

MOTION ESTIMATION IN THE FREQUENCY DOMAIN USING FUZZY C-PLANES CLUSTERING

Çiğdem Eroğlu Erdem, Güneş Karabulut, Evşen Yanmaz, Emin Anarım

Department of Electrical and Electronics Engineering
Boğaziçi University, İstanbul, 80815, Turkey

ABSTRACT

A recent work explicitly models the discontinuous motion estimation problem in the frequency domain and the motion parameters are estimated using a harmonic retrieval approach [1, 2]. The vertical and horizontal components of the motion are estimated from the locations of peak points of respective periodogram analysis and they are matched using a procedure proposed in [1]. In this work, we present a more efficient method for the motion parameter matching task, which eliminates the problems of the matching method described in [1]. The method described in this paper is based on fitting planes to the 3-D frequency domain data using the fuzzy c-planes clustering approach [3, 4]. The method is successfully tested on different video sequences.

1. INTRODUCTION

Motion estimation is an important problem in video processing that is utilized in applications such as video coding, object-based video manipulation, object based segmentation and dynamic scene analysis. Many studies have been carried out in an effort to estimate the motion of the objects in a scene [5, 6, 7, 8, 2, 9]. There are a number of approaches based on the correlation methods [5, 7] and the optical flow constraint [5, 6, 7]. These methods generally use regularization techniques assuming the smoothness of the motion field in order to obtain a unique solution. This smoothness assumption has an unwanted side effect of oversmoothing the motion discontinuities.

In a recent work [1, 2], the discontinuous motion estimation problem is handled in a harmonic retrieval framework. In this approach, the motion discontinuities are explicitly modeled and the velocity estimation is independent of the shape of the moving object and the density of motion discontinuities. The moving objects are assumed to have different translational motion parameters and the motion is assumed to be time-invariant in a short time interval. In order to find the motion parameters, first, horizontal and vertical parameter sets along with their corresponding sums are estimated from the peak locations of a periodogram analysis [1]. A parameter matching approach follows this step which is prone to errors. In this work, we present a more robust method for estimation of the motion parameters, which uses the 3-D frequency domain data and fits planes to them via the fuzzy c-planes clustering method [3, 4]. This approach has been indicated in [1] as future work but no results are given.

This work was supported by Scientific and Technical Research Council of Turkey (TÜBİTAK-BAYG) and Boğaziçi University Research Fund under project 99A207.

In Section 2, the theoretical background for modeling the motion estimation problem in the frequency domain [1] is introduced. In Section 3, the motion parameter estimation method using the fuzzy c-planes clustering approach is presented. In Section 4, the experimental results are presented and compared with the matching method of [1]. In Section 5, concluding remarks are given.

2. THEORETICAL BACKGROUND

In this section, the multiple discontinuous motion modeling approach described in [1] is briefly summarized.

Assuming there exists N objects (partitions) in the scene, (including the background) the image sequence is represented by [1]:

$$f(x, y; t) = \sum_{l=1}^N f_{vl}(x - u_l t, y - v_l t) + e(x, y; t) \quad (1)$$

where,

$$f_{vl}(x - u_l t, y - v_l t) = \sum_{l=1}^N f_{Il}(x - u_l t, y - v_l t) \bar{w}_{l0}(x - u_l t, y - v_l t) \quad (2)$$

is the sum of shifted ideal image of object l (f_{Il}), multiplied by its corresponding average support window over T frames (\bar{w}_{l0}) [1], and (u_l, v_l) denotes the velocity parameters of object l . The term

$$e(x, y; t) = \sum_{l=1}^N f_{Il}(x - u_l t, y - v_l t) d_l(x, y; t) \quad (3)$$

denotes the error, which is the sum of shifted ideal image of object l (f_{Il}) times the difference between the average support of object l over T frames and the support at time t (d_l). The 2-D Fourier transform of (1) gives:

$$F(w_x, w_y; t) = \sum_{l=1}^N F_{vl}(w_x, w_y) e^{-j(w_x u_l + w_y v_l)t} + E(w_x, w_y; t). \quad (4)$$

If we fix (w_x, w_y) , $F(w_x, w_y; t)$ is a 1-D signal consisting of N harmonics in noise where the l^{th} frequency of the harmonic is

$$w_l = -w_x u_l - w_y v_l, \quad l = 1, \dots, N. \quad (5)$$

For a specific (w_x, w_y) pair, the periodogram of $F(w_x, w_y; t)$ is given by:

$$I_F(w_x, w_y; w) = \frac{1}{T} \left| \sum_{t=0}^{T-1} F(w_x, w_y; t) e^{-j w t} \right|^2 \quad (6)$$

the peaks of which give the component frequencies $w = -w_x u_l - w_y v_l$, which is a function of the velocity values. In [1], the peaks of the periodograms of $F(w_x, 0; t)$, $F(0, w_y; t)$ and $F(w_x, w_y; t)$ are determined to estimate the motion parameter sets in the x and y directions, and the set consisting of the sum of these two velocity components, respectively. Then the motion parameter sets in the x and y directions are paired such that,

$$\sum_{i=1}^{\hat{N}} (a_i + b_i - c_i)^2 \quad (7)$$

is minimized where a_i, b_i and c_i are members of the sets consisting of the motion parameters in the x and y directions and their sums, respectively [1]. The parameter \hat{N} which indicates the number of moving objects in the scene is estimated as the maximum of the number of elements of the three velocity sets described above.

This approach may give erroneous results in some cases. For example, if we have three objects in our scene moving with velocities $(u_1, v_1) = (0, 2)$, $(u_2, v_2) = (2, 0)$ and $(u_3, v_3) = (2, 2)$, the above matching method of [1] will find the number of moving objects in the scene as $\hat{N} = 2$, since all the velocity sets will contain 2 elements. In the next section we propose another method based on fuzzy c-means clustering that eliminates the need for matching the velocity components using (7).

3. MOTION PARAMETER ESTIMATION USING FUZZY C-PLANES CLUSTERING

In this section, we present another approach for simultaneous motion parameter estimation and pairing which is based on the fuzzy c-planes clustering algorithm [3, 4]. This approach overcomes the difficulties of the pairing method of [1].

For each (w_x, w_y) pair, the corresponding dominant peaks of the periodogram of (6) are estimated and the 3-D data are formed as $h_i = (w_{x_i}, w_{y_i}, w_i)$. Note that since these 3-D points satisfy the equation

$$w_i = -w_{x_i} u_l - w_{y_i} v_l, \quad l \in \{1, \dots, N\} \quad (8)$$

they form N planes in 3-D. The plane parameters are the negatives of the motion parameters. In order to cluster these 3-D frequency domain points into N planes, we use the fuzzy c-planes algorithm, the steps of which are given below:

Let the number of 3-D data points be M .

1. Initialize the membership matrix consisting of elements u_{ij} , $i = 1, \dots, M, j = 1, \dots, N$. Here, u_{ij} denote the membership value of data point i to plane (cluster) j . The initialization is performed such that

$$0 \leq u_{ij} \leq 1, \quad \forall i, j, \quad (9)$$

and

$$\sum_{j=1}^N u_{ij} = 1, \quad \forall i. \quad (10)$$

2. The cluster centroids are estimated using the equation,

$$v_j = \frac{\sum_{i=1}^M u_{ij}^q h_i}{\sum_{i=1}^M u_{ij}^q}, \quad j = 1, \dots, N. \quad (11)$$

Here, q is a parameter that represents the amount of fuzziness of the clustering, which is chosen around 2 in our experiments. The larger the q is, the fuzzier the clustering becomes.

3. Then, the covariance matrix of each class (plane) is estimated using:

$$R_j = \frac{\sum_{i=1}^M u_{ij}^q (h_i - v_j)(h_i - v_j)^T}{\sum_{i=1}^M u_{ij}^q}, \quad j = 1, \dots, N. \quad (12)$$

4. The largest 2 eigenvectors (V_{j1} and V_{j2}) of R_j are found and the squared distance between the data point i and plane j is found from:

$$d_{ij}^2 = \|h_i - v_j\|^2 - \sum_{k=1}^2 ((h_i - v_j)^T V_{jk})^2, \quad \forall i, j. \quad (13)$$

5. The membership values of the data points are updated as:

$$u_{ij} = \frac{(d_{ij}^2)^{\frac{-1}{q-1}}}{\sum_{k=1}^N (d_{ik}^2)^{\frac{-1}{q-1}}}, \quad i = 1, \dots, M; j = 1, \dots, N. \quad (14)$$

6. The steps 2-5 are carried out until the change in the membership values between two successive iterations is below a certain error threshold ε , i.e.,

$$\|u_{ij}^s - u_{ij}^{s+1}\| < \varepsilon, \quad (15)$$

where s denotes the iteration counter.

4. EXPERIMENTAL RESULTS

The proposed method is first tested on a synthetically generated sequence, which contains three moving blocks on a stationary background. A few frames of this sequence are shown in Figure 1. The velocity vectors for the three blocks are $(u_1, v_1) = (-2, -2)$, $(u_2, v_2) = (-2, 0)$, $(u_3, v_3) = (0, -2)$ in the northwest direction. The theoretical motion planes in the 3-D frequency domain is shown in Figure 2. The top plane corresponds to the parameters $(u_1, v_1) = (-2, -2)$, and the other planes correspond to the parameters $(u_2, v_2) = (-2, 0)$, and $(u_3, v_3) = (0, -2)$. In Figure 3, several periodogram plots are given for different pairs. Note that the three peaks are clearly visible in Figure 3(a) and Figure 3(b) whereas only two peaks exist when $w_x = w_y$ and w_y is close to zero as shown in Figure 3(c) and Figure 3(d). The 3-D data extracted from the peak locations of the periodograms are shown in Figure 4. The estimated velocities are close to the true values as tabulated in Table 1 and as demonstrated in Figure 5. Using fuzzy c-planes clustering, the velocities of the three moving objects are correctly estimated whereas, this is not possible with the matching approach of [1] as discussed in Section 2.

We also tested the plane-fitting approach on the 'Taxi in garden' sequence which is shown in Figure 6. In this sequence, the white taxi moves with a horizontal velocity of 3 and a vertical velocity of -1 in the north-east direction. As a preprocessing step, we applied histogram equalization in order to enhance the contrast. Although the background is cluttered, the plane fitting approach was able to estimate the velocity vectors as (2.9948,-0.98736).

Original vectors	Estimated vectors
$(u_1, v_1) = (-2, -2)$	$(\hat{u}_3, \hat{v}_3) = (-2.018, -2.018)$
$(u_2, v_2) = (-2, 0)$	$(\hat{u}_1, \hat{v}_1) = (-2.060, -0.071)$
$(u_3, v_3) = (0, -2)$	$(\hat{u}_2, \hat{v}_2) = (0.009, -2.044)$

Table 1: The original and estimated velocity values of synthetic sequence containing three moving blocks.

Original vectors	Estimated vectors
$(u_1, v_1) = (2.7, 0.45)$	$(\hat{u}_1, \hat{v}_1) = (2.971, 0.618)$
$(u_2, v_2) = (-2.65, -0.35)$	$(\hat{u}_2, \hat{v}_2) = (-2.243, -0.334)$

Table 2: The original and estimated velocity values of the 'Hamburg Taxi' sequence.

Finally, we used the Taxi sequence which is shown in Figure 7 for testing the proposed approach. The true velocity values are determined by hand [1]. As seen in Table 2, the estimated values are quite close to the actual velocities.

5. CONCLUSIONS

In this work we presented an alternative approach for the pairing and the motion parameter estimation problem of [1]. The presented method is based on fitting planes to the 3-D data in the frequency domain using fuzzy c-planes algorithm. This approach eliminates the problems in the velocity parameter pairing stage of [1].

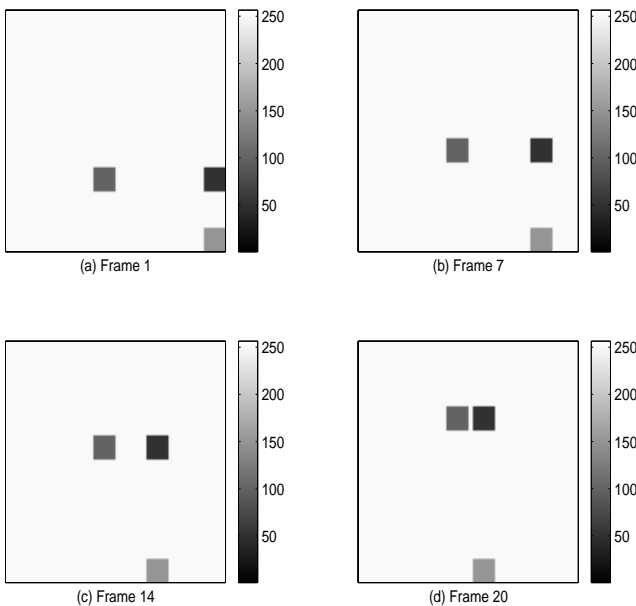


Figure 1: Several frames of the synthetic sequence containing three moving blocks.

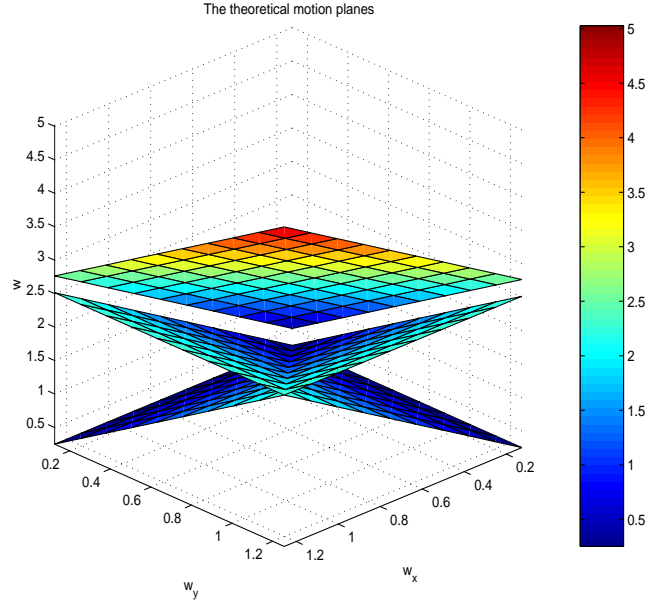


Figure 2: The theoretical motion planes for the synthetic sequence containing three moving blocks.

REFERENCES

- [1] W. Chen, G. Giannakis, and N. Nandhakumar, "A harmonic retrieval framework for discontinuous motion estimation," *IEEE Trans. Image Processing*, vol. 7, no. 9, pp. 1242–1256, 1998.
- [2] W. Chen, G. B. Giannakis, and N. Nandhakumar, "Spatio-temporal approach for time-varying image motion estimation," in *Proc. IEEE Int. Conf. on Image Processing (ICIP)*, 1994, vol. 2, pp. 232–236.
- [3] J. C. Bezdek, C. Coray, R. Gunderson, and J. Watson, "Detection and characterization of cluster substructure i. linear structure: Fuzzy c-lines," *SIAM Journal of Appl. Math.*, vol. 40, no. 2, pp. 339–357, 1981.
- [4] F. Hoppner, F. Klawonn, R. Kruse, and T. Runkler, *Fuzzy Cluster Analysis: Methods for Classification, Data Analysis and Image Recognition*, John Wiley and Sons, 1999.
- [5] J. L. Barron, D. J. Fleet, and S. S. Beauchemin, "Performance of optical flow techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43–77, 1994.
- [6] B. K. P. Horn and B. G. Schunk, "Determining optical flow," *Artificial Intelligence*, vol. 17, pp. 185–203, 1981.
- [7] A. M. Tekalp, *Digital Video Processing*, Prentice-Hall, New Jersey, 1995.
- [8] R. W. Young and N. G. Kingsbury, "Frequency-domain motion estimation using a complex lapped transform," *IEEE Trans. Image Processing*, vol. 2, no. 1, pp. 2–17, 1993.
- [9] C. Demiroglu and E. Anarım, "Sıklık düzleminde devinim parametrelerinin kestirimi," in *Sinyal İşleme ve Uygulamaları Kurultayı*, Ankara, 1999, pp. 336–339.

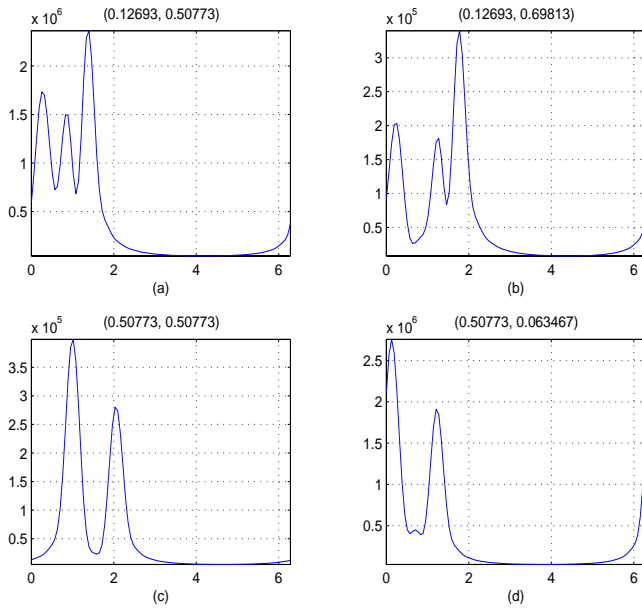


Figure 3: The periodogram plots for (a) $(w_x, w_y) = (0.13, 0.51)$, (b) $(w_x, w_y) = (0.13, 0.70)$, (c) $(w_x, w_y) = (0.51, 0.51)$, (d) $(w_x, w_y) = (0.51, 0.06)$

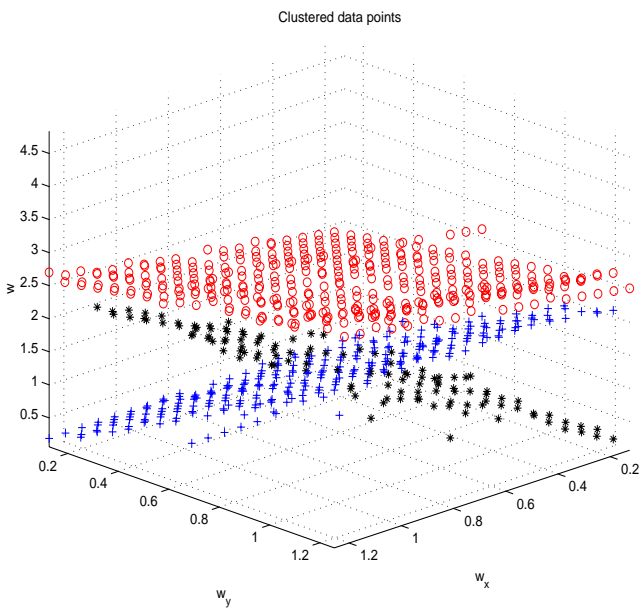


Figure 4: The data points obtained from the peak points of the periodogram for the synthetic sequence. The symbols 'o', '*' and '+' represent data points belonging to different planes estimated by the fuzzy c-planes clustering algorithm.

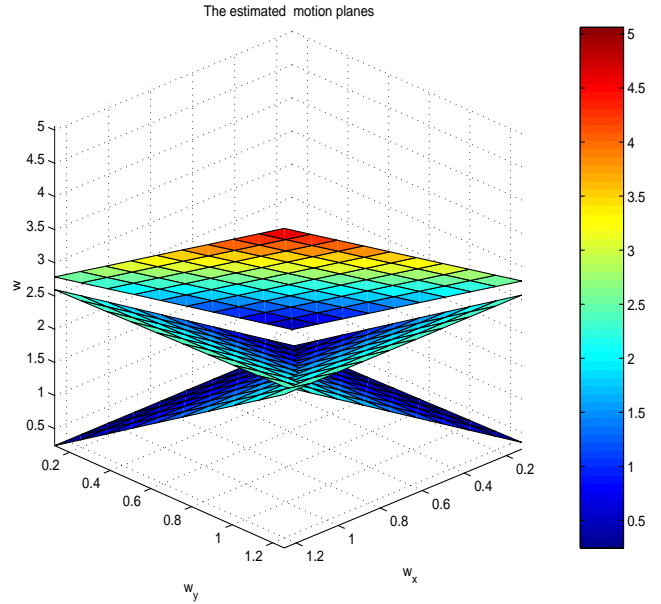


Figure 5: The motion planes estimated using the clustered data shown in Figure 4.

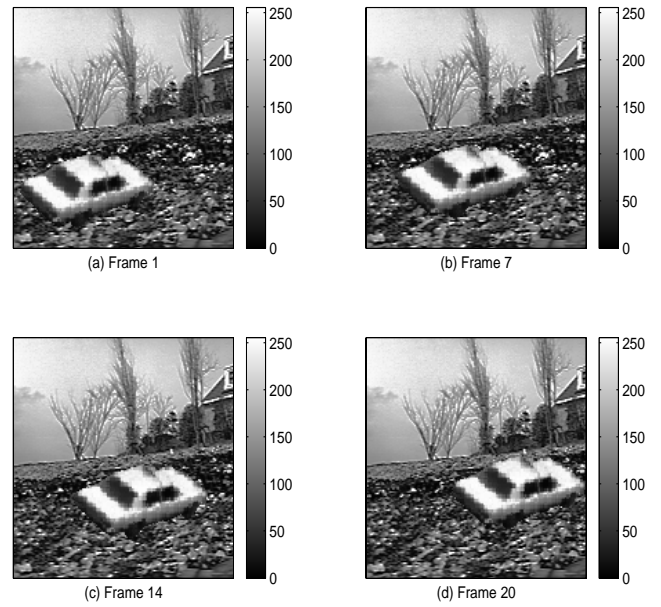


Figure 6: Several frames of the 'Taxi in garden' sequence.



Figure 7: The 1st (left) and the 20th (right) frames of the 'Hamburg Taxi' sequence.