Subspace Learning for Background Modeling: A Survey

Thierry Bouwmans

Subspace Learning for Background Modeling: A Survey

T. Bouwmans

Laboratoire MIA, Université de La Rochelle, Avenue M. Crépeau, 17000 La Rochelle, France

Tel Phone: (33) 05.46.45.72.02

Email address: tbouwman@univ-lr.fr

Fax number: (33) 05.46.45.82.40

Abstract: Background modeling is often used to detect moving object in video acquired by a fixed camera. Recently, subspace learning methods have been used to model the background in the idea to represent online data content while reducing dimension significantly. The first method using Principal Component Analysis (PCA) was proposed by Oliver *et al.* [1] and a representative patent using PCA concerns the detection of cars and persons in video surveillance [2]. Numerous improvements and variants were developed over the recent years. The purpose of this paper is to provide a survey and an original classification of these improvements. Firstly, we classify the improvements of the PCA in term of strategies and the variants in term of the used subspace learning algorithms. Then, we present a comparative evaluation of the variants and evaluate them with the state-of-art algorithms (SG, MOG, and KDE) by using the Wallflower dataset.

Keywords: Background modeling, subspace learning, principal components analysis

Short Running Title: Subspace learning for background modeling

INTRODUCTION

In many video applications, moving objects must be detected like in video surveillance [3], optical motion capture [4-6] and multimedia [7-9]. The common way to extract moving objects is to use background subtraction methods. These methods have different steps: background modeling, background initialization, background maintenance and foreground detection as shown in Fig. (**1**).

**Fig. (1). Background Subtraction: The pipeline**
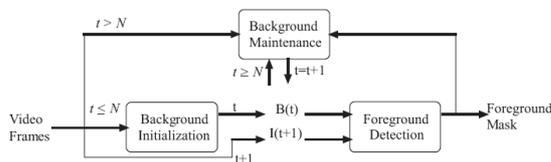
The background modeling describes the kind of model used to represents the background. Once the model has been chosen, the background model is initialized during a learning step by using N frames. Then, a first foreground detection is made and consists in the classification of the pixel as a background or as a foreground pixel. Thus, the foreground mask is applied on the current frame to obtain the moving objects. After this, the background is adapted over time following the changes which have occurred in the scene and so on. The background modeling is the key choice because it determines how the model will adapt to the critical situations [7, 10]: Noise image due to a poor quality image source, camera jitter, camera automatic adjustments, time of the day, light switch, bootstrapping, camouflage, foreground aperture, moved background objects, inserted background objects, multimodal background, waking foreground object, sleeping foreground object and shadows. These critical situations have different spatial and temporal properties. The main difficulties come from the illumination changes and dynamic backgrounds:

- Illumination changes appear in indoor and outdoor scenes. The Fig. (**2**). shows an indoor scene which presents a gradual illumination change. It causes false detections in several parts of the foreground mask as can be seen at the Fig. (**2d**). The Fig. (**3**). shows the case of a sudden illumination change due to a light on/off. As all the pixels are affected by this change, a big amount of false detections is generated (see Fig. (**3c**)).

- Dynamic backgrounds appear in outdoor scenes. The Fig. (**4**). shows four main types of dynamics

Fig. (2). Gradual illumination changes (Sequence Time of Day from [11])

| a) Dark | b) Middle | c) Clear | d) Foreground mask |
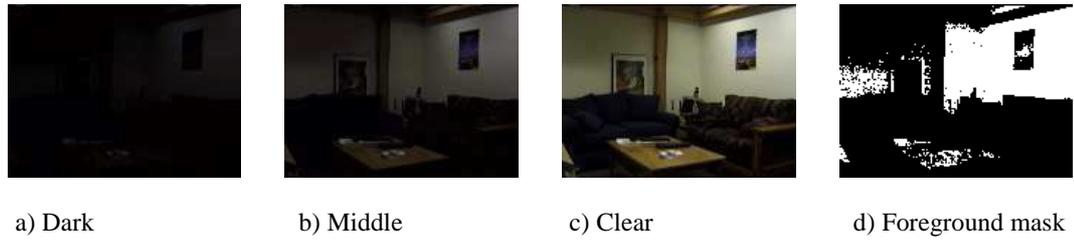
Fig. (3). Sudden illumination changes (Sequence Light Switch from [11])

| a) Light-on | b) Light-off | c) Foreground mask |
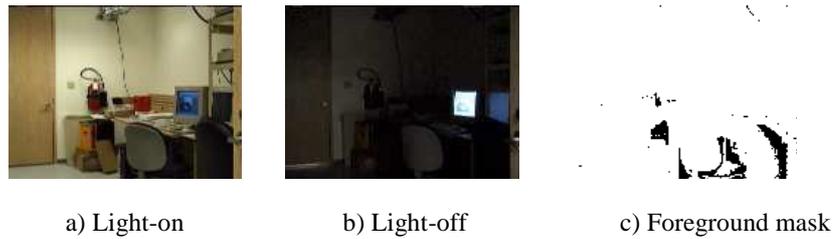
backgrounds: Camera jitter, waving trees, water rippling and water surface. The left column shows the original images and the right the foreground mask obtained by the MOG [12]. In each case, there is a big amount of false detections.

In the literature, many background modeling methods have been developed [13, 14] to be robust to these critical situations and can be classified in the following categories: Basic Background Modeling [15-17], Statistical Background Modeling [12, 18, 19], Fuzzy Background Modeling [20, 21], Background Estimation [10, 11, 22].

Reading the literature, one remark can be made: The statistical models offer more robustness to illumination changes and dynamic backgrounds. The simplest way to represent the background is to assume that the history over time of pixel's intensity values can be modeled by a Gaussian. Following this idea, Wren *et al*. [18] have proposed to use a single Gaussian (SG). Kim *et al*. [23] have generalized the SG using single general Gaussian (SGG) to alleviate the constraint of a strict Gaussian. However, a unimodal model cannot handle dynamic backgrounds when there are waving trees, water rippling or moving algae. To solve this problem, the Mixture of Gaussians (MOG) has been used to model dynamic backgrounds [12]. In the same idea, Allili *et al*. [24] have proposed the mixture of general Gaussians (MOGG). This model has some disadvantages. For example, background having fast variations cannot be accurately modeled with just a few Gaussians (usually 3 to 5), causing problems for sensitive detection. So, a non-parametric technique was developed for estimating background probabilities at each pixel from many recent samples over time using Kernel density estimation (KDE) [19] but it is time consuming. These five models consist in the first category models which are based on the Gaussian models. The second category uses more sophisticated statistical models as support vector machine (SVM) [25],

support vector regression (SVR) [26] and support vector data description (SVDD) [27].

Fig. (4). Dynamic backgrounds

a) Sequence Camera jitter from [28]
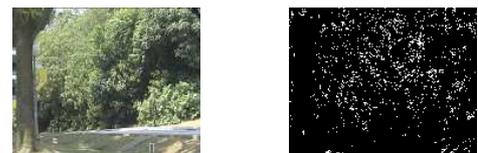
b) Sequence Campus from [29]

c) Sequence Water rippling from [29]

d) Sequence Water surface from [29]

Table 1. Statistical Background Modeling: An Overview

Categories	Methods	Authors - Dates
Gaussian Models	Single Gaussian (SG) (5) Single General Gaussian (SGG) (3) Mixture of Gaussians (MOG) (~100) Mixture of General Gaussians (MOGG) (3) Kernel Density Estimation (KDE) (21)	Wren *et al.* (1997) [18] Kim *et al.* (2007) [23] Stauffer and Grimson (1999) [12] Allili *et al.* (2007) [24] Elgammal *et al.* (2000) [19]
Support Vector Models	Support Vector Machine (SVM) (3) Support Vector Regression (SVR) (2) Support Vector Data Description (SVDD) (5)	Lin *et al.* (2002) [25] Wang *et al.* (2006) [26] Tavakkoli *et al.* (2006) [27]
Subspace Learning Models	Principal Components Analysis (SL-PCA) (15) Independent Component Analysis (SL-ICA) (2) Incremental Non Negative Matrix Factorization (SL-INMF) (2) Incremental Rank-(R_1,R_2,R_3) Tensor (SL-IRT) (1)	Oliver *et al.* (1999) [1] Yamazaki *et al.* (2006) [30] Bucak *et al.* (2007) [31] Li *et al.* (2008) [32]

The third category employs Subspace Learning methods. In [1], Subspace Learning using Principal Component Analysis (SL-PCA) is applied on N images to construct a background model, which is represented by the mean image and the projection matrix comprising the first p significant eigenvectors of PCA. In this way, foreground segmentation is accomplished by computing the difference between the input image and its reconstruction. In the same idea, many improvements of SL-PCA [10] were developed to be more robust and fast [33-45]. In the same category, Yamazaki *et al.* [30] and Tsai *et al.* [46] have used an Independent Component Analysis (SL-ICA). In another way, Bucak et al. [31, 47] have proposed an Incremental Non-negative Matrix Factorization (SL-INMF) to reduce the dimension. In order to take into account the spatial information, Li *et al.* [32] have used an Incremental Rank-(R_1,R_2,R_3) Tensor (SL-IRT). The Table **1** shows an overview of the statistical background modeling methods. The first column indicates the category model and the second column the name of each method. Their corresponding acronym is indicated in the first parenthesis and the number of papers counted for each method in the second parenthesis. The third column gives the name of the authors and the date of the related publication. The Gaussian models and support vector models are greatly designed for dynamics backgrounds and subspace learning models for illumination changes.

In this paper, we present a survey on subspace learning approaches which have attracted much attention recently to deal with illumination changes. These approaches can be divided into improvements and variants of PCA. The improvements consist to enhance the adaptation and the robustness by using incremental and robust PCA algorithms [33-45]. The variants consist to use an other subspace learning algorithms as the Independent Component Analysis (ICA) [30, 46], Incremental Non-negative Matrix Factorization (INMF) [31, 47] and Incremental Rank-(R_1,R_2,R_3) Tensor (IRT) [32].

The rest of this paper is organized as follows: In the Section 2, we firstly provide a background on subspace learning methods. Then, in the section 3, we remind the original PCA method proposed by Oliver *et al.* [1]. Then, we classify the modified versions found in the literature. Then,

the Sections 4, 5 and 6 review respectively SL-ICA, SL-INMF and SL-IRT algorithms applied to background modeling. Finally, in the Section 7, we present a comparative evaluation of the SL-PCA's variants with the state-of-art algorithms (SG, MOG, KDE) by using the Wallflower dataset.

2. BACKGROUND ON SUBSPACE LEARNING METHODS

The key problem in many data-analysis tasks is to find a suitable representation of the data. A useful representation can be obtained by reducing the dimensionality of the data so that further computational methods can be applied. Subspace learning methods have been developed in this context and so consist in reducing the dimension of the space significantly. The different methods have been classified by Skocaj and Leonardis [48] as reconstructive and discriminative methods:

- ***Reconstructive subspace learning:*** The reconstructive methods allow a well approximation of data and so provide a good reconstruction. Another advantage is that reconstructive methods are unsupervised techniques. Furthermore, reconstructive methods enable incremental updating which is very suitable for real-time application. These methods are task-independents. The most common reconstructive methods are the following: Principal Components Analysis (PCA) [49], Independent Component Analysis (ICA) [50] and Non-negative Matrix Factorization (NMF) [51]. PCA transforms a number of possibly correlated data into a smaller number of uncorrelated data called principal components. ICA is a variant of PCA in which the components are assumed to be mutually statistically independent instead of merely uncorrelated. The stronger condition allows to remove the rotational invariance of PCA, i.e. ICA provides a meaningful unique bilinear decomposition of two-way data that can be considered as a linear mixture of a number of independent source signals.

Table 2. Subspace Learning Methods: An Overview

Categories	Methods	Authors - Dates
Reconstructive methods	Principal Components Analysis (PCA)	Joliffe (2002) [49]
	Independent Component Analysis (ICA)	Hyvarinen and Oja (2000) [50]
	Non-negative Matrix Factorization (NMF)	Lee and Seung (2000) [51]
Discriminative methods	Linear Discriminant Analysis (LDA)	Tang and Tao (2006) [52]
	Canonical Correlation Analysis (CCA)	Hardoon et al. (2004) [53]

Non-negative matrix factorization (NMF) finds linear representations of non-negative data. Given a non-negative data matrix V, NMF finds an approximate factorization $V \approx WH$ into non-negative factors W and H. The non-negativity constraints make the representation purely additive, i.e allowing no subtractions, in contrast to PCA and ICA.

- Discriminative subspace learning: The discriminative methods are supervised techniques and allow a well separation of data and so provide a good classification. Furthermore, discriminative methods are spatially and computationally efficient. These methods are task-dependents. The most common discriminative methods are the following: Linear Discriminant Analysis (LDA) [52] and Canonical Correlation Analysis (CCA) [53]. LDA projects the data onto a lower-dimensional vector space such that the ratio of the between-class distance to the within-class distance is maximized. The goal is to achieve maximum discrimination. Canonical correlation analysis is a multivariate statistical model that facilitates the study of interrelationships among sets of multiple dependent variables and multiple independent variables. Canonical correlation simultaneously predicts multiple dependent variables from multiple independent variables.

The Table **2** shows an overview of the common subspace learning methods. The first column indicates the category and the second column the name of each method. Their corresponding acronym is indicated in the first parenthesis. The third column gives for the related recent survey the name of the author and the date.

All these methods are originally implemented with batch algorithms which require that the data must be available in advance and be given once altogether. However, this type of batch algorithms is not adapted for the application of background modeling in which the data are incrementally received from the camera. Furthermore, when the dimension of the dataset is high, both the computation and storage complexity grow dramatically. Thus, incremental methods are highly needed to compute in real-time the adaptive subspace for the data arriving sequentially [54]. Following these constraints, the reconstructive methods are the most adapted for background modeling. Furthermore, their unsupervised aspect allows to avoid a manual intervention in the learning step.

In the following sections, we survey the subspace leaning methods applied to background modeling: Principal Components Analysis, Independent Component Analysis , Non-negative Matrix Factorization and Incremental Rank-(R_1,R_2,R_3) Tensor. These four methods are reconstructive ones.

3. SUBSPACE LEARNING VIA PCA (SL-PCA)

3.1 Principle

Oliver *et al.* [1] have proposed to model each background pixel using an eigenbackground model. This model consists in taking a sample of N images $\{I_1, I_2,..., I_N\}$ and computing the mean background image μ_B and its covariance matrix C_B. This covariance matrix is then diagonalized using an eigenvalue decomposition as follows:

$$L_B = \Phi_B C_B \Phi_B^T \quad (1)$$

where Φ_B is the eigenvector matrix of the covariance of the data and L_B is the corresponding diagonal matrix of its eigenvalues.

In order to reduce, the dimensionality of the space, only M eigenvectors (M<N) are kept in a principal component analysis (PCA). The M largest eigenvalues are contained in the matrix L_M and the M vectors correspond to these M largest eigenvalues in the matrix Φ_M.

Once the eigenbackground images stored in the matrix Φ_M are obtained and the mean μ_B too, the input image I_t can be approximated by the mean background and weighted sum of the eigenbackgrounds Φ_M.
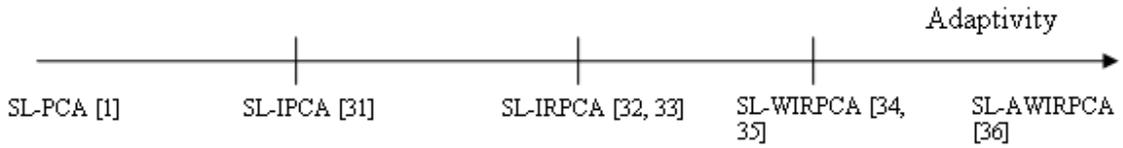
Fig. (5). Adaptivity of the SL-PCA Algorithms

So, the coordinate in eigenbackground space of input image I_t can be computed as follows:

$$w_t = (I_t - \mu_B)^T \Phi_M \qquad (2)$$

When w_t is back projected onto the image space, a reconstructed background image is created as follows:

$$B_t = \Phi_M w_t^T + \mu_B \qquad (3)$$

Then, the foreground object detection is made as follows:

$$\left| I_t - B_t \right| > T \qquad (4)$$

where T is a constant threshold.

3.2 Improvements

The eigenbackground model which we have called SL-PCA provides a robust model of the probability distribution function of the background, but not of the moving objects while they do not have a significant contribution to the model. So, the first limitation of this model is that the size of the foreground object must be small and don't appear in the same location during a long period in the training sequence. The second limitation appears for the background maintenance. Indeed, it is computationally intensive to perform model updating using the batch mode PCA. Moreover without a mechanism of robust analysis, the outliers or foreground objects may be absorbed into the background model [55-57]. The third limitation is that the application of this model is mostly limited to the gray-scale images since the integration of multi-channel data is not straightforward. It involves much higher dimensional space and causes additional difficulty to manage data in general. Thus, the different improvements founds in the literature attempt to solve these three limitations:

-   ***Alleviate the limitation of the size of the foreground object:*** Xu *et al*. [33, 34] proposed to apply recursively an error compensation process which reduces the influence of foreground moving objects on the eigenbackground model. An adaptive threshold method is also introduced for background subtraction, where the threshold is determined by combining a fixed global threshold and a variable local threshold. Results show more robustness in presence of moving objects. Another approach developed by Kawabata *et al*. [35] consists in an iterative optimal projection method to estimate a varied background in real time from a dynamic scene with foreground. Firstly, background images are collected for a while and then the background images are compressed using eigenspace method to form a database. After this initialization, a new image is taken and projected onto the eigenspace to estimate the background. As the estimated image is much affected by the foreground, the foreground region is calculated by using background subtraction with former estimated background to exclude the region from the projection. Thus the image whose foreground region is replaced by the former background is projected to eigenspace and then the background is updated. Kawabata et al. [35] proved that the cycle converges to a correct background image.

-   ***Dealing with the time requirement and the robustness:*** For the maintenance, some authors [36-44] have proposed different algorithms of incremental PCA. The incremental PCA proposed by [36] need less computation but the background image is contaminated by the foreground object. To solve this, Li *et al*. [37, 38] have proposed an incremental PCA which is robust in presence of outliers. However, when keeping the background model updated incrementally, it assigned the same weights to the different frames. Thus, clean frames and frames which contain foreground objects have the same contribution. The consequence is a relative pollution of the background model. In this context, Skocaj *et al*. [39, 40] used a weighted incremental and robust. The weights are different following the frame and this method achieved a better background model. However, the weights were applied to the whole frame without considering the contribution of different image parts to building the background model. To achieve a pixel-wise precision for the weights, Zhang and Zhuang [41] have proposed an adaptive weighted selection for an incremental PCA. This method performs a better model by assigning a weight to each pixel at each new frame during the update. Experiments [41] show that this method achieves better results than the SL-IRPCA [37, 38]. Wang *et al*. [42, 43] used a similar approach

using the sequential Karhunen-Loeve algorithm. All these incremental methods avoid the eigen-decomposition of the high dimensional covariance matrix using approximation of it and so a low decomposition is allowed at the maintenance step with less computational load. However, these incremental methods maintain the whole eigenstructure including both the eigenvalues and the exact matrix Φ_M. To solve it, Li *et al*. [44] have proposed a fast recursive and robust eigenbackground maintenance avoiding eigen-decomposition. This method achieves similar results than the SL-IPCA [36] and the SL-IRPCA [37, 38] at better frames rates. The Fig. (**5**). shows a classification of these algorithms following their robustness and their adaptivity.

- ***Dealing with the grey scale & the pixel-wise limitations:*** Han and Jain [45] have proposed an efficient algorithm using a weighted incremental 2-Dimensional Principal Component Analysis. It is shown that the principal components in 2D PCA are computed efficiently by transformation to standard PCA. To perform the computational time, Han and Jain [45] have used an incremental algorithm to update eigenvectors to handle temporal variations of background. The proposed algorithm was applied to 3-channel (RGB) and 4-channel (RGB+IR) data. Results [45] show noticeable improvements in presence of multimodal backgrounds and shadows. To solve the pixel-wise limitation, Zhao *et al*. [58] have employed spatio-temporal block instead of pixel. It is more robust to noise than the pixel wise approach [59]. Furthermore, their method consists in applying the candid covariance free incremental principal components analysis algorithm (CCIPCA) [60] which is fast in convergence rate and low in computational complexity than classical IPCA algorithms. Results [58] show more robustness robust to noise and fast lighting changes.

The Table **3**, Table **4** and Table **5** group by type these different improvements of the SL-PCA.

Table 3. Size of the foreground objects

Methods	Authors - Dates
Recursive Error Compensation (SL-REC-PCA)	Xu *et al*. (2006) [33, 34]
Iterative Optimal Projection (SL-IOP-PCA)	Kawabata *et al*. (2006) [35]

Table 4. Time requirement and the robustness

Methods	Authors - Dates
Incremental PCA (SL-IPCA)	Rymel *et al*. (2004) [36]
Incremental and robust PCA (SL-IRPCA)	Li *et al*. (2003) [37, 38]
Weighted Incremental and Robust PCA (SL-WIRPCA)	Skocaj *et al*. (2003) [39, 40]
Adaptive Weight Selection for Incremental PCA (SL-AWIPCA)	Zhang and Zhuang (2007) [41]
Sequential Karhunen-Loeve algorithm (SL-SKL-PCA)	Wang *et al.* (2006) [42, 43]
Fast Recursive Maintenance (SL-FRM-PCA)	Li *et al.* (2006) [44]

Table 5. Dealing with the grey scale & the pixel-wise limitations

Methods	Authors - Dates
Weighted Incremental 2D PCA (SL-WI2DPCA)	Han and Jain (2007) [45]
Candid Covariance Incremental PCA (SL-CCIPCA)	Zhao *et al.* (2008) [58]

4. SUBSPACE LEARNING VIA ICA (SL-ICA)

4.1 Principle

ICA generalizes the technique of PCA and has proven to be a good tool of feature extraction. When some mixtures of probabilistically independent source signals are observed, ICA recovers the original source signals from the observed mixtures without knowing how the sources are mixed. The assumption made is that the observation vectors $X = (x_1, x_2, ..., x_M)^T$ can be represented in terms of a linear superposition of unknown independent vectors $S = (s_1, s_2, ..., s_N)^T$:

$$X = AS \qquad (5)$$

where A is an unknown mixing matrix (M×N). The goal of ICA is to find a matrix W, so that the resulting vectors:

$$Y = WX \qquad (6)$$

recover the independent vectors S, probabilistically permuted and rescaled. W is roughly the inverse matrix of A.

4.2 Application to background modeling

In the context of background modeling, the ICA model is given by:

$$Y = WX_t \qquad (7)$$

where $X_t = (x_B, x_F)^T$ is the mixture data matrix of size 2*K in which K=M*N where M and N are the column and the row of the images of the sequence. $x_1 = (x_{11}, x_{12}, ..., x_{1K})$ is the first frame which can contain or not foreground objects and $x_2 = (x_{21}, x_{22}, ..., x_{2K})$ is the second frame which contains foreground objects. $W = (w_1, w_2)^T$ is the de-mixing matrix, in which $w_i = (w_{i1}, w_{i2})$ with i=1,2. $Y = (y_1, y_2)^T$ is the estimated source signals in which $y_i = (y_{i1}, y_{i2}, ..., y_{ik})$. Several ICA algorithms can be used to determine W. Yamazaki *et al.* [30] used a neural learning algorithm [61]. Tsai and Lai [46] used a Particle Swarm Algorithm (PSO) [62]. Once W is determined, there are two cases in the literature:

- The first case where x_1 contains foreground object like in Yamazaki et al. [30]. Then, the foreground mask for the frames x_1 and x_2 is obtained by thresholding respectively y_1 and y_2. The background image is obtained by replacing regions representing foreground objects in x_1 by the corresponding regions representing background in x_2.

- The second case where x_1 contains no foreground object like in Tsai and Lai [46]. Then, the foreground mask for the frames x_2 is obtained by thresholding y_2. The background image is y_1.

The ICA model has been evaluated on traffic scenes in [30] and has shown robustness in changing background like illumination changes. In [46], the algorithm has been tested on indoor scenes which present illumination changes too.

5. SUBSPACE LEARNING VIA INMF (SL-INMF)

5.1 Principle

The aim of non-negative matrix factorization (NMF), with rank r, is to decompose the data matrix $V \in R^{p \times q}$ into two matrices which are $W \in R^{p \times r}$ called the mixing matrix, and $H \in R^{r \times q}$ named as the encoding matrix:

$$V \approx WH \qquad (8)$$

So, the goal of the NMF consists in finding an approximate factorization that minimizes the reconstruction error. Different cost functions based on the reconstruction error have been defined in the literature, but because of its simplicity and effectiveness, the squared error is the most used:

$$F = \|V - WH\|^2 = \sum_{i=1}^{p} \sum_{i=1}^{q} (V_{ij} - (WH)_{ij})^2 \qquad (9)$$

where subscription ij stands for the ij[th] matrix entity.

5.2 Application to background modeling

Bucak *et al.* [31, 47] have proposed an incremental NMF algorithm to model the background. The initialization is made using N training frames. So, V is a vector column corresponding to a matrix of size $(p \times q) \times N$ where p and q are respectively the column and the row of the images. The matrices W and H are updated incrementally. The foreground detection is made by thresholding the residual error which corresponds to the deviation between the background model and the projection of the current frame onto the background model. The SL-INMF has similar performance to dynamic background and illumination changes than the SL-IPCA proposed by Li et al [37].

6. SUBSPACE LEARNING VIA INCREMENTAL RANK-(R1, R2, R3) TENSOR (SL-IRT)

The different previous subspace learning methods consider image as a vector. So, the local spatial information is almost lost. To solve this problem, Li *et al.* [32] have proposed to use a high-order tensor learning algorithm called incremental rank-(R$_1$,R$_2$,R$_3$) tensor based subspace learning. This online algorithm builds a low-order tensor eigenspace model in which the mean and the eigenbasis are updated adaptively. Denote $G = \left\{ BM_q \in R^{M \times N} \right\}_{q=1,2,...,t}$ as a scene's background appearance sequence with the q-th frame being BM_q. M and N are respectively the column and the row of the images. Denote p_{uv} as the u-th and the v-th pixel of the scene. The tensor-based eigenspace model for an existing $A = \left\{ BM_q^{uv} \in R^{I_1 \times I_2 \times t} \right\}_{q=1,2,...,t}$ ($I_1 = I_2 = 5$ corresponding to a K neigborhood of p_{uv} with $K = I_1 I_2 - 1 = 24$) consists of the maintained eigenspace dimensions (R$_1$,R$_2$,R$_3$) corresponding to the three tensor unfolding modes, the mode-n column projection matrices $U^{(n)} \in R^{I_n \times R_n}$, the mode-3 row projection matrix $V^{(3)} \in R^{(I_n . I_2) \times R_3}$, the column means $\overline{L}^{(1)}$ and $\overline{L}^{(2)}$ of the mode-(1,2) unfolding matrices $A_{(1)}$ and $A_{(2)}$, and the row mean $\overline{L}^{(3)}$ of the mode-3 unfolding

matrix $A_{(3)}$. Given the K-neighbor image region $I_{t+1}^{uv} \in R^{I_1 \times I_2 \times 1}$ centered at the u-th and v-th pixel p_{uv} of the current incoming frame $I_{t+1} \in R^{M \times N \times 1}$, the distance RM_{uv} (determined by the three reconstruction error norms of the three modes) between I_{t+1}^{uv} and the learned tensor-based eigenspace model is computed.

Then, Li *et al.* [32] made the foreground detection as follows:

p_{uv} is classified as background if $\exp(-\dfrac{RM_{uv}^2}{2\sigma^2}) > T$

p_{uv} is classified as foreground otherwise

where σ is a scaling factor and T denotes a threshold. Thus, the new background model $BM_{t+1}(u,v)$ at time t+1 is defined by Li *et al.* [32] as follows:

$BM_{t+1}(u,v) = H_{uv}$ if p_{uv} is classified as foreground

$BM_{t+1}(u,v) = I_{t+1}(u,v)$ otherwise

where $H_{uv} = (1-\alpha)MB_t(u,v) + \alpha I_{t+1}(u,v)$, MB_t is the mean matrix of $BM_{1:t}$ at time t and α is a learning rate factor.

Then, the tensor eigenspace model is updated incrementally and so on. The SL-IRT shows more robustness to noise than the SL-IPCA proposed by Li *et al.* [37].

7. PERFORMANCE EVALUATION

For the performance evaluation, we have compared the original version of the subspace learning models with the state-of-art algorithms which are composed by three gaussian models:

- **Single Gaussian:** The mean and covariance of pixel values are updated continuously [18]. Foreground detection is made by thresholding the difference between the current mean and the current value.
- **Mixture of Gaussians:** A mixture of K Gaussians models the background [12]. Each Gaussian is weighted according to the frequency with which it represents the observed background. The most heavily weighted Gaussians that together gives a sum over 50% of past data are considered background.
- **Kernel Density Estimation:** This model keeps a sample of intensity values for each pixel in the image and uses this sample to estimate the probability density function of the pixel intensity [19]. The density function
- is estimated using kernel density estimation technique. The foreground detection is made by thresholding the probability.

7.1 Wallflower dataset

For a comparative evaluation, several datasets exist and the most used is the Wallflower dataset provided by Toyama *et al.* [11]. It consists in a set of image sequences where each sequence presents a different type of difficulty that a practical task may meet. The performance is evaluated against hand-segmented ground truth. Three terms are used in evaluation: The false positive (FP) is the number of background pixels that are wrongly marked as foreground. The false negative (FN) is the number of foreground pixels that are wrongly marked as background. The total error (TE) is the sum of FP and FN. A brief description of the Wallflower image sequences can be made as follows:

- **Moved Object:** A person enters into a room, makes a phone call, and leaves. The phone and the chair are left in a different position.
- **Time of Day:** The light in a room gradually changes from dark to bright. Then, a person enters the room and sits down.
- **Light Switch:** A room scene begins with the lights on. Then a person enters the room and turns off the lights for a long period. Later, a person walks in the room, switches on the light, and moves the chair, while the door is closed.
- **Waving Trees:** A tree is swaying and a person walks in front of the tree.
- **Camouflage:** A person walks in front of a monitor, which has rolling interference bars on the screen. The bars include similar color to the person's clothing.
- **Bootstrapping:** The image sequence shows a busy cafeteria and each frame contains people.
- **Foreground Aperture:** A person with uniformly colored shirt wakes up and begins to move slowly.

7.2 Experiments and results

The Table **6** shows the performance in term of FP, FN and TE for each algorithm. The corresponding results are shown in Table **7**. The Fig. (**6**). shows the performance in term of FP and FN for each algorithm.

7.2.1 Gaussians models versus subspace learning models

From Table **7**, we can see that the Gaussian models give the biggest total errors with TE>20 000 instead of the subspace learning models with TE<20 000. The performance of the Gaussian models is due to their poor results on the sequence Light Switch. This is confirmed by the Fig. (**7**). which shows the performance without the sequence Light Switch. In this case, the best results are given by the MOG and the KDE which are better for multimodal backgrounds as shown for the sequence Waving Trees in Table **6**.

Table 6. Results on the Wallflower dataset

Sequence	Moved Objects	Time of Day	Light Switch	Waving Trees	Camou-flage	Boot-strap	Foreg. Aperture
Test image							
Ground truth							
SG Wren *et al.* [18]							
MOG Stauffer *et al.* [12]							
KDE Elgammal *et al.* [19]							
SL-PCA Oliver *et al.* [1]							
SL-ICA Tsai and Lai [46]							
SL-INMF Bucak *et al.* [31]							
SL-IRT Li *et al.* [32]							

Table 7. Performance Evaluation on the Wallflower dataset

Algorithm	Error Type	Moved Object	Time of Day	Light Switch	Waving Trees	Camou-flage	Bootstrap	Foreground Aperture	Total Errors	TE without LS	TE without C
SG Wren *et al.* [18]	FN	0	949	1857	3110	4101	2215	3464			
	FP	0	535	15123	357	2040	92	1290	35133	18153	28992
MOG Stauffer *et al.* [12]	FN	0	1008	1633	1323	398	1874	2442			
	FP	0	20	14169	341	3098	217	530	27053	11251	23557
KDE Elgammal *et al.* [19]	FN	0	1298	760	170	238	1755	2413			
	FP	0	125	14153	589	3392	933	624	26450	11537	22175
SL-PCA Oliver *et al.* [1]	FN	0	879	962	1027	350	304	2441			
	FP	1065	16	362	2057	1548	6129	537	17677	16353	15779
SL-ICA Tsai and Lai [46]	FN	0	1199	1557	3372	3054	2560	2721			
	FP	0	0	210	148	43	16	428	15308	13541	12211
SL-INMF Bucak *et al* [31].	FN	0	724	1593	3317	6626	1401	3412			
	FP	0	481	303	652	234	190	165	19098	17202	12238
SL-IRT Li *et al* [47].	FN	0	1282	2822	4525	1491	1734	2438			
	FP	0	159	389	7	114	2080	12	17053	13842	15448

Fig. (6). Overall performance
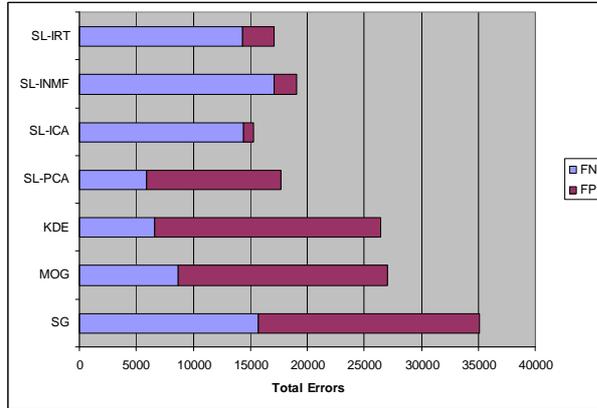
**Fig. (7). Overall performance without Light Switch**
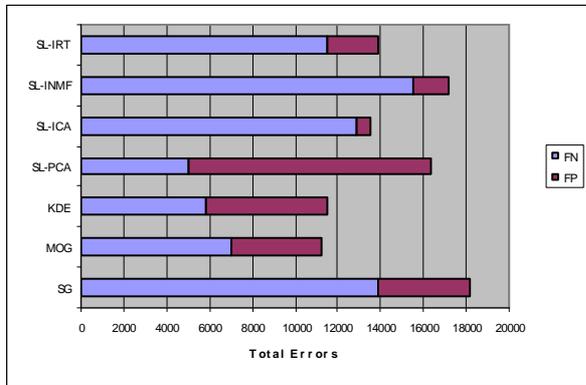
7.2.2 Subspace learning models

The subspace learning algorithms have different performances in term of TE. From Table **7,** we can see that the SL-ICA gives the smallest TE followed by the SL-IRT, SL-PCA and SL-INMF. This ranking has to be taken with precaution because a poor performance on one video influences the TE and then modifies the rank. The main interpretation is that some of them are more or less adapted for specific situations. For example, only the SL-PCA gives FP in the sequence Moved Objects due the fact that the model is not update overtime. In the same way, the SL-INMF gives the biggest total error due to its results on the sequence Camouflage. This is confirmed by the Fig. (**8**). which shows the performance without this sequence Camouflage. In this case, the SL-INMF is the second in term of performance. Furthermore, the SL-INMF is more robust than the SL-IPCA [37] in the case of illumination changes (see [31, 47]). The SL-ICA has globally good performance except for the sequence Boostrap by giving less true detection. The SL-IRT seems to be more efficient in the case of camouflage. The SL-PCA gives less FN than FP. For the SL-ICA, SL-INMF and SL-IRT, it is the contrary. We can remark that the SL-ICA provides very less FP than FN. It is

interesting in video-surveillance because it decreases false alarms.

Fig. (8). Overall performance without Camouflage
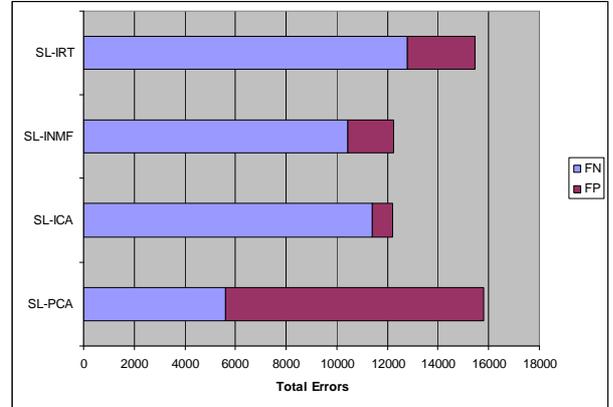
In resume, the actual subspace learning methods which are used in background modeling outperform the state-of-art algorithms in the case of illumination changes but can't deal with multimodal backgrounds. So, it will be interesting to investigate how to take into account robustly the multimodality with a subspace learning methods.

CURRENT & FUTURE DEVELOPMENTS

In this paper, we have attempted to provide a complete survey on background modeling methods based on subspace learning. Thus, we have proposed a classification of improvements and variants of the PCA algorithm proposed by Oliver *et al*. [1]. This survey has shown that subspace learning via PCA and its improvements are the most used. ICA, INMF and IRT have proved their efficiency in the case of illumination changes. The performance evaluation has shown that subspace learning models outperform some state-of art algorithms in the case of illumination changes.

For future investigations, discriminative subspace learning methods like LDA and CCA seem to be very interesting approaches. For example, LDA exists in several incremental versions as incremental LDA using maximum margin criterion [63] or using fixed point method [64] or sufficient spanning set approximations [65]. In the same way, Partial Least Squares (PLS) methods [66] give a nice perspective to model robustly the background.

ACKNOWLEDGMENTS

University, Taiwan) who have provided the results obtained by their algorithm called SL-ICA in this paper.

- Radu VASIU (University of Timisoara) who has provided an ERASMUS student and Cristina MARGHES who has implemented the algorithm called SL-IRT [32].

CONFLICT OF INTEREST:

The author declared no conflict of interest

REFERENCES

[1] Oliver N, Rosario B, Pentland A. A bayesian computer vision system for modeling human interactions. *Inter Conf on Vision Systems, ICVS 1999,* pages 255-272, Gran Canaria, Spain, January 1999.

[2] Ozer I., US7200266 (**2003**).

[3] Cheung S, Kamath C. Robust background subtraction with foreground validation for urban traffic video. J App Sig Proc EURASIP 2005, Issue 14, pages 2330-2340.

[4] Carranza J, Theobalt CM, Magnor, HS. Free-viewpoint video of human actors. ACM Trans Grap 2003; 22(3): 569-577.

[5] Horprasert TI, Haritaoglu CW, Harwood DL, Davis AP. Real-time 3D motion capture. *Workshop on Perceptual User Interfaces,* PUI 1998, San Francisco, California, pages 87-90.

[6] Mikic I, Trivedi M, Hunter E, Cosman P. Human body model acquisition and tracking using voxel data. Inter J Comp Vision 2003; 199-223.

[7] El Baf F, Bouwmans T, Vachon B. Comparison of background subtraction methods for a multimedia learning space. *Inter Conf on Signal Processing and Multimedia, SIGMAP 2007*, Barcelona, Spain, July 2007.

[8] Pande A, Verma A, Mittal A. Network aware optimal resource allocation for e-learning Videos. *The 6th inter conf on mobile Learning,* Melbourne Australia, October 2007.

[9] Warren J. Unencumbered full body interaction in video games. Thesis, MFA Design and Technology Parsons School of Design, New York, USA, April 2003.

[10] Chang R, Ghandi T, Trivedi, M. Vision modules for a multi sensory bridge monitoring approach. ITSC 2004: 971-976

[11] Toyama K, Krumm J, Brumitt B, Meyers B. Wallflower: Principles and practice of background maintenance. *Inter Conf on Computer Vision 1991*: 255-261, Corfu, Greece, September 1999.

[12] Stauffer C, Grimson W. Adaptive background mixture models for real-time tracking. *IEEE Conf on CVPR 1999*: 246-252.

[13] Piccardi M. Background subtraction techniques: A review. *The Inter Conf on Systems, Man and Cybernetics, SMC 2004*, pages 3199-3104, The Hague, The Netherlands, October 2004.

[14] Elhabian S, El-Sayed K, Ahmed S. Moving object detection in spatial domain using background removal techniques - state-of-art. CSENG 2008; 1(1): 32-54.

[15] Lee B, Hedley M. Background estimation for video surveillance. IVCNZ 2002: 315-320.

[16] McFarlane N, Schofield C. Segmentation and tracking of piglets in images. BMVA 1995: 187-193.

[17] Zheng J, Wang Y, Nihan N, Hallenbeck E. extracting roadway background image: A mode based approach. J Transport Res Report 2006; 1944: 82-88.

[18] Wren C, Azarbayejani A, Darrell T, Pentland A. Pfinder: Real-time tracking of the human body. IEEE Trans on Pat Analysis and Mach Intel 1997; 19(7): 780-785.

[19] Elgammal A, Harwood D, Davis L. Non-parametric model for background subtraction. *6th European Conf on Computer Vision, ECCV 2000*, pages 751-767, Dublin, Ireland, June 2000.

[20] Sigari M, Mozayani N, Pourreza H. Fuzzy running average and fuzzy background subtraction: concepts and application. Inter J of Computer Science and Network Security 2008; 8(2): 138-143.

[21] El Baf F, Bouwmans T, Vachon B. Type-2 fuzzy mixture of Gaus-sians model: Application to background modeling. *Inter Symposium on Visual Computing, ISVC 2008*, pages 772-781, Las Vegas, USA, December 2008.

[22] Messelodi S, Modena C, N. Segata, M. Zanin. A Kalman filter based background updating algorithm robust to sharp illumination changes. ICIAP 2005; 3617: 163-170.

[23] Kim H, Sakamoto R, Kitahara I, Toriyama T, Kogure K. Robust silhouette extraction technique using background subtraction. *10th Meeting on Image Recognition and Understand, MIRU 2007*, Hiroshima, Japan, July 2007.

[24] Allili M, Bouguila N, Ziou D. A robust video foreground segmentation by using generalized gaussian mixture modelling. *Fourth Canadian Conf on computer and robot vision, CRV 2007:* 503-509.

[25] Lin H, Liu T, Chuang J. A probabilistic SVM approach for background scene initialization. ICIP 2002; 3: 893-896.

[26] Wang J, Bebis G, Miller R. Robust video-based surveillance by integrating target detection with tracking. *IEEE Workshop on Object Tracking and Classification Beyond the Visible Spectrum in conjunction with CVPR 2006,* pages 137-142, New York, NY, June 2006.

[27] Tavakkoli A, Nicolescu M, Bebis G. A novelty detection approach for foreground region detection in videos with quasi-stationary backgrounds ISVC 2006, pages 40-49, Lake Tahoe, USA, November 2006.

[28] Bhandarkar S., Luo X., Fast and Robust Background Updating for Real-time Surveillance Systems. CVPR 2005, pages 1-6, San Diego, CA, June 2005.

[29] Li L., Gu I., Huang W., Tian. Q. Statistical modeling of complex background for foreground object detection. IEEE Transaction on Image Processing, 13(11):1459–1472

[30] Yamazaki M, Xu G, Chen Y. Detection of moving objects by independent component analysis. ACCV 2006; 467-478.

[31] Bucak S, Gunsel B, Gursoy O. Incremental non-negative matrix factorization for dynamic background modelling. *Inter Workshop on Pattern Recog in Information Sys*, Funchal, Portugal, June 2007.

[32] Li X, Hu W, Zhang Z, Zhang X. Robust foreground segmentation based on two effective background models MIR 2008: 223-228, Vancouver, Canada, October 2008.

[33] Xu Z, Shi P, Gu I. An eigenbackground subtraction method using recursive error compensation. Advances in Multimedia Information Proc PCM 2006; 4261: 779-787.

[34] Xu Z, Gu I, Shi P. Recursive error-compensated dynamic eigenbackground learning and adaptive background subtraction in video. Optical Eng 2008; 47(5).

[35] Kawabata S, Hiura S, Sato K. Real-time detection of anomalous objects in dynamic scene. *18th Inter Conf on Pattern Recognition, ICPR 2006*, Volume 3, pages 1171 - 1174, Hong-Kong, August 2006.

[36] Rymel J, Renno J, Greenhill D, Orwell J, Jones G. Adaptive eigen-backgrounds for object detection. *IEEE Inter Conf on Image Processing, ICIP 2004*, Suntec City, Singapore, October 2004.

[37] Li Y, Xu L,Morphett J, Jacobs R. An Integrated algorithm of incremental and robust PCA. *IEEE Inter Conf on Image Processing, ICIP 2003,* pages 245-248, Barcelona, Spain, September 2003.

[38] Li Y. On incremental and robust subspace learning. Pattern Recog PR 2004: 37(7): 1509-1518.

[39] Skocaj D, Leonardis A. Weighted and robust incremental method for subspace learning. *Inter Conf on Computer Vision, ICCV 2003*, pages 1494-1501, 2003.

[40] Skocaj D, Leonardis A. Incremental and robust learning of subspace representations. Image and Vision Computing IVC 2006: 1-12.

[41] Zhang J, Zhuang Y. Adaptive weight selection for incremental eigen-background modelling. ICME 2007, pages 851-854, Beijing, China, July 2007.

[42] Wang L, Wang L, Zhuo Q, Xiao H, Wang W. Adaptive eigenbackground for dynamic background modelling. Intelligent Computing in Signal Processing and Pattern Recognition, Lecture Notes in Control and Information Sciences, Volume 345, pages 670-675, 2006.

[43] Wang L, Wang L, Wen M, Zhuo Q, Wang W. Background subtraction using incremental subspace learning. ICIP 2007; 45-48.

[44] Li R, Chen Y, Zhang X. Fast robust eigen-background updating for foreground detection. ICIP 2006; 1833-1836.

[45] Han B, Jain R. Real-time subspace-based background modeling using multi-channel data. ISVC 2007, pages 162-172, November 2007.

[46] Tsai D, Lai C. Independent component analysis-based background subtraction for indoor surveillance. IEEE Trans Image Proc IP 2009, Volume 18, Issue 1, pages 158-167, January 2009

[47] Bucak S, Gunsel B. Incremental subspace learning and generating sparse representations via non-negative matrix factorization. Pattern Recog 2009; 42(5): 788-797.

[48] Skočaj D, Leonardis A. Canonical correlation analysis for appearance-based orientation and self-estimation and self-localization. CogVis Meeting, Zurich, Switzerland, January 2004.

[49] Jolliffe I. Principal component analysis. Springer Series in Statistics, Springer, NY, 2002.

[50] Hyvarinen A, Oja E. Independent component analysis: Algorithms and applications. Neural Networks 2000; 13: 411-430.

[51] Lee D, Seung H. Algorithms for non negative matrix factorization", Neural Infor Sys 2000; 942-948.

[52] Tang F, Tao H. Fast Linear discriminant analysis using binary bases. *18th Inter Conf on Pattern Recognition, ICPR 2006*, ,Volume 2, 2006

[53] Hardoon D, Szedmak S, Shawe S. Canonical correlation analysis: An overview with application to learning methods. Neural Computation 2004; 16: 2639-2664.

[54] Colmenarez, A., Gutta, S., Trajkovic, M.: US20030156759 (**2003**).

[55] Lipton, N.A., Haering, C.A.: US20040126014 (**2004**).

[56] Toyama, J.K., Krumm, L.B.: US20060034520 (**2006**).

[57] Porikli, F., Tuzel, O., Brinkman, D.: US20040239762 (**2004**).

[58] Zhao Y, Gong H, Lin L, Jia Y. Spatio-temporal patches for night background modeling by subspace learning. *Inter Conf on Pattern Recognition, ICPR 2008*, pages 1-4, Tampa, USA, December 2008.

[59] Zhang, Z., Lipton, A.,Venetianer, P., Yin W.: US20090060277(**2009**).

[60] Weng J, Zhang Y, Hwang W. Candid covariance free incremental principal components analysis. Transaction on PAMI 2003; 25: 1034-1040, August 2003.

[61] Bell A, Sejnowski T. An information-maximization approach to blind separation and blind deconvolution. Neural Comput 1995; 7(11):1129-1159.

[62] Schutte J, Groenwold A. A study of global optimization using particle swarms. J Global Optim 2005; 31(1): 93-108.

[63] Yan J, Zhang B, Yan S, *et al.* IMMC: Incremental maximum margin criterion, KDD 2004: 725-730 , Seattle, USA, August 2004.

[64] Chen D, Zhang L. An Incremental Linear Discriminant Analysis Using Fixed Point Method. ISSN 2006; 3971: 1334-1339.

[65] Kim.T, Wong S, Stenger B, Kittler J, Cipolla R. Incremental Linear Discriminant Analysis Using Sufficient Spanning Set Approximations. CVPR 2007, pages 1-8, Mineapolis, USA, June 2007.

[66] Rosipal R, Kramer N, Overview and Recent Advances in Partial Least Squares, SLSFS 2005, 3940: 34-35.