New Trends on Dynamic Object Segmentation in Video Sequences: A Survey

Ramírez A. Graciela and Chacón M. Mario I.

Estudio de las Nuevas Tendencias en la Segmentación de Objetos Dinámicos en Secuencias de Video

***Abstract*— This paper presents a survey of the latest methods developed for dynamic object segmentation in video sequences. After an exhaustive search of the latest seven years of research done in this area there were found more than 90 papers where segmentation task has been performed. These researches were classified according with the main method or feature used to segment. Once finalized the literature review it was found that Background Subtraction Methods are very common in segmentation tasks but also algorithms where Energy Minimization is used to improve segmentation results, as those based on Graph Cut, are very popular in segmentation literature. At the same time, different studies demonstrate that Motion information used in conjunction with Probabilistic algorithms could produce accurate segmentation results even when the dynamic object is partially occluded. Also, it was found that Depth information could achieve precise segmentation results; this is because most of the time the object to segment may not have consistent color and texture features but the object must occupy an integrated region in the space. This paper also describes a list of the common databases used to evaluate the performance of the different video segmentation.**

***Keywords*— Background Subtraction Methods, Dynamic Object Segmentation, Energy Minimization, Motion.**

Resumen— Este trabajo presenta un estudio de los métodos desarrollados recientemente para la segmentación de objetos dinámicos en secuencias de video. Después de una búsqueda exhaustiva de las investigaciones realizadas en esta área en los últimos seis años fueron seleccionadas más de 90 publicaciones en donde se desarrollan algoritmos de segmentación. Estos trabajos fueron clasificados de acuerdo al método o característica principal utilizada para la segmentación. Una vez finalizada la revisión de artículos se encontró que los Modelos de Sustracción de Fondo son muy utilizados en tareas de segmentación, así como también que algoritmos basados en Minimización de Energía son muy populares en la literatura para lograr mejoras en los resultados, tal es el caso de aquellos basados en Graph Cut. Al mismo tiempo, diferentes estudios demuestran que la información de Movimiento en conjunto con algoritmos de Probabilidad puede producir resultados de segmentación certeros aún y cuando el objeto dinámico se encuentre parcialmente ocluido. Además, se encontró que la información de Profundidad puede lograr resultados de segmentación muy precisos, esto debido a que comúnmente el objeto a segmentar pude no tener un color o textura consistente pero éste ocupa una región estable en el espacio. Al final de este trabajo se presenta una lista de las bases de datos que frecuentemente se utilizan para evaluar el desempeño de diferentes algoritmos de segmentación de objetos dinámicos en secuencias de video.

Palabras clave— Modelos de Sustracción de Fondo, Segmentación de Objetos Dinámicos, Minimización de Energía, Movimiento.

Manuscript received on January 09, 2013. This work was supported by Fondo Mixto de Fomento a la Investigación Científica y Tecnológica CONACYT-Gobierno del Estado de Chihuahua, under grant CHIH-2009-C02-125358.

G. Ramírez and M.I. Chacón are members of the Visual Perception Applications on Robotic Lab of Instituto Tecnológico de Chihuahua, Ave. Tecnológico 2909, Chihuahua Chih., C.P. 31310 México. Teléfono: +(52) 614-413-7474; fax: +(52) 614-413-5187. e-mail: gmramirez@itchihuahua.edu.mx, mchacon@ieee.org.

I. Introduction

In computer vision literature, segmentation means breaking a scene into non-overlapping compact regions, where each region constitutes pixels that are bounded together on the basis of some similarity or dissimilarity measure [1]. This process is the basic and most critical step towards more complex tasks such as object identification, industrial inspection, digital entertainment, tracking, human sequence evaluation, etc. Therefore, the basis for these high-level systems relies on the success of the dynamic object detection in the video sequences. If in the initial stage is produced an inaccurate segmentation result, the final high level system will carry over these errors producing failures in the final application.

During the last decade, there have been developed many algorithms to segment arbitrary objects from video sequences which can be categorized with respect to various criteria. Li and Ngan [2] explained seven common standards used to classify a segmentation algorithm: (1) *Data-based* mode: based on the data types used in the segmentation tasks, such as nature, human or medical videos; (2) *Interaction-based* mode: where two main categories can be found, supervised (initial user intervention defines the regions of the dynamic objects) or unsupervised; (3) *Feature-based* mode: based on the selection of features, for example segmentation based on color, texture, intensity, shape or motion features; (4) *Inference-based* mode: segmentation is performed based on “message passing” mode where particular features (color, texture, intensity, motion, etc.) are passed to a specific segmentation method; (5) *Space-based* mode: segmentation is performed in spatial or temporal methods; (6) *Class-based* mode: where specific objects are extracted of the video, such as face, cars, buildings, etc., and (7) *Semantic-specific* mode: defined as a process that divides an image into meaningful segments associated with some semantics. Of course, there is not a distinct boundary between them and a segmentation algorithm can be developed using a combination of different modes. In this paper, we classify different researches using the *Inference* and *Feature* modes. In the *Inference* mode we will focus on the classification of the main method used to achieve the segmentation results. We will present algorithms based on: a) background subtraction methods, b) energy minimization algorithms and c) clustering based segmentation methods. In some situations, the author(s) use(s) a combination of methods (for example a background subtraction method in conjunction with a graph-based segmentation algorithm); in these situations the classification will be made with the method where the author(s) made an important contribution. In the *Feature* mode classification we will describe researches based on: a) depth, b) motion, c) histogram and d) infrared information.

Looking for papers that present surveys of different segmentation methods it was found Faliu Yi and Inkyu Moon [3]. They describe a survey of graph-cut methods applied only to image segmentation, not to video segmentation tasks. It is a very good paper where the graph-cut algorithm is well explained. Another survey is presented by Thilagamani and Shanthi in [4]. As in the other case, in this paper is only exposed image segmentation algorithms not video segmentation methods. In [5] Candamo *et al.* published a survey of automatic-behavior-recognition techniques with focus on the surveillance of human activities, specifically in transit applications. In one section of Candamo's paper is presented a classification of different video segmentation algorithms. As there can be observed, the surveys found only exposed image segmentation methods or are not focused in a deep analysis of the latest algorithms developed for video segmentation tasks. Therefore it was considered appropriated to present in this paper the latest researches developed for the detection of dynamic object in video sequences identifying the different algorithms that could be implemented as well as those features that could be used to develop a robust segmentation algorithm. Mainly, (as explained previously) because this is a fundamental step in more sophisticated systems oriented to security or some other important application.

Commonly, in order to evaluate the performance of a new video or image segmentation algorithm, the authors opt to use the same database and compare their performance either visually or using some metrics as precision, recall, f-measure, etc. In this paper we present some common video and image databases used by the computer vision community to evaluate the performance of their new proposed methods.

The major contributions of this paper are:

1) to present a summary of the latest dynamic object segmentation algorithms.

2) to classify the different video segmentation algorithms using the Inference and Feature classification modes.

3) to present a list of the common databases used by the computer vision community to evaluate a segmentation algorithm.

Hence, if someone is interested in implementing a segmentation algorithm to identify dynamic objects in a video sequence, could be appropriate to review the recent publications presented in this paper and test the algorithm with some of the popular video databases that are shown at the end of this document.

In the following sections we describe how different types of algorithms were implemented for the identification of dynamic objects in video sequences. In most of the papers reviewed not only segmentation is performed but they also perform tracking or labeling of components, however we focus only on the dynamic object segmentation task. In section 2, we present the classification of different segmentation algorithms based on the main method used by the author(s), defined as *Inference-based* classification. In section 3, we classify the algorithms based in the main feature, defined as *Feature-base* classification. Section 4 describes the common databases used by the computer vision community. In section 5, we conclude this paper.

II. Inference-base Classification

Segmentation is a very important procedure because it is the initial step to diverse applications such as image interpretation, video analysis, digital entertainment, etc. However, it is a difficult task due to variations in the object of interest and its surrounding. Also, when is required to analyze a video sequence is necessary to consider many other factors as: dynamic background, momentary occlusions in the object of interest, camera motions, shadows, illumination variation, etc., that could affect the accuracy of the segmentation algorithm. In the following sections we will describe how different authors have implemented segmentation algorithms for video analysis and how they handle these segmentation issues.

A. Background Subtraction Methods.

One of the most common techniques used to segment dynamic objects from a scene is based on background subtraction methods where a comparison of an image that does not contain any object of interest (background) is performed against another observed image. Approaches based on background subtraction have become a popular choice when the illumination changes are gradual. This section describes different investigations where the objective is to build a robust background model and then perform background substation to achieve the segmentation of the dynamic objects.

One common method used for background modeling is based on Gaussian Mixture Models (GMM), first presented in [6], where each pixel is modeled by a mixture of K Gaussian distributions. The disadvantage of this model is the computational complexity of the algorithm making this technique unpopular for real-time purposes. Because of this reason, some authors implement a variation of this model that still preserve the benefits of GMM but reducing the computational cost. Appiah *et al.* [7] demonstrated that real time segmentation of moving objects from video sequences taken under variable lighting conditions can be achieved using a Field Programmable Gate Array (FPGA). The algorithm consists of two stages: background subtraction and connected components labeling. In the first part Appiah implemented a GMM algorithm followed by a temporal low pass filter for noise reduction. The results reported are compared against Pfinder and Stauffer's algorithms showing a slight improvement in the sensitivity measurement. The video sequences are from outdoor and indoor scenes taken from PETS 2000. The principal contribution of the paper is to show that a variation of a GMM algorithm can be implemented in real time.

Another FPGA implementation of a GMM algorithm is proposed by Genovese and Napoli [8]. The paper clearly describes how several researchers have developed a GMM based algorithm and how they propone to implement one variation in the FPGA.

Another alternative to reduce the computational cost of a GMM algorithm is proposed by Yuan *et al.* in [9] where a segmentation method based on Incremental Principal Component Analysis is used to identify moving object zones. Thus, the GMM is only applied in the moving zones reducing computational time. They tested the algorithm with their own images but didn't show a comparison of their results with other methods.

In some cases GMM is used in conjunction with other methods in order to obtain better results, that is the case reported in [10] where Support Vector Machine, SVM, is used in conjunction with GMM. In this paper the GMM models the background and a subtraction algorithm classifies the pixels as background or foreground. Then the SVM takes the foreground pixels and performs a second classification that identifies pixels that really belong to the background. The results show improvements when only GMM is used with outdoor videos.

The objective of algorithms based on background subtraction is to build a robust background model. Chan and Chien [11] created a multi background model where each pixel could have N candidates of background values. The algorithm continuously updates the multi background model and once the pixel is defined as background, morphological operations (opening and closing) are used to reduce noise.

In [12] Huang and Tan implemented a visual surveillance system to describe the motions of humans and vehicles and give a warning if these motions describe suspicious activity. The algorithm proposed constructs probability density functions of background, shadow and foreground based on GMM. Results in this work demonstrate that the proposed method can deal with dynamic scene and cast shadow effectively considering spatial information and Probability Density Functions (PDF) of shadow.

Chen *et al.* [13] proposed to initially perform an inexpensive analysis of the video applying only a simple algorithm and then use an inference procedure based on second order Markov model to identify which frames are necessary to perform deeper analysis. First, in order to detect moving objects in the scene a simple background subtraction algorithm is used. If a moving object is detected a robust method based in GMM is applied. This algorithm uses recursive equations to simultaneously select the appropriate number of components of the mixture model.

The concept of frame saliency is introduced in [14]. Frame saliency indicates the relevance of a frame to a GMM model in order to select a small set of salient frames for an efficiently and robust model training. Also the paper describes the development of a modified Expectation Maximization (EM) algorithm for saliency estimation and GMM learning initialization. This proposed GMM algorithm is applied in spatiotemporal segmentation. We can observe in the reported results that the proposed method is sensitive to moving background.

Nam *et al.* presented an intelligent video surveillance system in [15] where the analyzed information is exchanged between different systems. The moving object detection task is achieved with a background subtraction scheme and the background model is build using a Gaussian distribution applying hue and saturation values of the scene.

A new framework to identify moving objects, abandoned objects or removed objects using GMM as the basic framework is analyzed in [16]. To handle quick lighting changes in the scene the Phong shading model is used in the GMM algorithm. Another important feature that the method uses to segment moving objects is the incorporation of region level information to detect static regions, reduce foreground fragments and classify the type of object identified (abandoned or removed). In order to deal with camera motion or drastic light variation the frame level information is used to reset the GMM. Also the proposed GMM can handle stopped and slow moving objects because of the feedback mechanisms incorporated. The proposed method runs about 150fps for color images and 200fps for grayscale images at size 160x120 and works well in real time video surveillance. Only when the object is very similar to the background the algorithm produce erroneous results.

Zheng *et al.* presented a model based on GMM that incorporates neighborhood characteristic of a pixel (color and brightness) as a solution for bad segmentation results caused by light and shadow changes [17]. Segmentation results presented in this paper indicates that the algorithm has some problems when dynamic background is present.

Hishinuma *et al.* [18] proposed an algorithm based on GMM that is robust to changes in camera angles and pan/tilt motions without a pre-stored background model. Hishinuma take advantage of the internal processing capability of cell phones to implement a GMM based algorithm. Results demonstrate that the algorithm can produce accurate segmentation results and reveal that this algorithm could be implemented in mobile camera for intelligent sensor purposes.

Some authors in the literature have indicated that algorithms based on Kernel Density Estimation (KDE) could adapt better to complex background changes than those based on GMM. A distribution based on kernel is a generalization of a GMM algorithm which does not require parameter estimation (that is one of the advantages compared with GMM, it is not necessary to determine the number of distribution) [19].

Li *et al.* presented in [20] a kernel density estimation model with a function of cast shadow and reflection image suppression. Three KDE models are compared, the first one is based on pixel intensity information, the second on pixel color and intensity values and the third one is the proposed model that uses pixel gradient (magnitude and direction) and intensity information. The model is used in outdoor videos and shows that it can successfully suppress shadows and reflection.

Zhu *et al.* described in [21] a robust background model using a non-parametric KDE where the mean shift method is used to approximate the local maximum values of the density functions. Suppose the initial background Gaussian distribution has m Gaussians. The Means Shift algorithm (MS) is used to find all the local maximum values in the distributions to be estimated. The paper focuses on describing how the background is updated. In order to improve the final segmentation quality a Bayes classifier is used. The paper also presents comparisons of the proposed method versus others KDE methods showing improvements in convergence and running speed as well as in visual segmentation results.

Another model is presented by Zhu and Song in [19] [22] where in order to build a robust background model a recursive Bayesian estimator is used to update the background parameters. Once the algorithm determines that the background is robust, background subtraction is performed. Finally a local texture correlation operator is used to improve the segmentation results. The algorithm is compared with GMM algorithms showing a better in the proposed method.

The algorithm proposed in [23] assigns the label of foreground, background or shadow to each pixel based on a Bayesian MRF model. The pixel is characterized by its color components in the CIE L*u*v space and texture features. Results are presented using the recall and precision metrics.

A method that builds the background and foreground models for each new image based on non-parametric functions is described in [24]. For each new image, background and foreground models are built using non parametric functions. A Bayesian algorithm is used to track the moving objects and this information is feedback to the foreground model making the proposed method robust to illumination changes and to partial or total occlusions.

Segmentation algorithms are not only applied in surveillance systems to identify moving objects but also in other cases. For example a different application is shown in [25] where Lai *et al.* presented the idea of modify the view angle in sports videos (specifically in tennis videos) in order to enhance the experience of the TV watchers deciding their own viewing angle. The background is generated by calculating the pixel value with maximum probability in temporal and spatial distribution. Next, the foreground objects are segmented with a frame difference. Finally a Bayesian matting model is implemented to improve segmentation results.

Chiu and Tsai in [26] use the statistic mode to form the background model. The mode is based on the most frequently occurring gray level of observed consecutive image frames. The largest and second largest frequencies of pixel occurrences (first and second modes) are calculated and used in the background model. The segmentation results of this algorithm are compared against other methods based on GMM and Kernel Density Function using The Microsoft's Wallflower dataset. The findings show a better performance of the proposed method and it is able to process 153fps.

A surveillance system that detects moving intrusive ships in a farmed fish is presented by Hua Wu-Chih *et al.* in [27]. The background is constructed using the median scheme of previous N frames. One important contribution in this work is how it is reduced the influence of wave ripples in the segmentation results using brightness and chromatic distortion (brightness of wave ripple is larger than that of the background). Finally, morphological operations are implemented in order to make the ship region more accurate.

Tsai and Lai [28] proposed a fast background subtraction algorithm based on Independent Component Analysis (ICA) for home-care and health-care monitoring. In this model two images, one representing the background and other the foreground are used to form the data matrix of the ICA model. The ICA method measures the difference of the joint PDF and the product of the marginal PDF to separate highly correlated data. The Particle Swarm Optimization (PSO) algorithm is used to search for the best de-mixing matrix of the ICA model and separate the foreground objects of the image. Results show that the algorithm has some difficulties with shadows and also when the foreground region has some discontinuities because it is similar to the background.

A method based on background subtraction, defined as Visual Background extractor (ViBe), is presented in [29]. To classify a pixel the algorithm compares it to the closest values within the set of background samples that define a sphere centered on it. The pixel is classified as background if the cardinality of the set intersection of this sphere and the collection of background model samples is larger than or equal to a given threshold. The proposed algorithm outperforms methods based in GMM, Bayesian Histogram and first order low pass filter in computation speed and detection rate.

Schreiber and Rauter exposed in [30] how a parallel graphical processor unit (GPU) can be used to implement a real surveillance system based on background subtraction. Initially the background model is set equal to the first frame, then a matching condition between the current pixel and the new incoming pixel is calculated based on a threshold value. The time it takes to absorb a stopped object into the background is defined by the user. The method is compared with a GMM based method demonstrating that the proposed outperforms in adaptation speed, complexity and foreground separation. The algorithm was implemented on GPU using CUDA achieving a run time of 2.3ms on a 352x288 frame, almost 5 times faster than in the CPU version.

A color similarity metric is presented by Samatelo and Salles in [31] applied to a video surveillance system. This metric considers the magnitude and phase (based on intensity and chromaticity) of the difference between the background model and each new video frame. The algorithm is tested using the PETS2004 database and the percentage of correct detected objects is more than 90% in 12 scenes and only in 4 cases the result is below 50%.

Lucia Maddalena and Alfredo Petrosino [32] propose a method that can handle moving backgrounds, gradual illumination variations, camouflage and shadows using a Self-Organizing Neural Network (SOM). In the SOM each pixel is compared with an initial background model to determine if there exists a weight vector that best matches it, it this is found it means that the pixel belongs to the background and it is reinforced. Otherwise if no acceptable matching occurs the pixel is detected as belonging to the foreground. The algorithm is compared against other common algorithms indicating an improvement in the recall, precision and similarity measurements. One of the disadvantages of the proposed algorithm is that there exists many parameters to initialize and they could affect the segmentation results.

One of the most common problems in segmentation algorithms is when shadows are identified as part of the foreground causing incorrect segmentation results. In [33] is presented a technique that can detect achromatic and chromatic shadows even when the foreground is very similar to the shadowed regions. In this paper is explained the different kind of shadows and also is presented a comparison table of different detection algorithms developed for this task. First a moving object mask is generated using a simple background subtraction technique, then a luminance model is used to detect shadow areas where luminance, irradiance and reflectance vector of the moving area are calculated. Each identified area is partitioned into a set of segments using a graph based method to finally classify them as foreground or shadow after analyzing the intrinsic parameters of sub segments. The algorithm is tested with indoor and outdoor scenes and the shadow detection and discriminative rate is used as a performance metric achieving in the worst case a 0.72 value at a very low computational cost.

Even when the issue of sudden illumination change is known in the computer vision community, very few researches have been developed to deal with this particular problem. Vosters *et al.* [34] deal with this issue and proposed a model that requires a training stage in which an Eigen background model is trained using specified sequences that contains expected illumination changes of the background. The only issue reported is that moved background objects will be detected as foreground forever because the object's new location is not captured in the Eigen background model.

Lee *et al.* presented in [35] a segmentation method that use the difference between the current frame and the previous frame and compared it against a threshold to identify moving objects. A pixel that remains stationary for a long time is considered to be a reliable background pixel and is registered in the background model.

Ivan Huerta *et al.* [36] presented a segmentation model that combines three background subtraction models (based on color, edge and intensity information) and a temporal difference algorithm. Each model is designed to handle some common issues presented in video segmentation tasks such as camouflage and shadows. The segmentation results achieved by Huerta outperform most state-of-the-art algorithms using indoor and outdoor videos.

B. Energy Minimization Algorithms

1) Graph Cut Algorithm

In computer vision, image segmentation generally can be formulated as an energy minimization problem and Graph Cut theory has generated extensive interest due to its powerful capability as an energy minimization tool. Graph Cut is a technique used to improve the segmentation results by a previous segmentation algorithm. Graph Cut looks for a set of optimum segment boundary lines that separate interior and exterior markers. In a Graph Cut algorithm the image is treated as a graph with a set of vertices and edges. In the graph structure, vertices correspond to image pixels and edges represent the links between vertices in four directions. The objective is to minimize a function (energy function) who is modeling the boundary of the object [37].

As previously described, when the camera is in a static position and the scene in the video does not almost change it is common to use a background subtraction algorithm where the background is modeled using probabilistic methods. When the camera moves freely these kinds of methods do not have good performance. In [37] it is proposed a block based iterative appearance modeling technique that uses temporal model propagation and spatial model composition. The algorithm also employs Graph Cut to generate the final segmentation mask where the inputs of Graph Cut are the likelihood maps from the temporal and spatial models. The algorithm is implemented in real time and the camera moves freely. Precision, recall and f-measure results demonstrate the good performance of the algorithm. A temporal propagation model is also used by MinGang *et al.* in [38]. MinGang presents an interactive approach where a *video cutout*[1] system automatically propagates the segmentation results frame by frame based on the optical flow. In the final phase a Graph Cut algorithm is implemented to refine the propagated segmentation results.

Another example where Graph Cut is used is presented in [39] and [40] where a five-view camera system is used to capture the multi-view video data. Only one camera was selected as the key view to start the segmentation process using color, intensity orientation, motion and deep information to calculate Saliency Maps (SM). This information is used by Graph Cut to produce the final segmentation. The results presented shows that even when there is a moving object in the background it can be eliminated of the foreground region using the depth information in the SM. The videos used are from their own database and only use indoor scenes; the application that they propose is for video conferencing scenario where the camera is relatively near the person to be segment.

Liu *et al.* [41] presented an automatic human body detector based on human features. Face regions are identified to locate the human body and characteristics based on color distribution of the object and background are detected. A coarse to fine segmentation framework based on GMM and Graph Cut is proposed to deal with partly object detection. A background contrast removal is used to improve segmentation results. In this work is proposed a Self-Adaptive Initialization Level Set (SAILS) to handle problems where the color between the object and background is similar and also to speed up the evolution of the level set strategy. The system was evaluated with indoor and outdoor scenes considering different poses, sizes, movements and lighting conditions. The proposed method was compared versus region based Graph Cut, bi-layer segmentation, background cut and contour based level set and improvements in the results could be observed when using the proposed method. One of the disadvantages is that the system can detect multiple human bodies but only can segment one object at a time with stationary background. The advantage is that the proposed system could be implemented in real time scenarios as video surveillance systems.

Nicolas Papadakis and Aurelie Bugeau used in [42] motion information to build predicted sets using Gaussian velocity models characterizing the motion of each moving object. The Graph Cut algorithm is used in order to handle the occluded parts of the objects where the links generated are the key point for the segmentation of occluded parts.

A dynamic version of the Graph Cut algorithm was developed by P Kohli and P.H.S Torr in [43]. The user must provide segmentation cues or seeds in the first frame of the video sequence to build color histograms for the background and foreground. These histograms are later used to calculate the likelihood term of the energy function of the MRF for all the video sequence frames where the energy of the previous state is used to calculate the current one. The advantage with this algorithm is that the response time improves significantly compared with the best known static mincut algorithm. In [44]

[1] *Video-cutout* is a foreground segmentation technique used in video clip analysis.

and [45] is also implemented an interactive algorithm based on energy minimization procedure but using the geodesic distance.

An example of using Graph Cut algorithms combined with matting methods is exposed in [46]. Initially the algorithm needs a user input to indicate foreground and background regions. A bilayer segmentation step is performed where an opacity propagation algorithm is used to predict the foreground object in the next frame generating an Opacity Map (OM) in combination with the Graph Cut algorithm. At last, an accurate trimap (background, foreground and alpha matte) is generated based on a Local GMM. Results indicate that the algorithm fails with sharp illumination changes between frames and when the foreground object moves to fast.

Zhong *et al.* [47] implemented a dynamic object segmentation algorithm using a combination of a background subtraction model (bgs) and a matting algorithm. Initially, the dynamic object is detected with the bgs model; next, a heuristic algorithm defines the pixels seeds localization of background, foreground and unknown regions (unknown regions is the limit between the background and foreground where is not clear where one region ends and the other begins). Then, an energy minimization algorithm is implemented to identify the improved foreground region. The complete algorithm achieves a speed of 5 fps at the resolution of 320x240 pixels. Zhou *et al.* [48] also proposed an automatic pixel seed initialization in the matting algorithm followed by the graph cut algorithm.

Tang *et al.* presented a foreground prediction algorithm denominated Opacity Propagation [49]. This model takes the binary mask of the initial frame as an opacity map and propagates this information to the current frame by minimizing a cost function. The algorithm was tested in a variety of video sequences and compared with probability maps generated by Bayesian Estimation, Weighted Kernel Density Estimation, Coherence Strips and Local Classifiers showing improvements with respect to these methods. Results demonstrate that this method can handle abrupt illumination changes. One of the disadvantages of the proposed method is that the computational load is very high. Trying to overcome this issue some authors have implemented the minimization algorithm using a GPU. Vineet and Narayanan [50] explain how 60 Graph Cuts per second on 1024x1024 images were performed using a GPU.

The objective of Guillemaut and Hilton in [51] is to extract foreground regions and reconstruct them in 3D. The algorithm assumes multiple foreground layers corresponding to people or objects that are located at different depths considering a single background layer. The data is captured with multiple synchronized video cameras and an initial segmentation is performed with a simple subtraction scheme. The segmentation is performed using depth, color, contrast and multi-view consistency cues based on graph-cuts algorithms. Segmentation results presented demonstrate that the algorithm could produce accurate results but the main disadvantage is the time performance reported where in the best case is 63s per frame.

The segmentation procedure proposed by Jacobson *et al.* in [52] is based on normalized cuts and probability boundary (pB) algorithms for the detection of edges using color and texture information. A weight matrix is formed between each pair of pixels to measure pixel similarity that is used by the normalized cut algorithm to partition the image. Then, the image is further segmented by a k-means clustering algorithm. Regions with similar color and texture are merged on the assumption that they belong to the same object. This merge process is repeated until a small number of regions exist. The complexity of this algorithm is high, achieving a segmentation result of 60s per frame.

The method proposed by Pei Yin *et al.* in [53] generates correct segmentation results even in the presence of large background motion and with a nearly stationary foreground. The algorithm is based on diverse visual cues such as motion, motion context, color, contrast and spatial prior that are fused using a Conditional Random Field model. The segmentation task is achieved using a binary minimization algorithm defined as min-cut. The results using this method are compared against those based on stereo segmentation techniques achieving a very similar performance.

In [54] Cheng *et al.* presented an algorithm that was implemented for real time video analysis based on kernels to model the spatial-temporal characteristics of the background subtraction problem. The kernels used are defined as ILK (Implicit online Learning with Kernels), SILK (Sparse variant of ILK that incorporates spatial correlations) and SILK-GC where a Graph Cut algorithm is implemented. All these three variants are based on support vector machine methodology. The algorithms learn from previous scenes to predict a label for each pixel (foreground or background) and then minimize an energy function to optimize segmentation results. The proposed method was implemented in a GPU. In the results it can be observed that the algorithm had some difficulties with shadows detecting them as part of the foreground.

Civit and Escoda proposed an algorithm in [55] based on the Hierarchical Belief Propagation method and outliers reduction by regularization on over segmented region based on a cost minimization problem (the over segmented region is done using the k-means approach). Each pixel has a set of cost derived from its probabilities to belong to three classes: foreground, background and shadow. The probabilities are computed from chromatic distortion, color distance and brightness measures. Each pixel will be assigned to the model that has the lowest associated cost. The algorithm was implemented in a GPU GTX295 card with 1376x384 and 688x192 picture resolutions with a processing time of 140.6ms and 44.8ms per frame respectively.

Gulshan *et al.* [56] implement an algorithm that automatically segment humans from background given a bonding box. Using depth information a background subtraction algorithm takes a few frames to start segmenting accurately. First, local histograms of oriented gradients (HOG) are used to predict the initial segmentation results; then this segmentation mask is feed into Local GrabCut (using pixel color and edges information) to obtain the final segmentation results. The method is evaluated with their own database showing a performance of 88.5% in overlap score.

Zhang *et al.* [57] proposed a segmentation algorithm based on motion parameters, optical flow information, depth maps and a minimization algorithm to identify dynamic objects in video scenes. The disadvantage of Zhang’s algorithm is that the algorithm detected shadows as part of the foreground region.

Another example of segmentation using depth information is in [58] where Time-of-Flight (TOF) cameras are used to video segmentation tasks. The proposed method combines color and depth information in a probabilistic framework and Graph Cut is used to obtain the bilayer segmentation mask. After bilayer segmentation step a Matting algorithm is implemented in order to separate the foreground, background and "unknown" regions in the scene. The proposed method could be implemented in real time using GPU.

Dahan *et al.* [59] used the Kinect sensor where color and depht information are the inputs of the mean shift algorithm to produce initial detection of dynamic objects. An energy minimization algorithm based on the Graph Cut is used to obtain accurate segmentation regions. Visual results indicate the good performance of the segmentation algorithm presented here.

Segmentation algorithms based on Graph Cut have been very useful, however when high level cues are incorporated (information as shape or color distribution) the segmentation results in a harder energy minimization problem. In these cases, Branch and Mincut techniques are used as proposed in [60]. Shape information is relatively stable to changes in appearance and could be propagated to adjacent frames resulting in a good segmentation approach. Lee *et al.* [61] proposed a method that propagates the global shape (ground truth) for a small number of frames by using Branch and Mincut. Segmentation regions which are likely to be erroneous are identified (defined as questionable points) and a local reinforcement is performed where local appearance, shape of preliminary boundary and pairwise gradients are unified and are used to produce the final segmentation result. This result is used to update the set of shape template. The algorithm was tested using videos from different TV series and webcam sequences confirming its robustness to background clutter and camera motion. The performance decreases when the shape of the foreground changes abruptly. Another disadvantage of this algorithm is that it requires an initial ground truth set making the results dependent of these user's inputs. In [62] a similar approach is implemented where initially a manual segmentation is performed with approximately 5% of all frames. Then a segmentation of the incoming frames is calculated with a superpixel algorithm followed by a Graph Cut method. The disadvantage of this proposal is that 60s of video sequence are analyzed in 20 – 30min.

For a complete reference of how graph algorithms are been applied to vision problems refer to [63].

2) Image Foresting Transform (IFT)

In the IFT algorithm the user must indicate a seed (or marker) pixels of the object of interest and its surrounding background. IFT finds a minimum-cost path from internal and external seeds to each pixel. IFT process the image as a graph (as Graph Cut algorithm) and can handle partially occluded objects of arbitrary shape and size. The object does not need to have a homogeneous color or texture, and it may contain internal sharp boundaries between regions of very different looks and will often recover occluded objects once they become visible again. The interior and exterior markers will define the object segmentation in the next frame. IFT results indicate that is very efficient in object segmentation tasks since it uses a variant of Dijkstra's algorithm with the max path-cost function which can be implemented to run in $O(N)$ time for an image with N pixels, in contrast the best algorithm used in Graph Cut runs at $O(N^{2.5})$. An example of an IFT implementation is presented by Minetto *et al.* in [64] where the algorithm is compared against OpenCV Snake, and Zhong-Chang algorithms. Even when the object of interest is partially occluded the segmentation can be performed once the object becomes visible again. Also, when the background is very similar to the foreground or the background is dynamic the segmentation performance is acceptable.

3) Level Sets

Level set functions provide a simple framework for modeling the shape and evolution of curves. Curves merge, move, brake or disappear during the course of their evolution; level set method handles all these topological changes very easily [65]. In [66] is presented one method that can track multiple moving objects on dynamic and cluttered backgrounds and can handle inter-object occlusions problems. In this algorithm an initial segmentation is performed manually and once the background and foreground are separated an algorithm that combines region, boundary and shape information is used to improve edge regions. Minimization of the energy function is achieved using Level sets. The visual results presented in this paper are very accurate, but the computational cost is high.

Xiao *et al.* [67] show another example of an interactive method. In order to group similar regions (defined as region based information) a similarity map between the current frame and the training data is calculated using GMM. An energy function which incorporates edge and region based information is proposed using level set functions. Segmentation results presented are accurate but as previously mentioned the disadvantage to implement it in real time systems is that it requires an initial user input.

Another application of a level set algorithm is presented by Prisacariu and Reid [65]. Here is proposed a method based on the derivatives of a level set segmentation energy function with respect to the pose parameters of known 3D models. Segmentation and pose are recovered by nonlinear optimization functions as gradient descendent and conjugate gradients. The algorithm was implemented using a GPU achieving a real time implementation of the algorithm.

C. Clustering Based Segmentation Methods

A clustering algorithm is a very common method used to identify groups in a set of objects. In a clustering algorithm the objects in one specific group are more similar to each other that to those in a different group. Some authors have demonstrated that a clustering algorithm could be applied to segmentation tasks producing acceptable results.

Alpert *et al.* [68] implemented a method where pixels are gradually merged with its most similar neighbor to produce larger regions. Likelihood based on intensity and texture cues are defined to produce the merged regions. The results are compared against normalized cuts, mean shift, contour detection and hierarchical segmentation algorithms showing improvements in Alper's results.

Automatic analysis of video sports has gained much attention due to its usefulness to increase performance in athletes. Hung Mao-Hsiung *et al.* [69] propose a technique to segment and analyze these kinds of videos. Initially a PDF of color components (cb,cr) is used with the steepest ascent hill-climbing algorithm to generate clusters. They are merged into four color classes: red, green, blue and gray. The algorithm is compared against a GMM based method indicating that the algorithm proposed is significantly less computationally expensive and obtains comparable segmentation accuracy.

Silva and Scharcanski [70] select a set of points, defined as particles, and identify their trajectory based in a clustering and MS algorithm. A meta-clustering validation algorithm is implemented to compare the clusters formed in the previous step against meta-cluster prototypes in order to correct the segmentation labels. Finally a spatial filter is applied to eliminate groups of adjacent particles that are not significant. Results of the algorithm demonstrate that the parameter initialization of the MS algorithm could produce over-segmentation results. Kriechbaum *et al.* [71] also use motion information and the MS algorithm to segment moving objects. MS is applied to locate clusters to get the dominant colors of the image that can be used in the segmentation process. Kriechbaum also presents an overview of different natural based tracking techniques proposed in the last years.

Sundaram and Keutzer [72] indicate that intensity, color, texture and motion cues could be used to create an affinity matrix where its generalized eigenvectors could be calculated. A clustering algorithm is applied on the eigenvectors and 3D superpixels are obtained. This initial boundary is refined with ultrametric contour maps and it produces a segmentation result. This algorithm was implemented using CUDA for GPU programming. On average, the runtime for the segmentation was about 5 minutes for 200 frames of size 352x288.

A technique that can produce a background image based on a patch background initialization (PBI) algorithm is proposed in [73] by Colombari and Fusiello. First, camera noise is estimated using the median absolute difference algorithm. The spatial domain is subdivided into overlapping windows, on each window, cluster image patches are calculated using Sum of Squared Distances (SSD). A cluster representative is calculated with an average measure and the cluster with maximal length is selected as background. Finally the background is formed based on the background Tessellation algorithm.

III. Feature-based Classification

A. *Depth Information*

It is natural that the object of interest to segment may not have consistent color and texture features but must occupy an integrated region in the space [74]. Taking this into account there has been many researches that have used depth information in object segmentation tasks. Xia *et al.* [74] proposed to use the information provided by the Kinect to detect people in indoor environments. First, the image that is provided by Kinect is processed with the nearest neighbor interpolation algorithm to fill pixels that the sensor is not able to measure to then smooth them with median filter. A Canny edge detector is used to find edges in the depth array eliminating those that are smaller than a specified threshold. A binary head template is implemented to find regions that may contain a head. Finally a region growing algorithm finds the entire body of the person. The method is compared against Ikemura's algorithm showing that the proposal of this paper improves in precision, recall and accuracy results. The problem that is reported is that if the head is occluded or if the person is wearing a strange hat then the method will fail.

Some other authors proposed the use of stereo-vision techniques for image and video processing using depth information for segmentation tasks where the disparity concept is much employed. Disparity is defined as the relative displacement between the left and the right image points belonging to the same object point. Disparity can be an important feature, which provides layer information to segment video object from sequences. In [75], the disparity measure and Saliency Map (SM) function are used for segmentation task. The SM features used in this paper are color, depth, contour, texture, size, location and motion. One of the problems presented in this method is camouflage; also the algorithm is time consuming causing a non feasible real time implementation.

Wei *et al.* [76] utilized stereo video information to segment objects in video sequences. Results indicate that the use of depth, disparity and spatio-temporal information produce accurate segmentation results even when the object move slowly or when multiple objects overlapped.

In [77] depth information is used to identify different objects in the video sequence. Results indicate that the limitation of this proposal is that the moving objects cannot be too far away from the camera because of the similar depth values between the background and the moving object.

J. Ruiz-Hidalgo *et al.* [78] proposed to segment the depth map and color information into homogeneous regions using the Weighted Euclidean Distance, WEDM.

Multi-view Video Systems (MVV) have good applications in Three-Dimensional Television (3DTV) and Free Viewpoint Television (FTV) because it allows the user to change her/his viewing point and direction freely. One of the problems in a MVV system is the inconsistency color between different viewpoints caused by changes in scene illumination, camera calibrations, CCD noise, etc. In [79] is proposed a color correction algorithm of MVV based on depth information. Depth in each view is calculated by disparity estimation and the background and foreground are separated using this information (background has large depth and foreground has small depth). Once that background and foreground are separated, correction factors are obtained and the color correction algorithm is implemented.

B. *Motion Information*

The goal in motion-based segmentation is to partition images in a video sequence into segments of coherent motion. A method that combines motion and intensity information using non parametric distributions is presented by Herbulot *et al.* in [80]. The research demonstrated that when using these two features the segmentation results improves in contrast when they are used separately. The results are compared with an algorithm based on Graph Cut indicating a better performance in the Herbulot's method.

Feng Xu *et al.* [81] indicated that the use of motion information becomes a key issue in video segmentation tasks. Xu utilizes the Scale Invariant Feature Transform (SIFT) features among successive frames to estimate motion information to then over-segment the result using graph methods. These two tasks are executed iteratively to generate a segmentation mask. The performance of the segmentation algorithm is measured using the Receiver Operating Characteristic and compared versus Will's method showing a significant improvement.

Analyzing optical flow vectors, Ranchin *et al.* [82] detected moving objects in video sequences. A classical edge detector functions is used as a weighted total variation. Ranchin's method was tested with vehicular monitoring videos demonstrating the effectiveness of the algorithm.

Denman *et al.* [83] proposed to calculate optical flow information only in dynamic regions previously calculated by a simple background subtraction algorithm. Using this scheme the computational complexity could be reduced achieving a faster processing.

In [84] an initial segmentation is performed based on dense optical flow that describes motion vector and color information where pixels belonging to the same region have coherent motion. A Markov Random Field (MRF) formulates the foreground detection problem as a labeling problem where a likelihood energy function is evaluated for classification. Regions which have the same classification label and similar colors are merged constructing the final segmentation result. Once the foreground and background are separated the background model (based on Stauffer's GMM) is updated. Visual results indicate that the algorithm can successfully eliminate shadows.

Medical analysis systems are another very important application of dynamic object segmentation where suspicious tissues are identified from a video sequence taken with specialized equipment. In [85] motion analysis is used to perform segmentation to examine the whole digestive track. Two methods are used to extract motion features ARPS (Adaptive Rod Pattern Search) and BMSD (Bayesian Multi-Scale Differential) with optical flow information.

In [86] motion compensation information is used in conjunction with a group of pixels defined as macroblocks. This work is based on Fuzzy Logic for detecting moving objects and the system classify the object as a person, group of persons, motorcycle, cars, etc. The inputs of the fuzzy system are motion vectors. One of the advantages of fuzzy logic is that the system could adapt to changes causing an improvement in the object detection step (in this particular case changing the shape of the membership functions). Because of the image is treated with macroblocks the segmentation results could be considered as identification of moving zones.

Li [87] used optical flow algorithms to model motion patterns of dynamic objects and background. Initially, Li implemented a corner detection algorithm and then calculates its optical flow vector. Using the magnitude of the optical flow and the location of the feature points Li identified different groups of dynamic objects.

Schwarz *et al.* proposed the use of depth and intensity information produced by Time of Flight (ToF) cameras or the Kinect device to generate optical flow information [88]. In cases where body parts occlude each other, optical flow is used to identify occluding limbs from the body part achieving high accuracies in the estimation of full body poses. In this research is only considered static background, but the problem of a correct identification of body parts when they are been occluded is a very difficult task. Similarly, Abramov *et al.* [89] used depth and color information of the Kinect sensor in conjunction with optical flow to segment objects in video scenes. Later, in another research performed by Abramov *et al.* [90] optical flow information of stereo videos is calculated on a portable system with an integrated GPU. The algorithm is tested with video sequences acquired with moving cameras containing arbitrary moving objects.

Seamless video composition is the process of extracting foreground objects from a video sequence and pasting them into a new target video sequence, Zhi-Feng Xie *et al.* implemented this technique in [91]. Firstly, a foreground mask of previous frames is constructed using optical flow and propagates to the current frame and an initial trimap is formed. If the generated trimap is inefficient an interactive tool was developed to modify it. Then the discoloration and smudging artifacts are removed based on the mean and standard deviation of the definite foreground region and alpha matte information.

C. Histogram Information

When a moving camera is used in a security system is difficult to obtain accurate segmentation results due to constant changes in the background causing that subtraction background techniques cannot be employed. In [92] is proposed a spatial color histogram model based on color distribution and spatial information, defined as Spatiogram. A center voting method and Hough transform estimate the object location frame to frame. Once the object is localized a back projection technique is used for the segmentation task. The use of the voting scheme is feasible only after a modification of the spatiogram into a form similar to the R-Table of the generalized Hough transform. One of the advantages of this method is that it considers the existence of background regions which could have similar color and appearance with the object to be segmented. The disadvantage is that many parameters must to be initialized and also the algorithm is not completely automatic, in the beginning it needs a manually selection of the target to segment and model it using the proposed spatiogram method.

A method called pRAD (physics based Ridge Analysis of Distributions) is presented by Vazquez *et al.* in [93]. This method performs a histogram analysis exploiting the statistics of the ridges. This method is robust to discontinuities in histograms due to compression and noise caused by the acquisition method. Segmentation is based on the color information. The image is modeled by a set of segments that correspond to a Material Reflectance (MR) described by a distribution in histogram space where each MR is related to an object in the image. The results presented demonstrate that the chromatic information is an important cue on human segmentation tasks. The method is only tested with images, not with videos but due to the results demonstrate that this is a fast segmentation method is feasible to apply it in video sequences, especially because it can handle temporal variations due to camera movements, slight illumination changes and moving shadows.

Hu *et al.* [94] consider rainy video scenes where foreground and background pixels are separated using histogram change detection. Shadows and color reflection are removed using a diamond window mask and color analysis of moving objects.

TABLE I
MOST RELEVANT IMAGE AND VIDEO DATABASES USED TO COMPARE RESULTS IN SEGMENTATION ALGORITHMS

Dataset	**Location**
Wallflower dataset	http://research.microsoft.com/en-us/um/people/jckrumm/WallFlower/TestImages.htm
PETS database	http://www.cvg.rdg.ac.uk/slides/pets.html
CAVIAR database	http://groups.inf.ed.ac.uk/vision/CAVIAR/CAVIARDATA1/
Berkeley Segmentation	http://www.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/BSDS300/html/dataset/images.html
Caltech 101	http://www.vision.caltech.edu/feifeili/Datasets.htm
Caltech 256	http://authors.library.caltech.edu/7694/
CBCL StreetScenes	http://cbcl.mit.edu/software-datasets/streetscenes/
Imageparsing.com	http://www.imageparsing.com/
MSR Cambridge v2	http://research.microsoft.com/en-us/projects/objectclassrecognition/
TRECVid 2007	http://trecvid.nist.gov/trecvid.data.html
Microsoft Research dataset	http://research.microsoft.com/en-us/projects/i2i/data.aspx
Video Recognition Database	http://mi.eng.cam.ac.uk/research/projects/VideoRec/

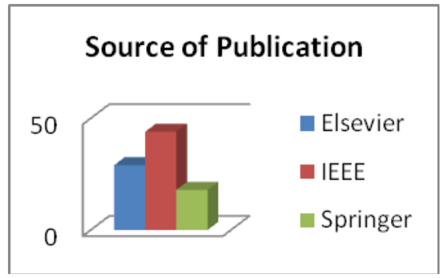
Finally the edge of the moving objects is refined using connected component labeling and morphological operations.

D. *Infrared Information*

Some authors have investigated the option to use the information obtained from infrared videos to segmentation tasks due to the fact that moving objects may have higher temperature compared with its surrounding region. In [95] human segmentation is performed in real time videos using temperature information. Camera information is normalized and incandescent points (light bulbs) are eliminated. The image is binarized and morphological operators are used to eliminate noise in the image. An algorithm is defined to validate the content of each region and verify if it contains one human or more. The method was tested with indoor and outdoor videos showing that it could be implemented in real time applications. The performance of the person segmentation hit was of 98%. One of the problems that were reported is that when the person is too close to the camera it could be not detected as a human.

IV. VIDEO AND IMAGE DATABASE FOR SEGMENTATION TASKS

In order to evaluate the different algorithms proposed for segmentation task many investigators have opted to use the same video database and compare the segmentation results. A very common video database used in the computer vision community is the wallflower's dataset developed by Tomaya *et al.* in [96]. In this dataset are considered issues related to dynamic background, sudden and gradual illumination changes and camouflage; then it is a very complete video dataset (in particular the waving tree video sequence is very common video used to test the segmentation performance). Another very popular dataset used to test segmentation algorithms is PETS where many real scenarios are considered, for example: university public spaces, indoor and outdoor people tracking, vehicle tracking, etc., even when this dataset is used to evaluate surveillance system, in the initial stage must be implemented a dynamic object segmentation algorithm. Similar to these databases, there exist more others that present



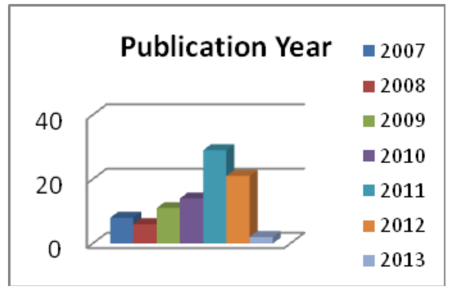


Fig. 1 Source of the consulted papers and it's publication year.

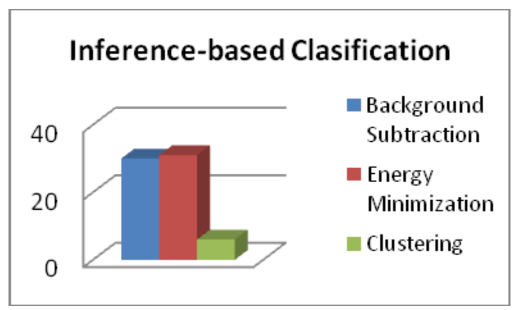


Fig. 2 Papers classification based on main method used to segment.

different challenges to the researches in order to develop robust segmentation algorithms. Table 1 summarizes these common databases used by the computer vision community to validate their new algorithm proposals. If a new segmentation algorithm is developed, it is a good option to test it with any of these databases to confirm that real or common scenarios were considered on its design.

V. CONCLUSIONS

After an exhaustive search of the latest seven years of research done in the area of dynamic object segmentation in video sequences, there were found more than 90 papers. Figure1 shows the publication year of the consulted papers and its source.

Considering the literature review of the different methods used to obtain a robust segmentation result it is difficult to determine which ones are better but most of them made a contribution that will improve future researches. First of all, trying to classify the papers with the method used for the segmentation task was hard because most of them are based on a combination of various methods and also a combination of several features. Figure 2 shows the distribution of the different papers reviewed, classified by the main method used to segment

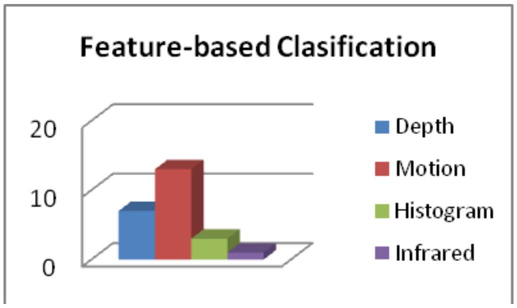


Fig. 3 Papers classification based on the feature used to segment.

defined as *Inference*-base classification. We can observe that Background Subtraction and Energy Minimization algorithms are very used among the computer vision community.

Figure 3 shows the paper distribution based on the main feature used to achieve the video segmentation results. It can be observed that motion and depth information has attracted much attention from researches.

As general conclusions we have:

1) The most common issues presented in video segmentation tasks are: shadows, camouflage, occlusion, dynamic background, gradual or sudden illumination changes and camera motion.
2) Background Subtraction algorithms are very used in segmentation tasks, particularly GMM (50% of the papers found based on background subtraction uses a GMM algorithm). The challenge on these kinds of algorithms is to build robust background models that could adapt appropriately to any possible background change. Specifically, in a GMM algorithm, each pixel could be modeled with n Gaussians. The disadvantage is that the computational time will increase as the number of Gaussians increase and also optimal parameters must be found for each Gaussian model.
3) Another popular algorithm is the one that is based on Energy Minimization, as Graph-Cut. It is an optimization method that is used in conjunction with other algorithms achieving very accurate segmentation results. The disadvantage on a Graph Cut algorithm is that most of them need an initial user definition of the foreground and background regions making them impractical for real time implementation. A variation on the algorithm must be considered to implement it with no user intervention.
4) One of the most common features used in video segmentation algorithms is motion that could be used in conjunction with probabilistic algorithms to deal with partial occlusion and camouflage problems. A common algorithm used to calculate motion information is Optical Flow where a brightness constancy assumption (BCA) is considered to calculate the pixel movement. If the information used by the BCA does not distinguish accurately the pixel movement, the motion information will be incorrect.
5) Depth information has demonstrated its robustness to environmental changes as illumination, dynamic background and camera motion, therefore it is considered convenient to use it in segmentation tasks. Most of the researchers that use depth information use the Kinect to acquire it because of its low cost compared with other depth sensors. Its disadvantage is that it can only be used indoors with a maximum distance between the sensor and the dynamic object of 4m approximately.
6) In order to implement the segmentation of moving objects and meet a real time execution various authors have chosen to implement their algorithms in a GPU.

REFERENCES

[1] Ajay K. Mishra, Yiannis Aloimonos, Loong-Fah Cheong, and Ashraf A. Kassim, "Active Visual Segmentation," Pattern Analysis and Machine Intelligence, IEEE Transactions On, vol. 34, no. 4, pp. 639-653, April 2012.

[2] Hongliang Li and King Ngi Ngan, Image/Video Segmentation: Current Status, Trends, and Challenges: Springer New York, 2011.

[3] Faliu Yi and Inkyu Moon , "Image segmentation: A survey of graph-cut methods," *International Conference on Systems and Informatics (ICSAI), IEEE* , pp. 1936- 1941 , May 2012.

[4] S. Thilagamani and N. Shanthi, "A Survey on Image Segmentation Through Clustering," *International Journal of Research and Reviews in Information Sciences*, vol. 1, no. 1, pp. 14-17, March 2011.

[5] J. Candamo, M. Shreve , D.B. Goldgof, D.B. Sapper, and R. Kasturi, "Understanding Transit Scenes: A Survey on Human Behavior-Recognition Algorithms," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 11, no. 1, pp. 206- 224 , March 2010

[6] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," in Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on., vol. 2, Fort Collins, CO, 1999.

[7] Kofi Appiah, Andrew Hunter, Patrick Dickinson, and Hongying Meng, "Accelerated hardware video object segmentation: From foreground detection to connected components labelling," Computer Vision and Image Understanding, vol. 114, no. 11, pp. 1282–1291, November 2010.

[8] M. Genovese and E. Napoli, "FPGA-based architecture for real time segmentation and denoising of HD video," Journal of Real Time Image Processing, pp. 1-13, Dec 2011.

[9] Yuan Yuan, Yanwei Pang, Jing Pan, and Xuelong Li, "Scene segmentation based on IPCA for visual surveillance," Neurocomputing, vol. 72, no. 10-12, pp. 2450–2454, June 2009.

[10] Jiaming Zhang and Chi Hau Chen , "Moving Objects Detection and Segmentation In Dynamic Video Backgrounds," in Technologies for Homeland Security, 2007 IEEE Conference on, Woburn, MA, May 2007, pp. 64- 69.

[11] Wei-Kai Chan and Shao-Yi Chien , "Real-Time Memory-Efficient Video Object Segmentation in Dynamic Background with Multi-Background Registration Technique," in Multimedia Signal Processing, 2007. MMSP 2007. IEEE 9th Workshop on, Crete, 2007, pp. 219- 222.

[12] Kaiqi Huang and Tieniu Tan, "Vs-star: A visual interpretation system for visual surveillance," Pattern Recognition Letters, vol. 31, no. 14, pp. 2265–2285, October 2010.

[13] Daozheng Chen, Mustafa Bilgic, Lise Getoor, and David Jacobs, "Dynamic Processing Allocation in Video," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 33, no. 11, pp. 2174-2187 , November 2011.

[14] Xiaomu Song and Guoliang Fan, "Selecting Salient Frames for Spatiotemporal Video Modeling and Segmentation," Image Processing, IEEE Transactions on, vol. 16, no. 12, pp. 3035- 3046, Dec 2007.

[15] Yunyoung Nam, Seungmin Rho, and Jong Hyuk Park, "Intelligent video surveillance system: 3-tier context-aware surveillance system with metadata," Multimedia Tools and Applications, vol. 57, no. 2, pp. 315-334, 2012.

[16] YingLi Tian , Andrew Senior, and Max Lu, "Robust and efficient foreground analysis in complex surveillance videos," Machine Vision and Applications, Oct 2011.

[17] Herong Zheng , Zhi Liu, and Xiaofeng Wang , "Research on the video segmentation method with integrated multi-features based on GMM," in Digital Image Processing, 2009 International Conference on, Bangkok, pp. 62- 66.

[18] Y. Hishinuma, T. Suzuk, K. Nakagami, and T. Nishitani, "Transformed domain GMM foreground segmentation for mobile video camera," in Image

Processing (ICIP), 2010 17th IEEE International Conference on, Hong Kong, 2010, pp. 2217- 2220.

[19] Qingsong Zhu and Zhan Song, "A Novel Recursive Bayesian Learning Method for Video Segmentation," in Information and Automation (ICIA), 2010 IEEE International Conference on, Harbin, 2010, pp. 1690- 1693.

[20] Zheng Li, Pohuang Jiang, Hong Ma, Jian Yang, and DongMing Tang, "A model for dynamic object segmentation with kernel density estimation based on gradient features," Image and Vision Computing, vol. 27, no. 6, pp. 817–823, May 2009.

[21] Qingsong Zhu, Guanzheng Liu, and Yaoqin Xie, "Dynamic Video Segmentation via a Novel Recursive Kernel Density Estimation," in Image and Graphics (ICIG), 2011 Sixth International Conference on, Hefei, Anhui, 2011, pp. 452- 455.

[22] Qingsong Zhu and Zhan Song, "A Novel Recursive Bayesian Learning-Based Method for the Efficient and Accurate Segmentation of Video With Dynamic Background," TRANSACTIONS ON IMAGE PROCESSING, IEEE, vol. 21, no. 9, pp. 3865-3876, Sept 2012.

[23] Csaba Benedek and Tamás Szirányi, "Bayesian Foreground and Shadow Detection in Uncertain Frame Rate Surveillance Videos," Image Processing, IEEE Transactions on, vol. 17, no. 4, pp. 608- 621, April 2008.

[24] Carlos Cuevas, Carlos R. del Blanco, Narciso García, and Fernando Jaureguizar, "Segmentation-tracking feedback approach for high-performance video surveillance applications," in Image Analysis & Interpretation (SSIAI), 2010 IEEE Southwest Symposium on, Austin, TX, 2010, pp. 41- 44.

[25] Jui-Hsin Lai, Chieh-Li Chen, Chieh-Chi Kao, and Shao-Yi Chien, "Tennis Video 2.0: A new presentation of sports videos with content separation and," Journal of Visual Communication and Image Representation, vol. 22, no. 3, pp. 271–283, April 2011.

[26] Wei-Yao Chiu and Du-Ming Tsai, "Dual-mode detection for foreground segmentation in low-contrast video images," Journal of Real Time Image Processing, Jan 2012.

[27] Wu-Chih Hua, Ching-Yu Yang, and Deng-Yuan Huang, "Robust real-time ship detection and tracking for visual surveillance of," Journal of Visual Communication and Image Representation, vol. 22, no. 6, pp. 543–556, Aug 2011.

[28] Du-Ming Tsai and Shia-Chih Lai, "Independent Component Analysis-Based Background Subtraction for Indoor Surveillance," Image Processing, IEEE Transactions On, vol. 18, no. 1, pp. 158-167, Jan 2009.

[29] Olivier Barnich and Marc Van Droogenbroeck, "ViBe: A Universal Background Subtraction Algorithm for Video Sequences," Image Processing, IEEE Transactions on, vol. 20, no. 6, pp. 1709- 1724, June 2011.

[30] David Schreiber and Michael Rauter , "GPU-based non-parametric background subtraction for a practical surveillance system," in Computer Vision Workshops (ICCV Workshops), 2009 IEEE 12th International Conference on, Kyoto, 2009, pp. 870- 877.

[31] J. L. A. Samatelo and E. O. T. Salles, "A New Change Detection Algorithm for Visual Surveillance System," Latin America Transactions, IEEE (Revista IEEE America Latina), vol. 10, no. 1, pp. 1221- 1226, Jan 2012.

[32] Lucia Maddalena and Alfredo Petrosino, "A Self-Organizing Approach to Background Subtraction for Visual Surveillance Applications," Image Processing, IEEE Transactions on, vol. 17, no. 7, pp. 1168- 1177, July 2008.

[33] Ariel Amato, Mikhail G. Mozerov, Andrew D. Bagdanov, and Jordi Gonzàlez, "Accurate Moving Cast Shadow Suppression Based on Local Color Constancy Detection," Image Processing, IEEE Transactions on, vol. 20, no. 10, pp. 2954- 2966, Oct 2011.

[34] Luc Vosters, Caifeng Shan, and Tommaso Gritti, "Real-time robust background subtraction under rapidly changing illumination conditions," Image and Vision Computing, vol. 30, no. 12, pp. 1004-1015, December 2012.

[35] Chao-Yang Lee, Shou-Jen Lin, Chen-Wei Lee, and Chu-Sing Yang, "An efficient continuous tracking system in real-time surveillance application," Journal of Network and Computer Applications, vol. 35, no. 3, pp. 1067–1073, May 2012.

[36] Ivan Huerta, Ariel Amato, Xavier Roca, and Jordi Gonz, "Exploiting multiple cues in motion segmentation based on background subtraction," Neurocomputing, no. 100, pp. 183–196, 2013.

[37] Taegyu Lim, Bohyung Han, and Joon H Han, "Modeling and segmentation of floating foreground and background in videos," Pattern Recognition, vol. 45, no. 4, pp. 1696–1706, April 2012.

[38] Chen MinGang, BoCong Sui, Yan Gao, and LiZhuang Ma, "Efficient video cutout based on adaptive multilevel banded method," Science China Information Sciences, vol. 55, no. 5, pp. 1082-1092, May 2012.

[39] Qian Zhang and King Ngi Ngan, "Multi-view video based multiple objects segmentation using graph cut and spatiotemporal projections," Journal of Visual Communication and Image Representation, vol. 21, no. 5-6, pp. 453–461, August 2010.

[40] Qian Zhang and King Ngi Ngan, "Segmentation and Tracking Multiple Objects Under Occlusion From Multiview Video," EEE Trans Image Process. , vol. 20, no. 11, pp. 3308-3313, Nov 2011.

[41] Qiang Liu, Hongliang Li, and King Ngi Ngan, "Automatic body segmentation with graph cut and self-adaptive initialization level set (SAILS)," Journal of Visual Communication and Image Representation, vol. 22, no. 5, pp. 367–377, July 2011.

[42] Nicolas Papadakis and Aurelie Bugeau, "Tracking with Occlusions via Graph Cuts," Pattern Analysis and Machine Intelligence, IEEE Transactions On, vol. 33, no. 1, pp. 144- 157, Jan 2011.

[43] P Kohli and P.H.S Torr, "Dynamic Graph Cuts for Efficient Inference in Markov Random Fields," Pattern Analysis and Machine Intelligence, IEEE Transactions On, vol. 29, no. 12, pp. 2079- 2088, Dec 2007.

[44] Xue Bai and Guillermo Sapiro, "Geodesic Matting: A Framework for Fast Interactive Image and Video Segmentation and Matting," International Journal of Computer Vision, vol. 82, no. 2, pp. 113-132, 2009.

[45] Varun Gulshany, Carsten Rotherz, Antonio Criminisiz, Andrew Blakez, and Andrew Zisserman, "Geodesic Star Convexity for Interactive Image Segmentation," in Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on, San Francisco, CA, June 2010, pp. 3129- 3136.

[46] Zhen Tang, Zhenjiang Miao, Yanli Wan, and Dianyong Zhang, "Video matting via opacity propagation," The Visual Computer, vol. 28, no. 1, pp. 47-61, 2012.

[47] Bineng Zhong *et al.*, "Background subtraction driven seeds selection for moving objects segmentation and matting," Neurocomputing, vol. 1, no. 103, pp. 132–142, 2013.

[48] Yu Zhou, An-wen Shen, and Jin-bang Xu, "Non-interactive automatic video segmentation of moving targets," Journal of Zhejiang University SCIENCE C, vol. 13, no. 10, pp. 736-749, October 2012.

[49] Zhen Tang, Yanli Wan, and Forrest F. Jesse, "Foreground prediction for bilayer segmentation of videos," Journal Pattern Recognition Letters, vol. 32, no. 14, pp. 1720-1734, Oct 2011.

[50] Vibhav Vineet and P. J. Narayanan, "CUDA Cuts: Fast Graph Cuts on the GPU," in Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on, Anchorage, AK, 2008, pp. 1- 8.

[51] Jean-Yves Guillemaut and Adrian Hilton, "Joint Multi-Layer Segmentation and Reconstruction for Free-Viewpoint Video Applications," International Journal of Computer Vision, vol. 93, no. 1, pp. 73-100, 2011.

[52] Natan Jacobson, Yen-Lin Lee, Vijay Mahadevan, Nuno Vasconcelos, and Truong Q. Nguyen, "A Novel Approach to FRUC Using Discriminant Saliency and Frame Segmentation," Journal IEEE Transactions on Image Processing, vol. 19, no. 11, pp. 2924-2934, Nov 2010.

[53] Pei Yin, Antonio Criminisi, John Winn, and Irfan Essa, "Bilayer Segmentation of Webcam Videos Using Tree-Based Classifiers," Journal IEEE Transactions on Pattern Analysis and Machine Intelligence , vol. 33, no. 1, pp. 30-42 , Jun 2011.

[54] Li Cheng, Minglun Gong, Dale Schuurmans, and Terry Caelli, "Real-Time Discriminative Background Subtraction," IEEE Transactions on Image Processing, vol. 20, no. 5, pp. 1401-1414, May 2011.

[55] Jaume Civit and Oscar Divorra Escoda, "Robust foreground segmentation for GPU architecture in an immersive 3D videoconferencing system," in Multimedia Signal Processing (MMSP), 2010 IEEE International Workshop on, Saint Malo, 2010, pp. 75- 80.

[56] Varun Gulshan, Victor Lempitsky, and Andrew Zisserman, "Humanising GrabCut: Learning to segment humans using the Kinect," in Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, Barcelona, 2011, pp. 1127- 1133.

[57] Guofeng Zhang, Jiaya Jia, Wei Hua, and Hujun Bao, "Robust Bilayer Segmentation and Motion/Depth Estimation with a Handheld Camera," Pattern Analysis and Machine Intelligence, IEEE Transactions on , vol. 33, no. 3, pp. 603- 617 , March 2011.

[58] Liang Wang *et al.*, "Automatic Real-Time Video Matting Using Time-of-Flight," International Journal of Computer Vision, vol. 97, no. 1, pp. 104-121, 2012.

[59] Meir Johnathan Dahan, Nir Chen , Ariel Shamir, and Daniel Cohen-Or, "Combining color and depth for enhanced image segmentation and retargeting," The Visual Computer, 2011.
[60] Victor Lempitsky, Andrew Blake, and Carsten Rother, "Image Segmentation by Branch-and-Mincut," Springer-Verlag Berlin Heidelberg, pp. 15-29, 2008.
[61] Soochahn Lee, Il Dong Yun, and Sang Uk Lee, "Robust bilayer video segmentation by adaptive propagation of global shape and local appearance," Journal of Visual Communication and Image Representation, vol. 21, no. 7, pp. 665–676, Oct 2010.
[62] E. Lomakina Rumyantseva, P. Voronin, D. Kropotov, D. Vetrov, and A. Konushin, "Video Tracking and Behavior Segmentation of Laboratory Rodents," Pattern Recognition and Image Analysis, vol. 19, no. 4, pp. 616-622, 2009.
[63] Pedro F. Felzenszwalb and Ramin Zabih, "Dynamic Programming and Graph Algorithms in Computer Vision," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 33, no. 4, pp. 721- 740, April 2011.
[64] R. Minetto *et al.*, "IFTrace: Video segmentation of deformable objects using the Image Foresting Transform," Computer Vision and Image Understanding, vol. 116, no. 2, pp. 274–291, Feb 2012.
[65] Victor A. Prisacariu and Ian D. Reid, "PWP3D: Real-Time Segmentation and Tracking of 3D Objects," International Journal of Computer Vision, vol. 98, no. 3, pp. 335-354, 2012.
[66] Mohand Saïd Allili and Djemel Ziou, "Active contours for video object tracking using region, boundary and shape information," Journal Signal, Image and Video Processing, vol. 1, no. 2, pp. 101-117, 2007.
[67] Chunxia Xiao, Jiajia Gan , and Xiangyun Hu, "Fast level set image and video segmentation using new evolution indicator operators," The Visual Computer, pp. 1-13, Feb 2012.
[68] Sharon Alpert, Meirav Galun, Achi Brandt, and Ronen Basri, "Image Segmentation by Probabilistic Bottom-Up Aggregation and Cue Integration," Pattern Analysis and Machine Intelligence,IEEE Transactions on, vol. 34, no. 2, pp. 315-327, Feb 212.
[69] Hung Mao-Hsiung , Hsieh Chaur-Heh , and Kuo Chung-Ming , "Generalized playfield segmentation of sport videos using color features," Pattern Recognition Letters, vol. 32, no. 7, pp. 987–1000, May 2011.
[70] Luciano S. Silva and Jacob Scharcanski, "Video Segmentation Based on Motion Coherence of Particles in a Video Sequence," Image Processing, IEEE Transactions on, vol. 19, no. 4, pp. 1036- 1049, April 2010.
[71] Andreas Kriechbaum , Roland Mörzinger, and Georg Thallinger, "A framework for unsupervised mesh based segmentation of moving objects," Multimedia Tools and Applications, vol. 50, no. 1, pp. 7-28, 2010.
[72] Narayanan Sundaram and Kurt Keutzer, "Long term video segmentation through pixel level spectral clustering on GPUs," in Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, Barcelona, España, 2011, pp. 475- 482.
[73] A. Colombari and A. Fusiello, "Patch-Based Background Initialization in Heavily Cluttered Video," Image Processing, IEEE Transactions on, vol. 19, no. 4, pp. 926- 933, April 2010.
[74] Lu Xia, Chia-Chih Chen , and J. K. Aggarwal, "Human Detection Using Depth Information by Kinect," International Workshop on Human Activity Understanding from 3D Data in conjunction with CVPR (HAU3D), June 2011.
[75] Ouyang Xingming and Wei Wei, "Video Object Segmentation Based on Disparity," Advances in Web and Network Technologies, and Information Management, vol. 5731, pp. 36-44, 2009.
[76] Jian Wei, Shigang Wang, and Liwei Chen, "Adaptive Stereo Video Object Segmentation Based on Depth and Spatio-Temporal Information," in Computer Science and Information Engineering, 2009 WRI World Congress on, Los Angeles, CA, 2009, pp. 140- 144.
[77] Y. Ma, S. Worrall, and A.M. Kondoz, "Automatic Video Object Segmentation Using Depth Information and an Active Contour Model," in Multimedia Signal Processing, 2008 IEEE 10th Workshop on, Cairns, Qld, Oct 2008, pp. 910- 914.
[78] J. Ruiz-Hidalgo, J.R. Morros, P. Aflaki, F. Calderero, and F. Marqués, "Multiview depth coding based on combined color/depth segmentation," Journal of Visual Communication and Image Representation, vol. 23, no. 1, pp. 42–52, July 2012.
[79] Yue Fei, Mei Yu, Feng Shao, and Gangyi Jiang, "A Color Correction Algorithm of Multi-view Video Based on Depth Segmentation," in Computer Science and Computational Technology, 2008. ISCSCT '08. International Symposium on, Shanghai, 2008, pp. 206- 209.
[80] Ariane Herbulot, Sylvain Boltz, Eric Debreuve, Michel Barlaud, and Gilles Aubert, "Space-Time Segmentation Based on a Joint Entropy with Estimation of Nonparametric Distributions," in SSVM'07 Proceedings of the 1st international conference on Scale space and variational methods in computer vision, 2007, pp. 721-732.
[81] Feng Xu, Kin-Man Lam, and Qionghai Dai, "Video-object segmentation and 3D-trajectory estimation for monocular video sequences," Image and Vision Computing, vol. 29, no. 2-3, pp. 190–205, Feb 2011.
[82] Florent Ranchin, Antonin Chambolle, and Francoise Dibos, "Total Variation Minimization and Graph Cuts for Moving Objects Segmentation," Scale Space and Variational Methods in Computer Vision, vol. 4485, pp. 743-753, 2007.
[83] Simon Denman, Clinton Fookes, and Sridha Sridharan, "Improved Simultaneous Computation of Motion Detection and Optical Flow for Object Tracking," Proceedings of the 2009 Digital Image Computing: Techniques and Applications, IEEE, pp. 175-182 , Dec 2009.
[84] Shih-Shinh Huang;, Li-Chen Fu, and Pei-Yung Hsiao, "Region-Level Motion-Based Background Modeling and Subtraction Using MRFs," Image Processing, IEEE Transactions on, vol. 16, no. 5, pp. 1446- 1456 , May 2007.
[85] Baopu Li, Max Q.-H. Meng, and Chao Hu, "Motion analysis for capsule endoscopy video segmentation," in Automation and Logistics (ICAL), 2011 IEEE International Conference on, Chongqing, 2011, pp. 46- 51.
[86] C. Solana Cipres, G. Fernandez Escribano, L. Rodriguez Benitez, J. Moreno Garcia, and L. Jimenez Linares, "Real-time moving object segmentation in H.264 compressed domain based on approximate reasoning," International Journal of Approximate Reasoning, vol. 51, no. 1, pp. 99–114, Dec 2009.
[87] Jianqin Yin Yanbin Han Wendi Hou Jinping Li, "Detection of the Mobile Object with Camouflage Color Under Dynamic Background Based on Optical Flow," Procedia Engineering, vol. 15, pp. 2201–2205, 2011.
[88] Loren Arthur Schwarz, Artashes Mkhitaryan, Diana Mateus, and Nassir Navab, "Human skeleton tracking from depth data using geodesic distances and optical flow," Image and Vision Computing, vol. 30, no. 3, pp. 217–226, March 2012.
[89] Alexey Abramov, Karl Pauwels, Jeremie Papon, Florentin Worgoten, and Babette Dellen, "Depth-supported real-time video segmentation with the Kinect," Applications of Computer Vision,IEEE, pp. 457-464, 2012.
[90] Alexey Abramov, Karl Pauwels, Jeremie Papon, Florentin Worgotter, and Babette Dellen, "Real-Time Segmentation of Stereo Videos on a Portable System With a Mobile GPU," Transactions on Circuits and Systems for Video Technology, IEEE, vol. 22, no. 9, pp. 1292- 1305, 2012.
[91] Zhi-Feng Xie , Yang Shen, Li-Zhuang Ma, and Zhi-Hua Chen, "Seamless video composition using optimized mean-value cloning," The Visual Computer, vol. 26, no. 6-8, pp. 1123-1134, 2010.
[92] Suryanto, Dae-Hwan Kim, and Hyo-Kak Kim, "Spatial color histogram based center voting method for subsequent object tracking and segmentation," Image and Vision Computing, vol. 29, no. 12, pp. 850–860, Nov 2011.
[93] Eduard Vazquez, Ramon Baldrich, Joost Van de Weijer, and Maria Vanrell, "Describing Reflectances for Color Segmentation Robust to Shadows, Highlights, and Textures," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 33, no. 5, pp. 917- 930, May 2011.
[94] Wu-Chih Hu , Yan-Ting Ye, Chao-Ho Chen, and Deng-Yuan Huang, "Video object segmentation in rainy situations based on difference scheme with object structure and color analysis," Journal of Visual Communication and Image Representation, vol. 23, no. 2, pp. 303-312, Feb 2012.
[95] Antonio Fernández Caballero, José Carlos Castillo, Juan Serrano Cuerda, and Saturnino Maldonado Bascón, "Real-time human segmentation in infrared videos," Expert Systems with Applications, vol. 38, no. 3, pp. 2577–2584, March 2011.
[96] Kentaro Toyama, John Krumn, and Barry Brumitt, "Wallflower: Principles and practice of background maintenance," *IEEE International Conference on Computer Vision*, 1999.

Graciela Ramírez Alonso was born in Chihuahua, Mexico in 1980. She received the BSc(2002) and the MSc.(2004) degrees in Electrical Engineering from the Chihuahua Institute of Technology, Chihuahua, Chih., Mexico.

Currently, she is a PhD student of the Institute of Technology, Chihuahua, Chihuahua, Mexico. Her research interest includes computer vision, video signal processing and pattern recognition.

Mario I. Chacón Murguía was born in Chihuahua, Mexico in 1959. He received the BSc(1982) and the MSc.(1985) degrees in Electrical Engineering from the Chihuahua Institute of Technology, Chihuahua, Chih., Mexico, and the Ph.D.(1998) in Electrical Engineering from New Mexico State University, Las Cruces NM, USA.

His professional experience includes positions as a Graduate Research and Teaching Assistant at New Mexico State University. He has also developed research projects for several companies. He currently works as a Research Professor at the Chihuahua Institute of Technology, in Chihuahua, Chih., Mexico. He has published more than 105 works in international and national journals and congresses and published 2 books. His current research includes visual perception, computer vision, image and signal processing using computational intelligence, and patter recognition.

Dr. Chacon is a Senior member of The IEEE, member of the IEEE Signal Processing Society, IEEE Computational Intelligence Society and member of the National Research System in Mexico.