

# Image Segmentation Based on Multi-Kernel Learning and Feature Relevance Analysis

S. Molina-Giraldo, A.M. Álvarez-Meza, D.H. Peluffo-Ordoñez,  
and G. Castellanos-Domínguez

Signal Processing and Recognition Group - Universidad Nacional de Colombia,  
Manizales, Colombia

{smolinag, amalvarezme, dhpeluffoo, cgcastellanosd}@unal.edu.co

**Abstract.** In this paper an automatic image segmentation methodology based on Multiple Kernel Learning (MKL) is proposed. In this regard, we compute some image features for each input pixel, and then combine such features by means of a MKL framework. We automatically fix the weights of the MKL approach based on a relevance analysis over the original input feature space. Moreover, an unsupervised image segmentation measure is used as a tool to establish the employed kernel free parameter. A Kernel Kmeans algorithm is used as spectral clustering method to segment a given image. Experiments are carried out aiming to test the efficiency of the incorporation of weighted feature information into clustering procedure, and to compare the performance against state of the art algorithms, using a supervised image segmentation measure. Attained results show that our approach is able to compute a meaningful segmentations, demonstrating its capability to support further vision computer applications.

**Keywords:** kernel learning, spectral clustering, relevance analysis.

## 1 Introduction

Image segmentation is an important stage in computer vision and image processing applications, it consists in splitting an image into disjoint regions such that the pixels have a high similarity according to a preset property or measure for each region and contrast difference among regions. The main goal is to obtain a proper segmentation that can be used in processes such as video object extraction [16], in which for partitioning the image into homogeneous regions that correspond to relevant objects, and then, to extract the moving object, the regions are merged according to temporal information of the sequence. Image segmentation is also used in object recognition systems [2]. Most of these systems partition the object to be recognized into sub-regions and attempt to characterize each region separately to simplify the matching process. Moreover, tracking systems that are region-based techniques [1] use the information of the entire objects regions. They track the homogeneous regions from the object by their

color, luminance or texture. At the end, a merging technique based on motion estimation is carried out to obtain the complete object in the next frame.

Several image segmentation methods have been proposed, which can mainly be divided into the following categories: Histogram-based algorithms, which consider the histogram of an image as a probability density function of a Gaussian, hence the segmentation problem is reformulated as a parameter estimation followed by pixel classification. However, the parameter estimation and the selection of a global threshold in 3D histograms for color images represent a difficult task and could slant the algorithm to work with only some images. The second category are Boundary-based algorithms, the basic idea of this approach is that changes in pixels values among neighboring pixels inside a region is not as significant as changes in pixels values on the boundary of a region, therefore regions can be identified when the boundaries are detected. The main drawback of this approach is that boundaries may not be not totally closed. Many post-processing algorithms have been created in order to connect open boundaries [3], however these algorithms always tend to attain over-segmented results. Finally, Grouping-based algorithms aim to group pixels in the same cluster if they have similar patterns or characteristics, while pixels grouped into different clusters have different characteristics. Nonetheless, traditional grouping algorithms, e.g., Kmeans and Expected Maximization, tend to fall into local optimal, whereas spectral clustering algorithms can converge in a global optimal and can be used on arbitrary datasets [8]. Conventional spectral clustering algorithms used for image segmentation, use as input, similarity matrices based only on pixel intensity. Mostly, this information source is not enough to obtain a good performance by the spectral clustering algorithm.

In this sense, we propose a methodology for image segmentation based on a grouping approach that incorporates multiple sources of information for each input pixel by using a Multiple Kernel Learning (MKL) framework and a relevance analysis for the automatic weight selection of the MKL approach. Taking into account the survey of unsupervised measures presented in [17], we propose to use the unsupervised measure FRC [13] as a feedback control to determine a proper free parameter for the employed kernel. Also, we propose to use the Kernel Kmeans technique as spectral clustering algorithm and a post-processing stage to relabel clusters that are spatially split. Finally, a supervised measure Probabilistic Rand Index, described in [15], is used to objectively evaluate the proposed algorithm performance.

This paper is structured as follow. In Sect. 2.1 we describe the incorporation of multiple sources of information into the image segmentation problem by means of MKL. Section 2.2 explains the automatic weight selection of the MKL approach. In Sect. 3 we present the proposed image segmentation methodology. Section 4, describes the experimental scheme used for the proposed methodology and we expose the performed experiments. Finally in Sect. 5 and 6 discussions and conclusions over attained results are exposed.

## 2 Theoretical Background

### 2.1 Image Analysis by Multi-Kernel Learning

Recently, machine learning approaches have shown that the use of multiple kernels instead of only one can be useful to improve the interpretation of data [12]. Given a set of  $p$  feature representations for each image pixel  $h_i = \{h_i^z : z = 1, \dots, p\}$ , based on the Multi-Kernel Learning (MKL) methods [7], the similarity among pixels can be computed via the function:

$$\kappa_{\omega}(h_i^z, h_j^z) = \sum_{z=1}^p \omega_z \kappa(h_i^z, h_j^z), \quad (1)$$

subject to  $\omega_z \geq 0$ , and  $\sum_{i=1}^p \omega_z = 1$  ( $\forall \omega_z \in \mathbb{R}$ ). Thereby, the input data can be analyzed from different information sources by means of a convex combination of basis kernels. Regarding to image segmentation procedures, each pixel of an image can be represented by including  $p$  different image features, which are properly combined by MKL as shown in (1), in order to enhance the performance of further spectral clustering stages. Nonetheless, as can be seen from (1), it is necessary to fix the  $\omega_z$  free parameters, to take advantage, as well as possible, of each feature representation.

### 2.2 MKL Weight Selection Based on Feature Relevance Analysis

We propose to select the weights values  $\omega_z$  in MKL by means of a relevance analysis over the original image features. This type of analysis is applied to find out a low-dimensional representations, searching for directions with greater variance to project the data, such as Principal Component Analysis (PCA). Although PCA is commonly used as a feature extraction method, it is useful to quantify the relevance of the original features, which also provides weighting factors taking into consideration that the best representation from an explained variance point of view will be reached [5]. Given a set of features ( $\eta_z : z = 1, \dots, p$ ) corresponding to each column of the input data matrix  $\mathbf{X} \in \mathbb{R}^{r \times p}$  (a set of  $p$  features describing a pixel image  $h_i$ ), the relevance of  $\eta_z$  can be identified as  $\omega_z$ , which is calculated as  $\mathbf{w} = \sum_{j=1}^d |\lambda_j \mathbf{v}_j|$ , with  $\mathbf{w} \in \mathbb{R}^{p \times 1}$ , and where  $\lambda_j$  and  $\mathbf{v}_j$  are the eigenvalues and eigenvectors of the covariance matrix  $\mathbf{V} = \mathbf{X}^T \mathbf{X}$ , respectively.

Therefore, the main assumption is that the largest values of  $w_z$  lead to the best input attributes, since they exhibit higher overall correlations with principal components. The  $d$  value is fixed as the number of dimensions needed to conserve a percentage of the input data variability.

## 3 Weighted Gaussian Kernel Image Segmentation

Taking into account the above mentioned techniques, we propose a new image segmentation methodology called Weighted Gaussian Kernel Image Segmentation (WGKS). The main goal of the proposed methodology is to properly identify the objects contained in an image.

The first step of WGKS is to conform a feature space from the original image  $\mathbf{H}_{n \times m}$ . In this sense,  $p$  different features are extracted for each pixel. Thus, a feature space  $\mathbf{X}_{r \times p}$  is obtained, with  $r = n \times m$ . A MKL framework is employed to identify the similarities among pixels, by combining the obtained features using a Gaussian kernel  $\mathbf{G}^z \in \mathbb{R}^{r \times r}$  as shown in (2)

$$\mathbf{G}^z(x_i^z, x_j^z) = \exp\left(-\frac{1}{2} \left(\frac{|x_i^z - x_j^z|}{|x_j^z| \sigma}\right)^2\right), \quad (2)$$

where  $\sigma$  is the kernel band-width and the term  $|x_j^z|$  in the denominator stands for comparing the samples  $x_i^z$  and  $x_j^z$  by means of a relative error. Therefore, as described in (1), a weighted gaussian kernel  $\mathbf{G} \in \mathbb{R}^{r \times r}$  can be inferred as  $\mathbf{G} = \sum_{z=1}^p \omega_z \mathbf{G}^z$ , where each  $\omega_z$  is estimated by a feature relevance analysis over the original input space.

In order to exploit the data representation obtained by  $\mathbf{G}$ , a Kernel Kmeans algorithm [6] is used to segment the original image  $\mathbf{H}$ . Moreover, the number of groups  $k$  is calculated from an eigenvalue analysis over a weighted linear kernel, which is computed as  $\mathbf{K}_L = \mathbf{X} \mathbf{W} \mathbf{X}^T$ , where  $\mathbf{W}_{p \times p} = \text{diag}(\mathbf{w}_{p \times 1})$ , in a similar way as described in [11].

## 4 Experiments

To verify the effectiveness of the proposed methodology, natural images drawn from the Berkeley Image Segmentation Database are tested [9]. The Database contains hand-labeled segmentations made by 30 human subjects for 300 color images of  $481 \times 321$  pixels. Natural images are in *jpg* format and human segmentation results in *seg* format. The images exhibit large variety of objects and real world scenes. For concrete testing, images are resized to  $97 \times 65$ , and the following features are extracted: RGB components, row position  $x$ , column position  $y$ , normalized rgb components, HSV components and YCbCr components. Thus, for each image an input feature space  $\mathbf{X} \in \mathbb{R}^{6305 \times 14}$  is obtained.

It is important to note that for all the provided experiments the Kernel-Kmeans (KN-Kmeans) technique is used as a spectral clustering approach. Moreover, the Probabilistic Rand Index (PR) is employed as a supervised segmentation measure. The PR allows to compare a test segmentation with multiple hand labeled ground-truth images, through soft nonuniform weighting of pixel pairs as function of the variability in the ground-truth set [15]. Consider a set of manual segmentations  $\{\mathbf{Y}_1, \dots, \mathbf{Y}_T\}$  of an image  $\mathbf{H} \{h_1, \dots, h_r\}$  consisting of  $r$  pixels. Let  $\mathbf{S}$  be the segmentation that is to be compared with the manually labeled set. The label of point  $h_i$  is denoted by  $l_i^{\mathbf{S}}$  in segmentation  $\mathbf{S}$ , and by  $l_i^{\mathbf{Y}_t}$  in the manually segmented image  $\mathbf{Y}_t$ , with  $\{t = 1, \dots, T\}$ . It is assumed that each label  $l_i^{\mathbf{Y}_t}$  can take values in a discrete set of size  $l^{\mathbf{Y}_t}$ , and correspondingly  $l_i^{\mathbf{S}}$  takes one of the  $l^{\mathbf{S}}$  values. The PR index chooses to model label relationships for each pixel pair by an unknown underlying distribution. It can be seen as if each human segmenter provides information about the segmentation  $\mathbf{Y}_t$  of the

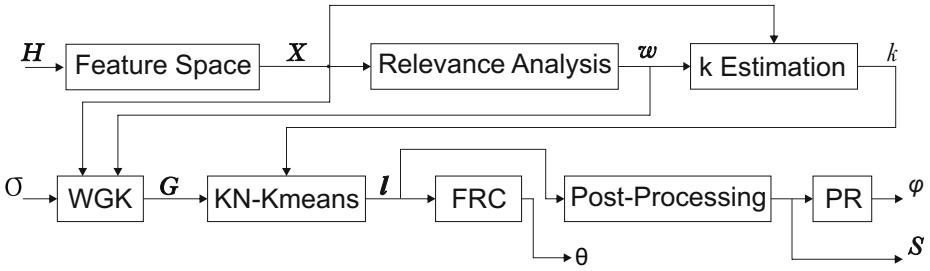


Fig. 1. WGKS scheme

image in the form of binary numbers  $\delta(l_i^{Y^t} = l_j^{Y^t})$ , for each pair of pixels  $(h_i, h_j)$ . Therefore, this measure allows us to compare the segmented image by an image segmentation algorithm against a set of ground truth images as

$$\phi(\mathbf{S}, \mathbf{Y}) = \frac{1}{T} \sum_{t=1}^T \frac{1}{\binom{r}{2}} \sum_{ij} [\delta(l_i^S = l_j^S) \delta(l_i^{Y^t} = l_j^{Y^t}) + \delta(l_i^S \neq l_j^S) \delta(l_i^{Y^t} \neq l_j^{Y^t})]. \tag{3}$$

being  $\phi$  the attained PR value. This measure is widely used because it can retain the uncertainty of the hand labeled segmentations, weighting it in a balanced way. Also, it has the capability to perform comparisons even if the number of groups of each segmented image is different [14].

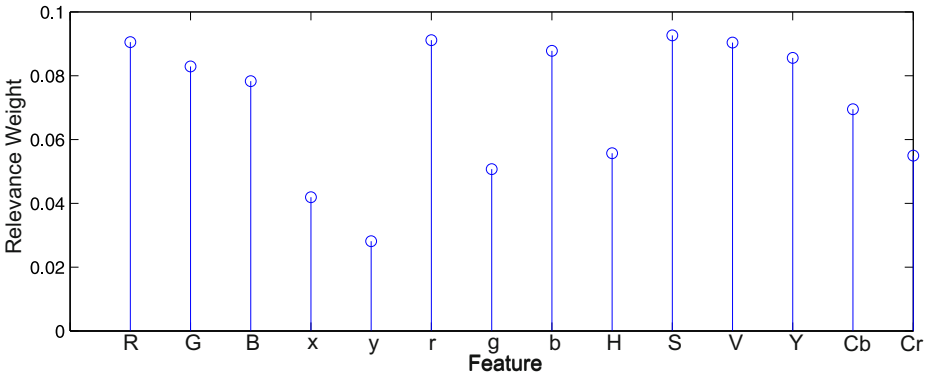
Note that, the  $\sigma$  free parameter of the Gaussian kernel is selected from the set  $\sigma = [0.15 \ 0.3 \ 0.45 \ 0.6 \ 0.75 \ 0.9]$ , using as a cost function the unsupervised measure FRC [13]. Given the optimum  $\sigma$  value according to FRC, a post-processing stage is employed, which consists in relabeling clusters that are split into different groups. The proposed WGKS scheme is shown in Fig.1.

Two different experiments are performed. The first one aims to prove the effectiveness of the proposed WGKS approach when incorporating more information into the segmentation process with automatic parameter selection. To this end, image 388016 (blond-girl) of the Berkeley dataset is used. The WGKS segmentation result over blond-girl image is compared against GKS (WGKS with all equal weights), and against traditional KN-Kmeans computing a gaussian kernel just over the RGB components. The attained segmentation results for blond-girl image are shown in Fig. 2, and the obtained relevance weights for WGKS are shown in Fig. 3.

The second kind of experiments are performed to compare the WGKS algorithm against a traditional image segmentation algorithm named Edge Detection and Image Segmentation System (EDISON), which is a low-level feature extraction tool that integrates confidence based edge detection and mean shift-based image segmentation [4]. The EDISON system has been widely used as a reference to compare image segmentation approaches [17,10]. For testing, the parameters of the EDISON system: scale bandwidth ( $b_s$ ) and color bandwidth ( $b_c$ ), are set as suggested in [10]. 50 randomly selected images from the Berkeley database are



**Fig. 2.** a) Original Image, b) Single Gaussian Kernel RGB, c) GKS (WGKS Equally Weighted), d) WGKS



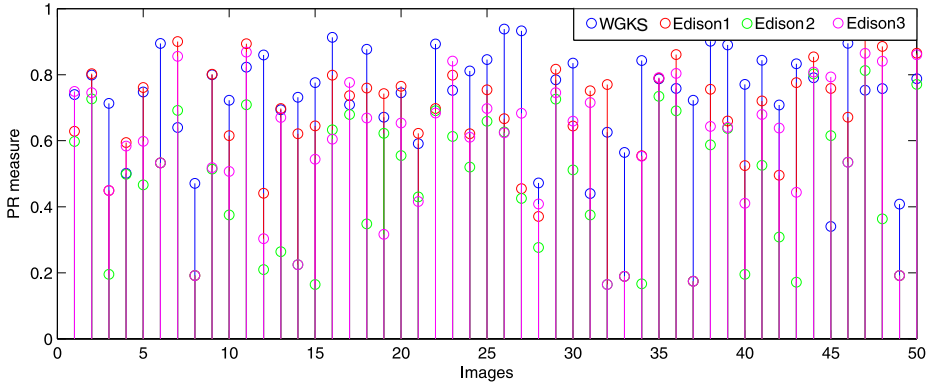
**Fig. 3.** Weight Selection by Relevance Feature Analysis for blond-girl Image 388016

used. The PR results for the second kind of experiments are presented in Fig. 4. Moreover, the mean estimated number of groups and the mean PR accuracy for all the 50 tested images are described in Table 1. Finally, some relevant results of the studied images are shown in Table 2 and Fig. 5.

**Table 1.** Segmentation Performance for 50 images drawn from the database

| Method  | k             | $\phi$               |
|---------|---------------|----------------------|
| EDISON1 | 79.68 ± 47.65 | 0.660 ± 0.191        |
| EDISON2 | 21.68 ± 25.53 | 0.473 ± 0.208        |
| EDISON3 | 54.68 ± 43.46 | 0.589 ± 0.205        |
| WGKS    | 9.862 ± 4.412 | <b>0.742 ± 0.142</b> |

k : Number of Groups,  $\phi$  : PR measure.



**Fig. 4.** Image segmentation results for 50 images. WGKS(b). EDISON1(r), EDISON2(g), EDISON3(m), are EDISON segmentation results using ( $b_s = 7, 7, 20$ ) and ( $b_c = 7, 15, 7$ ) respectively.

**Table 2.** Segmentation Performance for Images of Fig. 5

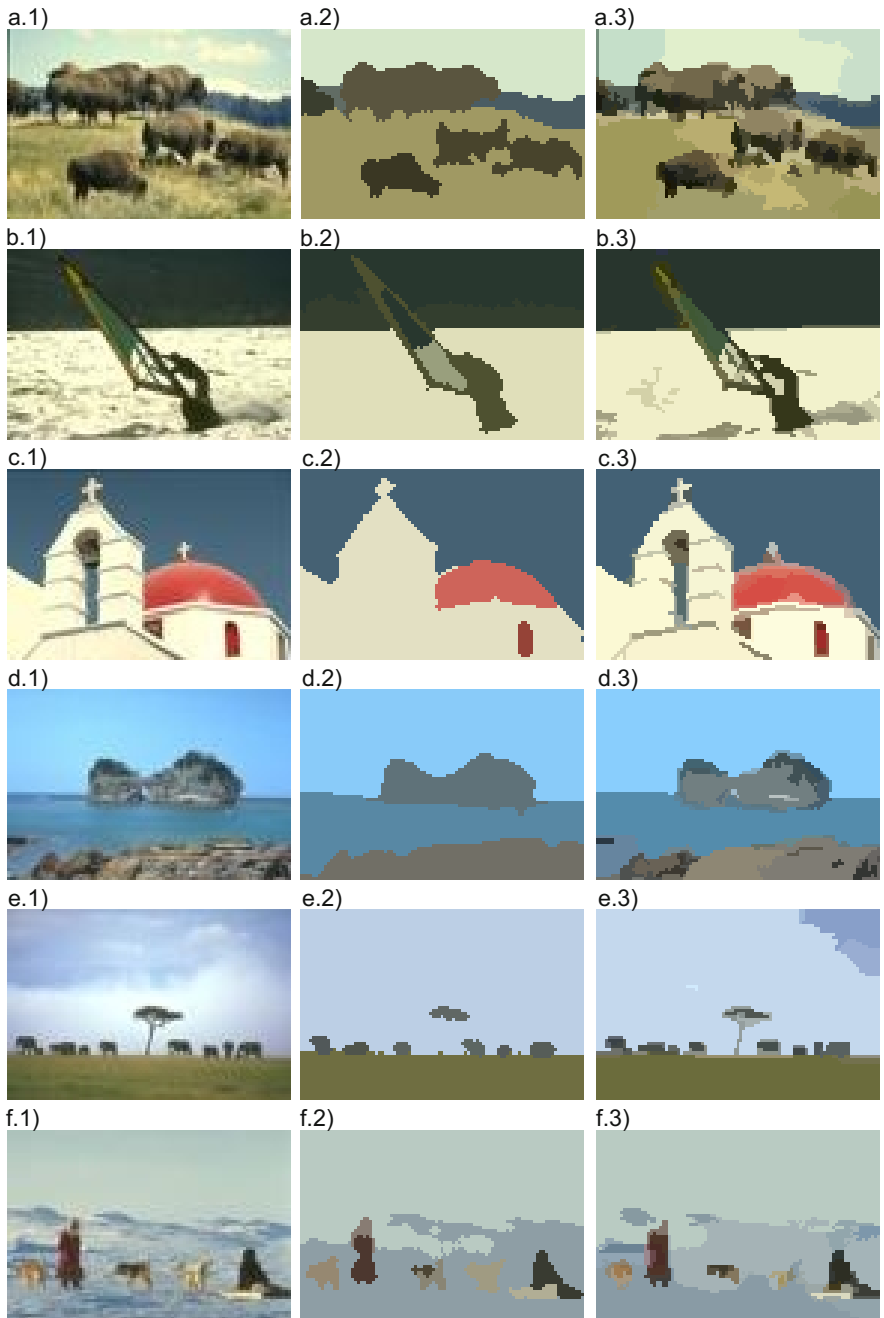
| Method |        | a     | b     | c     | d     | e     | f     |
|--------|--------|-------|-------|-------|-------|-------|-------|
| EDISON | k      | 100   | 61    | 76    | 43    | 32    | 46    |
|        | $\phi$ | 0.659 | 0.671 | 0.532 | 0.666 | 0.456 | 0.554 |
| WGK    | k      | 8     | 6     | 4     | 4     | 9     | 10    |
|        | $\phi$ | 0.890 | 0.894 | 0.897 | 0.936 | 0.932 | 0.842 |

$k$  : Number of Groups,  $\phi$  : PR measure.

## 5 Discussion

From the image segmentation results attained for the blond-girl image, it can be seen how the single Gaussian kernel based segmentation using only RGB components poorly performs, lacking of extra information that could improve the estimation of the number of groups and the Kn-Kmeans clustering (see Fig. 2 b). The latter can be corroborated by a PR measure of 0.055. When the spatial and color spaces information are incorporated into the spectral clustering algorithm based on MKL, the performance improves dramatically, obtaining a PR or 0.721 (see Fig. 2 c). Finally, using the proposed WGKS methodology, the best result is achieved obtaining a PR of 0.774. It can be explained by the estimated weights using the relevance analysis, which allows to identify the most relevant features, avoiding redundant information which could affect the pixel representation (see Fig. 2 d).

The results for the 50 images are exposed in Fig.4. For the EDISON system 3 different combinations of parameters are used. The first combination (EDISON1) is set as  $b_s = 7$  and  $b_c = 7$ , the second one as (EDISON2)  $b_s = 7$  and  $b_c = 15$  and the last one as (EDISON3)  $b_s = 20$  and  $b_c = 7$ . From the figure it can be observed that our methodology obtains the best results in most of the cases,



**Fig. 5.** Image segmentation results. (1) Original Images. (2) WGK. (3) EDISON.



obtaining the first place for 29 images, while EDISON 1 for 16. In Table 1 are exposed the mean results for the 50 images, it can be seen that the WGKS algorithm obtains the best results with the highest mean PR measure 0.742, furthermore, obtains the best stability for all images having the lowest standard deviation 0.142. It also can be seen that the EDISON system always obtains over-segmented results, generating a large amount of groups for each image, whereas the proposed algorithm can correctly identify the number of objects in the scene in most of the cases, it can be explained by the estimation of groups made by the eigenvalue analysis of the weighted linear kernel.

Image segmentation results attained by WGKS methodology, and EDISON system using  $b_s = 7$  and  $b_c = 7$  shown in Fig. 5 demonstrate that the proposed methodology produces more accurate and better segmentation results than the EDISON system, which clearly generate over-segmented images. The results in Table 2 expose that according to the PR measure, WGKS methodology generate very similar segmentations as those realized by each human person, identifying the objects present in the scene. By the other hand, EDISON system generates a large amount of groups for each image, hence, the PR measure penalizes the results, whereas the group estimation of our method was accurate for all images. It is important to note that all the approaches based on spectral techniques require a high computational cost due to the similarity matrix estimation.

## 6 Conclusions

We have proposed a grouping-based methodology for image segmentation called WGKS, which aims to incorporate different information sources by means of a MKL approach, each information source is weighted using a relevance analysis and a Kernel Kmeans algorithm is used to segment the resulting kernel. Experiments showed that the weighted incorporation of spatial and different color spaces information can enhance the data separability for further spectral clustering procedures. The attained results also showed that the estimation of the number of groups made by means of the eigenvalue analysis of the weighted linear kernel was accurate, supporting the performance of the spectral clustering algorithm. Moreover, the use of the FRC measure gave an effective feedback for the correct selection of the kernel bandwidth. As a future work, other different free parameter estimations are to be studied, as well as the extension for temporal analysis is to be designed such that the WGKS methodology can be performed and tested into a complete computer vision process. Furthermore, due to the complexity of the proposed WGKS, a GPU computation scheme could be proposed in order to achieve a real-time application over full size images.

**Acknowledgments.** This research was carried out under grants provided by a Msc. and two PhD. scholarships, and the project "ANÁLISIS DE MOVIMIENTO EN SISTEMAS DE VISIÓN POR COMPUTADOR UTILIZANDO APRENDIZAJE DE MÁQUINA", funded by Universidad Nacional de Colombia.

## References

1. Ozyildiz, E., Krahnstöver, N., Sharma, R.: Adaptive Texture and Color Segmentation for Tracking Moving Objects. *Pattern Recognition* 35(10), 2013–2029 (2002)
2. Besl, P., Jain, R.: Three-Dimensional Object Recognition. *ACM Computing Surveys* 17(1), 75–145 (1985)
3. Canny, J.: A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8(6), 679–698 (1986)
4. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(5), 603–619 (2002)
5. Daza-Santacoloma, G., Arias-Londoño, J.D., Godino-Llorente, J.I., Sáenz-Lechón, N., Osmá-Ruiz, V., Castellanos-Domínguez, G.: Dynamic Feature Extraction – An Application to Voice Pathology Detection. *Intelligent Automation and Soft Computing* 15(4), 667–682 (2009)
6. Dhillon, I., Guan, Y., Kulis, B.: Kernel K-Means – Spectral Clustering and Normalized Cuts. In: 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2004), pp. 551–556. ACM (2004)
7. Gonen, M., Alpaydin, E.: Localized Multiple Kernel Regression. In: 20th International Conference on Pattern Recognition (ICPR 2010), pp. 1425–1428. IEEE (2010)
8. Jung, C., Jiao, L., Liu, J., Shen, Y.: Image Segmentation Via Manifold Spectral Clustering. In: 2011 IEEE International Workshop on Machine Learning for Signal Processing (MLSP 2011), pp. 1–6. IEEE (2011)
9. Martin, D., Fowlkes, C., Tal, D., Malik, J.: A Database of Human Segmented Natural Images and its Application to Evaluating Segmentation Algorithms and Measuring Ecological Statistics. In: 8th IEEE International Conference on Computer Vision (ICCV 2001), vol. 2, pp. 416–423. IEEE (2001)
10. Pantofaru, C., Hebert, M.: A Comparison of Image Segmentation Algorithms. Tech. Rep. 336, Robotics Institute (2005)
11. Perona, P., Zelnik-Manor, L.: Self-Tuning Spectral Clustering. *Advances in Neural Information Processing Systems* 17, 1601–1608 (2004)
12. Rakotomamonjy, A., Bach, F.R., Canu, S., Grandvalet, Y.: SimpleMKL. *Journal of Machine Learning Research* 9, 2491–2521 (2008)
13. Rosenberger, C., Chehdi, K.: Genetic Fusion – Application to Multi-Components Image Segmentation. In: IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2000), vol. 6, pp. 2223–2226. IEEE (2000)
14. Unnikrishnan, R., Pantofaru, C., Hebert, M.: A Measure for Objective Evaluation of Image Segmentation Algorithms. In: IEEE Conference on Computer Vision and Pattern Recognition – Workshops (CVPR 2005 Workshops), pp. 34:1–34:8 (2005)
15. Unnikrishnan, R., Pantofaru, C., Hebert, M.: Toward Objective Evaluation of Image Segmentation Algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29(6), 929–944 (2007)
16. Wang, D.: Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking. *IEEE Transactions on Circuits and Systems for Video Technology* 8(5), 539–546 (1998)
17. Zhang, H., Fritts, J., Goldman, S.: Image Segmentation Evaluation – A Survey of Unsupervised Methods. *Computer Vision and Image Understanding* 110(2), 260–280 (2008)