Nearest neighbor classifiers for color image segmentation

Wojciech Bieniecki, Szymon Grabowski

Abstract – We present a class of simple algorithms for color image segmentation based on the Nearest Neighbor (1-NN) decision rule. The feature vector for each pixel in the image is constructed from color components in HSI space. Since processing all pixels with 1-NN rule is time-consuming, we decided that only some “crate” pixels will be classified straight with 1-NN, while the others will then be labeled according to their spatial neighborhood containing pixels already classified, and only in relatively rare cases sent to a “global” 1-NN classifier. We test the accuracy and computational efficiency of the algorithms applied to medical image segmentation.

Keywords – image segmentation, clustering, nearest neighbor.

I. INTRODUCTION

Segmentation is an important step in the process of image analysis. During this transform the image is divided into parts that correspond with objects or areas of the real world contained in the image. A complete segmentation defines a set of disjoint regions uniquely corresponding with objects in the input image, while a partial segmentation divides an image area to regions not directly corresponding with objects but, for example, with the classes of objects.

An example of such approach is clustering also known as pixel classification. In this group of algorithms, the image is processed globally by assigning each pixel to one of the defined classes, for example objects A, objects B, or background, without respect to object consistence. The technique is often based on thresholding the intensity level of each pixel. Finding an optimal range of intensity allows separating the background from the objects. More sophisticated clustering algorithms are based on image intensity histograms. Such a concept allows finding the optimal number of classes and automatically setting proper threshold values for pixel classification.

Clustering techniques were recently adopted for color images. In fact, color image segmentation requires building and analyzing the pixel distribution histogram in 3D space, which implies some difficulties.

Sometimes we have a priori knowledge about the image content: the number of object types and their color properties, and our task is to perform a quantitative analysis of the image: to measure the number of objects of specified classes and accurately establish their contours and area. In some cases, when we have some manually segmented pictures of the same group, we can use them as a training set for k-NN based image segmentation.

The non-parametric k Nearest Neighbors (k-NN) decision rule was introduced half a century ago [FH52] and is still one of the most effective classifiers from the point of accuracy. A sample to classify from a given vector space is assigned to the class most heavily represented among its k nearest neighbors from the reference (training, learning) set.

\[ \rho(x, x_i) = \|H - H_i\| + \|S - S_i\| + \|I - I_i\| + \|V_x - V_{x_i}\| \] (1)

Let us assume that in the color segmentation task we know beforehand the set of object type classes plus background. The samples are 4-feature vectors, where the features are color components H, S and I, and additionally the intensity gradient estimated at the given sample. We use the city-block distance (Eq. (1)) which appears computationally faster and usually at least as good as the Euclidean one.

The method has been successfully tested on color microscopic images for quantitative cancer cell nuclei measurement [BGS03].

II. TWO-LEVEL CLASSIFICATION

One of the drawbacks of the nearest neighbor classification are high computation costs. High resolution pictures require processing of hundreds of thousands pixels, and in the case of nearest neighbor methods the number of comparisons equals to the size of the test set multiplied by the size of the training set. We believe that currently (with the chosen features) the training set may contain hundreds of samples (labeled manually by an operator/expert), while increasing it will not lead to significant changes in classification accuracy. Still, in this case the process of classification may take up to a few minutes for large images.

The classification can be sped up by reference set reduction. It consists in replacing the original set with a smaller, but possibly representative one. We have successfully tested a stochastic reduction routine proposed by Skalak [Sk94]. One of its assets is reduction to the user-selected size. Other benefits of Skalak’s algorithm are simplicity (which enables further modifications), good accuracy (overfitting avoidance to some degree) and quick generation of the output set. In our experiments, the visual quality of color-segmented images with reference to Skalak’s reduced set is only negligibly worse than the quality of images obtained on the basis of 1-NN with full reference set, even though the specified reduction degree was about 15 [BGS03].

The specificity of the test set, which in our task contains all pixels of a given image, allows for another approach to processing time reduction. Normally, the k-NN classifier treats the test set as an unordered set of points, without a respect to their spatial dependencies. Nevertheless, in many image processing applications, the spatially close samples...
have tendency to belong to the same classes, which is justified by the supposition that natural images are discretizations (in spatial and color/intensity domains) of some continuous 2D functions. This simple observation inspired us for taking into account both the similarity of samples in the feature space and their spatial proximity.

In the task of cell nuclei recognition in a pathomorphological microscopic image (details on our project can be found in [BGS03]), the nuclei have circular shapes and are never very small on the image; moreover, detecting their boundaries with great precision is not crucial. These rather favorable assumptions make possible to design a classifier which processes only a fraction of all test pixels (before all the other samples) with plain 1-NN rule (which, as mentioned above, is relatively slow). The other samples are submitted for a two-level classifier: first a fast spatial neighborhood criterion is applied, and only if its decision is not confident enough, the plain 1-NN rule (with reference to the full or reduced training set) is launched. The idea of multi-stage (cascade, multi-level) classifiers is well-known in the field of pattern recognition; set) is launched. The idea of multi-stage (cascade, multi-level) classifiers is well-known in the field of pattern recognition; the specifics of our application make possible to achieve a significant processing cost reduction for the price of small or even negligible accuracy loss, if the details of the scheme are chosen appropriately.

Let the set of pixels to classify in the first pass (with 1-NN) be denoted by \( F \). Let the set of all pixels on the image be denoted by \( L \) (“lattice”). The general idea is to select \( F \) in such a (regular) way that each pixel \( x \) from \( L \setminus F \) has at least one neighbor in \( F \) (according to 4- or 8-neighborhood), which will serve as fast as possible criterion for acceptance or rejection of \( x \) at this level of classification. Such a broad concept poses a natural trade-off between speed and recognition accuracy.

We consider three different sets \( F \): chessboard (Fig. 1a), crate (Fig. 1b) and modified crate (Fig. 1c).\[ a) \quad b) \quad c) \quad d) \]
Fig. 1. Sets \( F \) (classified in the first pass): (a) chessboard, (b) crate, and (c) modified crate. In (d) it is shown that “crosses” cover the whole grid.

The ratios \( F/L \) are presented in Table I. Generally, the sparser set \( F \) is, the greater classification speedup can be reached, but also with the worse accuracy. Of the three sets, the modified crate is the sparsest.

<table>
<thead>
<tr>
<th>Method</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>all points</td>
<td>100</td>
</tr>
<tr>
<td>chessboard points</td>
<td>50</td>
</tr>
<tr>
<td>“crate” points</td>
<td>25</td>
</tr>
<tr>
<td>“modified crate” points</td>
<td>20</td>
</tr>
</tbody>
</table>

Note that the closest points in such set can be connected with chess knight movement. If each point of \( F \) is the center of a cross containing five points in \( L \), then it can easily be “seen” (Fig. 1d) that \( L \) may be covered with such crosses. Below we prove it formally (assuming, in order to avoid tedious handling of the image border cases, that lattice \( L \) is infinite).

**Convention:** let \( d_1 \) denote the Manhattan (city-block) distance.

**Definition:** The set \( S \) of five planar lattice points such that there exists a point \( o(x,y) \in S \) for which we have \( d(o,p) \leq 1 \) for any \( p \in S \), will be called a 5-cross (see Fig. 1d).

**Theorem 1:** The planar lattice \( L \) may be covered with non-overlapping 5-crosses.

**Proof:** To find the cover, it is enough to specify the locations of center points \( o_i \), \( i = 1, 2, ..., \) of all the used crosses. We show that all the center points \( o_i(x,y) \), \( i = -2, -1, 0, 1, 2, ... \) may conform to the formula

\[
y = 2x + 5m,
\]
where \( x = ..., -2, -1, 0, 1, 2, ...; m = ..., -2, -1, 0, 1, 2, ... \) to satisfy the requirement.

Note that correctness of the chosen set of center points requires meeting the two conditions:

a. there are no such 5-crosses \( S_i, S_j, i \neq j \), that \( p_i \in S_i \) and \( p_i \in S_j \) for some point \( p_i \);

b. each lattice point must belong to some 5-cross \( S_i \).

To prove condition a., it is enough to show that for any \( p_i(x,y) \) and any \( S_i, S_j, i \neq j \) and their center points \( o_i \) and \( o_j \), respectively, the inequality \( d(p_i,o_i) + d(p_i,o_j) \geq 3 \) holds, as it immediately implies the fulfillment of condition a. taking into account the triangle inequality required for any metric.

Trivially, for any center points \( o_i(x,y) \) and \( o_j(x',y') \), \( i \neq j \), if \( x = x' \), then \( d(o_i,o_j) \geq 5 \) (from formula (2)). Also it is easy to check that if \( |x - x'| = 1 \), then \( d(o_i,o_j) \geq 3 \). If \( |x - x'| = 2 \), then again \( d(o_i,o_j) \geq 3 \). Finally, if \( |x - x'| \geq 3 \), then obviously \( d(o_i,o_j) \geq 3 \) again.
We thus showed that for any center points \( o_i \) and \( o_j \), 
\( i \neq j \), 
\( d_i(o_i, o_j) \geq 3 \). According to basic metric properties (symmetry, triangle inequality), we obtain immediately 
\( d_i(o_i, o_j) \geq 3 \Rightarrow d_i(p_i, o_i) + d_i(p_i, o_j) \geq 3 \), for any \( p_i \),
which ends this part of the theorem proof.
To prove condition \( b \), let us notice that any lattice point 
\( p_i(x, y) \in L \) meets the formula
\[
y = 2x + 5m + r,
\]
for some \( m, r \), where
\[
x = \ldots, -2, -1, 0, 1, 2, \ldots; m = \ldots, -2, -1, 0, 1, 2, \ldots; r = 0, 1, 2, \ldots.
\]
Let \( R(p_i) = r \) and with any \( p_i \) let there be associated three more numbers, \( x, y \) and \( m \), according to Eq. (3). Therefore, we have five cases. It is easy to see that any lattice point 
\( p_i (x, y) \)
belongs to set \( S_i \) with center point \( o_i(x, y) \in S_i \) satisfying the formula
\[
y = 2x + 5m,
\]
such that
\[
\tilde{x} = x, \quad \tilde{y} = y, \quad \text{and} \quad \tilde{m} = m, \quad \text{if} \quad R(p_i) = 0 \quad \text{or} \quad R(p_i) = 1;
\]
\[
\tilde{x} = x + 1, \quad \tilde{y} = y, \quad \text{and} \quad \tilde{m} = m, \quad \text{if} \quad R(p_i) = 2;
\]
\[
\tilde{x} = x - 1, \quad \tilde{y} = y, \quad \text{and} \quad \tilde{m} = m + 1, \quad \text{if} \quad R(p_i) = 3;
\]
\[
\tilde{x} = x, \quad \tilde{y} = y + 1, \quad \text{and} \quad \tilde{m} = m + 1, \quad \text{if} \quad R(p_i) = 4.
\]
We showed that the proposed formula for placing the 5-crosses (or, equivalently, their center points) on the planar lattice meets the required conditions for lattice cover, which ends the proof.

We decided to terminate classification for any point from \( L \setminus F \) if both its neighbors (in 8-neighborhood sense) from \( F \) (the modified crate) belong to the same class. Below we prove that each point from \( L \setminus F \) has such two neighbors.

**Theorem 2:** For any point \( p_i(x, y) \in L \setminus F \), there exist two center points \( o_i(x, y) \) and \( o_j(x, y) \), \( i \neq j \), such that the following conditions are all met:

(a) \( d_i(p_i, o_i) + d_i(p_i, o_j) = 3 \); 
(b) \( |x - x_i| \leq 1; |y - y_i| \leq 1; \); 
(c) \( |x - x_o| \leq 1; |y - y_o| \leq 1 \).

**Proof:** Let one of such center points be the center point of the 5-cross to which \( p_i \) belongs; without loss of generality let it be \( o_i \). From the definition of a 5-cross, the twin (b) conditions are immediately fulfilled. Moreover, \( d_i(p_i, o_i) = 1 \). Subsequently, it is enough to show that there exists center point \( o_j(x_j, y_j) \), \( j \neq i \), for which \( |x - x_j| = 1 \) and \( |y - y_j| = 1 \).

Let us follow the conventions from the proof of Theorem 1. It is clear that any \( p_i(x, y) \in L \setminus F \) will meet formula (3), but \( r \) is now limited to 1, \ldots, 4. Assume \( r \) is 1. It is easy to see that
\[
o_i(x, y), \quad \text{where} \quad \tilde{x} = x + 1, \quad \tilde{m} = m, \quad \text{and} \quad \tilde{y} = 2\tilde{x} + 5\tilde{m},
\]
is the required center point. Analogously, the parameters for the remaining three cases \( (r = 2, 3, \text{or} 4) \) may be given without a difficulty, which ends the proof.

**The criteria of classification.** We have proved that for the \( F \) set constructed as “modified crate” any pixel to be classified (from \( \text{int}(L \setminus F) \) set, assuming \( L \) is finite and distributed over a rectangular area) has two geometrical neighbors. Image border pixels have one or two neighbors, depending on their position.

![Fig. 2. Majority (a) and unanimous (b) class voting for points not belonging to \( F \). Gray points that belong to \( F \) are assigned to one of the three classes with 1-NN rule, white points are classified according to their spatial neighborhood. Points '?' have to be classified by 1-NN in the second pass.](image)

If \( F \) is constructed as the chessboard or crate points (Fig. 1a or 1b, respectively), the criteria of classification may be defined as majority or unanimous voting (Fig. 2). In the case of unanimous voting, the pixel \( p \) is assigned to class \( c \) only if all of its spatial neighbors from \( F \) belong to the same class \( c \), otherwise pixel \( p \) must be classified with the training set. In majority voting variant, \( p \) is assigned to class \( c \), if \( c \) is the most frequent class among \( p \)'s geometrical neighbors from \( F \).

### III. THE EXPERIMENTAL RESULTS

The aim of the experiment was to compare several classification algorithms. We used two different test sets (images) with resolutions 700 x 525 pix and 1400 x 1050 pix, and three different training (reference) sets of sizes: 100, 200 and 667 points. This resulted in six combinations of training and test sets for each method.

Evaluation of the accuracy of classification methods in our application is a non-trivial question as human (expert) labeling of the many thousands of pixels is unrealistic. We therefore measured the agreement of our methods with the decisions made by pure 1-NN classifier which served as a yardstick. The measure of such an agreement was simply the fraction of pixels for which 1-NN and the particular method of ours yielded the same response. Of course, in addition to the accuracy, also speed measurements were performed.

The first comparison (Fig. 3) included the average fraction of points processed by 1-NN classifier in both passes. While the fraction of points handled in the first pass is fixed (and shown in Table 1), the result of the second pass depends mostly on the amount of noise in the image. It should be noticed that the overall fraction only quite weakly depends on the voting method. Not surprisingly, the pixels which are
submitted to 1-NN classification in the second pass are relatively infrequent.

The next test pertained to accuracy measurement (Fig. 4). This time we measured the percentage of points that matched full 1-NN classification. The performance achieved for modified crate variant (last bar in Fig. 4) is only negligibly worse than pure 1-NN classification (1.1%). Finally, we present a speed benchmark (Fig. 5). The processing routines have been implemented in C++ and compiled with MS Visual Studio on a Pentium III PC. We multiplied the sizes of test and training sets and divided the result by processing time. Average results have been expressed in megapixels (Mpix) per second. This test does not exactly reflect the computation complexity of the algorithms but presents the results in a legible way. Again, the winner was the version with the sparsest set $F$.

The image segmentation results obtained with the modified 1-NN classifier are promising. In respect of accuracy and computational efficiency taken together, they seem to be much better than pure 1-NN pixel classification. Nevertheless, the clustering methods using histogram analysis are still faster. On the other hand, nearest neighbor approach offers more accurate segmentation, because the feature vector may be constructed appropriately to the image specifics. In the future we intend to combine those two approaches to image segmentation in a hybrid algorithm.

REFERENCES

