THE RESOLUTION ENHANCEMENT BY SOFTWARE METHOD OF RANDOM SHIFT

Tomas Ponik, Jaroslav Placek

FMFI UK, Bratislava

Abstract

The goal of the paper is to describe one method of software resolution enhancement of records obtained by CCD cameras or other devices able to view and digitalize the image. The paper is divided into two parts. The first one explains the theoretical principle of the random shift method. The second part is dedicated to application HimRes which tests these method and creates resulti images with enhanced resolution.

Keywords: CCD camera, linear properties of the CCD camera, light-sensitive element, raster, center of gravity

Used terms: pixel, object, record, random shift, detection object, resulting image, actual pixel, covering, overlap (definitions are explained bellow).

Introduction

The CCD camera is nowadays a well-known devide used in many areas of science, development and industry. It is cheap and easy to use. Analog CCD cameras produce analog signal (PAL or NTSC) as the output which has the resolution of 768x576 pixels. This can be enough for simple image analysis but in many cases user wants to see further, to see details and other features in the image which is impossible with the use of analog cameras. The digital CCD camera with resolution up to 1280x1024 exists already but it‘s price is 40 times higher than ordinary analog one.

There is a way to enhance the resolution of camera image by snapping more pictures of the same scene and computing the new image using suitable features of these image series.

The explanation of used terms:

CCD camera – The device for viewing of the image.
CCD-elements – The elements of the CCD camera sensitive to light intensity which are created by MOS condensors or PN transitions.
*Light-sensitive element (*shortly *element)* – CCD-element, which represents pixel in raster image.
Pixel – Two-dimensional image element, which has a color. The definition of color is bellow.
Scene – The part of the real world viewed by te CCD camera or other device able to record the image.
Object – The part of the scene.
Record - Scene viewed by the CCD camera and digitalized to 2D space. Every record is with respect to other obtained records shifted, resp. rotated in the image plane.
Random shift – Informations of record positions in choosed coordinate system are not contained in the input data. The record capture is performed by random shift of the CCD camera, resp. of the viewed object.
Detection object – The circle or other geometric solid invariant to the operation of rotation which is used by determination the relative position of the record with respect to other records.
*Image with compressed sampling (*shortly *resulting image)* – Raster image with higher pixel density as on the records.
Actual pixel in resulting image (shortly *actual pixel*) – Actually processed pixel in resulting image.
The pixel color – Integer value representing the illuminance recorded by the image element. During creation of resulting image the domain for these values is the set of real numbers between 0 and 255.
The cover of actual pixel by the record pixels parts – Parts of the record pixels cover actual pixel if and only if the actual pixel surface is created by the unification of these parts (Figure 1).
Overlap of actual pixel by the record pixels – Record pixels overlap the actual pixel if and only if their parts cover actual pixel (Figure 1).

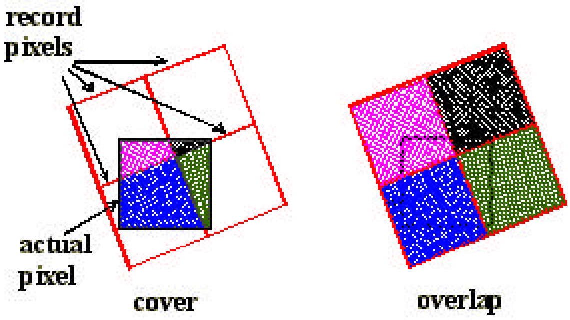


***Figure 1** Cover and overlap of actual pixel*

The coordinate system of image plane (shortly *Choosed coordinate system*) –The two-dimensional cartesian coordinate system is used in this paper. It is oriented according to the first captured record (Figure 2). The coordinate axis are parallel to record rows and columns. The center of the coordinate system is placed in the intersection of first row and first column of record.

For easier explanation only grayscale images are considered with black color evaluated as 255 and white color as 0. In the next explanation the resolution enhancement of factor two is considered.

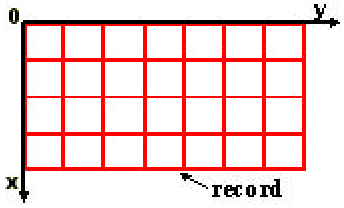


Figure 2 *Choosed coordinate system*

The method of random shift

The method is based on record averaging. The principle of the method is based on the as best as possible detection of relative positions of the record according to one selected record (usually the selected record is the first snapped) with the accuracy of the one tenth of the pixel.

Using of records positions information they are translated and rotated in the way that in selected coordinate system all the objects of the scene have identical coordinates for each snapped record of the scene. When imagined the records as overlays, then it is the same as translation and rotation of the overlays to get a feeling that it is one overlay.

The neigbouring pixels parts sizes of the record (label – *the contributions*) which cover the actual pixel are determined and the weighted average of color values of that contributions are calculated. Figure 3 illustrates one of the possible cases. S is the area of actual pixel, S_1, S_2, S_3, S_4 are areas of record pixles parts which cover the actual pixel. I is the contribution to resulting color value of actual pixel. I_1, I_2, I_3, I_4 are color values of record pixels overlapping the actual pixel. Then it is valid:

$$S = S_1 + S_2 + S_3 + S_4 \qquad (1)$$
$$I = I_1 * S_1 + I_2 * S_2 + I_3 * S_3 + I_4 * S_4 \qquad (2)$$

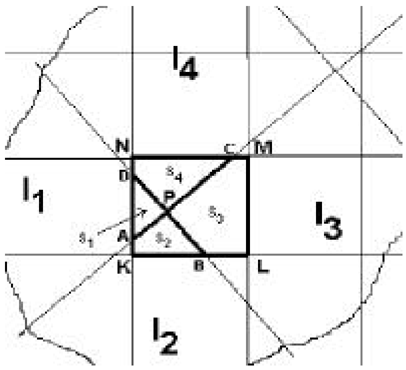


Figure 3 *Contribution of the record to resulting color of actual pixel*

When by this means achieved *record contributions to the resulting intensity of actual pixel* are averaged through all suitable records, the actual pixel color value is achieved.

For determination of relative position of the record according to the first snapped two detection objects are placed in the scene. After thresholding of the records and successive localization of these detection objects their centers of gravity are calculated. This will help to determine the relative positions of the records.

The determination of detection objects

In the next it is supposed that every detection object is created by pixels with darker color as their neighbours in the vicinity of detection objects.

The first thing to do is to isolate the detection objects from the background. The one of the effective and easy implementable methods fairly solving this problem is the method of thresholding. All the pixels lighter than certain threshold value are filled with white color (the background color). The rest of them have original color which is used in the calculation of the weighted average in the determination of the detection objects centers.

I suppose, that every detection objects is continuous. Because the thresholded image can contain in addition to detection objects others objects, it is necessary to find the way how to ignore them. In my application the detection objects are selected from all the detected objects by interactive user determination of one point from the detection object. For detection of the areas of the detection objects which are bounded by the background color I use the method of the area filling (seed fill).

When I know which pixles belong to detection objects I can determine its centers. Because the quality of the result strongly depends on the accuracy of the determination of that centers, for their computation I use several methods.

1. The method of simple average, where the center is the arithmetic average of the coordinates of all detection object pixels.

$$T = \frac{1}{N}\sum_{\Omega} X_i \, , \qquad (3)$$

T - the coordinates of the detection object center
N - the number of pixels of the detection object
X_i - the coordinates of the i-th pixel of the detection object (i = 1, ... N)
Ω - the set of all detection object pixels

2. The method of simple weighted average, where the center coordinates are calculated as the weighted average of all detection object pixles and the weight of every point of detection object is its color. E.g., in my application the darker is the point the more important it is in the calculation of the center.

$$T = \frac{1}{N_v}\sum_{\Omega} v_i X_i \, , \qquad (4)$$

T - the coordinates of the detection object center
N_v - the sum of weights of all the pixels of the detection object
v_i - the weight of the i-th pixel of the detction object
X_i - the coordinates of the i-th pixel of the detection object (i = 1, ... N)
Ω - the set of all detection object pixels

3. The method of quadratic weighted average, which is weighted too but the weight function is not linear but quadratic function of dependence from detection object pixels colors.

$$n_i \;=\; 2*P+1-\frac{2}{P}(255-F_i)^2 \qquad (5)$$

$$n_i \;=\; 2*P+1-\frac{2}{P}\left(255-F_i-\frac{P}{4}\right)^2, \qquad (6)$$

where
v_i - weight of i-th pixel of the detection object
P - threshold
F_i- color of i-th pixel of the detection object

The tests showed that the best results were obtained with the second method. In some cases the first (less often the third) method was more accurate. The accuracy of the methods was dependent on the shape, size and color of the detection objects.

To be able to determine relative position of the copys it is needed to know two detection objects at least. It means that in every copy I must be able to find the same detection objects identical to the objects in the scene. I suppose, that such pair of detection objects exists in every record. Then by the selection of the most suitable method I come out from following assumption: The smaller is the spread of the distances of identical pairs of detection objects centers (calculated by certain method) in the records the more accurate is that method. The spread means the size of the scale of distances of given pairs if centers calculed for all the records.

$$R=\max_{\forall i,j}\left\{\left|\,\mathrm{Vzdial}_i-\mathrm{Vzdial}_j\right|\right\} \qquad (7)$$

R – the scale of the distances of given pair of detection objects
$Vzdial_i$ – distance of given pair of detection objects in the i-th record
index i (j) represents i-th (j-th) record (if there are N records then i = 1, ..., N)
$\forall$i,j means that with indexes i,j go through all the copys

If I have more than two detection objects I perform described procedure for every pair. Then the best pair of the detection objects centers is the one, which scale if the distances was minimal. These centers I label by symbols T_1^i, T_2^i, where i is the index of the copy.

Transformation of records to equalize the detection objects coordinates

At first the image with doubled resolution is created in the way that every pixel in first record is divided into four pixels containing the same color as the pixel from the first record. I.e. by the resampling of the first record I get the base for image with enhanced resolution.

When I have the centers of the detection objects pair for each record I can move the records on the first record that the corresponding centers are identical. $T_1^1 = T_1^i$, $T_2^1 = T_2^i$ for each i.

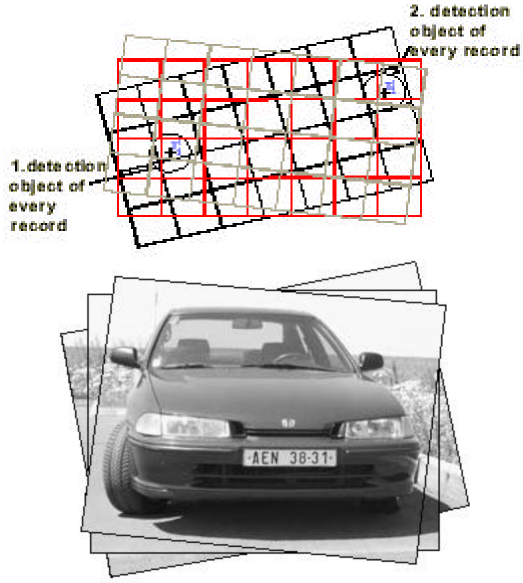


***Figure 4** The translation and rotation of the records on the first record*

To compute the pixels in the enhanced image it is needed to know the equations of the lines dividing two neighbouring rows (X-lines) or columns (Y-lines) in the records. It is necessary to determine these equations in such position of the record in which the detection objects T_1^i, T_2^i center coordinates are equal to detection objects centers coordinates T_1^1, T_2^1 of the first snapped copy. The example for one of the copys (red pixel grid) which is translated and rotated to validate $T_1^1 = T_1^i$, $T_2^1 = T_2^i$ is illustrated in the Figure 4. The gray pixel grid represents the enhanced image.All the records are translated in the way that the T_1^i coordinates for every record and T_1^1 coordinates for the first record have equal integer parts. Eg. if T_1^1 = (23.4,13.5) and in the second record is T_1^2 = (38.1,24.3), then the second record pixels are translated about 15=38-23 in *x*-direction and about 11=24-13 in *y*-direction.

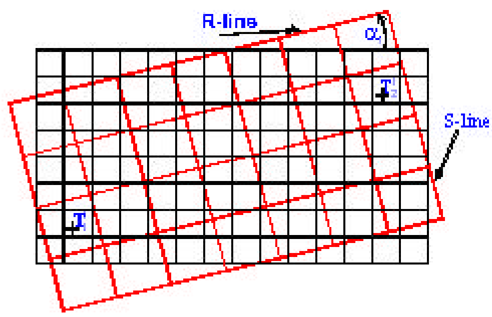


***Figure 5** The translation and rotation of the record to the resulting image*

Then the angle α_i of i-th record about which it is needed to rotate the record to validate $T_1^1 = T_1^i$, $T_2^1 = T_2^i$ is determined; and translate the records that $T_1^1 = T_1^i$. Finally the general equations of R and S-lines are determined and the pixels color in resulting image is calculated.

The computation of pixel intensity

Determination of general equations of X and Y-lines

I label the new x and y-coordinates of detection objects centers T_1^i, T_2^i by the symbols $T_1^i\langle x\rangle, T_1^i\langle y\rangle, T_2^i\langle x\rangle, T_2^i\langle y\rangle$. With known positions of every record detection objects centers it is possible to determine the position of its X and Y-lines for each record.

As it was mentioned before in the first step of the enhancement the new image with double resolution is created using the first record pixels. Then the line given by detection objects centers T_1^i, T_2^i in the i-th record will have the slope $k_i := \left(T_2^i\langle y\rangle - T_1^i\langle y\rangle\right) / \left(T_2^i\langle x\rangle - T_1^i\langle x\rangle\right)$ according to the selected coordinate system. Then for angle δ_i which includes this line with positive direction of x-axis is valid $d_i := arctan(k_i)$.

Let $\delta := \delta_1$ be angle which includes the line given by centers of detection objects in the first record with the positive directions of x-axis. Because the resulting image basis was created by resampling of the first record, R and S-liens in the first record have consistent direction with R and S-line of the resulting image. For angle which include R-lines of the record with positive direction of x-axis in the i-th record is valid $\alpha_i := \delta_i - \delta$. The same size has the angle which include the S-lines with positive direction of y-axis because R and S-lines are perpendicular to each other. Also it is valid that $\alpha_1 = \delta_1 - \delta = \delta_1 - \delta_1 = 0$.

The α_i angles of R and S-lines are known. To find their general equations for ever R and S-line one its point is selected. Let be the selected point from R line or S line respectively.

In this method the R-line points are calculated in the way that x-coordinate is determined and y-coordinate is calculated. For S-lines it is reversaly. The procedure is as follows:

The records are translated in a way that the centers T_1^i coordinates are equal to T_1^1 coordinates in integer parts. Then for coordinates By_j^i, and Bx_j^i of R and S-lines in the i-th record is valid

$$Bx_j^i = (x_{T_1^1}, y_{Bx_1^i}),\ y_{Bx_1^i} = (y_{T_1^1} + l_i^y) + (j-i) * s_i \quad (8)$$

$$By_j^i = (x_{By_1^i}, y_{T_1^1}),\ x_{By_1^i} = (x_{T_1^1} + l_i^x) + (j-i) * s_i \quad (9)$$

where

$x_{T_1^1}$, $y_{T_1^1}$ - are coordinates of the first center in the first record

$$l_i^y = \frac{r^y}{\cos a_i},\ l_i^x = \frac{r^x}{\cos a_i}, \quad \text{where} \qquad (10)$$

$$r^y = \left\langle y_{T_1^i} \right\rangle - y_{T_1^i},\ r^x = \left\langle x_{T_1^i} \right\rangle - x_{T_1^i},$$

while $\langle\ \rangle$ is the label of integer part

$$s_i = \frac{t}{\cos a_i}, \text{ where} \qquad (11)$$

t is the size of record pixel side

j - order number of R resp. S-line, j = 0, ..., number of R resp. S-lines

The calculation of y-coordinatesof R-lines is illustrated in FigureX.

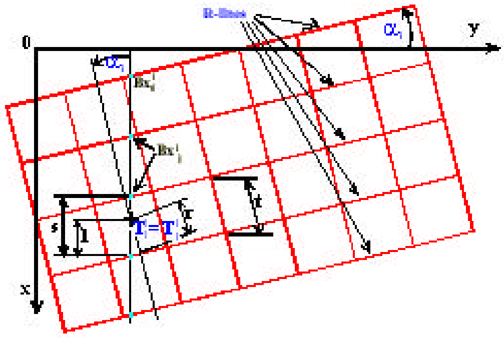


Figure 6 *The calculation of y-coordinates of R-lines of the record*

Because the angles α_i and points Bx_j^i, By_j^i of R and S-lines of the record are known it is possible to express their general equations

For R-lines:

$$tg\, a_i * x - y - tg\, a_i * x_{T_1^1} + y_{Bx_1^1} = 0 \qquad (12)$$

for S-lines:

$$x + tg\, a_i * y - x_{Bx_1^1} - tg\, a_i * y_{T_1^1} = 0 \qquad (13)$$

Now it is possible to determine the intersection of these lines.

The calculation of R and S-lines of the records and resulting image

R and S-lines in resulting image are parallel to coresponding coordinates axis. Now their intersections with R and S-lines of the records are determined.

When y resp. x-coordinates are set into equations of R resp. S-lines of the records the intersections can be calculated using the system of R and S-lines general equations. Finally the intersections of R and S-lines in the resulting image are computed. Because these lines are parallel to choosed coordinate system axis their intersections have the coordinate values equal to values in general equations of these lines.

The analysis of record pixels positions with respect to resulting image

The enhancement fulfilling the following is considered:

$$R \le S\frac{\sqrt{2}}{2}, \tag{14}$$

where
S - the length of record pixel side
R - the length of resulting image pixel side

This conditions reduces the number of possible types of relative positions of record pixels with recpect to the resulting image to four types. Even for doubling of the resolution these equation is valid.

The four types of actual pixel cover are possible:
1st type: Actual pixel is covered by part of one record pixel.
2nd type: Actual pixel is covered by two parts of neighbouring record pixels.
3rd type: Actual pixel is covered by three parts of neighbousring record pixels.
4th type: Actual pixel is covered by four parts of neighbousring record pixels.

The examples of every type are illustrated in the Figure 7.

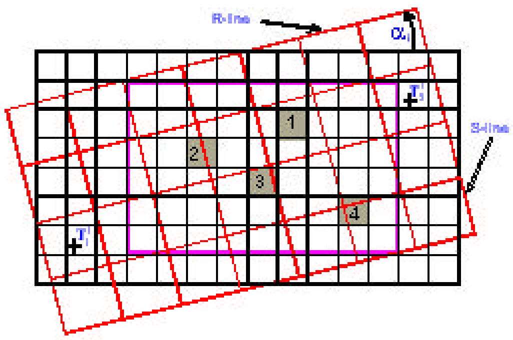


Figure 7 *Four possible types of resulting image pixel covering by record pixels*

The actual pixel type is determined by enumerating the number of record pixels overlapping it.
It is done by investigation of relative positions of all pairs of neigbouring R and S-lines with respect to the actual pixel. It is not necessary to investigate all the pairs. With suitable order of investigated pixels in resulting image it is possible to reduce this number.

The procedure in this method is as follows: The actual pixel vertices positions with respect to record R and S-lines are determined. From this information it is detected whether actual pixel is bounded by one, two or four pixels of the record. In case of one or two bounding pixels these pixels already overlap the actual pixel. However if this pixel is bounded by four record pixels it doesn‘t mean that all four pixels overlap it. There can occur the case with only three pixels overlaping it. The determination of these positions is done by determining the intersection of bounding record pixels position. If this intersection is in interior of investigated pixel the it is overlapped by four record pixels else it is overlaped by three.

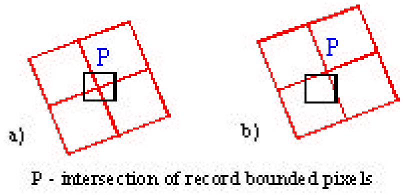


Figure 8 *Two possibilities of actual pixel overlap in case of four bounding pixels*

The calculation of pixels parts overlaping the actual pixel

Because It is already known how many pixels overlap actual pixel it is possible to determine the areas of parts of all overlapping pixels.

Because the areas are polygons (with at most 6 vertices) at first the coordinates of polygons vertices are calculated which are intersections of R and S-lines of the record and resulting image. The polygons are divided into triangles or trapuziums respectively. By enumeration of its areas and by addition the areas of the overlapping pixels parts are determined.

For acceleration of the method and reduction of number of arithmetic operations first the pixles parts are sorted according to their vertices number. Then the areas of all pixles parts are calculated except the last one with greates number of vertices. The area of that pixel part is computed by subdividing all the previously computed parts areas from area of the whole pixel. Example is illustrated in Figure 9. S_1, S_2, S_3 are record pixel parts overlapping the actual pixel, S_1 is the are of triangle ABC, S_2 is the area of the triangle DEF, S_3 is calculed by $S_3 = R^2 - S_1 - S_2$.

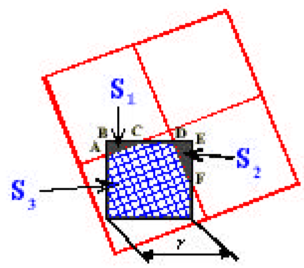


Figure 9 *The computation of the pixel part S_3*

The determination of actual pixel color

On a basis of known areas of the overlapping pixels parts of the records it is possible to determine the amount of contributions of record pixels to the resulting color of actual pixel.

The color which contributes to the resulting color of actual pixel willl be the weighted average of color of record pixels overlapping the actual pixel. The weights are areas of pixle parts covering the actual pixel.

The final color of the actual pixel is calculated so that the contributions of all suitable records are averaged. The suitable records are described in next section.

Up to now only double resolution was considered. Also non-integer enhancement can be achieved. You have to change the initial resampling to get the basis for enhanced image the further procedure is analogis to described one.

The further improvements of the method

The records used in this method have arbitrary relative positions. So there can occur the case of irregular distributed records which degrades the resulting image quality. The reason for this is the use of arithmetic average by the determination of actual pixel color which cases that by *irregular distribution of the records* the unbalanced image is created.

Under the term irregular distribution of the records we mean irregular distribution of the first detection object center coordinates of all records after the translation which consolides the integer parts of first detection objects. So of the records only that uniformly distributed in the plane are selected. The resulting image is determined only by such selected records labeled *suitable records*.

The next undesirable efect of record averaging is the smoothing if edges and depression of resulting image objects contrast cased by bad linear properties and noise of the light-sensitive elements of the CCD camera.

In effort of suppresing these defects there is pplied to resulting image folowing method which adds the information from first record to the result. This method is called *image stabilization*.

The method of image stabilization:

This method assumes that resulting image after reduction to records resulution must contain pixles with same color as on the records. The position of resulting image in choosed coordinate system is identical to first record position so the stabilization will use first record as reference. The pixel of first record covers exactly n^2 pixels of resulting image (n is a coeficient of integer resolution enhancement). Then must be valid that color computed by averaging of all resulting pixels covered by first record pixel equals to that first record pixel color. If resulting image doesn‘t fulfill these feature for some group of pixels the color of that group is modified to satisfy it. The color of pixels from this group are modified equally. If the color of one pixel from group is changed about one level, the further modification is possible after change of color of all pixels from that group.

The tests showed that the stabilization reduced the effect of smoothing and reduction of contrast in resulting image. In some cases another effect of partial undersampling occured and grouping of pixles was observed in the resulting image. This undesitable effect was most visible on places where average color of pixel group was most different from corrensponding pixel color in first record. So the application implements this method wit possibility of user definable maximum level of color modification. The artifitially constructed records and results obtained by random shift method are in the Figure 10. As it is visible the stabilized image after thresholding differs from scene only by one missing line with half-pixle width.

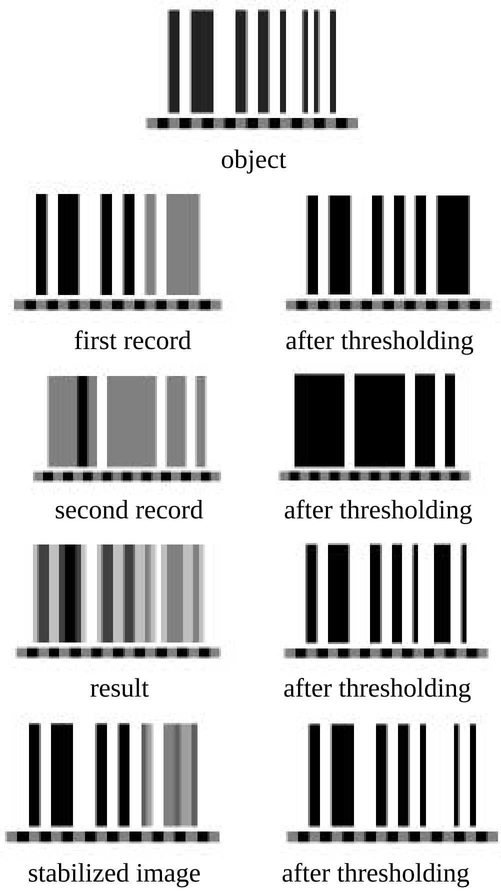


Figure 10 *Artificially constructed images and results*

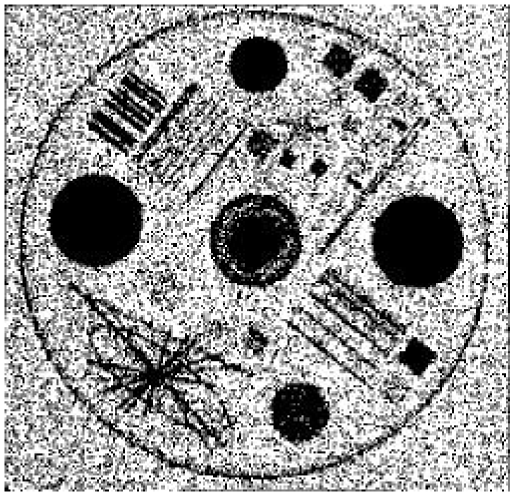
Conclusion

The paper described one of the methods for enhancement of viewed scene image. This method needed some additional modifications like image stabilization to get as best as possible results. During the tests images with 8-tuple resolution were obtained with quite nice results. This method can be further modified. Also as said before the result depends on the relative positions of records. One can construct the record series where the scene consists of regular black and white squares with square side size of one half of the record pixel. Whith composed such scene only images with one gray color can be recorded. From that information no improvement is possible when considering only translation of the record by the image snap. But if the records are snapped with rotated scene then it has more colors and the enhanced image can be computed. In the future work I want to investigate the best arrangement of the scene during the record to get even better results. Also I want to try to use other image capture devices. The length of the camcorder record can exceed hours and from these records viewing the same scene the image highly exceeding the PAL resolution can be reconstructed.

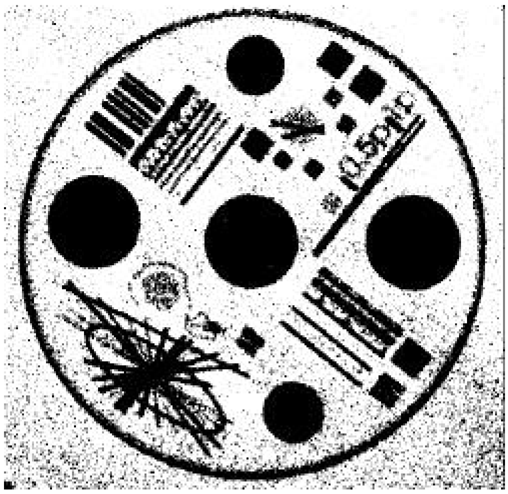
one of the seven records

result (without stabilization)

Figure 11 *The practical results*



one of the ten records



result (without stabilization)

Figure 12 *The practical results*

References

[Hlav98] Hlavac V., Sonka M., *Pocitacove videni*, Grada, 1992

[Ruzi95] Ruzicky E., Ferko A., *Pocitacova grafika a spracovanie obrazu*, Sapienta, Bratislava, 1995

[Zara98] Zára J., Benes B., Fekel P., *Moderni pocitacova grafika.* Oblasti, 1998