

지능형 로봇 시스템에서 하이브리드 실루엣 추출 방법을 이용한 인간의 몸 추출

Extraction of Human Body Using Hybrid Silhouette Extraction Method in Intelligent Robot System

김문환*, 주영훈**, 박진배*, 조영조***, 지수영***, 김혜진***

*연세대학교 전기전자과, **군산대학교, 전자공학부, ***한국전자통신연구원

Moon Hwan Kim*, Young Hoon Joo**, Jin Bae Park*, Young Jo Cho***, Su Young Chi***, Hye Jin Kim***

* Yonsei University (jmacs@control.yonsei.ac.kr, jbpark@yonsei.ac.kr)

**Kunsan, National University (yhjoo@kunsan.ac.kr)

*** ETRI, Intelligent Robot Research Division (youngjo@etri.re.kr, chisy@etri.re.kr, marisan@etri.re.kr)

ABSTRACT

This paper discusses a human body extraction method for mobile robot system. The skeleton features are used to analyze human motion and pose estimation. The intelligent robot system requires more robust silhouette extraction method because it has internal vibration and low resolution. The new hybrid silhouette extraction method is proposed to overcome this constrained environment. Finally, the experimental results show the superiority of the proposed method.

Key words : Skeleton model, temporal gradient, spatial gradient, hybrid silhouette, active contour.

1. Introduction

Human motion analysis is important research subject. It is receiving increasing attention for a wide spectrum of applications, such as man-machine interface, security surveillance, image retrieval and video indexing, and robot [1]. Especially, it is researched as a new method of human computation interaction (HCI) because human motion analysis concerns with the key techniques of HCI such as pose recognition and motion tracking. To analysis human motion, extracting features of human body from sequential image plays an important role. Without features from human body, it is not easy to analyze human motion.

Human motion analysis can be performed in two approaches. At first, the features of human body are acquired by capturing both position and motion information via sensors fixed on human joints. The sensors fixed on the human joints send major information to a main computer, and

then main computer analyze the major information then give the motion paymasters of human body. Without sensor, the color marks can be used instead of sensors. The main advantage of this approach is that it gives accurate features for motion analysis. However, it is not adapt to use in robot system or general situation because it needs sensing equipment and restrict environment.

Another approach extract features of human body from sequential images contain human body. It has been receiving extensive attention from computer vision and HCI researchers because it need only one camera and does not need main computer. Acquiring sequential images is performed in two ways as camera motion. The first one is to use camera fixed on environment. It gives sequential images including motionless background and human motion. The second one is to use camera fixed on moving object such as robot. This method gives sequential images including background and human motion. Most of researched is based

on the first method to capture sequential images. In this paper, however, we suppose the sequential image is acquired from camera fixed on robot which does not move. The sequential images from the supposed environment have following restrictions: 1) there exist background motion because robot has internal vibration. 2) The resolution of image is low. 3) It is not easy to accurate background model. 4) The position of camera is lower than human head. 5) The illumination condition is poor.

The objective of this work is to detect robust human body silhouette. In this paper, we presented a method of hybrid silhouette extraction method from sequential images restricted by robot system. First, the temporal gradient and spatial gradient are calculated from images. Then the hybrid silhouette is extraction by blending temporal gradient and spatial gradient. The motion region model is used to blend two gradient images.

2. Hybrid silhouette extraction method

The extraction of skeleton feature is based on silhouette of human body. There are three conventional approaches to silhouette generation: temporal differencing (two-frame or three frame) [7], background subtraction [3-6], and optical flow (see [8] for an excellent discussion).

The temporal differencing is very adaptive to dynamic environments, but generally does a poor job of extracting all relevant feature pixels. The background subtraction provide accurate silhouette of human body, but is extremely sensitive to change of image due to lighting and extraneous events. Optical flow can be used to detect features from image with background motion, but most optical flow computation methods are very complex and are inapplicable to embedded system such as robot system. However, these conventional approaches are not adaptable to extract silhouette from images restricted within robot system.

The hybrid silhouette extraction method is based on temporal differencing method. To overcome the poor extraction of relevant silhouette and restriction of robot system, we propose motion region model and compensation method for missing silhouette information.

2.1 Hybrid silhouette extraction

In this subsection, we will discuss about hybrid silhouette extraction method. Let sequential

image is define as

$$I(x, y; t) = [I_r(x, y; t), I_g(x, y; t), I_b(x, y; t)]^T = \mathbf{I} \cdot \mathbf{A}^T \quad (1)$$

Temporal gradient is temporal difference between consecutive two images. The temporal gradient and the spatial gradient including edge information are defined as

$$I_t = \frac{\partial I}{\partial t}$$

$$I_s = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right] \quad (2)$$

When we want to gradient information as key value of edge, gradient information should be mapped into natural number. In this approach, we simply use following mapping strategy,

$$\bar{I}_t = \|I_t\| \quad (3)$$

$$\bar{I}_s = \|I_s\| \quad (4)$$

In robot system, the temporal gradient \bar{I}_t can have three kind of undesired values: 1) \bar{I}_t can disappear because of camera problem. 2) \bar{I}_t has insufficient silhouette information. 3) \bar{I}_t has unnecessary motion information due to internal vibration of robot. To overcome first undesired situation, we should check sum of temporal gradient S_t defined as

$$S_t = \int_{\forall p \in I} f(\bar{I}_t, \gamma_t) dp$$

$$f(\bar{I}_t, \gamma_t) = \begin{cases} 1, & \text{If } \bar{I}_t > \gamma_t \\ 0, & \text{else} \end{cases} \quad (5)$$

where γ_t is minimum value to transform color image as binary image.

When S_t is zero, the old temporal gradient is replace to current temporal gradient.

The second and third undesired situation can be solved by using motion region model. To remove unnecessary motion information and complete insufficient silhouette information, we should use spatial gradient and history information of temporal gradient. History information of temporal gradient is the sum of consecutive some old temporal gradients. How to add spatial gradient and history information is key problem. The convex sum of temporal and spatial gradient

can generate reliable accurate silhouette. The convex sum denotes as

$$\eta \bar{I}_s(x, y) + (1 - \eta) \bar{I}_t(x, y) \quad (6)$$

where η is the convex sum parameter. By adjusting η , we can get accurate silhouette or imprecise silhouette. Generally, the convex sum parameter has static value. It is determined by experimental or manual method, generally. In this paper, but, convex parameter is determined by motion region model.

The motion region model is defined as $R \subset \mathbb{R}^2$. The initial motion region model has zero. Then the motion region mode is updated as following.

$$R(x, y; t+1) = \begin{cases} R(x, y; t) + \gamma_i, \bar{I}_t(x, y; t+1) > \gamma_i \\ R(x, y; t) + \gamma_d, \bar{I}_t(x, y; t+1) < \gamma_i \end{cases}$$

$$\gamma_i = \begin{cases} \bar{\gamma}_i, \bar{I}_s(x, y; t+1) > \gamma_s \\ \underline{\gamma}_i, \bar{I}_s(x, y; t+1) < \gamma_s \end{cases}$$

$$\gamma_d = \begin{cases} \bar{\gamma}_d, \bar{I}_s(x, y; t+1) > \gamma_s \\ \underline{\gamma}_d, \bar{I}_s(x, y; t+1) < \gamma_s \end{cases} \quad (7)$$

where γ_s is the binarization parameter of spatial image. $\bar{\gamma}_i$ and $\underline{\gamma}_i$ are upper and lower increasing parameters. $\bar{\gamma}_d$ and $\underline{\gamma}_d$ are upper and lower decreasing parameters. In subsection 2.2, the detailed description for parameters will be presented. Finally, the hybrid silhouette is calculated as

$$\bar{I}_b(x, y) = \eta \bar{I}_s(x, y) + (1 - \eta) \bar{I}_t(x, y)$$

$$\eta = \begin{cases} \eta_R, R(x, y) > \gamma_R \\ \eta_B, R(x, y) < \gamma_R \end{cases} \quad (8)$$

where η_R is convex sum parameter for motion region. η_B is the convex sum parameter for background. Determination of these parameters is also discussed in section 2.2. Figure 1 shows the procedure of hybrid silhouette extraction.

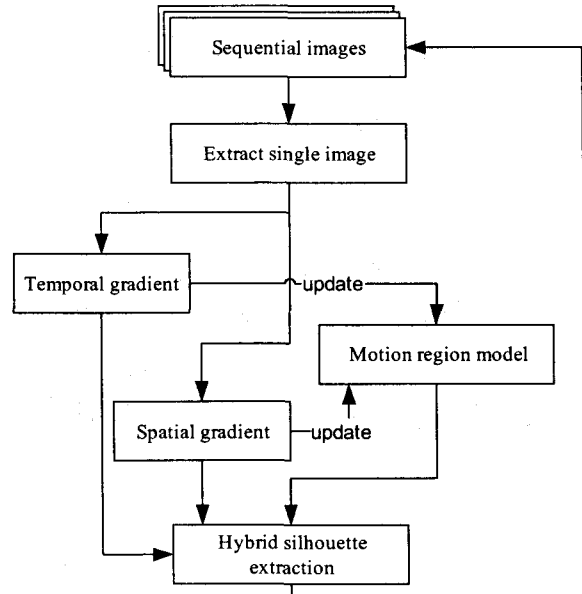


Figure 1. hybrid silhouette extraction process

3. Experimental results

The sequential images from intelligent robot system are obtained. The sequential images are captured as 320x240 resolutions with 24 bit color depth.

The hybrid silhouette extraction method is applied to sequential image. The parameters of extraction method are chosen manually. We perform experiments for three undesired situation as described in Section 2.1. Figure 2-4 show the results of hybrid silhouette extraction method under three undesired situation. We can check the accurate hybrid silhouette is extract in spite of unnecessary or insufficient silhouette information.

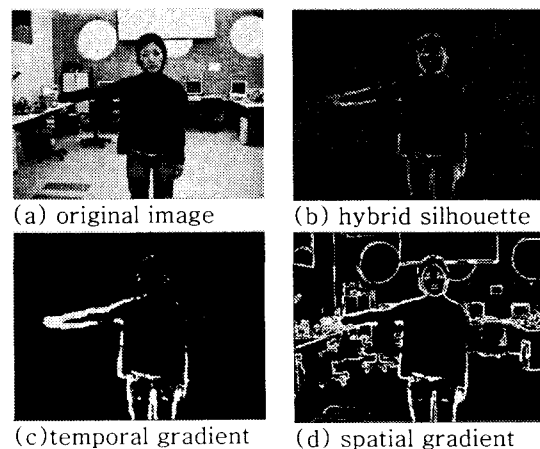


Figure 2. Hybrid silhouette extraction: \bar{I}_t has

insufficient silhouette

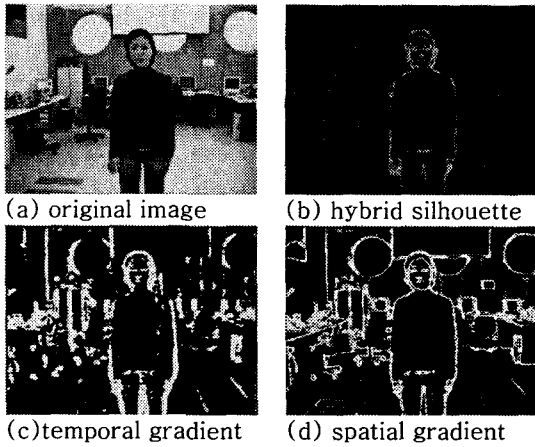


Figure 3. Hybrid silhouette extraction: \bar{I}_t has unnecessary motion information

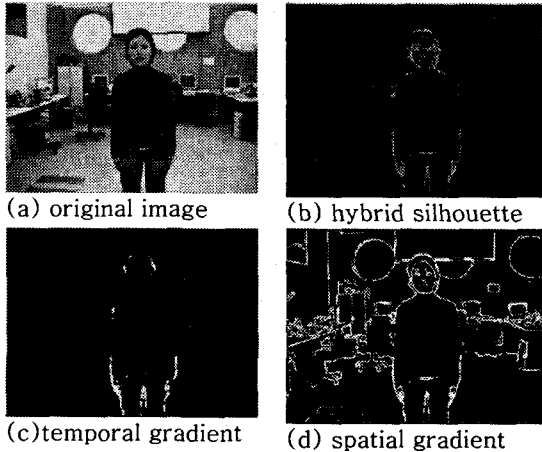


Figure 4. Hybrid silhouette extraction: \bar{I}_t is disappeared

4. Conclusion

In this paper, we proposed human body silhouette method by using hybrid silhouette extraction method in intelligent robot system. To overcome drawbacks occurring in robot system, temporal gradient and spatial gradient are blended via motion region model. The various experiments are performed to check performance of the proposed method. In the experiment results, we could check the superiority of the proposed method.

ACKNOWLEDGMENT

This work is a result of the URC project supported by MIC (Ministry of Information and Communication), KOREA.

References

- [1] J. K. Aggarwal and Q. Cai., "Human Motion Analysis: A Review," *Computer Vision and Image Understanding*, pp. 428-440, 1999
- [2] B. Fan; Z.-F. Wang, "Pose estimation of human body based on silhouette images," *International Conference on Information Acquisition Proceedings*, pp. 296-300, June, 2004.
- [3] Haritaoglu, R. Cutler, D. Hawood and L. Davis, "Backpack: Detection of people carrying objects using silhouettes," *Computer Vision and Image Understanding*, pp. 385-397, No. 3, 2001.
- [4] Haritaoglu, D. Harwood, and L. Davis, "A real time system for detection and tracking people" *Journal of Image and Vision Computing*, 1999.
- [5] Haritaoglu, D. Hawood and L. Davis, "Who? When? Where? What? A Real Time System for Detecting and Tracking People" , *Automatic Face and Gesture Recognition*, pp. 222-227, 1998.
- [6] Blake, M. Isard, and D. Reynard, "Learning to track curves in motion of contours," *Artificial Intelligence*, pp.101-133, 1995.
- [7] Anderson, P. Burt, and G. van der Wal., "Change detection and tracking using pyramid transformation techniques," *In Proceedings of SPIE - Intelligent Robots and Computer Vision*, Vol 579, pp. 72-78, 1985.
- [8] J. Barron, D. Fleet, and S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, pp. 42-77, 1994.