COMPUTER ANALYSIS OF OBJECTS' MOVEMENT IN IMAGE SEQUENCES: METHODS AND APPLICATIONS

João Manuel R. S. Tavares, Fernando J. S. Carvalho, Francisco P. M. Oliveira,

Ilda M. Sá Reis, Maria João M. Vasconcelos, Patrícia C. T. Gonçalves, Raquel R. Pinho, Zhen Ma

Faculdade de Engenharia da Universidade do Porto

Instituto de Engenharia Mecânica e Gestão Industrial

Rua Dr. Roberto Frias, s/n

4200-465 Porto

PORTUGAL

Email: tavares@fe.up.pt, fjc@isep.ipp.pt, franciscopmoliveira@yahoo.com.br, ildareis@ipb.pt, maria.vasconcelos@fe.up.pt, patricia.goncalves@fe.up.pt, rpinho@fe.up.pt, zhen.ma@fe.up.pt

ABSTRACT

Computer analysis of objects' movement in image sequences is a very complex problem, considering that it usually involves tasks for automatic detection, matching, tracking, motion analysis and deformation estimation. In spite of its complexity, this computational analysis has a wide range of important applications; for instance, in surveillance systems, clinical analysis of human gait, objects recognition, pose estimation and deformation analysis.

Due to the extent of the purposes, several difficulties arise, such as the simultaneous tracking of manifold objects, their possible temporary occlusion or definitive disappearance from the image scene, changes of the viewpoints considered in images acquisition or of the illumination conditions, or even non-rigid deformations that objects may suffer in image sequences.

In this paper, we present an overview of several methods that may be considered to analyze objects' movement; namely, for their segmentation, tracking and matching in images, and for estimation of the deformation involved between images.

Keywords: *Image Segmentation, Movement Tracking and Analysis, Matching, Templates, Deformable Models, Point Distribution Models, Level Set Methods, Kalman Filter*

1. INTRODUCTION

Computer analysis of objects' movement in image sequences is a very complex problem, which may involve tasks for automatic objects detection, matching, tracking and deformation estimation. Motivated by its wide range of significant applications, either in 2D or 3D domains, like in medical imaging based diagnosis, human gait analysis, surveillance systems, traffic analysis, object recognition, pose estimation and deformation analysis, the computer analysis of objects' movement has been evolving considerably over the last decades. For this analysis, many methods may be considered according to the needs of each application, but constrains associated with computational complexity as well as with computation time are usually present.

Although the performance of computer systems has been improving in the last years, the tracking systems, namely the ones that are able to capture, track and analyze the movement of objects in image sequences in a very fast way, often use some kind of simplifications in order to make straightforward and speed-up their computational process.

To analyze objects' movement along image sequences, we first need to detect (i.e. segment) the objects of interest in each image (i.e. find image regions or features corresponding to objects being tracked) and then track them through consecutive images, while maintaining the correct data association (i.e. matching features between consecutive images). Often, the deformation involved between two images of objects is also estimated.

Two main sources of difficulty in performing computational analysis of objects' movement from image sequences are: 1) changes in objects' appearance caused by variations of the considered viewpoints, illumination conditions, topology or non-rigid geometric deformations and 2) situations of total or partial objects occlusion that may occur.

This paper is organized as follows: in the next section, the segmentation of objects in images is considered, by briefly describing some usual methods and presenting some experimental results. In the third section, the problem of tracking objects along image sequences is introduced; thus, some works done in this domain are indicated and experimental results, in particularly using Kalman filtering, are presented. Then, methods to match objects' nodes between images are presented as well as an approach to estimate the deformation involved between two objects, and some experimental results obtained using them are also included. Finally, in the last section, some conclusions are addressed.

2. SEGMENTATION

In Computational Vision, the identification of objects represented in images is commonly known as segmentation. For this image analysis task computational methods are frequently used, based on templates matching, statistical modeling, deformable templates, deformable models or level set methods, [1]. In resume, to accomplish this operation, we can model the images' background or the objects to be segmented instead.

Template matching is used, for example, in [2] for the identification of the human eye in images. Thus, in the referred work, a template image of a human eye is used to search for it in an input image through image correlation, [3]. After this correlation procedure, the centres of regions in the input image that are more alike with the image template used will have the highest correlation values. In Figure 1, we can see the image template used and an example of the results found in [2]. This identification method can give satisfactory results, but presents several limitations, as the difficulty to deal with geometric transforms or illumination changes.

Figure 1. Image template (left) used in [2] to detect the human eye in an input image (right).

Also in [2], it is proposed a method to identify skin areas in an input image. For this, sample images of skin are used to build a statistical model for posterior skin segmentation. Using that model, pixels of areas in the input image that have high probability to be associated with human skin will have high values. In Figure 2, we can see an example of the use of that kind of segmentation, [2].

Figure 2. Skin regions found (right) in an input image (centre) using a statistical model built using sample skin images (some of them on the left).

Statistical approaches like point distribution models are frequently used to extract the most representative characteristics of objects from images. Then, active shape models and active appearance models use point distribution models to segment and recognize the modeled objects in new images. In these methods, models are built from training examples, represented as a set of labeled points, combining statistical shape with grey levels information for the object considered, [4-9].

Since the step of labeling the objects' points is usually manually made and so fastidious and quite time consuming, some approaches have appeared to automatically extract these landmarks points from the objects in images. For example, in [10-12] it is presented computational solutions to automatically extract landmarks from objects like hands and faces in order to smooth the construction of these models. The first modes of variation of the models built describe how the landmark points of the object considered can vary as well as the expected grey level around each of these points, Figure 3, Figure 4 and Figure 5. After the models are built, they can be employed to segment the modeled objects in new images through an active search, Figure 6 and Figure 7.

To conclude, this statistical modeling is only possible if we have a training set with images of the desired object, and it will be only adequate to segment shapes of the object modeled similar to the ones observed in that training set.

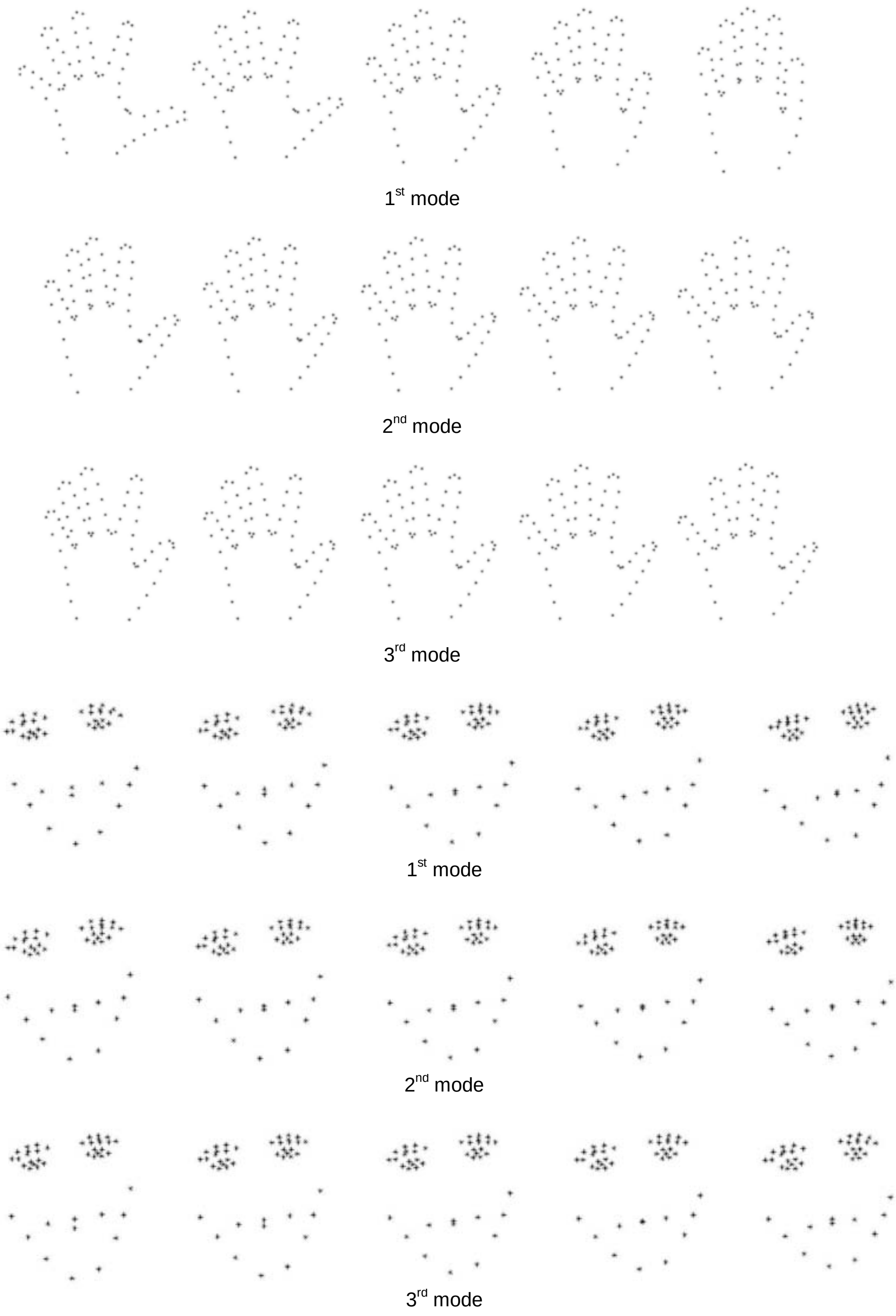


Figure 3. Effects of the first three modes of variation of models built for a hand (top) and for a face (bottom).

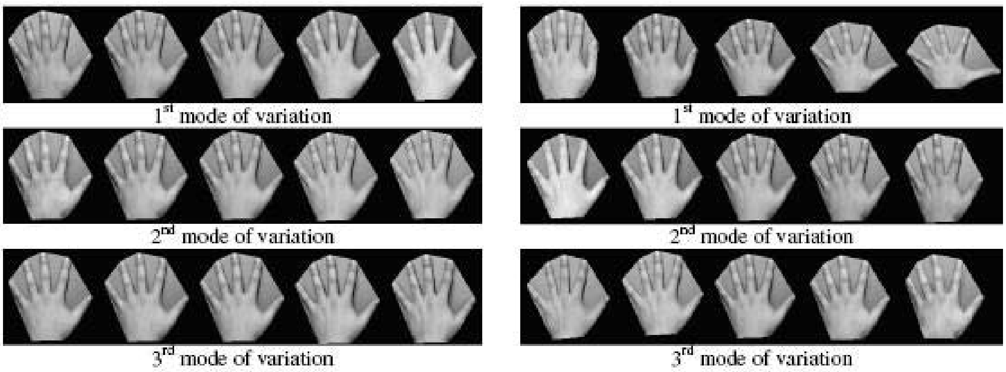


Figure 4. Effects of the first three modes of variation of a texture model (left) and of an appearance model (right) built for a hand.

Figure 5. Effects of the first three modes of variation of a texture model (left) and an appearance model (right) for a face.

Deformable templates are used, for example, in [13], for the segmentation of human eyes in images, Figure 8. In this methodology, a geometric template is built in function of the object to segment; in [13], the template is composed by one circle and two parabolas. Then, using image preprocessing operators, objects' specific features are enhanced in the input image, building energy fields that are later used to iteratively deform the template built in order to accomplish the segmentation. For this segmentation be successfully done, the template must be adequate and should be placed in the input image near the object to segment.

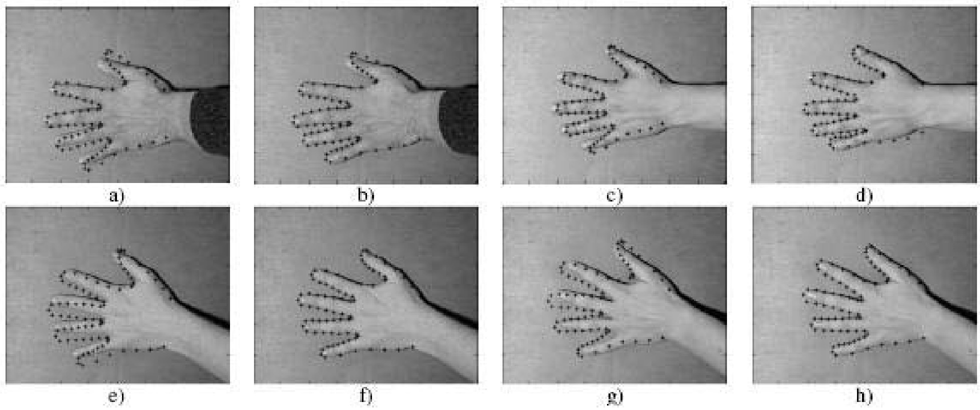


Figure 6. New images with initial position of the mean shape model built overlapped (a, c, e, g) and the final results of the segmentation process (b, d, f, h) using an active shape model.

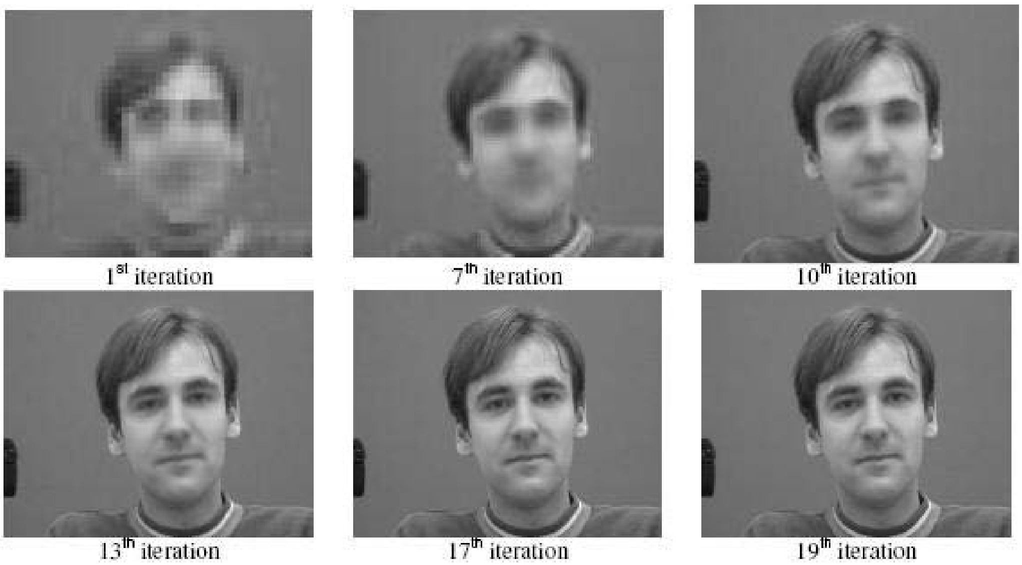


Figure 7. Iterations of the segmentation process using an active appearance model built for a face.

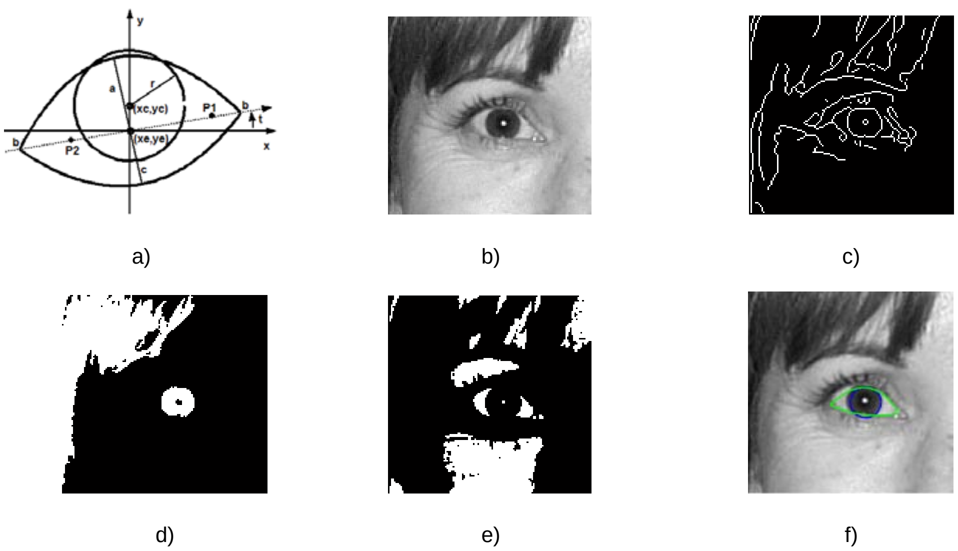


Figure 8. Detection (f) of the human eye in a image (b) using a deformable template (a); energy fields of intensity levels (b), edges (c), intensity valleys (d) and intensity peaks (e).

Usually, objects of free form are better segmented when deformable models are employed, [14]. In these methods, an elastic model is placed in the input image near the object to be segmented, and then the model is deformed in order to segment the desired object. This deformation is guided by image forces, computed by enhancing some particular characteristics of the object in the input image, like intensity edges. For instance, in [15], objects are modeled using the finite elements method, a virtual material is adopted for the object to segment, and the model built is then deformed to its final shape by image forces. More specifically, after manually defining a rough contour for the object to segment, this contour is modelled according to physical principles using the finite elements method. To move the physical model towards the border of the object to segment, the dynamic equilibrium equation is solved, that describes the equilibrium between the internal and external forces. The internal forces are defined by the physical characteristics adopted for the model, determined by the chosen virtual material and the selected level of interaction between the nodes (data points) of the model. The external forces are determined by the intensity, the edges and the distance from each pixel to the nearest edge. In Figure 9, we can see an example of the results obtained using this approach.

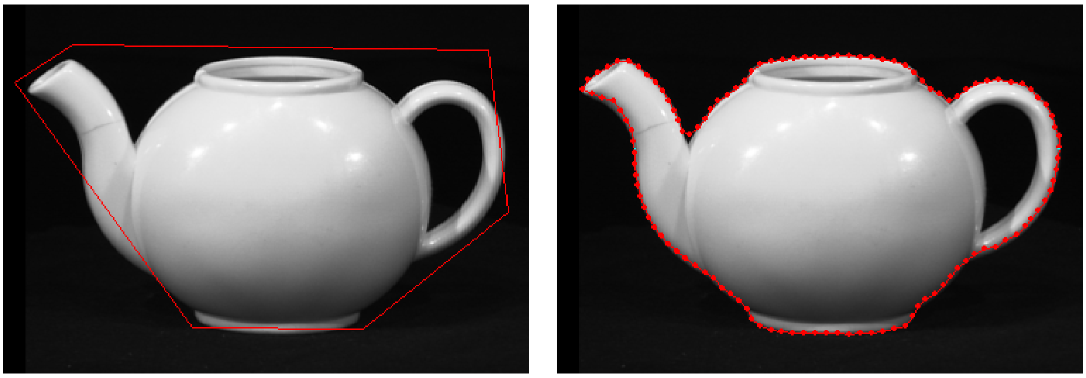


Figure 9. Initial contour user defined (left); result of the segmentation process considering a finite element model with 118 nodes and made of rubber (right).

Another way to perform the segmentation of objects in images is to use the level set method introduced by Sethian and Osher, [16]. The idea behind this method is to embed the moving contour into a higher dimensional level set function, Figure 10. The moving interface can be seen as the zero level set of the function. Then, instead of moving the contour points, we can track the zero level set of the function. The advantage of doing so is that the topological changes will be naturally handled and the geometric properties, like the curvature, can be implicitly calculated. Therefore, the computational complexity is greatly decreased.

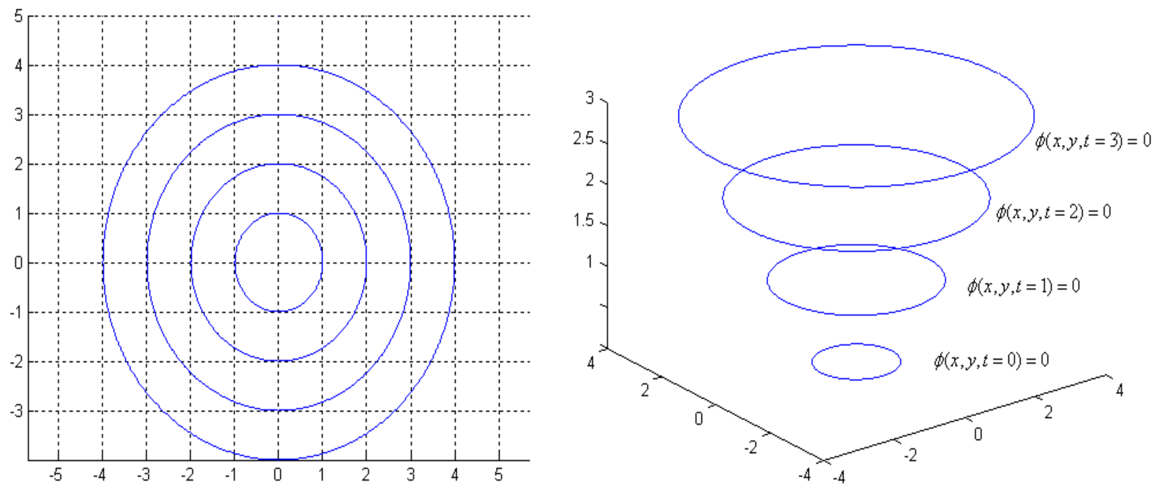


Figure 10. Circle moving in a 2D plane: moving traces (left), traces in the embedded space (right).

Level set methods are one of the most intensively studied approach in the past few decades, especially in the image processing fields, [17-22]. The main advantages of these methods are the ability of naturally handling topological changes and easily incorporating prior knowledge. Other characteristics like the sub-pixel accuracy could also be preferred in certain applications. The ability of naturally handling the topological changes is quite favorable for the tasks such as segmentation and motion tracking, where the topological structures can be quite complex or frequently changed. Segmentation algorithms based on the level set methods treat the procedure of segmentation as an evolution of the initial contours. To complete the segmentation task, proper speed functions should be defined to move the contour to the right position.

The works done by Malladi et.al, [23-27], and Caselles et.al, [28, 29], first applied the level set method to do the segmentation in medical images. Their speed model used the gradient information as stopping criteria. The principle of this speed model is intuitive: When the contour moves to the neighborhood of edges, the gradients of pixels are large, which make the values of speed function become small and therefore slow down the contour. The final edge of the interested objects is considered to be the contour after moving for certain time. An example of the use of this model is illustrated in Figure 11.

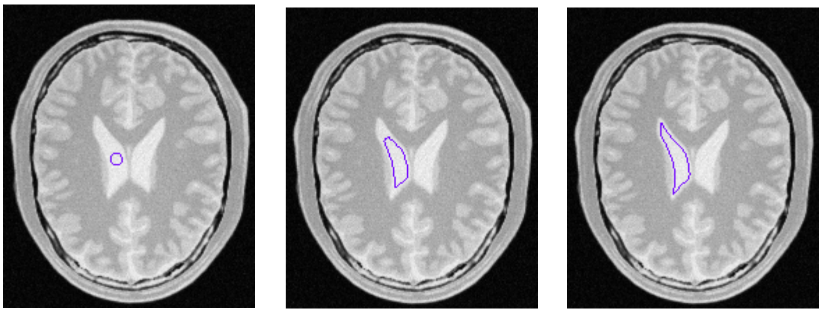


Figure 11. Using the Malladi's segmentation model: initial contour (left), moving contour (centre), final contour (right).

However, as pointed in [17, 30, 31], the above speed models suffered from the drawback of leaking. In others words, these models mainly relied on the gradient information. If the images are noisy or blurred, the contours are either stopped in wrong position or leaked into other objects. A problem is that the real images are usually not under the ideal conditions; for example, due to the heterogeneity of detect fields, many medical images suffer from partial volume effects.

To handle the leakage of the model used and to get more accurate results, different parameters such as the edge forces, [30], and area forces, [31], were incorporated into the equation model to stop the leakage. Algorithms with different regularizing schemes were also proposed. These algorithms usually combined the gradient information with different image cues such as image intensity, shape information and prior knowledge. It should be noticed that the moving equations of the contours related with most of these improved algorithms do not have intuitive meanings as the one in [25]. Instead, they are mainly derived from their corresponding energy functional.

Compared with the original segmentation models, the newly proposed ones have more sophisticated theoretical backgrounds and therefore are usually more effective. For example, in [32], prior shape information was incorporated to the geodesic active contour models [28], and through using the maximizing a posterior and principle component analysis techniques, combining with the level set methods, the proposed method can achieve much better results than the original geodesic active contour method. Another example like Chan-Vese's model [33] combined the Mumford-Shah model [34] with the level set methods to perform complex segmentations, see Figure 12.

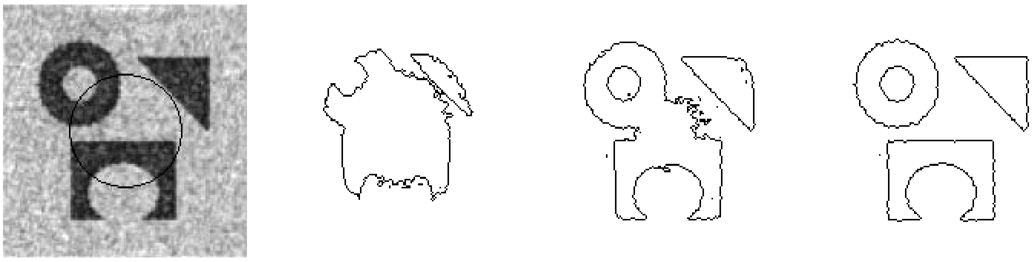


Figure 12. Segmentation example using Chan-Vese's model: initial contour (left), intermediate contours (centre), final contours (right).

3. TRACKING

Many current methods use a probabilistic representation of the uncertainty and stochastic filters to fuse and validate the information obtained from sensors, as well as to estimate parameters describing the dynamical system involved. For instance, in [35], the Kalman filter is used with a statistical background model to detect moving objects and a 3D coarse human shape model to constrain the shape of an upright human in complex situations. By doing it, multiple humans may be tracked in cases of occlusion, cast shadow or reflection. To update the position of the tracked persons, a search approach starts from the mean position predicted by the Kalman filter employed and is conducted for all possible positions within the neighborhood. If the tracked features overlap, a joint likelihood is used which turns the process computationally heavier. Nevertheless, in occlusion cases, the features follow the predictions of the stochastic filter. The data association problem is combined with filtering techniques to track ground targets using ground moving target indicator reports obtained from an airborne sensor in [36]. A method for the detection and tracking of multiple moving objects is presented in [37], based on particle filters to estimate the objects' states and sample based joint probabilistic data association filters to perform the matching between the features detected in the input sensor data and filters.

The Kalman filter is a widespread method for object tracking, but recently particle filters have become more usual [38]. The Kalman filter rests on the assumption that the disturbances and initial state vector are normally distributed. Under these circumstances, it is shown that the obtained mean of the conditional distribution of a state is an optimal estimator in the sense that minimizes the mean square error. However, if the normality assumption is dropped, then there is no guarantee that the filter will give the conditional mean of the state vector, [39]. So, particle filters were presented as a good alternative to the Kalman filter because they represent the conditional distribution with several particles which allows multimodal state distributions, [14]. However, these filters have revealed some problems too, such as difficulties on tracking multiple and articulated objects and, if the modeled system has reduced system noise or the observation has diminish variance, the filter may not perform well and the number of samples may collapse and lead to a single (wrong) peak. To overcome these difficulties, several solutions based on particle filters have been presented, as the scatter search particle filter, [40], and the kernel particle filter, [41]; nevertheless, particle filters still remain as an computational expensive solution, [42].

In each image of the sequence being analyzed, the matching between the features estimated by the filter used and the observations obtained is accomplished using data association techniques. Data association algorithms may include a hypothesis-validation step, which may be based, for example, on the Mahalanobis distance and its validation considering the chi-squared distribution, [43].

In [43], the tracking of line segments along image sequences is accomplished by using three independent Kalman filters: one, for the lines' centre point, another filter for the lines' length and a last one for the lines' direction. The matching step is done by using the Mahalanobis distance or geometric restrictions. To maintain the computational resources used as low as possible, it is employed a management model that decides if the tracking of a missing line segment should be maintained or stopped.

Many tracking approaches suppose that the features being tracked are permanent and only deal with temporary occlusion cases. In [44], human motion is captured with infra-red computer aided gait analysis systems, and accurate estimation of markers' positions play an important role in estimating the different joint angles for gait analysis. During occlusion cases, the 3D coordinates of each occluded marker are predicted and interpolated in each image using radial basis function neural networks.

The issue of multiple humans occlusion is considered in [45], through using the extended Kalman filter to estimate their path, but the system developed expects that merged blobs will eventually split in subsequent images. The approach proposed in [46], is also based on the idea of objects permanence, and the tracking during long periods is performed considering two levels: regions' level, where a customized genetic algorithm is used to search for optimal region tracks; and objects' level, where each object is identified by adopting adaptive appearance models, spatial distributions and inter-occlusion relationships.

Not so usual are approaches that consider the definitive disappearance of tracked features from the image scene. However, in many tracking applications, the definitive disappearance of features is not unusual; for instance, in surveillance systems people may move between different compartments. The simplest existing approaches to overcome occlusion situations can keep the track of occluded features during some previously defined number of occluded image frames or simply discard features whenever they are not detected in an image. In [35], when occlusion occurs the features follows the predictions of the Kalman filter used, and if it is occluded along a defined number of image frames, then its tracking is stopped. To increase the robustness of tracking people in image sequences, a prediction model is also used in [47] to deal with occlusion cases.

In [48], a two layer approach is used to deal with total occlusions as well as features merging and splitting: the first layer, produces a set of spatio-temporal strokes based on low-level operations, the second layer, performs a consistent classification of detected segments by using a statistical model based on Bayesian networks, and thus if an object is not detected during some instances of time then that label ceases.

In all the above approaches, the number of instances during which the tracking of missing features is maintained is user defined, and so no guarantee is given on the correctness and adequateness of that selection.

To track features' movement in image sequences was used in [38, 49] the Kalman filter combined with optimization techniques in the data association phase. With this method, the filter's robustness to occlusion and non-linear movements is improved. Moreover, in each image of the sequence, the quality of the matches between each feature predicted by the filter and each feature measured is established using the Mahalanobis distance, and the final set of matches is obtained by optimizing the sum of all involved Mahalanobis distances, [50]. By doing so, the best global correspondence set is always guaranteed. To simplify this correspondence process, it was also proposed in [51, 52], the use of an efficient approximation of the Mahalanobis distance. This combination of the Kalman filter with the optimization of correspondences allows good tracking results even if the filters restrictions are not satisfied (an often situation in many tracking applications).

In [43, 52-55], are used two management models, which can deal with the appearance, occlusion and disappearance of features during the tracking. The proposed management models handle the decision to keep the tracking of each occluded feature, taking into account its temporal behavior. Thus, features that keep appearing in the image sequence will obviously continue being tracked; however, if in the previous images there was not any measure data associated with a feature being tracked, then the tracking of that feature may be stopped.

In the first management model proposed, it was also used a management model which associated a confidence value to each new feature to be tracked, [43]. During the tracking, if the predicted state of a feature is associated with a new measured feature, its confidence value will increase, otherwise it will decrease. Thus, in this management model the decision of maintain or not the tracking of features is only based on the number of images during which it is visible or not. The second management approach proposed is based on an economics investment model instead – the net present value, [53]. The net present value approach was considered because many resemblances may be found between the evaluation of investment projects and the decision of maintain or not the tracking of occluded features. For instance, projects are managed attending to their own specificities and to the global market situation, while occluded features can be managed according to their behavior and to the global performance of the computational tracking system. Moreover, the net present value approach validates the use of cost functions to deal with the tracking of occluded features as well. Thereby, to keep or not the tracking of occluded features, it is considered the number of features being tracked, the quality of their previous matching and the tracking results. The simplicity of this management approach allows efficient and robust tracking results, with the computational cost being reduced strictly to the necessary.

For the first tracking example, [50], consider a sequence of 7 images in which 6 markers placed on a walking man's leg should be tracked, Figure 13. This image sequence was acquired using a high-speed camera with a frame rate of 200 frames/s, and the acquired images have size equal to 256x256 pixel^2. Using the Kalman filter, it was tracked the centres of the markers black dot, which in the image plane correspond to centres of blobs with size approximately equal to 3x3 pixel^2. In the first image of the sequence, only 5 markers (markers' labeling is observable in Figure 13) are visible, but in the next five images they are all visible. In the last image of the sequence, a marker was occluded but its tracking should continue, as it is suggested in Figure 14, but with higher uncertainty.

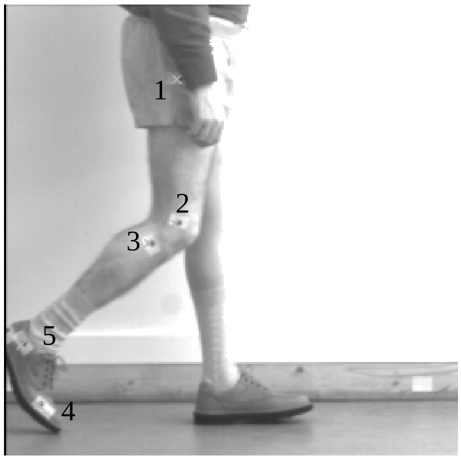


Figure 13. Tracking 6 markers in a human moving leg: first image of the sequence with 5 of the 6 markers to be tracked labeled.

Analyzing the results obtained for this first tracking example, we may notice that although the movement involved is non-linear, the error between the filters estimates and the introduced measures is always low (less than 3.35 pixels). Moreover, gradually the filter tends to get better results.

For the second tracking example, [52, 54, 55], consider a real image sequence of 547 frames with 3 mice in a lab environment, Figure 15 and Figure 16. To detect the regions of interest in each image of the sequence the background subtraction technique was used, and then the centre of each thresholded region was tracked. Several difficulties are associated with tracking the centre of the mice's regions in the acquired images. One of those difficulties arises from the fast movement of the mice, as they may go back and forth changing drastically their movement direction, Figure 15, or may move rapidly along a particularly direction, Figure 16. These nonlinear movements, which are not undertaken by the usual methods based on the Kalman Filter, can give rise to differences up to 45 pixels (images have 320x240 pixel^2) between the positions predicted by the filter and each corresponded measure (similarly behaviors were observed along xx and yy directions), Figure 17. The noise series is due to a noisy measure, which was instantly captured in image 293 and was not validated in the next measurement steps, and was consequently discarded by the management model used. Moreover, the proposed approach recovers well from the discrepancies observed, as we can see in Figure 17 where the relative maximums are quite often followed by relatively low values (below 10 pixels).

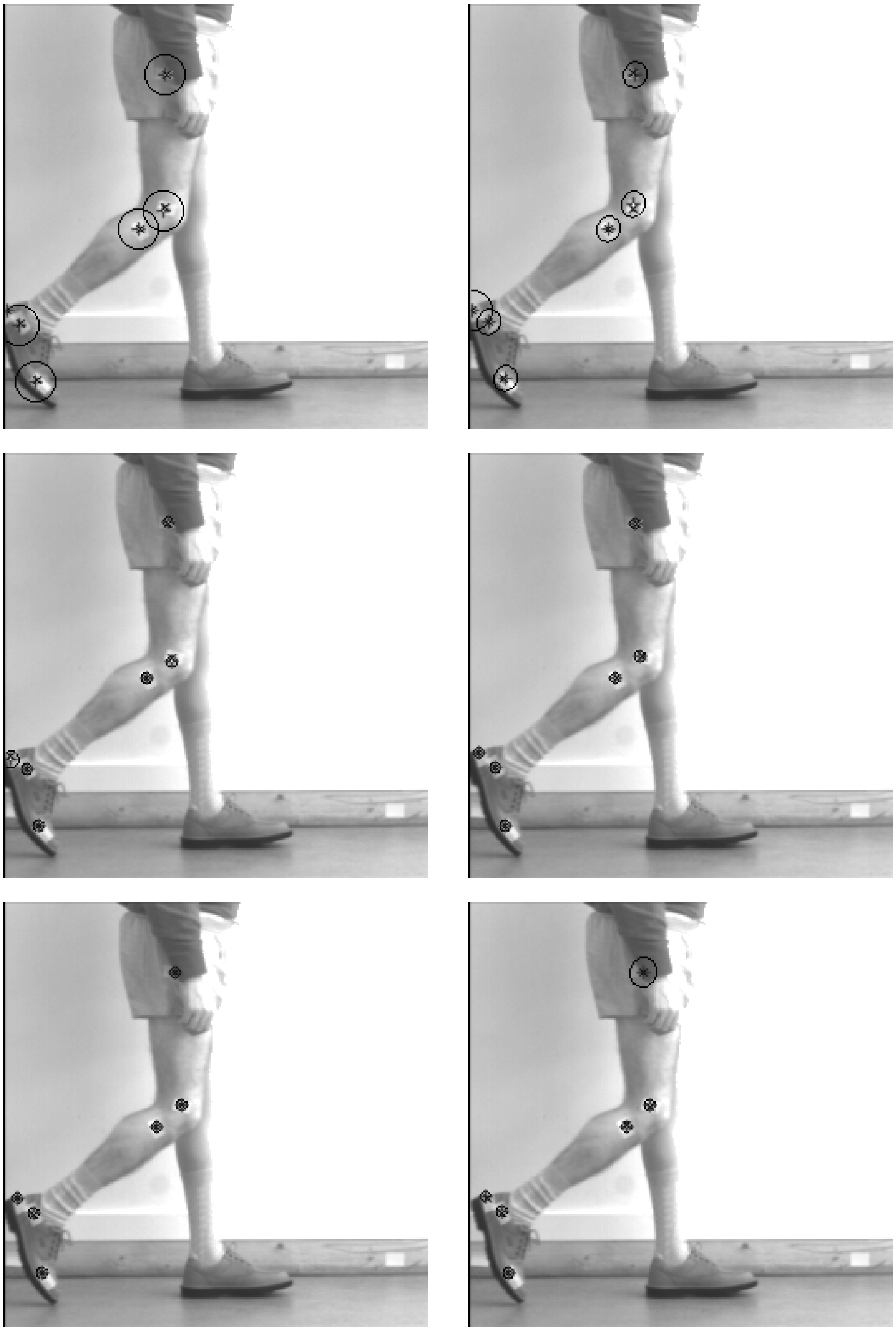


Figure 14. Tracking 6 markers in a 7 image sequence (first image in Figure 13): the Kalman Filter's search area is defined by solid ellipses, the predicted position for each marker is indicated by a + and the corrected position is represented with a x.

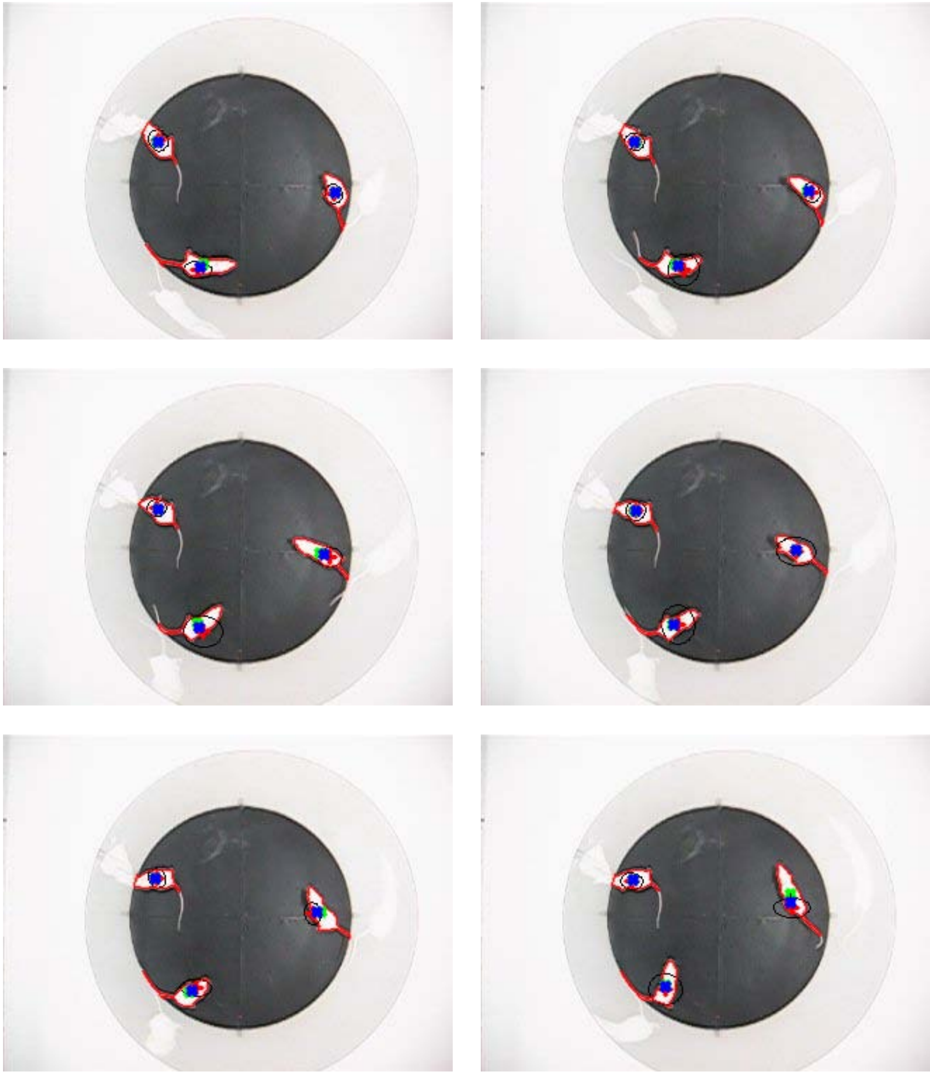


Figure 15. Tracking mice in a lab environment during 547 images: significant changes in movement's direction can be tracked correctly.

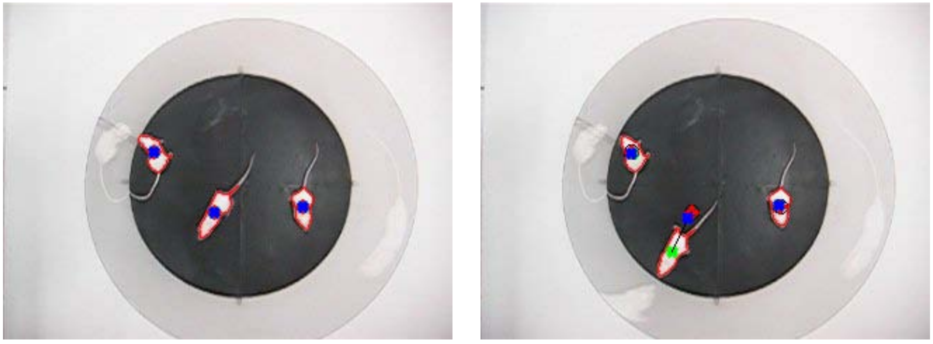


Figure 16. Tracking mice in a lab environment during 547 images: rapid movements can be tracked correctly.

For the next tracking example, [56], consider the real images acquired by an image surveillance system in a shopping centre, Figure 18 (images from [57]). In the second image, one of the persons previously tracked starts to enter a store and so the management model will stop her tracking after the sixth image. If she would come out of the store quickly, the management model could keep tracking her without any problem. However, if she comes out of the store only after several images, then the proposed approach would initialize her tracking as a new feature instead. This example shows that the method used is consistent and robust enough to be applied in long image sequences, such as those of image surveillance systems, where it is essential to free the computational resources associated with features whose tracking should be stopped.

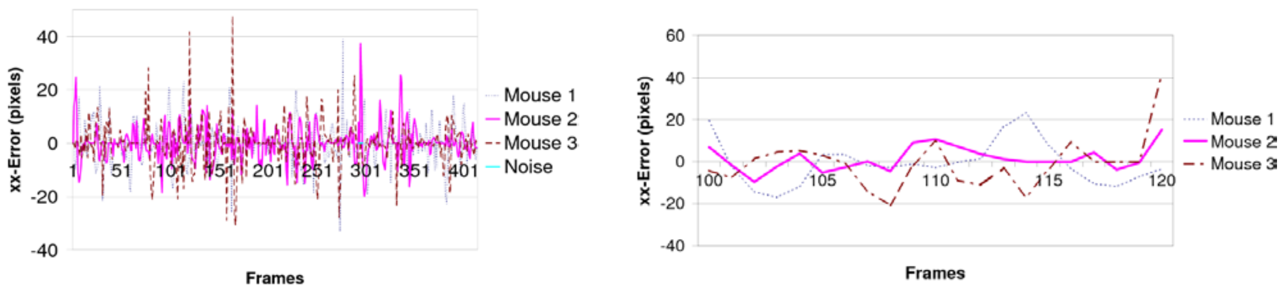


Figure 17. Tracking mice in a lab environment during 547 images: differences between the predicted positions and the associated measures along the xx direction (left); detail on the results of images 100 and 120 (right).

Figure 18. Tracking persons in a shopping centre: the used management model allows the correct tracking of features during long image sequences.

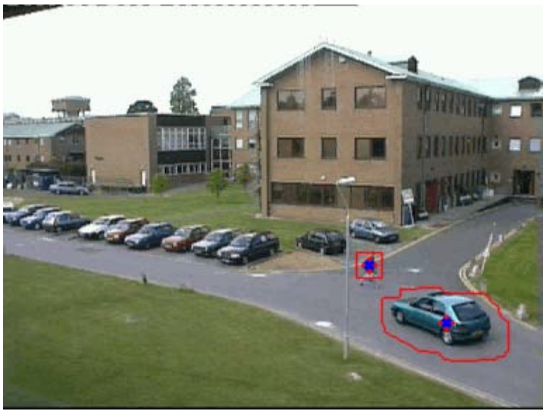
For the last experimental tracking example, [55], consider a sequence of real images of an outdoor campus scene from the PETS (Performance Evaluation of Tracking and Surveillance) 2001 datasets, Figure 19. To detect the regions of interest in this sequence, it was used the background subtraction technique, and to reduce the noise the obtained regions were then eroded and dilated. During the sequence, a person is partially occluded by a car. Thus, as they merged, the used image processing techniques only allow the detection of one region and the correspondent captured measure is attributed by the tracking system to the car. However, the persons' tracking is maintained during 3 frames if is used a low internal rate, equal to 0.02. Furthermore, with higher internal rates the person would be tracked for more time; for instance, with an internal rate of 4.0, the person would be continually tracked during 6 images. Then, after the person and the car split into two different regions, the tracking of the person is initialized and started. This example shows the influence of the internal rate of return in the tracking of missing features using the management model based on the net present value approach.

4. MATCHING AND SIMULATION

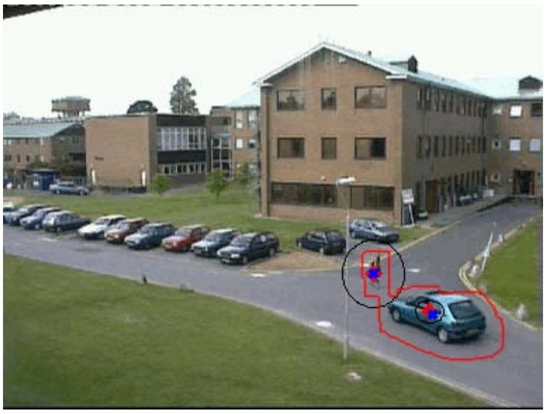
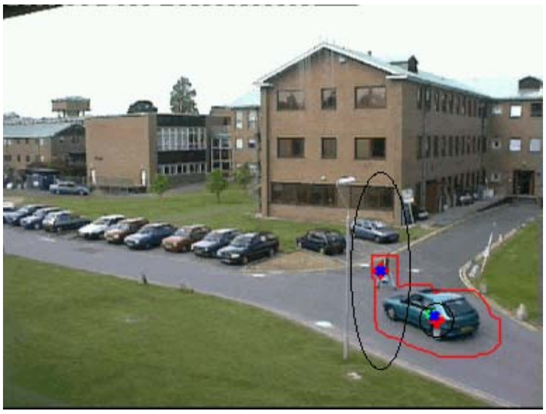
The determination of the correspondence between the data of two objects represented in images is a topic of raised importance and hard research in Computational Vision, [58, 59]; namely, because the applications that need the determination of correspondence between objects are particularly common. Some examples that can be referred are: tracking and movement analysis ([58, 60]), 3D reconstruction ([59, 61]), objects recognition ([11, 58]) and image registration ([11]).

Basically, the existing methods try to match the objects by using information that is image invariant, like curvature or displacements in a global coordinate space.

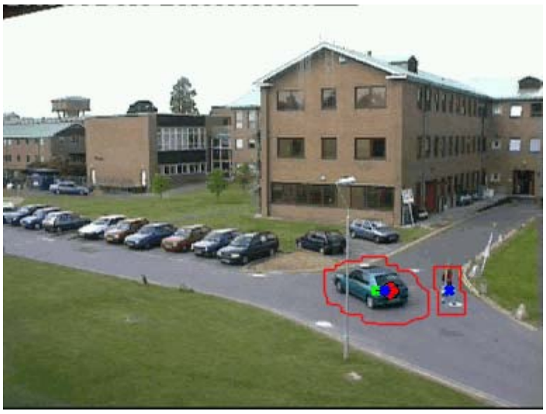
In [62], for example, it is presented a method to determine matches between points of contours represented in images, that considers curvature information and optimization of the global matching cost. Briefly, first, the angles defined by each set of three consecutive data points of each contour to be match are computed. Then, it is built the matrix of matching costs, by comparing the angles associated to each point of the contours. Finally, using an optimization algorithm, a global matching of minimum cost is found. Using this approach, the contours to match can have different numbers of points and the rigid transformation involved between them can also be estimated. In Figure 20, it is visible the matches found between the points of two contours and the estimated rigid transformation involved using the proposed method. This approach is more adequate for rigid objects because the higher the nonrigid deformations are, the more different will be the curvature values along the images.




(a) (b)

(c) (d)

(e) (f)

Figure 19. Tracking a person and a car with partial occlusion: in images (a)-(b) the person and the car are tracked separately; in (c)-(e) they are overlapped and the only captured measure is assigned to the car, so the person's tracking is continued with higher uncertainty until it is stopped; in (f) the two features are again tracked independently.

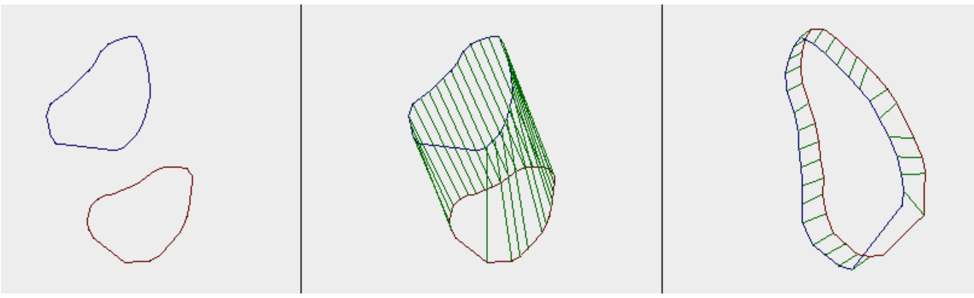


Figure 20. Two contours defined by 28 and 32 points (left), the matches found (centre) and after applying the rigid transformation estimated (right).

The matching between points of two objects can also be obtained by analyzing the displacements of those points in the respective eigenspace, commonly known as modal matching, [58]. For that, the eigenspace associated to each object is built and then the matches are found by searching for similar displacements in these spaces. For instance, in [58, 63, 64], the eigenspaces are built based on the geometrical shape of each object, then an affinity matrix is defined, whose elements, the costs of each possible matches, are found by computing the Euclidian distance between the eigenmodes associated to each object. In [63, 64], the best global matches are found by using an optimization technique in the search done on the affinity matrix. In this methodology, extra points can be matched as well by adding fictitious points in the search step, that are then matched with real points, [63-65]. This method is very fast, easy to implement and can achieve good matching results. In Figure 21, is presented an example of the results found using this method, [64].

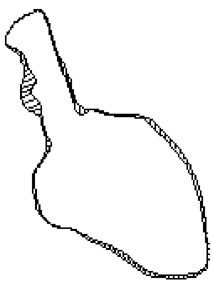


Figure 21. Matches found between two contours, one with 136 points and other with 139, using modal matching.

The optimization techniques used in [63, 64] do not consider the order of the contours' points, which sometimes can lead to crossed matches. To overcome this, it is proposed in [66-68] a new solution based on dynamic programming to achieve the best global matching without crossed matches. Using the solution proposed, the matching quality is improved and, additionally, the computation time decreases considerably.

In [58, 60, 63, 65, 69], a similar modal matching method is used, but each eigenspace is built from the finite element model of the associated object. The matches are then found in a similar away. This method is computationally more demand than the geometrical approach, but the matches found are more alike with the physical behavior expected for the objects.

In many applications, it is necessary to estimate the deformation involved between two objects or between two different instants of one object. One possible example is the estimation of the deformation of an object whose images were acquired using a large time step. In [58, 70-74], for instance, that estimation is accomplished according to physical principles by using the finite element method to model the objects, modal matching and optimization techniques to match the nodes of the models and the Lagrange equation of motion to estimate the nodal displacements. After segmenting the object in both images, the nodes of their physical models are matched and the deformation of one into the other is simulated by solving the dynamic equilibrium equation. Using the same approach as in [15] for the segmentation of an object, the internal forces in the equation are defined by the physical characteristics adopted for the model. However, the external forces are not determined as in [15]: using the fact that the final stage of the deformation is know, the external forces are defined in terms of the distance between the nodes of a contour and theirs correspondents in the other one. In Figure 22, it is presented an example of an experimental result using this physical methodology.

5. CONCLUSIONS

Computational analysis of the movement of objects represented in images is not a trivial task. A considerable part of the work done in Computational Vision is concerned with that goal. The main reason for this high effort is the huge number of possible and important applications for that analysis; namely, in medical area.

Usually, the analysis of objects' movement involves steps of objects segmentation in images, tracking of objects' features along image sequences and, very often, the matching of these features between consecutive images and even the estimation of the deformation involved between two images. In this paper, some methods were introduced to accomplish these tasks as well as some of their applications and results.

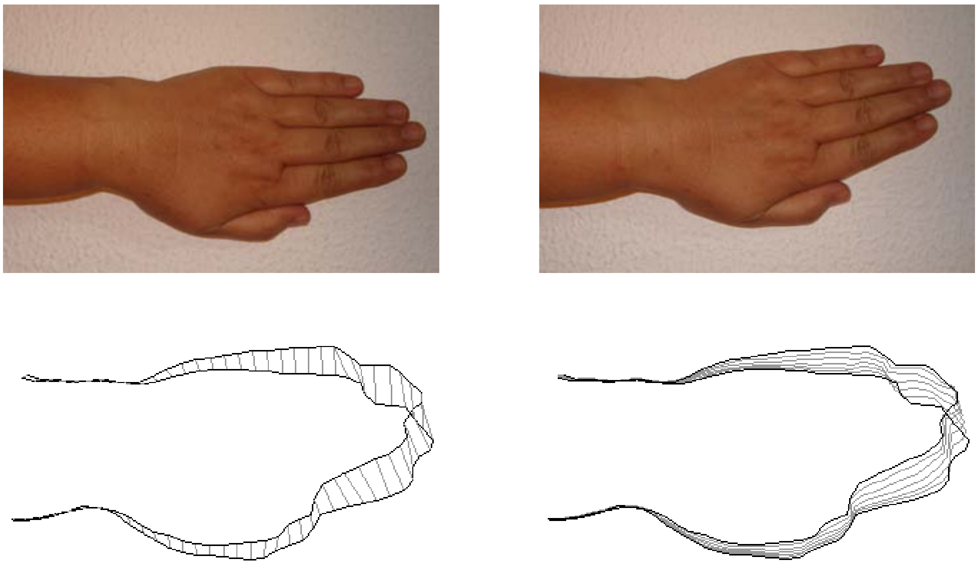


Figure 22. Original object in two different stages (top), contours and matches found between them (bottom left), and intermediate contours estimated by physical principles (bottom right).

ACKNOWLEDGMENTS

This paper was partially done in the scope of project “Segmentation, Tracking and Motion Analysis of Deformable (2D/3D) Objects using Physical Principles”, with reference POSC/EEA-SRI/55386/2004, financially supported by FCT - *Fundação para a Ciência e a Tecnologia* from Portugal.

The fourth, fifth and seventh authors would like to thank also the support of their PhD grants from FCT with references SFRH/BD/29012/2006, SFRH/BD/28817/2006 and SFRH/BD/12834/2003, respectively.

REFERENCES

[1] Y. J. Zhang, "A review of recent evaluation methods for image segmentation," in *Sixth International Symposium on Signal Processing and its Applications*, Kuala Lumpur, Malaysia, 2001.

[2] F. J. S. Carvalho and J. M. R. S. Tavares, "Metodologias para identificação de faces em imagens: Introdução e exemplos de resultados," in *Congreso de Métodos Numéricos en Ingeniería 2005*, Granada, España, 2005.

[3] R. Schalkoff, *Digital Image Processing and Computer Vision*: John Wiley and Sons Ltd, 1989.

[4] L. D. Cohen and I. Cohen, "Finite Elements for Active Contour Models and Ballons for 2-D and 3-D Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, pp. 1131-1147, 1993.

[5] T. Cootes, A. Hill, C. Taylor, and J. Haslam, "The use of active shape models for locating structures in Medical Images," *Image and Vision Computing*, vol. 12, pp. 355-366, 1994.

[6] T. Cootes and C. Taylor, "Active Shape Models: Smart Snakes," in *British Machine Vision Conference*, Leeds, UK, pp. 267-275, 1992.

[7] T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active Shape Models - Their Training and Applications," *Computer Vision Image Understanding*, vol. 61, pp. 38-59, 1995.

[8] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 681-685, 2001.

[9] J. Sung, T. Kanade, and D. Kim, "A Unified Gradient-Based Approach for Combining ASM into AAM," *International Journal of Computer Vision*, vol. 75, pp. 297-309, 2007.

[10] M. J. Vasconcelos and J. M. R. S. Tavares, "Automatic Modelling Image Represented Objects using a Statistic based Approach," in *The European Simulation and Modelling Conference (ESM)*, Porto, Portugal, pp. 7-12, 2005.

[11] M. J. M. Vasconcelos and J. M. R. S. Tavares, "Methodologies to Build Automatic Point Distribution Models for Faces Represented in Images," in *CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications*, Coimbra, Portugal, pp. 435-440, 2006.

[12] M. J. M. Vasconcelos and J. M. R. S. Tavares, "Methods to Automatically Built Point Distribution Models for Objects like Hand Palms and Faces Represented in Images," *Computer Modeling in Engineering & Sciences*, (in press).

[13] F. J. S. Carvalho and J. M. R. S. Tavares, "Eye detection using a deformable template in static images," in *VIPimage - I ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing*, Porto, Portugal, pp. 209-215, 2007.

[14] A. Blake and M. Isard, "Active contours: the application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion," Springer, 1998.

[15] P. C. T. Gonçalves, J. M. R. S. Tavares, and R. N. Jorge, "Segmentation of objects in images using physical principles," in *VIPimage - I ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing*, Porto, Portugal, pp. 329-334, 2007.

[16] S. Osher and J. A. Sethian, "Fronts Propagating with curvature-dependent speed: algorithms based on Hamilton-Jacobi formulations," *Journal of Computational Physics,* vol. 79, pp. 12-49, 1988.

[17] J. S. Suri, S. Singh, S. Laxminarayan, X. Zeng, K. Liu, and L. Reden, "Shape Recovery Algorithms Using Level Sets in 2D/3D Medical Imagery: A State-of-the-Art Review," *IEEE Transactions on Information Technology in Biomedicine,* vol. 6, pp. 8-28, 2002.

[18] J. A. Sethian, *Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry, Fluid Mechanics, Computer Vision, and Materials Science*, 2 ed.: Cambridge University Press, 1999.

[19] T. McInerney and D. Terzopoulos, "Deformable Models in Medical Image Analysis: A Survey," *Medical Image Analysis,* vol. 1, pp. 91-108, 1996.

[20] D. L. Pham, C. Xu, and J. L. Prince, "Current Methods in Medical Image Segmentation," *Annual Review of Biomedical Engineering,* vol. 2, pp. 315-338, 2000.

[21] Y. Qu, T.-T. Wong, and P. A. Heng, "Image Segmentation Using the Level Set Method," in *Deformable Models*, J. S. Suri and A. Farag, Eds.: Springer, pp. 95-161, 2007.

[22] D. Cremers, M. Rousson, and R. Deriche, "A Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape," *International Journal of Computer Vision,* vol. 72, pp. 195-215, 2007.

[23] R. Malladi and J. A. Sethian, "An O(N log N) algorithm for shape modeling," *Proceedings of the National Academy of Sciences,* vol. 93, 1996.

[24] R. Malladi and J. A. Sethian, "Level Set and Fast Marching Methods in Image Processing and Computer Vision," in *IEEE International Conference on Image Processing*, Lausanne, Switzerland, pp. 16-19, 1996.

[25] R. Malladi, J. A. Sethian, and B. Vemuri, "A Topology Independent Shape Modeling Scheme," in *SPIE Conference on Geometric Methods in Computer Vision II*, San Diego, California, USA, 1993.

[26] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Evolutionary Fronts for Topology-Independent Shape Modeling Recovery," in *Third European Conference on Computer Vision*, Stockholm, Sweden, pp. 3-13, 1994.

[27] R. Malladi, J. A. Sethian, and B. C. Vemuri, "Shape modeling with front propagation: a level set approach," *Transactions on Pattern Analysis and Machine Intelligence,* vol. 17, pp. 158-175, 1995.

[28] V. Caselles, F. Catté, T. Coll, and F. Dibos, "A Geometric Model for Active Contours in Image Processing," *Numerische Mathematik,* vol. 66, pp. 1-31, 1993.

[29] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic Active Contours," *International Journal of Computer Vision,* vol. 22, pp. 61-79, 1997.

[30] S. Kichenassamy, A. Kumar, P. Olver, A. Tannenbaum, and A. Yezzi, "Conformal curvature flows: From phase transitions to active vision," *Archive for Rational Mechanics and Analysis,* vol. 134, pp. 275-301, 1996.

[31] K. Siddiqi, Y. B. Lauzière, A. Tannenbaum, and S. W. Zucker, "Area and length minimizing flows for shape segmentation," *IEEE Transactions on Image Processing,* vol. 7, pp. 433-443, 1998.

[32] M. E. Leventon, W. E. L. Grimson, and O. Faugeras, "Statistical Shape Influence in Geodesic Active Contours," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)*, Hilton Head Island, SC, USA, pp. 316-323, 2000.

[33] T. F. Chan and L. A. Vese, "Active Contour without Edges," *IEEE Transactions on Image Processing,* vol. 10, pp. 266-277, 2001.

[34] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision,* vol. 1, pp. 321-331, 1987.

[35] T. Zhao and R. Nevatia, "Tracking multiple humans in complex situations," *IEEE Transactions on Image Processing,* vol. 26, pp. 1208-1221, 2004.

[36] L. Lin and Y. Bar-Shalom, "New assignment-based data association for tracking move-stop-move targets," *IEEE Transactions on Aerospace and Electronic Systems,* vol. 40, pp. 714-725, 2004.

[37] A. Almeida, J. Almeida, and R. Araujo, "Real-time tracking of multiple moving objects using particle filters and probabilistic data association," *Automatika*, vol. 46, pp. 39-48, 2005.

[38] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, pp. 174-188, 2002.

[39] P. S. Maybeck, *Stochastic Models, Estimation and Control*, vol. 3: Academic Press, 1982.

[40] J. Pantrigo, A. Sanchez, K. Gianikellis, and A. Montemayor, "Combining particle filter and population-based metaheuristics for visual articulated motion tracking," *Electronic Letters on Computer Vision and Image Analysis*, vol. 5, pp. 68-83, 2005.

[41] C. Chang and R. Ansari, "Kernel particle filter for visual tracking," *IEEE Signal Processing Letters*, vol. 12, pp. 242-245, 2005.

[42] T. Petrie, "Tracking bouncing balls using kalman filters and condensation," University of Colorado, Colorado, USA, 2004.

[43] J. Tavares and A. Padilha, "Matching lines in image sequences with geometric constraints," in *RecPad’95 - 7th Portuguese Conference on Pattern Recognition*, Aveiro, Portugal, 1995.

[44] H. Lakany, G. Hayes, M. Hazlewood, and S. Hillman, "Human Walking: Tracking and Analysis," in *IEEE Electronic & Communications, Colloquium on Motion Analysis and Tracking*. vol. 5, pp. 1-14, 1999.

[45] R. Rosales and S. Sclaroff, "Improved Tracking of Multiple Humans with Trajectory Prediction and Occlusion Modelling," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Workshop on the Interpretation of Visual Motion*, Santa Barbara, California, USA, 1998.

[46] Y. Huang and I. Essa, "Tracking multiple objects through occlusions," in *IEEE Conference on Computer Vision and Pattern Recognition 2005 (CVPR 2005)*, San Diego, USA, pp. 1051-1058, 2005.

[47] R. Plänkers and P. Fua, "Articulated Soft Objects for Video-based Body Modelling," in *International Conference on Computer Vision and Pattern Recognition*, Vancouver, Canada, pp. 394-401, 2001.

[48] P. Jorge, A. Abrantes, and J. Marques, "On-line tracking groups of pedestrians with bayesian networks," in *Workshop PETS* Prague, Czech Republic, 2004.

[49] G. Welch and G. Bishop, "An introduction to kalman filter," University of North Carolina at Chapel Hill, 1995.

[50] R. Pinho, J. Tavares, and M. Correia, "Human movement tracking and analysis with kalman filtering and global optimization techniques," in *II International Conference On Computational Bioengineering*, Lisbon, Portugal, 2005.

[51] R. R. Pinho, J. M. R. S. Tavares, and M. V. Correia, "Efficient Approximation of the Mahalanobis Distance for Tracking with the Kalman Filter," in *CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications*, Coimbra, Portugal, 2006.

[52] R. R. Pinho, J. M. R. S. Tavares, and M. F. V. Correia, "Efficient Approximation of the Mahalanobis Distance for Tracking with the Kalman Filter," *International Journal of Simulation Modelling*, vol. 6, pp. 84-92, 2007.

[53] R. Pinho, M. Correia, and J. Tavares, "An improved management model for tracking multiple features in long image sequences," in *6th WSEAS International Conference on Signal Processing, Computational Geometry Artificial Vision (ISCGAV'06)*, Crete, Greece, 2006.

[54] R. R. Pinho, J. M. R. S. Tavares, and M. F. P. V. Correia, "An Improved Management Model for Tracking Missing Features in Computer Vision Long Image Sequences," *WSEAS Transactions on Information Science and Applications*, vol. 4, pp. 196-203, 2007.

[55] R. R. Pinho, J. M. R. S. Tavares, and M. V. Correia, "An Efficient and Robust Tracking System using Kalman Filter," in *VIP Symposia on Internet related research with elements of M+I+T++* Venice, Italy, 2006.

[56] R. R. Pinho, J. M. R. S. Tavares, and M. V. Correia, "A Movement Tracking Management Model with Kalman Filtering, Global Optimization Techniques and Mahalanobis Distance," in *ICCMSE 2005 - International Conference of Computational Methods in Sciences and Engineering*. vol. 4A Loutraki, Korinthos, Greece: Brill, pp. 463-466, 2005.

[57] CAVIAR: Context Aware Vision using Image-based Active Recognition, EC's Information Society Technology's programme project IST 2001 37540, http://homepages.inf.ed.ac.uk/rbf/CAVIAR/, accessed in 2004.

[58] J. M. R. S. Tavares, "Análise de Movimento de Corpos Deformáveis usando Visão Computacional." vol. PhD: Faculdade de Engenharia, Universidade do Porto, 2001.

[59] J. Maciel, "Global Matching: Optimal solution to correspondence problems." vol. PhD: Instituto Superior Técnico, Portugal, 2002.

[60] J. Tavares, J. Barbosa, and A. Padilha, "Matching Image Objects in Dynamic Pedobarography," in *RecPad 2000 - 11th Portuguese Conference on Pattern Recognition*, Porto, Portugal, 2000.

[61] T. Azevedo, J. Tavares, and M. Vaz, "Development of a Computer Platform for Object 3D Reconstruction using Active Vision Techniques," in *VISAPP 2006 - First International Conference on Computer Vision Theory and Applications*, Setúbal, Portugal, 2006.

[62] F. P. M. Oliveira and J. M. R. S. Tavares, "Matching contours in images using curvature information," in *VIPimage - I ECCOMAS Thematic Conference on Computational Vision and Medical Image Processing*, Porto, Portugal, pp. 375-379, 2007.

[63] L. F. Bastos and J. M. R. S. Tavares, "Matching of Objects Nodal Points Improvement using Optimization," *Inverse Problems in Science and Engineering*, vol. 14, pp. 529-541, 2006.

[64] L. F. Bastos and J. M. R. S. Tavares, "Objects Matching Improvement using Optimization Techniques in a Geometric Modal Methodology," in *WSCG'2004 - 12-th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision'2004*, Plzen, Czech Republic, 2004.

[65] L. F. Bastos and J. M. R. S. Tavares, "Improvement of Modal Matching Image Objects in Dynamic Pedobarography using Optimization Techniques," in *AMDO 2004 - Third International Workshop on Articulated Motion and Deformable Objects* Palma de Mallorca, Spain: Springer, 2004.

[66] F. P. M. Oliveira and J. M. R. S. Tavares, "Algorithm of dynamic programming for optimization of the global matching between two contours defined by ordered points," in *ICCES08 - International Conference on Computational and Experimental Engineering & Sciences* Honolulu, Hawaii, pp. 861-862, 2008.

[67] F. P. M. Oliveira and J. M. R. S. Tavares, "Algorithm of dynamic programming for optimization of the global matching between two contours defined by ordered points," *Computer Modeling in Engineering & Sciences*, vol. 31, pp. 1-11, 2008.

[68] F. P. M. Oliveira and J. M. R. S. Tavares, "Optimization of the global matching between two contours defined by ordered points using an algorithm based on dynamic programming," *Journal of Computer Science,* vol. 7, pp. 12-20, 2008.

[69] J. M. R. S. Tavares and L. F. Bastos, "Improvement of Modal Matching Image Objects in Dynamic Pedobarography using Optimization Techniques," *Electronic Letters on Computer Vision and Image Analysis,* vol. 5, pp. 1-20, 2005.

[70] P. C. T. Gonçalves, R. R. Pinho, and J. M. R. S. Tavares, "Physical Simulation using FEM, Modal Analysis and the Dynamic Equilibrium Equation," in *CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications*, Coimbra, Portugal, pp. 197-204, 2006.

[71] R. R. Pinho and J. M. R. S. Tavares, "Morphing of Image Represented Objects using a Physical Methodology," in *SAC 2004 - The 19th ACM Symposium on Applied Computing* Nicosia, Cyprus: ACM, 2004.

[72] R. R. Pinho and J. M. R. S. Tavares, "Transitional Object's Shape Simulation by Lagrange's Equation and Finite Element Method," in *The IASTED International Conference on Applied Simulation and Modelling (ASM 2004)*, Rhodes, Greece, 2004.

[73] R. Pinho and J. Tavares, "Dynamic Pedobarography Transitional Objects by Lagrange's Equation with FEM, Modal Matching and Optimization Techniques," in *ICIAR 2004 - International Conference on Image Analysis and Recognition*, Porto, Portugal, pp. 92-99, 2004.

[74] P. C. T. Gonçalves, J. M. R. S. Tavares, and R. M. N. Jorge, "Segmentation and Simulation of Objects Represented in Images using Physical Principles," *Computer Modeling in Engineering & Sciences*, vol. 32, pp. 45-55, 2008.