Recognition of Human Posture from 3D Surface Data of Laser Range Sensor Using Human Ellipsoid Tree Model and Blackboard Model

For Human Surveillance

Shigeyoshi Nakajima, Yuuki Fukui, Yukie Hashimoto, Takashi Toriu, and Hiromitsu Hama
Graduate School of ENG,
Osaka City Univ. (OCU)
Osaka City, JAPAN
{ nakajima; toriu }@info.eng.osaka-cu.ac, { fukui; yhashimoto}@info.eng.osaka-cu.ac ,
hama@ado.osaka-cu.ac.jp

***Abstract*— In this paper we propose a new method to recognize human posture from 3D surface data using laser range sensor (LRS). Recognition of human posture from data of a sensor is an important technique and very useful for a surveillance or other purposes. There are many sensors used for human posture recognition. We selected a laser range sensor and a tilting moving table to measure positions of 3D points on a human surface. We employed a black board system to create instances of human ellipsoid tree models (HETM). The best evaluated instance showed us a posture of a human. The experimental results show that the proposed method can recognize human 3D postures properly in most cases.**

Keywords; 3D surface data; laser range sensor; human posture; ellipsoid; blackboard model; surveillance.

I. INTRODUCTION

Recognition of human posture is an important and useful. It will help surveillance system in a public space and a home of a solitary person. There are many recent works to detect human posture or object using surveillance camera or other sensor with 2D data structure[1-5]. Other researchers worked with 3D structures [6-8]. F. Gomez-Caballero et al. [9] used a time of flight (TOF) camera. TOF camera is an immediate sensor to measure a 3D surface. But a view angle of TOF is narrow and its price is expensive. Other researchers used Laser range sensor or Laser Detection and Ranging (LADER). S. Kumar et al.[10] approximated series of 3D points by hyperquadrics curved surface. L. Navarro-Serment [11] projected 3D surface data to 2D planes perpendicular to others and extracted features from data on the 2D planes. T. Nakamura et al.[12] employed straight homogeneous generalized cylinder (SHGC) method. But SHGC is too flexible to express 3D surface of human bone structure. L. Chevalier et al.[13] shows the method of superquadric modeling and superellipsoid of 3D objects. Their method was sophisticated. But robustness is required in recognition of human body from LRS data. Y.Yang and X. Shi [14] employed a blackboard (BB) model for a cardiac diagnostic system because BB provides robust recognition.

Mikuriya et al.[1] used skeleton structure but didn't use of thickness of human body parts because they dealt with 2D image. We use 3D bone structure to recognize 3D surface data of LRS. We select ellipsoid rather than generalized cylinder, superquadric or superellipsoid. Human bones are solid tissues. Their shape don't change in human posture (except slight dilatation and constriction of a thorax according to breath) But there are soft tissue on the bones, muscles, adipo, viscera and skin. Also there are clothes on a bare human body. So a human surface differs from the bone structure but the difference is small compared to a total shape. And LRS scanning is rather coarse compared to high resolution 2D image data. The number of points is almost 2000 to 3000 for one person in our experiments. And LRS scanning data sometime partially vanish. Occlusion is one of the reasons. A type of material is another reason. Hair often vanishes in LRS scanning. And a fur coat does. Robust posture recognition method must deal with such cases. We constructed human ellipsoid tree models (HETM) according to a structure of bones of a human. An optimization process leads HETM to a proper 3D structure of a human posture. Ellipse is less consumable of computation cost than generalized or superellipsoid in optimization process. We use also Blackboard (BB) model. BB leads initialization and optimization of instances of HETM to accuracy and low computation.

The rest of this paper structured as follows. In section 2, a detail of the proposed method is described. In section 3, there are experimental results. Section 4 shows consideration. Section 5 shows conclusions and future works.

II. METHOD

A. 3D Surface Data Using LRS and Tilting Table

To capture 3D surface data of a human body, we used a LRS censor (URG-04LX-UG01 made by HOKUYO AUTOMATIC CO.,LTD.) using active infra-red ray on a tilting table. Fig. 1 shows a black LRS on a white tilting table.

Figure 1. LRS on a Tilting Table

Fig. 2 shows data of a human body lying on a floor. We took also data of a floor only as background data. A differential processing gives us data of a human body only as Fig. 3.

We took shots of three types of postures. One is a straight posture. Also we tool shots of postures of straight and spread limbs (arms and legs) and postures with bending limbs. And we took shots of 8 directions of body axis for each posture.

B. Human Ellipsoid Tree Model (HETM)

We designed a data structure of a human body for recognition of a posture. Superellipsoid and superquadric are sophisticated structure. But there are cloths on a human bare body. So sophisticated structure are easily affected such artifacts, e.g. frills, button and collar of cloths. We selected simple ellipsoids for robust recognition. There are 11 main bone parts in a human body. A cranial bone (head), a thorax (chest), a pelvis (hip), 2 femora (thigh), 2 sets of a tibia and a fibula (shin), 2 humeri (upper arm) and 2 sets of a radius and a ulnas (forearm). We ignored a hand bone and a foot bone because they are smaller than the parts described above and they often change their shapes. We made a rooted tree which consisted of 11 ellipsoids corresponding to each bone part. We decided that a chest is the root of the tree. Because it is the largest part. A head, a hip and 2 upper arms are child nodes of the chest node. 2 thighs are child nodes of a hip. A shin is a child node of a thigh. A forearm is a child mode of a shin. Fig. 4 shows a total view of a HETM.

A child node part and its parent node part connect to each other at each specific point in their coordinates. We call such the connecting point as a joint. A joint can change its angle. When a thigh changes its direction the connecting shin also moves in conjunction.

C. Blackboard Model

We employed blackboard model for recognition of a human posture. Fig. 5 shows our blackboard model. In an initial state, there are only LRS points and HETM in BB. The BB estimates values of a center and body axis of a human body in bottom up way. Then HETM creates several instances according to the center, the body axis with top-down estimation. Then in next stage BB changes variables of the instances to optimize them for LRS points.

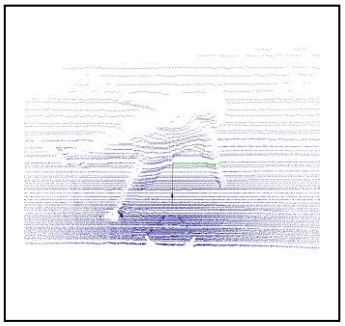


Figure 2. 3D Points of a Human Body and a Floor

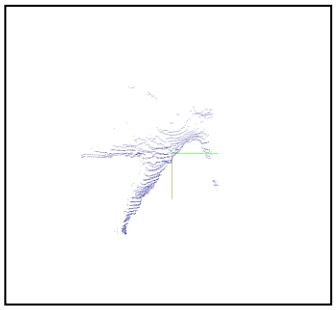


Figure 3. 3D Points of a Human Body

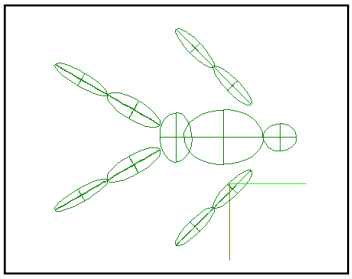


Figure 4. Human Ellipsoid Tree Model

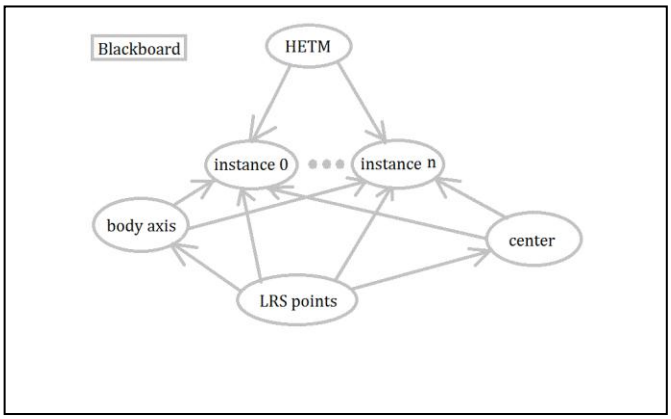


Figure 5. Blackboard Model

Identify applicable sponsor/s here. If no sponsors, delete this text box. (sponsors)

D. Center and Body Axis

BB estimates the center of a human body from LRS points in bottom up way. The initial value of the center $\vec{M}$ is the mean of LRS points $\vec{L}_i$ *(i=1,…,N).*

$$\vec{M} = \frac{\sum_{i=1}^{N} \vec{L}_i}{N} . \qquad (1)$$

And then BB estimates the body axis on a coronal plane. BB selects a coronal plane before it. BB prepares candidates of planes. For example a plane perpendicular to *x* axis is expressed as ($x=M_x$). M_x is the *x* component of $\vec{M}$.

There are other 8 planes. There are total 9 planes.

$$y=M_y .$$
$$z=M_z .$$
$$x+y=M_x+M_y .$$
$$x-y=M_x-M_y .$$
$$z+y=M_z+M_y .$$
$$z-y=M_z-M_y.$$
$$x+z=M_x+M_z .$$
$$x-z=M_x-M_z. \qquad (2)$$

If there is an body axis $\vec{b}$, BB must select a plane *p* nearest to $\vec{b}$ among the 9 planes. BB calculates the mean distance from a plane to $\vec{L}_i$. A plane with the lowest mean distance is selected as *p*. A value *d* is the difference angle between $\vec{b}$ and *p*. The maximum value of *d* is 12.45 deg. In *p* , BB selects a direction from 4 directions (0, 45, 90, 135 deg) as the approximated body axis $\vec{b}'$. BB calculates a standard deviation of feet of perpendicular of all $\vec{L}_i$ along each direction. BB selects the direction with the maximum standard deviation as the approximated body $\vec{b}'$. The maximum difference from *b'* to the optimal direction on p is 22.5 deg (i.e. 45/2 deg). The difference angle from a direction of a unit vector with vertical 12.45 deg to a direction of a unit vector with horizontal 22.5 deg is 25.5 deg. So $\vec{b}'$ is different from $\vec{b}$ up to 25.5 deg. The difference of 25.5 deg seems to be trivial as an initial difference of a HETM instance.

E. Initial Angles of Joints

Before creation of HETM instances BB decides initial angles of joints from LRS points. BB makes a cylinder which center is of $\vec{M}$. The diameter of the cylinder is 1/3 of a range of feet of a perpendicular of all $\vec{L}_i$ along $\vec{b}'$. If more than 10% of of $\vec{L}_i$ are out of the cylinder BB estimates the body is not ramrod-straight (some arms and legs are far from the body axis). So BB makes an initial posture of an instance like Fig. 4. If less than 10% of of $\vec{L}_i$ are out of the cylinder BB estimates the body is ramrod-straight like as Fig. 6(c). Fig. 6(a) shows a LRS shot of a ramrod-straight posture and a cylinder. Fig. 6(b) shows a LRS shot of a spread posture and a cylinder.

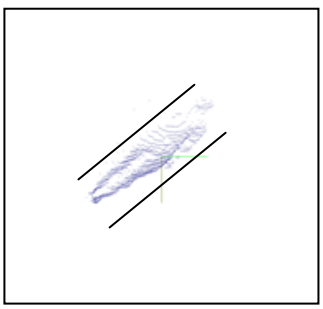
F. Instantiation of HETM

According the value of $\vec{M}$, $\vec{b}'$ and a straight flag, instances of HETM are created. The total size of the instance is adjusted to the range of feet of a perpendicular of all $\vec{L}_i$ along $\vec{b}'$. But the sign of body axis (Which direction is a head direction?) is still unknown. So 2 instances are created with a positive direction and a negative direction. A wrong direction instance will be rejected and a proper direction instance will be selected as a correct answer after optimization with evaluation function.

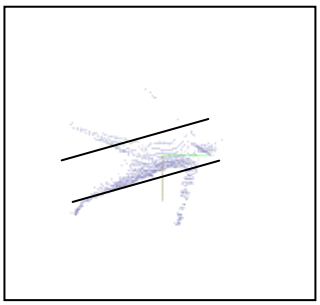
G. Evaluation Function

An instance h_j is evaluated by function *f*(h_j). An instance h_j has ellipses ε_{jk} (k =1,…,11). A distance d_{ijk} from $\vec{L}_i$ to ε_{jk} can be calculated. Then the distance d'_{ij} from $\vec{L}_i$ to h_j is evaluated as the minimum about *k*.

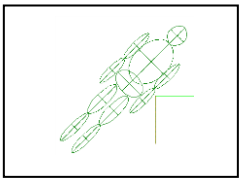
$$d'_{ij} = \min_k (d_{ijk}) \qquad (3)$$



(a) Ramrod-Straight Posture with Cyliner



(b) Spread Posture with Cylinder



(c) HETM Instance for Ramrod-straight Case

Figure 6. Ramrod-straight or Not

Function $f(h_j)$ is mean value of d'_{ij}.

$$f(h_j) = \frac{\sum_{i=1}^{N} d'_{ij}}{N} \qquad (4)$$

Each h_j is evaluated using $f(h_j)$ and changes its parameters (a position and directions) to optimize, i.e. to reduce a value of $f(h_j)$ in processes described below.

H. Optimization of Center

The center position of the HETM is the center of the chest node. It is initialized to $\overrightarrow{M}_i$. BB changes the center position to reduce $f(h_j)$. BB uses steepest descent method for reduction. First BB changes the position with gradient descent method down to 1mm. In this center optimization process directions of ellipsoids, sizes of ellipsoids and angles of joints don't change.

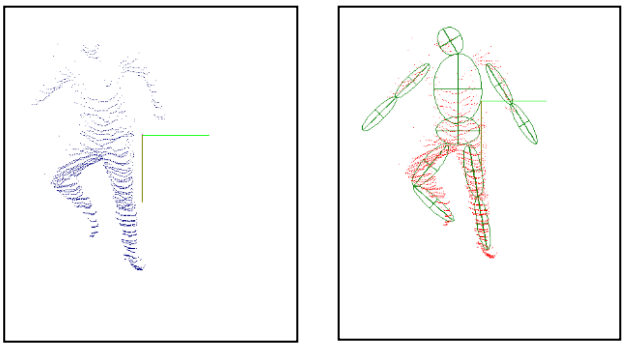
I. Optimization of Directions and Joint Angles

After center optimization, a HETM instance changes a direction of a chest and joint angles between parent nodes and child nodes about roll pitch yaw. It will be executed in a breadth first order.

Fig. 7 shows one of the results of a bending leg shot after changes of joints. Fig. 7(a) shows the LRS points. Fig. 7(b) shows a HETM instance after changes of angles of knee joints. One knee is almost straight. But another knee bends. The result of the HETM instance slightly misaligns in Fig. 7(b). But the misalignment is in an acceptable error range in this case.

J. Selection

After optimization BB selected a instance h_j with the lowest $f(h_j)$ as a correct answer. There are 2 instances according the sign of a head direction.Fig. 8 shows another HETM instance with the opposite side comparing to Fig. 7(b). The legs widely misalign. Then the value of $f(h_j)$ of Fig. 8 is higher than the value of Fig. 7(b).



(a) LRS Points (b) HETM Instance

Figure 7. LRS Points and HETM Instance of Bending Leg Shot after Changes of Joint Angles

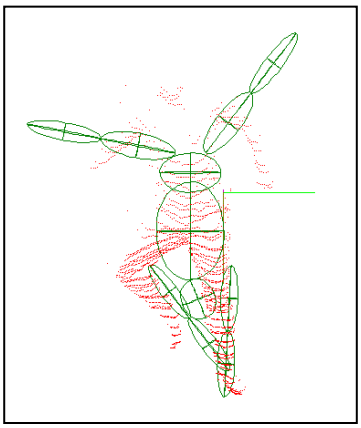


Figure 8. HETM Instance of Bending Leg Shot with Opposite Side

III. Experimental Result

A. Direction of Head

Fig. 9 shows the results of $f(h_j)$ (mm) of shots of straight postures.

In Fig. 9 black lines show the values of $f(h_j)$ of instances with proper directions of heads. Gray lines show instances with opposite sides. The highest value of black lines in Fig. 9 is 35 mm. The lowest value of gray lines in Fig. 9 is 40.1 mm. We can discriminate correct instances from wrong instances with threshold value 38mm in Fig. 9. The success of discrimination means a recognition of a proper direction of a head.

Fig. 10 shows the results of shots of spread postures. Fig. 11 shows the results of shots of bending legs postures. The highest value of black lines in Fig. 10 and Fig. 11 is 45.4mm. The lowest value of black lines in Fig. 10 and Fig. 11 is 50.5mm. We can discriminate correct instances from wrong instances with threshold value 38mm in Fig. 10 and Fig. 11.

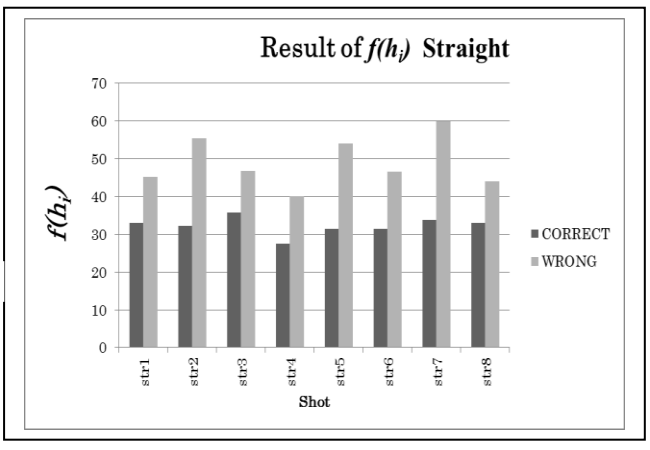


Figure 9. Result of Straight Postures

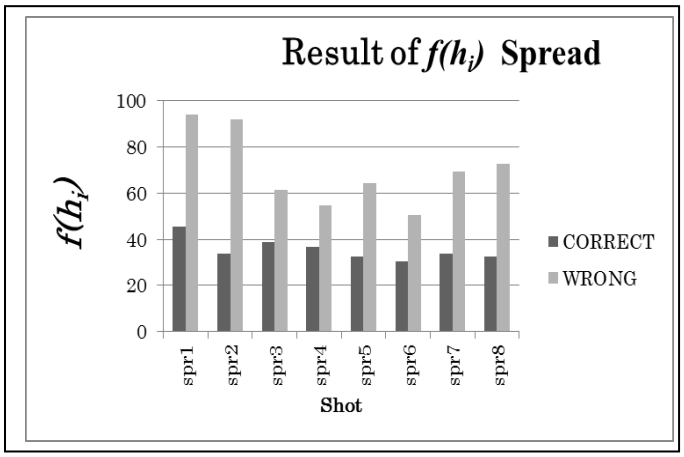


Figure 10. Result of Straighe Postures

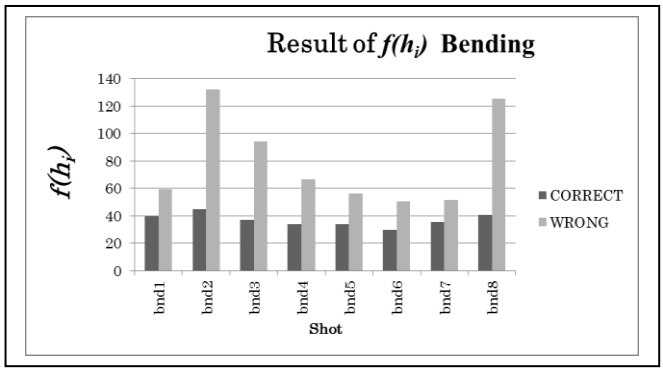


Figure 11. Result of Bending Limb Postures

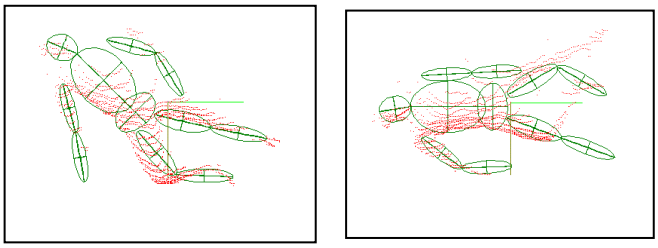
BB can discriminate a correct instance from a wrong instances if BB knows whether the shot is straight or not as described in Fig. 6 and Ⅱ E. The accuracy was 100 %.

But the lowest value of gray lines in Fig. 9 is lower than the highest of black lines in Fig. 10 and Fig. 11. We cannot discriminate properly without knowledge of straightness.

B. Angle of Joints

Fig. 12 shows 2 instances. We think that Fig. 12 (a) shows correct limbs. But we think that Fig. 12 (b) shows wrong limbs. In both shots the subject bended his right knee and left elbow and didn't bend his right elbow and left knee. But in Fig. 12(b) the instance assume a right foot as a left shin. The number of such apparently wrong instance with correct head direction is 1 in 24 shots. The accuracy was 95.8 %.

In the case of Fig. 12(b) the instance assumed that the left arm was straight. It differs from the shot. But it is the result of partial vanishments of LRS data caused by occlusion. We didn't count it as a failure. We dealt results of partial vanishments as the same in other shots. Partial vanishments didn't affect total postures in our experiments.



(a) Correct Limbs (b) Wrong Limbs

Figure 12. Instances with Correct Limbs and Wrong Limbs

C. Misalignment

Fig. 13 is an enlarged part of Fig. 12(a). It seems that there is a misalignment in the result. We think that the instance assumed the right foot as a part of right sin. But we don't have reasonable criteria to evaluate such alignments of LRS data now. We will try to evaluate alignment in future

IV. CONSIDERATION

A. HETM

We ignored hands and feet in design of HETM. We thought that hands and feet are smaller than other parts such as thighs and forearm. But Fig. 12(b) and Fig. 13 showed us that feet affected the result of recognition of posture much. And there are many body types. Sometime there are soft tissues with large amount in a human body, i.e. breasts, belly and hip with much muscle and adipo to change a human appearance in LRS data. Also sometimes people wear individualistic cloths, for example a big hat, a big collar, long sleeves and a long skirt. And sometimes people have bags. We must improve the HETM to apply this method to such cases.

B. Evaluation Function $f(h_j)$

We used mean value of distance as an evaluation function $f(h_j)$ in (4). But there are other variations of evaluation, e.g. the percentage of points over 30.0mm. We need to investigate performance of other functions to solve cases like Fig. 12(b).

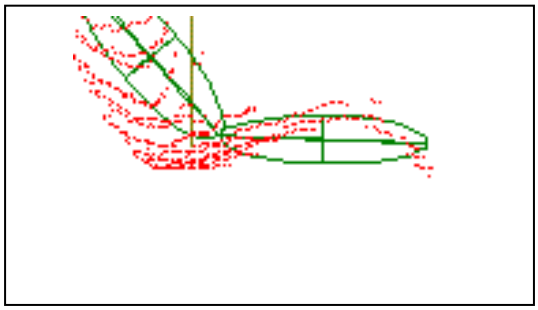


Figure 13. Misalignment

C. Optimization

BB optimized instances in a fixed order. Angles of joints changed in breadth-first order. There are more flexible ways to optimization e.g. A* search algorithm. But an improvement of accuracy often increases computation cost. We will investigate trade-off of accuracy and computation with an improved method.

After an initialization of an instance the relative size of child node in comparison to its parent node didn't change. It is also one of expected works.

D. Threshold of Evaluation Function

If Fig. 9,10 11 each threshold range was narrow. Also we couldn't decide a common threshold value for all postures. We must reduce values of evaluation functions of correct instances with improving the method.

V. CONCLUSION

We proposed a method to recognize a human posture from LRS 3D data using HETM and BB. The accuracies of the experimental results were very good. An execution was sudden and we didn't bother about time of the executions.

In future we will improve the method to be applied to data of other people with different body types, different cloths and belongings.

ACKNOWLEDGMENT

This work is partially supported by SCOPE: Strategic Information and Communications R&D Promotion Program (102307010). We also thank Assoc. Prof. Yougo Takada and a Ph.D. student Mr. Tomonori Tajiri in Osaka City University who let us use their 3D range sensor.

REFERENCES

[1] Takashi Mikuriya, Shinya Yoshitsugu, and Masaki Oshita, "Posture estimation from a single image using the structure of the human body", *Proc. of Meeting on Image Recognition and Understanding 2007*, Hiroshima , Japan, pp.1576-1581, 2007.

[2] Y. Hashimoto, Thi Thi Zin, T. Toriu and H. Hama "Fast multi-slit composition method using attitude measurement unit for pedestrian detection", *Inovative Computing, Information and Control Express Letters Part B: Application*, Vol.2, No.3, pp.2767-2766, June 2011.

[3] T. T. Zin, Hideya Takahashi and Hiromitsu Hama, "Robust person detection in far infrared images – methods based multi-slits and GC movement patterns -", *Inovative Computing, Information and Control Express Letters*, Vol.5, No.3, pp.751-761 , March 2009.

[4] M. Murshed, M. Kabir and P. Chae, "Moving object tracking – an edge segment based approach", *International Journal of Inovative Computing, Information and Control*, Vol.7, No.7(A), pp.3963-3979 , July 2011.

[5] D. Fernandez, I. Parra, M.A. Sotelo, and P.A. Revenga, "Bounding box accuracy in pedestrian detection for intelligent transportation systems", *Proc. of 32nd Annual Conference on Industrial Electronics* , pp. 3486 - 3491, Madrid SPAIN, Nov. 2006.

[6] M. Bertozzi et al., "Pedestrian detection by means of far-infrared stereo vision", *Computer Vision and Image Understanding*, Vol.106, Issue 2-3, pp.194–204, May 2007.

[7] T. Komori, H. Nagahara, A. Shimada and R. Taniguchi, "Complementary SLAM using monocular and Binocular vision", Proc. of Symposium Hinokuni 2011, p.C-4-2 , Fukuoka JAPAN, March 2011 .

[8] A. Howard , L. H. Matthies, A. Huertas, M. Bajracharya and A. Rankin, "Detecting pedestrians with stereo vision: safe operation of autonomous ground vehicles in dynamic environments", Proc. of the 13th. International Symposium of Robotics Research, Nov. 2007.

[9] F. Gomez-Caballero, T. Shinozaki and S. Furui, "User Identification Using Time-of-Flight Camera Image Streams", *Information Processing Society of Japan Technical Report*, Vol.2, pp.615-616, March 2010.

[10] S. Kumar, H. Song, D. Goldgof and K. Bowyer, "On Recovering Hyperquadrics from Range Data", *Institute of Electrical and Electronics Engineers Trans. on Pattern Analysis and Machine Inteligence"*, Vol.17, No.11, pp.1079-1083, Nov. 1995.

[11] L. Navarro-Serment, C. Mertz, and M. Hebert, "Pedestrian Detection and Tracking Using Three-Dimensional LADAR Data", *The International Journal of Robotics Research October 2010*, Vol.29, pp.1516-1528, May 2010.

[12] T. Nakamura, M. Asada and Y. Shirai, " Recovery of an SHGC Shape and Pose from Shading and Contour Based on Weak Lambertian Assumption", *Journal of Institute of Electronics, Information and Communication Engineers* , Vol.J77-D-Ⅱ,No. 7 1226-1235, Jul. 1994.

[13] L. Chevalier, F. Jaillet and A.Baskurt, "Segmentation and superquadric modeling of 3D objects", *Journal of Winter School of Computer Graphics (WSCG'03)* , Pilsen, Czeck Rep. , Vol. 11, No. 2, pp. 232-239, Feb. 2003.

[14] Y. Yang and X. Shi, "Cadiac Arrhythmia Interpretation System Based on Blackboard Model", *Biomedical Soft Computing and Human Sciences*, Vol.5, No.2, pp.17-22, 2000.