IRSTI 50.05.13

**A.M. Ashiraliyev[1]** – main author
**[2]Guo Ling (**orcid 0000-0001-8227-3528**)**

*[1]Master, [2]PhD, Associate Professor*
*Nanjing University of Science and Technology, Nanjing, China*
*E-mail: [1]aiadil12@mail.ru*

# DIGITALIMAGE CORRECTION BASED ON EDGE BAND AND HOUGH TRANSFORM

**Abstract.** Taking photo of a musical score usually results in a skewed image due to incorrect placement of the score, and the angle of the score effects on the accuracy of data recognition. The Hough transform method was used for data recognition, in order to improve the Hough transform method, fitting interpolation algorithms were added, which speeds up the process of determining the angle and ensures the accuracy of determining image correction based on the edge band.

**Keywords**:musical score, image processing, Hough transform, image correction, image recognition.

**Introduction.** The object of the research is images of musical scores obtained using mobile terminals such as mobile phones or tablets. Image quality has an important impact on the complexity of algorithm design, as well as on the efficiency and accuracy of recognition. Due to the shooting angle, score images often have the following problems:

1. The image is tilted due to the fact that between the angles of the coordinate axis of the

image tilt angle of the photographing device and the direction of the notch line are not equal to 0 when the image is acquired. Thetilt deviation is as shown in Figure 1.

2. Deformation of the image shift caused by the fact that the image plane of the shooting device is not parallel to the plane of the sheet, as shown in Figure 2.


Figure 1. Image with a curved angle

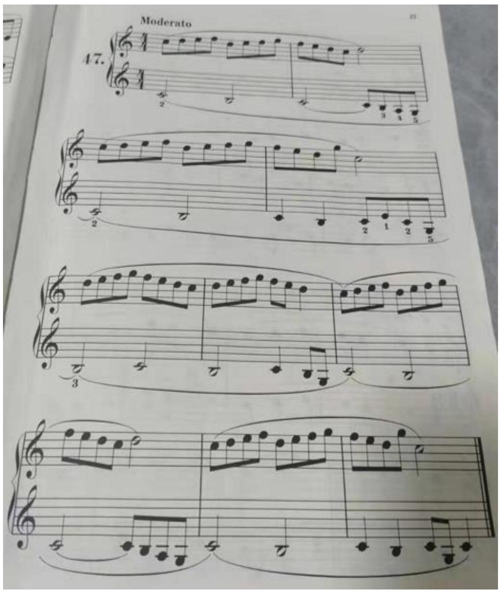


Figure 2. Deformation caused by shear

When shooting a musical score, it's very difficult to keep the imaging plane of the shooting device completely parallel to it. And a big tilt angle will severely affect theaccuracy in subsequent recognitionof musical notes. Therefore, it was necessary to correct the slope of the image to ensure the relative level of spectral lines and facilitate subsequent positioning and line removal operations.

Spectral lines have obvious linearity characteristics, and the most commonly used feature detection method is the Hough transform, which was first proposed by Paul Hough and later refined by Richard Doda and Peter Hart in 1972, extending detectable objects from straight lines to the detection of complex shapes, especially circles and ellipses[1]. The principle of the Hough transform is to detect objects of a certain shape using statistical voting algorithms. The process consists of calculating the local maximum value of the cumulative result in the parameter space for a set corresponding to a certain shape as a result of the Hough transform. Since the description parameters are calculated based on statistical global information, which has good fault tolerance and reliability in situations where the boundary of the area is disturbed by noise or hidden by other targets, which leads to some border violations.

The Hough transform maps a point in the image space to a sinusoidal curve in the parameter space, and the mapping curves of several collinear points intersect at a point in the parameter space (θ, γ). The transformation equation has the following form：

$$\gamma = x\cos\theta + y\sin\theta \quad \theta \in [0, \pi) \tag{1}$$

Where (x, y) are the coordinates of the spatial points of the image; (θ,γ) are the coordinates of the spatial points of the parameter; γ is the distance from the origin to the straight line, which is the polar diameter; θ is the minimum angle between the normal direction of the straight line and the right direction of the x axis, which is the polar angle.

Many scientists use Hough transforms to detect spectral lines based on their characteristics, and the image is corrected by the slope of straight lines. The Hough transform algorithm has a good effect of detecting discontinuous straight lines, but it takes a lot of time and requires the transformation of each foreground point in the parameter space. Due to the disadvantages of detecting spectral lines in using the Hough transform mentioned above, the following improvements have been developed：

1. Narrowing the range of detection angles. When the image is actually captured, the tilt angle

basically varies in a small range. There is no need to detect all straight lines within [π, 0).

2. In addition to the many parallel lines in the image of the score, there are many musical notes.

If the image of the score is directly transformed by the Hough transformation, a lot of time will be spent on calculations to count useless points, therefore, before performing the Hough transformation, it is necessary to pre-select possible line points, thereby eliminating the note points and preserving the line points.

Edges are the most important feature of an image, which are often present between objects and the background, as well as between the objects themselves[2]. Edge models are classified according to their grayscale profiles and are mainly divided into step models, slope models and strip edge models. Spectral lines can be considered as a model of the band edges. In light of the above-mentioned improvement method, the boundary band algorithm was first used to pre-select spectral line points in the image of a musical score, and then the Hough transform was used to detect pre-selected points on a straight line, and the results were adjusted using quadratic interpolation. The algorithm process is as follows：

1. Setting the width W of the strip edge, W=5；

2. For each foreground point (x, y) is denoted as the coordinates of the foreground point, I(x, y)

is the value of the current gray point, and the gray difference between the current point and the adjacent position is calculated, $\Delta I1 = I(x, y) - I(x, y - W)$ $\Delta I2 = I(x, y) - I(x, y + W)$, if the modulus of the difference is greater than 100, we consider it as a point of the spectral line and set it as 0.

After the edge band selection algorithm, most of the points obtained are points of spectral lines. Although there are a small number of points that slightly affect the results and accuracy of detection, as shown in Figure 3.

3. The resulting image is convolved with a Gaussian kernel of size 5 * 5 and a variance of 0.5, which eliminates the clustering process and reduces the calculation time.

4. The coordinates of the local maxima in space are recorded, and then the polar angle and polar diameter are interpolated and set respectively. Taking the interpolation fit of the polar diameter as an example, the mathematical model of the interpolation fit is a square curve, see Formula 2. Information about the neighborhood of the extreme coordinates is used to interpolate the extreme diameter. Formulas 3, 4 and 5 are used for fitting and interpolation.

Figure 3. The result of the edge band algorithm

$$\gamma = a * (x - c)^2 + b \tag{2}$$

where a, b, c are the adjusted parameters, x is the abscissa corresponding to the local maximum value of the space, i.e. required polar diameter, and γ is the value of the function curve.

$$a = \frac{(\gamma_{x+1} + \gamma_{x-1} - 2 * \gamma_x)}{2} \tag{3}$$

$$c = \frac{(\gamma_{x-1} - \gamma_{x+1})}{2 * a} - 0.5 \tag{4}$$

$$b = \gamma_x - a * c * c \tag{5}$$

where $\gamma_x$ is the statistical cumulative value corresponding to x, $\gamma_{x+1}$ is the statistical cumulative value corresponding to the right neighborhood of x, $\gamma_{x-1}$ is the statistical cumulative value corresponding to the left neighborhood of x, and c is the value of the final fit result.

After the improvements mentioned above, the detection time and detection accuracy have been greatly improved. To test the effectiveness of the algorithm, the image is rotated at a certain angle, which is used as the initial value, and compares it with the angle calculated using the above-mentioned improved method. The results are shown in table 1, which demonstrates the detection error is mainly within 0.007 degrees.

Table 1

Accuracy of spectral line detection by improved algorithm

| Initial value | -5.5 | -3.5 | -1.5 | -0.5 | 0.5 | 1.5 | 3.5 | 5.5 |
|---|---|---|---|---|---|---|---|---|
| Detected value | -5.50114 | -3.50005 | -1.50216 | -0.50090 | 0.49979 | 1.50049 | 3.49956 | 5.49222 |
| Error rate | 0.00114 | 0.00005 | 0.00216 | 0.00090 | 0.00021 | 0.00049 | 0.00044 | 0.00778 |

After the determination of the angle of inclination of the spectral line was completed, all images had to be rotated in the opposite direction by the appropriate angle in order to restore it to its horizontal position. Each pixel of the image is treated as a vector. According to the properties of the rotation matrix, the operation is performed on the vector. The length of the vector before rotation is equal to the length after rotation. In accordance with this property, the transformation ratio between the coordinates of the points before and after rotation was obtained.

The transformation relationship is as follows：

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \tag{6}$$

where (x, y) is the initial pixel coordinate, (x', y') is the corrected coordinate, and θ is the rotation angle of the image.

Formula 6 is a transformation formula obtained by taking the position of the center of the image as the center of rotation. Under normal circumstances, the origin of the image coordinates is in the upper-left corner of the image, so the coordinates of the image had to be converted to coordinates with the center of the image as the origin. It is assumed that the width of the original image is w, and the height is h, $(x_0, y_0)$ is a point in the original coordinate, and the point after the transformed coordinate is $(x_1, y_1)$. The conversion ratio is as follows：

$$x_1 = x_0 - w/2$$
$$y_1 = y_0 - h/2 \tag{7}$$

The length and width of the rotated image changes, so it’s necessary to calculate the length and width of the new image as follows：

$$w_{new} = \sin\theta * w + \cos\theta * h$$
$$h_{new} = \cos\theta * w + \sin\theta * h \tag{8}$$

Where w and h are the width and height of the original image, and $w_{new}$ and $h_{new}$ are the width and height of the new image after the rotation. The value represents the rotation angle of the image.

There is the display relationship between the coordinates in the new and original images:

$$\begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & w_{new}/2 \\ 0 & 1 & h_{new}/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} con\theta & -\sin\theta & 0 \\ \sin\theta & con\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -w/2 \\ 0 & 1 & -h/2 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ y_0 \\ 1 \end{pmatrix} \quad (9)$$

Where $(x_0, y_0)$ is the coordinate of the point of the original image, $(x_2, y_2)$ is the coordinate of $(x_0, y_0)$ in the new coordinate system of the image after the transformation of linear rotation, θ is the angle of rotation of the image. w, h are the width and height of the original image, respectively, $w_{new}$, $h_{new}$ are the width and height of the new image after rotation.

After the image slope has been calculated, the image slope correction is performed using formula 9. Figure 4 represents the image after correction. The calculation of the rotation angle θ is obtained using an improved algorithm, where θ = 3.49956.


Figure 4. Calibration results

**Conclusion.** The research mainly considered the problems of tilt in the image of the score and its correction. Taking into account the situation when the image slope cannot be processed by subsequent algorithms, some disadvantages of the basic Hough transform have been eliminated, as well as fitting interpolation and edge band algorithms have been added, which significantly speeds up the process of determining the angle and ensures the accuracy of the definition. During the research 326 pictures of musical scores where used and only 4 pictures were taken for introducing the results.

**References**

1. Duda R O, Hart P E. Use of the Hough transformation to detect lines and curves in pictures[R]. Sri International Menlo Park Ca Artificial Intelligence Center, 1971.
2. Mo Deju, Liang Guanghua. Digital Image Processing [M]. Beijing University of Posts and Telecommunications Press, 2010.

*Material received 17.02.22.*

**А.М. Әшіралиев, Гуо Лин**

*Нанкин ғылым және технология университеті, Нанкин қ., Қытай*

## ЖИЕК ЖОЛАҒЫ ЖӘНЕ ХАФ ТҮРЛЕНДІРУ НЕГІЗІНДЕ ЦИФРЛЫҚ КЕСКІНДІ ТҮЗЕТУ

**Аңдатпа.** Музыкалық партитураны түсіру әдетте партитураның дұрыс орналаспауына байланысты еңкейтілген кескінге әкеледі, музыкалық партитураның бұрышы деректердің тану дәлдігіне әсер етеді. Деректерді тану үшін Хаф түрлендіру әдісі қолданылды, Хаф әдісін жақсарту үшін фитингтік интерполяция алгоритмдері қосылды, бұл бұрышты анықтау процесін жылдамдатады және жиек жолағы негізінде кескінді түзетуді анықтаудың дәлдігін қамтамасыз етеді.

**Тірек сөздер:** музыкалық партитура, кескінді өңдеу, Хаф түрлендіруі, кескінді түзету, кескінді тану.

***А.М. Аширалиев, Гуо Линь***

*Нанкинский университет науки и технологий, г. Нанкин, Китай*

## ЦИФРОВАЯ КОРРЕКЦИЯ ИЗОБРАЖЕНИЯ НА ОСНОВЕ КРАЕВОЙ ПОЛОСЫ И ПРЕОБРАЗОВАНИЯ ХАФА

**Аннотация.** Съемка партитуры обычно приводит к наклону изображения из-за неправильного размещения партитуры, угол наклона партитуры влияет на точность распознавания данных. Для распознавания данных использовался метод преобразования Хафа, с целю улучшения метода Хафа добавлены подгоночные алгоритмы интерполяции, что ускоряет процесс определения угла и обеспечивает точность определения коррекция изображения на основе краевой полосы.

**Ключевые слова:** музыкальная партитура, обработка изображений, преобразование Хафа, коррекция изображения, распознавание изображений.