

Research on interactive 3D motion real-time image reconstruction technology

Baichuan-Cai , Rongfang-Mei², Wenjuan-Wang ,Yang bo , Zhang yi

(Yibin Vocational and Technical College, Yibin Sichuan,644003)

Keywords: the number of shape based; missing feature points; reconstruction of the characteristics of details; difference of curve endpoints

Abstract. The traditional 3D motion real-time image reconstruction technology, without considering the missing reconstruction feature points brought by wrong estimation for the shape based parameter in the algorithm, effects the authenticity and definition of 3D motion real-time image reconstruction. To this end, the interactive object 3D motion real-time image reconstruction method is put forward, the 3D action structure of missing feature recovered from dynamic image sequences is utilized, in the minimum feature space, through the characteristics of the key details to predict reconstruction, so as to determine the dynamic shape heart of motion model, using curve endpoint interpolation plane controlled 3D motion real-time image HC-B zier generation method to compensate for wrong estimation of shape base, and compensate the defects of 3D motion real-time image brought by wrong estimation of shape base, like features is missing, object reconstruction effect is poor, robustness is weak, and the definition is low.

Introduction

The 3D motion real-time image reconstruction problem based on object dynamic 2D image sequence, is a hot topic in the field of pattern recognition. The 3D motion information only belongs to specific objects is recovered from a series of image sequence [1, 2]. With the development of biometric identification technology, 3-D motion objects real-time image reconstruction technology is becoming more and more popular [3]. However, the object is vulnerable to be affected by the outside factors, background is complex, and noise is loud. The traditional 3D motion real-time image reconstruction based on two dimensional image sequence is assuming that the number of objects' shape base is known, these algorithms do not consider the lack of reconstruction feature points caused by wrong estimation of shape based parameter, therefore, it is difficult to use the traditional reconstruction algorithm to implement the method directly [4, 5]. At present in the related fields of computer vision and pattern recognition, how to recover object 3D motion real-time image information in change scenarios, from a set of 2D object dynamic image sequences, is still a difficult problem.

In order to solve this problem, the interactive object 3D motion real-time image reconstruction method is put forward, the 3D action structure of missing feature recovered from dynamic image sequences is utilized, in the minimum feature space, through the characteristics of the key details to predict reconstruction, so as to determine the dynamic shape heart of motion model, using curve endpoint interpolation plane controlled 3D motion real-time image HC-B zier generation method to compensate for wrong estimation of shape base, and compensate the defects of 3D motion real-time image brought by wrong estimation of shape base, like features is missing, object reconstruction effect is poor, robustness is weak, and the definition is low.

The reconstruction method of 3-D motion real-time image

The current mainstream 3-D motion real-time image reconstruction method based on two-dimensional dynamic image sequence: motor recovery method based on feature correspondence (Feature Based Motion Recovery), is on the basis of the corresponding relationship of the characteristics between frames to extract two-dimensional feature points information which can be used to complete the three-dimensional transformation, and then being transformed into 3D

information, and generally include the following steps:

The representative feature points in each image $\begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix} \begin{matrix} i=1,\dots,F \\ j=1,\dots,P \end{matrix}$ is extracted to compose the feature points calculation matrix $W_{2F \times P}$ (F is the number of image frame for operations, P is the number of effective feature points of a frame contains non rigid image), through the above basis can calculate the 3D motion characteristic matrix $\tilde{S}_{3 \times P}$ and the action characteristics rotation matrix $R_{i3 \times 3}$.

In general, three-dimensional shape features of most objects can be regarded as the weighted linear combination of shape based, so it can be: $\tilde{S}_i = \sum_{l=1}^K \omega_l S_l$. Among them, ω_l is the weighted coefficient, S_l is shape based matrix, K it the number of shape base, the number is different along with different values of non rigid body, most of the algorithms are based on estimation method.

Considering the camera position and other factors, it has the following formula:

$$\begin{bmatrix} u_{i1}, \dots, u_{iP} \\ v_{i1}, \dots, v_{iP} \end{bmatrix} = \bar{R}_i \left(\sum_{l=1}^K \omega_l S_l \right) + \bar{T}_i e_n^T \quad (1)$$

Among them, \bar{R}_i is the effective deformation of rotation matrix $R_{i3 \times 3}$, \bar{T}_i is the effective deformation of translation vector $T_{i3 \times 1}$, $e_n^T = [1, \dots, 1]_{1 \times n}$.

If the amount of computation is too large, then the following formula can be utilized to simplify the computation, thus reducing the amount of computation, improve the speed of computation.

$$\begin{cases} \bar{u}_{ij} = u_{ij} - \frac{1}{P} \sum_{j=1}^P u_{ij} \\ \bar{v}_{ij} = v_{ij} - \frac{1}{P} \sum_{j=1}^P v_{ij} \end{cases}$$

Formula (1) can be transformed into:

$$\begin{bmatrix} \bar{u}_{i1}, \dots, \bar{u}_{iP} \\ \bar{v}_{i1}, \dots, \bar{v}_{iP} \end{bmatrix} = \bar{R}_i \left(\sum_{l=1}^K \omega_l S_l \right) \quad (2)$$

Collecting all extracted valid 2D feature point information, and superposition operation is processed, there are:

$$\begin{bmatrix} \bar{u}_{11} & \dots & \bar{u}_{1P} \\ \bar{v}_{11} & \dots & \bar{v}_{1P} \\ \dots & \dots & \dots \\ \bar{u}_{F1} & \dots & \bar{u}_{FP} \\ \bar{v}_{F1} & \dots & \bar{v}_{FP} \end{bmatrix} = \begin{bmatrix} \omega_{11} \bar{R}_1 & \dots & \omega_{1K} \bar{R}_1 \\ \dots & \dots & \dots \\ \omega_{F1} \bar{R}_F & \dots & \omega_{FK} \bar{R}_F \end{bmatrix} \begin{bmatrix} S_1 \\ \dots \\ S_K \end{bmatrix} \quad (3)$$

It can be seen through the calculation above, the characteristics calculation matrix w of objects movement is rank type convergence. With some constraint conditions to process singular value decomposition for w can obtain both M and B , and the presence of a $3K \times 3K$ order nonsingular transformation matrix Q can help the factorization get correct results.

Predictive reconstruction algorithm of key details characteristics

Through the above analysis we can see that, the traditional 3D motion real-time image reconstruction technology based on two-dimensional image sequence have many drawbacks, a 3-D motion real-time image reconstruction algorithm based on interaction is proposed in this paper, according to the relationship between the key detail characteristics, combining with 3D action structure of incomplete features recovered in dynamic image sequences, \ to realize the reconstruction of 3D action real-time image.

Determination of dynamic centroid of the object model

In general, the initial center of construction model is the geometric midpoint of specified object. However, the center of the specified object geometric midpoint is accurate, only when the specified image is static. Therefore, the object is regarded as a 3D motion manifold with its centroid viewed

as the center, the position change of the centroid in space is utilized to describe the translational motion of object. The centroid coordinates can be determined by the following formula:

$$C_i = \begin{bmatrix} a_i \\ b_i \end{bmatrix} = \sum_{j=1}^P \lambda_{ij} \begin{bmatrix} u_{ij} \\ v_{ij} \end{bmatrix} \quad (4)$$

Among them, C is the sum of spatial position parameters weights of all key detail characteristics in i image, λ_{ij} is the noise presenting in the image and outlier disturbance weight parameters.

With the mean variance algorithm can obtain the value of the disturbance weight parameter. In order to make the procedure more simple and convenient, noise reduction is applied to the direction component deviation σ_{ijr}^2 of all key details characteristics. The sum of mean variance parameters Q_{ij} can be obtained by computation with the following formula

$$\sigma_{ijr}^2 = \sigma_{ij1}^2 + \sigma_{ij2}^2 \quad (5)$$

The value of λ_{ij} is able to obtain through operation:

$$\lambda_{ij} = \left(\frac{\sum_{l=1}^P \sum_{i=1}^F (\sigma_{ij1}^2 + \sigma_{ij2}^2)}{\sum_{l=1}^F \sum_{i=1}^P (\sigma_{il1}^2 + \sigma_{il2}^2)} \right)^{-1} \quad (6)$$

Through the above formula, it can be learnt that weight coefficient is related to direction component variance parameters σ_{ij1}^2 and σ_{ij2}^2 , for the cases of larger outlier and noise interference coefficient, the accuracy of treatment results is stronger.

Realization of 3D action real-time image reconstruction algorithm

This paper presents a three-dimensional motion real time image HC-Bézier generating method based on curve endpoint interpolation plane control to optimize the traditional method, the improved algorithm is described as follows:

The new curve has the basic properties of convex hull, symmetry can be combined with the method, the curve description formula is:

$$\begin{cases} b(0) = \frac{1}{2}(V_0 + V_1), & b'(0) = \frac{3+\lambda}{2}(V_1 - V_0), \\ b''(0) = [\frac{\pi^2}{4}(n-2) - 3 - 3\lambda](V_1 - V_0) \\ b(1) = \frac{1}{2}(V_{n-1} + V_n), & b'(1) = \frac{3+\lambda}{2}(V_n - V_{n-1}), \\ b''(1) = [3 - \frac{\pi^2}{4}(n-2) + 3\lambda](V_n - V_{n-1}) \end{cases} \quad (7)$$

The V_0 , v and V_n respectively represents the control point formed by the curve in the action image convex hull at control endpoints, control edge superposition interpolation is processed for 3D action plane curve endpoints to get the contour curve definition of i frame action image :

$$b_{n,i}(t) = \sum_{j=0}^{n_i} b_{n_i,j}(t; \lambda_i) \mathcal{V}_{j+\sum_{s=1}^{i-1} (n_s-1)}, \quad t \in [0,1] \quad (8)$$

Wherein, $i = 1, 2, \dots, l$, when $i = 1$, $\sum_{s=1}^{i-1} (n_s-1) = 0$

Therefore, each curve segment of the action combination curve construct a plane control basis function with integral recursive polynomial, the steps is described as:

(1) extract the edge contour curve of 3D motion image as edge processing prior knowledge, the correlated contour points initialization is processed in image area, reduce the input of subsequent contour feature iterative error and redundant feature points;

(2) by using the control points $V_{ij} \in \mathbf{R}^3$ ($i = 0, 1, \dots, l_1(m-1) + 1$; $j = 0, 1, \dots$, and the parameters define a combination surface λ composed of $l_1 \times l_2$, $m \times n$ order surface patches:

$$b_{ij}(u, v) = \sum_{k=0}^m \sum_{l=0}^n b_{mk}(u; \lambda_u) b_{nl}(v; \lambda_v) \mathcal{V}_{k+(m-1)(i-1)+l+(n-1)(j-1)} \quad (9)$$

(3) the gray region $M \times M$ of three-dimensional motion real-time image, is set to be a 3×3 model, using manual initialization to scan the region for point $P(x_0, y_0)$, arbitrary point $Q(x_i, y_i)$ in lasso model can be obtained, through the gray level difference comparison between two points, can obtain the enhanced regional maximum value of motion profile surface feature points:

$$pixel_A = \max \left(\sum_{i=1}^8 (Q - P) \right) \quad (10)$$

Each outline point which can produce the maximum grey value are recorded to avoid the interference of redundant noise point, the maximum gray level difference point corresponding to P value at this time is $K(x_0, y_0)$.

(4) similarly, with $K(x_0, y_0)$ as the center, the lasso region is extended to the 8 region model, the arbitrary point in its neighborhood is marked as $R(x_i, y_i)$, then the maximum gray level difference between the two pixels is:

$$pixel_B = \max(\sum_{i=1}^8 (R - K)) \quad (11)$$

(5) if $pixel_A < pixel_B$, update contour points constituted of 3-D motion surface, and recorded point P becomes K , otherwise not update, the above process is applied to every pixel, until all the pixels have been executed.

Experimental results and simulation

The facial action expression is used as an example, with the traditional method and proposed method to process 3D motion real time reconstruction for image sequence (as shown in Fig. 1). As can be seen by comparison, the traditional method of factorization have obvious characteristic points missing at the eyes, nose and lips and other parts, while the method in this paper can be very good to make up for this deficiency.

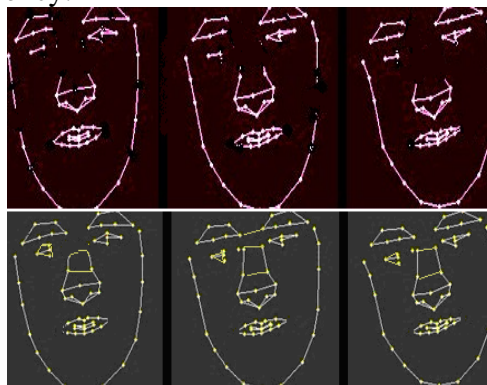


Figure 1 tendency chart of the definition using different algorithms

In order to analyze the performance quantitatively, aiming at the back projection error $\sigma = \frac{\|W - W_r\|_F}{\|W\|_F} \times 100\%$ to compare 3D motion real-time image reconstruction results, w is the original measurement matrix, w_r is back projection measurement matrix of reconstruction results. The error trend as shown in figure 2:

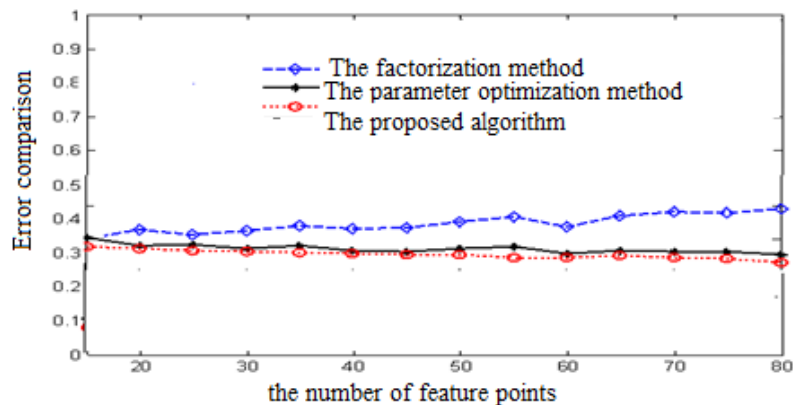


Figure 2 tendency chart of error comparison

The specific statistical results are shown as table 1:

Table 1 result statistics

methods	Traditional iteration method	The proposed algorithm
the back-projection error	0.23	0.16
The number of feature points	61	87

It is shown that the proposed method increase the robustness of the algorithm and accuracy of 3D dimension action recovery action, the reconstructed dynamic change of 3D action is more continuous.

Conclusions

The interactive object 3D motion real-time image reconstruction method is put forward in this paper, the 3D action structure of missing feature recovered from dynamic image sequences is utilized, in the minimum feature space, through the characteristics of the key details to predict reconstruction, so as to determine the dynamic shape heart of motion model, using curve endpoint interpolation plane controlled 3D motion real-time image HC-B zier generation method to compensate for wrong estimation of shape base, and compensate the defects of 3D motion real-time image brought by wrong estimation of shape base, like features is missing, object reconstruction effect is poor, robustness is weak, and the definition is low.

References

- [1] Tang Tao, Li Yi, Chen Tao, Li Zhiyong. A new method of extracting local affine invariant feature of the image [J]. Computer simulation. 2007.7.23-24
- [2] J. Yan and M. Pollefeys. A factorization-based approach for articulated non-rigid shape, motion and kinematic chain recovery from video. IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(5), May 2008.
- [3] M. Paladini, A. D. Bue, and M. Stosic. Factorization for non-rigid articulated structure using metric projections. In 2009 IEEE Conference on Computer Vision and Pattern Recognition, p 2898-2905, 2009
- [4] Ali Khamene, Shahriar Negahdaripour. Motion and structure from multiple cues: imagemotion, shading flow, and stereo disparity [J]. Computer Vision and Image Understanding, Volume 90, Issue 1, April 2003, 99-127.
- [5] Fadi Dornaika. Self-calibration of a stereo rig using monocular epipolar geometries[J]. Pattern Recognition, Volume 40, Issue 10, Oct. 2007, 2716-2729.