Feature Matching using Co-inertia Analysis for People Tracking

Srinidhi Mukanahallipatna Simha, Duc Phu Chau, François Bremond

HAL Id: hal-00909566

https://hal.inria.fr/hal-00909566

Submitted on 26 Nov 2013

Feature matching using co-inertia analysis for people tracking

Srinidhi MUKANAHALLIPATNA SIMHA[1]
Duc Phu CHAU[1] and Francois BREMOND[1]

[1] *STARS team, INRIA, Sophia Antipolis Mediterranee, France*
{*srinidhi.mukanahallipatna_simha, duc − phu.chau, francois.bremond*}*@inria.fr*

Abstract: Robust object tracking is a challenging computer vision problem due to dynamic changes in object pose, illumination, appearance and occlusions. Tracking objects between frames requires accurate matching of their features. We investigate real time matching of mobile object features for frame to frame tracking. This paper presents a new feature matching approach between objects for tracking that incorporates one of the multivariate analysis method called Co-Inertia Analysis abbreviated as COIA. This approach is being introduced to compute the similarity between Histogram of Oriented Gradients (HOG) features of the tracked objects. Experiments conducted shows the effectiveness of this approach for mobile object feature tracking.

1 Introduction

Computer vision is a fast growing field these days. Research in recent years has focused more on ways of movements of the user, understanding the user's act or behavior, and then reacting appropriately. Hence there is a need for *tracking* the user/person in videos/images. One of the ways to track the user involves matching features between frames.

Matching the visual appearances of the user over consecutive image frames is one of the most critical issues in visual target tracking. The general process in feature tracking of objects is to find the distance/similarity between them in the feature space. Similar to many other computer vision problems such as object recognition, two important factors play a major role: visual features that characterize the target in feature space, and the similarity/distance between them to determine the closest match. The uniqueness in this tracking approach, different from those recognition tasks, is that the matching is constructed between two consecutive frames and hence it demands a more computationally efficient solution. A simple distance metric like euclidean distance for simple nearest neighbor matching will work if we can identify strong features which are invariant to lighting changes and local deformation and those are discriminative from the false positives. But when strong features cannot be easily specified, the choice of distance/similarity metric largely influences the matching performance. Hence, finding an ideal metric for visual tracking becomes critical.

This paper does not intend to present new image feature extraction method. Rather, we introduce a multivariate analysis method to find the object similarity/matching between frames known as co-inertia analysis (COIA). This method has been introduced recently to solve statistical problems in ecology (Doledec and Chessel, 1994) and is quiet unknown in the computer vision community. The proposed method for tracking has been found out to be efficient especially against illumination and appearance changes in objects.

1.1 Background

There are several strategies to match two data tables namely *the within* and *between* principal components analysis and the linear discriminant analysis. If two tables are linked by the same individuals, one can find a structure, a *co-structure* to study the relationship between the two set of variables.

Let's call X and Y two continuous variables measured on the same individuals. Let's call $\bar{x}$ and $\bar{y}$ the means of X and Y respectively. Let's call $v(x)$ and $v(y)$ the descriptive variances of X and Y respectively. A measure of the relationship between X and Y is provided by the descriptive covariances:
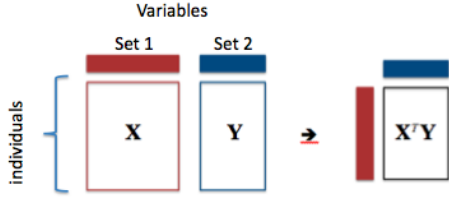
Figure 1: Basic strategy to match two data tables

$$cov(x,y) = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \quad (1)$$

The covariance can be negative or positive: that depends on the sense of the relationship. All the variance-covariance information can be gathered in a matrix (which can be called the covariance matrix).

$$\begin{pmatrix} v(x) & cov(x,y) \\ cov(y,x) & v(y) \end{pmatrix}$$

This matrix is symmetric: $cov(x,y) = cov(y,x)$ and $cov(x,x) = v(x)$. One can divide the covariance between X and Y by the square roots of the variances of X and Y (i.e., by the standard deviation of X and Y). By doing so, we obtain the coefficient of correlation between X and Y.

$$cor(x,y) = \frac{cov(x,y)}{\sqrt{v(x)}\sqrt{v(y)}} \quad (2)$$

The closer the coefficient to either -1 or 1, the stronger the correlation between the two variables.

The organization of this paper is as follows: after the introduction, related work is discussed in section 2. Section 3 describes the Co-Inertia Analysis. The proposed tracking system along with the algorithm is presented in section 4. The experiments are reported and discussed in section 5 followed by conclusion and future work.

2   Related work

The typical feature matching techniques are Mean-shift, Particle filter and Template matching. Mean shift algorithm (Comaniciu et al., 2003) uses color histogram to describe the target region for which the amount of calculation needed is small resulting in good real time implementation. But Mean shift algorithm finds its disadvantage when the image is gray scale or less textured and also when there is scale variations in the image. This results in losing the target easily.

Particle filter (N.Johnson and D.C.Hogg, 1996) for nonlinear filtering algorithm based on bayesian estimation has the unique advantage of processing the parameter estimation. This algorithm does not find its usage universally since it has a problem of convergence besides being slow.

Template matching (Feng et al., 2008) is one of the principle techniques in visual tracking. This algorithm judges the matching degree based on the similarity of the adjacent pixels. This calculation is easy and fast but need to be calculated for the whole image. Problem of this method occurs when the target is deformed, rotated or occluded in which case it fails to track the target.

The existing feature matching methods, besides euclidean metric, includes the Matusita metric (Hager et al., 2004), the Bhattacharyya coefficient (Comaniciu et al., 2003), the Kullback-Leibler divergence (El-gammal et al., 2003), the information-theoretic similarity measures (Viola and Wells., 1995), and a combination of those (Yang et al., 2005). In (Jiang et al., 2011) the author presents an adaptive metric into differential tracking method where learning of optimal distance metric is automatic for accurate matching. In (Yang et al., 2005) the author describes about a new similarity measure which allows the mean shift algorithm to track more general motion models in an integrated way using fast gauss transform leading to efficient and robust non parametric spatial feature tracking algorithm.

The proposed method calculates the coefficient of relationship between the physical objects in the current frame to the physical objects in the previous frame. This is enabled by getting the Histogram of Oriented Gradients (HOG) features (Dalal and Triggs, 2005) of the physical objects in the first place and then using COIA to find the coefficient of relationship between them for tracking the objects.

COIA has been used in Human computer interaction and Biometrics community for finding audio-visual speech synchrony (Goecke and Millar, 2003) (Eveno and Besacier, 2005). For the first time we present COIA as a feature matching technique of mobile objects for tracking in computer vision domain.

3   Co-Inertia Analysis

Coinertia analysis (COIA) is a relatively new multivariate statistical analysis for coupling two (or more) sets of parameters by looking at their linear combinations. It was introduced for ecological studies by Doledec and Chessel (Doledec and Chessel, 1994).

As it appears to be relatively unknown in the Computer Vision community, we will first give some background information. In COIA, the term inertia is used as a synonym for variance. The method is related to other multivariate analysis such as canonical correspondence analysis , redundancy analysis, and canonical correlation analysis (CANCOR) (Gittins, 1985). COIA is a generalization of the inter-battery analysis by Tucker (Tucker, 1958) which in turn is the first step of partial least squares (skuldsson, 1988).

COIA is very similar to CANCOR in many aspects. It also rotates the data to a new coordinate system and the new variables are linear combination's of the parameters in each parameter set. However, here, it is not the correlation between the two sets that is maximized but the co-inertia (or co-variance) which can be decomposed.

Compromise between the correlation and the variance in either set is found by COIA. It aims to find orthogonal vectors - the co-inertia axes - in the two sets which maximize the co-inertia value. The number of axes is equivalent to the rank of the co-variance matrix. The advantage of COIA is its numerical stability. The number of parameters relative to the sample size does not affect the accuracy and stability of the results (Doledec and Chessel, 1994). The method does not suffer from co-linearity and the consistency between the correlation and the coefficients are very good (Dray et al., 2003). Thus, COIA is a very well suited multivariate method in our case.

The co-inertia value is a global measure of the co-structure in the two sets (mobile objects). The two parameter sets vary accordingly (or inversely) if the value is high and the sets vary independently if it is low. The correlation value gives a measure of the correlation between the co-inertia vectors of both domains.

Furthermore, one can project the variance onto the new vectors of each set and then compare the projected variance of the separate analysis with the variance from the COIA (see the appendix of (Doledec and Chessel, 1994) for the theory). The ratio of the projected variance from the separate analysis to the variance from the COIA is a measure of the amount of variance of a parameter set that is taken by the co-inertia vectors. It is important to compare the sum of axes, not axis by axis, because the variance projected onto the second axis depends on what is projected onto the first axis, and so on. Often it is sufficient to look at the first 2-3 axes because they already account for 90-95% of the variance. There are many possibilities of giving inputs to COIA to find the relationship in general as described in (Dray et al., 2003). Finally, a measure of overall relatedness of the two

domains, in our case mobile objects, based on the selected parameters is given by the RV coefficient (Heo and Gabriel., 1997).

In addition, COIA computes the weights (coefficients) of the individual parameters in the linear combination's of each set, so that it becomes obvious which parameters contribute to the common structure of the two sets and which do not. As has already been pointed out, these weights are much more stable than the weights in a CANCOR analysis. Finally, a measure of overall relatedness of the two domains based on the selected parameters is given by the RV coefficient (Heo and Gabriel., 1997).

One of COIAs biggest advantages is that it can be coupled with other statistical methods, such as correspondence analysis and PCA. That is, these methods are performed on the data of the two domains separately and then a COIA follows. In fact, (Dray et al., 2003) shows that seen in this context, COIA is a generalization of many multivariate methods. In our analysis, it means that we can use both shape PCs as input for a COIA without being restricted as in the case of CANCOR.

This multivariate method for finding the similarity between two tables i.e two mobile objects between frames in our case has been successfully tried and showcased with effective results in comparison with the state of the art.

3.1 Defining the relationship between two data tables

Different statistics such as Pearson correlation coefficient or covariance can be used to measure the relation between two variables. The purpose of this section is to define a statistic that measures the relation between two (or more) sets of variables. Let's call (X, Q_X, D) a statistical triplet, where X is a dataset containing p variables measured on n individuals. Q_X and D represent the weights of the p variables and n individuals, respectively. If all the variables of X are centered, the inertia I_X is the sum of variances. If D is the diagonal matrix $(n \times n)$ of individual weights $D = diag(w_1, ..., w_n)$ and if Q $(p \times p)$ is a metric of this hyperspace, then inertia of the "cloud of individuals" around the reference point o is simply

$$I_o = \sum_{i=1}^{n} w_i \|X_i - o\|_{Q_X}^2 = \sum_{i=1}^{n} w_i \|X_i\|_{Q_X}^2 \quad (3)$$

In general

$$I_X = trace(X Q_X X^T D) \quad (4)$$

This total inertia is a global measure of the variability of the data. It is the weighted sum of square distances measured with Q_X, between the points of X (n

individuals) and the reference point o. If Q is the Euclidean metric and D the diagonal matrix of uniform weights and if o is the centroid of the cloud, the inertia is simply a sum of variances. The individuals X_i can be projected on a Q_X-normed vector $\mathbf{u}$ and the projected inertia is expressed by

$$I(\mathbf{u}) = \mathbf{u}^T Q_X X^T D X Q_X \mathbf{u} \tag{5}$$

The total inertia can be easily decomposed on a set of p orthogonal Q_X-normed vectors $\mathbf{u}_k$:

$$I_o = \sum_{k=1}^{p} I(\mathbf{u}_k) = \sum_{k=1}^{p} \mathbf{u}_k^T Q_X X^T D X Q_X \mathbf{u}_k$$

$$= \sum_{k=1}^{p} \|X Q_X \mathbf{u}_k\|_D^2 \tag{6}$$

Let (Y, Q_Y, D) be a statistical triplet, where Y is a dataset containing q variables measured on n individuals. Q_Y and D represent the weights of the q variables and n individuals, respectively. If all the variables of Y are centered, the inertia I_Y is the sum of variances. In the same way as explained previously, the inertia for this data will be

$$I_Y = trace(Y Q_Y Y^T D) \tag{7}$$

and can be decomposed as above on a set of vectors $\mathbf{v}_k$ and it is not difficult to study the common geometry of the two datasets. Co-Inertia is a global measure of the co-structure of individuals in the two data hyperspaces. It is high when two structures vary simultaneously (closely related) and low when they vary independently (not closely related). It is defined by

$$COI = \sum_{k=1}^{p} \sum_{j=1}^{q} (u_k^T Q_X X^T D Y Q_Y v_j)^2 =$$

$$\sum_{k=1}^{p} \sum_{j=1}^{q} ([X^k]^T D Y^j)^2$$

$$= trace(X Q_X X^T D Y Q_Y Y^T D) \tag{8}$$

If the datasets are centered, then inertia is a sum of variances and co-inertia is a sum of square covariances.

3.2 Principle of Co-Inertia Analysis

The co-inertia criterion measures the concordance between two data sets, and a multivariate method based on this statistic has been developed. Co-inertia analysis (Doledec and Chessel, 1994) is a symmetric coupling method that provides a decomposition of the co-inertia criterion on a set of orthogonal vectors. It is defined by the analysis of the datasets defined. Different types of data lead to different transformations (centering, normalization,...) of X and Y and to different weights Q_X and Q_Y. Co-inertia analysis aims

to find vectors v_1 and u_1 in the respective spaces with maximal co-inertia. If X and Y are centered, then COIA maximizes the square covariance between the projection of X on u_1 and the projection of Y on v_1:

$$P(u_1) = (X Q_x).u_1$$

$$cov^2(P(u_1), P(v_1)) = corr^2(P(u_1), P(v_1))$$

$$\times var(P(u_1)) \times var(P(v_1)) \tag{9}$$

This square covariance can be easily decomposed, showing that COIA finds a compromise between the correlation, the variance of individuals in their respective spaces viewpoint. The second and further pairs of vectors $(u_2, v_2...)$ maximize the same quantity but are subject to extra constraints of orthogonality.

$$RV = \frac{COI(X,Y)}{\sqrt{COI(X,X)}\sqrt{COI(Y,Y)}} \tag{10}$$

The RV-coefficient is the coefficient of correlation between the two tables X and Y. This coefficient varies between 0 and 1: the closer the coefficient to 1, the stronger the correlation between the tables.

4   The Proposed tracking algorithm

The proposed tracking algorithm needs a list of detected objects in a temporal window as input. This is enabled using a processing chain for image acquisition, background subtraction (A.T. Nghiem and Thonnat, 2009), classification, object detection and tracking system (Chau et al., 2011). The size of the temporal window is a parameter. This tracker system first computes the link similarity between any two detected objects appearing in a given temporal window based on Co-inertia analysis. The trajectories that include a set of consecutive links from previous stages are then computed to get the global similarity. Noisy trajectories are removed through a filter.

To get the similarity link of the objects, first Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005) descriptors of the detected physical objects in the current frame are calculated and compared with the HOG descriptors of the detected Physical Objects (PO) in the previous frame. This is where the usage of COIA comes in to the picture. COIA is used to compare the similarity between the physical objects. For each detected object pair in a given temporal window, the system computes the link similarity (i.e. instantaneous similarity) defined using COIA. A temporal link is established between these two objects when their link similarity is greater or equal to a threshold. At the end of this stage, we obtain a weighted graph (Chau et al., 2011) whose vertices are

the detected objects in the considered temporal window and whose edges are the temporally established links associated with the object similarities.
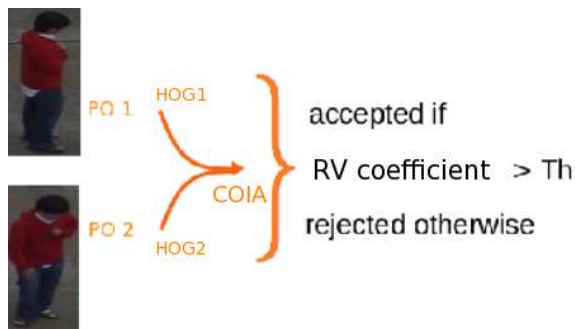

Figure 2: Physical Object (PO) similarity measure criteria.

Algorithm 1: Calculating F2F link between objects using COIA

Require: Objects Detected from A Detector

1: **function** COMPUTEF2FCOIALINK(ObjectT, imageT, imageTmN)
Check (objectT != NULL)
bboxTmN = Previous.getbbox;
bboxT = Present.getbbox;
ObjectTmN = getROI(imageTmN,bboxTmN);
ObjectT = getROI(imageT,bboxT);
ObjectTFeature = HOG(ObjectT);
ObjectTmNFeature = HOG(ObjectTmN);
COIAScore =
$COIA(ObjectTFeature, ObjectTmNFeature)$;
return *COIAScore;*

In this paper HOG descriptor is calculated on the whole detected objects (whole bounding box) but other possibilities are taking a patch in the detected object and calculating HOG on the same for finding similarity is also possible. Initial trial on finding feature similarity with color histograms between mobile objects using COIA has been accomplished but will be consolidated by combining different features for efficiency and will be portrayed in future. Different features like Local Binary Patterns, Local ternary patters etc. can be used as input to COIA.

As shown in Figure 2, the RV coefficient will be higher for the same physical objects and lower for the different physical objects which will enable us to track the correct objects. The long term similarity score between the object and the trajectory is calculated as explained in (Chau et al., 2011). To supplement the frame to frame tracking and long term tracking, using COIA to measure the similarity between

---

Algorithm 2: Calculating COIA Score

Require: Mobile Objects features matrices

function COIA(ObjectTFeature.matrix, ObjectTmNFeature.matrix)
Check dimensions of the input matrices;
W = combine the two feature matrices;
[u,w,v] = SVD(W);
x = ObjectTFeature.matrix;
y = ObjectTmNFeature.matrix;
COI = trace(w);
Calculate I_x from equation 4;
Calculate I_y from equation 7;
RV = $\frac{COI}{\sqrt{I_x \times I_y}}$;
return *RV;*

Algorithm	PO1-PO1	PO1-PO2	PO1-PO3	PO1-PO4	PO1-PO5
COIA	**1.0**	**0.8932**	**0.8663**	**0.7161**	**0.8429**
Correlation	1.0	0.5532	0.2259	0.0287	0.2712
Intersection	1.0	0.8675	0.8585	0.7466	0.8212
Chi-square	0	0.2779	0.3746	0.8114	0.4584
Bhattacharyya Distance	0	0.1338	0.1564	0.2397	0.1738

Table 1: Comparison of Histogram similarities (COIA, Correlation, Intersection) and distances (Chi-square and Bhattacharyya)

mobile objects using HOG features, we make use of dominant color, 2D distance, 2D shape ratio and 2D area of the mobile objects as described in (Chau et al., 2011) for better efficiency.

Figure 3: From left PO1, PO2, PO3, PO4, PO5

Table 1 signifies the behavior of COIA. In the Figure 3, PO1 and PO2 are same but with different poses and the rest of the PO's are different from PO1. As explained the HOG feature of each PO is calculated and COIA is used to find the similarity between PO's using the corresponding feature matrices. As expected, the similarity between PO1 and PO2 should be higher than any other PO combination from PO1. To summarize COIA, Correlation and Intersection should be higher for PO1-PO2 and lower for chi-square and Bhattacharyya distance than others except PO1-PO1 obviously. The main motivation of Table 1 is to showcase, for the first time in tracking mobile objects, that

COIA behaves as expected compared to other algorithms in finding similarity between mobile objects. All the values were calculated using OpenCV.

5   Experimental results

We present some real-time object tracking results using the proposed algorithm. The objective of this experimentation is to show the effectiveness of the proposed algorithm in tracking system. People detection algorithm with Background subtraction based on Mixture of Gaussian and Local Binary Pattern with Adaboost is used for detection.

HOG descriptor of the OpenCV is used to get the features of the detected objects which goes as input to the COIA routine to calculate the coefficient of relationship between them. The videos used for testing the tracker belongs to a public benchmark dataset ETISEO (Etiseo,) and private dataset Hospital. The algorithm has been implemented in C++ but was coded in Matlab initially to verify the proposed idea. The proposed method runs on Intel Xeon 8 core 2GHz CPU.

In order to evaluate the tracking performance, we use the three tracking evaluation metrics defined in the ETISEO project (Etiseo,). The first tracking evaluation metric M1 measures the percentage of time during which a reference object (ground truth data) is correctly tracked.

The second metric M2 computes throughout time how many tracked objects are associated with one reference object. The third metric M3 computes the number of reference object IDs per tracked object. These metrics must be used to obtain a complete performance evaluation. The metric values are defined in the interval [0, 1]. The higher the metric value is, the better the tracking algorithm performance.

All the result images shown contains green bounding box representing tracking of the person associated with id and blue ellipse shows the person with tracking/trajectory path of the person.

The Figure 5 shows the tracked image sequences of Hospital dataset. This dataset is challenging in terms of shadow and pose change of the object (person). The proposed algorithm for tracking produces a perfect tracking through out the video sequence as mentioned in Table 2. Figures 6, 7 and 8 illustrates the tracking of objects for ETISEO public dataset video sequence. This dataset is very challenging because of many illumination/contrast changes in the frames. As shown in Figure 6, the video sequence is of low light settings and there are transitions from light switched off to on and vice-versa which makes the tracking very challenging. The proposed algorithm for tracking again produces perfect tracking results for this video sequence in-spite of challenging lighting/contrast conditions. The results are shown in Table 2.

The contrast of an object is defined as the color intensity difference between the object and its surrounding background. Figure 4 shows the tracking result in the sequence ETI-VS1-BC-13-C4 (from the ETISEO dataset) at some frames in two cases: the first row illustrates the tracking result using the proposed COIA method; the second row illustrates the tracking result using (P.Bilinski et al., 2009). From frame 1400 to frame 1425, the object contrast does not change and the tracking results are good in both cases. From frame 1425 to frame 1525, the object contrast decreases, and from frame 1525 to frame 1600 it increases. While the proposed COIA based tracker still ensures a good performance during these two periods, the (P.Bilinski et al., 2009) tracker cannot keep good tracking result. When the object contrast changes, the distance between HOG features of the physical objects in (P.Bilinski et al., 2009) decreases and hence tracking fails (change in bounding box color). So the variation of object contrasts influences the tracking performance.

In these experiments, tracker results from seven different teams (denoted by numbers) in ETISEO (Video understanding Evaluation) project (Etiseo,) have been presented: T1, T8, T11, T12, T17, T22, T23. Because names of these teams are hidden, we cannot determine their tracking approaches. Table 3 presents performance results of the considered trackers. The tracking evaluation metrics of the proposed tracker gets the highest values in most cases compared to other teams.

In Table 4 (Zamir et al., 2012) Standard CLEAR MOT (Kasturi, 2009) has been used as evaluation metrics. MOTA (Motion Object Tracking Accuracy) measures False positives, false negatives and ID-Switches . The average distance between the ground truth and estimated target locations has been defined as MOTP (Motion Object Tracking Precision). MOTP signifies the ability of the tracker in estimating the precise location of the object, regardless of its accuracy at recognizing object configurations, keeping consistent trajectories, and so forth. Therefore, MOTA has been widely accepted in the literature as the main gauge of performance of tracking methods. Video sequence View-001-S2L1 Figure 9 from PETS 2009 dataset has been used to compare the proposed tracking method with other tracking algorithms.
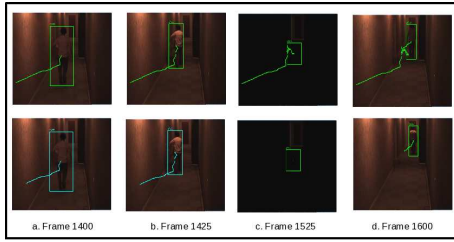
Figure 4: Top row proposed tracking algorithm in this paper, bottom row tracking result from (P.Bilinski et al., 2009)

Dataset	M1	M2	M3
Hospital	**1.0**	**1.0**	**1.0**
ETISEO	**1.0**	**1.0**	**1.0**

Table 2: Tracking results of two video sequences - Figure 5 and 7

Figure 5: Tracking of a patient in the hospital video sequence.

6 Conclusion and Future work

This paper presents a new approach to track physical/mobile objects, for the first time in this domain, using COIA which measures the similarity between them. The proposed method has been successfully validated with public datasets as mentioned in the previous section and shows promising results. This method has been validated with challenging video sequences to show the significance of the approach. We propose to use other features like color histograms, Local Binary Patterns etc. and combination of multiple features for the detected objects to find the similarity between them using COIA. We would also like to

Metric	Proposed method	T1	T8	T11	T12	T17	T22	T23
M1	**0.6**	0.48	0.49	0.56	0.19	0.17	0.26	0.05
M2	**1.0**	0.8	0.8	0.71	1.0	0.61	0.35	0.46
M3	**1.0**	0.83	0.77	0.77	0.33	0.80	0.33	0.39

Table 3: Tracking results of ETI-VS1-BE-18-C4 sequence - Fig 8

PETS2009view001	MOTA	MOTP
(Berclaz et al., 2011)	80.00	58.00
(Shitrit et al., 2011)	81.46	58.38
(Andriyenk and Schindler, 2011)	81.84	73.93
(Henriques et al., 2011)	84.77	68.74
Proposed method	**88.42**	**65.60**

Table 4: Tracking result comparison of PETS sequence - Fig 9
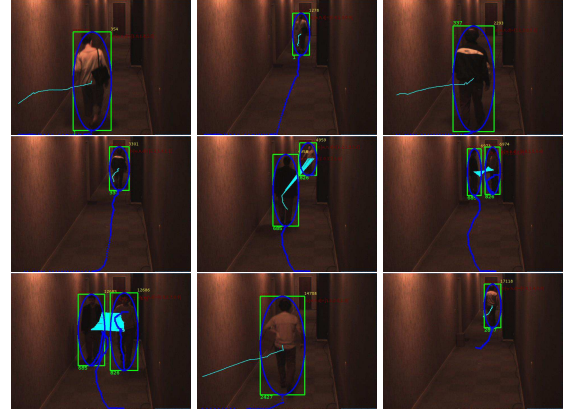

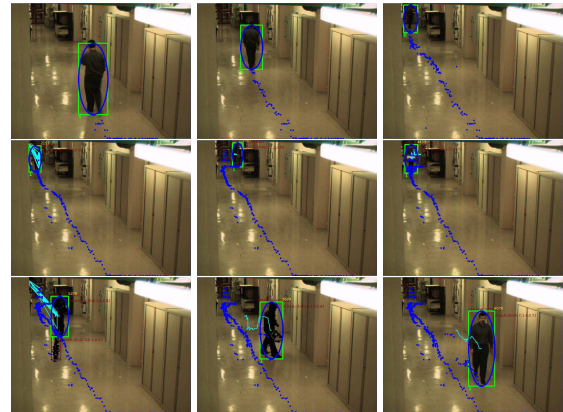Figure 6: People Tracking in the ETISEO ETI-VS1-BC-13-C4 video sequence.

Figure 7: People Tracking in a ETISEO video sequence.

propose in future to come up with a different weighting strategy for the features of the objects in finding the similarity using COIA.

REFERENCES

Andriyenk, A. and Schindler, K. (2011). Multi-target tracking by continuous energy minimization.. *CVPR*.

A.T. Nghiem, F. B. and Thonnat, M. (2009). "Controlling Background Subtraction Algorithms for Robust Object Detection". *ICDP*.

Berclaz, J., Fleuret, F., Turetken, E., and Fua, P. (2011).

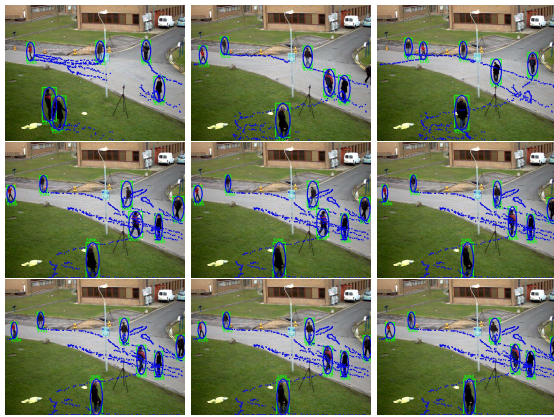Figure 8: People Tracking in the ETISEO ETI-VS1-BE-18-C4 video sequence.

Figure 9: People tracking on PETS View-001-S2L1 video sequence.

Multiple object tracking using k-shortest paths optimization.. *TPAMI*.

Chau, D., Bremond, F., and Thonnat, M. (2011). A multi-feature tracking algorithm enabling adaptation to context variations. *4th International Conference on Imaging for Crime Detection and Prevention 2011 (ICDP 2011)*, pages P30–P30.

Comaniciu, D., Ramesh, V., and Meer., P. (2003). Kernel-based object tracking. *TPAMI*, 5(25):564–577.

Dalal, N. and Triggs, B. (2005). Histograms of oriented gradients for human detection. *CVPR*, 1:886 –893 vol. 1.

Doledec, S. and Chessel, D. (1994). Co-inertia analysis: an alternative method for studying species-environment relationships. *Freshwater Biology*, 1:277–294.

Dray, S., Chessel, D., and Thioulouse, J. (2003). Co-inertia analysis amd the linking of ecological tables. *Ecology*.

Elgammal, A., Duraiswami, R., and Davis., L. S. (2003). Probabilistic tracking in joint feature-spatial spaces. *CVPR*, 1:I–781–I–788.

Etiseo. "http://www-sop.inria.fr/orion/ETISEO/".

Eveno, N. and Besacier, L. (2005). A speaker independent liveness test for audio-video biometrics. *9th European Conference on Speech Communication and Technology*, pages 232–239.

Feng, Z. R., Lu, N., and Jiang, P. (2008). Posterior probability measure for image matching. *Pattern Recognition*, 41:2422–2433.

Gittins, R. (1985). Canonical Analysis. *Springer-Verlag, Berlin, Germany,*.

Goecke, R. and Millar, B. (2003). Statistical Analysis of the Relationship between Audio and Video Speech Parameters for Australian English. *AVSP*.

Hager, G. D., Dewan, M., and Stewart., C. V. (2004). Multiple kernel tracking with ssd. *CVPR*, 1:I–790–I–797.

Henriques, J., Caseiro, R., and J., B. (2011). Globally optimal solution to multi-object tracking with merged measurements.. *ICCV*.

Heo, M. and Gabriel., K. (1997). A permutation test of association between configurations by means of the RV coefficient,. *Communications in Statistics - Simulation and Computation,*, 27:843–856.

Jiang, N., Liu, W., and Wu, Y. (2011). Adaptive and discriminative metric differential tracking. *CVPR 2011*, pages 1161–1168.

Kasturi, R. (2009). Framework for performance evaluation of face, text, and vehicle detection and tracking in video: Data, metrics, and protocol.. *TPAMI*.

N.Johnson and D.C.Hogg (1996). Learning the distribution of oject trajectories for event recognition. *Image and Vision computing*, 14:583–592.

P.Bilinski, F.Bremond, and M.Kaaniche. (2009). Multiple object tracking with occlusions using HOG descriptors and multi resolution images.. *ICDP, London (UK),*.

Shitrit, H., Berclaz, J., Fleuret, F., and P., F. (2011). Tracking multiple people under global appearance constraints.. *ICCV*.

skuldsson, A. H. (1988). Partial least square regression. *Journal of chemometrics*, 2:211–228.

Tucker, L. (1958). An inter-battery method of factor analysis. *Psychometrika*, 23:111–136.

Viola, P. and Wells., W. M. (1995). Alignment by maximization of mutual information. *ICCV*, page 0:16.

Yang, C., Duraiswami, R., and Davis., L. (2005). Efficient mean-shift tracking via a new similarity measure. *CVPR*, 1:176–183.

Zamir, A. R., Dehghan, A., and Shah, M. (2012). Gmcp-tracker: Global multi-object tracking using generalized minimum clique graphs. *ECCV*.