Study on the Actions Matching Characteristics of Music Rhythm

Zhang Lin

Tengzhou secondary vocational education center school, Tengzhou Shandong China
37088914@qq.com

Keywords: music rhythm; action; matching; feature fusion

Abstract: In order to improve the real-time detection and dynamic feature matching recognition ability of movements in music rhythm playing, a method of motion precision point feature matching in music rhythm playing based on high and low tone envelope profile feature detection is proposed. The sound division quasi-partition feature matching method is used to detect the action image acquisition and block fusion in music rhythm playing, the dynamic feature fusion technology is used to recognize the quantitative feature marking of action mark points in music rhythm playing, and the edge outline eigenvalues of action in music rhythm playing are extracted, and the dynamic detection of action mark points in music rhythm singing is realized by combining tone fuzzy block fusion technology. The edge similarity transformation is used to segment the movement mark points of music rhythm singing, and the dynamic feature matching of music rhythm action mark points is realized according to the segmentation results. The simulation results show that the method has high accuracy and good image fusion, which improves the marking and recognition ability of the movement feature points in music rhythm singing.

1. Introduction

In the computer vision environment, the image processing technology is used to monitor the music rhythm singing video, combined with the feature detection results of the music rhythm singing video, the music rhythm singing action guidance is carried out, the automatic planning and action marking ability of the music rhythm singing are improved, and the music rhythm singing movement precision point feature matching is carried out in the music rhythm singing video. And follow up the action guidance and design of physical training according to the results, and the research on the matching methods of motion precision points in music rhythm playing has attracted a lot of attention[1].

Traditionally, the characteristic match of the feature of the action sound in the music rhythm performance is set up on the basis of information fusion and feature extraction of the image, the characteristic parameter of the motion image area in the music rhythm performance is obtained, the envelope detection of the motion image area in the music rhythm performance is realized, so that the action area in the music rhythm playing can be accurately positioned, the fixed point mark is realized, the self-adaptive performance of the action sound quasi-point characteristic matching in the music rhythm performance is not good, the calculation cost is large[2]. In this paper, a method for matching the feature of the motion sound in the music rhythm performance based on the high-bass envelope contour feature detection is presented. the method comprises the following steps of: carrying out motion image acquisition and blocking fusion detection in the music rhythm performance by adopting a sub-sound quasi-partition characteristic matching method, and carrying out quantitative characteristic labeling identification of the action mark point in the music rhythm performance by adopting the dynamic characteristic fusion technology, the edge contour characteristic value of the motion in the music rhythm performance is extracted, the dynamic detection of the music rhythm singing action mark point is realized by combining the tone fuzzy block fusion technology, the edge similarity transformation is adopted to perform the music rhythm singing action mark point characteristic segmentation, And the dynamic characteristic matching of
the music rhythm action mark point is realized according to the segmentation result. Finally, the
simulation experiment analysis is carried out, and the superiority of the method in improving the
matching ability of the feature matching point in the music rhythm performance is shown.

2. Music rhythm singing action image acquisition and division quasi-partition feature
matching

2.1. Image acquisition of music rhythm singing action

In order to realize the feature matching of action point in music rhythm playing, firstly, the
feature analysis model of action image in music rhythm playing is constructed, the sub-block fusion
of action image in two-dimensional music rhythm playing is detected by using split-tone
quasi-partition feature matching method, and the action feature point mark in music rhythm playing
is carried out by using sharpening template enhancement technology[3]. The two-dimensional pop
field distribution model of action feature points in music rhythm playing is shown in Fig. 1.

According to the two-dimensional pop field distribution structure of the action image in music
rhythm playing shown in Fig. 1, the pixel feature components of the action image in music rhythm
playing are obtained, and the template matching function \( f(g_i) \) of the action image in music
rhythm playing is constructed as follows:

\[
f(g_i) = c_i \sum \frac{\rho_i \rho_j}{\sqrt{\varrho_i} + \varepsilon} + \sum \frac{\rho_i}{\sqrt{\varrho_i} + \varepsilon}
\]  

(1)

In this way, the background difference component of the three-dimensional dynamic motion
pitch feature matching of the motion image in the music rhythm performance is obtained, and the feature decomposition is performed in the three-dimensional feature distribution area of the motion image in the music rhythm performance[4], and the color feature decomposition of the motion image in the music rhythm performance is carried out by RGB decomposition method, and the RGB decomposition formula is:

\[
f(G_i) = a_1 + a_2x + a_3y + a_4z + \sum gU(g'_i, p_i)
\]

\[
g(G_i) = b_1 + b_2x + b_3y + b_4z + \sum gU(g'_i, p_i)
\]

\[
h(G_i) = c_1 + c_2x + c_3y + c_4z + \sum hU(g'_i, p_i)
\]

(2)

Where \( \Phi(T_n) \) is given by the following formula:

\[
\Phi(T_n) = \gamma T_H + \theta T_0 + \omega T_0
\]  

(3)
Partition detection and information fusion processing of motion in two-dimensional music rhythm performance are carried out by means of sub-resolution quasi-partition feature matching method:

\[
\begin{pmatrix}
    x \\
y \\
z \\
o \\
o \\
o
\end{pmatrix} = \gamma \begin{pmatrix}
    H + r \cdot I \\
    P \\
    \gamma \Theta \omega
\end{pmatrix} \begin{pmatrix}
    a \\
b \\
c
\end{pmatrix}
\]

(4)

In the music rhythm playing, the action mark points of the action image are random. Combined with the pixel space fusion technology, the dynamic feature matching and recognition of the action mark points in the music rhythm playing are carried out[5].

2.2 Feature matching of sound division and quasi-partition of image

The motion image acquisition and block fusion detection in music rhythm playing are carried out by using the method of sub-tone quasi-partition feature matching[6]. The Gibbs prior energy function of action image in music rhythm playing is defined as shown in the following formula:

\[
\sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{C}} \sum_{k \in \mathcal{Y}} V(\mathbf{Y}, \mathbf{P}) = \sum_{i \in \mathcal{C}} \sum_{j \in \mathcal{C}} \left\{ -\mathbf{p}_i \cdot \mathbf{y}_j \right\}
\]

(5)

In the formula, \( \mathbf{p}_i = [\mathbf{p}^1, \mathbf{p}^2, \mathbf{p}^3, \mathbf{p}^4]^T \), which is called the fractal coefficient of motion precision point feature matching in the music rhythm playing. The multifractal method is used to smooth the image filtering of the music rhythm singing action, and the color gradient decomposition method is used in the neighborhood to fuse the action image area in the music rhythm playing, and the smoothing coefficient set is constructed[7]. The block size of the action in the music rhythm playing is described as follows:

\[
\lim_{P \to +\infty} \kappa^P_{\delta}(f)(x, y) = \max_{(s, t) \in \mathcal{B}(x, y)} f(s, t) = \delta^P_{\delta}(f)(x, y)
\]

(6)

\[
\lim_{P \to -\infty} \kappa^P_{\varepsilon}(f)(x, y) = \min_{(s, t) \in \mathcal{B}(x, y)} f(s, t) = \varepsilon^P_{\delta}(f)(x, y)
\]

(7)

Wherein, \( \delta \) and \( \varepsilon \) represent the prior feature coefficient and correlation coefficient of action image in music rhythm playing, respectively. In the local region Potts model, the segmentation and quasi-partition feature matching method is used to fuse the action image in music rhythm playing[8], and the information fusion output is obtained as follows:

\[
P(\mathbf{Y}_{w_1} \mid \mathbf{X}_{w_1}; \mathbf{X}_{w_1}, \mathbf{Y}_{w_1} \mid \mathbf{P}) = \frac{1}{Z(\mathbf{P})} \frac{p(\mathbf{Y}_{w_1} \mid \mathbf{X}_{w_1}; \mathbf{P})}{Z(\mathbf{P})}
\]

(8)

\[
Z(\mathbf{P}) = \sum_{(i, j) \in \mathcal{Y}} P(\mathbf{Y}_{w_1} \mid \mathbf{X}_{w_1}; \mathbf{Y}_{w_1} \mid \mathbf{P})
\]

Wherein, \( \mathbf{P} \) represents the template matching feature quantity of action precision point feature matching in music rhythm playing, smoothed and de-noised the action image in music rhythm playing, and obtains that the average pixel set of action image in music rhythm playing after denoising is expressed as follows:

\[
\mathbf{x}_T = \frac{1}{T} \sum_{i=1}^{T} \mathbf{x}_i
\]

(9)

Wherein, \( \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_T \) are the sub-block template matching set of music rhythm singing action image, and \( T \) is the texture width of action precision point feature matching. The statistical shape model of action image in music rhythm playing is established[9]. Combined with multifractal method, the edge pixel set of music rhythm singing action is obtained as follows:

\[
F = \frac{p(x, y) - p(x, y)^{1/2}}{\sum_{k=1}^{v(x)} k(x, y)}
\]

(10)

Wherein

\[
p(x, y) = \frac{k(x, y)}{v(x)} \quad v(x) = \frac{\sum_{k=1}^{v(x)} k(x, y)}{v(x)}
\]

(11)
According to the above analysis, the division and quasi-partition feature matching of music rhythm singing action is carried out, and the dynamic feature fusion technology is used to identify the quantitative feature marking of action mark points in music rhythm playing.

3. The feature matching optimization of music rhythm singing movement

3.1 Dynamic feature fusion

On the basis of the above mentioned motion image acquisition and block fusion detection in music rhythm playing, this paper designs an action precision point feature matching algorithm in music rhythm playing. In this paper, a motion precision point feature matching method based on high and low tone envelope outline feature detection is proposed to extract the color boundary information feature quantity of action image in music rhythm playing. The brightness characteristics of super pixels of action image in music rhythm playing are obtained as:

\[
P(Y) = \frac{\exp\left(-\beta \sum_{c \in C} V_c(Y)\right)}{\sum_{c} \exp\left(-\beta \sum_{c \in C} V_c(Y)\right)}
\]  

In the formula, \( \sum_{c \in C} V_c(Y) \) is the total number of pixels in the music rhythm singing action. \( C \) is the neighborhood group defined in the block matching. Finally, the super pixel pair neighborhood for the adaptive fusion segmentation of the action image in the music rhythm playing is obtained as follows:

\[
J = \sum_{k=1}^{N} \sum_{l=1}^{N} u_{k,l}^* d(x_k, y_l) + \beta \sum_{k=1}^{N} \sum_{l=1}^{N} u_{k,l}^* d(\overline{x_k}, \overline{y_l})
\]  

From this, the multi-scale fusion and segmentation model of action image in super-pixel music rhythm playing is obtained. The fuzzy feature components of action image in music rhythm playing are expressed as follows:

\[
w(d) = f(\|x - x_j\|) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{(x_i - x_j)^2}{2}\right)
\]  

The pixel fusion model of action mark point area in music rhythm playing is expressed as follows:

\[
\beta_i = \exp\left(-\frac{|x_i - x_j|}{2\sigma^2}\right) \frac{1}{\text{dist}(x_i, x_j)}
\]  

The RGB value matching method is used to fuse the action image in two-dimensional music rhythm playing, and the dynamic quantitative feature marking recognition of action in music rhythm playing is carried out[10].

3.2 Feature segmentation and dynamic feature matching of action marker points in music rhythm singing

The deep residual learning method is used to segment the feature of music rhythm singing action mark point, initialize the horizontal pixel set of action in music rhythm playing, and use the active outline lasso method to reconstruct the image with high resolution, and the reconstructed image is regarded as follows:

\[
P(\phi) = \int \frac{1}{2} (|\nabla \phi|^2 - 1)^{2} dx
\]  

Wherein, \( E^{LBF} \) is the local gray energy term of the action in the music rhythm playing, and
$E^{LGFE}$ is the local gradient energy term. The extraction results of the local dynamic feature quantity of the action image in the music rhythm playing are described as follows:

\[
E^{LGFE}(\phi, f_1, f_2) = \\
\lambda_1 \left[ \int K_o(x-y) \left[ I - f_1(x) \right]^2 H(\phi)dy \right] dx \\
+ \lambda_2 \left[ \int K_o(x-y) \left[ I - f_2(x) \right]^2 \left( 1 - H(\phi) \right) dy \right] dx
\] (17)

\[
E^{LGFE}(\phi, f_1^G, f_2^G) = \\
\lambda_1 \left[ \int K_o(x-y) \left[ I^G - f_1^G(x) \right]^2 H(\phi)dy \right] dx \\
+ \lambda_2 \left[ \int K_o(x-y) \left[ I^G - f_2^G(x) \right]^2 \left( 1 - H(\phi) \right) dy \right] dx
\] (18)

The feature detection method of high and low tone envelope outline is used to detect the action marking points of music rhythm singing dynamically. The human action is analyzed by two-dimensional manifolds, and the feature distribution function of the edge outline of the action mark points is obtained as:

\[
H_\varepsilon(z) = \frac{1}{2} \left[ 1 + \frac{2}{\pi} \arctan \left( \frac{z}{\varepsilon} \right) \right]
\] (19)

\[
\delta_\varepsilon(z) = \frac{1}{\pi} \left( \frac{\varepsilon}{\varepsilon^2 + z^2} \right), z \in R
\] (20)

The edge similarity transformation is used to divide the characteristics of the music rhythm singing action marker points, and the dynamic feature matching of the music rhythm action marker points is optimized to the $f_1, f_2, |\nabla f_1|, |\nabla f_2|$ respectively according to the segmentation results:

\[
\begin{align*}
    f_1(x) &= \frac{K_o(x) \ast H_\varepsilon(\phi(x)) I(x)}{K_o(x) \ast H_\varepsilon(\phi(x))} \\
    f_2(x) &= \frac{K_o(x) \ast \left[ 1 - H_\varepsilon(\phi(x)) \right] I(x)}{K_o(x) \ast \left[ 1 - H_\varepsilon(\phi(x)) \right]}
\end{align*}
\] (21)

\[
\begin{align*}
    f_1^G(x) &= \frac{K_o(x) \ast H_\varepsilon(\phi(x)) I^G(x)}{K_o(x) \ast H_\varepsilon(\phi(x))} \\
    f_2^G(x) &= \frac{K_o(x) \ast \left[ 1 - H_\varepsilon(\phi(x)) \right] I^G(x)}{K_o(x) \ast \left[ 1 - H_\varepsilon(\phi(x)) \right]}
\end{align*}
\] (22)

According to the regression analysis model, the nonlinear regression function is used to adjust adaptively, and the dynamic detection of music rhythm singing action mark points is realized by combining tone fuzzy blockage fusion technology, so as to improve the dynamic feature matching and feature resolution ability of music rhythm singing action mark points.

4. Simulation results

In order to verify the application performance of this method in the realization of motion point feature matching in music rhythm playing, the experimental analysis is carried out. The experiment is programmed on Matlab 2012 platform. The collection pixel level of action mark points in music rhythm playing is 2000, the number of dynamic feature points is set as 24, the edge similarity feature quantity $\lambda$ is set to 0.01, and the dynamic feature matching recognition is carried out by
124×120 pixel points. The frame acquisition number of the image is 230~400 frames. The collection results of action features in music rhythm playing are shown in Fig. 2.

Based on the results of action features in music rhythm playing in Fig. 2, the dynamic feature fusion technology is used to identify the quantitative feature tagging of action mark points in music rhythm playing, and the edge outline eigenvalues of action in music rhythm playing are extracted, as shown in Fig. 3.

According to the result of feature extraction, the dynamic feature matching of action mark points in music rhythm playing is realized, and the tracking results are shown in Fig. 4.

The analysis of Fig. 4 shows that the method can effectively realize the matching and recognition of the feature of the motion point in the music rhythm playing, and test the accuracy of the feature matching of the action point in the music rhythm performance by different methods. The comparative results are shown in Table 1. The analysis Table 1 shows that the accuracy of the method in the matching of the action point feature in the music rhythm playing is higher.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td>0.965</td>
<td>0.865</td>
<td>0.879</td>
</tr>
<tr>
<td>400</td>
<td>0.989</td>
<td>0.879</td>
<td>0.904</td>
</tr>
<tr>
<td>600</td>
<td>1</td>
<td>0.924</td>
<td>0.927</td>
</tr>
</tbody>
</table>

5. Conclusions

In the music rhythm singing video, the sound precision point feature matching of the music
rhythm singing action is carried out, and the action guidance and design of music training according to the results are followed. In this paper, a method of motion precision point feature matching in music rhythm playing based on the high and low tone envelope outline feature detection is proposed. The sound division quasi-partition feature matching method is used to detect the action image acquisition and block fusion in music rhythm playing, the dynamic feature fusion technology is used to recognize the quantitative feature marking of action mark points in music rhythm playing, and the edge outline eigenvalues of action in music rhythm playing are extracted, and the dynamic detection of action mark points in music rhythm singing is realized by combining tone fuzzy block fusion technology. The edge similarity transformation is used to segment the movement mark points of music rhythm singing, and the dynamic feature matching of music rhythm action mark points is realized according to the segmentation results. It is found that the method in this paper has good accuracy and dynamic feature matching ability in music rhythm playing.

References


