

On Skeletons Attached to Grey Scale Images

M. Karimi Behbahani,¹ Arash Rafiey,² Mehrdad Shahshahani³

Institute for Studies in Theoretical Physics and Mathematics
Tehran, Iran

ABSTRACT

In [2], [3] and [4] a procedure for the reduction of a binary image to a skeleton for the purpose of the automated recognition and classification of images is described. In this work we extend the construction of skeletons to grey level or color images. A hierarchy of decorated graphs are constructed which capture the visual content of the original image to different levels of detail. The correspondence of the decorated graphs to the image is demonstrated by an inverse algorithm which allows the approximate reconstruction of the original image.

KEYWORDS: Image Processing, Computer Vision, Skeletons, Object recognition.

1 Introduction

Pixel by pixel comparison of images does not lead to an acceptable recognition algorithm in computer vision. A small perturbation, a rotation or translation of an image often results in one which may appear very different on a pixel by pixel comparison basis, but perceptually is identical with the original. To alleviate this problem, a method was developed in [2], [3] and [4] to assign certain skeletons or shock graphs to binary images. These objects are invariant under small perturbations, rotations or translations, and yet may contain the crucial information necessary for the identification of an object.

In this paper we modify the algorithm in above mentioned papers, in order to apply it to grey level and color images. The key idea is to replace the distance function with an en-

ergy function which not only takes into account the distance to the “edge”, but incorporates the gradient vector field of the intensity function(s). Broadly speaking, the skeleton is defined as the set of points where the Laplacian of the energy is greater than a specified threshold. Empirical evidence shows that the information contained in the local maxima is insufficient for an acceptable recognition algorithm.

To evaluate the information contained in the skeletons we introduce the refined notion of a decorated skeleton. From the decorated skeletons we can approximately reconstruct the original image. The goodness of a reconstructed image depends on the amount information stored in the decorated skeleton. We propose a hierarchy of decorated skeletons which encode the original image to different levels of detail and fidelity. Our algo-

¹Names of authors appear in alphabetical order. behbahani@ipm.ir,

²ar2rafiey@cc.sbu.ac.ir

³Also visiting Department of Statistics, Stanford University. shahshah@stat.stanford.edu, or sl2z@aol.com

rithms have linear or log times linear complexity which allows for fast implementation.

2 Grey Scale Image Skeletons

The construction of skeletal graphs in [2], [3] and [4] is modified here to accommodate grey scale and color images. The modification is based on the introduction of an energy function for a grey scale image. Let $I(x, y)$ denote the intensity of the image at pixel (x, y) and $d(., .)$ denote the standard Euclidean distance function. Then the energy of a pixel (x, y) is defined as

$$E(x, y) = \min_{(x', y')} \left[d((x, y), (x', y')) + \frac{\alpha}{||\text{grad}I(x', y')||} \right],$$

where $\alpha > 0$ is a positive parameter. The resulting skeleton is robust relative to small perturbations of α . Generally speaking, larger values of α tend to increase the complexity of the skeletal graphs.

The calculation of the energy function by directly using its definition and seeking the minimum through exhaustive search is computationally intensive. A major simplification is achieved by a modified version of the method proposed by Borgefors [1] for the calculation of the distance of a point in an image to the edge of the given shape. The complexity of this method is linear in the number N of pixels and gives a satisfactory approximation to the actual minimum. To describe the necessary modification we recall that Borgefors assigns the values 0 or ∞ to each pixel according as the point is an edge

or not. Consider the masks

$+\sqrt{2}$	$+1$	$+\sqrt{2}$	$+\sqrt{2}$	$+1$	$+\sqrt{2}$
$+1$	0			0	$+1$

Following [1], to calculate the energy at pixel designated 0 we make use of the forward mask and evaluate

$$E(i, j) = \min\{E(i, j), E(i, j-1) + 1, E(i-1, j-1) + \sqrt{2}, E(i-1, j) + 1, E(i-1, j+1) + \sqrt{2}\}$$

The forward mask is passed over the entire image from left to right and top to bottom once. Then the backward mask is applied from right to left and bottom to top by making use of the formula

$$E(i, j) = \min\{E(i, j), E(i, j+1) + 1, E(i-1, j+1) + \sqrt{2}, E(i-1, j) + 1, E(i-1, j-1) + \sqrt{2}\}$$

This determines the required minimum distance (approximately). In our modification we assign the value

$$\frac{\alpha}{||\text{grad}I(x, y)||}$$

to the pixel (x, y) . Note that for binary images this yields the Borgefors 0, ∞ matrix. Passing the same masks over the image yields a satisfactory approximation for the energy function.

The negative of the divergence of the gradient of the energy function is calculated using the expression

$$-\text{div}\left(\frac{\partial E}{\partial x}, \frac{\partial E}{\partial y}\right)(x, y) \\ = \lim_{C \rightarrow (x, y)} \left[\frac{1}{a_C} \int_C < -\text{grad} E, \mathbf{N} > dc \right],$$

where C is a small closed curve containing (x, y) in its interior, dc is the line element on C , and a_C is the area enclosed by C . A threshold T is introduced and a “thick skeleton” $\tilde{S}$ defined by the inequalities

$$(x, y) \in \tilde{S} \quad \text{if and only if} \\ -\text{div}\left(\frac{\partial E}{\partial x}, \frac{\partial E}{\partial y}\right)(x, y) \leq T.$$

The thickness of the skeleton $\tilde{S}$ and its complexity depend on the threshold T and also affect the reconstruction which is discussed in the next section. A thinning algorithm based on the divergence of the energy function, similar to the procedure described in [2], is used to convert $\tilde{S}$ to a skeleton S . Figures 1 through 5 demonstrate the dependence of the skeleton on the choice of α and T (see Table of Figures for explanation).

It is evident that the skeletons are quite complex and it is desirable to simplify their structure and obtain a hierarchy of skeletons reflecting different levels of detail. This can be achieved by different methods. One approach is by reducing the size of the image by a factor of 4 or 16 and then applying the algorithm to obtain the skeleton. Then a simplification algorithm eliminates “irrelevant” branches and loops. This algorithm is ad hoc and calculates lengths of branches and loops. A threshold determines which branches and

loops should be eliminated. A second approach reduces the images by the same factors and quantizes the levels to four or eight and then applies the algorithm. By a similar procedure the small loops and branches are eliminated. Figures 8 and 9 show the skeletons attached to Figure 6 after the reduction of the image and the elimination of irrelevant branches and loops of the skeletons, while Figure 7 is the (unsimplified) skeleton.

An important feature of the assignment of a skeleton to an image is that small perturbations or rigid motions of the latter do not affect the former. Figures 11 show the application of an affine transformation close to the identity to Figure 10. The resulting skeleton, which are *almost* identical, are shown in Figures 13 and 14. As the affine transformation moves away from the identity the skeleton will differ more from the original one. The new skeleton may differ from the original in the creation or annihilation of loops or connected components. Strictly speaking the change in the skeleton is not a continuous function since at some transition points loops or connected components are created or annihilated which is not a continuous process. The meaning and significance of transition points are not clear.

3 Reconstruction

What kind of information is stored in the skeleton of an image? In order to gain some understanding of this question we developed an inverse algorithm for the reconstruction an image from the skeleton to evaluate which features are preserved and which are not. To appreciate the difficulties involved in an inverse algorithm note that the skeleton assigned to

a binary image consisting of a single disc (a circle with its interior) is point regardless of its radius. Therefore without additional information about the size of the disc it is not possible to reconstruct it. This suggests that by specifying an additional parameter (e.g., a distance function on the skeleton) we may be able to reconstruct an approximation to the original image. Rather than describing the algorithm for the binary image we introduce a more general algorithm for grey level or color images with an obvious simplification in the binary case.

The inverse algorithm for grey scale images requires the introduction of the notion of decorated skeleton. By the *decorated skeleton* we mean a triple (S, E, I) where E and I are the energy and intensity of the point with coordinates (x, y) in the original image. Points not on the skeleton are assigned zero energy and intensity. The general idea is to solve a heat equation

$$\frac{\partial I}{\partial t} = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2},$$

with initial condition determined by the decorated skeleton to obtain the grey level image when t becomes large. How large t should be depends on the the energy function which gives information about the distance to the edge. The non-uniformity in t creates complications in the actual calculations. To circumvent this problem we make use of a simpler procedure based on successive applications of two “reverse” masks to obtain an approximation to the original image. The reverse masks are

$-\sqrt{2}$	-1	$-\sqrt{2}$	$-\sqrt{2}$	-1	$-\sqrt{2}$
-1	0			0	-1

In the application of the reverse masks we make use of the formula

$$E(i, j) = \max\{E(i, j), E(i, j-1) - 1, E(i-1, j-1) - \sqrt{2}, E(i-1, j) - 1, E(i-1, j+1) - \sqrt{2}\},$$

for the forward path. Now let (i', j') be the point that the above maximum is attained. Set

$$I(i, j) = I(i', j')$$

pass the mask over the entire decorated skeleton from left to right and top to bottom to complete the forward pass. Similarly in the backward pass we make use of

$$E(i, j) = \max\{E(i, j), E(i, j+1) - 1, E(i-1, j+1) - \sqrt{2}, E(i-1, j) - 1, E(i-1, j-1) - \sqrt{2}\},$$

and let (i', j') be the point where the maximum is attained. Set $I(i, j) = I(i', j')$ as before and pass the mask over the entire image from right to left and top to bottom to obtain the reconstructed image.

Simplifying the skeletons, as described earlier, affects the information content of the skeletons. Figures 12, 13 and 18 show skeletons with different complexities attached to the same image. The reconstructed images are shown in Figures 15, 16 and 19. Note that while the fidelity depends on the complexity of the skeleton, for the general recognition problem it suffices to make use of the simplified skeletons. In other words, the simplified skeletons retain essential information for the application to recognition problems.

Table of Figures

Figure No.	Description	α	T	Reduction Factor	Comment
1	Saturn	-	-	1	Original
2	Saturn	5	-9	1	Skeleton
3	Saturn	2	-9	1	Skeleton
4	Saturn	5	-15	1	Skeleton
5	Saturn	5	-5	1	Skeleton
6	Female Face	-	-	1	Original
7	Female Face	5	-9	1	Skeleton
8	Female Face	5	-9	$\frac{1}{4}$	Simplified Skeleton
9	Female Face	5	-9	$\frac{1}{16}$	Simplified Skeleton
10	Male Face	-	-	1	Original
11	Male Face	-	-	$\frac{1}{4}$	Distorted Image
12	Male Face	5	-9	1	Skeleton of 10
13	Male Face	5	-9	$\frac{1}{4}$	Simplified Skeleton of 10
14	Male Face	5	-9	$\frac{1}{4}$	Simplified Skeleton of 11
15	Male Face	-	-	-	Reconstruction from 12
16	Male Face	-	-	-	Reconstruction from 13
17	Male Face	-	-	1	4-Level Quantization
18	Male Face	5	-9	1	Simplified Skeleton of 17
19	Male Face	-	-	-	Reconstruction of 18
20	Blood Cells	-	-	1	Original
21	Blood Cells	5	-9	1	Simplified Skeleton of 20
22	Blood Cells	-	-	-	Reconstruction of 21

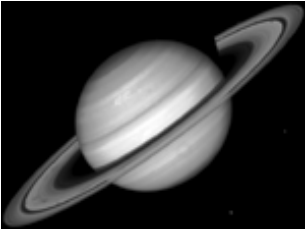


Fig. 1



Fig. 2



Fig. 3



Fig. 4



Fig. 5



Fig. 6

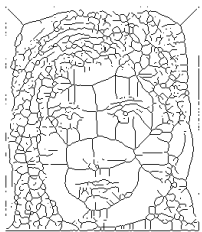


Fig. 7

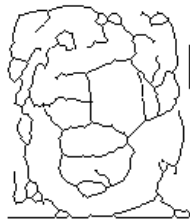


Fig. 8



Fig. 9



Fig. 10



Fig. 11



Fig. 12



Fig. 13



Fig. 14



Fig. 15



Fig. 16



Fig. 17

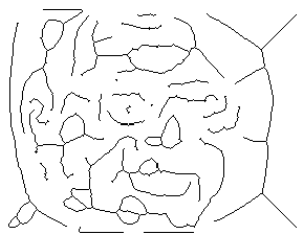


Fig. 18



Fig. 19

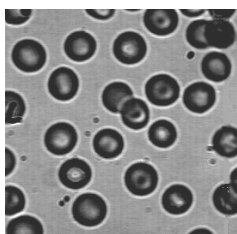


Fig. 20

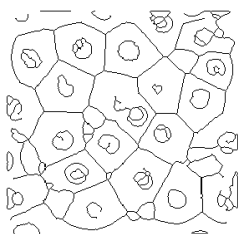


Fig. 21

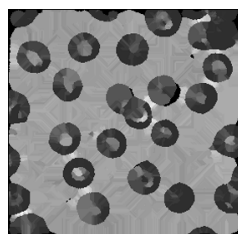


Fig. 22

References

- [1] G. BORGEFORS , Distance transformations in arbitrary dimensions, *Computer Vision, Graphics, and Image Processing*, **27**, 1984, pp. 321-345.
- [2] P. DIMITROV ,C. PHILLIPS, AND K.SIDDIQI - Robust and Efficient Skeletal Graphs, *Conference on Computer Vision and Pattern Recognition (CVPR'2000)*, Hilton Head, South Carolina, USA, **1**, 2000, pp. 417-423.
- [3] A. SHOKOUFANDEH, AND S. DICKINSON - Graph Theoretical Methods in Computer Vision, in *Theoretical aspects of Computer Science* (G. B. Khosrovshahi, A. Shokoufandeh, A. Shokrollahi, eds.), Springer, 2002, pp. 148-174.
- [4] K. SIDDIQI, A. SHOKOUFANDEH, S. DICKINSON, S. W. ZUCKER, Shock Graphs and Shape Matching, *International Journal of Computer Vision*, **30**, (1999), pp. 1-22.