퍼지 및 지능시스템학회 논문지 2005, Vol. 15, No. 7, pp. 852-857

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

*연세대학교 전기전자과, **군산대학교전자공학부 , ***한국전자통신연구원

요 약

본 논문은 지능형 로봇 시스템을 위한 인간 몸의 하이브리드 실루엣 추출 기법을 제안한다. 지능형 로봇은 내부적인 진동과 낮은 해상도로 인해 강인한 실루엣 추출을 필요로 한다. 이를 극복하기 위해서 본 논문에서는 하이브리드 실루엣 추출 기법을 제안하였다. 하이브리드 실루엣은 영상간의 공간차 및 시간차 정보를 고려 생성되며 움직임 영역 모델을 통해 두 정보간의 중요성에 가중치를 준다. 최종적으로 실험결과를 통해 제안된 기법의 우수성을 확인 하였다.

Abstract

This paper discusses a human body extraction method for intelligent robot system. 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. The temporal and gradient information is combined as hybrid silhouette. The motion region model is used to adjust combining parameters in hybrid silhouette. 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 spectrumof 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 sequentialimage 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 ro-

접수일자 : 2005년 9월 15일
완료일자 : 2005년 12월 5일
본 연구는 ETRI 지능로봇연구단의 URC 과제의 일부로 정보통신부로부터 지원 받았습니다.

bot 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.

It is known that the pose and motion of a human body can be determined by the potionand motion information of the joints [2]. How to obtain the position of joint from the sequential image is important to analysis human motion. However, most of the existing method does not consider restrictions occurring robot system. Therefore, it is need to develop a new method to extract key information from images restricted by robot system.

The objective of this work is to give position of the joint in the image as skeleton features. In this paper,we presented a method of skeleton feature extraction method from sequential images restricted by robot system. First, the hybrid silhouette generation method is considered to extract accurate silhouette which is robust to background motion. Then the adaptive skeleton model is applied to extract features of human body. The energy function is defined to find precise features in human body. Finally, color based hand detection method is used to compensate hand position of skeleton model.

In remainder of this paper, we discuss hybrid silhouette generation method in Section 2, and in Section 3 the adaptive skeleton model and energy functions are described. The color based hand detection method is introduced in Section 4. In Section 5 we illustrate the experimental results. Finally, we conclude our paper and give some remarks in Section 6.

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 adaptiveto 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 embeddedsystem such as robot system. However, these conventionalapproaches 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

$$I(x,y;t) = [I_r(x,y;t), I_g(x,y;t), I_b(x,y;t)]^T = [p(t)]^T. \tag{1}$$

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

$$I_t = \frac{\partial I}{\partial t}, \quad I_s = \left[\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y}\right]. \tag{2}$$

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,

$$\overline{I_t} = \|I_t\| \tag{3}$$

$$\overline{I_s} = \|I_s\| \tag{4}$$

In robot system, the temporal gradient $\overline{I_t}$ can have three kind of undesired values: 1) $\overline{I_t}$ can disappear because of camera problem. 2) $\overline{I_t}$ has insufficient silhouette information. 3) $\overline{I_t}$ 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

$$S^t = \int_{\forall p \in I} f(\overline{I_t}, \gamma_t) dp$$

$$f(\overline{I_t}, \gamma_t) = \begin{cases} 1, & \text{If } \overline{I_t} > \gamma_t \\ 0, & \text{else} \end{cases} \tag{5}$$

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

When S_t is zero, the old temporal gradient is 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

$$\eta \overline{I_s}(x,y) + (1-\eta)\overline{I_t}(x,y) \tag{6}$$

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 $R \subset \Box^2$. The initial motion region model has zero. Then the motion region mode is updated as following,

$$R(x,y;t+1)=\begin{cases} R(x,y;t)+\gamma_i, \overline{I_t}(x,y;t+1)>\gamma_t \\ R(x,y;t)+\gamma_d, \overline{I_t}(x,y;t+1)<\gamma_t \end{cases}$$
$$\gamma_i=\begin{cases} \overline{\gamma}_i, \overline{I_s}(x,y;t+1)>\gamma_s \\ \underline{\gamma}_i, \overline{I_s}(x,y;t+1)<\gamma_s \end{cases}$$
$$\gamma_d=\begin{cases} \overline{\gamma}_d, \overline{I_s}(x,y;t+1)>\gamma_s \\ \underline{\gamma}_d, \overline{I_s}(x,y;t+1)<\gamma_s \end{cases} \quad (7)$$

where γ_s is thebinarization parameter of spatial image. $\overline{\gamma}_i$ and $\underline{\gamma}_i$ are upper and lower increasing parameters. $\overline{\gamma}_d$ 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

$$\overline{I_b}(x,y)=\eta\overline{I_s}(x,y)+(1-\eta)\overline{I_t}(x,y)$$
$$\eta=\begin{cases} \eta_R, R(x,y)>\gamma_R \\ \eta_B, R(,xy)<\gamma_R \end{cases} \quad (8)$$

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.

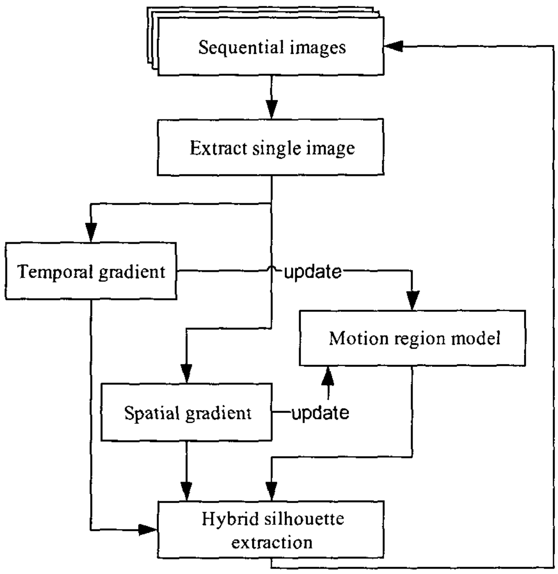


Figure 1. hybrid silhouette extraction process

2.2 Choices of parameters

There exist many parameters in hybrid silhouette extraction method. Unfortunately, these parameters are not easy to be determined via automatic or intelligent method. Therefore, we provide a guideline on appropriate relative choices of parameters. The upper and lower increasing parameters have the following constraint

$$\overline{\gamma}_i > \underline{\gamma}_i > 0. \quad (9)$$

Similarly, the upper and lower decreasing parameters have the following constraint

$$0 < \overline{\gamma}_d < \underline{\gamma}_d. \quad (10)$$

The increasing and decreasing parameters have to comply to the following constraint

$$|\overline{\gamma}_i| > \alpha|\overline{\gamma}_d|$$
$$|\underline{\gamma}_i| < |\underline{\gamma}_d| \quad (11)$$

where α means temporal memory length. When α is two, the motion region model have recent two past motion information. The convex sum parameter η_R and η_B have following constraint

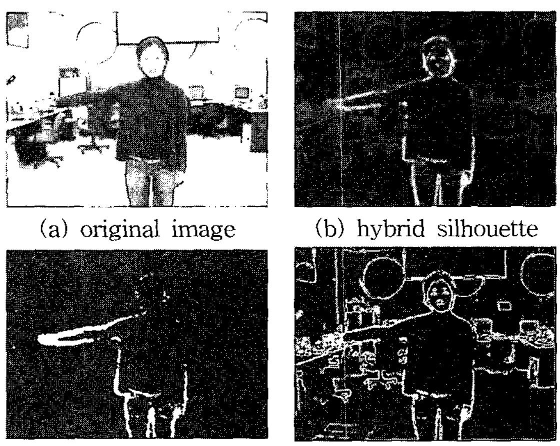
$$\eta_R > 0.5 > \eta_B. \quad (12)$$

Large $\overline{\gamma}_i$ values lead to make spatial gradient complete insufficient silhouette whereas small $\overline{\gamma}_i$ values yield less complete silhouette. However, large $\overline{\gamma}_i$ values also yield redundant silhouette information.

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 thee 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.

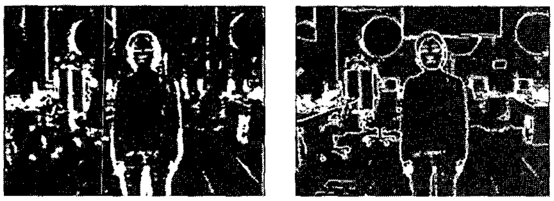


(a) original image (b) hybrid silhouette

(c)temporal gradient (d) spatial gradient

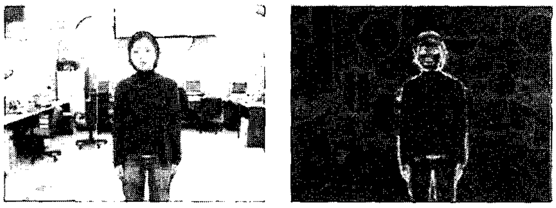
Figure 2. Hybrid silhouette extraction: $\overline{\ }$ has insufficient silhouette

(a) original image (b) hybrid silhouette

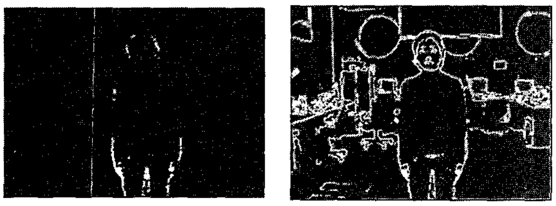


(c)temporal gradient (d) spatial gradient

Figure 3. Hybrid silhouette extraction: $\overline{I_t}$ has unnecessary motion information



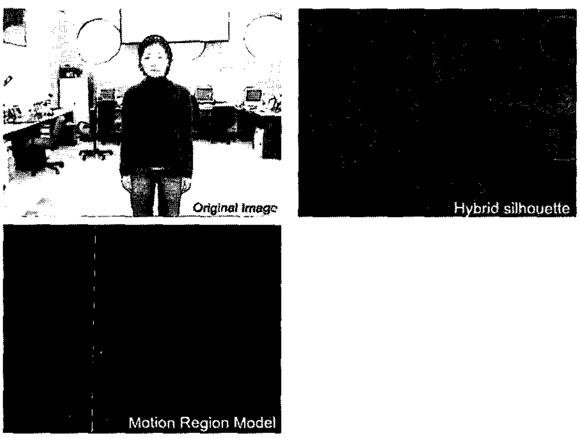
(a) original image (b) hybrid silhouette



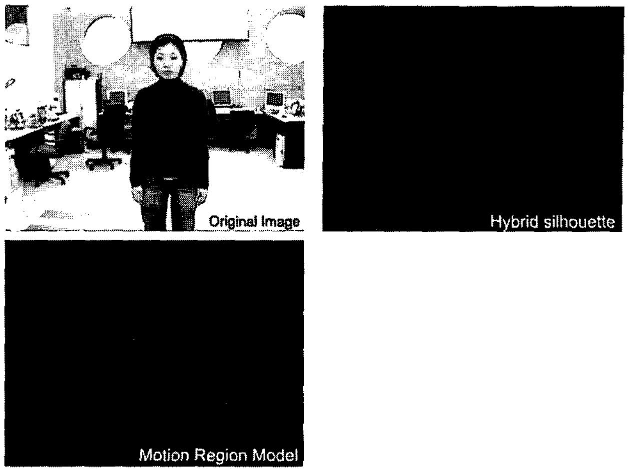
(c)temporal gradient (d) spatial gradient

Figure 4. Hybrid silhouette extraction: $\overline{I_t}$ is disappeared

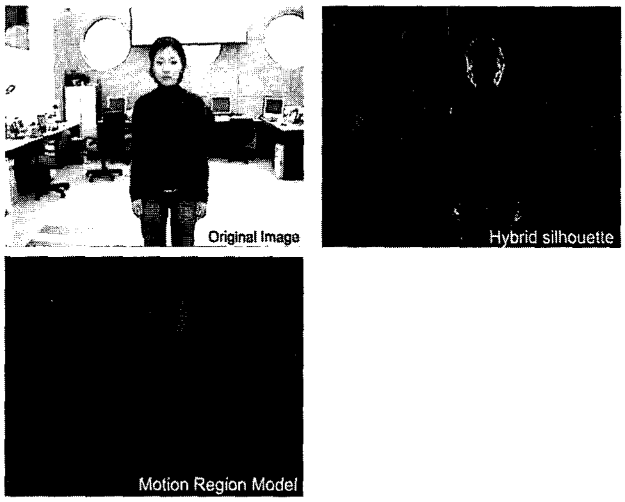
Figure 5 shows Transition of motion region and hybrid silhouette. Figure 5 (a) is initial state of hybrid and motion region model. Figure (b)-(g) show the change of motion region model. we could confirm



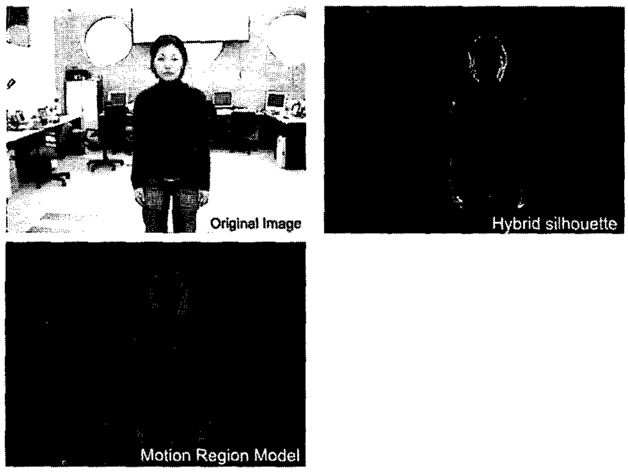
(a) Hybrid silhouette of frame 1



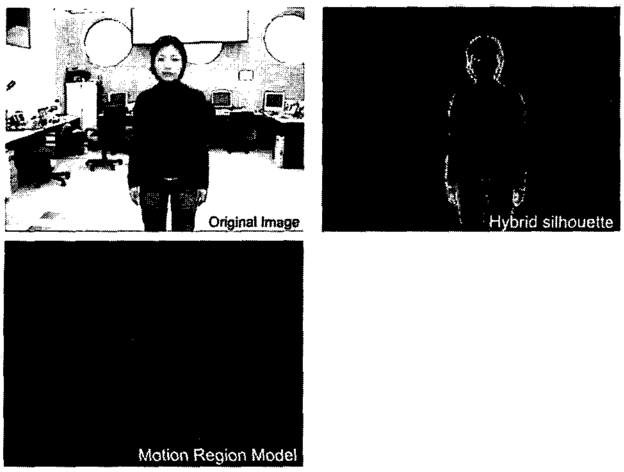
(b) Hybrid silhouette of frame 2



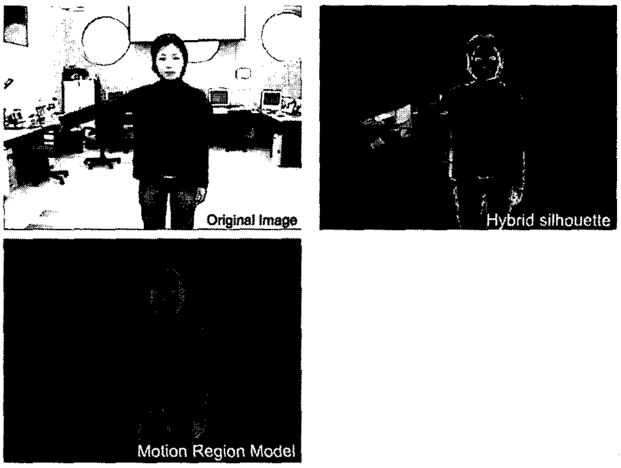
(c) Hybrid silhouette of frame 3



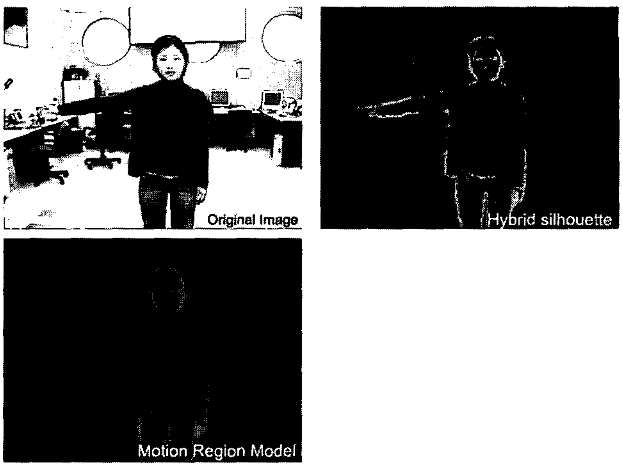
(d) Hybrid silhouette of frame 4



(e) Hybrid silhouette of frame 5



(f) Hybrid silhouette of frame 6



(g) Hybrid silhouette of frame 7

Figure 5. Changes of Hybrid silhouette and motion region model.

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 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.

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] I. 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] I. Haritaoglu, D. Harwood, and L.'Davis, "A real time system for detection and tracking people" Journal of Image and Vision Computing, 1999.

[5] I. Haritaoglu, D. Hawood and.L. Davis, "Who? When? Where? What? A Real Time System for Detecting and Tracking People", Automatic Face and Gesutre Recognition, pp. 222-227, 1998.

[6] A. Blake, M. Isard, and D. Reynard, "Learning to track curves in motion of contours," Artificial Intelligence, pp.101-133, 1995.

[7] C. 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.

저 자 소 개

김문환(Moon Hwan Kim)
2004년 제 14권 제 5호 참조

주영훈(Young hoon Joo)
2005년 제 15권 제 1호 참조

박진배(Jin Bae Park)
2004년 제 15권 제 1호 참조

지수영(Su Young Chi)
ETRI 지능형로봇연구단 팀장

조영조(Young-Jo Cho)
2004년 제 14권 제 5호 참조
ETRI 지능형로봇연구단 단장

김혜진(Hye Jin Kim)
ETRI 지능형로봇연구단 HRI 팀