Abstract: Data Matrix Codes (DMC) are established by these days in many areas of industrial manufacturing thanks to their concentration of information on small spaces. In today’s usually order-related industry, where increased tracing requirements prevail, they offer further advantages over other identification systems. This underlines in an impressive way the necessity of a robust code reading system for detecting DMC on the components in factories. This paper provides a solution for the Data Matrix Codes recognition. Using the Mean Shift Algorithm, the main information of the DMC pattern are computed and based on these, a virtual identification pattern is constructed. Seeking and matching over the image, the Data Matrix Code is located. We concentrate on Data Matrix Codes in industrial environment, punched, milled, lasered or etched on different materials in arbitrary orientation.

Key–Words: DMC Scanning Method, Industrial Data Matrix Codes, Data Matrix Code Recognition, Mean Shift

1 Introduction

The industrial Data Matrix Code (DMC) is a two-dimensional matrix bar-code consisting of dots (modules) arranged in a square. The information to be encoded can be text or raw data [4]. Error correction codes are added to increase the robustness of the code: even if they are partially damaged, they can still be read. As more data is encoded in the symbol, the number of modules (rows and columns) increases from \(8 \times 8\) to \(144 \times 144\). For industrial purposes, Data Matrix Codes are directly marked on the parts by different techniques like milling, punching or lasering.

In Fig. 1, it is presented the diagram of the DMC system recognition. The image captured by the video camera is thresholded and labeled. Applying the Mean Shift algorithm to the labeled objects from the image, one estimates the distance between the modules, the Data Matrix pattern’s orientation and the nominal area of the modules. Having this information, one can create a virtual identification pattern and a virtual timing pattern. Based on these patterns, one looks into the labeled objects structure and are held only the objects (called modules), which fall within the respective patterns. Using the virtual identification and timing pattern, a matching process will preserve only those modules, which are consistent with the two patterns. Knowing the position of the pattern, the Data Matrix Code can be scanned.

2 DMC acquisition and labeling

After the gray image is thresholded, in the obtained image there are a set of objects. An object is a group of pixels that are connected to each other. In Fig.2, the first image indicates the connectivity using \(N_4(p)\) and the second image indicates \(N_8(p)\) respectively. Each object is labeled with a different value, from ‘1’ to the number of objects.
3 Mean Shift algorithm

Mean shift is a non-parametric feature-space analysis technique, a so-called mode seeking algorithm. It locates the maxima of a density function given discrete data sampled from that function [1]. By iteratively moving an analysis window starting from an initial estimate \( x \), it detects the modes (maxima) of the density.

Let a kernel function \( K(x_i - x) \) be given. This function determines the weight of nearby points for re-estimation of the mean. Typically one uses the Gaussian kernel on the distance to the current estimate \( x \), which is the length of the segment is equal with half of the angles. In our case the dimension \( V \) is the number of elements from inside \( V \) [5].

\[
y = \frac{\sum_{x_i \in N(x)} K(x_i - x) x_i}{\sum_{x_i \in N(x)} K(x_i - x)},
\]

where \( N(x) \) is the neighborhood of \( x \), a set of points for which \( K(x) \neq 0 \). The mean-shift algorithm now sets \( x \leftarrow m(x) \), and repeats the estimation until \( m(x) \) converges, Fig. 3 [7].

If dense regions (or clusters) are present in the feature space, then they correspond to the mode (or local maxima) of the probability density function. For each data point, Mean shift associates it with the nearby peak of the data-set’s probability density function. For each data point, Mean shift defines a window around it and computes the mean of the data point. Then it shifts the center of the window to the mean and repeats the algorithm till it converges. After each iteration, one can consider that the window shifts to a more denser region of the data-set.

4 DMC orientation estimation with Mean Shift

Let’s consider the labeled image of a Data Matrix Code and all the coordinates of the objects from the image. One starts from the first module to the last one from the image and calculates the angle of the straight line created by the modules, using the equation:

\[
\alpha_{ij} = \arctan \frac{x_j - x_i}{y_j - y_i},
\]

where \((x_i, y_i)\) and \((x_j, y_j)\) are the coordinates of modules \( i \) and \( j \).

The calculates angels are stored in a connected list in increasing order.

The analysis window’s size should not be bigger than half of the angles. In the example from Fig. 5(b) the the size is chosen as half of the space taken by the angles’ value.

The general expression for the non-parametric density estimation is:

\[
P(x) \approx \frac{k}{N \cdot V},
\]

where \( V \) is the volume of the \( x \) position, \( N \) is the total number of the elements, and \( k \) is the number of elements from inside \( V \) [6].

If we assume that the region surrounding the examples \( k \) is a Hypercube with sides of length \( h \), centered in an estimated point \( x \), its volume is given by \( V = h^D \), where \( D \) is the number of dimensions, Fig. 4.

In our case the dimensiom \( D = 1 \), this means that \( V \) is the length of the segment, and \( k \) is the number of the values contained in the segment [5].

Next, based on Fig. 5(b), the Mean Shift steps are described [2]:

- in the first step, the analysis segment is centered in the center of the data base coverage. The length of the segment is equal with half of the coverage. Using the equation (3), the density
center of the points from the segment is computed. The center of the segment is moved in the position calculated by the equation, Fig. 5 (b),

- in the second step, the segment would exceed the coverage of the angels. That is why, it is re-calculated its dimension using the new values form the segment and its real length.
- the procedure is repeated until it converges to mean, $3^{th} - 6^{th}$ steps.

We can notice that the peak position of the function densities stabilizes after each iteration. The Mean shift does not require a pre-parametrization process, the algorithm is adapting according to the values of the area subject to estimation.

5 Estimating the distance and the nominal area of the modules with Mean Shift

To estimate the distance between the modules and the modules’ nominal area we use the same method presented in the previous chapter. To create the structure of the distances we calculate the Euclidean distance between each object and its eight nearest neighbors. In Fig. 6 it is presented the histogram of the distances from the image and the ongoing of the Mean Shift, implemented to estimate the nominal area of the Data Matrix Code modules.

To better understand the method previously described, Fig. 8 presents the Data Matrix modules position in a 3-D space. The three dimension of the space are given by the areas, distances, and the angles of the modules. The points represented with red in the figure, are the Data Matrix Code modules. The other objects from the image are represented with blue.
DMC pattern recognition

For the DMC pattern recognition, we start from its specifications, which consist of two adjacent edges marked entirely with modules and two adjacent edges marked only in the odd positions. It is also known that the pattern has a square shape and of course its orientation, the distance from the modules and the modules’ nominal area. With these inputs, a virtual pattern is built for the finder pattern and one for the timing one. Knowing that, the number of rows and columns of the DMC pattern varies between $8 \times 8$ and $144 \times 144$, consider the size of the DMC pattern of $8 \times 8$. Using the matrix of coordinates, relations (4, 5), a pattern is built that overlaps the structure of coordinates of the objects from the Data Matrix Code image. Finally, it is searching for the objects that matches to the virtual pattern.

$$L_F = \begin{bmatrix} L_{8,1} & \cdots & - & - \\ L_{7,1} & \cdots & - & - \\ \vdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ L_{1,1} & L_{1,2} & \cdots & L_{1,7} & L_{1,8} \end{bmatrix}, \quad (4)$$

where: $L_{1,2} = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$, $L_{2,1} = \begin{bmatrix} x_0 - d \sin \alpha \\ y_0 + d \cos \alpha \end{bmatrix}$,

$L_{1,8} = \begin{bmatrix} x_0 + 7d \cos \alpha \\ y_0 + 7d \sin \alpha \end{bmatrix}$

$L_{8,1} = \begin{bmatrix} x_0 - 7d \sin \alpha \\ y_0 + 7d \cos \alpha \end{bmatrix}$

(5)

The same operation is done for the timing pattern. It is known that the timing pattern has in its structure modules marked alternatively. Using this information, a matrix pattern of $4 \times 4$ is built, relations (6, 7), and it is searching in the structure of the coordinates from the image.

$$L_T = \begin{bmatrix} L_{1,8} & - & - & - & - \\ L_{1,6} & - & - & - & - \\ L_{1,4} & - & - & - & - \\ L_{1,2} & - & - & - & - \\ - & L_{1,2} & L_{1,4} & L_{1,6} & L_{1,8} \end{bmatrix}, \quad (6)$$

where: $L_{2,1} = \begin{bmatrix} x_0 - d(\cos \alpha + 2 \sin \alpha) \\ y_0 - d(\sin \alpha - 2 \cos \alpha) \end{bmatrix}$,

$L_{1,4} = \begin{bmatrix} x_0 + 2d \cos \alpha \\ y_0 + 2d \sin \alpha \end{bmatrix}$,

$L_{4,1} = \begin{bmatrix} x_0 - 2d(\cos \alpha + 2 \sin \alpha) \\ y_0 - 2d(\sin \alpha - 2 \cos \alpha) \end{bmatrix}$.

(7)

For the safe recognition of the Data Matrix pattern, the two patterns are rotated with
(90°, 180°, 270°), Fig. 9 - 10, due to the square form of the Data Matrix pattern.

The objects coordinates that satisfy the conditions to fit simultaneously into the built patterns are chosen as candidates for a valid DMC.

From all candidates, one selects the patterns that have the margins two by two at equal distances and the edges orientated towards the exterior. Fig. 11(a, b) presents the possibilities of the finder and timing patterns, and in Fig. 11(c) presents the selected patterns.

The detected patterns have a 8 × 8 modules size, but the size of the Data Matrix code being unknown, the other parts of the fingerprint are not known either. Drawing imaginary lines over the four detected edges, their intersection leads us to the discovery of the other two corners of the pattern. Thus we know the four corners and the position of in the Data Matrix Code pattern.

7 Conclusion

This article offers a solution for finding the orientation of the industrial Data Matrix Code. Based on the objects’ coordinates from the image, we use the Mean Shift algorithm to estimate the pattern orientation, the distance between the modules and their area. The Mean Shift algorithm does not need to be parameterized, is adapting by the information from the data set subject to be estimate.

Form the technical data information of the DMC pattern, it is also known, that the DMC is composed of a finder pattern and a timing one, and the size of the patterns varies from 8 × 8 up to 144 × 144 [4]. Using the estimated data about the modules, it builds two virtual patterns. The finder pattern consists of a matrix 8 × 8 size, while the timing pattern size is 4 × 4. It searches in the structure of the images, the objects that overlap simultaneously the two given patterns. The finder and the timing patterns that are diagonally opposite and form a square represent the Data Matrix Code final pattern.

Fig. 12 - 14 present some tests of the proposed method and its results. The finder pattern modules are pointed out with blue color, and the timing pattern modules with red. In Fig. 12(a, b) - 14(a, b), there it is shown the possibility of different systems of modules that could match the built patterns, but there are selected only the ones that meet the condition to form a Data Matrix pattern, Fig. 12(c) - 14(c).

Note that there are different types of materials tested, and the codes of the patterns are of different sizes and with different orientations. The Data Matrix pattern size is precisely recognized in all presented figures. Its orientation and even the material on which it is marked does not influence the recognition system. Having the position of the Data Matrix Code one can scan inside of it for extracting the desired information [3]. The identification of the Data Matrix Codes, whose image is acquired with a video camera oriented non-perpendicularly on the fingerprint code area, is possible due to the method’s ability of recognizing the pattern. This method offers flexibility of use, with the capability of recognition of the Data Matrix Codes marked on non-flat surfaces.
(a) Possible finder patterns  (b) Possible timing patterns

(c) Industrial DMC pattern

Figure 14: Iron, pattern size 1.6 cm

Acknowledgements: This work was partially supported by the strategic grant POSDRU 6/1.5/S/13, (2008) of the Ministry of Labour, Family and Social Protection, Romania, co-financed by the European Social Fund Investing in People.

References:


