TR-91-012

Keep-Sliding Toboggan Segmentation

Xu Yao and Yi-Ping Hung

Table of Contents

Keep-sliding Toboggan Segmentation

Xu Yao and Yi-Ping Hung

Institute of Information Science
Academia Sinica
Nankang, Taipei

Abstract:

Image segmentation is one of the most important problems in computer vision. Recently, Fairfield proposed an interesting approach to image segmentation using toboggan enhancement followed by naive contrast segmentation, which is a non-iterative, linear execution time method. The way it operates can be thought of as a man tobogganing in the *first derivative terrain*, i.e., the graph surface of a discontinuity measure computed by the first derivative of the image intensity. The segmentation results it produced appeared equal in quality to that of other complex optimal region growing methods. In this paper, an improved version of Fairfield's method, called keep-sliding toboggan segmentation, is presented. With our method, the toboggan will keep sliding on a plain in the derivative terrain, where the original toboggan method will stop sliding. Therefore, our method produces far less regions than the original one does. Other improvements are achieved as follows. In stead of being followed by contrast segmentation post-process, our keep-sliding tobogganing process is preceded by a pre-filtering process which suppresses small fluctuations in the first derivative terrain. Because of this pre-filtering operation, our tobogganing process can automatically merge the regions having small inter-contrast. Also, a new discontinuity measure is proposed to allow the detection of small target regions without over-segmenting the images. Experimental results indicate that the segmentations produced by the keep-sliding toboggan method are less noisy, and therefore, it is more appropriate to use them as initial segmentations for higher level image segmentation techniques.

1. Introduction

One of the most important tasks in computer vision is image segmentation. Its goal is to divide an image into meaningful regions to represent objects in the scene. This problem has been studied for more than twenty years [Hara85][Kana80][Zuck76][Rose82]. Recently, Fairfield proposed an interesting approach to image segmentation using *toboggan enhancement* followed by naive *contrast segmentation*, which is a non-iterative, linear execution time method [Fair90]. His idea is motivated by anisotropic diffusion [Pero87] and adaptive smoothing [Sain89]. He has shown that the segmentation results his method produced appeared equal in quality to that of other complex optimal region growing methods. The way it operates can be thought of as a man tobogganing in the first derivative terrain, i.e., the graph surface of a discontinuity measure computed by the first derivative (e.g., see figure 1(b)).

To describe the **toboggan enhancement** method, let us consider a one-dimensional (1-D) example. Figure 1(a) shows the 1-D signal $x(n), n = 1, \ldots, 30$. Let $x^*(n)$ denote the enhanced value of $x(n)$ after applying the toboggan enhancement algorithm and $x'(n)$ denote a discontinuity measure of $x(n)$. Figure 1(b) shows the graph of $x'(n)$, which is computed by the centered difference approximation of the first derivative, i.e., $x'(n) = |x(n+1) - x(n-1)|/2$.

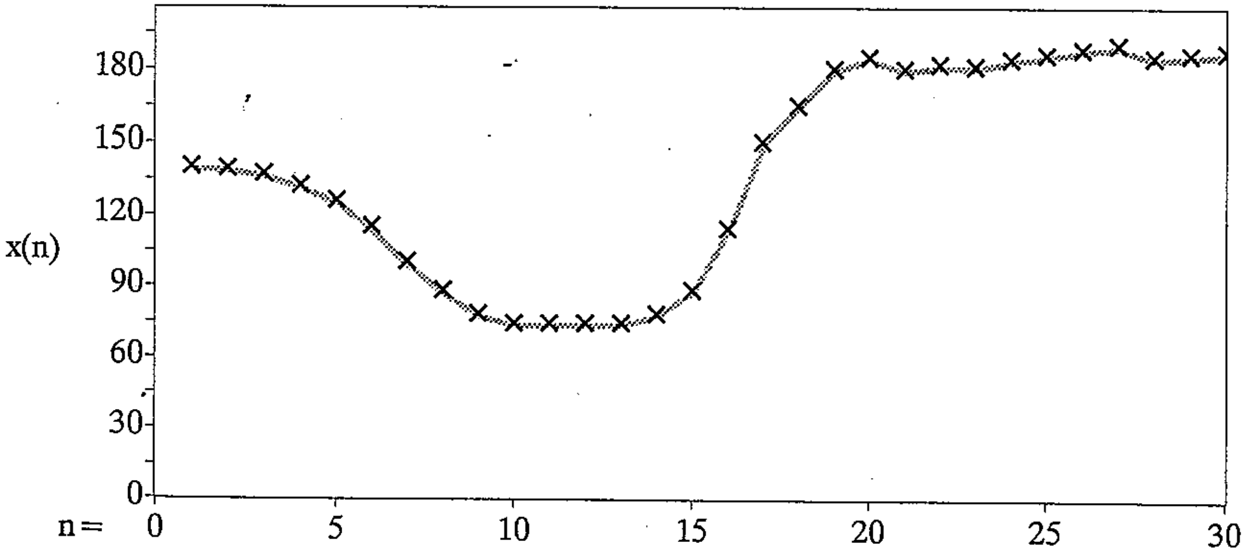

Figure 1(a). A one–dimensional signal to be segmented.

Consider figure 1(b). Imagine a toboggan located at site $n=14$ which has two neighbors $n=13$ and $n=15$. The toboggan will slide to the site $n=13$ because $x'(13)$ is smaller than $x'(14)$. Once the toboggan is at $n=13$, it will slide to site $n=12$ because $x'(12)$ is smaller than $x'(13)$. Since site $n=12$ is a local minimum, i.e., $x'(12) \leq x'(13)$ and $x'(12) \leq x'(14)$, the toboggan will stop sliding. Then, we assign $x^*(14)=x^*(13)=x^*(12)=x(12)$. Figure 1(c) displays the enhanced signal $x^*(n)$ after the toboggan enhancement.

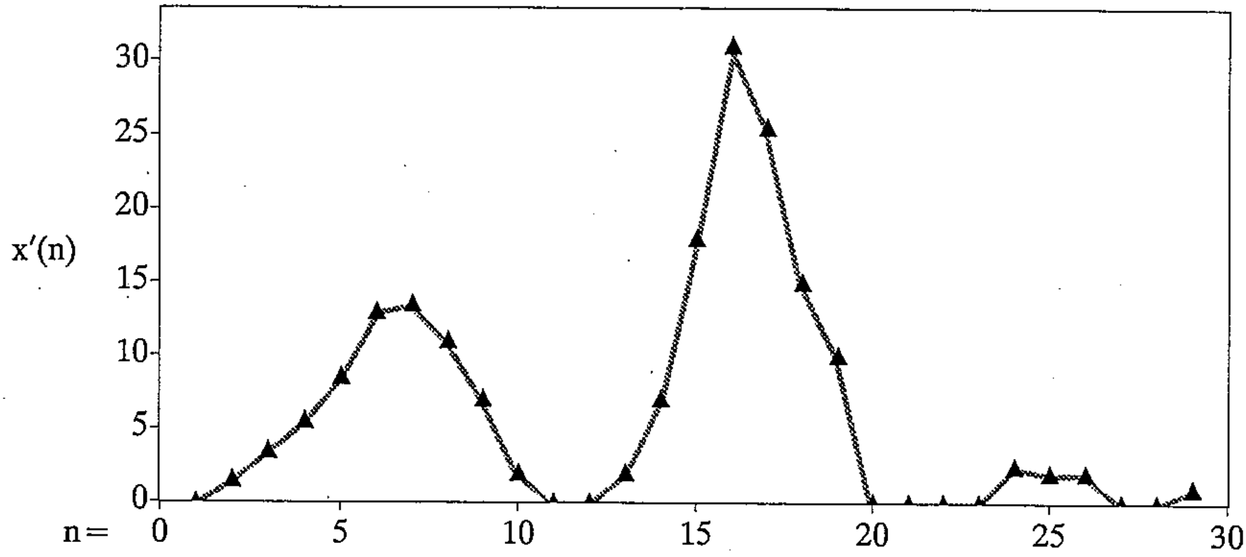

Figure 1(b). The first derivative of the signal in figure 1(a) computed by $x'(n) = |x(n+1) - x(n-1)|/2$.

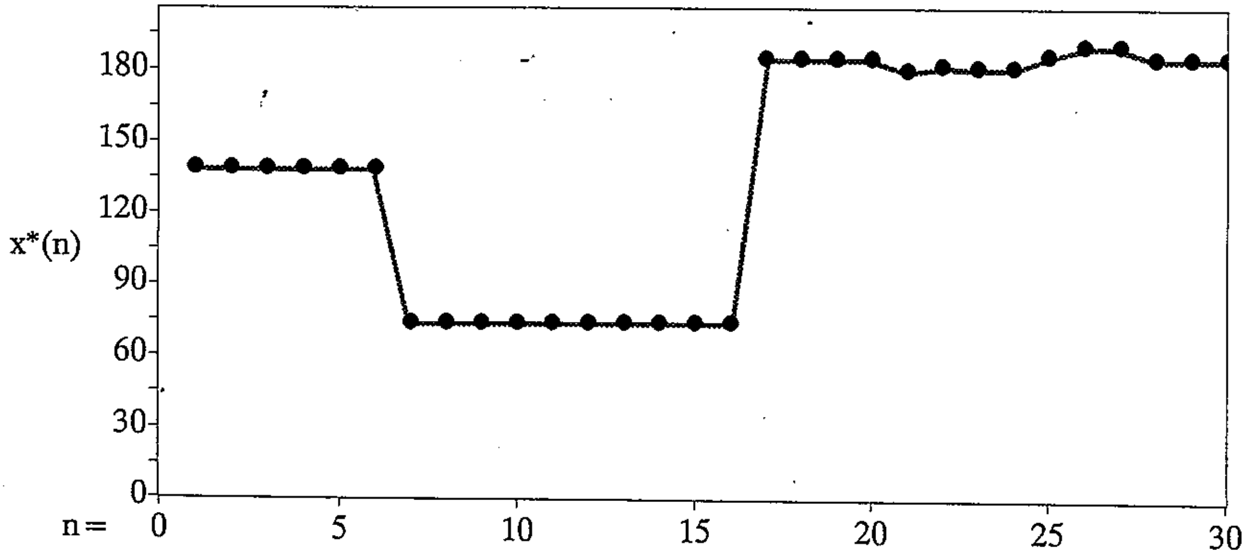

Figure 1(c). The results after the toboggan contrast enhancement.

Now let us consider a two-dimensional (2–D) signal $x(m, n)$. Figure 2 is a discontinuity measure $x'(m, n)$ of $x(m, n)$, for example, the value computed by Prewitt operator. Let $least(m, n)$ denote the smallest discontinuity value among the eight neighbors of site (m, n). Suppose the toboggan is at site $(3, 5)$, where $x'(3, 5) = 11$. It will slide to site $(4, 6)$ because $least(3, 5) = x'(4, 6) = 5 < x'(3, 5) = 11$. Therefore the toboggan will slide to site $(4, 6)$. With one more step, the toboggan will stop sliding at site $(5, 5)$ because none of the neighbors of site

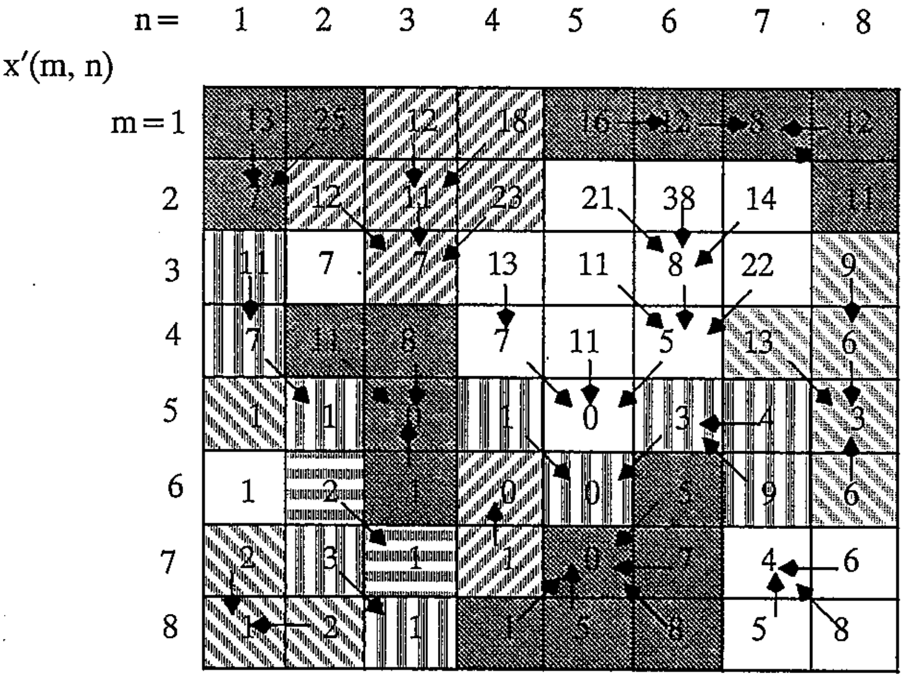

Figure 2. The segmentation result marked on the first derivative terrain.

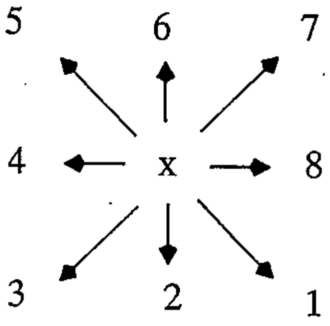

Figure 3. The visiting order for searching for the smallest neighbor.

(5, 5) has smaller discontinuity value than itself, i.e., $least(5, 5) \geq x'(5, 5)$. The sites having been traversed by the toboggan will then form a region $R1 = \{ (3, 5), (4, 6), (5, 5) \}$.

Let us consider another starting point, site *(3, 4)*, where $x'(3, 4) = 13$. Site *(3, 4)* has two neighbors having the discontinuity value of $least(3, 4) = 7$, which are site *(3, 3)* and site *(4, 4)*. In principle, the toboggan can slide to any of these two neighbors. Here, we simply select the one that is visited first according to the visiting order shown in figure 3. Hence the toboggan will slide to site *(4, 4)*. Continue the tobogganing process, another region $R2 = \{ (3, 4), (4, 4), (5, 5) \}$ will be formed. Since *R1* and *R2* has a non–empty intersection *(5, 5)*, *R1* and *R2* will be merged into one region. Final enhancement result is shown in figure 2 where different colors (or patterns) represent for different regions.

Figure 4(a) shows an image of 256 by 256 pixels. Figure 4(b) is the enhanced image of figure 4(a) using the toboggan enhancement algorithm described above. Note that the enhanced image implies by itself an over–segmentation of the original image (see figure 4(c)).

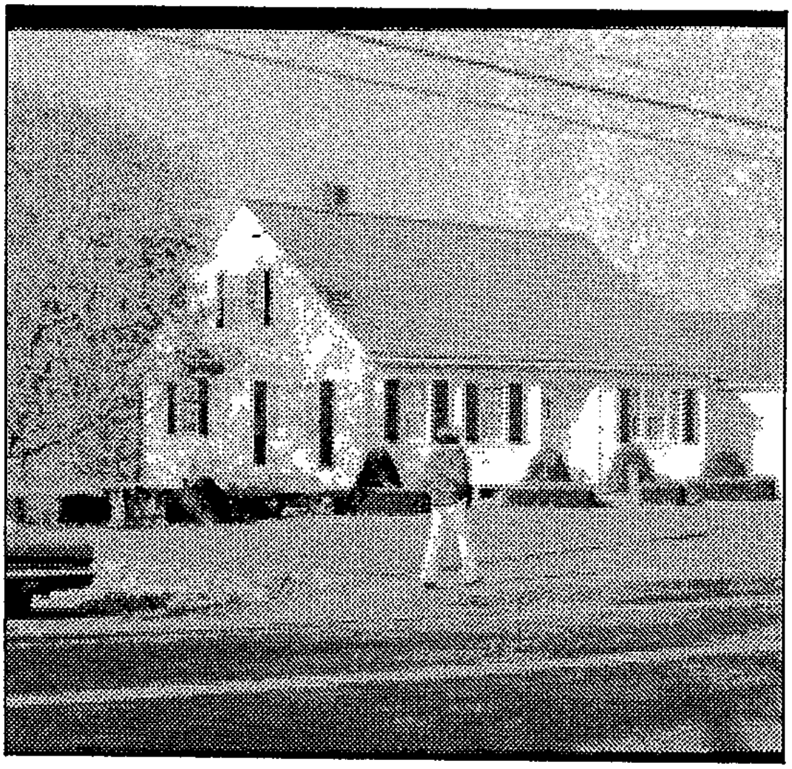

Figure 4(a). The original image of 256 x 256 pixels.

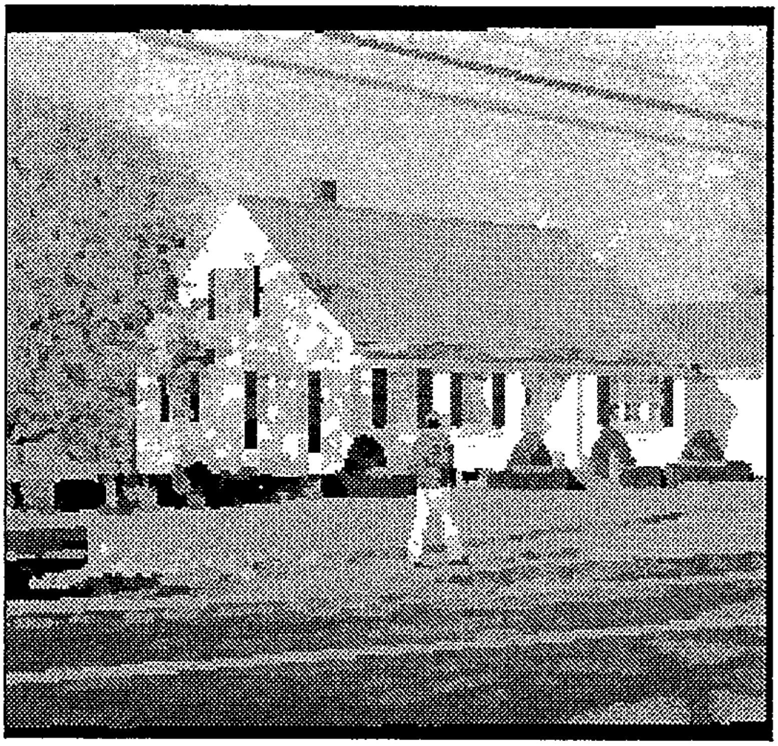

Figure 4(b). The enhanced image of figure 4(a) using the toboggan contrast enhancement algorithm.

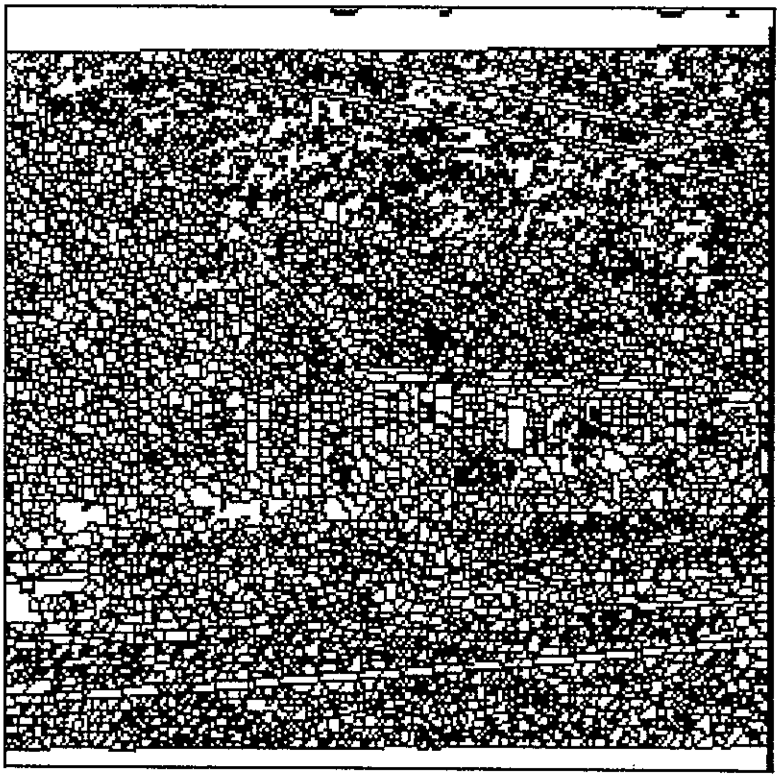

Figure 4(c). The segmentation result obtained in figure 4(b).

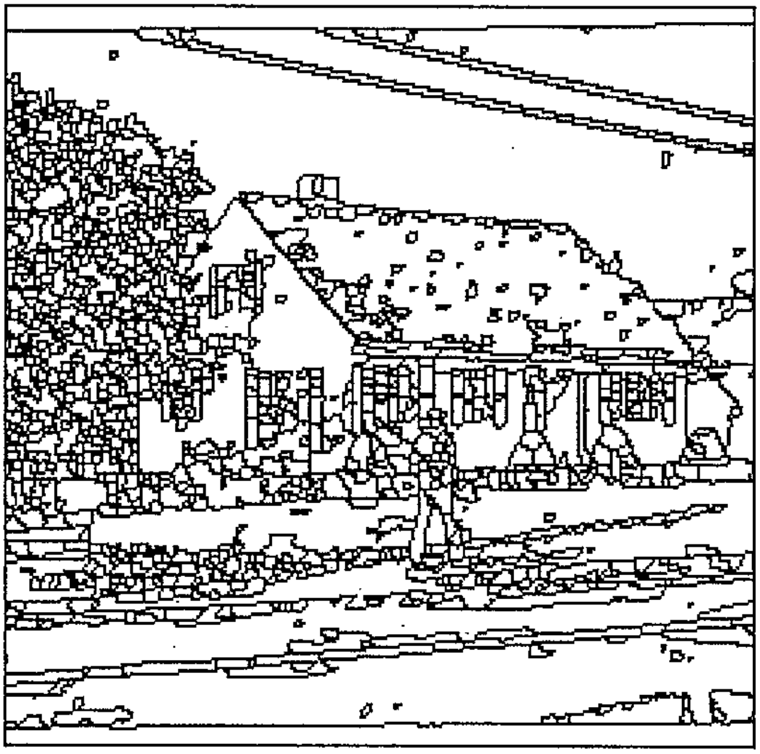

Figure 4(d). The segmentation result obtained by applying the contrast segmentation process, with $s = 8$, to the image in figure 4(b).

Fairfield has shown that much better segmentation can be achieved by applying a simple **contrast segmentation** post–process to the image produced by the toboggan enhancement algorithm. The contrast segmentation process is a simple method for image segmentation, which requires a pre–specified threshold s. In this process, two neighboring sites are assigned to the same region if the absolute difference of their intensity is smaller than the threshold s. For example, figure 4(d) shows the segmentation result obtained by applying the contrast segmentation post–process, with $s = 8$, to the image in figure 4(b). As a whole, the original **toboggan segmentation** method is a two–stage process which combines *toboggan enhancement* and *contrast segmentation*.

In this paper, we propose a modified version of the toboggan segmentation method, called "keep–sliding toboggan segmentation", whose performance is better than that of the original algorithm. Details of our improved algorithm are described in section 2. Experimental results are shown in section 3. Section 4 gives some concluding remarks.

2. Keep–sliding Toboggan Segmentation

The segmentation results obtained from the toboggan methods is intended to be used as an initial segmentation for higher level segmentation methods which incorporate prior knowledge. In this paper, our objective is to improve the toboggan method so as to (i) prevent unnecessary over–segmentation, and (ii) enhance the capability of detecting small target regions. To achieve this goal, we introduce four modifications to the original algorithm, which are:

(1) the keep–sliding stopping rule for tobogganing,

(2) the pre–filtering process which prevents over-segmenting the image into regions having small *inter–contrast* (i.e., contrast between the neighboring regions),

(3) the toboggan splitting technique for solving the multi–path problem, and

(4) a better discontinuity measure for detecting small targets.

Motivation and details for these modifications are described in the following subsections.

2.1 Stopping Rule for Tobogganing

Consider the lower–left part of figure 2. Many contiguous sites have the same discontinuity value of $x'(m, n) = 1$ or $x'(m, n) = 0$. With the original toboggan algorithm, the toboggan will stop sliding at site (m, n) if $x'(m, n) = least(m, n)$. Therefore, many regions are formed due to the flatness of the terrain, as shown in figure 2. This type of over–segmentation can be prevented by changing the stopping rule of the toboggan.

Our new stopping rule is to stop sliding only when $x'(m, n)$ is strictly smaller than $least(m, n)$. In other words, the toboggan will keep sliding even when $x'(m, n) = least(m, n)$. Figure 5 shows the segmentation results after changing the stopping rule to the keep–sliding scheme. It can be seen that the number of regions is less than that in figure 2. For example, a toboggan at site $(1, 4)$ will eventually slide to site $(7, 5)$ by traversing through the discontinuity

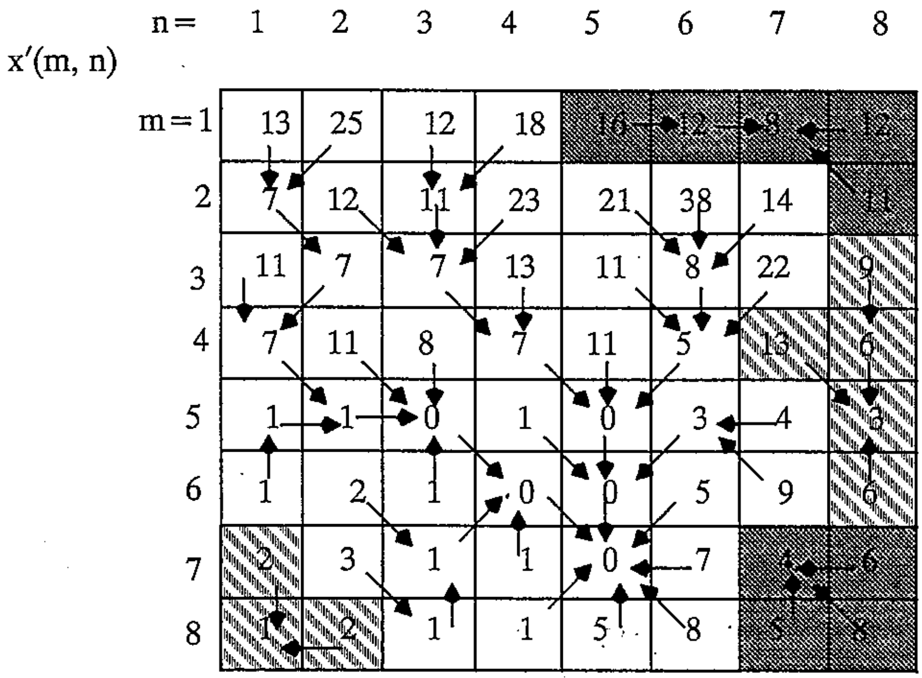

Figure 5. The result obtained by using the keep–sliding stopping rule (see figure 2).

value of *18–11–7–7–0–0–0* when using the new stopping rule, while it will stop at site *(3, 3)* if using the original stopping rule. As another example, consider a toboggan at site *(6, 1)*. It will slide to site *(7, 5)* by traversing through the discontinuity values of *1–1–1–0–0–0* when using the new stopping rule, but will go nowhere if using the original rule. Note that if more than one neighbors have the discontinuity value of *least(m, n)*, we have to decide how to proceed the tobogganing. This issue will be discussed in section 2.3.

2.2 Pre–filtering the Derivative Terrain

In the previous section, we have shown that certain type of undesired over–segmentation can be prevented by using the keep–sliding stopping rule. Here we introduce a pre–filtering process which, when cooperating with the keep–sliding stopping rule, can prevent over–segmenting regions having small inter–contrast. First, consider the 1–D example shown in figure 1. Many regions are formed in the range between $n = 17$ to $n = 29$ (see figure 1(c)) after toboggan contrast enhancement. More precisely, seven regions are formed within that range,

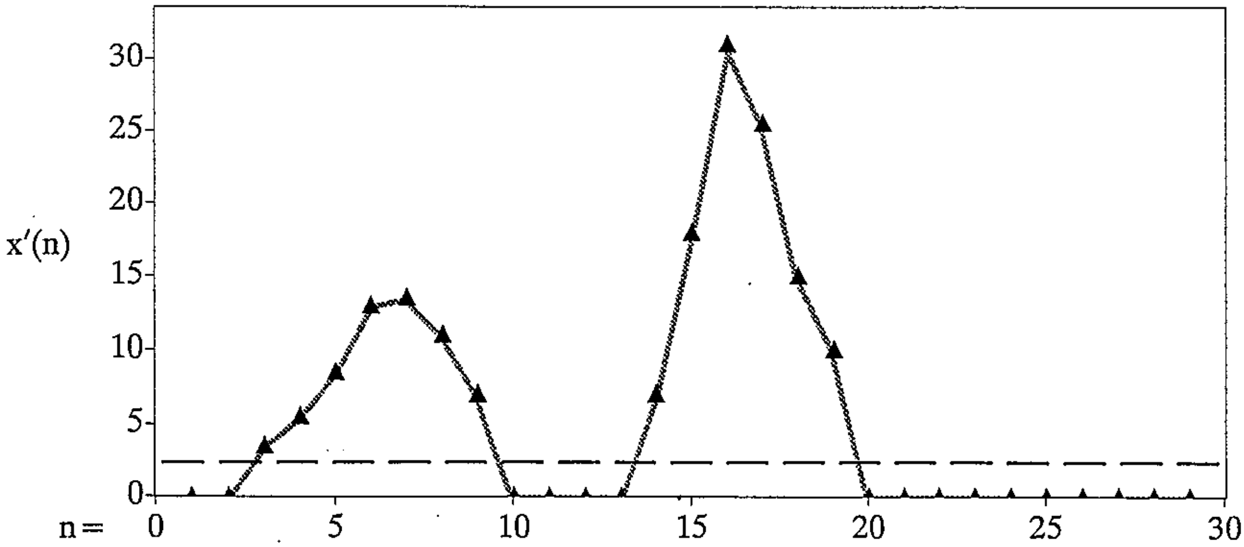

Figure 6(a). The first derivative of the signal in figure 1(a) after pre–filtering with the threshold $T=2.5$.

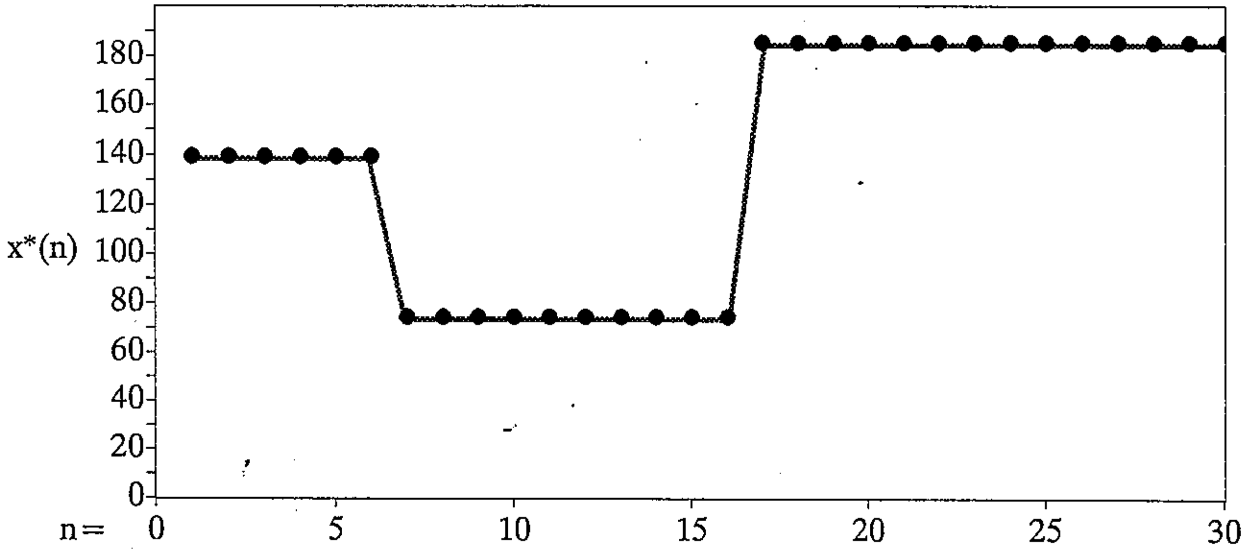

Figure 6(b). The enhanced signal obtained by using the keep–sliding tobogganing process on figure 6(a).

which are {17, 18, 19, 20}, {21}, {22}, {23, 24}, {25}, {26, 27}, {28, 29}. However, if we force $x'(n)$ to be zero when $x'(n) < 2.5$, and adopt the keep–sliding stopping rule, all of these small regions will be merged together as shown in figure 6(b). The idea is to force all the sites having the discontinuity values smaller than a threshold T to have zero discontinuity value. Those sites having high discontinuity values will not be affected by this operation. This operation will be referred to as a *pre–filtering* process on the derivative terrain. Note that this pre–

filtering process should be used together with the keep–sliding stopping rule. Otherwise, the toboggan will stop sliding on the contiguous zero value terrain and generate too many small regions.

The purpose of the pre–filtering process is to enforce *merging of small inter–contrast regions* taken place right in the tobogganing process, instead of doing it in the contrast segmentation post–process. According to our experience, the pre–filtering process and the contrast segmentation post–process generate similar, though not exactly the same, segmentation results. We choose to use the pre–filtering process in this paper.

2.3 Toboggan Splitting Technique

As mentioned in section 2.1, if more than one neighbors have the discontinuity value of *least(m, n)*, we have to decide how to proceed the tobogganing process. This situation will be referred to as a multi–path problem. Consider the derivative terrain in figure 2 again. The toboggan at site *(6, 7)* has two least discontinuity neighbors (or two paths) to slide, which are *(5, 6)* and *(5, 8)*. This multi–path problem occurs frequently in ordinary 2–D images. If the toboggan selects only one path to slide (either randomly or according to a pre–specified order as shown in figure 3), site *(5, 6)* and site *(5, 8)* will be classified into two different regions, as shown in figure 2. However, if we split the toboggan into two, at site *(6, 7)*, and let the two toboggans slide independently, two different regions will be produced by the two toboggans (one attracted by site *(6, 5)* and the other by site *(5, 8)*). We may want to merge these two regions together since they are generated by the toboggan split at site *(6, 7)*.

Notice that the toboggan should not split every time it meets the multi–path problem. For example, in figure 2, the toboggan at site *(2, 6)* also has two least discontinuity neighbors. Because $x'(2, 6) = 38$ is quite a large discontinuity value, site *(2, 6)* should be treated as a region boundary point. However, if we use the toboggan splitting technique there, two obviously different regions will be merged. Consider another derivative terrain, as shown in figure 7, where an obvious edge is located along the diagonal sites. The toboggan at site *(3, 2)*

x'(m, n)	n = 1	2	3	4
m = 1	0	0	8	18
2	0	10	20	12
3	7	20	10	0
4	24	6	0	0

Figure 7. A derivative terrain containing an edge along the diagonal sites.

will meet the multi–path problem. If we use the splitting technique there, the two obviously different regions will be merged. This is an over–merge. To avoid it, if site *(m, n)* has multiple least discontinuity neighbors, the toboggan at site *(m, n)* will be split only when the discontinuity value at site *(m, n)* is smaller than a threshold μ. Usually, we simply set $\mu = 1$.

2.4 Discontinuity Measure

Discontinuity measure plays a very important role in the toboggan segmentation method. By choosing it more carefully, better segmentation can be produced. In our experiments, two discontinuity measures are tested to see their effects on the capability of segmenting small objects. One is computed by Prewitt operator and the other by an augmented Prewitt operator described below.

Due to the smoothing effect of the Prewitt operator, the toboggan segmentation method (both Fairfield's and ours) will miss some small, though high–contrast, target regions in the images. That is because useful informations can be lost in the calculation of discontinuity values when using Prewitt operator. Hence, small target regions will not be able to be segmented out. Consider figure 8(a). Sites *(3, 3)*,*(3, 4)*, *(4, 3)*, *(4, 4)* form a high–contrast region, which is the small target region to be detected. If we apply the toboggan method (either the original one or the keep–sliding one) on the discontinuity terrain computed by Prewitt operator, this small high–contrast target region will not be segmented out after toboggan

x(m, n)	n= 1	2	3	4	5	6	7	8
m=1	50	48	50	53	53	49	44	46
2	51	53	56	56	51	43	40	42
3	50	64	114	116	69	51	43	41
4	47	62	97	91	60	48	42	37
5	43	47	46	45	39	38	37	35
6	39	46	45	40	38	38	39	43
7	36	42	46	46	44	42	41	42
8	33	39	45	47	44	41	41	41

Figure 8(a). The original intensity image containing a high-contrast target region indicated by the shaded area.

enhancement since the discontinuity values around this target region do not form a valley in the discontinuity terrain (see figure 8(b)).

To detect small target regions, we compute the discontinuity value using the *augmented Prewitt operator* defined as follows:

> Set $x'(m, n) = 0$ if the intensity at a site (m, n) is larger or smaller than all of its eight neighbors; otherwise, simply compute $x'(m, n)$ by Prewitt operator.

With this discontinuity measure, small targets we are interested in can be segmented out easily using the toboggan process. For example, figure 8(c) shows that the small high-contrast region marked in figure 8(a) can be segmented out by toboggan method if the discontinuity measure is computed by the augmented Prewitt operator. Finally, we want to mention that, while the discontinuity measure defined by Prewitt operator is not good for detecting small targets, the discontinuity measure defined by *the distance to closest extreme* (DCE) [Fair90] is

x′(m, n)	n=1	2	3	4	5	6	7	8
m=1	34	34	38	45	48	45	42	42
2	25	12	16	15	13	5	1	4
3	29	13	14	14	14	5	2	3
4	27	17	19	21	19	7	4	2
5	35	11	11	14	11	4	1	2
6	25	2	0	1	1	1	1	3
7	31	3	0	1	2	1	0	2
8	28	3	2	0	1	0	0	4

Figure 8(b). The segmentation result obtained by using the first derivative terrain computed by Prewitt operator.

x′(m, n)	n=1	2	3	4	5	6	7	8
m=1	34	34	38	45	48	45	42	42
2	25	12	16	15	13	5	0	4
3	29	13	14	0*	14	5	2	3
4	27	17	19	21	19	7	4	2
5	35	11	11	14	11	4	1	0
6	25	2	0	1	1	1	1	3
7	31	3	0	1	2	1	0	2
8	28	3	2	0	1	0	0	4

Figure 8(c). The segmentation result obtained by using the first derivative terrain computed by the augmented Prewitt operator.

bad even for *large* region having intensity of the exponential form $e^{-|x|}$. The result produced by using the augmented Prewitt operator is more appropriate for our objective.

3. Experimental Results

This section will show some image segmentation results using the new *keep-sliding toboggan segmentation* method described above. This modified toboggan segmentation method is controlled by two parameters, T and μ, mentioned in sections 2.2 and 2.3 . The first parameter T is the pre-filtering threshold which forces small fluctuations in the first derivative terrain to be zero so that regions having small inter-contrast will not be over-segmented when using the keep-sliding stopping rule. The second parameter μ is the splitting threshold which controls the toboggan splitting. Usually, we simply set $\mu = 1$ to avoid segmenting contiguous zero-valued terrain into multiple regions.

Figures 9(a) – 9(d) show the segmentation results using the keep-sliding toboggan method with the pre-filtering process and the augmented Prewitt operator. In the pre-filter-

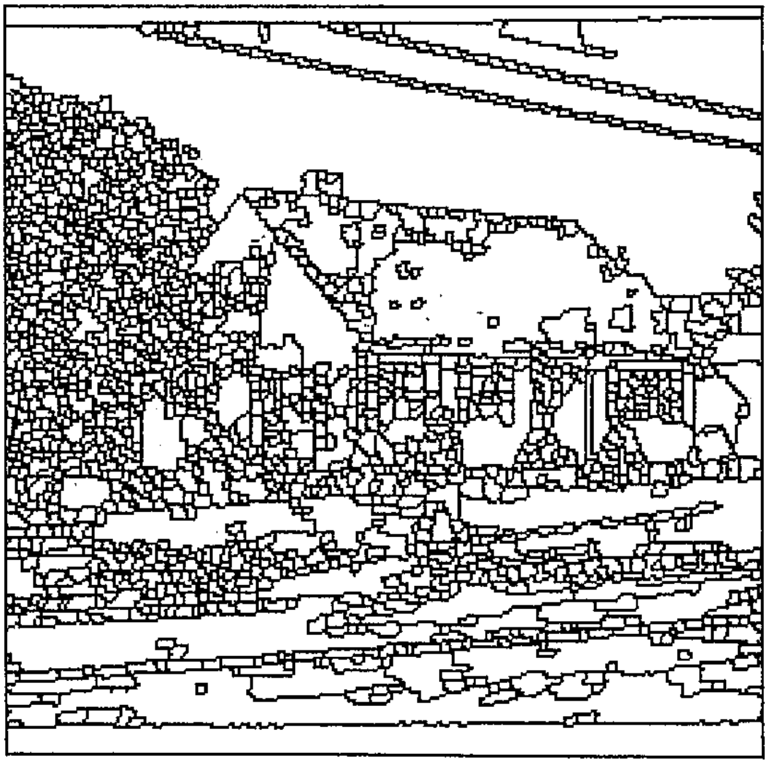

Figure 9(a). The segmentation obtained by using the keep-sliding toboggan segmentation method. ($T=1$, $\mu=1$)

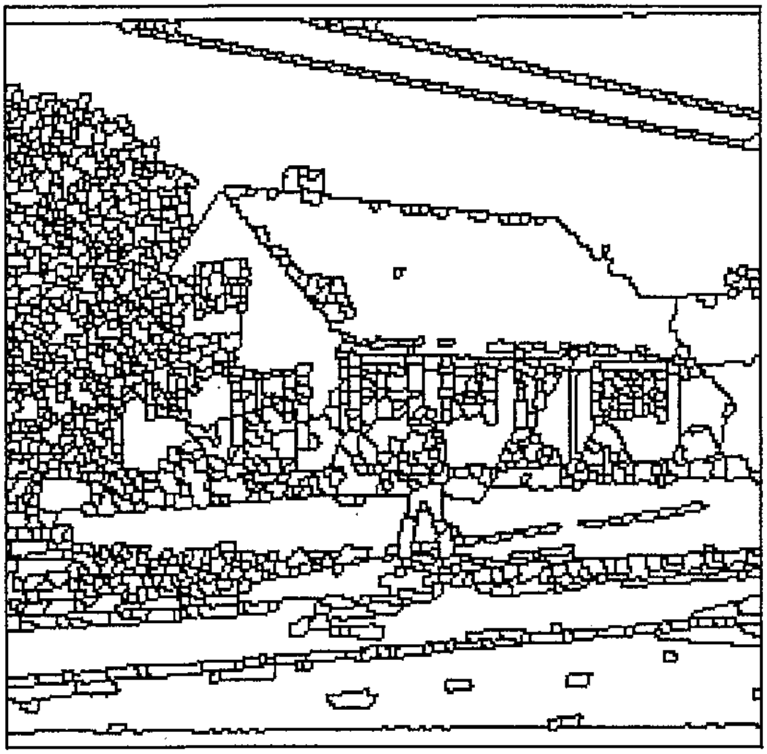

Figure 9(b). The segmentation obtained by using the keep–sliding toboggan segmentation method. ($T=2$, $\mu=1$)

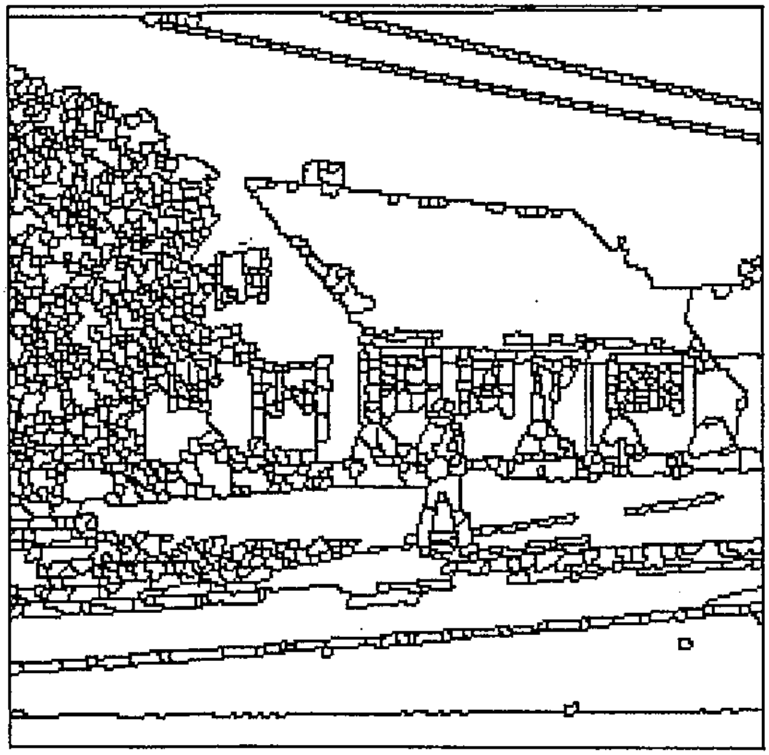

Figure 9(c). The segmentation obtained by using the keep–sliding toboggan segmentation method. ($T=3$, $\mu=1$)

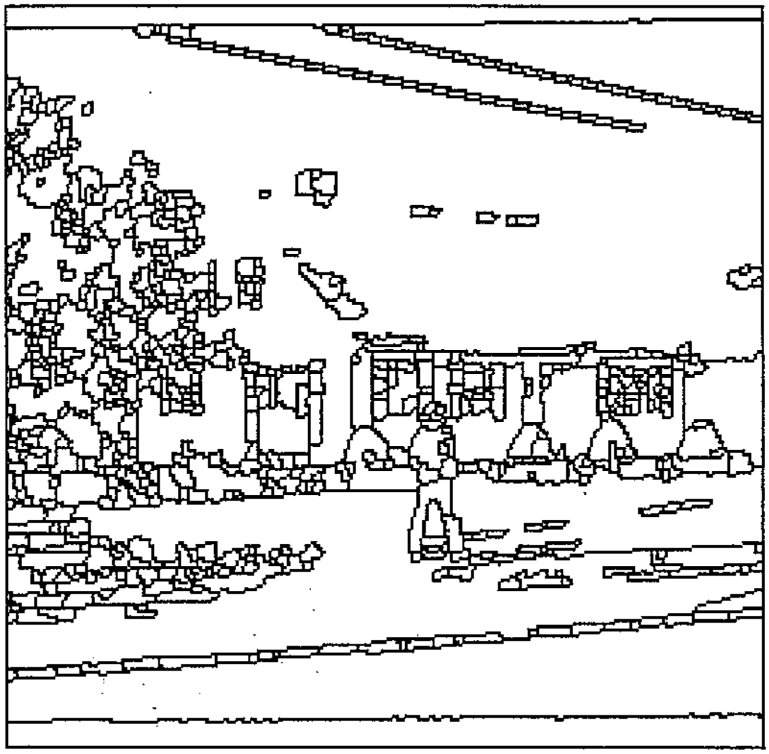

Figure 9(d). The segmentation obtained by using the keep–sliding toboggan segmentation method. ($T=5$, $\mu=1$)

ing process, the threshold T used is 1, 2, 3, 5, respectively, while the splitting threshold μ is set to 1 in all the four experiments. It can be easily seen, from figures 9(a)–9(d), that small T gives finer segmentation, while larger T may merge regions that should not be merged. Since the segmentation results obtained from the toboggan segmentation method is only intended to be used as an initial segmentation for higher level region merging, we do not want the regions to be over–merged. That is, we prefer over–segmentation to over–merging. We found that it would be a good practice to let $T = 1$ or 2, and let $\mu = 1$. Without optimizing the codes, it takes about 10 seconds to process a 256 by 256 image on a SPARCstation.

4. Concluding Remarks

In this paper we present an improved version of the toboggan segmentation method proposed by Fairfield. With our algorithm, unnecessary over–segmentation is diminished and the capability of small target detection is enhanced. It should be kept in mind that we prefer over–segmentation to over–merging since our objective is to produce an initial segmentation for region merging process. To prevent unnecessary over–segmentation, we adopt in our algorithm the keep–sliding stopping rule, the pre–filtering process, and the toboggan splitting technique. To enhance the capability of detecting small target region, we compute the first derivative terrain using the augmented Prewitt operator (or Sobel operator).

In section 3, we have shown some experimental results using different pre–filtering threshold values. Notice that the pre–filtering process introduced in this paper and the contrast segmentation post–process used by Fairfield have similar effect on merging regions having small inter–contrast. The difference is that the pre–filtering process achieves this goal by suppressing small fluctuations in the first derivative terrain, while the contrast segmentation process achieves similar goal by merging regions having small inter–contrast using the intensity values. According to our experience, the pre–filtering process usually achieves more applaudable segmentation. However, we do not claim that the pre–filtering process is indispensable since our purpose of using the toboggan segmentation method is only to produce an initial segmentation for further region merging, and the contrast segmentation post–process is itself a region merging process.

Another issue concerns with the positioning quality of the region boundary point. As an example, consider a pixel locating at a thick intensity edge, i.e., a site near a thick ridge in the first derivative terrain. To which region this pixel will be classified is strongly dependent on the visiting order for searching for the smallest neighbor (see figure 3). To get higher quality segmentation, we may want to fine–tune the region boundaries using the original intensity data after the toboggan segmentation and before the higher level region merging.

Figure 10 shows a first trial on applying the region merging technique to the segmented image obtained from the toboggan segmentation method. The segmentation result is not perfect because we have not completed the development of the region merging process yet and are still working on it. In any case, because the initial segmentation available for

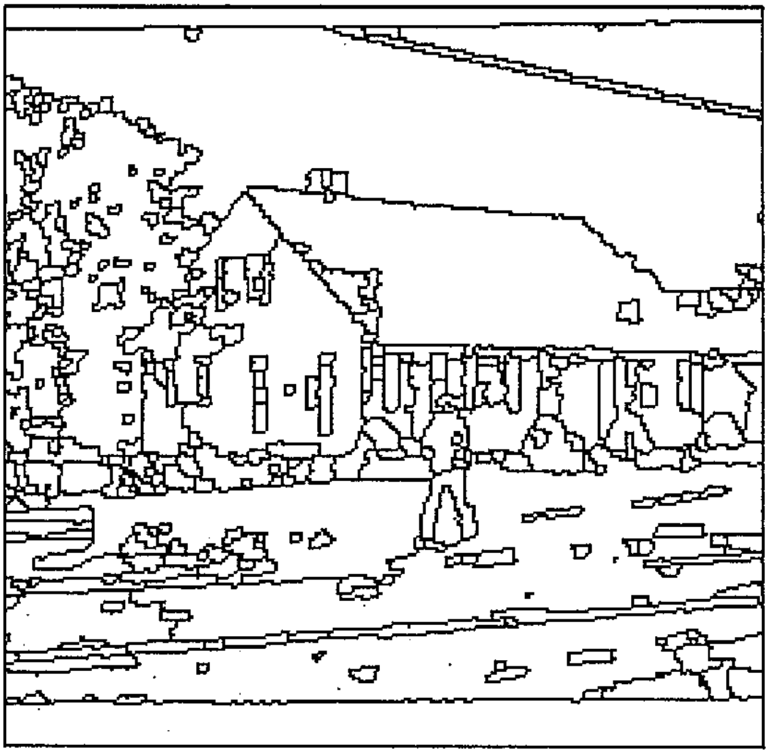

Figure 10. The segmentation result obtained by applying a region merging process to the initial segmentation in figure 9(a).

region merging is quite reliable, the task of developing the region merging algorithm has become much easier.

REFERENCES

[Fair90] J. Fairfield, *"Toboggan Contrast Enhancement for Contrast Segmentation"*. In Proceedings of IEEE Computer Vision and Pattern Recognition, pp.712–716, 1990.

[Hara85] R.M. Haralick and L.G. Shapiro, *"Image Segmentation Techniques"*, Computer Vision, Graphics, and Image Processing, vol. 29, pp.100–132, 1985.

[Kana80] T. Kanade, *"Region Segmentation: Signal vs. Semantics"*, Computer Vision, Graphics, and Image Processing, vol. 13, pp.279–297, 1980.

[Pero87] P. Perona and J. Malik, *"Scale Space and Edge Detection using Anisotropic Diffusion"*. In Proceedings of IEEE workshop on Computer Vision, pp.16–22, Miami, 1987.

[Rose82] A. Rosenfeld and A.C. Kak, *Digital Picture Processing*, Academic Press, New York, 1982.

[Sain89] P. Saint–Marc, J.S. Chen, and G. Medioni, *"Adaptive Smoothing: A General Tool for Early Vision"*, In Proceedings of IEEE Computer Vision and Pattern Recognition, pp. 618–624, 1989.

[Zuck76] S.W. Zucker, *"Region growing: childhood and adolescence,"* Computer Graphics and Image Processing 5:382–399, 1976.