

Multiple Representation of Complex Intensity Changes for Image Segmentation

Pedro Medina-Rodríguez and Enrique Fernández-García

Dpto. de Informática y Sistemas. Universidad de Las Palmas de G.C.
35017 Las Palmas de Gran Canaria. SPAIN

Abstract. In this paper we present a method for finding a multiple representation of complex intensity changes that is useful in image segmentation. In computer vision, it is often required to separate objects from background in images taken under conditions of poor and non uniform illumination. In these situations, it is well-known that many of the intensity changes do not correspond to the object's boundaries but to other factors (mainly shadows and brightness variation) that complicate the segmentation process. With this in mind, we propose a segmentation process that proceeds in stages where one of the most important is the local description of the intensity changes giving rise to a multiple representation where many clues are accentuated. Other higher level stages act to select, manipulate, and combine clues in agreement with certain strategies (decision rules) which can be defined based on previous knowledge or heuristic information in order to segment the entities of interest.

1 Introduction

Despite a wide variety of different segmentation techniques [1,2] no general theory of segmentation exists. The general problems of segmentation involve processing arrays of numeric values representing brightness (or/and colour, depth, etc.) in order to extract features of boundaries and regions over local areas. Traditionally, the segmentation of images has been approached by methods based on discontinuity detection (edge-based methods) or similarity in some characteristics of the image (region-based methods). Unfortunately, in real images, due to noise and other factors, neither the edge-based methods nor the region-based methods give anywhere near perfect results [3]. Another aspect to be considered is that the low level processes can only produce partitions of the image on a nonsemantic basis [4], in such a way, that they may not necessarily correspond to the entities of interest.

In applications in which an image consists of only objects and background in conditions of uniform illumination, one may attempt to segment the image by global thresholding. Threshold selection is usually based on the information contained in the gray level histogram of the image. However, problems appear when the object area is small compared to the background area or when both the object and the background assume some broad range of gray levels in cases of uneven background and poor illumination. In such cases, the histogram is no longer bimodal and the segmentation

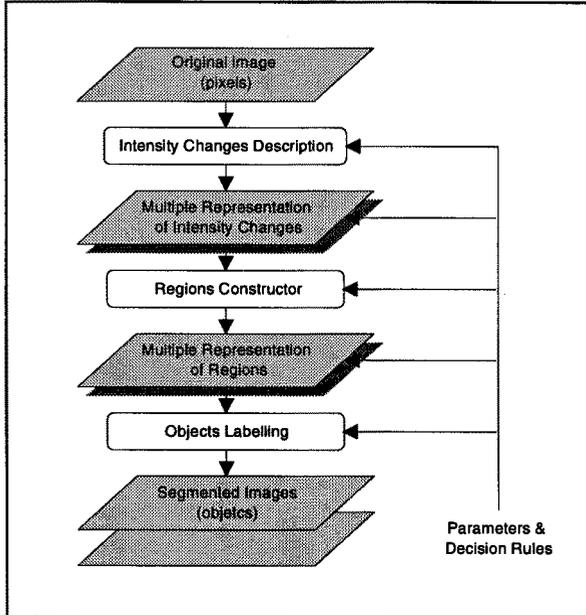


Fig. 1. The global segmentation method.

methods based on a threshold level computed from the global image histogram are often unsuccessful so there is a need for other methods [5,6,7].

Yanowitz and Bruckstein [8] have suggested a method for obtaining a threshold surface which is determined by interpolating the image gray levels at points where the gradient is high, indicating probable object edges. Thus, the gradient map and the gray level distribution give the necessary information for calculating an adaptive threshold level at each point of the image. This method has some drawbacks: it is necessary to apply a validation process to the segmented images, in order to eliminate some segmentation errors (“ghost” objects and stains) and, the edge detection at multiple scales is not considered. It’s well known that the significant changes in the image can occur at multiple resolutions: fine and coarse resolutions. Therefore, it’s necessary to use different size operators in order to detect all the intensity changes at their own scale.

In this paper, we propose a segmentation method that allow us to separate objects from background in images where variable gray levels of objects and background and changing light conditions are present. In such cases, shadows and brightness could appear to complicate the segmentation process. Fig.1 shows a block diagram of the proposed global segmentation method where the multiple representation idea plays an important role. The method has been organised in several stages which can be classified in two types: processing stages and multiple representation stages. The *processing stages* operate on the entry data according to the algorithms that in each case are applied, and whose behaviour is determined by parameters that are adjusted and associated to each particular segmentation objective. Basically, three processing stages can be distinguished: local description of the intensity changes, regions

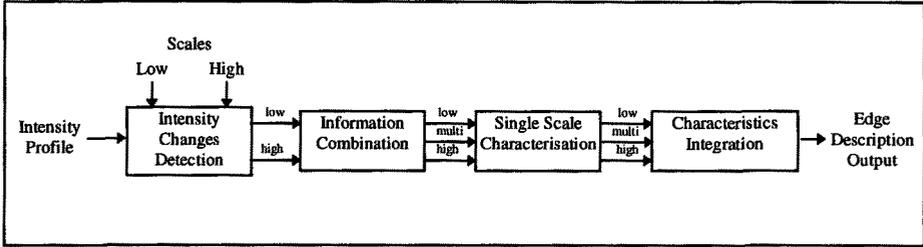


Fig. 2. Local description scheme of the intensity changes.

constructor and objects labelling. The entry data to these stages are the primitive elements (pixels, intensity changes, linear regions) of the preceding representation stage.

On the other hand, the *multiple representation stages* store new representations of the original image in which a number of clues are accentuated. Other processes act to select, manipulate, and combine clues in agreement with certain strategies (decision rules) which can be defined based on previous knowledge or heuristic information of the entities which we want to segment. In this way, we can see that each stage produces increasingly more useful descriptions of the image.

The aim of this paper is to describe the first processing stage which is an expansion process that acts on the original image in order to obtain a multiple representation of the intensity changes. This process consists in the local description of the intensity changes detected at multiple scales whose purpose is to capture all the significant properties of objects present in an image since that the first clues about physical properties of the scene are provided by the intensity changes. The proposed method acts on one-dimensional profile of intensities. Its extension to the two-dimensional case is carried out by exploring the image in two perpendicular directions, such as rows and columns which give two representations of the image that must be combined according to some criteria.

2 Local Description of the Intensity Changes

In this section, we present a method for locating and describing the intensity changes in an image that are due to diverse factors among which the most important are the changes in reflectance of visible surfaces and changes in illumination, in such a way that its geometrical structure, sharpness and contrast give enough information about the physical edges in the scene [9]. Thus, we consider that it is interesting to describe the intensity changes via a set of attributes which will help us in the detection and characterisation of regions from which the entities of interest can be obtained. An edge in an image corresponds to an intensity discontinuity in the scene. Although it may correspond to an edge of an object in space, it needs not [10]. It might well be the image of a shadow (illumination discontinuity) or surface mark (reflectance discontinuity). Fig. 2 shows a block diagram of the local description of the intensity changes that we are to address.

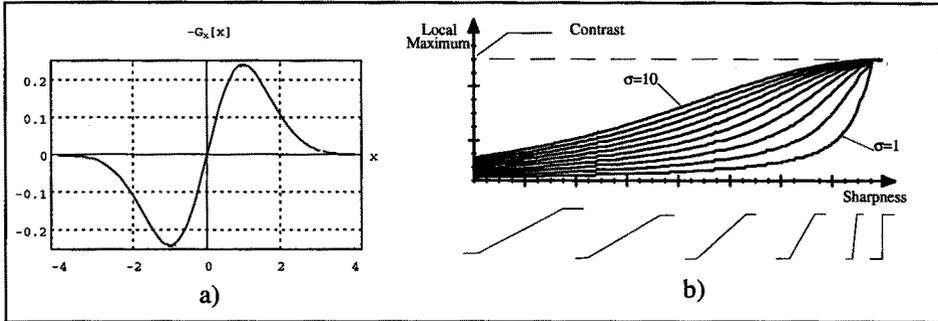


Fig. 3. a) Differential operator and b) Local maximum of peaks Vs sharpness.

2.1 Basic Detection

From the very early days of computer vision, edge detection was recognised as necessary, with a simple implementation of a gradient function. Noise sensitivity forced the inclusion of a smoothing step before differentiation. Various edge detection methods have appeared in the literature [11-17]. In this sense we are not interested in presenting a new edge detector but in capturing information about the physical edges by detecting and characterising the intensity changes by using a simple 1D edge detector, such as the first derivative of the Gaussian although others edge detectors can be used. The first derivative of the Gaussian approximates very closely an edge detector that is optimal given certain assumptions [12]. Its extension to the two-dimensional case is carried out by exploring the image in two perpendicular directions, such as rows and columns.

The detection of intensity changes is carried out by convolving the intensity profile, $I[x]$, with a differential kernel, $g'(x, \sigma) = -(x/\sigma^2)g(x, \sigma)$, obtained from the first derivative of the Gaussian as shown in Fig. 3a.

$$g(x, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

This convolution implements the basic process of smoothness and differentiation that can be found in many of the edge-detection methods. The smoothing operation serves two purposes. First, it reduces the effect of noise on the detection of intensity changes. Second, it sets the resolution or spatial scale at which the intensity changes are detected [9]. On the other hand, the physical interpretation of the convolution process is the subtraction of the mean intensities of two neighbouring regions. The standard deviation or scale σ of the Gaussian determines the size of the regions in which the mean intensities are calculated.

As a result of the convolution process a differential profile is obtained, in which the intensity changes result as peak-shaped no-null zones. The maximum value of the peaks depends both on the size of the filter and the transition sharpness supposing that there are no interference between adjacent transitions. The graph in Fig. 3b shows the

value of the local maximum of each one of the peaks for transitions of different sharpness when filters of different sizes are applied. The maximum degree of sharpness is obtained by a step-type transition. We can appreciate that for a fixed degree of sharpness and in the absence of adjacent transitions, the maximum value of the peaks increases depending on the size of the filter until a maximum value is reached which coincides with the contrast of the transition.

On the other hand, it is well-known that the intensity changes produced in an image occur at different scales or spatial resolution (coarse and fine resolutions) in such a way that its detection demands the use of filters of different sizes. In real situations, it is difficult to find a single size of filter that can detect, in an optimal manner, all of the transitions. Thus, we must use small-sized filters for the detection of transitions which correspond to fine details and large-sized filters for that corresponding to coarse details, which in many cases appears as diffuse transitions as a consequence of bad illumination conditions. Therefore, it is necessary to apply a set of filters of different sizes and their response combine in a convenient way to make the detection of intensity change at its own scale effective. Thus, in our approach we only use a small-sized filter (low-scale) for the detection of transitions which correspond to fine details and a large-sized filter (high-scale) for that corresponding to coarse details.

2.2 Combining Information from Multiple Scales

Next step is the combination of the differential responses of each individual scales in order to obtain a new differential profile called *multiscale differential profile* which serves as a reference for searching the intensity changes detected at the different scales. The solution that we have adopted for obtaining each point of the multiscale differential profile is based on taking the maximum response of the individual scales, both sign and magnitude. This can be expressed as:

$$I'(x)_{om} = \text{MAX} (I'(x)_{\sigma_1}, I'(x)_{\sigma_2}).$$

Where $I'(x)_{\sigma_1}$, $I'(x)_{\sigma_2}$ are the differential profiles that correspond to the "n" individual scales; $I'(x)_{om}$ is the new multiscale differential profile obtained from the combination process and, $\text{MAX}(a,b)$ is a function which gives the value of "a" if $|a| \geq |b|$ or "b" in otherwise. Thus, each no null zone represents the presence of an intensity change detected by the scale that gives the best response.

Fig. 4 shows an example of the process for the detection of intensity changes in a discrete range of spatial scales defined by the standard deviation σ of the Gaussian. We can distinguish between transitions of high and low spatial frequencies that correspond to coarse and fine details, respectively. A total of five graphs are shown corresponding in consecutive order to: a) the intensity profile; b) multiscale differential profile obtained by combining the results of each of the filters according to the criterion of maximum response; c) and d) differential profiles obtained by convolving the profile of intensities with filters of two different sizes: $\sigma = 0.33$ and 3.0 , respectively. Each value of the standard deviation define an operator of $\pm 3 \cdot \sigma_n + 1$ pixels. We can see for the diffuse transitions how the response of the filter increases depending on its size, reaching a value very closely to the transition contrast for the

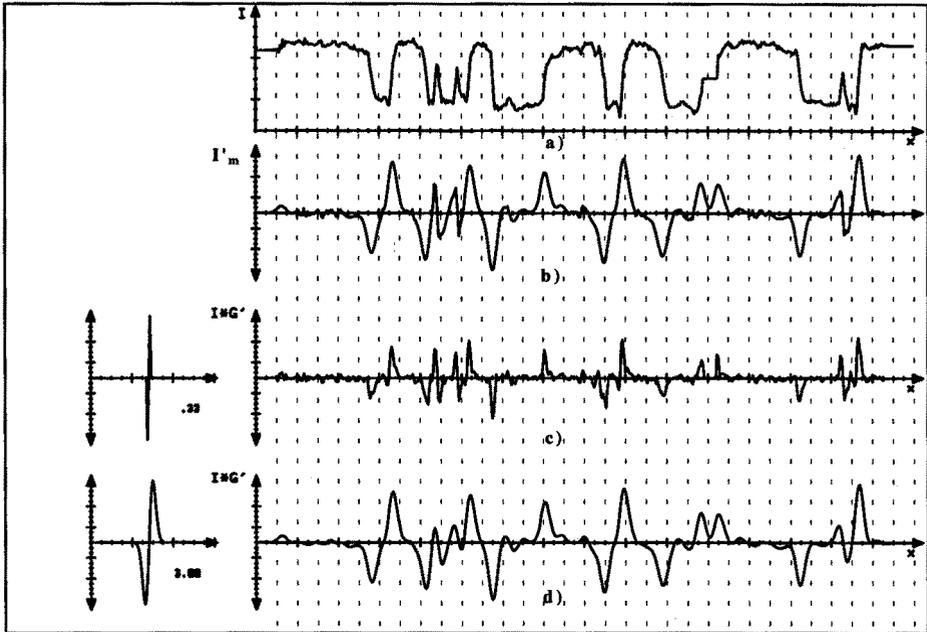


Fig. 4. Detection of intensity changes at multiple scales. a) Profile of intensities; b) Multiscale differential profile; c) and d) Differential profiles at scales $\sigma=0.33$ and 3.0 .

largest of the filters. On the contrary, for fine transitions of high spatial frequency the response from the filters is reduced since interference between adjacent transitions of different polarities appear, and the positions of the local maximum are displaced, giving rise to a possible dislocation of the edge points [18].

In the multiscale differential profile we can see how the transitions are accentuated by no-null zones except the transitions composed of several smaller transitions of the same polarity (staircase type) which perhaps give rise to only one peak or no-null zone depending of the largest scale value. In these cases, subtransitions can not be detected by analysing only the peaks of the multiscale differential profile. Therefore, an integration process of the information obtained at each individual scale is needed.

2.3 Characterising Intensity Changes

The intensity changes description consists of the extraction of a set of characteristics which is associated to each intensity change. According to Fig. 5, each no-null zone with a differential profile defines the spatial limits where the process of detection and characterisation of an intensity change at each scale is carried out. Each no-null zone or peak is characterised by the following attributes: the spatial limits of the transition (x_i-x_f), the position of the edge point (x_{ep}), the grey level of the edge point (g_{ep}), the minimum grey level (g_{min}), the contrast (con) and the tune level ($tune$). The *transition contrast* can be positive or negative, depending on whether or not it is a

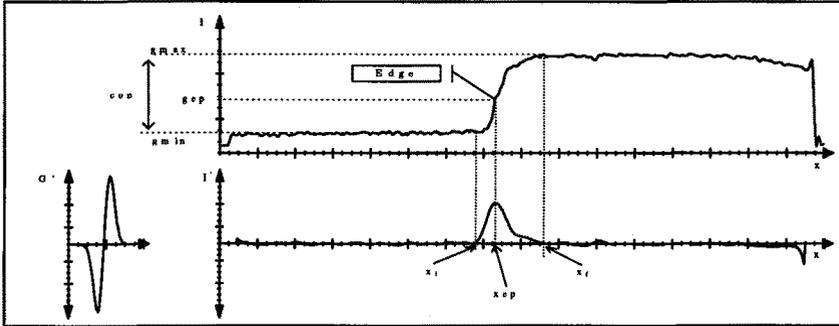


Fig. 5. Characterisation of one intensity change.

transition region from dark to light or conversely, respectively. The tune level is an attribute defined for each transition as the ratio between the local maximum of each gradient peak and the contrast of the transition.

Next, in order to obtain a complete characterisation of the intensity changes an integration process is needed where the results of the characterisation process at each scale are combined in a convenient way allowing us to attack two problems: the detection of subtransitions in staircase type transitions and dislocation of the edge points at high scales. The integration of characteristics completes the characterisation of the detected intensity changes at multiscale differential profile by two basic actions: addition of a new attribute to the intensity changes and refinement of attributes by using the lowest scale in which the intensity change is detected.

According to Fig. 6 transitions can be classified in two types: single and multiple. A transition is single when, at a lower scale, new transitions of the same polarity do not appear, and in the opposite case, it is said to be multiple. The multiple transitions are coded in such a way that they take into account, on the one hand, the coarse transition (father) detected at the multiscale differential profile, and on the other hand, the fine transitions (children) detected at the lowest scale in which it is detected. The type attribute is codified by a byte (BA) in which can be distinguished the first nibble (A) and the second nibble (B) giving rise to the following codification: Single Transition ($BA=00$); Global Multiple Transition ($BA=1n$), where "n" is the number of local transitions and, Local Multiple Transition ($BA=20$).

Once the transitions are classified the system is ready to generate the final

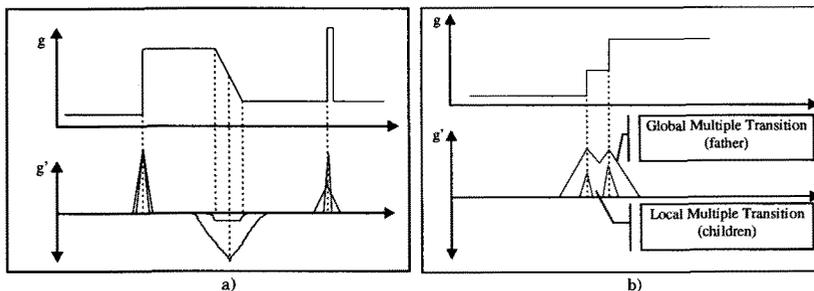


Fig. 6. Types of transitions. a) Single transitions and b) Multiple transitions.

description of the intensity changes according to the following way: if the transition is single then its final description will be taken from the lowest scale in which the intensity change has been detected, thus, we try to reduce the negative effects of the dislocation of the edges when the spatial scale is increased due to the interference of the adjacent transitions or the superimposition of various factors responsible for the intensity change, for instance, a change in reflectance and shadows. Otherwise, that is, if it is a multiple transition with "n" subtransitions, "1+n" characterisations will be generated: one characterisation for the global multiple transition detected at the multiscale profile and "n" characterisations for the local multiple transitions (subtransitions) detected at single scale. The global multiple transitions take their attributes from the characterisation at multiscale profile, while the local multiple transitions take their attributes from the characterisation at the lowest scale in which they have been detected.

Finally, the intensity changes are represented by a data structure of N+1 blocks where we distinguish between 1 head block and N data blocks. The head block supplies global information related to the description process applied to the horizontal or vertical scan image such as: number of intensity changes (*nic*), direction of exploration (*dir*) and, the system parameters that correspond to the minimum contrast (*pcont*) and the values of the low and high spatial scales (*ps1,ps2*). The remaining blocks are the local attributes associated to each intensity changes which supply information related to the 1D spatial limits of the transitions (*xi, xf*), 2D position of the edge point in the image (*col, row*), grey level of the edge point (*gep*), minimum grey level (*gmin*), contrast (*con*), tune level (*tune*), type of transition (*type*) and selection mark (*sel*).

HEAD BLOCK		DATA BLOCKS	
Block 0	Values	Block i	Values
ic[0].nic	1 ... n	ic[j].xi	1 ... 512
ic[0].dir	0: Hor. 1:Vert.	ic[j].xf	1 ... 512
ic[0].pcont	0 ... 255	ic[j].col	1 ... 512
ic[0].ps1	0.33 ... xx.yy	ic[j].row	1 ... 512
ic[0].ps2	0.33 ... xx.yy	ic[j].gép	0 ... 255
-		ic[j].gmin	0 ... 255
-		ic[j].con	-255 ... +255
-		ic[j].tune	0.0 ... 1.0
-		ic[j].type	BA 00 -> Single Transition 1n -> Global Multiple Transition 20 -> Local " " n = Number of Local Mul. Trans.
-		ic[j].sel	Selection mark

Table 1. Representation of the local description of the intensity changes.

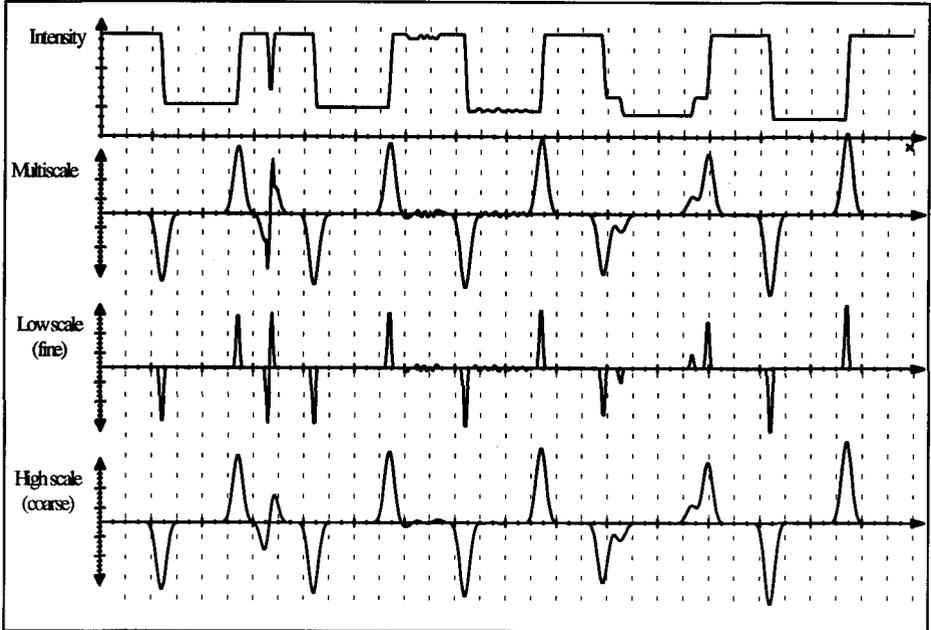


Fig. 7. The detection process applied to the intensities along a horizontal scan line in an image (row=340).

LOCAL DESCRIPTION OF THE INTENSITY CHANGES										
Head Block	nic	dir	pcont	ps1	ps2					
	18	0	20	.33	3.00					
Data Blocks.	xi	xf	col	row	gcp	gmin	con	tune	type	sel
1	1	1	1	340	128	32	192	80	0	1
2	38	41	39	340	186	72	-152	75	0	1
3	86	89	87	340	110	72	152	75	0	1
4	105	107	106	340	184	104	-120	100	0	1
5	108	110	109	340	184	104	120	100	0	1
6	134	137	135	340	184	64	-160	75	0	1
7	182	185	183	340	104	64	160	75	0	1
8	230	233	231	340	182	56	-168	75	0	1
9	278	281	279	340	98	56	168	75	0	1
10	310	337	318	340	190	48	-176	75	12	-1
11	317	320	318	340	190	88	-136	75	20	1
12	328	331	329	340	78	48	-40	75	20	1
13	367	393	384	340	122	48	176	75	12	-1
14	373	376	374	340	58	48	40	75	20	1
15	383	386	384	340	122	88	136	75	20	1
16	422	425	423	340	178	40	-184	75	0	1
17	470	473	471	340	86	40	184	75	0	1
18	512	512	512	340	128	32	-192	80	0	1

Table 2. Attribute values of the intensity changes.

3 Experimental results

We now look more closely at some of the results obtained by the proposed method for detecting and characterising intensity changes in order to obtain a multiple representation of them. Fig. 7a shows an intensity profile obtained from a horizontal scan line (row 340) of a synthetic image of 512x512 pixels where different transitions can be seen. Fig. 7c, 7d show the results after applying the convolution processes with differential operators of different sizes meanwhile Fig. 7b, shows the combination of the low and high scales giving rise to the multiscale differential profile. The different attribute values after applying the characterisation and the integration processes are shown in the Table 2.

The *head block* data tell us that eighteen intensity changes have been detected in the intensity profile by using the following system parameter values: contrast threshold ($pcont=20$), low scale ($ps1=0.33$) and high scale ($ps2=3.00$). The *data blocks* show the local attributes of each intensity change or the results of the description process where we can see that one artificial transition has been added to each extreme of the intensity profile in order to obtain a complete sequence of regions (background included). All transitions are single except two of them which are global multiple transitions where two local multiple transitions can be distinguished.

Fig. 8 shows an original image of 512x512 pixels where we can see two types of entities: objects and background. Shadows and brightness can be distinguished due to an uneven background and poor illumination. The image has been scanned horizontally (Fig 9) and vertically (Fig. 10). Fig. 9a - 9b show two edge images which correspond to the horizontal edge detection at the scales $\sigma=0.33, 3.0$, respectively. In the same way, Fig. 10a - 10b show two edge images which correspond to the vertical edge detection. Low contrast intensity changes have been eliminated by thresholding. The threshold value defined by the system parameter $pcont$ has been fixed at 20. Fig. 9c and Fig. 10c show the result of combining the two edge images according to the proposed method which was described in section 2.2. Only the position attribute of edges (*column and row*) has been visualised.

On the other hand, following the hypothesis that intensity changes originate in the boundaries of the entities of interest, the extraction and characterisation of regions is carried out by analysing the sequence of intensity changes and by associating a set of attributes to each region. Given the one-dimensional character of the proposed method, for each exploration direction of the image, we obtain a new representation of the image where the primitive elements are linear regions in the form of segments. These elements give rise to a new representation of the image in which the primitive elements can be selected and/or combined in order to obtain significant segmented entities. It should be noted that in non-controlled environments where the general illumination conditions are not uniform, the intensity changes produced in the image are due to multiple factors and not all of them correspond to the physical boundaries of the objects. In these situations, at a gray-tonality level, it is very difficult to establish the criteria for making decisions as to whether a region corresponds to an object or not.

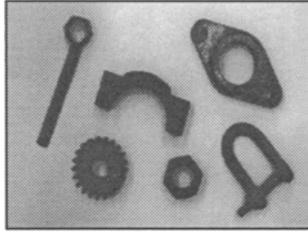


Fig. 8. Original image of 512x 512 pixels.

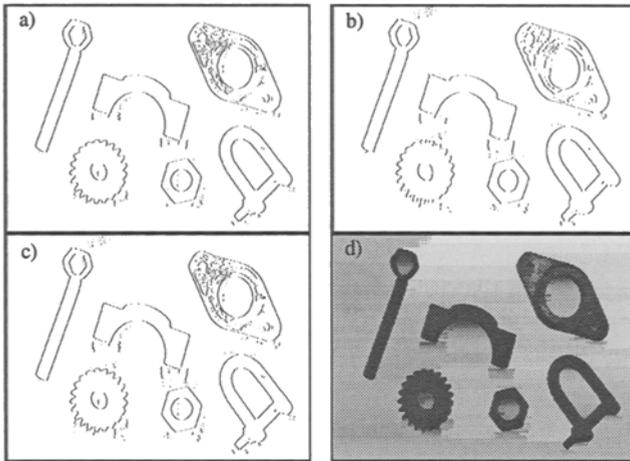


Fig. 9. Local description of the intensity changes along a horizontal scan image. a), b) Edges images obtained at scales $\sigma=0.33$ and 3.0 and minimum contrast=20, respectively; c) Combination of edge information and d) Horizontal regions.

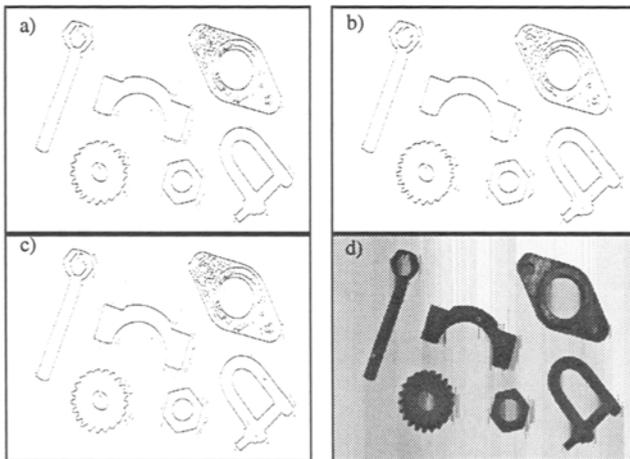


Fig. 10. Local description of the intensity changes along a vertical scan image. a), b) Edges images obtained at scales $\sigma=0.33$ and 3.0 and minimum contrast=20, respectively; c) Combination of edge information and d) Vertical regions.

Fig. 9d and 10d show a representation of gray level attribute of regions which were obtained from the multiple representation of the intensity changes by other processes not discussed in this work. With this in mind the segmented entities can be obtained from the multiple representation of the regions where other higher level processes act to obtain the entities of interest in agreement with certain strategies (decision rules) which can be defined based on previous knowledge or heuristic information of the entities which we want to segment.

4 Conclusions

This new image segmentation method is based on a multiple representation scheme, where multiple representation is obtained by expansion processes which act at the lowest levels of the vision system in order to produce a dimensional increase of the sensorial process. In this way, a wide variety of decision clues are accentuated giving rise to a multiple representation in which reduction processes based on decision rules act in order to reduce the syntactic information and increase the semantic information.

We have described the processes to obtain the multiple representation of the intensity changes which the most outstanding are: intensity changes detection at multiple scales, information combination, characterisation processes at each single scale and, characteristics integration for obtaining a final description of the intensity changes. After applying these processes a new representation of the original image is obtained which is a complete representation in the sense that the original image can be recovered from the attributes associated to each intensity changes (primitive elements) without an important lost of information. On the other hand, it constitutes the starting point to generate a multiple representation of regions (no discussed in this paper) in which a set of decision rules based on previous knowledge or heuristic information can be laid down so that the segmentation results can be obtained on a semantic basis. The method has been tested successfully in real images obtained by a TV camera in applications of artificial vision oriented to the automatic objects manipulation by robots.

References

1. R.M. Haralick and L.G. Shaphiro: Image segmentation techniques. *Comput. Vision Graphics Image Process.*, 29, , pp. 100-132, (1985).
2. N.R. Pal and S.K. Pal: A review on image segmentation techniques. *Pattern Recognition*, 26, 9, pp. 1277-1294, (1993).
3. R Kasturi and R. C. Jain: *Computer vision: Principles*. IEEE Computer Society Press., pp. 65-76, (1991)
4. A.M. Nazif and M.D. Levine: Low level image segmentation: An expert system . *IEEE Trans. Pattern Analysis Mach. Intell.*, 6, 5, pp. 555-577, (1984).
5. C.K. Chow and T. Kaneko: Automatic boundary detection of the left ventricle from cineangiograms. *Computers and Biomedical Research*, 5, , pp. 388-410, (1972).

6. Y. Nakagawa and A. Rosenfeld: Some experiments on variable thresholding. *Pattern Recognition*, 11, , pp. 191-204, (1979).
7. P.K. Sahoo, S. Soltani, A.K.C. Wong and, Y.C. Chen: A survey of thresholding techniques. *Comput. Vision Graphics Image Process.*, 41, , pp. 233-258, (1988).
8. S.D. Yanowitz and A.M. Bruckstein: A new method for image segmentation. *Comput. Vision Graphics Image Process.*, 46, , pp. 82-95, (1989).
9. E.C. Hildreth: Edge detection. A.I. Memo N° 858 MIT, pp. 1-21, (1985).
10. Vishvjit S. Nalwa and Thomas O. Binford: On Detecting Edges. *IEEE Trans. Pattern Analysis Mach. Intell.*, 8,6, pp.699-714, (1986).
11. L.S. Davis: A survey of edge detection techniques. *Computer Graphics and Image Processing*, 4, , pp. 248-270, (1975).
12. J. Canny: A computational approach to edge detection. *IEEE Trans. Pattern Analysis Mach. Intell.*, 8, 6, pp. 679-698, (1986).
13. V. Torre and T.A. Poggio: On edge detection. *IEEE Trans. Pattern Analysis Mach. Intell.*, 8, 2, pp. 147-163, (1986).
14. R. Deriche: Fast algorithms for low-level vision. *IEEE Trans. Pattern Analysis Mach. Intell.*, 12, 1, pp. 78-87, (1990).
15. J.S. Chen and G. Medioni: Detection, localization, and estimation of edges. *IEEE Trans. Pattern Analysis Mach. Intell.*, 11, 2, pp. 191-198, (1989).
16. S. Castan, J. Shao and, J. Shen: New edge detection methods based on exponential filter. *IEEE Proc. 10th Int. Conf. on Pattern Recognition*, 1, , pp. 709-711, (1990).
17. Ziou, D. and Tabbone, S.: A multi-scale edge detector. *Pattern Recognition*, 26,9, pp.1305-1314, (1993).
18. Lu, Y. and R.C. Jain: Behaviour of edges in scale space. *IEEE Trans. Pattern Analysis Mach. Intell.*, 11, 4, pp. 337-356, (1989)