Using Top-Down and Bottom-Up Analysis for a Multi-Scale Skeleton Hierarchy

Gunilla Borgefors[1], Giuliana Ramella[2], Gabriella Sanniti di Baja[2]

[1]Centre for Image Analysis, Swedish University of Agricultural Sciences, Uppsala, Sweden
Fax: +46 18 553447; email: gunilla@cb.uu.se
[2]Istituto di Cibernetica, Italian National Research Council, Arco Felice, Naples, Italy
Fax: + 39 81 5267654; email: gr, gsdb@imagm.na.cnr.it

Abstract Multi-scale skeletons can be conveniently employed in the matching phase of a recognition task. The multi-scale skeletons are here obtained by first computing the skeleton at all levels of a resolution structure and then establishing a hierarchy among skeleton components at different scales, using a parent-child relationship. Although subsets of the skeleton expected to represent given pattern subsets may consist of different number of components at different scales, a component preserving decomposition is obtained that produces a hierarchy in accordance with human intuition.

1 Introduction

Pattern recognition can be based on pattern decomposition and description. In this respect, the skeleton is a convenient tool to facilitate the matching phase of a recognition process, in the case of patterns whose shape can be perceived as the superposition of elongated regions [1,2]. The skeleton is a linear subset of the pattern, centred within the pattern, and is characterised by the same topological and geometrical structure. Skeleton branches are in correspondence with the elongated regions constituting the pattern. Thus, the spatial relationships among pattern subsets can be easily derived, e.g., while tracing the corresponding skeleton branches. This would not happen if the morphological skeleton , e.g., [3], is used, since it does not generally reflect the topological properties of the pattern.

If skeleton components are hierarchically ranked, a better pattern description becomes available and recognition is facilitated [4]. Moreover, the complexity of the matching phase can be reduced by using multi-scale skeletons, e.g., [5,6]; in fact, one can initially match only lower scale skeletons, which represent the most significant pattern subsets, and thus reduce the number of comparisons among higher scale skeletons, necessary to achieve an exact match.

In this paper, we use the multi-scale skeletons obtained by simultaneously extracting the skeleton at all levels of a resolution pyramid; moreover, a hierarchical skeleton decomposition is obtained at all resolution levels by identifying and ranking skeleton subsets, based on their permanence in the skeleton at the various scales.

A first step towards a hierarchical decomposition of the multi-scale skeletons has been taken recently [7]. Analogously to [7], the resolution structure we use here to obtain the multi-scale skeletons is the AND-pyramid. The AND-pyramid is easy to implement but, as the resolution decreases, the pattern is shrunk and narrow regions of the initial pattern may either completely vanish or become disconnected (see Figure 1, where the AND-pyramid of a test pattern is shown). Since skeletonization is a

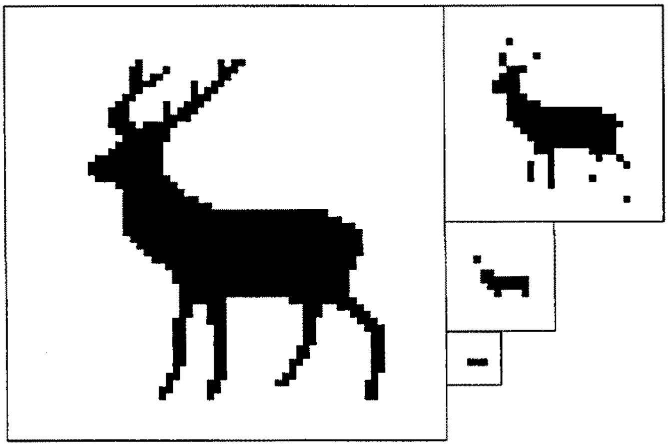

Figure 1. The four significant levels of the AND-pyramid of a test $2^6 \times 2^6$ image.

topology preserving process, the multi-scale skeletons computed on the AND-pyramid will have the same fate. Thus, subsets of the skeleton expected to represent given pattern subsets may consist of different number of components at different scales. We here address and solve the problem of identifying a set of non-connected skeleton fragments as belonging to the same component by inferring connectedness information from the higher scale skeletons onto the lower scale ones during a bottom-up analysis of the pyramid. In this way we can establish a correspondence among skeleton subsets at various scales, even when the subsets do not consist of the same number of connected components. Another novel feature of the current skeleton decomposition is that small noise components are absorbed by adjacent more significant components. The identification of disjunct components and the removal of noise components both contribute to reduce the number of components in the skeleton decomposition to the most significant ones and, accordingly, simplify the skeleton hierarchy.

2 Skeleton Hierarchy

Let P be a $2^n \times 2^n$ binary picture, where black and white pixels respectively constitute the pattern and its complement. We assume that all the pixels on the border of P are white and store P in the highest resolution level (also called the first level, or the bottom level) of an AND-pyramid. The next, lower, resolution level of the pyramid is built from the first level. All pixels with four black *children* in the first level are set to black. Similarly, the third level is built from the second, and so on until all (n+1) resolution levels are obtained. Indeed, resolution levels sized $2^2 \times 2^2$, $2^1 \times 2^1$ and $2^0 \times 2^0$ are not meaningful for skeletonization purposes and we will consider the $2^3 \times 2^3$ pixel image as the last pyramid level (also called the top level). The AND-pyramid is easy to compute, but it is *not* shape preserving and, when the resolution decreases, some of the subsets present in the initial pattern either completely vanish or become disconnected, [8]. Thick regions appear at all resolution levels and constitute the most significant pattern components.

Skeletonization is accomplished simultaneously at all resolution levels. Any skeletonization algorithm can be used. We favour algorithms requiring two distinct phases, respectively tailored to the identification of an at most two-pixel wide set of skeletal pixels, and to the reduction of this set to unit width, see, e.g., [9]. Using this approach, we can postpone the second skeletonization phase, till after the desired hierarchy has been built at all levels of the pyramid. This is preferable to guarantee that the parent-child relationship among skeleton components at different scales can be correctly established. Even though we are aware that the term skeleton should be used only for the set resulting after both skeletonization phases have been accomplished, to avoid lengthy periphrases in the following we will refer also to the original sets of skeletal pixels as the skeletons.

2.1 Top-Down Approach

The main idea that in [7] guided the construction of the skeleton hierarchy by means of a top-down process was the observation that skeleton components present at lower resolution levels were definitely also present at higher levels. The process used in that work can be summarised as follows. The connected components of the skeleton are identified at the lowest resolution level (2^3x2^3 level). Each component is *parent* of a *child* component at the next level (2^4x2^4), *grandparent* of a *grandchild* at level 2^5x2^5, and so on; a suitable process allows one to establish the parent-child relationship and to identify all the descendants. As the structure of the skeleton is generally more and more complex as soon as the resolution increases, not all skeletal pixels at the successive levels are assigned to descendants of components in the top level. Pixels not assigned to child components at level 2^4x2^4 are grouped into connected components that are interpreted as new parent components, directly originating at level 2^4x2^4. Their children and grandchildren can be found in the higher levels. Similarly, new parent components and their descendants can be found at all levels. During the process, skeleton components at each single pyramid level are assigned a *permanence number*, counting the number of levels up to the most remote corresponding ancestor component. At each level, the most significant components are those with the largest permanence. The maximal permanence is equal to the number of levels of the resolution pyramid.

Three sub-processes are done to establish the parent-child relationship. In fact, due to the discrete nature of the AND-pyramid and the skeleton, a child component might have some of its pixels slightly shifted compared to their "expected" positions. The black children of the pixels in a parent component do not exhaust the black pixels expected to constitute the child component. The first sub-process is termed *projection*. Every skeletal pixel p in a (parent) skeleton component projects, at the immediately higher resolution level, over a 2x2 set *(quadruplet)*, generally including both white pixels and skeletal (black) pixels; skeletal pixels in the quadruplet Q associated to p are assigned to the corresponding child component. The second sub-process is termed *expansion* and is active whenever Q includes only white pixels. Expansion interprets as belonging to the current child component the skeletal pixels possibly found in the four quadruplets placed North, East, South and West of Q, provided that these pixels have not already been marked as belonging to any other child component. Either projection or expansion is accomplished. Both processes involve pixels placed in a pair of successive levels. The third sub-process, termed *propagation*, is accomplished after projection (or expansion) has been performed from all pixels at a given level onto the successive higher resolution level. Propagation involves only pixels on the latter level. For every pixel p ascribed to a child component by projection or expansion, propagation assigns also the black neighbours of p to the same component, provided that they have not already been assigned to other components.

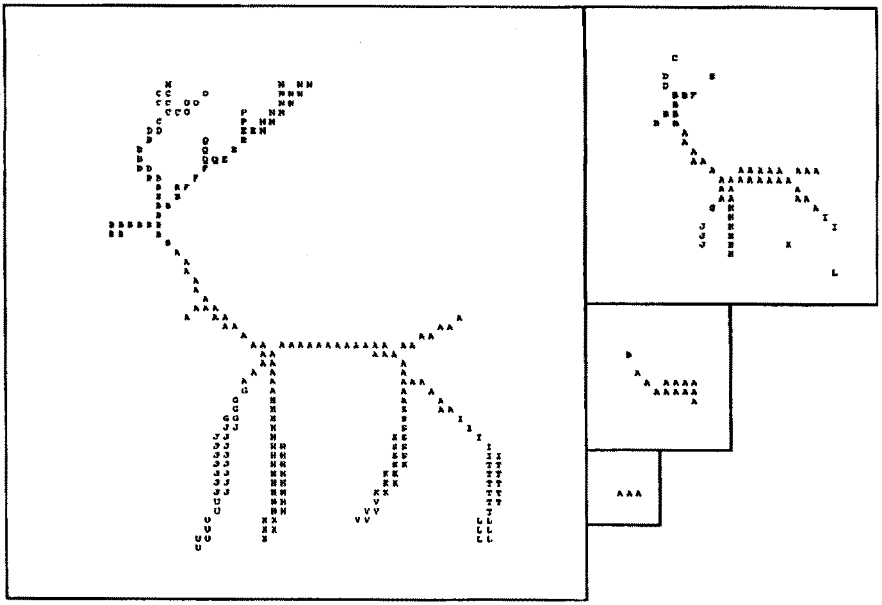

Figure 2. Skeleton hierarchy obtained by using the "old" top-down process [7].

At any pyramid level, the skeleton components with the highest permanence are processed first to identify the corresponding child components. Components with smaller and smaller permanence are then processed and assign to their child components only pixels that have not yet been included in other child components with larger permanence. The criteria used to establish the parent-child correspondence favours skeleton components that are more stable, that is components that are present at many pyramid levels.

In Figure 2, the hierarchy produced on the skeleton of the test pattern by the top-down process introduced in [7] is shown. The same letter is used to denote skeletal pixels belonging to components enjoying the parent-child relationship. At each level, letter A denotes the skeleton component with the highest permanence, letters B, C-L and M-W are components with smaller and smaller permanence. We note that a number of components larger than the intuitively expected one characterises the hierarchy. Skeletal pixels that are grouped into a unique connected component onto a given level, correspond to pixels constituting distinct components onto the immediately smaller resolution level, due to unavoidable pattern disconnections occurred while building the AND-pyramid. Since the hierarchy is built starting from the lower resolution levels where disconnections are likely to occur, at higher levels subsets of the skeleton representing regions of the pattern perceived as a whole are segmented into a number of components with different permanence that, hence, are assigned to different hierarchy levels (see, for example, the skeleton subsets in correspondence with the legs of the deer in Figure 2). The top-down process used to establish the parent-child relationship could not recover information on connectedness, lost when storing the pattern in the AND-pyramid.

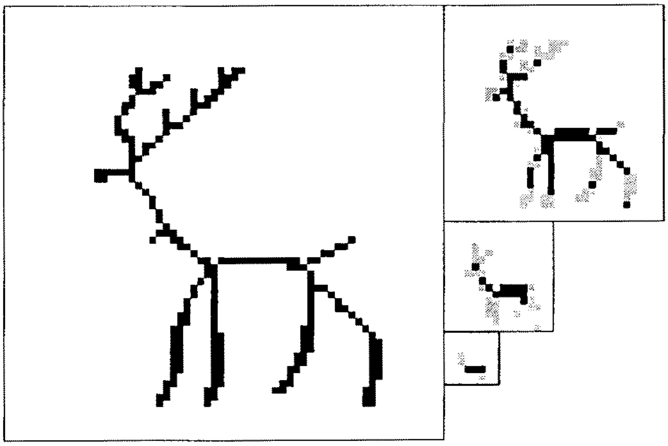

Figure 3. Gray squares denote the extra pixels added to the skeletal pixels (black squares) by the bottom-up OR-projection process.

2.2 Bottom-Up Component Restoring Process

To recover the information on how skeleton fragments should be grouped into components, the pyramid has to be checked bottom-up, so that skeleton subsets that are connected at a given level can transfer connectedness information onto the successive smaller resolution levels. The OR logic operation is used for this purpose.

An *OR-projection* process is accomplished simultaneously for all pairs of successive levels, $2^k x 2^k$ and $2^{k-1} x 2^{k-1}$, $3<k\leq n$. Any level $2^k x 2^k$ is partitioned into 2x2 blocks of pixels (quadruplets) and, for every quadruplet with at least a skeletal pixel, its parent pixel on level $2^{k-1} x 2^{k-1}$ is changed to black, (if it was white).

As an effect of the OR-projection, a number of extra pixels are added to the skeletal pixels found during skeletonization (see Figure 3). The extra pixels modify both the number of connected components and the structure of the set of the skeletal pixels on level $2^{k-1} x 2^{k-1}$, making it resemble more closely the set of skeletal pixel on level $2^k x 2^k$. We distinguish two types of extra pixels. Extra pixels of type 1 link skeletal pixels on level $2^{k-1} x 2^{k-1}$, that would otherwise been grouped into a number of distinct components. Extra pixels of type 1 should be kept if we like a more intuitive skeleton decomposition, that is able to identify a subset of the skeleton as a unit even when, due to resolution problems, the represented region appears as disconnected at that level of the AND-pyramid. Extra pixels of type 2 correspond to regions of the pattern that are totally absent at level $2^{k-1} x 2^{k-1}$. Extra pixels of type 2 should be removed. To this aim, topology preserving removal operations can be repeatedly applied to the extra pixels. For any of them, the connectivity number C_8, as defined in [10], is used to count the number of components of black pixels (skeletal and extra pixels) in its neighbourhood. Extra pixels having $C_8\leq 1$ are sequentially removed. These pixels are, in fact, not necessary for connectedness maintenance. Removal is iterated as far as removable extra pixels are found.

The above sketched removal process removes all type 2 extra pixels, but also reduces to unit width the set of extra pixels of type 1. This might prevent the identification of all pixels expected to constitute the child components, when

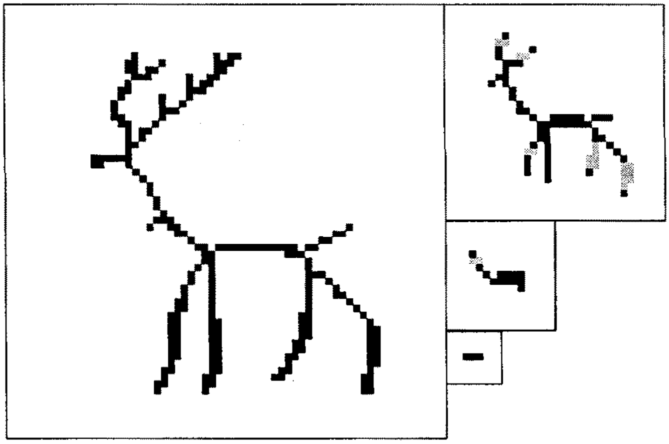

Figure 4. Gray squares denote the extra pixels restoring skeleton connectedness.

establishing the parent-child relationship. Indeed, the process sketched in Section 2.1 uses the two-pixel wide set of skeletal pixels, rather than the unit wide skeletons, to build the hierarchy. Thus, the set of type 1 extra pixels should not be reduced to unit width. We proceed as follows. Rather than actually removing the extra pixels for which it is $C_8 \leq 1$, we *mark* them as pixels candidates for removal. (Note that, when computing C_8, marked pixels in the neighbourhood of any extra pixel are interpreted as if they were already white pixels.) Marking is iterated as long as extra pixels that can be marked are found. Then, the marker is simultaneously removed from all pixels having a horizontal/vertical neighbour that is a non-marked extra pixel. Only pixels that are still marked after this process are type 2 pixels and are finally removed. In Figure 4, the effect of the removal process is shown.

The process described in Section 2.1 is then applied to build the hierarchy on the skeleton modified to restore connectedness. The result are shown in Figure 5, where the extra pixels have all been removed to facilitate the comparison of the performance of the new algorithm with that of [7], shown in Figure 2. In the bottom level, letters A, B, C-H and I-Q denote components with permanence 4, 3, 2 and 1, respectively. The number of components has significantly been reduced with respect to Figure 2.

A few noise components, defined as components consisting of single pixels, still affect the hierarchy. They can be found at the periphery of other components (e.g., at level $2^6 \times 2^6$ the tips of the horns of the deer labelled I or M, both adjacent to components with permanence 2) or in between components whose parents were indeed adjacent at the immediately lower resolution level (e.g., at level $2^6 \times 2^6$ pixel labelled N in between component B, with permanence 3, and D, with permanence 2, in the horns of the deer). Noise components can be originated at any level and are all characterised by permanence equal to 1. They should be absorbed by adjacent components to simplify the structure of the resulting decomposition. This process is performed at each level of the pyramid, before identifying the connected components directly originating at that level, i.e., the components with permanence equal to 1. Each skeletal pixel not already assigned to any component that has *all* its skeletal neighbours already assigned to some component is assigned to the neighbouring component having the highest permanence. Finally, reduction to unit width is performed on all pyramid levels by means of an iterative thinning, based on topology

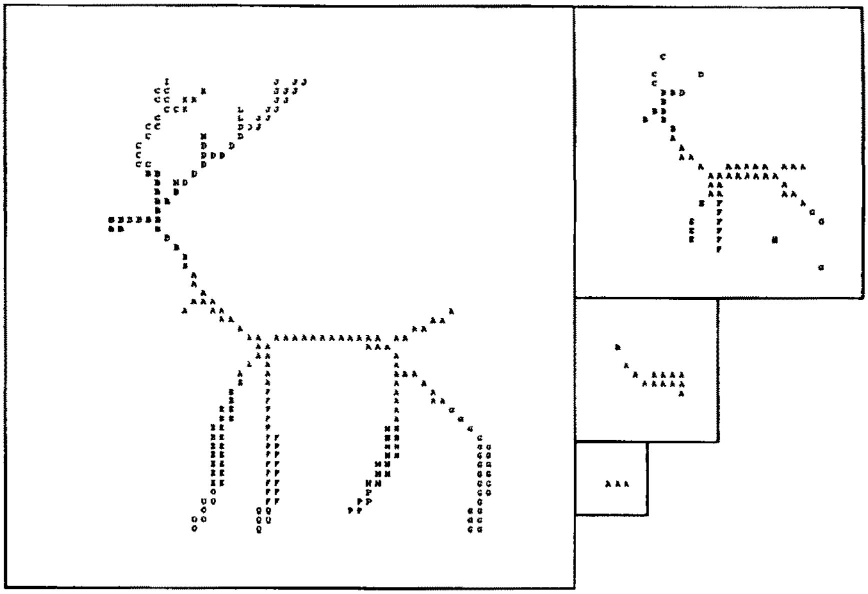

Figure 5. Skeleton hierarchy obtained by applying the top-down decomposition after using a bottom-up connectedness restoring process.

preserving removal operations. At each iteration, removal is active only on pixels having a given permanence, starting from the lowest permanence. In this way, removal of pixels belonging to components with lower permanence (and hence having smaller significance) is favoured. The number of iterations required is equal to the maximum permanence in the hierarchy, i.e., is equal to the number (n-2) of pyramid levels.

The resulting hierarchy for the unit wide skeleton can be seen in Figure 6, where again the extra pixels, that are only used to establish correspondences between components, are not shown. Letters A, B, C-H and I-N denote components with decreasing permanence (4, 3, 2 and 1 in the bottom level, respectively). A smaller number of components, all significant, has been obtained and the hierarchical decomposition is more in accordance with human intuition.

3 Conclusion

A method to hierarchically rank components of multi-scale skeletons computed at all resolution levels of an AND-pyramid has been presented. The proposed method uses both top-down and bottom-up processes to identify skeleton components. Subsets of the skeleton expected to represent given pattern subsets are taken as a whole even if they consist of different number of connected sets of skeleton pixels at different scales. This is achieved by inferring connectedness information from the higher resolution scales, using a bottom-up analysis of the pyramid, in addition to the more intuitive top-down process used to build the hierarchy. Moreover, small noise components are absorbed by adjacent more significant components. The decomposition of the skeleton obtained in this way is in accordance with human intuition.

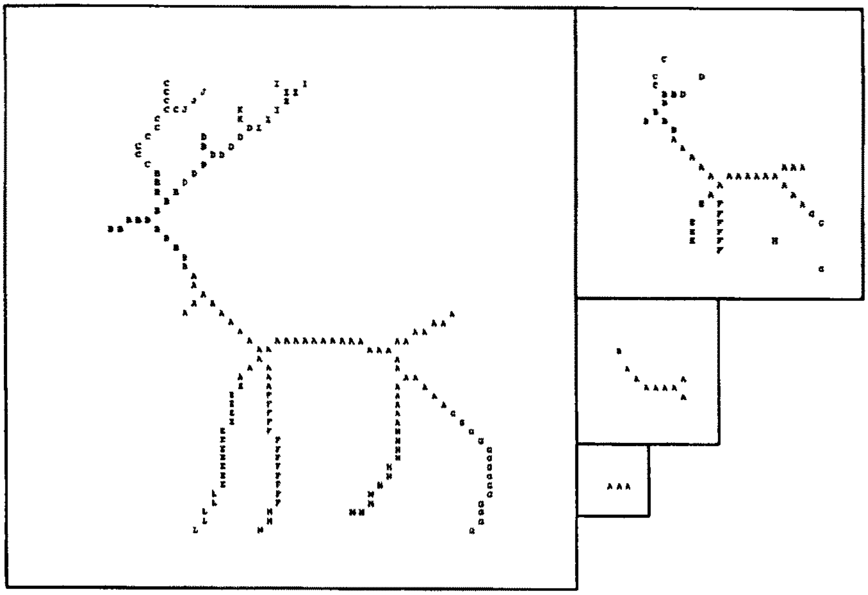

Figure 6. The resulting skeleton hierarchy.

References

1 H.Blum, R.N.Nagel, "Shape description using weighted symmetric axis features", *Pattern Recognition*, **10**, 167-180 (1978).

2 G. Sanniti di Baja, E. Thiel, "(3,4)-weighted skeleton decomposition for pattern representation and description", *Pattern Recognition*, **27**, 1039-1049 (1994).

3 A.Toet, "A hierarchical morphological image decomposition", *Pattern Recognition Letters*, **11**, 267-274 (1990).

4 H. Rom, G.Medioni, "Hierarchical decomposition and axial shape description", *IEEE Trans. on PAMI*, **15**, 973-981 (1993).

5 A.R. Dill, M.D. Levine, P.B. Noble, "Multiple resolution skeletons," *IEEE Trans. on PAMI* **9**, 495-504 (1987).

6 S.M.Pizer, W.R.Oliver, S.H.Bloomberg, "Hierarchical shape description via the multiresolution symmetric axis transform," *IEEE Trans. on PAMI* **9**, 505-511 (1987).

7 G.Borgefors, G.Ramella, G.Sanniti di Baja, "Multi-scale skeletons from binary pyramids", *Proc. 3rd Int.Workshop on Visual Form*, May 97, Capri, Italy.

8 G.Borgefors, G.Ramella, G.Sanniti di Baja, "Multiresolution representation of shape in binary images", in *Discrete Geometry for Computer Imagery*, S.Miguet, A.Montanvert and S.Ubéda (Eds.), LNCS **1176**, Springer-Verlag, 51-58 (1996).

9 C.Arcelli, G.Sanniti di Baja, "A one-pass two-operations process to detect the skeletal pixels on the 4-distance transform," *IEEE Trans. on PAMI* **11**, 411-414, 1989.

10 S.Yokoi, J.I.Toriwaki, T.Fukumura, "An analysis of topological properties of digitized binary pictures using local features", *CGIP*, **4**, 63-73 (1975).