Human Body Tracking Based on Discrete Wavelet Transform

SHYANG-LI CHANG[#], CHEN-CHIEN HSU[*], TSUNG-CHI LU[#], TI-HO WANG[#]
[*]Department of Electrical Engineering, Tamkang University
151 Ying-Chuan Rd., Tamsui, Taipei County, 25137
TAIWAN

[#]Department of Electronic Engineering, St. John's University,
499 Tam King Rd., Sec. 4, Tam-Sui, Taipei County, 25135
TAIWAN

Abstract: - A novel human body tracking system based on discrete wavelet transform is proposed in this paper based on color and spatial information. The configuration of the proposed tracking system is very simple, consisting of a CCD camera mounted on a rotary platform for tracking moving objects. By using the position information of objects in the image frame captured by the camera, the rotary platform is controlled to position the tracking object around the central area of images to improve tracking efficiency. Thanks to the use of discrete wavelet transform, computations can be significantly reduced while achieving real-time tracking.

Keywords: - Object tracking, human-body tracking, monitoring, discrete wavelet transform, image processing, colour image.

1 Introduction

Image tracking is getting more and more popular over the past years because of the advancements of automation technologies. It has been widely applied in surveillance systems, robot localization, human-computer interaction, etc. In the early stage, very simple tracking algorithms were only available because of the constraints of computer speed. For example, objects were generally affixed with a specific geometric shape or color, based on which tracking is performed [1]. Thanks to recent improvements on computation capabilities, many complex tracking algorithms have been developed with better efficiency for implementation of image analysis and tracking systems. For example, position information of tracking objects in an image frame has been used as feedback signals to control robotic arms [2].

Among the researches of object tracking [3]-[9], human-body tracking is the most attractive one although difficulties still exist because the shapes and dynamics of humans are complicated and the backgrounds are cluttered. Over the past years, many applications of people tracking systems such as surveillance, human-computer interface, people-counting system, etc, have been attempted based on a popular method of background subtraction to segment and track moving objects in real-time surveillance. For example, a W4 surveillance system [10], which detected and tracked people using background subtraction and shape analysis to achieve real-time performance on generally used PC platform was proposed. Segmentation methods using background subtraction, however, have difficulties in image sequences from moving camera or sequences including instantaneous change of illumination or shadow. Jang [11] proposed a model-based tracking method using Kalman filters to predict and identify objects under a simple background. In complex environments, targets may disappear totally or partially due to occlusion by other objects. To solve this problem, a structural Kalman filter was proposed [12] for preventing tracking objects being occluded. Though worked satisfactorily to some extents in providing better identification rate, the method, however, has difficulties in achieving real-time tracking because of demanding computations required.

To overcome difficulties in achieving real-time tracking and improving tracking efficiency, a novel colour image real-time human body tracking system based on discrete wavelet transform is proposed in this paper for identifying targets based on their color and spatial information. The configuration of the proposed tracking system includes a CCD camera mounted on a rotary platform for tracking moving objects. By using the position information of objects in the image frame captured by the camera, the rotary platform is controlled to position the tracking object around the central area of images to improve tracking efficiency. Thanks to the use of discrete wavelet transform, computations, which are critical in real-time tracking, can be significantly reduced. Existing researches of human-body tracking were generally based on greyscale images, ignoring the facts that colours of the clothes that humans wear are salient features for identification purpose. As a result, we will incorporate colour information to establish important features for identifying the human body in the proposed approach.

The rest of this paper is organized as follows. In Section 2, preliminary backgrounds, including coordinate systems and wavelet transform, are introduced. The proposed human body tracking system is described in Section 3. Experimental results of the proposed algorithms are presented in Section 4. Finally, conclusions are drawn in Section 5.

2 Preliminaries

In this section, colour coordinate systems and wavelet transform will be introduced. How these techniques are incorporated into the proposed tracking system is also discussed.

2.1 Color coordinate systems

There are several color coordinate systems, each having its particular merits, which are being used in color image processing. The human visual system can be modelled as using three perceptual attributes: luminance, hue, and saturation to distinguish one color from another. Many other image processing tasks, such as enhancement and restoration, require that only the luminance component be processed [16]-[17], [19]. The RGB (red-green-blue) coordinate system is commonly used for representing digital color images. However, color coordinate systems related to human visual system's perceptual attributes (luminance, hue, and saturation) are often useful for processing color images. Although researches have been conducted toward the development of color measurement techniques, there is no color coordinate system that is universally accepted as full compliance to human perception.

The YIQ color coordinate system adopted in color television broadcasting is commonly used for color image processing. The Y value represents the luminance, and the I and Q values jointly describe the hue and saturation. In the YIQ system, one could define hue and saturation as follows:

$$\text{Hue} = \tan^{-1}(\frac{Q}{I}) \quad (1)$$

$$\text{Saturation} = \sqrt{I^2 + Q^2} \quad (2)$$

In a tracking system, variation of luminance information might result in the change of color information, which generally causes failures in tracking objects. As a result, we should always minimize the impact on the luminance due to the change of color information to prevent tracking failures. Because luminance (Y) and color information (I, Q) are independently represented in the YIQ color coordinate system, it is therefore a preferred system for use in tracking systems and will be adopted in this paper. However, color images extracted by an image extraction component are generally represented in the RGB color coordinate system. Transformation from RGB to YIQ is therefore essential, which can be performed via a simple, linear conversion as follows:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (3)$$

2.2 Discrete wavelet transform

Recently, wavelet transform has been extensively applied in the areas of image processing, image compression, edge detection, and texture analysis [13]-[18]. The realization of wavelet transform is straightforward, which requires only simple filters. A 2-dimensional wavelet transform decomposes an image into 4 sub-images as shown in Fig. 1, where filters are first applied in one dimension (e.g. X axis) and then in the other (e.g. Y axis). Because down-sampling is performed at these two stages, the size of the sub-images becomes 1/4 as large as the original image. Observing these four sub-images in Fig. 2, we found that the wavelet transform preserves not only the frequency features but also spatial ones. For an original image, it can be decomposed into four different bands (LL, HL, LH, HH) via the discrete wavelet transform. These sub-bands contain different frequency characteristics with the use of high-pass and low-pass filters. The high-pass filter extracts the high-frequency portions (e.g. edges of the object). On the other hand, the low-pass filter gives the low-frequency information representing the most energy of an image and rejects the noise of an image as well. The basic idea is to use the wavelet transform to reduce the resolution of each frame of the sequence for reducing the computational cost. Because of implementation consideration, the Haar basis wavelet transform will be used in this paper due to its simplicity and speed efficiency, where only the low-frequency part is used for processing due to the consideration of low computing cost and noise reduction issue. The original image of 240×320 is pre-processed via a 2-level discrete wavelet transform to obtain a lowest-frequency sub-image (i.e. LL2 in Fig. 2(d)) for further processing in the proposed tracking system. As a result, the image size of the sub-image LL2 has been reduced to 60×80, which represents 1/6 of the size of the original image.

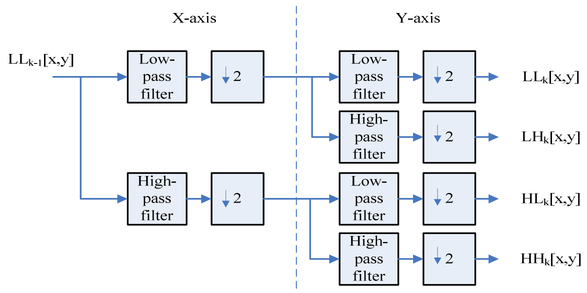


Fig. 1 Two-dimensional discrete wavelet transform.

3. The proposed human body tracking system

The objective of tracking is to closely follow objects in each frame of a video stream such that the object position as well as other information is always known. To overcome difficulties in achieving real-time tracking and improving tracking efficiency, a novel colour-image real-time human body tracking system based on discrete wavelet transform is proposed in this paper, where a CCD camera is mounted on a rotary platform for tracking moving objects. Procedures in tracking moving objects via the proposed approach can be illustrated via a flowchart shown in Fig. 3. Assume color images captured by a CCD camera having a size of 240×320 are sent to a computer for further processing via an image acquisition card. A 2-level discrete wavelet transform for the images is performed, where the lowest-frequency sub-image (i.e. LL2 in Fig. 2(d)) is only used for subsequent processing.

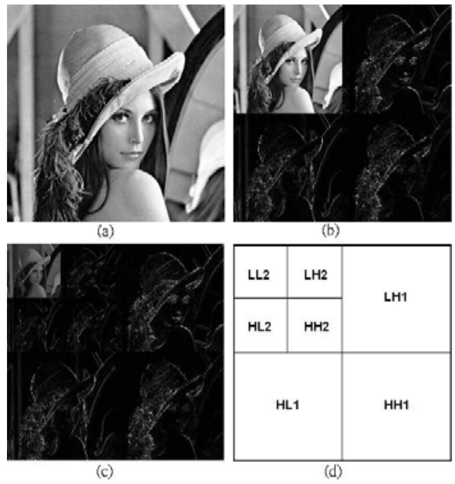


Fig. 2 (a) Original image (b) first-level DWT (c) second-level DWT (d) sub-bands of second-level DWT.

As far as moving object detection is concerned, we need to detect objects entering into the image frame first and subsequently identify the coordinates of the object. There are many approaches for detecting moving objects. Among them, the background subtraction is the most popular and efficient method to detect moving objects in real-time surveillance. To begin with, an image without moving objects is assumed as the background image. When a target object moves into the image frame, there will be difference in terms of luminance values between the image with the object and the background image. By subtracting these two images, the object can be detected. For this reason, the Y component (luminance) of the image is only required. Note that background images, which might slightly vary from time to time, are critical when using the adopted background subtraction method. As a result, the luminance value of background images might subject to variation, which is not desired. To reduce this adverse effect of background variation, we adopt a dynamic background to obtain the background image for using in the proposed approach, where 8 consecutive image frames are averaged to form the background image when there is no object entering the image. That is,

$$B[x,y] = \frac{1}{8}\sum_{k=1}^{8} Y_{n-k}[x,y] \qquad (4)$$

, where $B[x,y]$ represents the greyscale value of the position x,y in the coordinate system, and $Y_n[x,y]$ is the greyscale value of the Y component of the position x,y in the nth image frame. Difference image can be obtained by subtracting the background image from the current image as:

$$D[x,y] = \left|Y_n[x,y] - B[x,y]\right| \tag{5}$$

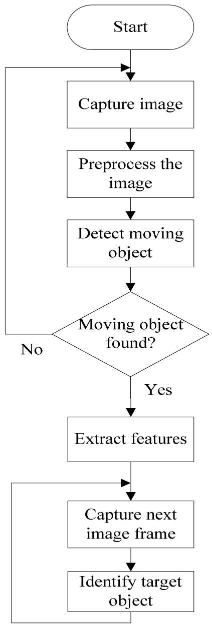


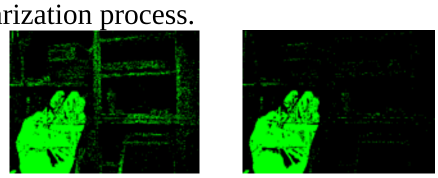
Fig. 3 Flowchart of the tracking procedures.

3.1 Moving object detection

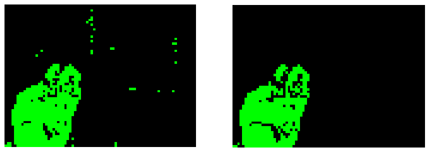
Based on the difference image obtained, we can detect the object entering the image. To facilitate the detection process, we further threshold the difference image by a value *T* via a binarrization process. If the pixel value in the difference image is greater than the threshold value *T*, it is then set to 255, indicating that it belongs to the object. Otherwise, the pixel value is set to 0.

$$D[x,y] = \begin{cases} 0 & \text{, if } D[x,y] < T \\ 255, & \text{if } D[x,y] \geq T \end{cases} \tag{6}$$

Care should be taken in choosing the threshold value *T*. If *T* is too large, pixels belonging to the object might not be counted. On the other hand, noises might be occurring if *T* is too small. Figure 4 shows images via different threshold values *T* with and without wavelet transform. It is clear that noises are significantly reduced in the images if discrete wavelet transform is performed, where high-frequency components are filtered out as shown in (c) and (d). A larger threshold value *T*=30 also generates better results in eliminating noises during the binarization process.



(a) (b)



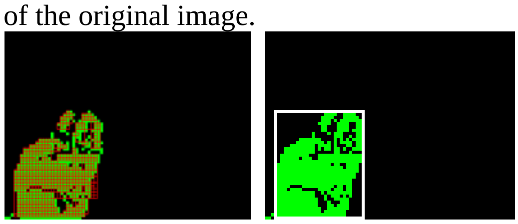
(c) (d)

Fig. 4 Images via different threshold values with and without wavelet transform; (a)(b) without wavelet transform; (c)(d) via 2-level DWT; (a)(c) *T* = 20; (b)(d) *T*=30.

3.2 Bounding box for the moving object

By scanning the image after the binarization process as shown in Fig. 4(d), we shall detect the object. If only one moving object is assumed during the detection, the coordinate projection method is the most simple and effective way. However, this is not the case in general, and we have to assume that more than one object might be entering the image frame. As a result, procedures to generate bounding boxes for the moving object need to be created for clear positioning of the object.

After the completion of the process via the pseudo codes, there are pluralities of tiny boxes which are overlapped each other for approximating the object as shown in Fig. 5a. By eliminating the overlapped portions of these tiny boxes, a rectangular frame encompassing the object can be obtained as shown Fig. 5b. Note that objets entering into the image frame might not be unique. One of the methods to single out the target object is to use the area as a criterion. If an object having an area smaller than T_a, it is regarded as a noise, where T_a is an experimental value set to A/100 here and A is the area of the original image.



(a) (b)

Fig. 5 Bounding boxes and rectangular frame surrounding an object.

3.3 Feature extraction and object identification

When the coordinates of the target object become available, features associated with the object need to be extracted for comparison during the tracking. Traditionally, methodologies for object tracking can be classified into two types: template-based and contour-based approaches. Template matching [20]-[22] is the most straightforward and useful method to search for the target in an image, which can be easily implemented. Generally, the target of interest is chosen as a template in advance, and the most similar block with respect to the template is then located from an image. On the other hand, contour-based methods [23]-[24] consider the outline of the target as image information instead of the content of the target in template-based methods. We will use template-based method to meet the requirements of real-time tracking.

4 Experimental results

We implemented the proposed tracking system in Windows XP PC with Pentium 2.0G CPU, 1024MB RAM under Borland C++ builder 5 software environment as the implementation platform. The resolution of each color images is 320x240 pixels. We can achieve real-time processing at about 25 frames per second. Figure 6 shows the experimental results in tracking a walking man via the proposed approach. When the man enters the center of scene, the tracking system treat the man as the target object and uses it's positional information to actuate the motor of the rotary platform to keep the moving object in the center of scene. As demonstrated in Fig. 6, satisfactory performances have been achieved via the proposed approach.

5 Conclusions

With the aim at single human-body tracking, a novel colour image real-time human body tracking system based on discrete wavelet transform is proposed in this paper for identifying the target based on color and spatial information. To improve tracking performances, discrete wavelet transform is used to pre-process the image for reducing computations required and achieving real-time tracking. Thanks to the setup of the CCD camera which is mounted on a rotary platform, position information of the target object in the image can be used to control the movement of the platform for locking the tracking object around the certral area of the image frame via the proposed approach. Therefore, tracking performance is significantly improved. The experiments results have shown that the proposed tracking system is capable of real-time tracking human objects in about 25 frames per second.

Acknowledgements:
This work was partially supported by the National Science Council, Taiwan, under Grant NSC NSC-94-2218-E-129-002.

Reference:

[1] J. Feddema and O. Mitchell, "Vision-guided servoing with feature-based trajectory generation," *IEEE Trans. on Robotics and Automation*, Vol. 5, PP. 691-700, 1989 .

[2] B. Barmish, *New Tools for Robustness of Linear Systems*, Macmillan Publihing Company, New York, 1994.

[3] H. Kang, D. Kim, and S. Y. Bang, "Real-time Multiple People Tracking Using Competitive Condensation," *Proc. of 28th IEEE Conference on Decision and Control*, Vol. 1, pp. 413-416, Aug. 2002.

[4] D. Ramanan, D. Forsyth, and D. Zisserman, "Tracking people and recognizing their activities," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005)*, Vol. 2, pp. 20-25, June 2005.

[5] I. A. K. Saeed and N. V. Afzulpurkar, "Real Time, Dynamic Target Tracking using Image Motion," *Proc. of the 2005 International Conference on Mechatronics*, pp. 241-246 , Taipei, Taiwan, July 2005.

[6] G. Zhu, Q. Zeng, and C. Wang, "Efficient edge-based object tracking," *Pattern Recognition*, Vol. 39, pp. 2223-2226, 2006.

[7] H. Ning, T. Tan, L. Wang, and W. Hu, "Kinematics-based tracking of human walking in monocular video sequences," *Image and Vision Computing*, Vol. 22, pp. 429-441, May 2004.

[8] F. H. Cheng and Y.L. Chen, "Real time multiple objects tracking and identification based on discrete wavelet transform," *Pattern Recognition*, Vol. 39, pp. 1126-1139, 2006.

[9] M. Amiri, H.R. Rabiee, F. Behazin, and M. Khansari, "A new wavelet domain block matching algorithm for real-time object tracking," *Proc. of International Conference on Image Processing*, Vol. 3, pp. 961-964, Sept. 2003.

[10] I. Haritaoglu, D. Harwood, and L. Davis, "W4: Real-time surveillance of people and their activities," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 22, NO. 8, pp. 809-830, August 2000.

[11] D.S. Jang, G.Y. Kim, and H.I Choi, "Model-based tracking of moving object," *Pattern Recognition*, Vol. 30, pp. 999-1008, 1997.

[12] D.S. Jang, S.W. Jang, and H.I. Choi, "2D human body tracking with Structural Kalman filter," *Pattern Recognition*, Vol. 35, pp. 2041-2049, 2002.

[13] A.K. Liu, W.Y. Tseng, and S.Y. Chang, "Wavelet Analysis of AVHRR Images for Feature Tracking," *Geoscience and Remote*

Sensing Symposium, Vol. 1, pp. 85-87, May 1996.

[14] A.K. Liu, S.Y. Wu, and W.Y. Tseng, "Wavelet analysis of Satellite Images for Coastal Monitoring," *IEEE Journal of Oceanic Engineering*, Vol. 22, pp. 9-17, Jan. 1997.

[15] T. C. Wang and N. B. Karayiannis, "Detection of microcalcifications in digital mammograms using wavelets," *IEEE Trans. On Medical Imaging*, Vol. 17, No. 4, pp. 498-509, Aug. 1998.

[16] R.C. Gonzalez and R.E. Woods, *Digital Image Processing*, 2nd edition, Addison Wesley, 2002.

[17] K. R. Castelman, *Digital Image processing*, Prentice-Hall, 1996.

[18] A. Arerbuch, D. Lazar, and M. Israeli, "Image Compression Using Wavelet Transform and Multiresolution Decomposition," *IEEE Trans. Image Processing*, Vol. 5, No. 1, pp.4-15, 1988.

[19] R. S. Ledley, M. Buas, and T. J. Golab, "Fundamentals of true-color image processing," *Proc. Intl. Conf. on Pattern Recognition*, Vol. 1, pp. 791-795, 1990.

[20] R. Brunelli and T. Poggio, "Template Matching: Matched Spatial Filters and Beyond," *MIT AI Memo* 1549, July 1995.

[21] N.Sube and T.S. Huang, "Improving visual matching," *IEEE Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 58-65, 1997.

[22] J. Lewis "Fast Normalized Cross-Correlation," *Vision Interface*, pp. 120-123, 1995.

[23] S. Birchfield, "An Elliptical Head Tracker," *IEEE Conf. on Signals, Systems, and Computers*, Vol. 2, pp. 1710-1714, 1997.

[24] S. Birchfield, "Elliptical Head Tracking Using Intensity Gradiants and Color Histograms," *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 232-237, 1998

Fig. 6 Image sequences in tracking a walking man (every 25^{th} frame is shown).