ISSN : 0974 - 7435 Volume 10 Issue 14

2014

BioTechnology

An Indian Journal

FULL PAPER

BTAIJ, 10(14), 2014 [8273-8280]

Human skeleton three-dimensional reconstruction-based sports features analysis

Xiaocheng Sun

Jilin Engineering Vocational College, Siping 136000, Jilin, (CHINA)

ABSTRACT

Sports video analysis is helpful for athlete techniques improvement, human body three-dimensional reconstruction on key motions can effective help coaches and athletes to make progress. The paper firstly states human body sports capturing and three-dimensional reconstruction research background and research status. And then, start from human skeleton three-dimensional reconstruction feature points extraction and matching algorithm, detail Harris Corner algorithm application in moving image feature point extraction, and make principle statements on human skeleton three-dimensional reconstruction's initialization process and basic matrix refining process. Finally, apply Harris Corner algorithm, eight-point algorithm, gold standard algorithm, Sampson correction algorithm and linear triangular algorithm to carry on feature point extracting 、matching and three-dimensional reconstruction to do experiments on men aerobics and men martial arts events one frame image, finally it gets relative satisfied experiment result.

KEYWORDS

Three-dimensional reconstruction; Sports video; Feature point extraction; Sampson correction; Linear triangular method.

INTRODUCTION

With sports high technological development, three-dimensional simulation technology of athletes moving image is increasing maturity, carries on sports features exploration on athletes in different time phases, which can let athlete to make targeted training, and is helpful for his sports technical improvements. On the basis of stating human body motion capturing and three-dimensional reconstruction research status, the paper analyzes current sports videos human skeleton three-dimensional reconstruction, in the hope of exploring more convenient and accurate algorithm, which provides theoretical basis for sports numeral analysis.

For human body moving image three-dimensional simulation researches, lots of people have made efforts, efforts results show that athletes training techniques and sports ability have a qualitative leap, from which Cheng Xuan and others (2013)proposed a kind of texture information-based three-dimensional human body sports recovering method, and provided a robust、adaptive tracking model, and experiment in HumanEva-Ⅱ test set showed that the algorithm could get better result[1]; Li Jie and others (2013)put forward a kind of improved factorization human body moving images three-dimensional dynamic simulation algorithm, applied optics fluid concept reasonable calculating on sports shape base numbers, carried on binding calculation on the quantity of shape base, so that implemented human body moving image sequence three-dimensional dynamic simulation modeling[2]; Zhang Dong-Hai (2013)utilized computer image's iconology method to do computer automatic identification on frame calibration, it implemented sports technical analysis's three-dimensional reconstruction frame's computer automatic identification[3].

The paper based on formers research, analyzes sports video image's human skeleton three-dimensional reconstruction technology, in the paper, it states Harris Corner algorithm, eight-point algorithm, gold standard algorithm, Sampson correction algorithm and linear triangular algorithm, and applies above algorithm to do human body three-dimensional sports simulation on aerobics and martial arts process representative motions, in the hope of providing theoretical references for sports targeted training and precision training.

HUMAN MOVEMENT CAPTURE AND THREE-DIMENSIONAL RECONSTRUCTION RESEARCH BACKGROUND AND STATUS OVERVIEW

Present sports video analysis provides scientific data basis for sports events technical development, and also provides relative scientific references for athletes' targeted training, while athletes or sports apparatus real-time detection and contour extraction are key techniques in video analysis, which are also core contents of sports analysis. The paper analyzes human body sports video image's feature points extraction and three-dimensional reconstruction technology, in the hope of providing more scientific platform for computer-assisted sports teaching, states the chapter research contents about human body sports tracking, in the hope of learning current the technique research status and directions that can improve, which provides basis for the paper provided sports initialization algorithm and feature points extraction algorithm scientific extraction.

Research background overview

With computer processing capacity and visual theories research level have obviously improved in current academic environment and technical environment, let human body movement tracking and analysis to move forward from theoretical research to practical, it has more broad development prospects in intelligent human-computer interaction、animation and games 、clinical medicine diagnosis、intelligent video monitoring、identity authentication、virtual reality and augmented reality、video analysis and sports as well as other aspects.

Davis LS. and others (2000)pointed out in intelligent monitoring fields, traditional monitoring system was used in venues of higher safety requirements, adopted several computer controlled video cameras to monitor, though most important occasions equipped with video cameras, it still needed special working staff to monitor display so that find abnormal events, or checked conditions after the events, it cannot effective stop some events that should not happen, and due to storage devices capacity was limited, it could not store all information for a long time, ideal intelligent video monitoring system was able to continuous unremitting make real-time monitoring, and make automatic real-time analysis of video data, detect moving human body, analyze his behaviors, if behaviors were abnormal or there were criminal behaviors happened, system could correct and timely raise the alarm, so that effective prevented emergency and crimes occurrence[4]. Similarly in abnormal intense sports competitive process, in order to make real-time narration on human body sports features, and make more real-time scientific statements for audiences, analysis of sports videos are also necessary.

In case that current computer technology, animation analog simulation and games software main carriers move forward towards physical sports, in order to more realistic let people to experience sports intensity and visual effects, realistic sports feature analysis plays crucial roles, applies intelligent monitoring system to make real interpretation on human movement features in videos, and then utilize decoding data to make human body three-dimensional reconstruction, it can more scientific and realistic reflect movement process, and provides firm basis for current computer technology and cultural development tide organic combination.

In competitive sports scientific training, human body three-dimensional movement plays crucial roles, the technique can analyze video sequence's human body movement trajectory of peoples interests, it makes technical supports for explaining human body movement behavior extremum 、group cooperation and getting human body movement model, such technical support can improve competition result and assist movement training.

In dance teaching and golf as well as tennis competitive sports events individuation training, accurate tracking of human movements and accurate extracting of human movement's movement link points decide technical training quality, therefore in above fields, human body three-dimensional movement analysis and reconstruction technology have greater applied prospects.

Li Hao-Jie (2006)pointed out video human body sports analysis and researches can provide more plentiful real-time information for sports competition rebroadcasting, as adding competition commentary 、 objective evaluation on competitions、reconstruction competition important fragments and others, so that enhance sports competition rebroadcasting effects, let audience to get more valuable information during watching television relay process, which is also important application values in business field[5].

Research status overview

In the 70s of the 20th century, psychologist Johansson firstly started working surrounding human body movement, in his laboratory, attached bright spots on research objects articulation points, and set laboratory into dark environment, then let sportsman to make each kind of motions in the dark. Psychologist Johansson in his work " Visual motion perception" stated experiment conclusion was that for static light spot set, human vision sensory perceptual system could not get any significant information, and for movements generated light sport set sequence, it was supposed to distinguish movement shape as walking, running and so on, and set and judged sportsman gender[6]. Also in view of this, psychologist Johansson put forward movement identification problem, in successive human movement analysis fields, people more focused on sports identification that promoted the technical development.

Ding Hai-Shu and others (1992)put forward that with computer technology, image processing technology, VLSI technology and high resolution sensor technology constantly updating and development, relative scientific researchers constantly joined, participated research and development, human movement tracking and analysis techniques got further perfection and development, by far, some system had already arrived at practical phase, as utilized landmarks that attached in human feature pints to do feature points tracking, such method had already been widely used, but landmarks that attached in human body surface sometimes might be covered or followed by human body surface deformation, mark point would generate malposition, subjects would somehow have feeling of discomfortableness, therefore label-free human movement capture attracted researchers' attentions[7].

In current human movement tracking, usually it adopts model method that is to project human body model and match to images, by virtue of human body model geometric structure and sports model prior theory, it carries on top-down capturing.

Top-down capturing predicts next frame movement model according to previous frame image results, and then on the basis of predicting image and real image comparison, it corrects prediction model, let it to arrive at more reasonable and scientific prediction model. Top-down movement capturing processing procedures is as Figure 1 shows, in Figure 1, symbols description is as TABLE 1 shows.

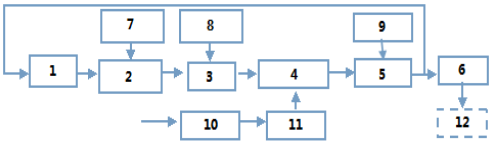

Figure 1: Top-down human movement capture processing procedures structure

TABLE1: Figure 1 each symbol description

Symbol	Description	Symbol	Description	Symbol	Description
1	Last frame or initial posture	5	Optimize prediction posture	9	Kinematics constraint
2	Predict current frame posture	6	Tracking result	10	Current frame image
3	Model projection	7	Movement model	11	Feature extraction
4	Match and get objective function	8	Human model	12	High level processing

Apply as Figure 1 showed top-down human movement capturing processing procedures, it has following three shortcomings:

It is hard to implement automatic initialization.

Error accumulative effect generated automatic recovering difficulties.

Processing procedures consumed long time and hard to implement real-time detection.

But by unremitting efforts from 1973 to now, label-free human movement capturing has already made larger progress, in the measuring technique aspect, it adopts multiple video cameras methods, in estimation technical aspects, it adopts model-based 3D tracking method, former can avoid movement process sheltering problem, the latter can accurate get 3D posture information, which is current human movement capture and three-dimensional reconstruction development orientations.

HUMAN SKELETON THREE-DIMENSIONAL RECONSTRUCTION'S FEATURE POINTS EXTRACTION AND MATCHING ALGORITHM ANALYSIS

In human movement capturing and human model 3D tracking process procedures, the principle is scientific getting two-dimensional video sequence, in two-dimensional video images, it extracts human movement feature points, targeted feature points and human movement matching features, it carries on three-dimensional reconstruction, and gets human skeleton 3D movement status, in the hope of creating more clearly and comprehensive information for scientific sports analysis.

In human skeleton three-dimensional reconstruction, feature points selection, feature points extraction and feature points and sports matching relations are key links, the paper puts aside kinematics and physiology feature points selection technology, focuses on stating feature points extraction and matching algorithm, which builds basis for human movement three-dimensional reconstruction. This chapter divides two sections to state image's movement objects feature points extraction and matching as well as their algorithm, explores pictures and movement objects correlation.

Feature points extraction algorithm

The paper adopted feature points extraction and matching algorithm is famous Harris Corner algorithm, the algorithm is implemented on the principle that based on SSD dissimilarity degree function, from which SSD function and NCC function are used as comparison criterion on image regions. If use w to express image window, use I and I' to express intensity of image points inside the window, then dissimilarity degree function-based two images dissimilarity degree calculation is as formula (1)shows:

$$\mathrm{SSD}=\iint_{w}\left[I(x,y)-I'(x,y)\right]^2 dxdy \tag{1}$$

Define NCC similarity degree measurement function expression is as formula (2)shows:

$$\begin{cases}\mathrm{NCC}=N(I',I)\Big/\sqrt{N(I',I)N(I,I)}\\ N(A,B)=\iint_{V}\left[A(x,y)-\overline{A}\right]\left[B(x,y)-\overline{B}\right]dxdy\end{cases} \tag{2}$$

Movement video feature points extraction and matching are discussed on the basis of formula (1) and formula (2), in image feature points most important feature is that one feature point located window should be different from its human adjacent windows, when reference feature points and adjacent feature points have a small deviation, it can get two images approximate dissimilarity degree function, as formula (3) shows:

$$\begin{cases}\mathrm{SSD}(u,v)=\begin{bmatrix}u & v\end{bmatrix}\mathbf{M}\begin{bmatrix}u\\ v\end{bmatrix}\\ \mathbf{M}=\sum_{x,y}w(x,y)\begin{pmatrix}I_x^2 & I_xI_y\\ I_xI_y & I_y^2\end{pmatrix}\\ \mathrm{SSD}\approx\Delta^T M\Delta\end{cases} \tag{3}$$

In formula (3), $w(x,y)$ represents weight, in general, the weight takes 1, we know that moving windows have image gray changes, therefore it can apply formula (3) $\mathbf{M}$ matrix feature value to analyze, if use λ_1 and λ_2 to represent matrix $\mathbf{M}$ feature value, then it can apply as Figure 2 showed oval $\mathrm{SSD}(u,v)$ attributes to analyze, in Figure 2, short axis

direction shows picture gray fastest changing direction, long axis direction represents slowest gray changing direction, two feature values get bigger show that gray change gets bigger, on the contrary, it gets smaller.

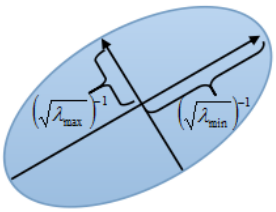

Figure 2 : $\mathbf{M}$ matrix feature value oval $\mathrm{SSD}(u,v)$ attribute analysis

Apply Figure 2 oval analysis method to analyze Figure 3 gray change status, it is clear that :in a region's $\mathbf{M}$ matrix two feature values relationship λ_2 is far larger than λ_1, that is to say, original image vertical gray change in the region is far larger than horizontal gray change, it can be thought as edge region, by analogy, when both two feature values are very big, region is angular point, when both two feature values are very small, region is plane, when two feature values gap is very big, region is edge.

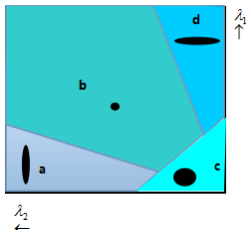

Figure 3 : $\mathbf{M}$ matrix feature value attributes

According to above analysis, targeted at b region's points, it makes analysis, introduces corner cost function as formula (4) shows, in formula, k represents experience constant, the constant generally takes number within [0.004, -0.006].

$$\begin{cases} R = \det\mathbf{M} - k(trace\mathbf{M})^2 \\ \det\mathbf{M} = \lambda_1\lambda_2 \\ trace\mathbf{M} = \lambda_1 + \lambda_2 \end{cases} \quad \textbf{(4)}$$

And then according to cost function, calculate Figure 3 four regions corner cost value, obtained result is that R_a is less than zero, R_b is larger than zero, R_c absolute value is small, R_d is less than zero.

Feature points matching algorithm

According to 3.1algorithm, it can extract image feature points, in order to match feature points, the paper utilizes NCC similarity degree function to do matching and judgment on feature points, if set image's feature points as (a,b), target is to look for image's one matching point, then its algorithm steps are as following shows :

STEP1. In feature point (a,b) located region, take the point as center or a rectangular window, window length and width ratio is set as ten-to-one.

STEP2. In STEP1 obtained rectangular window, look for all feature points, and respectively calculate each feature point and (a,b) NCC similarity degree, and NCC window size can be freely defined.

STEP3. Look for all points that NCC is larger than threshold values, in general threshold value is set between 0.7~0.8, if the region all points don't have points that larger than threshold value, then is thought that matching fails.

STEP4. Look for NCC maximum point from all points in NCC that are larger than threshold value, and think the point is (a,b) matching point.

The matching algorithm is applicable in case that continuous image sequence and image angle of view get closer, and the paper researched sports video meets such condition, therefore the paper selects Harris Corner to do feature points' extraction and matching on video images.

MOVEMENT'S HUMAN SKELETON THREE-DIMENSIONAL RECONSTRUCTION'S MOVEMENT INITIALIZATION PROCESS AND EXPERIMENT RESULT ANALYSIS

Basic matrix refining process

It has definition of basic matrix as formula (5)shows:

$$\mathbf{x}'^{\mathrm{T}}\mathbf{F}\mathbf{x} = \mathbf{0} \tag{5}$$

In formula (5), $\mathbf{x}' \leftrightarrow \mathbf{x}$ shows adjacent two pictures any one pair of matching points, if it can give enough numbers of matching points, equation (5)basic solves unknown matrix $\mathbf{F}$, here $\mathbf{F}$ is basic matrix, if it has n group of matching points then it can construct a homogeneous linear equation set, and then apply SVD分decomposition, and then it can get equation set least square solution, the algorithm is called eight-point method, but eight-point method obtained matching quantity is far larger than eight pairs, in order to refine basic matrix, the paper adopts non-linear iteration method to refine initial matrix, finally selects minimized Sampson cost function to solve relative accurate basic matrix, from which Sampson cost function is as formula (6)shows:

$$\sum \frac{\mathbf{x}'^{\mathrm{T}}\mathbf{F}\mathbf{x}}{\left(\mathbf{x}'^{\mathrm{T}}\mathbf{F}\right)_1^2 + \left(\mathbf{x}'^{\mathrm{T}}\mathbf{F}\right)_2^2 + \left(\mathbf{F}\mathbf{x}\right)_1^2 + \left(\mathbf{F}\mathbf{x}\right)_2^2} \tag{6}$$

Firstly apply linear algorithm to calculate $\mathbf{F}$ initial status, then apply, Levenberg-Marquardt iteration minimum as formula (6) showed Sampson cost function to get a refined basic matrix, the algorithm is called gold standard algorithm.

In calculating initial basic matrix $\mathbf{F}$, it needs to select best $\mathbf{F}$ according to one judgment condition, the paper introduces RANSAC searching engine in the hope of reducing calculation times, RANSAC robust estimation algorithm steps are as following:

STEP1. Randomly select s pieces of data points from sample S as model examples.

STEP2. Define model distance threshold value t range's data point set Si, point set Si is sample consistency set.

STEP3. If point set Si size is larger than one threshold value T, then use Si all estimated models and end algorithm.

STEP4. If point set Si size is less than T, select a new sub set and repeat steps 1-3.

STEP5. Through N times of experiments, select maximum consistency set Si, and use Si all points to re-estimate model.

In RANSAC robust estimation, sample times adopt adaptive algorithm, firstly set sample counter as 0, and set sample quantity as infinite great, then when sample times are larger than sample counter, it needs to select a sample and calculate inner points number, let ε =1-(inner points number)/(total points number), and take $p = 0.99$, by ε and formula (7), it solves sample times, and on this basis, let sample counter to add 1, finally ends adaptive algorithm.

$$N = \frac{\log(1-p)}{\log\left[1-(1-\varepsilon)^s\right]} \tag{7}$$

1) Basic matrix refining process programming:
2) Calculate every video image feature point.
3) Calculate on feature points matching.
4) RANSAC robust estimation.
5) Non-linear estimation.
6) Guide matching.

In distance measurement, it needs to provide a current estimation for $\mathbf{F}$ from RANSAC samples, use distance d to measure one pair of matching points meet epipolar geometric approximate degree, the paper adopted points and epipolar lines' Euclidean distance to measure distance that is reprojection error method.

The paper adopted searching window size is 100 pixel; internal threshold values is 1.5 pixel, guiding matching is proceeding in epipolar line 1.5 pixel nearby searched parameters.

Initialization process principle analysis

Initialization process targeted structure and movement, in the paper human skeleton three-dimensional reconstruction research, it adopts bi-camera different angles to implement by simultaneously shooting human movement, and therefore it needs to initialize main frame and structure. In initialization process, it needs to select two images that fit for structure movement initialization, by the two images basic matrix initialization video camera matrix, main frame initialization is newly-established world center and first image video camera center overlapping, video camera matrix is as formula (8) shows:

$$\mathbf{P}_1\left[\mathbf{I}|\mathbf{O}\right];\mathbf{P}_2=\left[\left[\mathbf{e}\right]_x\mathbf{F}+\mathbf{e}\mathbf{a}^{\mathrm{T}}|\mathbf{e}\right] \tag{8}$$

If main frame initialization correctly ends, then it can directly apply linear triangular method to calculate space 3D point. In general, measured points have errors, therefore ray from point back projection is not collinear, and non-collinear points are called undesirable measurement points. For undesirable measurement points, it needs to make adjustment, the paper adopts Sampson approximately to correct measurement points, as Figure 4 shows Sampson correction schematic graph.

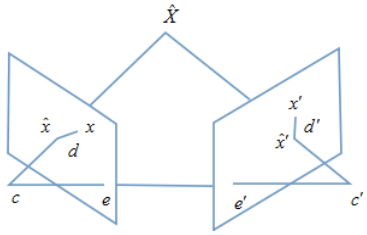

Figure 4 : Sampson correction schematic diagram

Sampson correction computational formula is as formula (9)shows:

$$\begin{bmatrix}\hat{\mathbf{x}}\\ \hat{\mathbf{y}}\\ \hat{\mathbf{x}}'\\ \hat{\mathbf{y}}'\end{bmatrix}=\begin{bmatrix}\mathbf{x}\\ \mathbf{y}\\ \mathbf{x}'\\ \mathbf{y}'\end{bmatrix}-\frac{\bar{\mathbf{x}}'^{\mathrm{T}}\mathbf{F}\bar{\mathbf{x}}}{\left(\mathbf{F}\bar{\mathbf{x}}\right)_1^2+\left(\mathbf{F}\bar{\mathbf{x}}\right)_2^2+\left(\mathbf{F}^{\mathrm{T}}\bar{\mathbf{x}}'\right)_1^2+\left(\mathbf{F}^{\mathrm{T}}\bar{\mathbf{x}}'\right)_2^2}\begin{bmatrix}\left(\mathbf{F}\bar{\mathbf{x}}'\right)_1\\ \left(\mathbf{F}\bar{\mathbf{x}}'\right)_2\\ \left(\mathbf{F}\bar{\mathbf{x}}\right)_1\\ \left(\mathbf{F}\bar{\mathbf{x}}\right)_2\end{bmatrix} \tag{9}$$

When undesirable measurement points get through correction, then can apply linear triangular method to initialize its corresponding three-dimensional space point, take two equations from every image point, totally give four secondly unknown four equations to compose into linear equation sets, the equation sets least square solution is the homogeneous expression of required initialized three-dimensional space point coordinate.

Established initial 3D point structure, but current structure and video camera simultaneously carry on a shooting transformation as Figure 5 shows; generated model meets constraint as formula (10) shows:

$$\mathbf{x}_i=\mathbf{P}\mathbf{X}_i=\left(\mathbf{P}\mathbf{H}_{\mathbf{p}}^{-1}\right)\left(\mathbf{H}_{\mathbf{p}}\mathbf{X}_i\right) \tag{10}$$

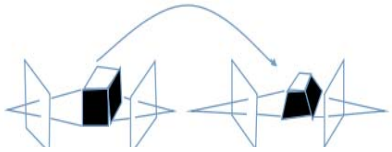

Figure 5 : Shooting transformation process schematic diagram

Three-dimensional reconstruction result display

The paper adopts Harris Corner algorithm to make feature points extraction and matching on video sequence images, applies eight-point algorithm, gold standard algorithm and RANSAC searching engine to refine basic matrix, carries on points extraction and three-dimensional reconstruction on men aerobic and men martial arts movement process human skeleton that is obtained after main frame, structure and video camera calibration, obtained athletes feature points extraction result is as Figure 6 and Figure 7 show.

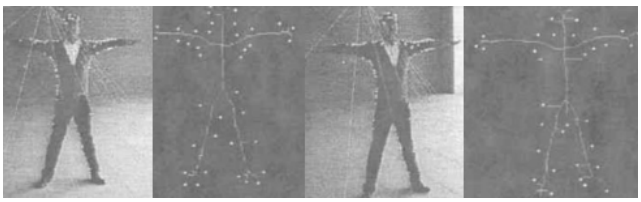

Figure 6 : Men's aerobics one frame movement feature point's extraction result

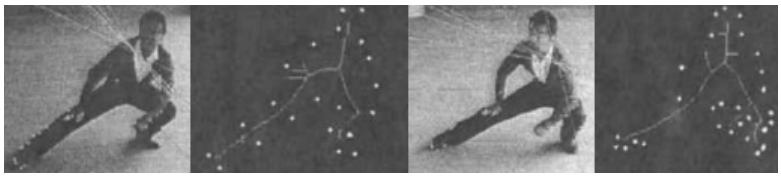

Figure 7 : Men's martial arts one frame movement feature point's extraction result

Carry on basic matrix refinement, Sampson correction and linear triangular method on extracted feature points then can get human skeleton three-dimensional reconstruction, as Figure 8 shows men aerobics and men martial arts athletes' extracted feature points corresponding three-dimensional space points.

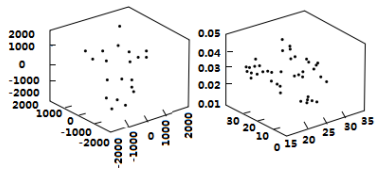

Figure 8 : Human skeleton three-dimensional space points experiment result

CONCLUSION

The paper firstly states human body sports capturing and three-dimensional reconstruction research background and research status, it provides basis for current such technique research orientations, and paves the way for sports video sequence image targeted analysis. Start from human skeleton three-dimensional reconstruction feature points extraction and matching algorithm, detail Harris Corner algorithm application in moving image feature point extraction, secondly make principle statements on human skeleton three-dimensional reconstruction's initialization process and basic matrix refining process, it provides theoretical basis for image and human movement correlations and three-dimensional reconstruction algorithm exploration. Apply Harris Corner algorithm, eight-point algorithm, gold standard algorithm, Sampson correction algorithm and linear triangular algorithm to carry on feature point extracting、matching and three-dimensional reconstruction to do experiments on men aerobics and men martial arts events one frame image, obtained result is relative satisfactory.

REFERENCES

[1] Huang Yin hua, Ouyang Liu qing, Kang Chang fa et al.; On the Chinese soccer fan culture[J], Journal of Wuhan Institute of Physical Education, **36(6)**, 7-9 **(2002)**.

[2] Lin Zi-Yong; A discussion on the Chinese football fans in cultural perspective[J], Sports Sciences Researches, **9(2)**, 44-49 **(2005)**.

[3] Fu Dao-Hua, Zhang Pei-Zhi, Meng Xian-Lin; Study on economic culture function of soccer fans and factors of aberration actions and prevention measurements[J], China Sport Science and Technology, **42(6)**, 33-37 **(2006)**.

[4] Bi Bo; The analyzes to basic integrant part of the soccerfan culture's connotation[J], Sports & Science, **28(5)**, 68-70, 67 **(2007)**.

[5] Liu Kai; Analysis of soccer globalization on world soccer sports development[J], Bulletin of Sport Science & Technology, **16(12)**, 18-20 **(2008)**.

[6] Zhao Gui-Sheng, Han Xin-Jun, Chen Jian-Sheng, Xie Lun-Li, Suo Yan-Jun, Hu Xiao-Hua; Reason analysis and countermeasures on misbehavior of football fans[J], Bulletin of Sport Science & Technology, **19(2)**, 10-12 **(2011)**.

[7] Zhou Xiujun, Mao Zhichen; The deviant behavior classification and the trend analysis of chinese soccer fans on the soccer field[J], Sports & Science, **32(6)**, 103-106 **(2011)**.

[8] Qiu Jun, Li Kai-Xian; Sun bao-jieinstitute of physical education and research, tsinghua university, Beijing China, Formation of fan's impermissible behaviors and preventing measurements in sports competition[J], China Sport Science, **24(12)**, 18-22 **(2004)**.