Object Detection and Tracking in Contourlet Domain

Nguyen Thanh Binh and Tran Anh Dien

Faculty of Computer Science and Engineering,
Ho Chi Minh City University of Technology, Vietnam
ntbinh@cse.hcmut.edu.vn, dientrananh@gmail.com

Abstract. This paper describes a method for the moving object detection and tracking in video sequences using contourlet transform. For the contourlet transform to be translation-invariant a 2D cycle spinning is implemented on subbands Δ_1 and Δ_2. Cycle spinning for edge detection is implemented. The shape of object may change from this frame to other frame. The 3D moving object is combined two parts: a 2D shape change and 2D motion. The 2D motion of the object, we use the minimum Hausdorff distance from the model to the image to find where object moved to. With 2D shape change of the object, we use distance from the image to the transformed model to select set of image pixels of the next model. For performance evaluation, we compared the proposed method based on the contourlet transform using cycle spinning with the similar methods based on the complex wavelet transform and wavelet transform.

Keywords: edge detection, contourlet tranform, cycle spinning, object tracking.

1 Introduction

Wavelet transforms domain methods are used for object tracking. Several algorithms have been proposed to solve the problem of tracking [3, 4]. Every tracking method requires an object detection mechanism either when the object first appears in the video. A common for object detection is to use information of single frame.

Y. Wang [17] proposed an algorithm that derives the objects based on the motion between frames. This tracking is limited and not able to handle some complex situations such as object starts moving, object stops moving and objects move together. A lot of existing methods first perform computationally expensive spatial segmentation based on a moving object region tracking [15]. This is not necessary in a lot of applications, where only moving objects need to be tracked.

The Discrete wavelet transform (DWT) provides a fast, local, sparse and decorrelated multiresolution analysis of images. DWT have limited such as shift-sensitivity and poor directionality [9]. Several researchers have provided solutions for minimizing these disadvantages. Recently new X-let multiscale transforms have been developed such as curvelet, contourlet [1,2,5,6,7,13,14] which integrate the concept of directionality in better way. Techniques to reduce these drawbacks have been proposed; new multiscale transforms have been designed such as the contourlet transform to integrate the concept of directionality in a more useful fashion.

P.C. Vinh et al. (Eds.): ICCASA 2012, LNICST 109, pp. 192–200, 2013.

We propose an implementation of the contourlet transform and described a method for moving object tracking, using contourlet transform with cycle spinning. The shape of object may change from this frame to other frame. The approach consists of two steps. First contourlet coefficients are used for detection of object and second tracking of object in the sequence of frames. The rest of the paper is organized as follows: In section 2, we described the basic concepts of contourlet transform. Details of the proposed algorithm are given in section 3. In section 4, the results of the proposed method for detection, tracking are shown and compared to other methods and finally in section 5, we presented our conclusions.

2 The Contourlet Transform

The contourlet construction presented here is based on the work of Do and Vetterli [1]. Contourlets constitute a new family of frames that are designed to represent smooth contours in different directions of an image. Contourlet is easily applied in image processing because its representation is a fixed transform. Contourlet allows for a different number of directions at each scale and aspect ratios. This feature allows an efficient contourlet-based approximation of a smooth contour at multiple resolutions. The discrete contourlet transform is a multiscale and directional decomposition using a combination of Laplacian pyramid (LP) and directional filter bank (DFB) [1].

The idea of the contourlet construction [1] is: let $a_0[n]$ be the input image, the output after the LP step is I bandpass images $b_i[n]$, i =1, 2,..., I and a lowpass image $a_I[n]$. Each bandpass image $b_i[n]$ is decomposed by an ℓ_i -level DFB into $2^{\ell i}$ bandpass directional images $c_{i,k}^{(\ell_i)}[n]$, for $k = 0, 1,\ldots, 2^{\ell i}-1$.

In the discrete contourlet transform, the multiscale and directional decomposition steps are decoupled. So, we have different numbers of directions at different scales.

Contourlet decomposition proceeds through two main steps: first, Laplacian pyramid multiscale decomposition is performed; then directional filter bank decomposition is used to link point discontinuity to linear structures. In more detail, an image is decomposed into a low pass image and bandpass images by the LP decomposition. Each bandpass output is further decomposed by the DFB step. The output of the DFB step consists of smooth contours and directional edges. In this paper, each directional subband at each level consists of 2^n element, where n is a positive integer. Fig 1 shows a contourlet decomposition.

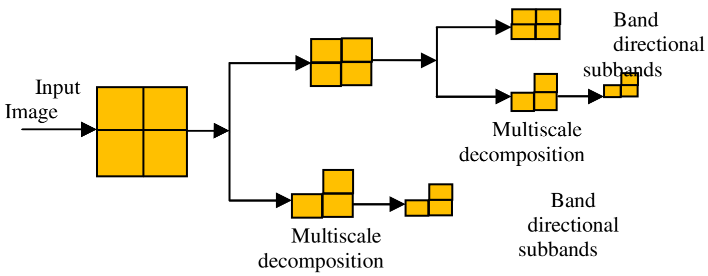


Fig. 1. Contourlet decomposition

3 The Proposed Method

In this section, we present an approach to implement of the contourlet transform and described a method for moving object tracking, using contourlet transform. The shape of object may change from this frame to other frame. The approach consists of two steps. First contourlet coefficients are used for detection of object and second tracking of object in the sequence of frames. A video sequence contains a series of frames. Each frame can be considered as an image. If an algorithm can track moving objects between two digital images, it should be able to track moving objects in a video sequence. The algorithm starts with two steps.

3.1 The Detection of Object Algorithm

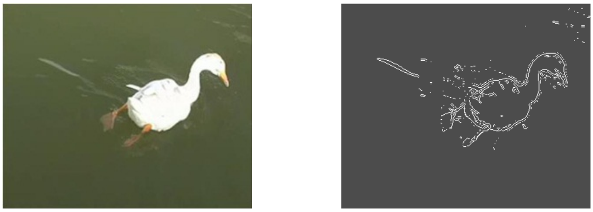
The goal of edge detection is to divide the given image into regions that belong to distinct objects in the depicted scene. This process consists of three steps [1,2,13] apply contourlet transform to the image, retain the coefficients where the signal-to-noise ratio is high, and reduce the coefficients where the signal-to- noise ratio is low.

A 2-D cycle spinning is implemented on subbands Δ_1 and Δ_2 so that the contourlet transform is translation-invariant. The transformed data are shifted, edge detected, and unshifted. Applying the following procedure,

$$\hat{s} = \frac{1}{K_1 K_2} \sum_{i=1, j=1}^{K_1, K_2} S_{-i,-j}(T^{-1}(\theta[T(S_{i,j}(x))])) \qquad (1)$$

where K_1 and K_2 are the maximum number of shifts, we expect an improvement for the estimation $\hat{s}$ compared to the image without cycle spinning. For the contourlet transform, if the input image has size $N \times N$, with $N = 2^K$, then after K shifts in each direction, the output becomes repetitive and so the maximum numbers of shifts will be K in each direction. If one decomposes an image of size $N \times N$ using the contourlet transform, then the number of decomposition levels in the Δ_1 stage will be at most K, and therefore, the maximum number of shifts is K in the row and column directions.

The threshold for the contourlet coefficients can be calculated using statistical properties of noise or blur. After thresholding the contourlet coefficients, the image can be reconstructed. Donoho and Johnstone [10] have described a threshold that depends on standard deviation of contourlet coefficients. Here, to compute the threshold value for image edge detection, we use a combination of three parameters: the contrast ratio (ratio between standard deviation and mean of contourlet coefficients), the absolute median of contourlet coefficients, and a level dependent parameter. Fig. 2 shows the detection in duck images with the method presented above.



(a) Original image. (b) Detected image

Fig. 2. Duck image

3.2 The Tracking Algorithm

In all the computations, the frame rate is adequate and the shape change from one time frame to the next is required to be small. The 3D moving object is combined two parts: a 2D shape change and 2D motion. To change 2D shape in image frame, we use 2D geometric models to capture them. With 2D motion of the object, we use the minimum Hausdorff distance [11] from the model to the image to find where object moved to.

Hausdorff distance [11] defined distance between two point sets P and Q as

$$H(P, Q)=\max (h(P, Q), h(Q, P)) \tag{2}$$

where $h(P,Q) = \max_{p\in P} \min_{q\in Q} \|p-q\|$ and $\|.\|$ denotes Euclidean distance.

Daniel[11] defines the partial distance as:

$$h_k(P,Q) = \underset{p\in P, q\in Q}{K^{th}} \min \|p-q\|$$

For some transformation group G [11], the natural definition of a distance is simply the minimum with respect to that group action $D_G(P,Q) = \min_{g\in G} H(g(P),Q)$, if $D_G(P,Q)$=0, two shapes are same otherwise change in shape is measured.

With 2D shape change of the object, we use distance from the image to the transformed model to select set of image pixels of the next model. The model M_t consist of m point moved to next time frame I_{t+1}. The location of the object in the new image frame is computed by the minimum valued d of distance from M_t to the frame I_{t+1}

$$d = \min_{g\in G} h_k(g(M_t), I_{t+1}) = \min_{g\in G} \underset{p\in M_t}{K^{th}} \min_{q\in I_{t+1}} \|g(p)-q\|$$

which d identifies the transformation $g^*\in G$ of M_t which minimizes the rank order.

4 Experiments and Results

We have applied the procedure described in section 3 and observed good performance in our detection and tracking experiments as briefly demonstrated in this section.

Hard thresholding is applied to the coefficients after decomposition in contourlet domain. In case of contourlet transform with cycle spinning, the input image is of size $N \times N$, with $N = 2^K$. In the case $K = 8$ (i.e., $N = 256$) after 8 shifts in each direction, the transform output repeats and so the maximum number of shifts will be 8 in each direction. We apply the same approach to the contourlet transform.

For the tracking part, we determine the object in each frame of the movie. The object area is determined in the first frame by hand. In this experiment, we use mouse to select the object area in the first frame. The object area is determined in the first frame, the tracking algorithm need to track the object from frame to frame.

For performance evaluation, we compared the proposed method based on the contourlet transform with the similar methods based on the complex wavelet transform and wavelet transform. The comparison of results with other methods has been done using our program on the same video and at similar scales. Here, we report the results on some video clips. Our experimental approach was as follows.

Our experiment is on car and player video clips with frame size 288 by 352. The proposed method processes this video clip at 28 frames/second. We have experimented on the video up to 1000 frames. Here, we report the results upto 1000 frames and starting from frame number: 100, 200, 300, 400, 500, 600, 700, 800, 900 and 1000. Some results achieved as shown in fig. 3 and fig 4.

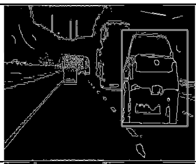
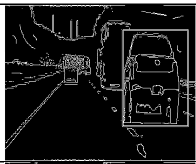
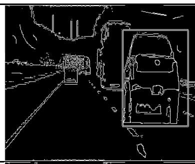
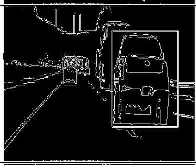
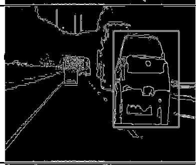
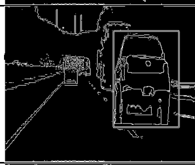
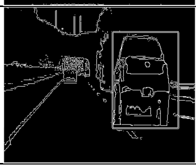
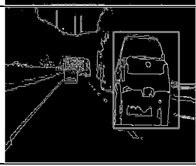
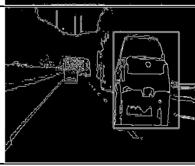
Frame No	Wavelet Domain	Complex wavelet Domain	Proposed method
100			
200			
300			

Fig. 3. Tracking in Car video clips upto 1000 frames

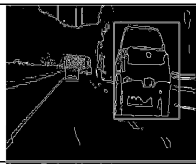
400			
500			
600			
700			
800			
900			
1000			

Fig. 3. (*Continued*)

Frame No	Wavelet Domain	Complex wavelet Domain	Proposed method
100			

Fig. 4. Tracking in Play video clips upto 1000 frames

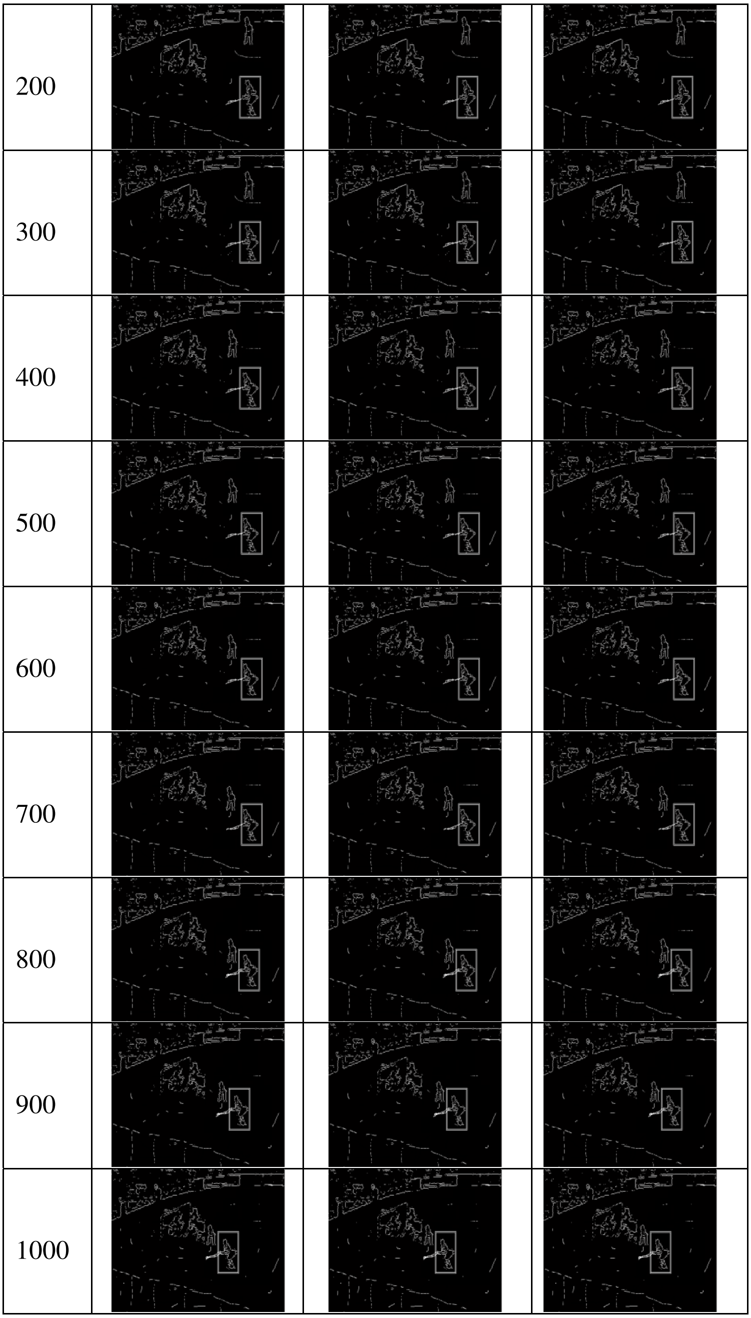


Fig. 4. (*Continued*)

In these figures, we observe that the proposed method performs better than the other two methods. Other experiments also show that the proposed method works well and better than the other ones. In wavelet domain, tracking is high execution time. In complex wavelet domain, tracking is more accurate but is also medium execution time. The proposed method is also more accurate and faster other method. As above mentioned, contourlet allows for a different number of directions at each scale and aspect ratios. This feature allows an efficient contourlet-based approximation of a smooth contour at multiple resolutions.

5 Conclusions

In this paper we have constructed a method for detection and tracking of object that provides a sparse expansion for typical images having smooth contours. We use contourlet coefficients and cycle spinning to detection and tracking the object in the sequence of frames. We compared the proposed method based on the contourlet transform with the similar methods based on the complex wavelet transform, and wavelet transform. The proposed method is more accurate and faster other method. However, if the quality of the frame in video is very bad, such as strong noise, blur, etc., then the estimation ability is reduced. To avoid this problem, we should try to reduce noise and blur before the application of the proposed algorithm. Unlike the other methods, the algorithm does not rely upon many properties of object such as size, shapes, colors, etc…. In all of computations, it has been assumed that the frame rate is adequate and the size of the object may be change between adjacent frames.

References

1. Do, M.N., Vetterli, M.: The Contourlet Transform: An Efficient Directional Multiresolution Image Representation. IEEE Transactions on Image Processing 14, 2091–2106 (2005)
2. Eslami, R., Radha, H.: Translation-invariant contourlet transform and its application to image denoising. IEEE Transactions on Image Processing 15(11), 3362–3374 (2006)
3. Stamou, G., Krinidis, M., Loutas, E., Nikolaidis, N., Pitas, I.: 2D and 3D motion tracking in digital video. In: Bovik, A.C. (ed.) Handbook of Image and Video Processing. Academic Press (2005)
4. Moeslund, T.B., Granum, E.: A survey of computer vision based human motion capture. Computer Vision and Image Understanding 81, 231–268 (2001)
5. Liu, K., Guo, L., Chen, J.: Contourlet transform for image fusion using cycle spinning. Journal of Systems Engineering and Electronics 22(2), 353–357 (2011)
6. Raghavendra, B.S., Bhat, P.S.: Contourlet Based Multiresolution Texture Segmentation Using Contextual Hidden Markov Models. In: Das, G., Gulati, V.P. (eds.) CIT 2004. LNCS, vol. 3356, pp. 336–343. Springer, Heidelberg (2004)
7. Li, Y.-Q., He, M.-Y., Fang, X.-F.: SAR Image Segmentation Algorithm Using Mean Shift on Contourlet Domain. Computer Engineering 33(22), 48–50 (2007)
8. Contourlet Toolbox, Matlab source code, `http://www.ifp.uiuc.edu/~minhdo/software/`
9. Gopinath, R.A.: The Phaselet Transform – An Integral Redundancy Nearly Shift-Invariant Wavelet Transform. IEEE Trans. on Signal Processing 51, 1792–1805 (2003)

10. Donoho, D.L., Johnstone, I.M.: Ideal spatial adaptation by wavelet shrinkage. Biometrika 8, 425–455 (1994)
11. Huttenlocher, D.P., Noh, J.J., Rucklidge, W.J.: Tracking Non-Rigid Objects in Complex Scenes. In: Proceedings of 4th International Conference on Computer Vision, Berlin, May 11-14, pp. 93–101 (1993)
12. Stamou, G., Krinidis, M., Loutas, E., Nikolaidis, N., Pitas, I.: 2D and 3D motion tracking in digital video. In: Bovik, A.C. (ed.) Handbook of Image and Video Processing. Academic Press (2005)
13. Binh, N.T., Minh, L.N.: Adaptive medical image edge detection in contourlet domain. In: Proceedings of the 4th International Conference on the Development of Biomedical Engineering, pp. 238–241 (2012)
14. Binh, N.T., Khare, A.: Object tracking of video sequences in curvelet domain. International Journal of Image and Graphics 11(1), 1–20 (2011)
15. Masoud, O., Papanikolopoulos, N.P.: A novel method for tracking and counting pedestrians in real-time using a single camera. IEEE Transactions on Vehicular Technology 50, 1267–1278 (2001)
16. Prakash, O., Khare, A.: Tracking of Non-Rigid Object in Complex Wavelet Domain. Journal of Signal and Information Processing 2, 105–111 (2011)
17. Wang, Y., Van Dyck, R.E., Doherty, J.F.: Tracking Moving Objects in Video Sequences. In: Proc. Conference on Information Sciences and Systems, Princeton, NJ (March 2000)