Image Processing and Analysis: Applications and Trends

João Manuel R. S. Tavares 1*

1 Dept. of Mechanical Engineering (DEMec), Faculty of Engineering (FEUP), University of Porto (UP), Porto, PORTUGAL
Institute of Mechanical Engineering and Industrial Management (INEGI), Porto, PORTUGAL
(Email: tavares@fe.up.pt)
*Corresponding Author

Abstract
The computational analysis of images is challenging as it usually involves tasks such as segmentation, extraction of representative features, matching, alignment, tracking, motion analysis, deformation estimation, and 3D reconstruction. To carry out each of these tasks in a fully automatic, efficient and robust manner is generally demanding. The quality of the input images plays a crucial role in the success of any image analysis task. The higher their quality, the easier and simpler the tasks are. Hence, suitable methods of image processing such as noise removal, geometric correction, edges and contrast enhancement or illumination correction are required.

Despite the challenges, computational methods of image processing and analysis are suitable for a wide range of applications. In this paper, the methods that we have developed for processing and analyzing objects in images are introduced. Furthermore, their use in applications from medicine and biomechanics to engineering and materials sciences are presented.

Keywords
Segmentation, Matching, Registration, 3D Reconstruction, Neuronal Networks, Deformable Models, Level Set Methods, Stochastic Filters, Volumetric Methods

1 Introduction
The computational analysis of objects in images is a very challenging issue as it usually involves automatic tasks for segmentation, that is, the detection of the objects represented, extraction of representative features from the objects, matching between images, rigid and non-rigid alignment of images, temporal tracking and motion analysis of features in image sequences,
deformation estimation between two objects, as well as the 3D shape reconstruction of the objects from these images. Although, to carry out each of these tasks in a fully automatic, efficient and robust manner is generally demanding, some of these tasks often appear associated. For example, to analyze the behavior of organs from sequences of medical images, first the input images should be segmented, then suitable features of the organs under analysis should be extracted and tracked along the image sequences and finally the motions involved should be tracked and analyzed.

The quality of the input images plays a crucial role in the success of any computational image analysis task, as the higher their quality is, the easier and simpler the task can be. Hence, to improve the original quality of the input images, suitable methods of computational image processing, such as noise removal, geometric correction, edges and contrast enhancement and illumination correction or homogenization, are required.

Despite the inherent difficulties, computational methods of image processing and analysis provide a wide range of important applications for our society. Applications regarding 2D, 3D or even 4D data can be easily found in surveillance, virtual reality, medicine, engineering, biomechanics, bioengineering and materials sciences.

In this paper, the computational methods of image processing and analysis that we have developed in order to analyze objects from images are introduced; particularly, those which have been used for image segmentation, matching, alignment, tracking, as well as for 3D shape reconstruction from images. Furthermore, their use in applications from medicine and biomechanics to engineering and materials sciences are going to be presented and discussed.

This paper is organized as follows: in the next section, segmentation of objects in images is introduced with some of the methods that we have applied and some of their results. In the third part we talk about, the methods which we have been working on to match object nodes between images, to register objects in images as well as to estimate the deformation involved between two objects in images together with some of their experimental results. In the fourth section, the problem of tracking objects along image sequences is introduced showing some of our works in this domain and their respective results are presented. The 3D reconstruction of object shapes from 2D images is presented in the fifth section, along with some experimental results. Finally in the last section our conclusions.

2 Segmentation

In the computational vision domain, the identification of objects represented in images is commonly known as segmentation. For this task computational methods that are based on template matching, statistical modeling, deformable templates, deformable models, level set methods or neuronal networks [1-2] are frequently used. In short, to accomplish a segmentation task, one can model the objects in images or the backgrounds of the images.

Template matching is used, for example, in [3] for the identification of human eyes in images. Thus, a suitable template image of a human eye is used to search for an eye in input images through image correlation. After the image correlation, the centers of the regions in the input images that are more alike, to the image template used, will have the highest correlation.
values, Figure 1. This method is very simple and straightforward, but presents several limitations to deal with, for example, geometric deformations or differences in illumination.

![Figure 1](image1.png)

*Figure 1.* Image template (left) used to detect human eyes in images (right).

Additionally, [3] describes a method to identify skin areas in input images that uses sample images of skin to build a statistical model. Then the model identifies pixels of regions in the input images that have high probability to be skin, Figure 2.

![Figure 2](image2.png)

*Figure 2.* Skin regions found (right) using a statistical model built from sample skin images (left) in an original image (center).

Statistical approaches, like the one used to build point distribution models, are frequently employed to extract the most representative characteristics of objects from images. Then, improved models, such as active shape models and active appearance models can segment the objects modeled in new images. In these statistical modeling approaches, the models are built from training examples, described by a set of labeled points, combining the geometrical and appearance information of the objects [4]. Thus, the models built infer the mean shape and appearance of the objects as well as the admissible deviations relatively to that mean configuration.

Figure 3 shows the segmentation process of a hand by using an active shape model. Recently, this statistical modeling was used to analyze and simulate the vocal track shape during speech production from magnetic resonance (MR) images [5].

![Figure 3](image3.png)

*a) b) c) d)*

*Figure 3.* New image with the initial position of the mean shape model built overlapped (a) and the results of the segmentation process after 9, 19 and 29 iterations (b, c, d).

Therefore, the statistical modeling of objects is only feasible if a training set with images of the object under study is available, and it is only adequate to segment configurations of that object similar to the ones observed in the training set.

Another common methodology for object segmentation employs geometric templates that are defined in function of the objects to be segmented. For example, in [6] this geometry is used for the segmentation of human eyes. Thus, the template used is composed of one circle and two parabolas. Then, image preprocessing operators enhance specific features of the objects in the input images in order to define the energy fields to interactively deform the
template and accomplish the segmentation, Figure 4. For this segmentation to be successfully, the templates must be adequately defined and should be placed in the input images near the objects to be segmented.

![Figure 4](image)

**Figure 4.** Detection of a human eye by using a deformable template: (a) template used; images of the energy fields based on intensity levels (b), i.e. the original image; edges (c); intensity valleys (d); intensity peaks (e) and resulting segmentation (f).

Usually, objects of free form are better segmented by deformable models [1-2]. In these methods, elastic models are placed in the input images near the objects to be segmented, and then the models are deformed in order to perform the segmentation. This deformation process is driven by image forces, computed by enhancing some particular characteristics of the objects, like intensity edges. For instance, in [7], after manually defining rough contours for the objects to be segmented in the input image, the contours are modeled according to physical principles using the finite elements method (FEM). Then, the contours modeled move toward the borders of the objects driven by the dynamic equation that describes the equilibrium between the internal and external forces applied at the nodes (data points) of the models. The internal forces are defined by the physical characteristics adopted for the models. And these characteristics turn are defined by the virtual materials chosen and the level of interaction selected between the nodes of the models. The external forces are calculated considering the intensity of the image pixels, edge values and distances from each pixel to the nearest edge. Figure 5 shows an example of the results obtained using this approach.

![Figure 5](image)

**Figure 5.** Initial contour user defined (left); the result of the segmentation process considering a finite element model 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 [1-2]. The idea behind this method is to embed the moving contours into higher dimensional level set functions. The moving interfaces can be seen as the zero level set of the functions used. Then, instead of moving the points of the contours, one can track the zero level set of the functions. 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 [1-2]. Therefore, the computational complexity is greatly decreased. An example of the use of this kind of models is illustrated in Figure 6.

However, the original level set models suffered from the drawback of leaking [1-2]. In other words, these models mainly relied on the gradient information. If the images are noisy or
blurred, the contours either stop in wrong positions or leak into erroneous objects. This is a serious problem as the real images are not usually acquired under ideal conditions; for example, many medical images suffer from partial volume effects due to the heterogeneity of the acquired magnetic field.

Compared with the original segmentation models, the models that we have been working on recently have more sophisticated theoretical backgrounds and therefore, are usually more effective. For example, in [8] prior shape information was incorporated into an active contour model to segment the pelvic floor in MR images of the female pelvic cavity, Figure 7.

Another common approach to segment objects in images is based on artificial neuronal networks. The fundamental paradigm of neural networks is to construct a composed model using a considerable number of units, known as neurons that constitute very simple processing units, with a great number of connections between them. The information among the neurons employed in the network is transmitted through the associated synaptic weights. The flexibility of the artificial neural networks as well as their capacity to learn and to generalize the learned information is very attractive and important, justifying their wide use.

We have used neuronal networks to segment regions in images from the Material Science areas. For example, in [9-10], we used neuronal networks to identify material microstructures from metallographic images; a neuronal network is used in [11] to identify the delaminated regions in composite plates in radiographic images coming from drilling operations; in [12], the secondary phases of a nickel alloy are identified in SEM images by a neuronal network, Figure 8; and finally, a neuronal network is used in [13] to quantify the porosity of a synthetic material from optical microscopic images. There are two main problems with the use of artificial neuronal network: the need of a training set of images, well representative of the objects to be segmented; the topologies of the neuronal networks are established in terms of the application, therefore restricting their applicability and adaptability.

3 Matching, Registration and Simulation

Matching the data of two objects, or of two configurations of one object, represented in images is a topic of great importance and
intense research in Computational Vision; particularly, due to the huge number of potential applications. Some examples that should be mentioned are: tracking and movement analysis, 3D reconstruction, object recognition and image registration. Basically, the existing methods try to match the objects by using information that is an invariant image like curvature or displacements in global coordinate spaces.

Figure 8. Original SEM image (left) and the segmentation obtained by a neuronal network (right).

For example [14] introduces a method to determine matches between points of contours represented in images that consider curvature information and optimization of the global matching cost. Figure 9 shows the matches found between the points of two contours and the estimated rigid transformation involved using the proposed method.

Figure 9. Two contours defined by 28 and 32 points (left), the matches found (center) and after applying the estimated transformation (right).

Additional to the curvature information [15] takes into account the distances to the centroid points in order to obtain more reliable results. Despite the good results that these approaches can achieve, they are more adequate for rigid objects because the higher the non-rigid deformations are, the more different the geometrical parameters will be.

The matching between points of two objects can also be obtained by analyzing the displacements of those points in the eigenspaces, using the well known modal matching approach [16]. In the modal approach, the eigenspace associated to each object is built and then the matches are found by searching for similar displacements in those spaces. Thus, in [16-18], 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 correspondence, are found by computing the Euclidian distance between the eigenmodes associated to each object. In [17-18], the best global matches are found by using an optimization technique in the search carried out on the affinity matrix. With this methodology, when two objects to be matched have a different number of points, the extra points can also be matched by considering fictitious points in the search step that are then matched with real points [17-19]. Figure 10 shows an example of the results found using this method [18], which is very fast and can achieve good matching results. However, the tuning of the method parameters is highly dependent on the shape of the object.

The optimization techniques used in [17-18] do not consider the order of the contour points, which sometimes can lead to crossed correspondences. To overcome this a new solution based on dynamic programming was proposed in [14-15]. Using this solution, the matching quality is improved and, additionally, the tuning of the parameters of the matching
method becomes easier and the computation time decreases considerably.

Figure 10. Matches found between two contours, one with 136 and other with 139 points, using modal matching.

In [16-17, 19-21], a similar modal matching approach is used, but each eigenspace is built from the finite element model of the associated object, and then the matches are found in an analogous way. This method is computationally more demanding than the geometrical approach, but the matches found are more alike to the physical behavior expected for the objects.

The registration of objects in images, that is, the alignment of the objects represented, is a very important task in image analysis. For example, in medicine, the registration of organs represented in medical images is crucial for a patient’s prognosis. Usually, the registration of objects is carried out by taking into consideration the characteristic features of the objects, such as the maximum curvature points, their matching and the estimation of the involved transformation. However, some difficulties arise when, for example, the key and invariant features are not easily identified, the objects are particularly occluded, the deformations involved are highly non-linear or the shape of the objects are very distinct.

Since 2008, we have been developing new methods to register objects in images [14-15, 22-23]. In [23] a registration method based on the previous matching of the contours of the objects and on optimization was applied to the alignment of pedobarographic images. These images convey the interaction between the sole of the foot and the ground that has clinical relevance in gait and posture analysis. The same kind of images were registered in [22] by using a global method based on the Fourier transform and its properties, Figure 11.

Figure 11. From left to right: template image, source image and overlapped images before and after the registration.

In many applications, it is necessary to estimate, that is, to simulate, the deformation involved between two objects or between two different configurations of one object. One possible example is the estimation of the deformation of an object whose images were acquired using a very large time step. In [7, 24-27] that estimation is accomplished according to physical principles by using FEM to model the objects, modal matching and optimization techniques to match the nodes of the models built and the Lagrange equation of motion to estimate the nodal displacements. Figure 12 shows an experimental result using this physical methodology. The main difficulty of this methodology is the selection of adequate virtual materials for the objects, as the simulated behaviors strongly depend on the properties of the materials.
Computer analysis of objects in motion in image sequences is a very complex problem, as it may involve tasks for automatic detection of objects, matching, tracking and deformation estimation. Motivated by its wide range of significant applications, considering either 2D or 3D data, like in medical imaging based diagnosis, human gait analysis, surveillance systems, traffic analysis, recognition of objects, pose estimation and deformation analysis, the computer analysis of objects in motion has evolved considerably over the last decades. For this analysis, many solutions may be used to meet the needs of each application, but constrains associated with computational complexity as well as with computation speed are usually assumed.

To analyze objects in motion along image sequences, one first needs to detect (i.e. segment) the objects of interest in each image (i.e. find in the images regions or features representative of the objects to be tracked) and then track them through consecutive images, while maintaining the correct data association (i.e. matching features between consecutive images). Two main sources of difficulty in performing computational analysis of objects in motion from image sequences are: 1) change in appearance of the objects caused by change of viewing angle, illumination conditions, topology or non-rigid deformations; and 2) situations where there is total or partial occlusion of the objects.

The Kalman filter is a widely used method for tracking objects, which is based on the assumption that the disturbances and initial state vector are distributed normally [1, 28]. Under these circumstances, the obtained mean of the conditional distribution of a state is an optimal estimator. However, if the normality assumption is dropped, then there is no guarantee that the filter will give the adequate conditional mean of the state vector. Therefore, 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. However, these filters have revealed some problems too, such as difficulties in tracking multiple and articulated objects. Additionally, if the modeled system has reduced system noise or the measured features have a very low variance, particle filters may not perform successfully or even collapse as 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, and the kernel particle filter; nevertheless, particle filters are still an expensive computational solution [1, 28].

Recently, also to overcome the linear restriction of the Kalman filter, the Unscented Kalman filter was proposed. The latter filter has been shown to be more adequate than the Kalman filter for problems when the motion involved is
non-linear. However, in problems that have linear or near-linear motions, the Kalman filter continues to be a good option, particularly when combined with a suitable data matching strategy [29].

In [30], the tracking of line segments along image sequences is carried out with three independent Kalman filters: one, for the center point of the lines, another filter for the length of the lines and the last one for the direction of the lines. The matching step is done by using the Mahalanobis distance or geometric restrictions. To track motion features in image sequences in [28] the Kalman filter combined with optimization techniques in the data association phase was used. With this strategy, the robustness of the filter 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 calculated using the Mahalanobis distance, and the final set of matches is obtained by optimizing the sum of all Mahalanobis distances involved [31].

In order to speed up the global tracking process, the use of an efficient approximation of the Mahalanobis distance in the matching phase was proposed in [32-33].

In [30, 33-35], two management models that can deal successfully with the appearance, occlusion and disappearance of features during the tracking were used. The proposed management models handle the decision to keep the tracking of each occluded feature, taking into account its historical behavior. Thus, features that keep appearing in the image sequence will obviously continue to be tracked. However, if in the previous images there was not any data associated with a feature being tracked, then the tracking of that feature may be stopped and the associated computational resources freed.

The first management model proposed associates a confidence value to each new feature to be tracked [30]. 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. Then, if the confidence value is lower than a previously defined value, the tracking of the associated feature is stopped and its resources freed. Thus, in this management model the decision to 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 [34]. The net present value approach was considered because many resemblances may be found between the evaluation of investment projects and the decision to maintain or not the tracking of occluded features. Thus, the decision to keep the tracking of occluded features or not, the number of features being tracked, the quality of their previous matching and the tracking results are taken into consideration. The simplicity of this management approach allows efficient and robust tracking results, with the computational cost being kept strictly to the necessary.

For an illustrative example of tracking features along image sequences, let’s consider a sequence of 7 images in which 6 markers placed on a leg of a man walking should be tracked, Figure 13. Using the Kalman filter, the center of the markers, which in the images corresponds to centers of blobs of a size approximately equal to 3x3 pixel$^2$, were tracked. In the first image of the sequence, only 5 markers (marker labels can be seen in Figure
were visible, but in the next five images all the markers were visible. In the last image of the sequence, marker 1 was occluded, but its tracking would continue, although with less certainty. Analyzing the results obtained, we may notice that despite the non-linearity of the motion involved, the errors between the estimations of the filter and the real features were always low (inferior to 3.35 pixels). Moreover, the filter gradually tended to get better results as we can see by the area of the ellipses included in the images.

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

5 3D Reconstruction
In this section, we introduce the methodologies we have used to reconstruct the 3D shape of objects from 2D images. The following methodologies are commonly used [36-37]:

- For external shapes: active techniques, that is, with the projection of some kind of energy on the surface of the objects or with relative motion between the imaging acquisition system and the objects; passive techniques, that is, without energy projection or relative motion; and methods of space carving;
- For internal shapes: 2D segmentation, for example, the segmentation of the contours of the objects, and then the interpolation of the segmentation features in order to obtain the surfaces of the objects.

Figure 14. Tracking 6 markers in a 7 image sequence (first image in Figure 13): the filter’s uncertainties are represented by the ellipses, the estimated positions are indicated by + and the corrected positions are represented with x.

Usually, the reconstruction of the shape of 3D objects from 2D images involves tasks of camera calibration, data segmentation, matching, triangulation and interpolation. In addition, typical difficulties are due to geometric distortions, bad or unstable illumination conditions, occlusion occurrence,
image noise, complex shapes, etc. In [38-39], the organs of a female pelvic cavity are reconstructed in 3D from MR image slices through contour segmentation and blend, Figure 15.

**Figure 15.** The organs are reconstructed in 3D (right) from a set of 2D MR image slices (left).

Seeing that the 3D reconstruction of objects from images using active vision techniques is under great demand for many engineering applications, we have developed a computational platform that integrates many of these well known techniques [40].

Recently, particularly when reconstructing the external shapes of smooth objects, like human anatomical structures, we have applied space carving methods [36-37]. These methods work on the volumetric spaces of the objects and do not require a matching process between the images used, which is usually very complex with smooth objects. Thus, typically, the 3D models are built from a sequence of images, acquired using a turntable device. For this, the bounding volumes of the objects are defined and then are carved by removing the non-photo consistent voxels, Figure 16.

**Figure 16.** Steps of a space carving method in the 3D reconstruction of a hand model from a sequence of 2D images.

Despite the considerable work that has already been done, much work still has to be done; particularly, to make the methods more robust and adaptable to different applications. In addition, the validation of the methods developed in real scenarios assumes a crucial role. For these goals to be overcome, the contribution of researchers from other scientific fields and of potential end-users is essential and most welcome.

**6 Conclusions**

In this paper, we have briefly presented the methods that we have been working on to process and analyze objects represented in images. We saw, briefly, methods and applications regarding tasks of segmentation, matching, registration, simulation, tracking and 3D reconstruction of objects.

**Acknowledgments**

The methods presented in this paper have been partially developed in the scope of projects “Methodologies to Analyze Organs from Complex Medical Images – Applications to Female Pelvic Cavity”, “Aberrant Crypt Foci and Human Colorectal Polyps: Mathematical Modelling and Endoscopic Image Processing” and “Cardiovascular Imaging Modeling and Simulation - SIMCARD”, with references PTDC/EEA-CRO/103320/2008, UTAustin/MAT/0009/2008 and UTAustin/CA/0047/2008, respectively, financially supported by FCT - Fundação para a Ciência e a Tecnologia, Portugal.

**References**


