E dependence on the color numbers g .

If CI method is applied to the approximation of the image corresponding to the terminal values g, E_g on a convex curve during a hierarchy construction, then as a result of approximation error minimizing, the convexity may be violated for the previous value E_{g-1}. Therefore, we have to return to the approximation of the image in $g-1$ colors, then, to approximation in $g-2$ colors, etc. In order not to worry too much about maintaining convexity when programming, it is most convenient to program the monitoring and correction of convexity violations at the lowest level of the cluster merging operation. This is achieved by embedding in the elementary procedure merging clusters the procedure for detecting a subset of pixels with convexity violation and immediate correcting convexity violations on the found pixel subset, using Ward's original method. Due to the limited number of involved image approximations with a different number of colors and the localization of the pixel sets processed by Ward's original method, the restoration of convexity, in our experience, does not take much time.

Regarding the CI-method, it should be noted that out of the three discussed methods, this is the most effective method for reducing the approximation error E, which, with the proper data structure, belongs to high-speed methods, improves the approximation in E and visually manifests itself more noticeably, the worse the quality original approximation of the image. Characteristically, the CI-method works well even for a segmental hierarchy, despite the fact that the condition of pixel connectivity in clusters limits the minimization of the approximation error E [17].

The joint implementation of the modernized Ward's clustering and the clustering by CI method ensures the actual real-life minimization of the approximation error E to suboptimal values. But for the terminal minimization of E close up to optimal values and obtaining the optimal image approximations themselves [18], as well as for converting one binary hierarchy of suboptimal approximations to another, one more clustering method is needed. This is so-called K-meanless method, which is a corrected and modernized K-means method.

3.4. K-means and K-meanless methods

K-means method [4,5,19, 20] is one of the simplest methods of cluster analysis, developed in the last century for calculations on arithmometers, which were poorly suited for extracting square roots from numbers. The main disadvantage of K-means method is that this venerable method in the formula for increment $\Delta E_{correct}$ of the total squared error $E_{correct}$, caused by reclassifying a set of pixels from one pixel cluster into another, omits the dependence of $\Delta E_{correct}$ on the number of pixels to be reclassified as well as the pixel numbers in the donor and acceptor clusters, that in accurate formula (5) are included in the corresponding coefficients expressed via square roots [7]. Due to the indicated disadvantage, K-means method, generally speaking, does not provide the calculation of the available minimum for E, since ends before it can be reached. The insufficiency of minimizing E by K-means method is aggravated if only individual pixels are reclassified from cluster to cluster, and when larger sets of pixels are taken into account, the role of ignored coefficients increases.

A feature of the K-means method is that it provides a slight reduction in the approximation error E compared to E reduction providing by Ward's and CI methods. This becomes a disadvantage if K-means is applied to non-suboptimal image approximations.

In image processing, K-means is the most popular cluster analysis technique. Often, the application of this method exhausts the developer's familiarity with the detection of objects by pixel clustering. In fact, the essentially negative experience in E-minimization of the traditional application of K-means method is replicated in many papers and, in our purely subjective opinion, only confuses each another generation of researchers.

S.D. Dvoenko in the process of experimental research guessed to replace K-means with K-meanless method using the functional E to be minimized itself [21]. The programmatic implementation of K-meanless method uses the same data structure as K-means method. Therefore, in most applied problems oriented to minimizing E using K-means method [22,23], it makes sense to replace it with the modernized K-meanless method in order to improve the E minimizing results and expanding the application area of software implementation.

4. Conclusion

The novelty of the study lies in the fact that the following findings are proposed.

1. Experimental confirmation of the efficiency of image approaching by an approximation hierarchy described by a convex sequence of approximation errors. Theoretically, this is obvious if the sequence of minimal errors of optimal approximations is convex, as it is suggested.
2. Refinement and modernization of classical methods of cluster analysis in combinations with each other.
3. Formalization, interpretation and quantitative assessment of the effect of simplifying object detection when merging the target image and the image with object examples into single joint image.
4. Statement and implementing a software experiment using the example of iris detection.

The study is carried out within the framework of the pixel clustering model, which, in addition to the modernized methods, provides for the following developments.

1. Formalization of the concepts of structured images, objects and superpixels (image elements), distinguished by a computer from each other.
2. Statement and development of a solution to the problem of approximating a non-hierarchical sequence of optimal approximations using a binary hierarchy of approximations, which, like optimal approximations, is described by a convex sequence of approximation errors and contains the optimal approximation in a given number of colors [16,18].
3. Definition and examples of computing an irregular hierarchy of superpixels, providing error-free obtaining of a series of optimal approximations by merging superpixels [18].
4. Computations using the Algebraic Multilayer Network (AMN) data structure, which provides a high-speed implementation of the image pixel clustering model in terms of dynamic Sleator-Tarjan trees and cyclic graphs [18].

The theoretical significance of the study lies in the fact that it is useful for cluster analysis, in which systems of methods for minimizing the approximation error E or, equivalently, the standard deviation σ are not perfectly developed. The practical significance of the study lies in the fact that the relevance of strengthening software implementations of classical E minimization methods in publicly available software tools such as MatLab is substantiated.

Probably, some programmers will be confused by the brief description of the methods given in the paper, since they simply do not know how to work with arbitrary sets of pixels as simply as with individual pixels. This is supported by AMN. However, it is possible to practically verify results, as source codes and executable modules of programs the author makes freely available to whom it may concern on his web pages of the ResearchGate and MachineLearning sites. Among the source codes, the code of the program [24] may be of interest, as it implements a perfect data structure for executing Ward's method by image parts.

In this paper, we paid attention to the fact that the method of dividing an image into parts in the modernized Ward's method speeds up calculations and in a predictable way affects the resulting variants of the *correct* image approximation hierarchy, which is described by a convex sequence of approximation errors. Optimization of the choice of splitting the image *into parts*, i.e. splitting the set of N pixels into clusters, is the subject of ongoing research. The results are planned to be published in the next paper.

5. Acknowledgements

This research was funded within the framework of the budgetary theme FFZG-2022-0006 (Theoretical and Technological Basis for Operational Processing of Big Heterogeneous Data Streams in Sociocyberphysical Systems).

6. References

[1] I. G. Khanykov, M. V. Kharinov, Ward's Pixel Clustering for Color Image, Proc. of the 27th Int. Conf. on Computer Graphics and Vision (Graphicon'2017), Lomonosov Moscow State University, 2017, Perm, pp. 232–235.
[2] V. A. Nenashev, I. G. Khanykov, Formation of Fused Images of the Land Surface from Radar and Optical Images in Spatially Distributed On-Board Operational Monitoring Systems, J. Imaging, 7(12) (2021) 251. doi: 10.3390/jimaging7120251.
[3] V. A. Nenashev, I. G. Khanykov, A. P. Shepeta, Device for multiple-angle synthesis of complex image of the Earth's surface, 2021. Patent No. RU 2756904 C1, Filed August 24th., 2020, Issued June 10th., 2020.
[4] S. A. Ajvazyan, V. M. Buhshtaber, I. S. Enyukov, L. D. Meshalkin, Applied Statistics: Classification and Dimension Reduction, Finance and Statistics, Moscow, 1989.
[5] I. D. Mandel, Cluster Analysis, Finance and Statistics, Moscow, 1988.
[6] F. Murtagh, L. Pierre, Ward's hierarchical agglomerative clustering method: which algorithms implement Ward's criterion, Journal of classification, 31(3) (2014) 274–295.
[7] M. Kharinov, Reclassification formula that provides to surpass K-means method, arXiv preprint, arXiv:1209.6204 [cs.CV], 2012.
[8] R. Nock, F. Nielsen, Statistical Region Merging, IEEE Trans. Pattern Anal. Mach. Intell. 26(11) (2004) 1452–1458.
[9] M. V. Kharinov, Development of dynamic data structures for an automatic image recognition system, Ph.D. thesis, SPIIRAS, St. Petersburg, 1993.
[10] Yu. I. Monich, V. V. Starovoitov, D. I. Samal, Experimental complex of personality recognition programs based on the iris of the eye, Elektronics Info, No 5 (2010) 65–68.
[11] I. A. Matveev, Optimization of the circular path as a method for selecting and refining the boundaries of the iris in the image of the eye, Proceedings of the Russian Academy of Sciences. Theory and control systems, No. 5 (2011) 95–101.
[12] N. P. Grishenkova, D. N. Lavrov, Review of methods for identifying a person by the iris of the eye, Mathematical structures and modeling, 1 (29) (2014) 43–64.
[13] K. A. Gankin, A. N. Gneushev, I. A. Matveev, Segmentation of the iris image based on approximate methods with subsequent refinements, Proceedings of the Russian Academy of Sciences. Theory and control systems, No 2 (2014) 80–94.
[14] I. Solomatin, I. Matveev, Detecting visible areas of iris by qualifier of local textural features, J. Mach. Learn. Data Anal., 1(14) (2016) 1919–1929.
[15] K. Malinowski, K. Saeed, An iris segmentation using harmony search algorithm and fast circle fitting with blob detection, Biocybernetics and Biomedical Engineering, 42(1) (2022) 391–403.
[16] M. V. Kharinov, A. N. Buslavsky, Object Detection in Color Image, in: Proceedings of the 14th International Conference on Pattern Recognition and Information Processing (PRIP'2019), Publishing Center of BSU, Minsk, 2019, pp. 43–47. doi: 10.13140/RG.2.2.28493.28640.
[17] M. V. Kharinov, I. G. Khanykov, Optimization of piecewise constant approximation of segmented image, Proceedings of SPIIRAS (2015) 3(40), 183–202.

[18] M. V. Kharinov, Clustering for Detection of Binary Object Hierarchy Using Modernized Classical Clustering Methods, Proceedings of the 15th International Conference on Pattern Recognition and Information Processing (PRIP'2019), Publishing Center of BSU, Minsk, 2021. pp. 198–201.
[19] Hugo Steinhaus, Sur la division des corps matériels en parties, Bull. Acad. Polon. Sci., (1956) C1 III IV, 801–804.
[20] I. V. Pustoshkina, D. A. Pustoshkin, J. Rodrigues, D. Gupta, A. Khanna, A. Shankar, C. Seo, G. P. Joshi, Automatic Vehicle License Plate Recognition Using Optimal K-Means With Convolutional Neural Network for Intelligent Transportation Systems, IEEE Access (2020) 8, 92907–92917.
[21] S. D. Dvoenko, Meanless k-means as k-meanless clustering with the bi-partial approach, in: Proceedings of the 12th International Conference on Pattern Recognition and Information Processing (PRIP'2019), Publishing Center of BSU, Minsk, 2014, pp. 50–54.
[22] A. K. Jain, Data clustering: 50 years beyond K-means, Pattern recognition letters (2010) 31(8), 651–666.
[23] J. Blöomer, C. Lammersen, M. Schmidt, C. Sohler, Theoretical Analysis of the k-Means Algorithm – A Survey, arXiv preprint, arXiv:1602.08254 [cs.DS], 2016.
[24] M. V. Kharinov, A. N. Buslavsky, Certificate of state registration of the program for computers No. 2018614015 dated 09/18/2018: “The program of hierarchical clustering of pixels of color image by Ward's method with buffering and taking into account the repetitions of the minimum increments of the total quadratic error,” RIA Registration number: AAAA-G18-618122790062-6.