

GLOBAL SKELETON

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
MIDDLE EAST TECHNICAL UNIVERSITY

BY

MURAT GENÇTAV

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF MASTER OF SCIENCE
IN
COMPUTER ENGINEERING

JANUARY 2011

Approval of the thesis:

GLOBAL SKELETON

submitted by **MURAT GENÇTAV** in partial fulfillment of the requirements for the degree of **Master of Science in Computer Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Adnan Yazıcı
Head of Department, **Computer Engineering**

Prof. Dr. Sibel Tari
Supervisor, **Computer Engineering Department, METU**

Examining Committee Members:

Assoc. Prof. Halit Oğuztüzün
Computer Engineering Dept., METU

Prof. Dr. Sibel Tari
Computer Engineering Dept., METU

Asst. Prof. Tolga Can
Computer Engineering Dept., METU

Asst. Prof. İlkay Ulusoy
Electrical and Electronics Engineering Dept., METU

Dr. Hacer Yalım Keleş
Hacer Yalım Keleş Private Company

Date:

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last Name: MURAT GENÇTAV

Signature :

ABSTRACT

GLOBAL SKELETON

Gençtav, Murat

M.Sc., Department of Computer Engineering

Supervisor : Prof. Dr. Sibel Tari

January 2011, 99 pages

A novel and unconventional shape description scheme is proposed which captures the hierarchy of parts and medial descriptors. Both the parts and the medial descriptors are extracted simultaneously, in a complementary fashion, using a real valued function defined over the shape domain. The function arises out of both global and local interactions within the shape domain and it is related to an extension of a linear elliptic PDE with an integral term. The part hierarchies, extracted via level sets and watersheds of the function, are organized into proper binary trees, and the medial descriptors, extracted via ridges and watersheds of the function, are organized as rooted depth-1 trees. The medial descriptors (we named global skeleton) consist of two distinct medial abstractions. The limbs and prominent boundary features are captured in the form of conventional skeletons. Secondly, the coarse structure of the shape is captured in the form of a watershed region, which is a powerful tool in respect to both stability and representation of prominent shape properties. Additionally, as an important technical contribution that addresses part matching, the randomized hierarchy tree is introduced that endows the part hierarchy tree with a probabilistic structure.

Keywords: Shape Representation, Matching, Skeletons, Hierarchical Shape Decomposition

ÖZ

GLOBAL İSKELET

Gençtav, Murat

Yüksek Lisans, Bilgisayar Mühendisliği Bölümü

Tez Yöneticisi : Prof. Dr. Sibel Tari

Ocak 2011, 99 sayfa

Bu çalışmada, uzuvların hiyerarşisini ve simetri betimleyicilerini yakalayan yeni bir şekil betimleme şeması önerilmiştir. Uzuvlar ve simetri betimleyicileri eşzamanlı ve birbirini tamamlayıcı bir şekilde, şekil alanı üzerinde tanımlı ve reel bir fonksiyonu kullanılarak çıkarılmıştır. Bu fonksiyon şekil alanındaki global ve yerel etkileşimlerden ortaya çıkmış olup, lineer eliptik bir kısmi diferansiyel denkleme bir integral terim eklenmesiyle elde edilen denklemin çözümüdür. Uzuv hiyerarşileri fonksiyonun seviye kümeleri ve watershed bölgelerinden çıkarılıp ikili ağaç yapısında; simetri betimleyicileri ise tepe noktaları ile watershed bölgelerinden çıkarılıp tek derinlikli ağaç yapısında temsil edilmiştir. Global iskelet adı verdiğimiz simetri betimleyicileri iki ayrı simetri soyutlamasından oluşmaktadır. Çevresel uzantılar ve önemli sınır ayrıntıları geleneksel iskelet betimleyicileri formunda yakalanmıştır. İkinci olarak, şeklin kaba gövdesi yine bir watershed bölgesi formunda yakalanmıştır. Bu bölge kararlılığı ve önemli şekil özelliklerini temsil etmesi bakımından güçlü bir araç olarak değerlendirilmektedir. Bunlara ek olarak, uzuv eşlemeye yönelik önemli bir teknik katkı olarak uzuv hiyerarşisi ağacına olasılıksal bir yapı kazandıran rastgele hiyerarşi ağacı sunulmuştur.

Anahtar Kelimeler: Şekil Temsili, Eşleme, İskeletler, Hiyerarşik Şekil Parçalama

To my wife

ACKNOWLEDGMENTS

I would like to thank my advisor Prof. Dr. Sibel Tari for her supervision and invaluable helps throughout this study. It is always a pleasure to work with her and get benefit from her vision and knowledge in every step of my research. I also express my pleasure to TÜBİTAK (The Scientific and Technological Research Council of Turkey) for supporting me financially.

TABLE OF CONTENTS

ABSTRACT	iv
ÖZ	v
ACKNOWLEDGMENTS	vii
TABLE OF CONTENTS	viii
LIST OF TABLES	xi
LIST OF FIGURES	xii
CHAPTERS	
1 INTRODUCTION	1
1.1 Skeletons	1
1.1.1 Instabilities and Strategies	3
1.2 The Silhouette Database	5
1.3 Contributions and Thesis Organization	7
2 THE DISTANCE TRANSFORM OF TARI [9, 10]	8
2.1 Mathematical Formulation	8
2.2 Discussion	10
2.2.1 Comparison to MAT	10
2.2.2 Central and Peripheral Regions	10
2.2.3 Hierarchical Shape Decomposition	12
2.2.4 Connection to the Tari, Shah and Pien Method [11]	12
3 THE PRELIMINARY PART HIERARCHY TREE	15
3.1 Extracting The Preliminary Part Hierarchy Tree	15
3.2 Observations	20
3.2.1 Under- and Over-Partitioning	23
3.2.2 Instabilities in Node Depths	23

3.3	The Coarse Structure	26
4	GLOBAL SKELETON	31
4.1	Detection of Skeleton Points	31
4.2	Building Global Skeleton	33
4.2.1	Extraction of Skeleton Branches	33
4.2.1.1	Pruning	35
4.2.2	Constructing the Coordinate Frame	40
4.3	Discussion	45
4.3.1	Comparison with Aslan and Tari [2]	45
4.3.1.1	The ω Function vs. the Edge Strength Function	46
4.3.1.2	Describing Spatial Relations	46
4.3.1.3	Coarse Forms of Shapes	47
4.3.1.4	Topological Categorization of Shapes	48
4.3.1.5	Pruning Strategy	48
4.3.1.6	Summary of the Comparison	50
5	MATCHING	51
5.1	Randomized Hierarchy Tree	51
5.1.1	Randomization Procedure	52
5.1.2	Matching Experiments	53
5.2	Matching Global Skeletons	64
5.2.1	The Context Tree	64
5.2.2	Tree Editing Distance Algorithm by Zhang and Shasha [12]	66
5.2.3	Matching Strategies for Different Types	69
	Matching Two Single-Body Shapes:	69
	Matching Two Multiple-Body Shapes:	69
	Matching an Intermediate-Form Shape to a Shape of a Main Type:	69
	Matching Two Intermediate-Form Shapes:	70
5.2.4	Experimental Results	70

6	CONCLUSION	74
	REFERENCES	76
A	GLOBAL SKELETONS OF THE SHAPES IN THE DATABASE	77
B	MATCHING GLOBAL SKELETONS: INTRA-CATEGORY DEMONSTRATIONS	80
C	MATCHING GLOBAL SKELETONS: RETRIEVAL RESULTS	95

LIST OF TABLES

TABLES

Table C.1	Retrieval results for category 1.	95
Table C.2	Retrieval results for category 2.	95
Table C.3	Retrieval results for category 3.	96
Table C.4	Retrieval results for category 4.	96
Table C.5	Retrieval results for category 5.	96
Table C.6	Retrieval results for category 6.	97
Table C.7	Retrieval results for category 7.	97
Table C.8	Retrieval results for category 8.	97
Table C.9	Retrieval results for category 9.	98
Table C.10	Retrieval results for category 10.	98
Table C.11	Retrieval results for category 11.	98
Table C.12	Retrieval results for category 12.	99
Table C.13	Retrieval results for category 13.	99
Table C.14	Retrieval results for category 14.	99

LIST OF FIGURES

FIGURES

Figure 1.1 MAT of a rectangle shape: a) the shape b) the nearest distance surface c) some level curves on the surface d) skeleton	2
Figure 1.2 Maximal disc formulation	2
Figure 1.3 Instability due to boundary perturbation	3
Figure 1.4 Instabilities in ligature regions caused by small changes on the boundary. (Adapted from Figure 14 in [13])	5
Figure 1.5 The silhouette database by Aslan [2]	6
Figure 2.1 Distance transforms for a human figure: a) the figure b) Euclidean distance transform c) distance transform of Tari [9, 10] (note that dark blue curves have negative values)	11
Figure 2.2 Separation of central and peripheral regions using zero level curve (bold, dark-red curve)	13
Figure 2.3 (Taken from [9]) The effect of α . a) Plot of ω for $\alpha \leq \frac{1}{ \Omega }$. The locations of the zero-crossings and the extrema are not affected by different choices of α . b) The solid blue line is Euclidean distance function d . The dashed line is ω when $\alpha = \frac{2}{ \Omega }$ and the dash-dotted line is ω when $\alpha = 2$. Here, we see that data fidelity criterion dominates.	13
Figure 3.1 The preliminary part hierarchy tree of a turtle shape. Central region is the black region at node 11 and peripheral region is the black region at node 21.	16
Figure 3.2 Part of the preliminary part hierarchy tree of a dog-bone shape. Disconnected central region is split into two nodes (11 and 12).	16
Figure 3.3 The preliminary part hierarchy tree of a shape with disconnected peripheral region. Peripheral region is split into three nodes (21, 22 and 23).	17

Figure 3.4 Saddle point on a surface (from Wikipedia).	18
Figure 3.5 A sample run of decomposition algorithm on a ring-like peripheral region where $s_1 > s_2$. p_1, p_2 are saddle points and b_1, b_2 are corresponding watershed boundaries.	20
Figure 3.6 The preliminary part hierarchy tree for a human shape. The first number put in the left hand side of the nodes indicates the partitioning level (splitting saddle point value) of the part and the second number is the node identifier.	21
Figure 3.7 Segmentations of the shapes in the database. Central regions are gray and peripheral parts are colored.	22
Figure 3.8 Under-partitioning problem. Front legs of the horse (node 2112) are not separated.	23
Figure 3.9 Over-partitioning problem. The node 2111 is divided into two parts unin- tuitively.	24
Figure 3.10 The preliminary part hierarchy tree for a different human shape.	25
Figure 3.11 The preliminary part hierarchy tree for a cat shape.	27
Figure 3.12 The preliminary part hierarchy tree for another cat shape.	28
Figure 3.13 Saddle point order vs. saddle point value. (Red dots indicate locations of part centers.)	29
Figure 3.14 Coarse structures (black regions) of some shapes in the database.	30
Figure 3.15 Ribbon-like vs. blob-like parts. The blob-like head is mostly included by the coarse structure, while ribbon-like limbs (arms and legs) are excluded.	30
Figure 4.1 Sample skeletons (K) extracted from (a) human, (b) hand, (c) horse, and (d) turtle shapes	32
Figure 4.2 Coarse contour (yellow) embedded on the skeleton of the hand shape.	33
Figure 4.3 Skeleton branches (red lines) as groupings of connected skeleton points which track evolution of prominent protrusions. Skeleton points are included by the enclosing contexts (bright colored regions), but excluded by the coarse struc- ture (darker region on the shape center).	34
Figure 4.4 Illustration of (a) gradient magnitude, (b) saliency measurements for skele- ton points of a star shape with boundary noise	37

Figure 4.5	A sample run of the skeleton extraction algorithm for the right leg of a frog shape. (a) Skeleton points placed within the enclosing context and outside of the coarse structure. (b) One of the initial groupings (group of connected skeleton points) shown with critical points: p_{center} (blue), p_{ref} (yellow) and p_{start} (red). (c) Weighted distances of the skeleton points in the initial grouping. (d) End points. (e) Candidate skeleton branches extracted from the initial grouping. (f) Extracted skeleton branches after processing all of the initial groupings and selecting salient branches that represent prominent protrusions.	39
Figure 4.6	Star shape as a sample single-body shape. Shape boundary (outer contour) and the level contours (inner contours) of the central region are depicted.	40
Figure 4.7	A dog bone shape as a sample multiple-body shape. Shape boundary (outer contour) and the level contours (inner contours) of the two central regions are depicted.	42
Figure 4.8	Another dog bone shape as a sample intermediate-form shape. Shape boundary (outer contour) and the level contours (inner contours) of the central region are depicted.	42
Figure 4.9	Two horse shapes of different types.	43
Figure 4.10	Illustration of the constructed polar coordinate system for a single-body shape. Red and green lines are polar axes originating from the center of the central region (black square). The radial coordinate r of a skeleton branch is the length of the position vector \vec{s} (dashed arrows). The angular coordinates θ_1 and θ_2 are the counterclockwise angles between the vector \vec{s} and two polar axes.	43
Figure 4.11	Illustration of the constructed local coordinate systems for a multiple-body shape.	45
Figure 4.12	Comparison of coordinate frames. a) (Taken from [1]) Illustration of the constructed coordinate frame in the work of Aslan and Tari [2] for a hand shape. b) Illustration of our coordinate frame on the same hand shape.	47
Figure 4.13	(Taken from [1]) a) The skeleton branches of a hand shape in [2]. b) The level curves for the evolving shape boundary and the states of the hand shape at the times the skeleton branches indicated by the numbers in (a) terminated.	49

Figure 4.14 Global skeleton representation for the hand shape that is used for illustrations in Figure 4.13	49
Figure 5.1 The preliminary part hierarchy tree of a heart shape.	53
Figure 5.2 The preliminary part hierarchy tree of a cat shape.	56
Figure 5.3 The preliminary part hierarchy tree of a horse shape.	56
Figure 5.4 The preliminary part hierarchy tree of a cat shape.	57
Figure 5.5 (First example) Matchings between leaf parts of cat and horse shapes, whose preliminary part hierarchy trees were given in Figures 5.2 and 5.3. Front legs of the cat are encircled to imply their parent in the preliminary part hierarchy tree.	58
Figure 5.6 (Second example) Matchings between leaf parts of two cat shapes, whose preliminary part hierarchy trees were given in Figures 5.2 and 5.4.	58
Figure 5.7 The preliminary part hierarchy tree of a human shape.	60
Figure 5.8 The preliminary part hierarchy tree of an occluded human shape.	60
Figure 5.9 Matching human shapes (in Figures 5.7 and 5.8): The winning randomized hierarchy tree pair with node associations.	61
Figure 5.10 The preliminary part hierarchy tree of a horse shape.	62
Figure 5.11 (Third example) Matchings between leaf parts of a human shape and its occluded version, whose preliminary part hierarchy trees were given in Figures 5.7 and 5.8.	63
Figure 5.12 (Fourth example, first trial) Optimal matchings between leaf parts of horse shapes, whose preliminary part hierarchy trees were given in Figures 5.3 and 5.10. Confusion appears in matching third legs and tails.	63
Figure 5.13 (Fourth example, second trial) Sub-optimal matchings between leaf parts of horse shapes, whose preliminary part hierarchy trees were given in Figures 5.3 and 5.10. No confusion appears in this case.	63
Figure 5.14 The symbolic context tree representation for a single-body shape.	65
Figure 5.15 The symbolic context tree representation for a multiple-body shape.	66
Figure 5.16 Topology of a context tree.	67
Figure 5.17 Distance matrix.	71

Figure 5.18 Matching a star to a turtle. Placements and lengths of the skeleton branches are very similar. Coarse structures may provide a clear distinction.	71
Figure 5.19 Matching turtles. The turtle on the right has extra branches which increase the editing distance.	72
Figure 5.20 Matching a cat to a horse.	72
Figure A.1 Global skeletons.	77
Figure A.2 Global skeletons.	78
Figure A.3 Global skeletons.	79
Figure B.1 Matching: Category 1	81
Figure B.2 Matching: Category 2	82
Figure B.3 Matching: Category 3	83
Figure B.4 Matching: Category 4	84
Figure B.5 Matching: Category 5	85
Figure B.6 Matching: Category 6	86
Figure B.7 Matching: Category 7	87
Figure B.8 Matching: Category 8	88
Figure B.9 Matching: Category 9	89
Figure B.10 Matching: Category 10	90
Figure B.11 Matching: Category 11	91
Figure B.12 Matching: Category 12	92
Figure B.13 Matching: Category 13	93
Figure B.14 Matching: Category 14	94

CHAPTER 1

INTRODUCTION

Skeletons have been found useful as a shape representation method being invariant under translation, rotation and scale. They are medial axis representations for shapes. The organization of the skeleton branches is widely used for matching two shapes and measuring their similarities by the help of some properties associated with parts. However, skeleton analysis has been suffering from several instabilities due to boundary noise, numerical issues related with discrete nature of data etc. In the course of this chapter, reader can find some background information such as the emergence of skeletons with Blum's famous Medial Axis Transform (MAT), main issues related with skeleton analysis and some strategies proposed to overcome these difficulties. Then in the final section, reader can find main contributions and the general organization of the thesis.

1.1 Skeletons

Skeletons are introduced into the literature by Blum with his famous MAT [3] in sixties. According to his definition, medial axis or skeleton is the locus of discontinuities in the nearest distance surface. A sample run of MAT on a rectangle shape is depicted in Figure 1.1. As can be seen, the locus of discontinuities emerge as the corner points and the self-intersections of the level contours. If the distance values of skeleton points are kept, one can recover the original shape. This is a unique property of MAT, which however brings some instability problems that will be explained soon.

Another formulation of medial axis transform can be made in terms of maximal discs. A maximal disc is a disc which is completely contained inside the shape and which is not con-

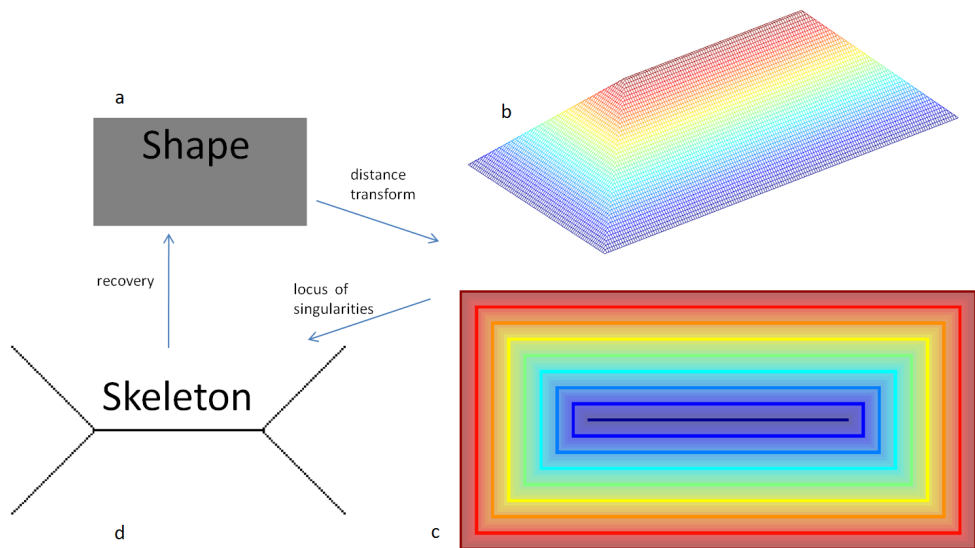


Figure 1.1: MAT of a rectangle shape: a) the shape b) the nearest distance surface c) some level curves on the surface d) skeleton

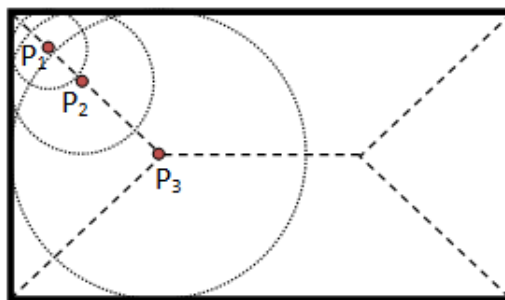


Figure 1.2: Maximal disc formulation

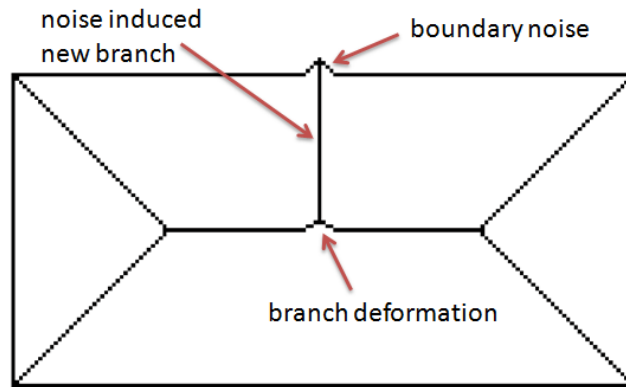


Figure 1.3: Instability due to boundary perturbation

tained in such another disc [4]. The loci of centers of all maximal discs form the medial axis or skeleton of the shape (Figure 1.2).

An important property of skeletons according to Blum's definition can be explained using maximal discs formulation. As Figure 1.2 reveals, each maximal disc, which is responsible of one skeleton point at its center, touches the shape boundary at two or more tangent points. This makes a skeleton point be equidistant from at least two boundary points. Thus it is formed as a result of the interaction between two or more boundary points.

1.1.1 Instabilities and Strategies

Yet MAT may seem promising and very simple to compute and use, researchers has been faced with many theoretical and implementation issues. As a consequence, a great number of studies were proposed to handle these issues.

Maybe the most studied issue is the sensitivity to boundary perturbations. The problem becomes apparent even in the shape with a very small boundary noise in Figure 1.3. It causes a long and unintuitive new branch and deformation on the old branch. Remember that the skeleton preserves every details of the boundary including the noise. In the literature, there is a great endeavor to tackle this problem. Among different strategies some examples are: Pre-smoothing, pruning after skeletonization, curve evolution and smoothed distance functions.

Presmoothing is defined as filtering out the noise present in the shape boundary before skeletonization. Pruning after skeletonization is another technique which detects and eliminates noise induced skeleton branches based on a measure of significance.

The fundamental formulation of curve evolution was introduced by Osher and Sethian [7]. The work of Kimia et al. [5] was the first one that utilizes the method to extract shape skeleton. Curve evolution is defined as deforming the boundary curve by moving the points on it, where the velocity of a point depends on two components: morphology based component and curvature based smoothing component. Here, the curvature dependent motion is introduced for noise removal throughout the evolution of the boundary curve. Evolving curve generates a surface, which is a smoothed version of the nearest distance surface used in MAT. Since the corners are rounded out by smoothing, points of high curvature on the level curves serve as the substitute for the skeleton in the new surface. The new skeleton is disconnected, and significance of branches is proportional to their survival over scales (under increasing weights of smoothing component over morphology component). Noise induced branches are pruned according to a significance threshold.

An alternative approach, which is closely related to the curve evolution method, is suggested in the work of Tari, Shah and Pien [11]. The basic tool is the edge strength function whose level curves represent deformation of the boundary curve under the influence of morphology and curvature dependent motions. The method is computationally simpler and faster than the curve evolution. Another advantage over curve evolution is that it can be directly applied to gray scale images without extracting the shape outline.

Another challenging issue emerges in ligature regions. A ligature region is a segment of skeleton where skeleton branches of two or more parts are connected. Ligature regions are extremely sensitive to addition or deletion of parts and also to deformations in the boundary of the shape (Figure 1.4), whereas their contribution to the reconstruction and representation of shape boundary is still little. This fact suggests the idea that the skeleton branches can physically be disconnected, and the connection may be handled in a higher level of abstraction. In the following paragraphs, two works that follow the idea are mentioned briefly.

In their recent work [6], Macrini et al. proposed a method for identification and regularization of ligature regions based on geometric properties of skeletons. The method yields a set of disconnected skeleton branches, which they call bones, after excluding ligature regions and

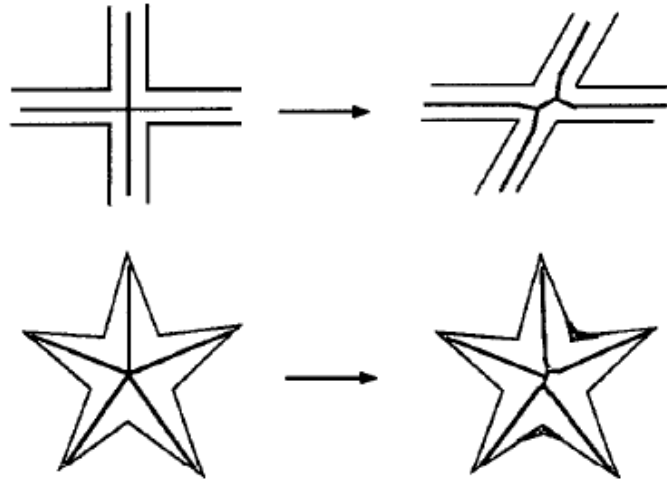


Figure 1.4: Instabilities in ligature regions caused by small changes on the boundary. (Adapted from Figure 14 in [13])

restoring deformed segments on non-ligature regions that represent shape parts. Physically removed ligature segments are then utilized for connecting bones in constructing bone graphs.

In another work [2], Aslan and Tari utilized disconnected nature of the skeleton which is extracted using the method of Tari, Shah and Pien [11]. After grouping the skeleton points into disconnected skeleton branches representing parts, they described spatial organization of branches using a polar coordinate frame.

1.2 The Silhouette Database

The main database used throughout this thesis is constructed by Aslan [2]. It consists of 56 silhouettes from 14 different categories (4 per category). To demonstrate the performance of the work under visual transformations, silhouettes within the same categories are selected such that they have differences in orientation, scale, articulation and boundary details. The database is shown in Figure 1.5.

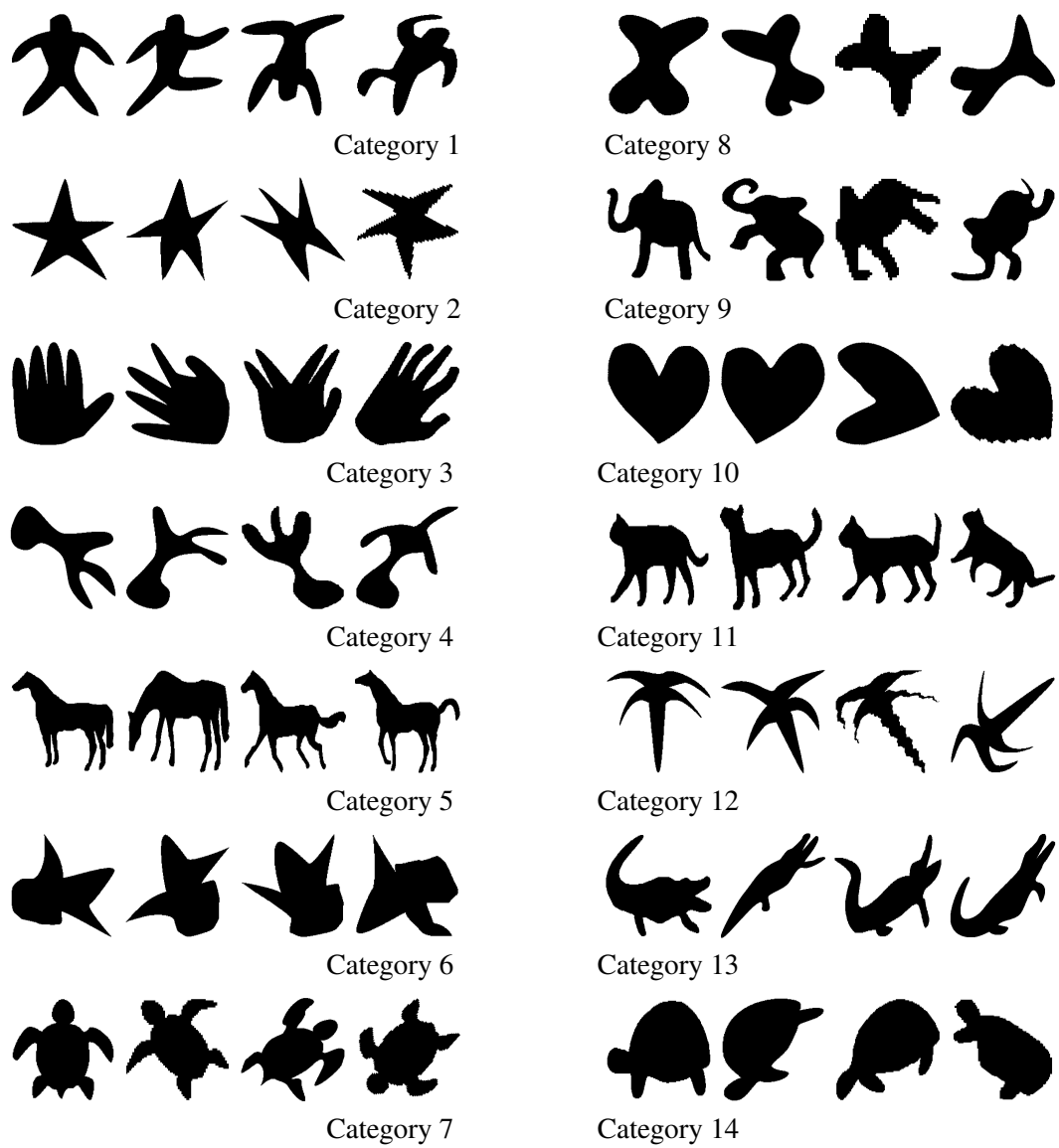


Figure 1.5: The silhouette database by Aslan [2]

1.3 Contributions and Thesis Organization

The main contribution of this thesis work is a novel and unconventional shape description scheme which defines the followings:

1. Description of part partitioning hierarchy via a proper binary tree.
2. Global Skeleton: A new medial descriptor which captures limbs and prominent boundary features in the form of conventional skeletons and the coarse structure of the shape in the form of a watershed region. Resulting descriptions are organized in rooted depth-1 trees.

In addition to the shape description scheme, randomized hierarchy tree is proposed that addresses instabilities in node depths of part hierarchy tree by endowing it with a probabilistic structure.

The thesis is organized as follows: In Chapter 2, the distance transform of Tari proposed and detailed in [9] and in [10] is reviewed, which is the basic tool of the thesis work. Then in Chapter 3, a hierarchical shape decomposition scheme is proposed that extracts a hierarchical structure called the preliminary part hierarchy tree utilizing the level sets of the distance transform of Tari. Global skeleton and its extraction method is explained in Chapter 4. In Chapter 5, first a randomization procedure is proposed to obtain randomized hierarchy tree from the preliminary part hierarchy tree. Then, some illustrative experiments on matching randomized hierarchy trees are discussed. The final part of the Chapter 5 explains the matching method used to match global skeletons and discusses the matching results. Finally, Chapter 6 is a general discussion about the work done and the future work.

CHAPTER 2

THE DISTANCE TRANSFORM OF TARI [9, 10]

This chapter reviews the distance transform of Tari [9, 10] that combines long-range interactions along the shape contour, short-range interactions in the shape interior, and global interaction throughout the shape interior in order to obtain an interior specific distance surface. This surface is the main tool of the thesis work used for construction of both the randomized hierarchy tree and global skeleton.

2.1 Mathematical Formulation

Let ω be a function defined on a discrete shape domain Ω as the minimizer of an energy $E(\omega)$. Let this energy be the sum of two energies: $E_{Bdy}(\omega)$ to model long-range interactions along the shape boundary and $E_{Reg}(\omega)$ to model region based energy. Here, $E_{Reg}(\omega)$ is also a sum of two energies, one models the global ($E_{Reg}^G(\omega)$) and the other the local (short-range) interactions ($E_{Reg}^L(\omega)$) within the shape domain Ω .

$$\begin{aligned} E(\omega) &= E_{Reg}(\omega) + w_{Bdy}E_{Bdy}(\omega) \\ &= E_{Reg}^G(\omega) + E_{Reg}^L(\omega) + w_{Bdy}E_{Bdy}(\omega) \end{aligned} \quad (2.1)$$

where w_{Bdy} is a constant that determines the weight of boundary based energy E_{Bdy} relative to region based energy E_{Reg} . $E(\omega)$ may be considered as the sum of each energy terms at each pixel:

$$E(\omega_{i,j}) = \sum_{i,j \in \Omega} E_{Reg}^G(\omega_{i,j}) + E_{Reg}^L(\omega_{i,j}) + w_{Bdy}E_{Bdy}(\omega_{i,j}) \quad (2.2)$$

Lets start with E_{Bdy} . As already mentioned in Chapter 1, formation of skeleton points is a result of the interactions among boundary points. Considering this fact, Tari modeled long-range interactions among boundary points utilizing Euclidean distance surface. This surface

codes minimum distance to the nearest boundary points at all points in the shape domain Ω . Skeleton points emerge as the ones which are equidistant from at least two boundary points. Eventually, E_{Bdy} is expressed as follows:

$$E_{Bdy}(\omega_{i,j}) = (\omega_{i,j} - d_{i,j})^2 \quad (2.3)$$

where d is Euclidean distance function. The condition to minimize 2.3 becomes:

$$(\omega_{i,j} - d_{i,j}) = 0 \quad (2.4)$$

Obviously, 2.4 states that, in the absence of other terms, the minimizer function ω should be close to Euclidean distance function d .

Lets proceed with region based energies and first consider E_{Reg}^G defined as follows:

$$E_{Reg}^G(\omega_{i,j}) = \frac{1}{|\Omega|} \left(\sum_{(i,j) \in \Omega} \omega_{i,j} \right)^2 \quad (2.5)$$

which is minimized when the condition below is satisfied.

$$\frac{1}{|\Omega|} \sum_{(i,j) \in \Omega} \omega_{i,j} = 0 \quad (2.6)$$

The condition 2.6 is satisfied when the global sum of the minimizer ω over the entire shape domain Ω is zero, independent of the pixel location (i, j) . As an important result of the global influence through E_{Reg}^G , function ω attains both positive and negative values. This result is at the heart of the work of Tarı and of this thesis, and will be discussed in the next section.

Finally, Tarı defines the second region based term E_{Reg}^L as follows:

$$E_{Reg}^L(\omega_{i,j}) = -0.5(\omega_{i+1,j} \cdot \omega_{i-1,j} + \omega_{i,j+1} \cdot \omega_{i,j-1}) \quad (2.7)$$

It is obvious that 2.7 is minimized when the gradient of ω is zero within the shape domain Ω . In case of competence with the other two criteria, this condition will impose smoothness on ω . Formally, the condition takes this form:

$$(-2 + 4) * \omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1} = 0 \quad (2.8)$$

With all three competing terms, the first order derivative w.r.t. $\omega_{i,j}$ takes the following form:

$$\begin{aligned} \frac{\partial E}{\partial \omega_{i,j}} &= \frac{1}{|\Omega|} \left(\sum_{(i,j) \in \Omega} \omega_{i,j} \right) + (w_{Bdy} - 2) \omega_{i,j} - d_{i,j} \\ &+ (4\omega_{i,j} - \omega_{i-1,j} - \omega_{i+1,j} - \omega_{i,j-1} - \omega_{i,j+1}) \end{aligned} \quad (2.9)$$

The minimizer $\omega_{i,j}$ satisfies the following equation:

$$\begin{aligned} d_{i,j} &= -\left(-4\omega_{i,j} + \omega_{i-1,j} + \omega_{i+1,j} + \omega_{i,j-1} + \omega_{i,j+1}\right) \\ &+ \alpha\omega_{i,j} + \frac{1}{|\Omega|} \left(\sum_{(i,j) \in \Omega} \omega_{i,j} \right) \end{aligned} \quad (2.10)$$

where $\alpha = (w_{Bdy} - 2)$.

To summarize, the new distance function ω simultaneously satisfies the following three competing criteria:

1. being close to d
2. having zero mean
3. being smooth

which means that the function ω is an oscillating smooth approximation of Euclidean distance function d .

2.2 Discussion

This section discusses the distance transform of Tarı [9, 10] by emphasizing some of its instrumental aspects that are utilized in this thesis.

2.2.1 Comparison to MAT

A distinguishing property of the function ω is that it attains positive and negative values as mentioned in the previous section. Complemented with the other two criteria (smoothness and being close to function d), ω becomes an oscillating smooth approximation of Euclidean distance transform (Figure 2.1).

2.2.2 Central and Peripheral Regions

A critical level curve, which is the inner zero level curve, emerges as an important instrument to robustly split the shape domain into two regions: the central region and the peripheral

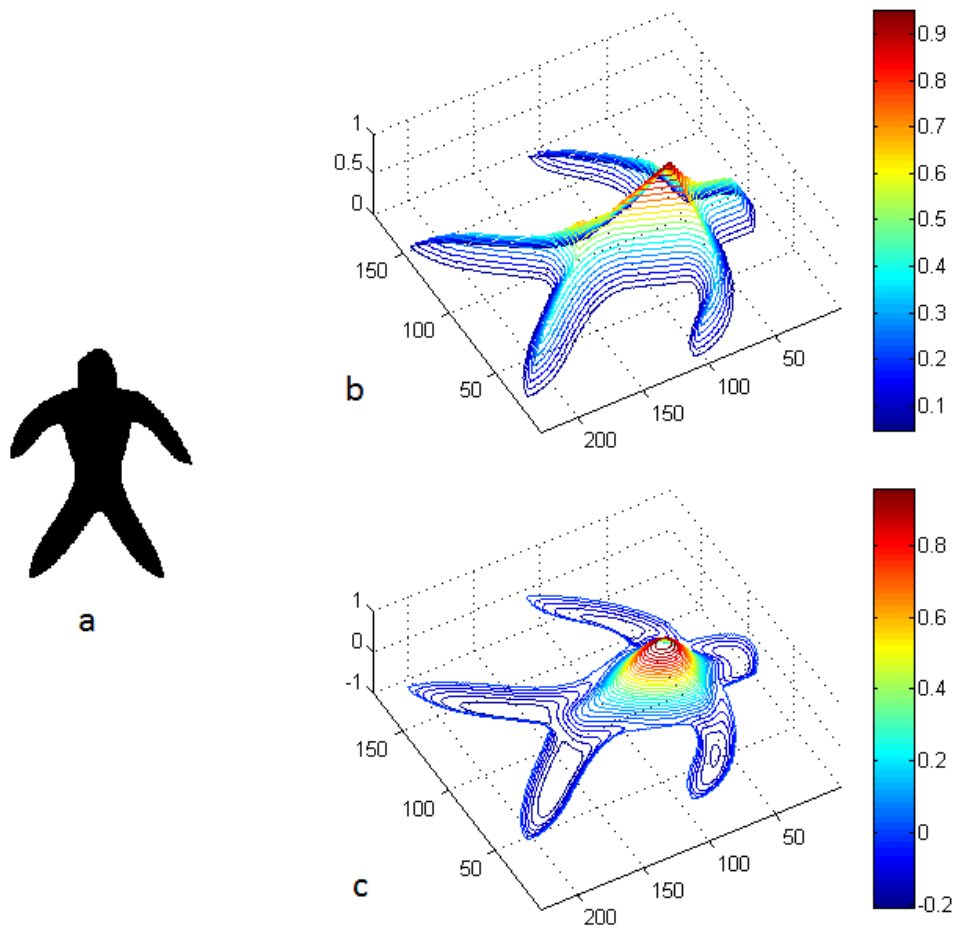


Figure 2.1: Distance transforms for a human figure: a) the figure b) Euclidean distance transform c) distance transform of Tari [9, 10] (note that dark blue curves have negative values)

region (Figure 2.2). Here, the choice of α is important from the aspect that it determines the importance of the first condition (being close to d) relative to the second condition (having zero mean). The effect of different choices of α is shown in Figure 2.3. This figure in 2D is highly illustrative and advisory for a good choice of α . When the value of α doesn't exceed $\frac{1}{|\Omega|}$, the location of neither the zero-crossings nor the extrema points are affected. Thus to obtain the least deformable zero level curve, one has to choose $\alpha \leq \frac{1}{|\Omega|}$. In this thesis, the value of α is chosen as $\frac{1}{|\Omega|^2}$.

2.2.3 Hierarchical Shape Decomposition

In fact, the capability of shape segmentation is not limited with the separation of the central and peripheral regions. Instead, this is just a starting point. Central and peripheral regions can be decomposed into subregions in themselves iteratively resulting in a very intuitive hierarchical decomposition of the shape into its parts. Main tools used for the hierarchical shape decomposition are the level sets of the function ω . Reader can find the details about the hierarchical shape decomposition method developed in this thesis work in Chapter 3.

2.2.4 Connection to the Tari, Shah and Pien Method [11]

Let us drop the term E_{Reg}^G that models the global interaction within the shape domain Ω from the energy $E(\omega)$. Then the minimizer of the reduced energy satisfies the following equation:

$$-\left(-4\omega_{i,j} + \omega_{i-1,j} + \omega_{i+1,j} + \omega_{i,j-1} + \omega_{i,j+1}\right) + \alpha\omega_{i,j} = d_{i,j} \quad (2.11)$$

which is the discretization using central difference approximation, of the PDE below:

$$\begin{aligned} (-\Delta + \alpha)\omega_{i,j} &= d_{i,j} \\ \text{s.t. } \omega_{i,j} &= 0 \text{ for } (i,j) \in \delta\Omega \end{aligned} \quad (2.12)$$

where Δ denotes the Laplace operator. If the right hand side function is chosen as a constant function, one obtains the PDE proposed by Tari, Shah and Pien (TSP) [11] to obtain the smoothed distance function that they call the edge strength function. This function was already mentioned among different strategies that try to overcome instabilities in skeleton analysis.

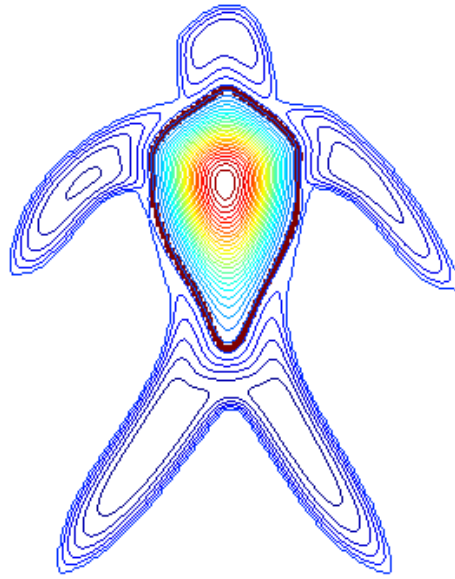


Figure 2.2: Separation of central and peripheral regions using zero level curve (bold, dark-red curve)

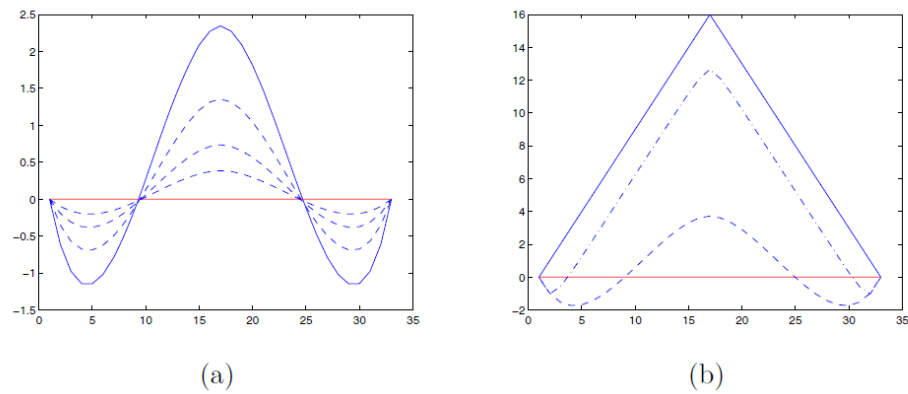


Figure 2.3: (Taken from [9]) The effect of α . a) Plot of ω for $\alpha \leq \frac{1}{|\Omega|}$. The locations of the zero-crossings and the extrema are not affected by different choices of α . b) The solid blue line is Euclidean distance function d . The dashed line is ω when $\alpha = \frac{2}{|\Omega|}$ and the dash-dotted line is ω when $\alpha = 2$. Here, we see that data fidelity criterion dominates.

The connection between distance functions of Tari [9, 10] and of TSP [11] plays an important role in this thesis work. Indeed, the level curves of the distance function of Tari [9, 10] is analogous to the level curves of the distance function of TSP [11] and also shares its advantages such as smoothing out the boundary noise. Exploiting this connection, the skeleton extraction method of TSP [11] will be applied to the distance function of Tari [9, 10] when constructing global skeleton.

CHAPTER 3

THE PRELIMINARY PART HIERARCHY TREE

This chapter introduces the preliminary part hierarchy tree, which represents the hierarchical organization of the shape parts. The preliminary part hierarchy tree is a strict binary tree. The nodes at different levels correspond to the parts at different partitioning hierarchy. The preliminary part hierarchy tree is in fact an intermediate step which will be followed by a more robust tree structure called the randomized hierarchy tree. The randomization procedure and the resulting randomized hierarchy tree are not in the scope of this chapter and will be covered in Chapter 5. In the final part of this chapter, the coarse structure is proposed as an alternative central structure which is the heart of the global skeleton idea.

3.1 Extracting The Preliminary Part Hierarchy Tree

The preliminary part hierarchy tree is the hierarchical structure that reveals the preliminary partitioning hierarchy of a shape into its parts. The root node of the tree represents the shape itself. In the first level of partitioning hierarchy, the shape splits into two main regions named central and peripheral regions. For $s \leq 0$, let $\{\chi_s = (x, y) \in \Omega : \omega(x, y) < s\}$ denote a lower level set of ω . For $s \geq 0$, let $\{\chi^s = (x, y) \in \Omega : \omega(x, y) > s\}$ denote an upper level set of ω . Central region is defined as the upper zero level set χ^0 and the peripheral region as the lower zero level set χ_0 . Central and peripheral regions for a turtle shape are shown in Figure 3.1, at the two depth one nodes 11 and 21 respectively.

If the upper (or lower) zero level set is disconnected, the shape splits into more than one central (or peripheral) regions resulting in a tree with more than two depth one nodes. A dog-bone shape is a good example of shapes having disconnected central region (Figure 3.2).

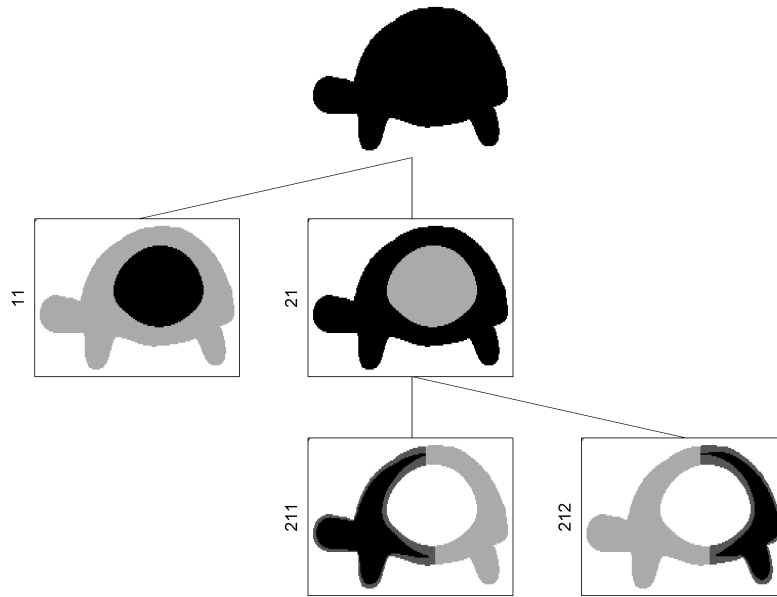


Figure 3.1: The preliminary part hierarchy tree of a turtle shape. Central region is the black region at node 11 and peripheral region is the black region at node 21.

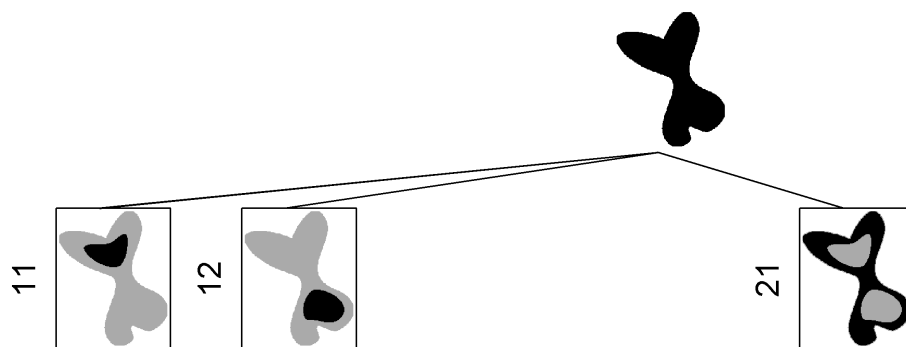


Figure 3.2: Part of the preliminary part hierarchy tree of a dog-bone shape. Disconnected central region is split into two nodes (11 and 12).

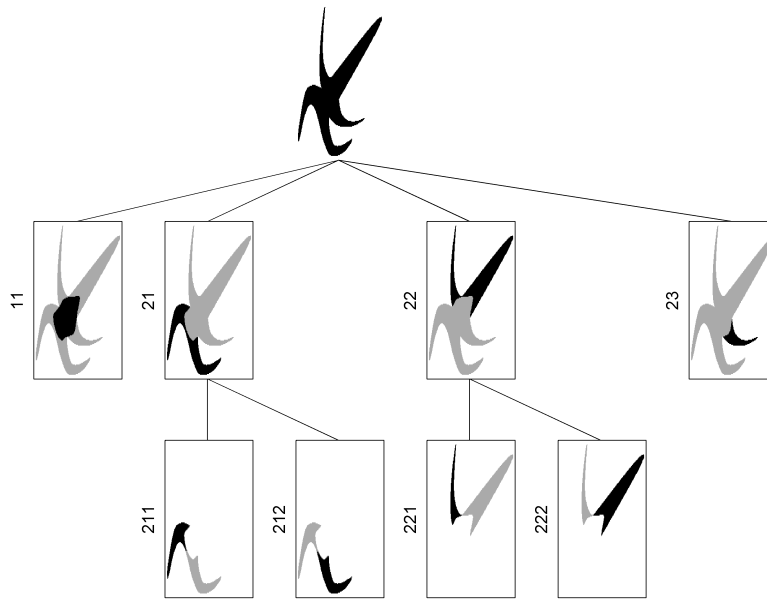


Figure 3.3: The preliminary part hierarchy tree of a shape with disconnected peripheral region. Peripheral region is split into three nodes (21, 22 and 23).

Indeed, this shape consists of two similar main structures (at two ends) which are combined with a narrow neck between them. Intuitively, one expects to obtain two central regions corresponding to these two main structures. An example for shapes having disconnected peripheral region is shown in Figure 3.3.

It is convenient here to mention the tools that will be used for hierarchical decomposition of peripheral region. These tools are saddle points on the ω surface and watershed boundaries passing through saddle points. Saddle points are special set of points where

- gradient magnitude is zero ($|\nabla\omega| = 0$),
- and the sign of the determinant of the Hessian matrix is negative ($\frac{\partial^2}{\partial x^2}\omega \frac{\partial^2}{\partial y^2}\omega - \left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\omega\right)\right)^2 < 0$).

The distinctive property of the saddle points is that they are local minimum points in one direction in their neighborhood while being local maximum in another direction (Figure 3.4).

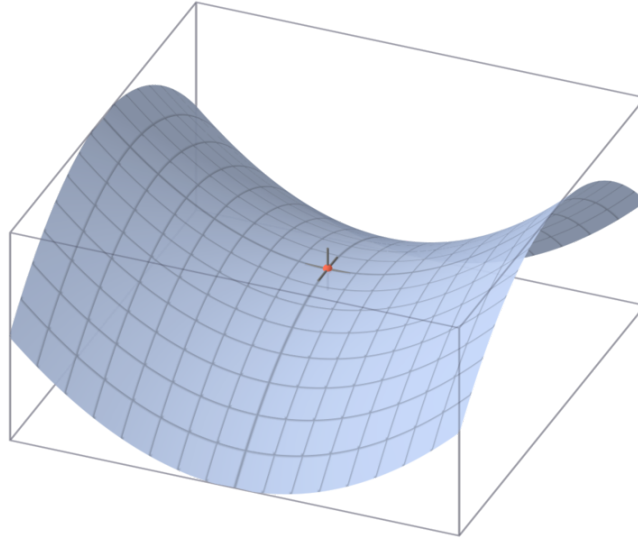


Figure 3.4: Saddle point on a surface (from Wikipedia).

Another important property, which is utilized in partitioning, is that a saddle point is the self-intersection point of the level curve having the same level with the saddle point.

After the initial partitioning through zero level curve, decomposition procedure proceeds with partitioning peripheral region into two parts. If the region is connected, it is a ring-like structure having at least two saddle points with the function values s_1 and s_2 such that $0 > s_1 \geq s_2$.

If $s_1 = s_2$, then the lower level set χ_{s_2} ($= \chi_{s_1}$) is a union of two disjoint lower level sets $\chi_{s_2}^1$ ($= \chi_{s_1}^1$) and $\chi_{s_2}^2$ ($= \chi_{s_1}^2$). If $s_1 \neq s_2$, χ_{s_2} is a proper subset of χ_{s_1} . In this case, since χ_{s_1} is connected, the saddle point s_2 must be used to split the region. In either case, the resulting level sets $\chi_{s_2}^1$ and $\chi_{s_2}^2$, which we call seed parts, constitute the next level of partitioning hierarchy. In Figure 3.1, the seed parts are shown as the black regions at the nodes 211 and 212.

If peripheral region is already not a ring-like structure, only one saddle point with the greatest function value is sufficient for partitioning. Partitioning of nodes 21 and 22 (which build up peripheral region with node 23) in Figure 3.3 are examples of this phenomenon.

Hierarchical decomposition of seed parts continues until reaching a seed part on which ω function doesn't have any saddle point more. The node that corresponds to such a seed part becomes a leaf node. Then, the minimum point (maximum for central regions) that is located within the enclosing context of a leaf part is detected as the part's center point.

One may note from Figure 3.1 that seed parts (black regions at nodes 211 and 212) don't cover the whole shape part that they represent. Somehow, we have to determine the boundary of the shape part represented by a seed part. This issue is handled by calculating the watershed zone that encloses the seed part. We call this zone as the enclosing context. In a tree node, the enclosing context can be seen as the union of black and dark gray areas. Light gray area is the complement region. Union of the complement and the enclosing context gives the parent.

In implementation level, saddle points and the corresponding watershed boundaries passing through them are calculated before the decomposition procedure. Thus partitioning the seed part and the enclosing context occurs simultaneously and efficiently (i.e. when a saddle point is used for splitting a seed part, the corresponding watershed boundary is used for splitting the enclosing context).

To obtain a simple organization, three constraints are considered:

1. Decomposition into too small seed parts is not allowed by inhibition of using saddle points and corresponding watershed boundaries that result in at least one seed part that is smaller than 0.05% of the shape's full size.
2. Decomposition into too small enclosing contexts is not allowed by inhibition of using saddle points and corresponding watershed boundaries that result in at least one enclosing context that is smaller than 0.5% of the shape's full size.
3. Partitioning is not allowed using a watershed boundary that doesn't touch the central region.

If the saddle point with the greatest function value is inhibited to satisfy these constraints, then the next saddle point (if exists) with the greatest function value is used for partitioning.

A sample partitioning at one level is shown in Figure 3.5, where the peripheral region (node 11) is ring-like and saddle point values are different ($s_1 > s_2$). In this case, two watershed boundaries b_1 and b_2 (corresponding to saddle points with values s_1 and s_2 respectively) are used for splitting enclosing contexts, while only the saddle point s_2 is used for splitting seed parts.

Due to our binary splitting procedure, starting from the depth one nodes, each subtree is a proper binary tree. However, there is no guaranty that the whole tree is a binary tree, because

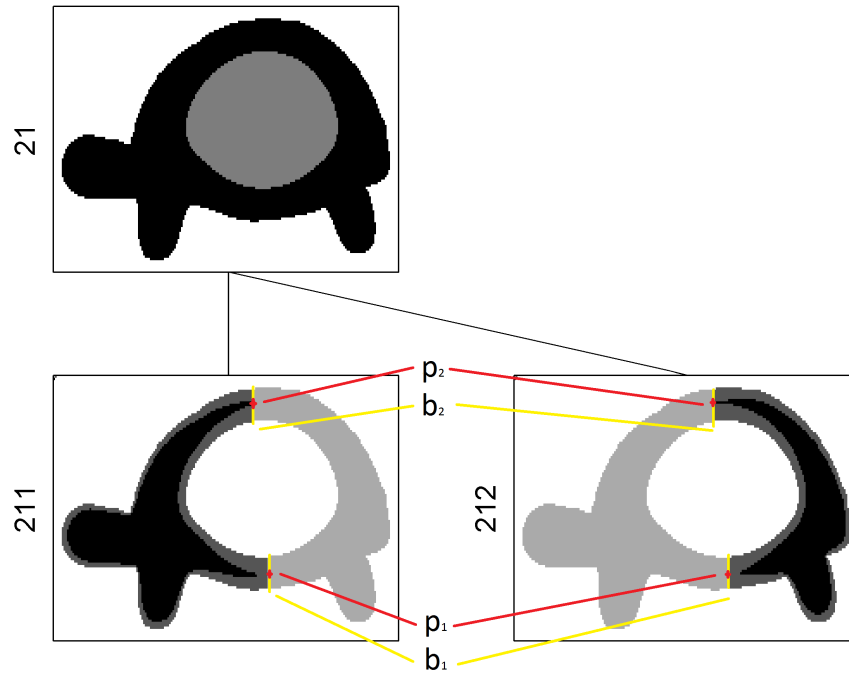


Figure 3.5: A sample run of decomposition algorithm on a ring-like peripheral region where $s_1 > s_2$. p_1, p_2 are saddle points and b_1, b_2 are corresponding watershed boundaries.

central and peripheral regions don't have to be connected, which means root may have more than two children.

An interesting property is that peripheral regions are split into two pieces along the minor axis of the central region. Preliminary part hierarchy trees in Figure 3.1 and 3.6 may be illustrative to observe this phenomenon.

Enclosing contexts of the leaves in the preliminary part hierarchy tree build up the ultimate segmentation of the shape into its parts. Segmentations of the shapes in the database are shown in Figure 3.7.

3.2 Observations

In this section, three main observations associated with the preliminary part hierarchy tree will be discussed.

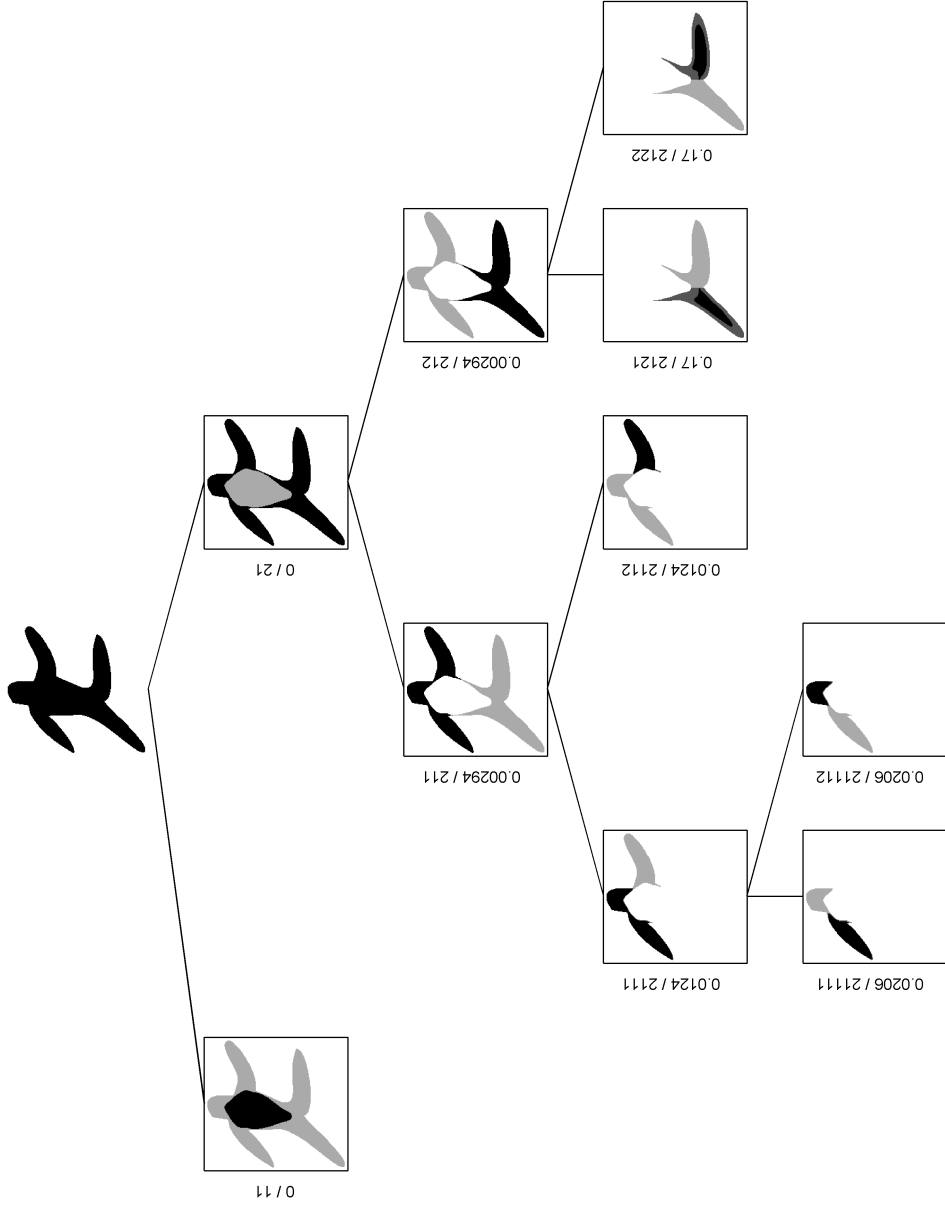


Figure 3.6: The preliminary part hierarchy tree for a human shape. The first number put in the left hand side of the nodes indicates the partitioning level (splitting saddle point value) of the part and the second number is the node identifier.



Figure 3.7: Segmentations of the shapes in the database. Central regions are gray and peripheral parts are colored.

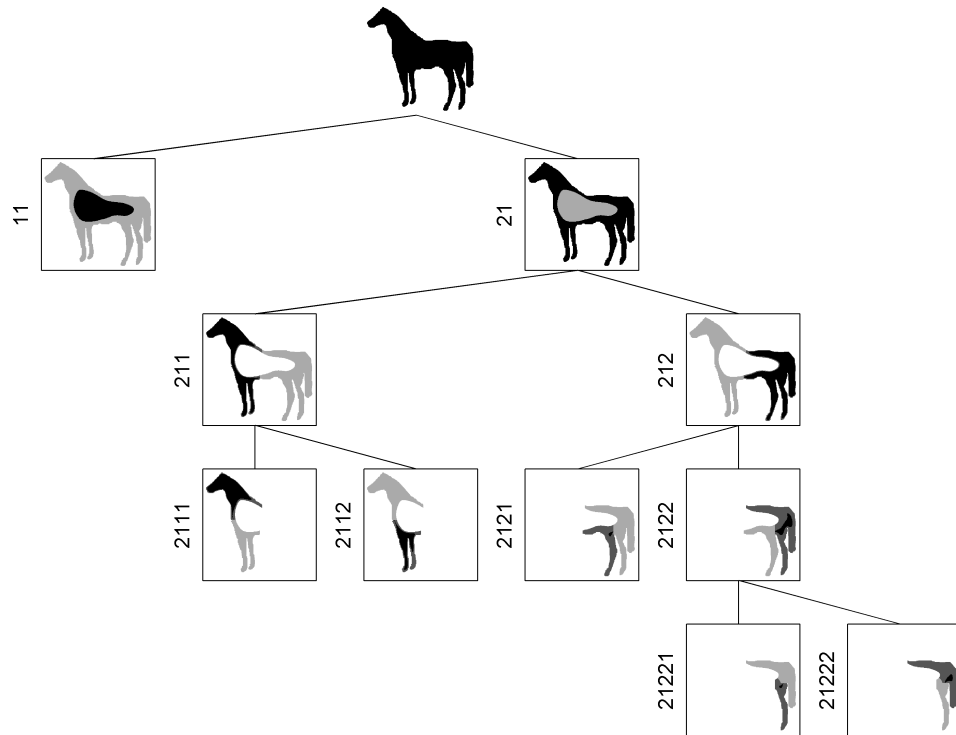


Figure 3.8: Under-partitioning problem. Front legs of the horse (node 2112) are not separated.

3.2.1 Under- and Over-Partitioning

The first phenomenon is observed in the tree which belongs to a horse shape shown in Figure 3.8. At the node 2112, front legs of the horse are no further separated. We call this under-partitioning. The phenomenon arises because there is no saddle point between two structures which are expected to be separated.

In the second tree (in Figure 3.9) extracted from a heart shape, we see that the node 2111 is unintuitively divided into two parts (over-partitioned). In this case there is an extra saddle point separating these two parts.

3.2.2 Instabilities in Node Depths

The depth of a node (except for depth one nodes) merely depends on the order of a saddle point, at which the partitioning occurs that produces the node, in a list of possible saddle points. Due to its discrete nature, the order may vary even with infinitesimal differences

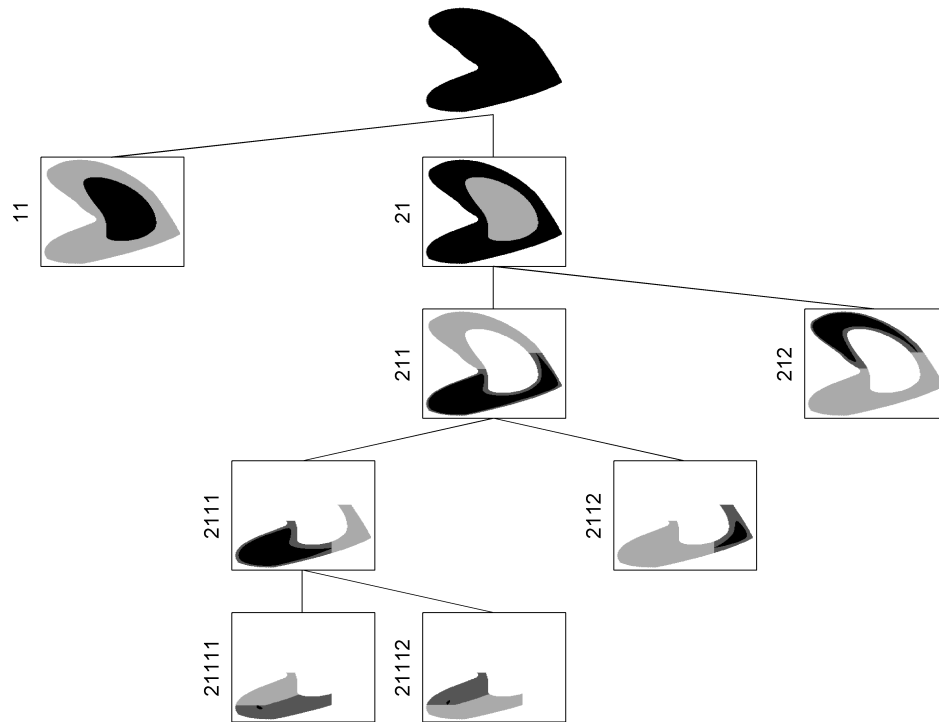


Figure 3.9: Over-partitioning problem. The node 2111 is divided into two parts unintuitively.

between the function values of saddle points. Consequences are discussed on sample trees in the following paragraphs.

A tree that is extracted from a human silhouette was demonstrated in Figure 3.6. The nodes corresponding to two arms (2112,21111) are, unintuitively, at different levels. This is due to the binary splitting procedure that merely takes into account the order of the two saddle points: 1) The saddle point separating the right arm (at 2112) from the part that represents the union of the left arm and the head (2111) and 2) the saddle point that separates the left arm (21111) from the head (21112). In fact, the relative difference between the values of these saddle points is very slight and this suggests that the head and two arms should be simultaneously separated.

Another tree that is extracted for another human silhouette is shown in Figure 3.10. In this case, the tree levels of the right and left arms switch, because of the change in the order of two saddle points. Last two examples suggest that consecutive saddle points that have closer values should be handled carefully, because this implies that a slight change in their value may change the order.

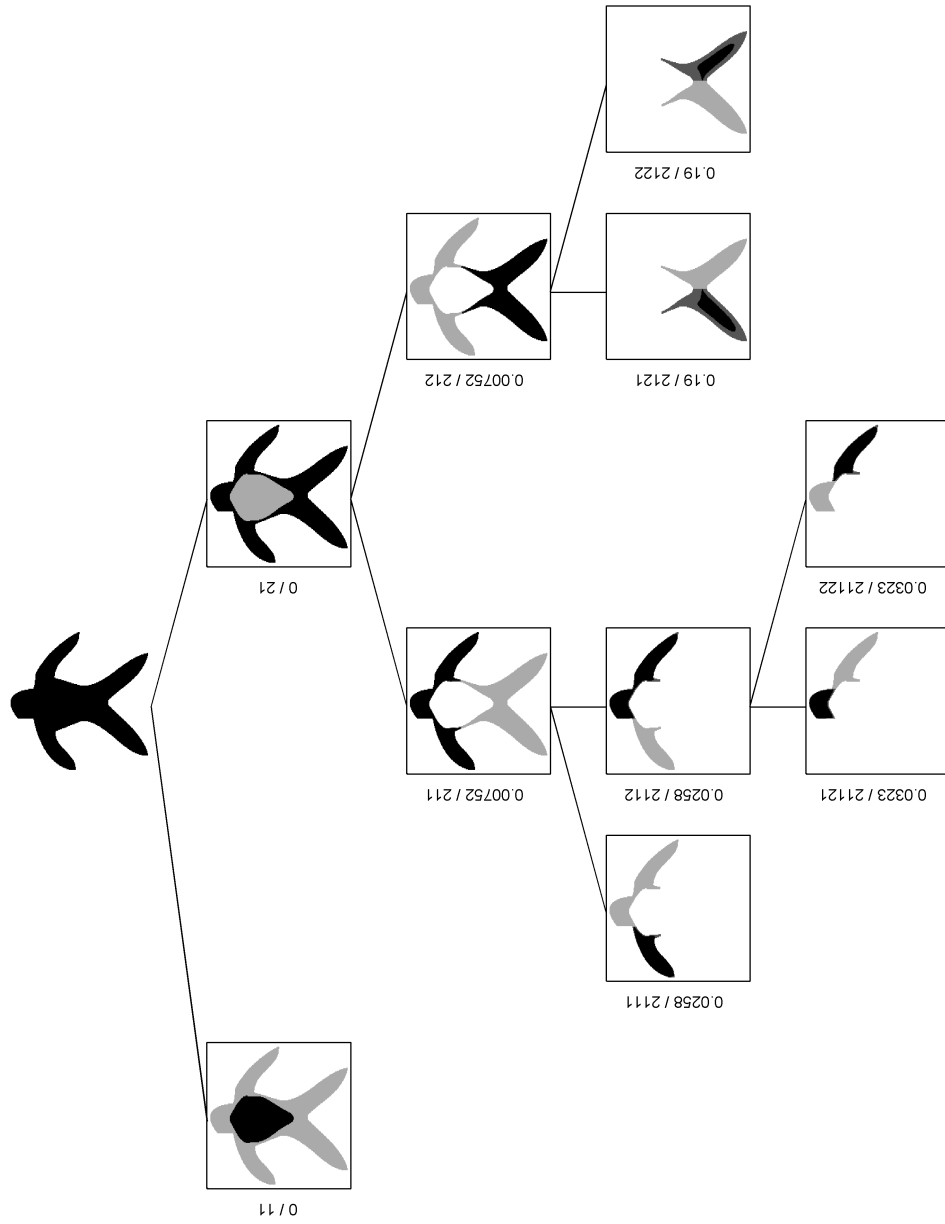


Figure 3.10: The preliminary part hierarchy tree for a different human shape.

A spurious partitioning may also cause inconsistencies in node depths. In a sample case in Figure 3.11, the part at the node 212 is split into two parts by a spurious division. Consequently, the legs and the tail moves one level down in the hierarchy. A reference tree for comparison is provided in Figure 3.12. It is important here to note that spurious divisions occur at a saddle point the value of which is very close to that of the previous saddle point that produces the parent (the node 212 in Figure 3.11). Moreover, important characteristics of the parent is preserved in one of the child nodes (the node 2121 in Figure 3.11). This suggests that the aforementioned child node may replace the parent without too much loss of information.

Some other demonstrations are given in Figure 3.13. In the first row of the figure, the parts are color coded according to their level in the tree. In the second row, the parts are color coded according to the values of saddle points. Consider the star shapes for example. Intuitively five arms should be at the same level. Obviously, saddle point values are more reliable.

Clearly, the value of the saddle point seems to be quite robust in indicating the actual depth of a node relative to other nodes in the tree. However, their order can not be used directly in assigning levels to the nodes robustly, and one should be careful about consecutive saddle points that have relatively closer values. It is mainly this observation that led us to the randomized hierarchy tree, proposed to overcome the inconsistency related with node depths. Details about the randomized hierarchy tree will be given in Chapter 5.

3.3 The Coarse Structure

An alternative central structure, which we call the coarse structure, is proposed in this thesis work. Indeed, the coarse structure is one of the most stable parts of a shape with the central region with respect to visual transformations and boundary noise. The advantage of the coarse structure over the central region is that it is a discriminative structure revealing the identification of the shape up to some degree. As the name reveals, it is a coarse version of the shape which keeps main characteristics of the shape and hides boundary and articulation details. However, the central region is generally just an elongated blob, which hides most of the main characteristics of the shape.

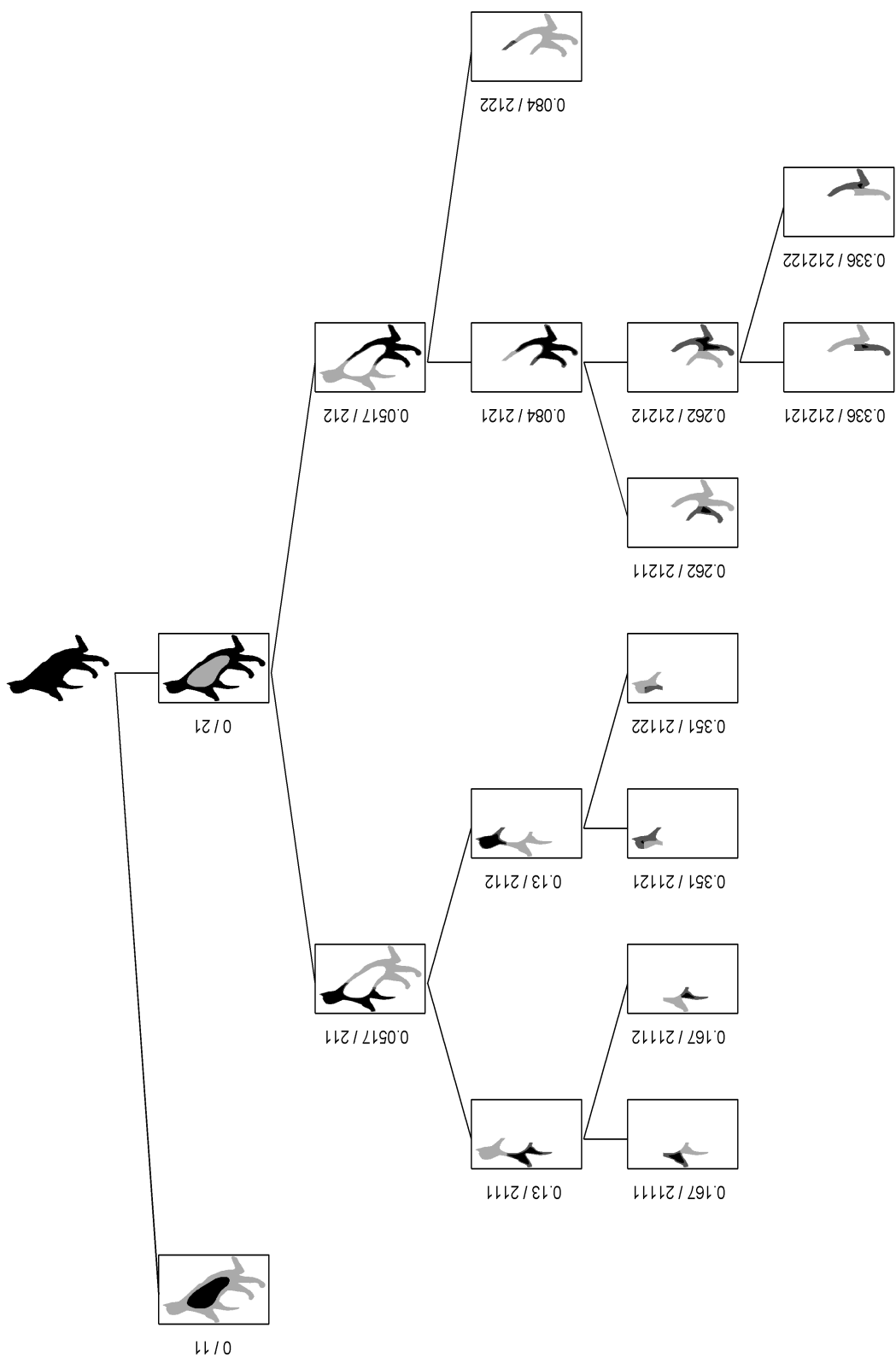


Figure 3.11: The preliminary part hierarchy tree for a cat shape.

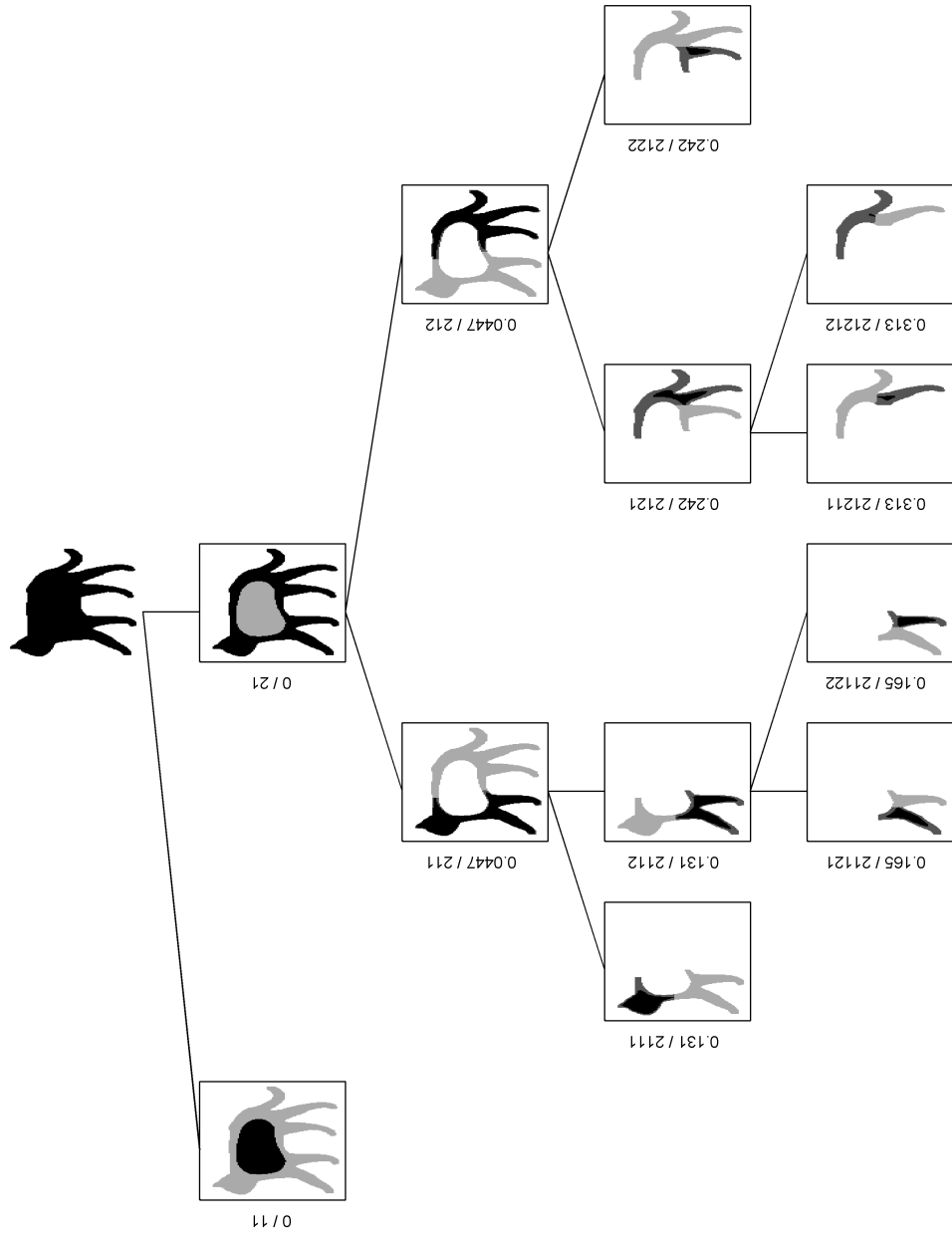


Figure 3.12: The preliminary part hierarchy tree for another cat shape.

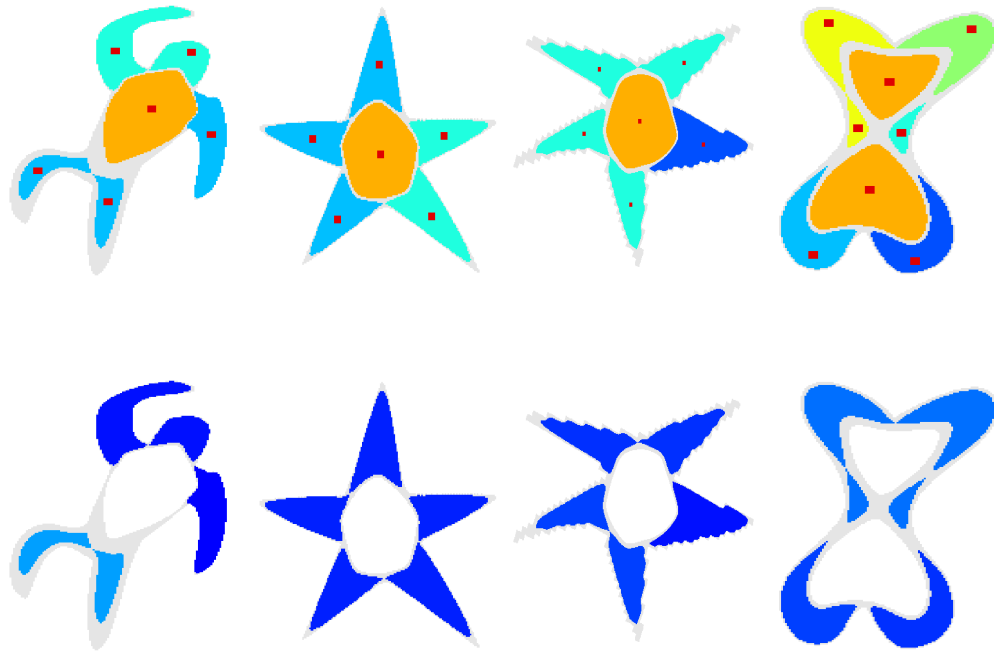


Figure 3.13: Saddle point order vs. saddle point value. (Red dots indicate locations of part centers.)

Remember from Figure 2.3 that extrema of the ω function are also stable reference points for $\alpha = \frac{1}{|\Omega|^2}$. The coarse structure is surrounded by a contour which consists of the points on the ω surface that are local minimum points at least in one direction.

Extraction of the coarse structure is as straightforward as that of central region. In fact, it is a watershed region of $-\omega$ (negated surface) which overlaps with the central region. Demonstrations are given in Figure 3.14.

An interesting property of the coarse structure is that while it is excluding ribbon-like articulations (limbs), it tends to include blob-like protrusions. This phenomenon is illustrated in Figure 3.15.

Extraction of such a structure is very promising in the sense that it can provide a way of extracting noise free shape features. In this work, the coarse structure is utilized in extracting global skeletons; particularly in separation of part skeletons from main body and in construction of global coordinate frames.

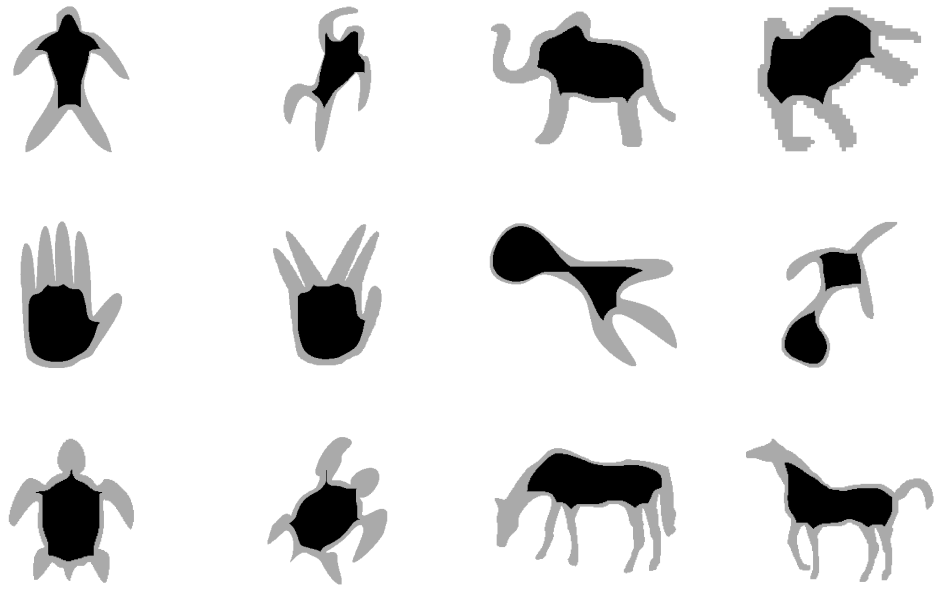


Figure 3.14: Coarse structures (black regions) of some shapes in the database.



Figure 3.15: Ribbon-like vs. blob-like parts. The blob-like head is mostly included by the coarse structure, while ribbon-like limbs (arms and legs) are excluded.

CHAPTER 4

GLOBAL SKELETON

This chapter describes the extraction method of global skeleton. Firstly, all the critical points of the ω function (in the sense of Tari, Shah and Pien [11]) are detected. The critical points falling inside the coarse structure (as defined in Section 3.3) is excluded considering that the coarse structure is better represented as a region than by axis. Secondly, an algorithm is employed which is developed for extracting major skeletons of parts (that track evolution of prominent protrusions). Finally, a global coordinate frame, which is to be used in defining spatial relations, is attached.

4.1 Detection of Skeleton Points

In Chapter 2, the connection of the distance function of Tari [9, 10] to the distance function in [11] was emphasized and application of the skeleton extraction method in [11] was proposed exploiting this connection.

TSP method defines a skeleton point as a maximum curvature point along a level curve. Since its computation is less sensitive to noise, they prefer to detect points of minimum $|\nabla\omega|$ along level curves which correspond approximately to the maximum curvature points.

Skeleton (K) is the set of zero-crossings of $d|\nabla\omega|/ds$ where s denotes arc-length along the level curves.

$$\frac{d|\nabla\omega|}{ds} = \omega_{\eta\xi}$$
$$\omega_{\eta\xi} = \frac{(\omega_y^2 - \omega_x^2)\omega_{xy} - \omega_x\omega_y(\omega_{yy} - \omega_{xx})}{|\nabla\omega|^2} \quad (4.1)$$

where η is the direction of the inward normal and ξ is the direction perpendicular to η .

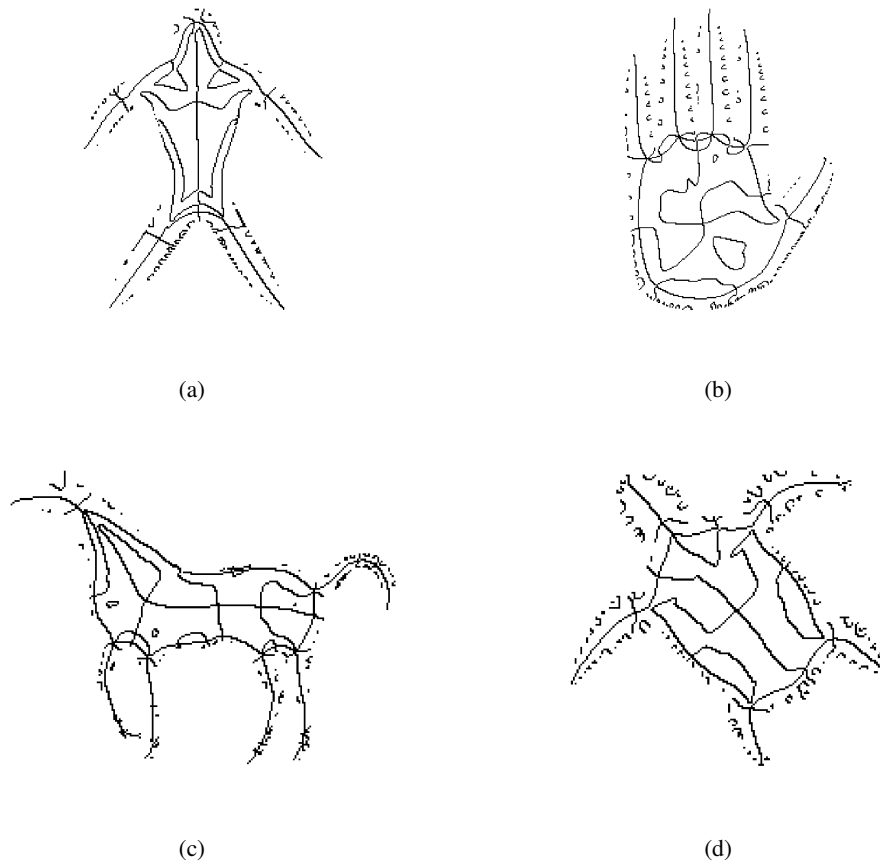


Figure 4.1: Sample skeletons (K) extracted from (a) human, (b) hand, (c) horse, and (d) turtle shapes

In Figure 4.1, sample skeletons are depicted. The skeleton may seem very crowded especially in the ligature zones. Nevertheless, peripheral parts have their skeletons clearly separated from the ligature zones. When examined carefully, it will be noted that the skeleton develops a closed contour (yellow contour in Figure 4.2) which appears to be a coarse form of the actual shape boundary and which is the thing that makes clear the distinction between ligature and part skeletons. This is actually the boundary of the coarse structure previewed already in Chapter 3.

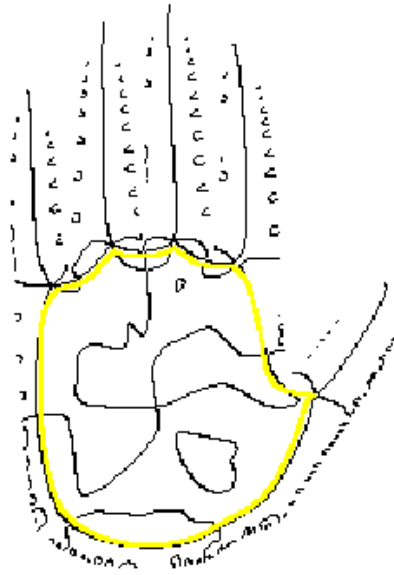


Figure 4.2: Coarse contour (yellow) embedded on the skeleton of the hand shape.

4.2 Building Global Skeleton

Building process includes extracting noise free skeleton branches and constructing a coordinate frame using principle axes of the coarse structure.

4.2.1 Extraction of Skeleton Branches

A skeleton branch can be defined as a grouping of connected skeleton points which track evolution of a prominent protrusion, and which are included by the enclosing context of a leaf part, but excluded by the coarse structure of the shape. This definition is illustrated in Figure 4.3.

The set of skeleton points (K) includes skeleton points which track evolution of protrusions and also skeleton points which track evolution of indentations. Former is the set of skeleton points $K+$ where $d^2 |\nabla\omega| / d s^2$ is positive and the latter is the set of skeleton points $K-$ where $d^2 |\nabla\omega| / d s^2$ is negative. This test can be computed using the formula given in the work [11]:

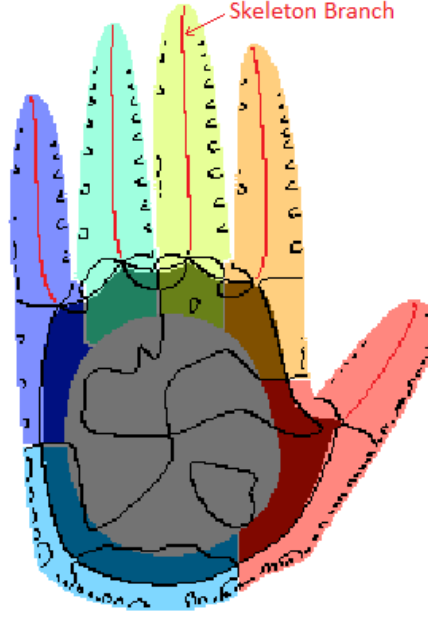


Figure 4.3: Skeleton branches (red lines) as groupings of connected skeleton points which track evolution of prominent protrusions. Skeleton points are included by the enclosing contexts (bright colored regions), but excluded by the coarse structure (darker region on the shape center).

$$d^2 |\nabla\omega| / d s^2 = \omega_{\eta\xi\xi} + \frac{\omega_{\xi\xi} (\omega_{\xi\xi} - \omega_{\eta\eta})}{|\nabla\omega|}, \quad (4.2)$$

where

$$\begin{aligned} \omega_{\xi\xi} &= \frac{\omega_y^2 \omega_{xx} - 2\omega_x \omega_y \omega_{xy} + \omega_x^2 \omega_{yy}}{|\nabla\omega|^2} \\ \omega_{\eta\eta} &= \frac{\omega_x^2 \omega_{xx} - 2\omega_x \omega_y \omega_{xy} + \omega_y^2 \omega_{yy}}{|\nabla\omega|^2} \\ \omega_{\eta\xi\xi} &= \frac{\omega_x \omega_y^2 \omega_{xxx} + \omega_y (\omega_y^2 - 2\omega_x^2) \omega_{xxy} + \omega_x (\omega_x^2 - 2\omega_y^2) \omega_{xyy} + \omega_x^2 \omega_y \omega_{yyy}}{|\nabla\omega|^3}. \end{aligned} \quad (4.3)$$

Obviously, the second derivative test is computationally not straightforward and it requires high resolution to give reliable results. Because of this reason, a skeleton tracing method is proposed instead of relying on the above formulas and grouping the points in the set $K+$. In the method, all we did is tracing paths made by connected skeleton points in the set K within an enclosing context of a part. An important constraint is that most of the skeleton points should be in the set $K+$.

The algorithm for extraction of skeleton branches is described below:

- For each leaf part in the preliminary part hierarchy tree do:
 - Extract the skeleton points (of the set K) that are located within the enclosing context of the part but outside of the coarse structure.
 - Determine initial groupings by labeling connected skeleton points.
 - For each group do:
 - * Find the nearest point (p_{ref}) to the part center (p_{center}) on the boundary of the coarse structure.
 - * Find the nearest point (p_{start}) of the group to p_{ref} .
 - * Determine the points p_{end}^i ($i = 1..n$) that are the end points of the maximally weighted shortest paths starting from the point p_{start} . Here, the weight of a path is measured as the product of the number of included skeleton points in K and the number of included skeleton points in $K+$. With the help of this weight, our algorithm prefers paths that include more points in $K+$.
 - * For each p_{end}^i starting with the furthest one (according to the weighted distance) do:
 - Back trace the path in direction of decreasing distance from p_{end}^i until the path ends.
 - Make the group of visited points a candidate branch ($candidate^i$).
 - Remove visited points to ensure that branch points do not overlap.
 - * Pruning: Qualify a branch $candidate^i$ as a skeleton branch, if
 - it represents a protrusion,
 - and it is prominent.

4.2.1.1 Pruning

The last step of the algorithm is a pruning procedure that determines desired skeleton branches among candidates. It involves two sub-procedures. First one is for elimination of indentation induced branches (whose points are in $K-$), and the second one is for elimination of noise induced branches.

Elimination of Indentation Induced Branches

The procedure that eliminates indentation induced branches is as follows:

- Eliminate a candidate branch, if
 - more than 45% of the path consists of skeleton points in K ,
 - or the length of a connected component that consists of skeleton points in K exceed 30% of the path length.

Elimination of Noise Induced Branches

There exist many branches that are induced by boundary noise and that should be pruned. Based on the idea that noise induced branches do not survive much in the interior of the shape domain and hence they are short relative to the prominent branches, Aslan and Tari [2] employed a length threshold for pruning. A shortcoming of this method is that it prunes some prominent branches that are shorter than the threshold.

In this work, an alternative method is proposed to measure prominence of a skeleton branch. The idea is deduced from the fact that prominent protrusions, which are ribbon-like structures, reveal higher curvature of level curves on the ridge points together with lower gradient magnitudes. In the case of less prominent (blob-like) protrusions such as boundary noise, we observe lower curvature and higher gradient. Gradient magnitudes on the skeleton points of a noisy star shape are shown in Figure 4.4-a.

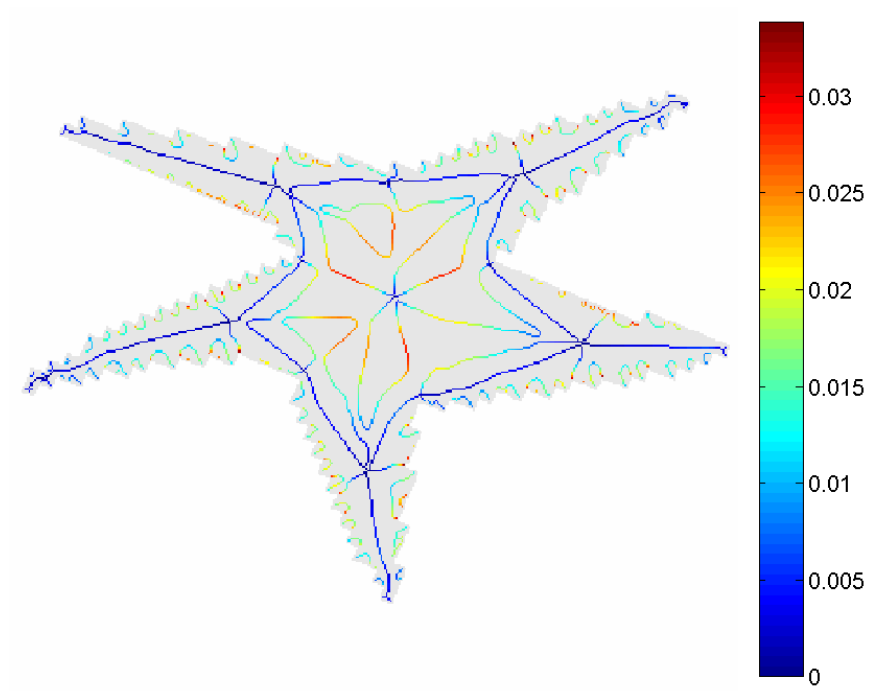
Based on this observation, our method measures saliency $sal(p)$ for all skeleton points p . The measure is inversely proportional to the gradient magnitude ($|\nabla\omega|$) on p and in the range of $(0, 1]$. Here, since the gradient magnitude is in the range $[0, M)$ (M is any positive number), we use the following formula to map gradient magnitudes to point saliency measure:

$$sal(p) = e^{-c|\nabla\omega_p|} \quad (4.4)$$

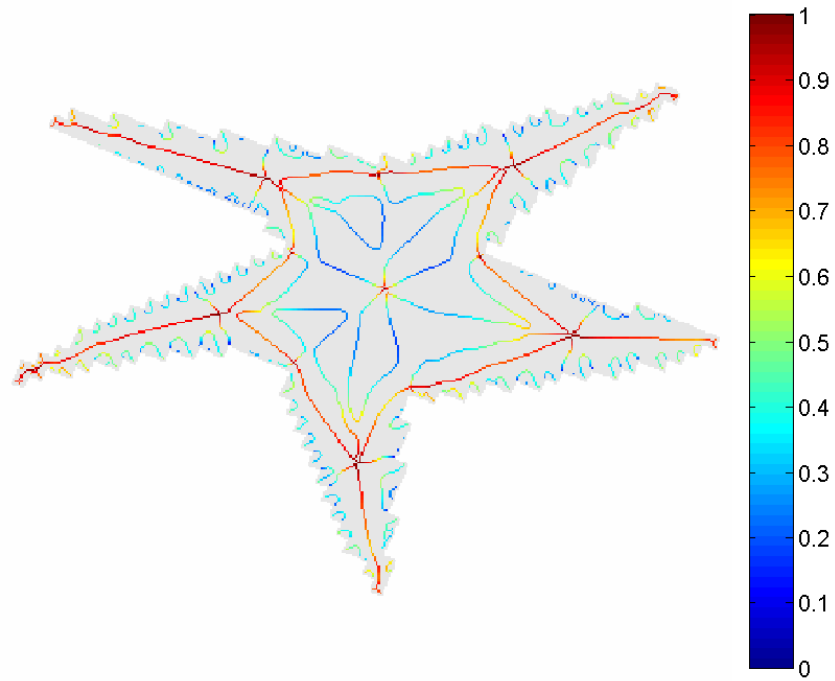
where

$$c = -\frac{\ln 0.5}{\frac{1}{N} \sum_{i=1}^N |\nabla\omega_{p_i}|} \quad (4.5)$$

and p_i is the i^{th} skeleton point and N is the number of all skeleton points. Saliency measurements for the skeleton points of the star shape are shown in Figure 4.4-b.



(a)



(b)

Figure 4.4: Illustration of (a) gradient magnitude, (b) saliency measurements for skeleton points of a star shape with boundary noise

The next step is determining saliency of the whole candidate branch. This is achieved by averaging saliency values of included skeleton points:

$$sal(candidate^i) = \frac{1}{N^i} \sum sal(p_j) \quad (4.6)$$

where $p_j \in candidate^i$ and N^i is the number of skeleton points included by the $candidate^i$.

A possible method to eliminate noise induced candidate branches and qualify only those which represent prominent protrusions is as follows:

- Eliminate the branch $candidate^i$, if
 - $sal(candidate^i) < 0.5$

The above method solves most of the problem. However, saliency of some very small noise induced branches may exceed the threshold. Such problems are solved by employing a (saliency weighted) length threshold which is smaller than that required in the case without saliency threshold. A smaller length threshold then reduces the risk of eliminating prominent branches. Saliency weighted length of a branch can be computed as follows:

$$swlen(candidate^i) = \sum sal(p_j) \quad (4.7)$$

where $p_j \in candidate^i$.

Finally, the procedure becomes as follows:

- Eliminate the branch $candidate^i$, if
 - $sal(candidate^i) < 0.5$,
 - and $swlen(candidate^i)$ is less than 1% of the boundary length of the coarse structure.

The algorithm is illustrated in Figure 4.5 in the context of a human shape.

The reason why we allow more than one branches per leaf part is the possibility that a leaf part may include more than one prominent branches in the following cases:

1. The part is under-partitioned (see Chapter 3 for details). Therefore, it may consist of more than one parts each having its own skeleton branch.

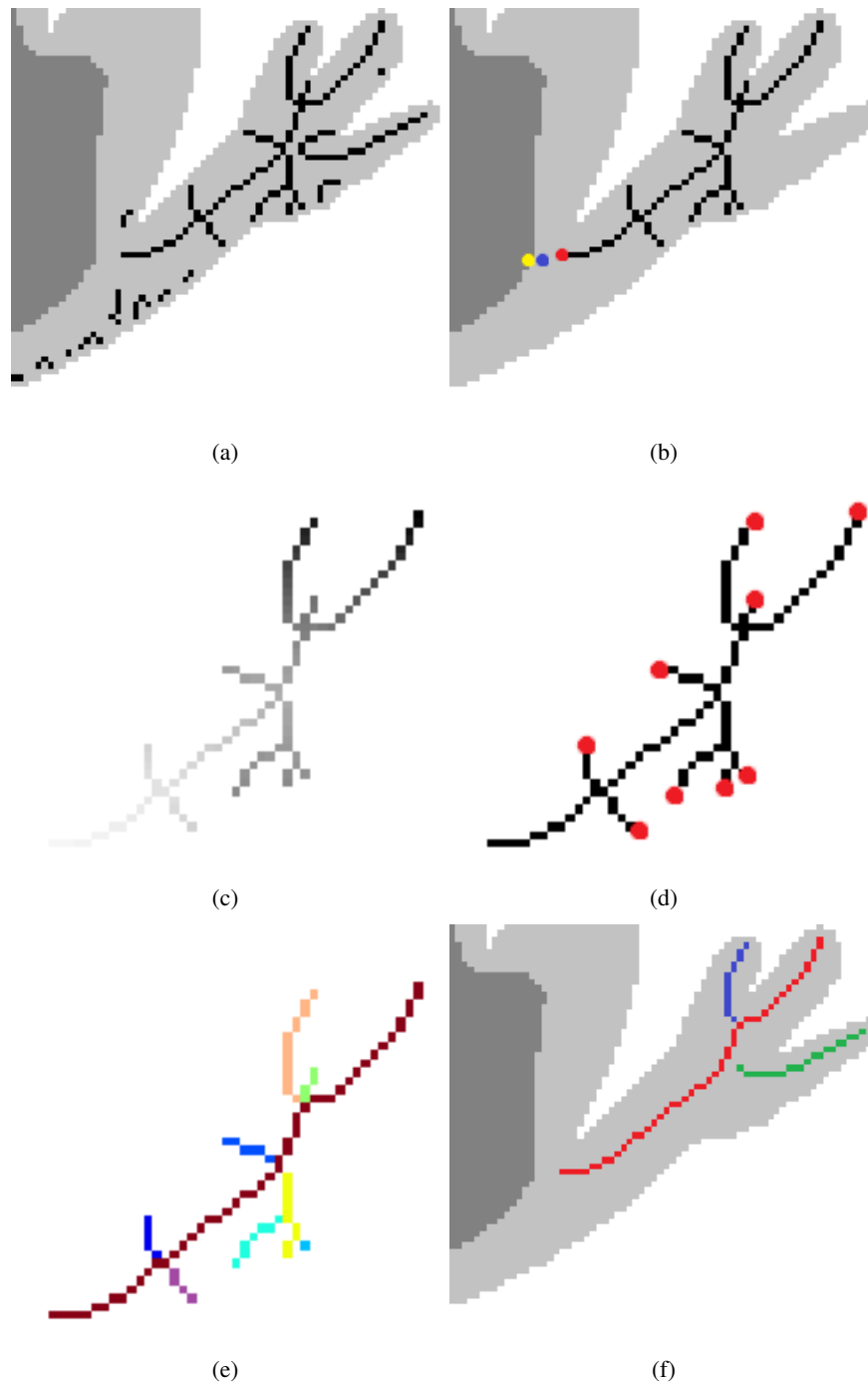


Figure 4.5: A sample run of the skeleton extraction algorithm for the right leg of a frog shape. (a) Skeleton points placed within the enclosing context and outside of the coarse structure. (b) One of the initial groupings (group of connected skeleton points) shown with critical points: p_{center} (blue), p_{ref} (yellow) and p_{start} (red). (c) Weighted distances of the skeleton points in the initial grouping. (d) End points. (e) Candidate skeleton branches extracted from the initial grouping. (f) Extracted skeleton branches after processing all of the initial groupings and selecting salient branches that represent prominent protrusions.

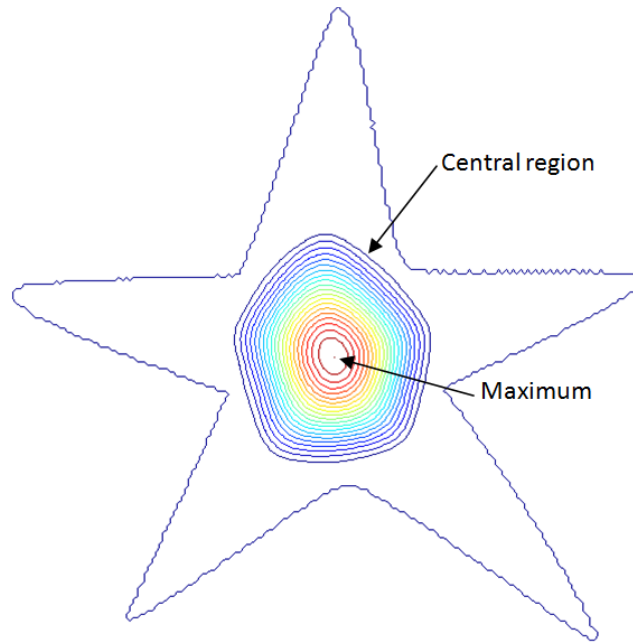


Figure 4.6: Star shape as a sample single-body shape. Shape boundary (outer contour) and the level contours (inner contours) of the central region are depicted.

2. The part may develop a minor protrusion which is not detected as a separate part, such as an ear within the context of a head. Skeletons of such protrusions, which are qualified as prominent using the aforementioned length threshold, are also included.

4.2.2 Constructing the Coordinate Frame

Describing the spatial organization of primitives is quite instrumental for pattern matching purposes, and used in a great number of applications. In the presence of robust reference points, an easy way is setting up a coordinate system and determining the relative placement of these points with respect to the constructed system.

In this work, shapes in the database are categorized into three types according to their structural properties and different coordinate frames are constructed for different types.

First category, which we call *single-body*, consists of the shapes that have one central region with one global maximum point on the ω surface. Single-body shapes are considered as the shapes having single main body (Figure 4.6).

The second category, which we call *multiple-body*, consists of the shapes in the database

that have two separate central regions. This kind of shapes have two main bodies connected to each other with a narrow neck. A dog bone shape is a good example of multiple-body category (Figure 4.7).

The last category, which we call *intermediate-form*, is a transition category between the two main categories single and multiple-body. Its members represent intermediate-forms during possible deformations of shapes from one main category to another. Intermediate-form shapes have only one central region (like single-body shapes) but two (local) maximum points in the central region. This means that there are two central sub-regions (two watershed regions of $-\omega$ within the central region) which constitute the unified central region. Two example cases are as follows:

1. A shape with intuitively two main bodies connected with a relatively wide and short neck may develop one non-separated central region with two local maximum points corresponding to two main bodies. A specific example to this case is another dog bone shape in the database (Figure 4.8).
2. Another example case emerges in horse shapes in the database. As shown in Figure 4.9, one of the horse shapes develops one central region with one global maximum, whereas another one develops one central region with two local maxima. In this case, an intuitive determination of the actual number of main bodies is also very hard, because there is no prominent neck that separates potential two main bodies.

For single-body shapes, a global polar coordinate frame is designed. The polar axis is selected as the principle axis of the coarse structure which is an excellent tool for extraction of such a sensitive feature. Actually, there are two polar axes each has the same origin which is the weighted centroid of the central region, but each has its own ray. The rays have the same orientation parallel to the principle axis but reverse directions.

The radial (r) and the angular (θ) coordinates of a skeleton branch are measured with respect to its starting point p_{start} . Firstly, the position vector \vec{s} is defined that connects the origin to p_{start} . Then, the radial coordinate r becomes the magnitude $|\vec{s}|$, and the angular coordinate θ becomes the counterclockwise angle between the vector \vec{s} and a polar axis. Since there are two polar axes, two corresponding angular coordinates θ_1 and θ_2 are measured. The coordinate system is illustrated on a star shape in Figure 4.10.

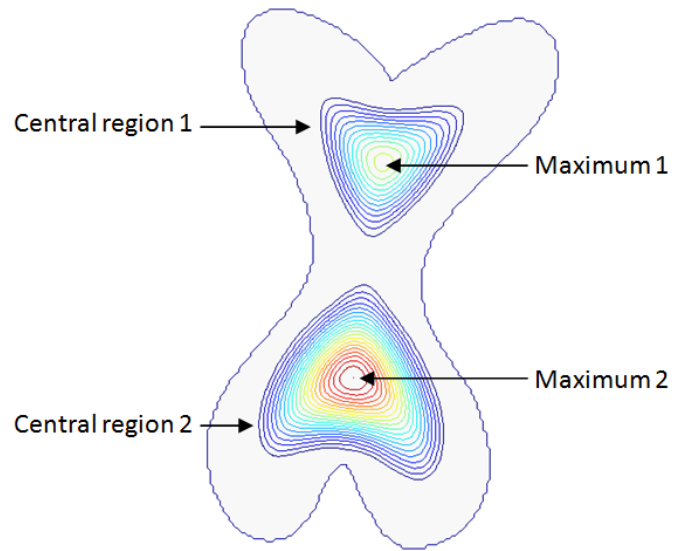


Figure 4.7: A dog bone shape as a sample multiple-body shape. Shape boundary (outer contour) and the level contours (inner contours) of the two central regions are depicted.

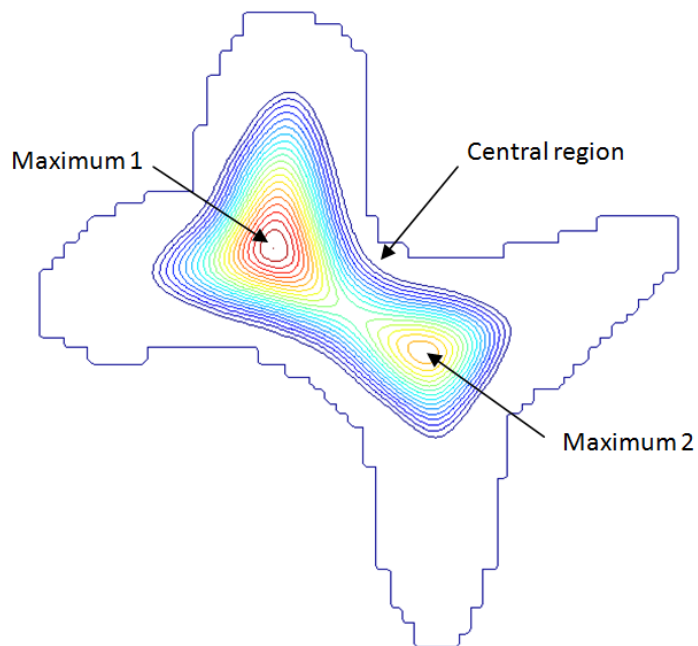


Figure 4.8: Another dog bone shape as a sample intermediate-form shape. Shape boundary (outer contour) and the level contours (inner contours) of the central region are depicted.

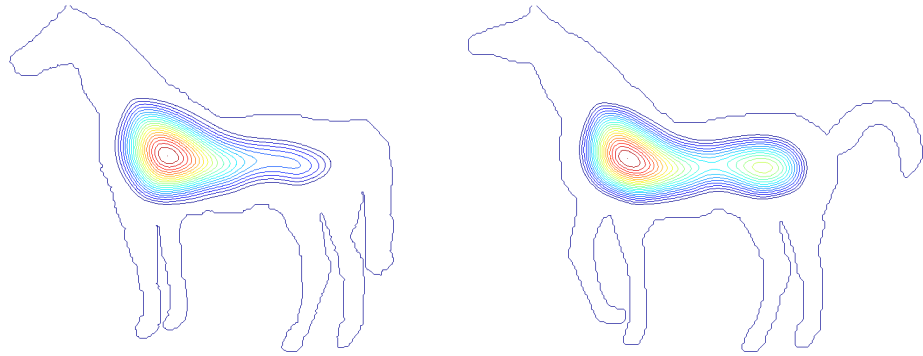


Figure 4.9: Two horse shapes of different types.

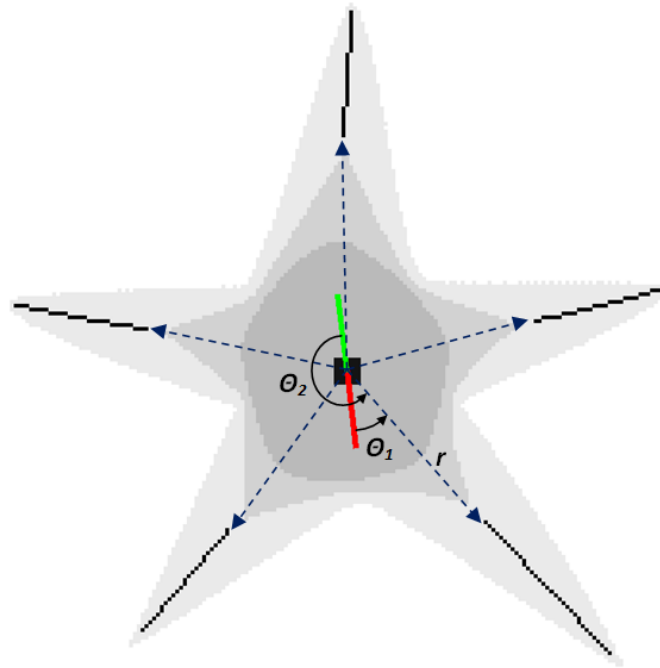


Figure 4.10: Illustration of the constructed polar coordinate system for a single-body shape. Red and green lines are polar axes originating from the center of the central region (black square). The radial coordinate r of a skeleton branch is the length of the position vector \vec{s} (dashed arrows). The angular coordinates θ_1 and θ_2 are the counterclockwise angles between the vector \vec{s} and two polar axes.

Using one global coordinate system is not a reliable way of describing spatial organization in multiple-body shapes. Generally, the spatial organization of peripheral parts in the context of a main body (central region) is considered to be robust up to a significant degree. However, the relative positions of main bodies tend to vary with the deformations in the connecting neck. For example, the position of a head relative to the human body may easily vary with the movements of the neck, while the positions of the ears relative to the head remains fixed.

This observation suggests that the spatial relations of the peripheral parts should be described within the context of the corresponding central region. Therefore, two separate coordinate frames, each of which describes the spatial relations in the appropriate context, are defined. The two origins are located at the weighted centroids of the two central regions. In this case, polar axis is determined in a different manner. Firstly, a reference point is extracted, which is the weighted centroid of the binding part(s) which connects two central regions. Then, two rays are sent from the reference point to the weighted centroids of two central regions. Finally, the directions of the rays become the directions of the polar axes of the corresponding coordinate frames. In this case, there exists only one polar axis per coordinate frame, since the directions are now distinctive with the help of the reference point. Construction of coordinate frames for multiple-body shapes is illustrated in Figure 4.11. Polar coordinates are determined in the same way as in single-body shapes except that measurements are made separately within the contexts of central regions and only one angular coordinate is obtained per skeleton branch.

It is suitable to represent intermediate-form shapes multiply in order to be able to match them to both single and multiple-body shapes. In the first representation for matching to single-body shapes, one global coordinate frame is constructed using the method developed for single-body shapes. The second representation of intermediate-form shapes is analogous to the representation of multiple-body shapes. In this case, the two coordinate frames are built on the weighted centers of the two sub-central regions and the reference point is selected as the saddle point where the central region is partitioned into two sub-central parts.

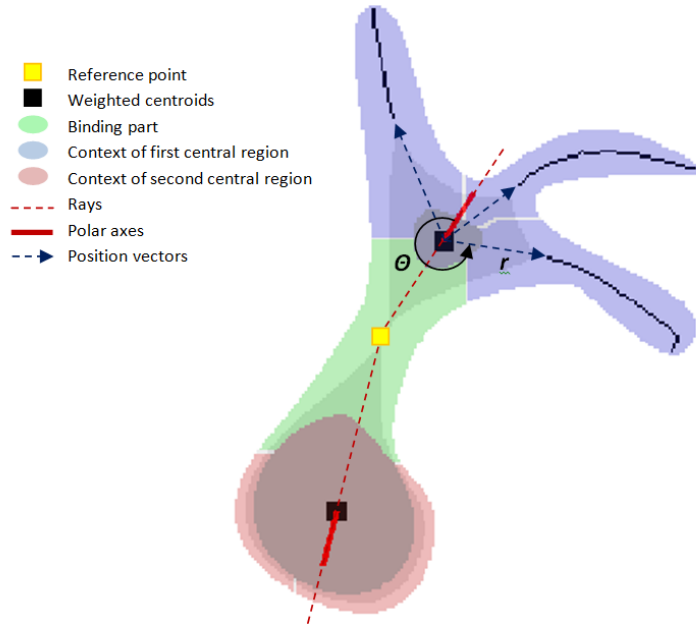


Figure 4.11: Illustration of the constructed local coordinate systems for a multiple-body shape.

4.3 Discussion

The name global skeleton is given to emphasize the distinctive property of the proposed medial descriptor that it involves not only skeletal branches like conventional medial descriptors, but also a medial region, i.e. the coarse structure. The coarse structure is utilized in discrimination of skeleton branches from inconsistent ligature zones and in extracting reference axes for construction of robust coordinate frames (for type-one shapes). Resulting global skeletons for database shapes are depicted in Appendix A. Skeleton branches are black colored lines, whereas the dark gray region is the central region and mid-gray region is the coarse structure with its boundary emphasized with pink color. Finally, red lines are the polar axes with the origin indicated using a black square. Multiple representations of type-two shapes are shown one under the other.

4.3.1 Comparison with Aslan and Tarı [2]

Global skeleton is best comparable to the disconnected skeleton work of Aslan and Tarı [2], since both are related from many aspects. First of all, the main tools (the ω function for global

skeleton and the edge strength function for disconnected skeleton) are strongly connected as shown in Chapter 2. Then, the two works describe spatial relations similarly by constructing object centered polar coordinate frames. Finally, both of the works define structures which are coarse forms of the shapes. These structures are stated to be least deformable under visual transformations of the shapes such as translation, rotation and scaling. The two works are compared in the following subsections.

4.3.1.1 The ω Function vs. the Edge Strength Function

The ω function has an important advantage over the edge strength function. The advantage is the ease of choosing the parameter α , which is already mentioned in Chapter 2. In contrast to this, the choice of the smoothing parameter ρ that results in reliable edge strength functions for stable shape descriptions is hard. A too small value causes insufficient diffusion, that is the inner points are not affected from the diffusion. As a consequence, the desired shape analysis can not be done. Another consideration in determination of the level of smoothing is that the shapes should be represented at the same level of details. The strategy that they use to find a reliable solution is starting with a small value of ρ and increasing it until reaching an edge strength function with a single extremum whose absolute value is above a predefined threshold.

4.3.1.2 Describing Spatial Relations

The two works describe spatial relations similarly by constructing object centered polar coordinate frames (Figure 4.12). Reference axes are determined differently in the work of Aslan and Tarı [2] by utilizing directions of major negative skeleton branches (whose elements are in the set K^-) which generally connects critical points (extremum and saddle points) of the edge strength surface to the prominent indentations of shape boundary. Here, stability is negatively affected from the sensitivity in the ligature zones. In order to overcome the resulting ambiguities they employ multiple-representations for problematic shapes. In contrast, our method seems to be robust in such cases, since it doesn't involve handling skeleton sections that fall in ligature zones. However, our method may also become sensitive for the shapes that have two principal axes that are almost equally significant, i.e. they have close eigenvalues. The solution might be again multiple-representation using two different principal axes.

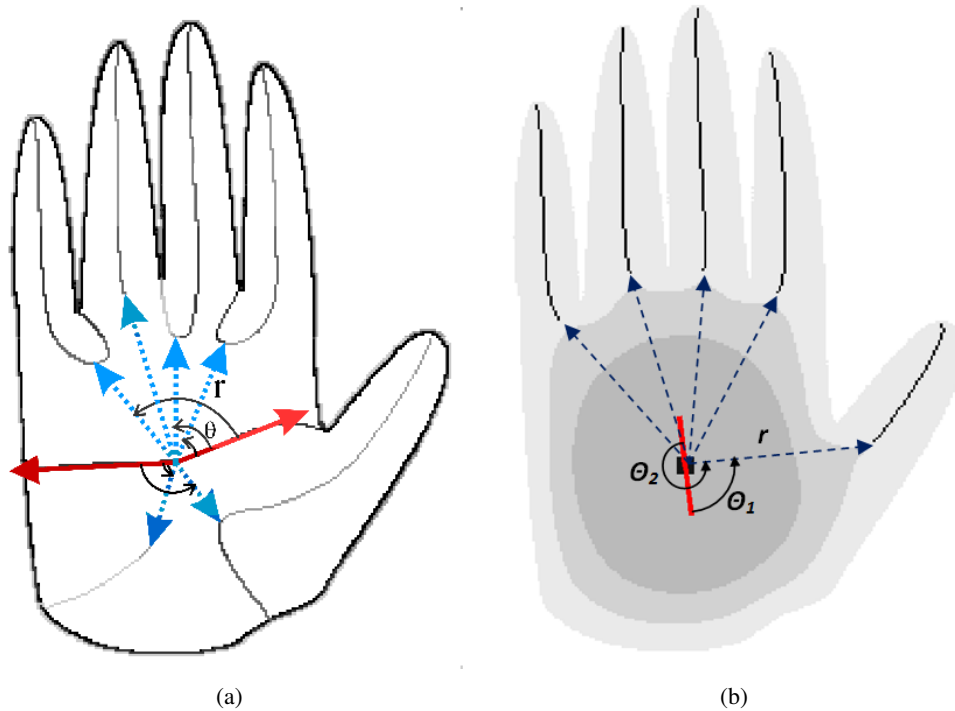


Figure 4.12: Comparison of coordinate frames. a) (Taken from [1]) Illustration of the constructed coordinate frame in the work of Aslan and Tarı [2] for a hand shape. b) Illustration of our coordinate frame on the same hand shape.

Another difference is noticed in handling multiple-body shapes. In the work [2], a single coordinate system is constructed which has the origin on the saddle point that separates two blobs. In a previous section (Section 4.2.2), we had emphasized that using one global coordinate system for multiple-body shapes may fail due to the deformations in the connecting neck, and considering this fact we choose to use separate coordinate systems centered on object's main bodies in order to describe spatial relations of parts that are in their own contexts.

4.3.1.3 Coarse Forms of Shapes

The coarsest structure in the work of Aslan and Tarı [2] is defined as the state of the shape during the evolution, when all minor branches have terminated at junction points. In Figure 4.13 taken from [1], major (thick) and minor (thin) skeleton branches of a hand shape are depicted in part (a), and two different states of the shape are given: (1) after the termination of three fingers, and (2) after the termination of all the minor branches. The state (2) is stated as the coarsest form. For a comparison, in Figure 4.14, the global skeleton of the same hand shape is depicted with the central region (darkest gray) and our coarse structure (mid-gray).

Obviously, the coarsest form of the hand shape (the state (2)) in [2] mostly resembles the central region extracted in our work. The advantage of the coarse structure over the central region (and thereby over the coarsest structure in [2]) was discussed in Section 3.3. The coarse structure, like the central region, is stable under visual transformations. Additionally, it keeps discriminative shape properties. However, the coarsest structure in [2], similar to the central region, is just an elongated blob.

4.3.1.4 Topological Categorization of Shapes

Another point of comparison is categorization of the shapes according to their topological properties. Remember that we divide the shapes in the database into three types according to the topological properties of the central regions. The types consist of two main types (single-body and multiple-body) and a transition type (intermediate-form) which provides a smooth transition between the main types. In the other work, there are also two main types of shapes like ours, but without an intermediate type the transition seems to be dangerously sharp. They measure the prominence ratio of the blobs (analogous to main bodies in our terms) during the initial analysis. Then they force the shapes that have a prominence difference greater than a simple threshold (corresponds to a small value of ratio) to have eventually one blob (one global maximum) by applying a sufficient amount of diffusion. The other shapes that have two blobs (two local maxima) with nearly the same prominence are not forced to have one blob. Here, the problem is that very similar shapes that have close prominence ratio values which are on the different sides of the threshold will fall into different categories.

4.3.1.5 Pruning Strategy

In the work of Aslan and Tari [2], the raw length measure of the skeleton branches is used as an indicator of branch prominence. Considering this, they prune branches shorter than a length threshold. As discussed previously in this chapter, using merely raw length measure causes pruning prominent branches erroneously. We developed a different strategy that reduces the risk of pruning prominent branches. We first introduced the saliency measure for skeleton points based on the idea that skeleton points of the prominent branches reveal low gradient magnitudes. The saliency measure became inversely proportional to the gradient magnitude and was in the range (0,1]. Then we measured saliency of the whole branches based on

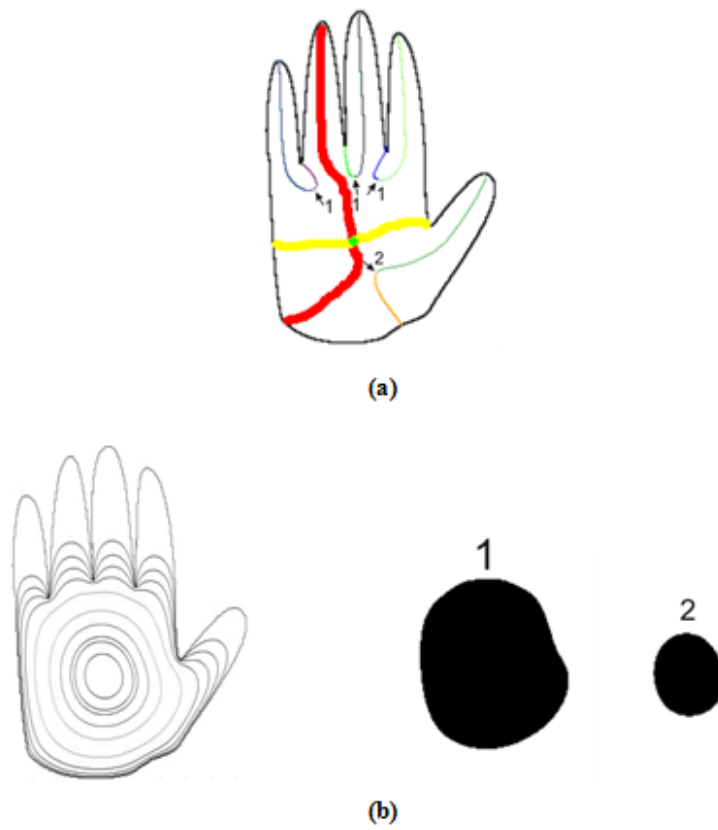


Figure 4.13: (Taken from [1]) a) The skeleton branches of a hand shape in [2]. b) The level curves for the evolving shape boundary and the states of the hand shape at the times the skeleton branches indicated by the numbers in (a) terminated.

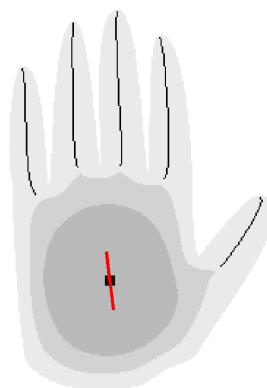


Figure 4.14: Global skeleton representation for the hand shape that is used for illustrations in Figure 4.13

the individual skeleton point saliency values that are included by the branch in two different ways: average and total. Finally, we put two different criteria corresponding to two different measures so that a branch that doesn't satisfy any of the criteria would be pruned.

4.3.1.6 Summary of the Comparison

Global skeleton is compared above to the work of Aslan and Tari [2] from different aspects. Briefly, global skeleton has the following advantages over the other work:

1. Finding a stable solution for the ω function is easier than that of the edge strength function.
2. Extraction of reference axes for representing spatial relations is not affected from skeletal ligature inconsistencies.
3. Context dependent description of spatial relations in multiple-body shapes improves stability against deformations in the necks that connect the main bodies, which results in changes in relative positions of main bodies and eventually in relative positions of the peripheral parts from different contexts.
4. Among the topological shape categories, the intermediate-form provides a safe and smooth transition between the main categories.
5. Our pruning scheme reduces the risk of erroneous pruning of prominent branches.

CHAPTER 5

MATCHING

First part of this chapter is dedicated to randomized hierarchy tree. First a randomization procedure is proposed to obtain samples of the randomized hierarchy tree from the preliminary part hierarchy tree. Then, some illustrative experiments on matching randomized hierarchy trees are discussed. The last part is dedicated to matching global skeletons.

5.1 Randomized Hierarchy Tree

Several instabilities in node depths of the preliminary part hierarchy tree were explained in Chapter 3. Then, our observations led us to conclude:

1. The value of the saddle point seems to be quite robust in indicating the actual depth of a node relative to other nodes in the tree.
2. Direct use of the saddle point order is not reliable in assigning levels to the nodes, and one should be careful about consecutive saddle points that have relatively closer values.

At this point, the question may arise: How can we determine whether two consecutive saddle points are closer? In fact, such a determination results in a discretization of the hierarchical levels which assigns hard explicit depths to the nodes, just like the basic algorithm that extracts the preliminary part hierarchy tree. Here, no doubt, the same instabilities will arise. Therefore, we decided to adopt a probabilistic approach. In this approach, instead of deciding whether the values of consecutive saddle points are closer, we just measure the relative difference between them. Then, this measure is utilized to derive probabilities to introduce randomness in determination of node depths.

5.1.1 Randomization Procedure

Let ϵ be the relative difference between the values of two consecutive saddle points. It is the difference between the saddle point values of the children and their parent divided by the saddle point value of the children and attains a value in the range $0 < \epsilon \leq 1$. On the one hand, a small value of ϵ indicates closer saddle point values. This implies that the local tree structure is not stable, since a slight change in saddle point values may change their order. On the other hand, a large value of ϵ indicates that the consecutive saddle points are well separated and hence the local structure is stable.

Randomization procedure starts with depth two nodes of the preliminary part hierarchy tree and iterates through their children. Two alternative events are considered for each pair of siblings:

1. The pair of siblings maintain their depth (no change in the local tree structure).
2. They inherit the depth of their parent (change in the local tree structure) by replacing their parent and becoming the children of their grandparent.

Let p be the probability of the change in the local tree structure and $1 - p$ the probability of no change. We require that p approaches 1 as ϵ approaches the smallest value ($=0$). Equally, p should approach 0 as ϵ approaches to its largest value ($=1$). To estimate the probability p , the function $e^{-4\epsilon}$ is used. For the largest possible value of ϵ which is 1, there is less than 2% chance for change in the local tree structure ($e^{-4} = 0.018$).

The preliminary part hierarchy tree of a heart shape shown in Figure 5.1 is appropriate for a brief explanation of the procedure. In a non-graphical format, the tree can be expressed as $root \rightarrow (11; 21 \rightarrow (211 \rightarrow (2111; 2112); 212))$. In any preliminary part hierarchy tree, ϵ is equal to 1 for depth two nodes, since saddle point values of depth one nodes are fixed to 0. As a result of this, two depth two nodes 211 and 212 replace their parent and become children of their grandparent (the root) with probability $p = 0.018$, whereas the local tree structure is preserved with probability $(1 - p) = 0.982$. In the next hierarchy level, $\epsilon = 0.208$ for nodes 2111 and 2112. This implies that with probability $q = 0.435$ nodes 2111 and 2112 replace their parent, while with probability $(1 - q) = 0.565$ the local structure is preserved.

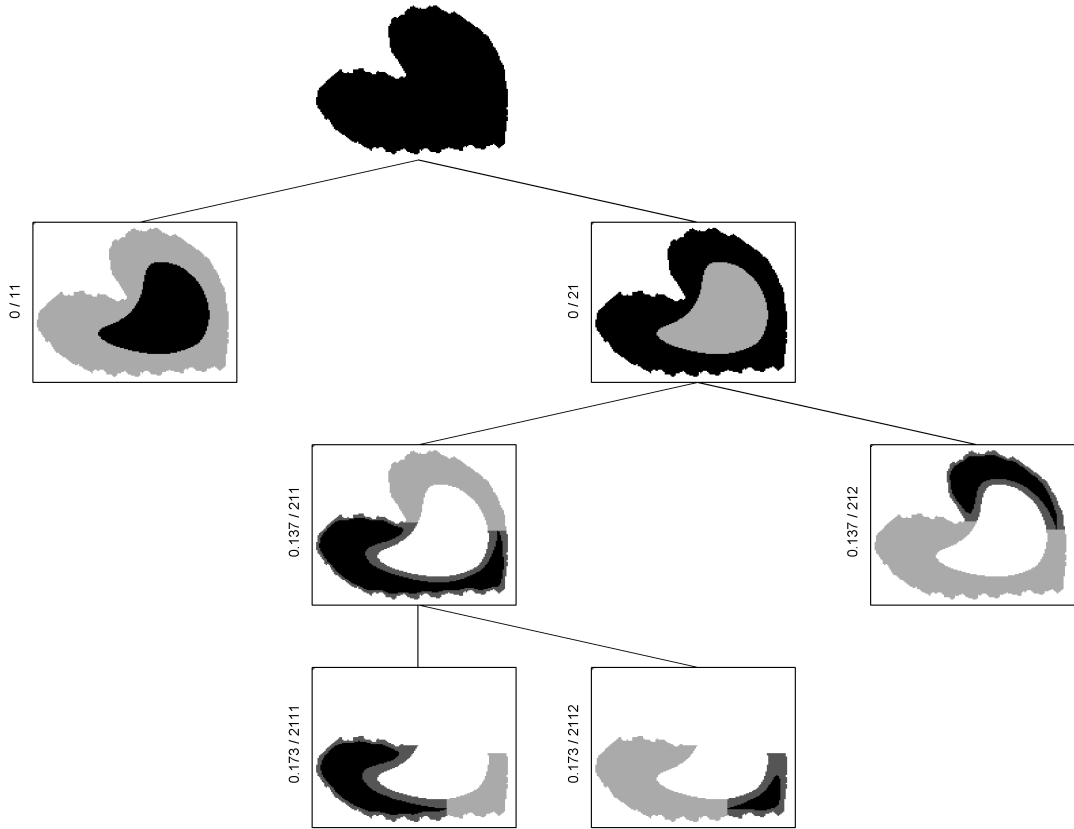


Figure 5.1: The preliminary part hierarchy tree of a heart shape.

Four possible organizations for the given preliminary part hierarchy tree are listed below with corresponding probabilities:

1. The entire structure is preserved with probability $(1 - p)(1 - q) = 0.555$.
2. $root \rightarrow (11; 21 \rightarrow (2111; 2112; 212))$ with probability $(1 - p)q = 0.427$.
3. $root \rightarrow (11; 211 \rightarrow (2111; 2112); 212)$ with probability $p(1 - q) = 0.01$.
4. $root \rightarrow (11; 2111; 2112; 212)$ with probability $pq = 0.007$.

5.1.2 Matching Experiments

Running the randomization procedure, we can obtain samples from a randomized hierarchy tree. Matching two randomized hierarchy trees is formulated as finding the maximum weighted clique in a weighted association graph derived from a pair of samples from the

respective randomized hierarchy trees. For each matching task, 100 pairs are sampled and matched. The match with the highest score is selected.

The approximate algorithm proposed in [8] is used to solve the maximum clique problem. Consider two attributed trees (V_1, E_1, α_1) and (V_2, E_2, α_2) , where (V, E) is the rooted tree and α is a function that assigns a feature (or attribute) vector $\alpha(u)$ to each node $u \in V$. The weighted tree association graph (wTAG) is the weighted graph (V, E, w) , where:

- $V = V_1 \times V_2$,
- any two nodes in the graph (k, m) and (l, n) are considered adjacent if the hierarchical relationship between k and l in the first tree is the same as that between m and n in the second tree,
- and w is a function which assigns a positive weight to each node $(u, v) \in V$.

Any node (u, v) of the wTAG is a possible match between nodes u and v , and the corresponding node weight $w((u, v))$ is a measure of similarity between them. Similarity can be estimated using various features associated with tree nodes. In order to keep illustrations simple and to reveal the power of the randomized hierarchy trees we used two basic features. The first one is the function value of the local extremum within a part. For the parts at the leaf nodes (with a single local extremum) the extremum value is obtained at the associated part center. For such parts, this feature provides a rough estimate of the width. Accordingly for a part in a higher level of hierarchy, the feature should give the width estimate of the widest leaf part, where the former is the ancestor of the leaf part. The second feature is the area of the enclosing context of a part.

Both features need to be normalized to make them invariant under scaling and to prevent, as possible, any feature bias in measuring similarity. The first feature is normalized by dividing it by the maximum value of the ω function within the shape domain. The second feature is normalized by dividing it by the area of the coarse structure. Experiments show that the range of the normalized second feature differences is more than twice that of the normalized first feature differences. In order to remove such a bias, the second feature differences are halved. Then each feature difference is converted to a similarity value x in the range $(0, 1]$ using the exponential function e^{-4x} . Finally, the node weight is determined as the weighted sum of

computed feature similarities. In implementation, two features are equally weighted (with 0.5).

In the first example, two shapes are selected, which are from similar categories: cat and horse. The preliminary part hierarchy trees of two shapes are demonstrated in Figures 5.2 and 5.3. Obviously, the structures of two trees are identical, except that the front legs of the horse shape are not partitioned. The selected maximum weighted clique (taken over 100 sample comparisons) contains 11 associations: $root \Leftrightarrow root$, $11 \Leftrightarrow 11$ (central regions), $21 \Leftrightarrow 21$ (peripheral regions), $211 \Leftrightarrow 211$ (front bodies), $2111 \Leftrightarrow 2111$ (heads), $2112 \Leftrightarrow 2112$ (front legs), $212 \Leftrightarrow 212$ (rear bodies), $2121 \Leftrightarrow 2122$ (fourth legs+tails), $21211 \Leftrightarrow 21221$ (fourth legs), $21212 \Leftrightarrow 21222$ (tails), $2122 \Leftrightarrow 2121$ (third legs). The match between leaf parts of two shapes are shown in Figure 5.5.

In the second example, the cat shape used in the first experiment (Figure 5.2) is compared to another cat shape whose preliminary part hierarchy tree is depicted in Figure 5.4. This is a more challenging example including the following differences between two preliminary part hierarchy trees:

- The spurious division of the rear body of the second cat causes an erroneous shift in the levels of the parts of the rear body.
- The fourth leg and the tail of the second cat is not separated.
- The head is divided into two parts in the tree of the second cat (which is not observed in the tree of the first cat).

Despite the differences between two trees the selected maximum weighted clique contains all of the correct associations: $root \Leftrightarrow root$, $11 \Leftrightarrow 11$ (central regions), $21 \Leftrightarrow 21$ (peripheral regions), $211 \Leftrightarrow 211$ (front bodies), $2111 \Leftrightarrow 2111$ (heads), $2112 \Leftrightarrow 2112$ (front legs), $21121 \Leftrightarrow 21121$ (first legs), $21122 \Leftrightarrow 21122$ (second legs), $212 \Leftrightarrow 2121$ (rear bodies), $2121 \Leftrightarrow 21211$ (fourth legs+tails), $2122 \Leftrightarrow 21212$ (third legs). In this clique, the parts that have no corresponding parts in the other tree are excluded (parts 21111 and 21112 of the second cat, and parts 21211 and 21212 of the first cat). Moreover the shifted parts of the second cat's rear body are successfully matched to their counterparts in the tree of the first cat as a result of randomized reorganization. At the winning reorganized tree sample, level three nodes of the second cat 2121 and 2122 replace their parent 212. Therefore the selected clique does not

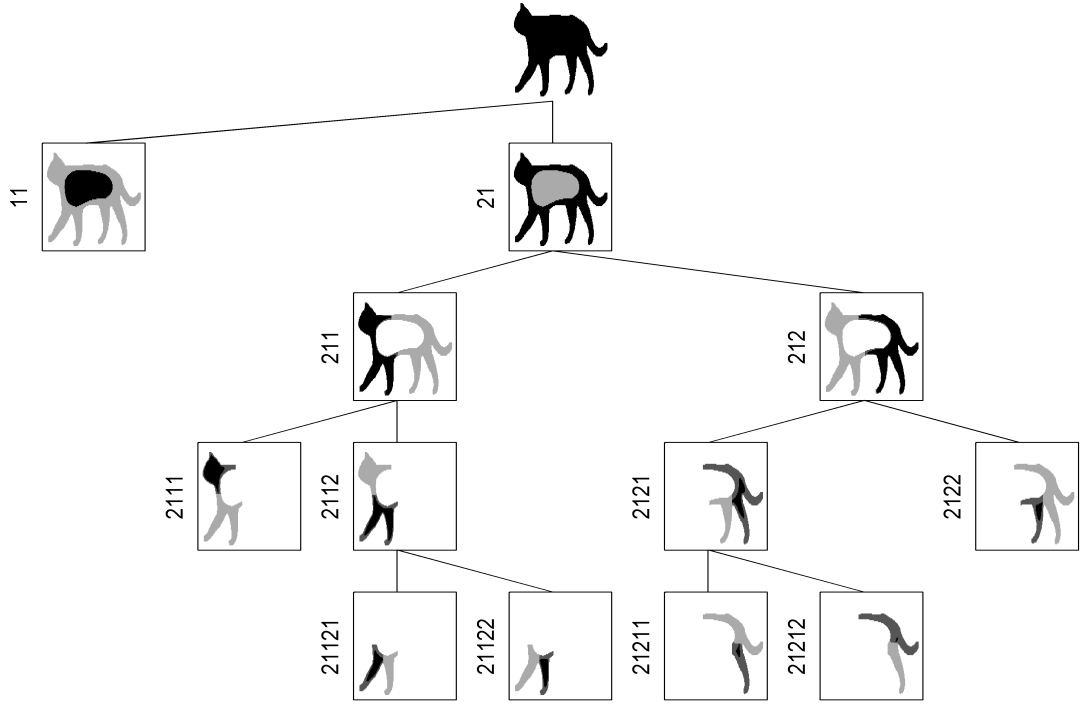


Figure 5.2: The preliminary part hierarchy tree of a cat shape.

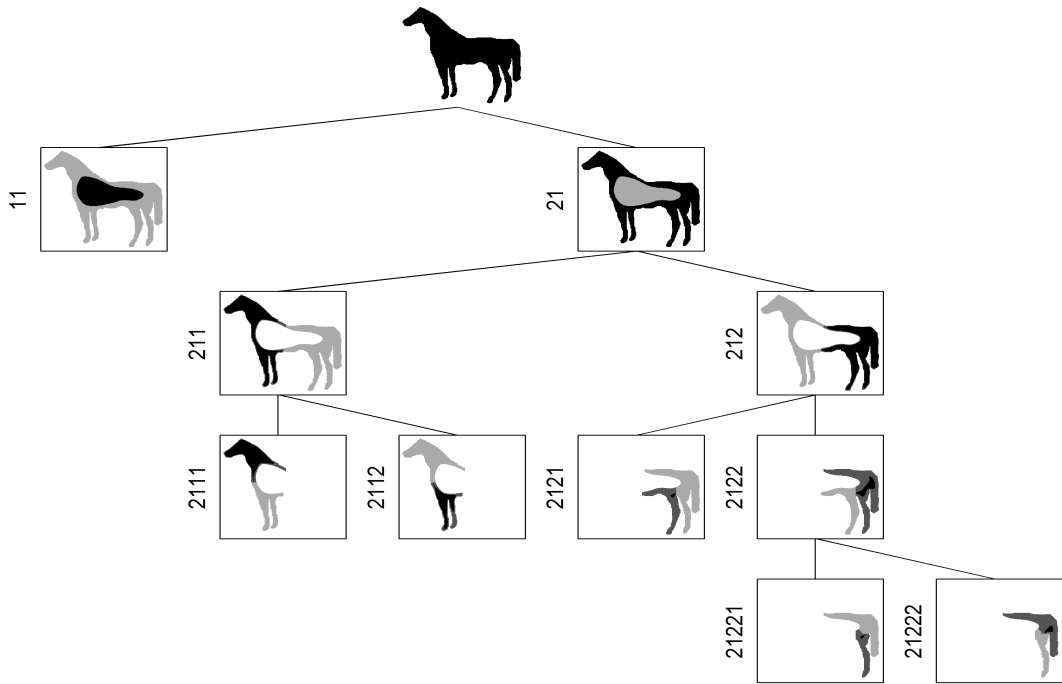


Figure 5.3: The preliminary part hierarchy tree of a horse shape.

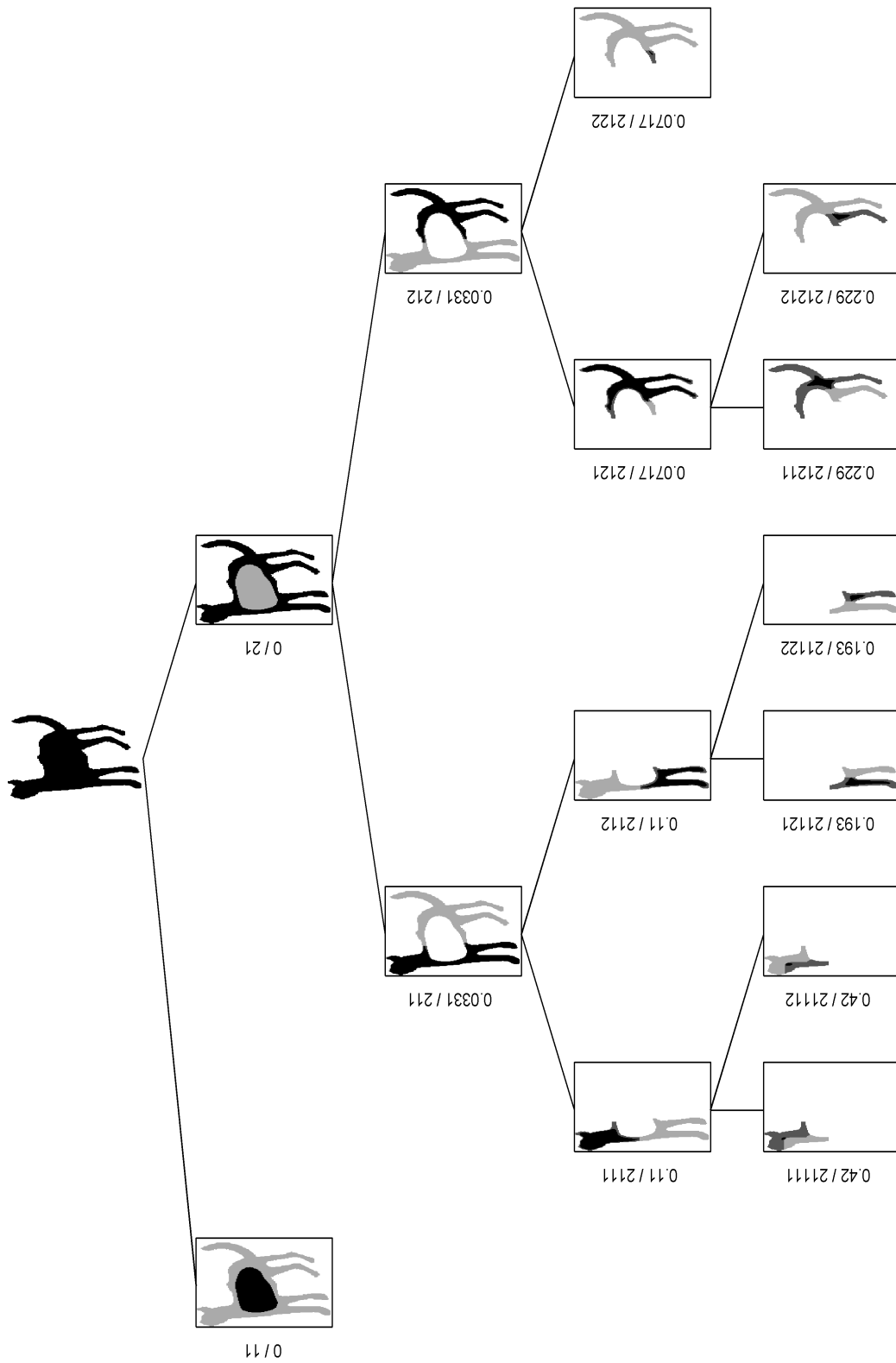


Figure 5.4: The preliminary part hierarchy tree of a cat shape.

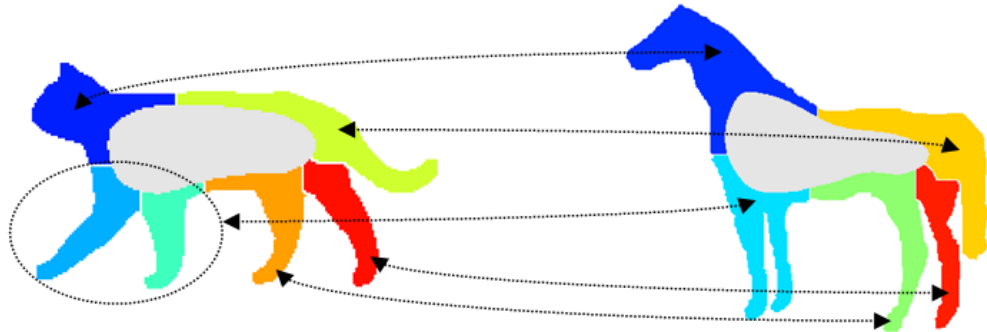


Figure 5.5: (First example) Matchings between leaf parts of cat and horse shapes, whose preliminary part hierarchy trees were given in Figures 5.2 and 5.3. Front legs of the cat are encircled to imply their parent in the preliminary part hierarchy tree.

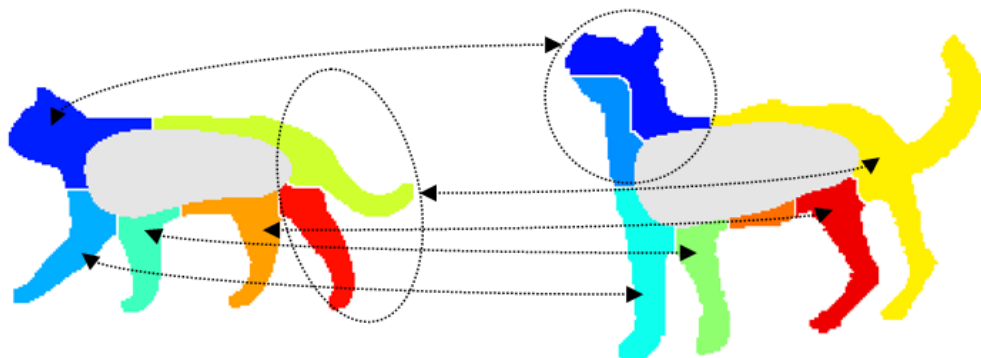


Figure 5.6: (Second example) Matchings between leaf parts of two cat shapes, whose preliminary part hierarchy trees were given in Figures 5.2 and 5.4.

include the parent node 212. The match between leaf parts of two shapes are shown in Figure 5.6.

In the third example, two human shapes are compared whose preliminary part hierarchy trees are shown in Figures 5.7 and 5.8 respectively. The difference is that the lower body of the second shape is shifted down twice due to an occlusion. Demonstrating the capability of the randomization procedure, the winning clique includes all of the correct associations: $root \Leftrightarrow root$, $11 \Leftrightarrow 11$ (central regions), $21 \Leftrightarrow 21$ (peripheral regions), $211 \Leftrightarrow 211$ (upper bodies), $21111 \Leftrightarrow 21111$ (arms on the left), $21112 \Leftrightarrow 21112$ (heads), $2112 \Leftrightarrow 2112$ (arms on the right), $212 \Leftrightarrow 21212$ (lower bodies), $2121 \Leftrightarrow 212121$ (legs on the left), $2122 \Leftrightarrow 212122$ (legs on the right). Two successive changes that occur at the reorganization of the winning randomized tree are as follows: First, nodes 2121 and its sibling 2122 replace their parent 212 and become the children of the node 21. Second, the lower body (21212) and its sibling replace their parent 2121 and become the children of their grandparent (which is node 21 after the first change). Consequently, the nodes 212 and 2121 doesn't appear in the selected clique. The winning randomized tree pair and the node associations are shown in Figure 5.9.

In the fourth example, two horse shapes are compared. The preliminary part hierarchy trees are given in Figures 5.3 and 5.10 respectively. The selected clique contains the following 11 associations: $root \Leftrightarrow root$, $11 \Leftrightarrow 11$ (central regions), $21 \Leftrightarrow 21$ (peripheral regions), $211 \Leftrightarrow 211$ (front bodies), $2111 \Leftrightarrow 21111$ (heads), $2112 \Leftrightarrow 2112$ (front legs), $212 \Leftrightarrow 2122$ (rear bodies), $2121 \Leftrightarrow 21221$ (third leg \Leftrightarrow tail), $2122 \Leftrightarrow 21222$ (fourth leg + tail \Leftrightarrow third leg + fourth leg), $21221 \Leftrightarrow 212222$ (fourth legs), $21222 \Leftrightarrow 212221$ (tail \Leftrightarrow third leg). In this case, a confusion appears in matching the parts of the rear body. The confusion is a result of the partitioning hierarchy. First split in the rear body separates the third leg of the first horse from the remaining parts (fourth leg + tail), whereas first split separates the tail of the second horse from the remaining parts (third leg + fourth leg).

Interestingly, in another trial, a sub-optimal fully correct matching is obtained that contains 10 associations instead of 11. In the sub-optimal solution, the node 2122 (fourth leg + tail) of the first tree, and the node 21222 (third leg + fourth leg) are both replaced by their children. Consequently, the source of the problem is removed and correct correspondences between the third legs and the tails are found. When we check the clique weights (that indicate the total similarity between the associated parts), we see that the clique with 11 associations has a total

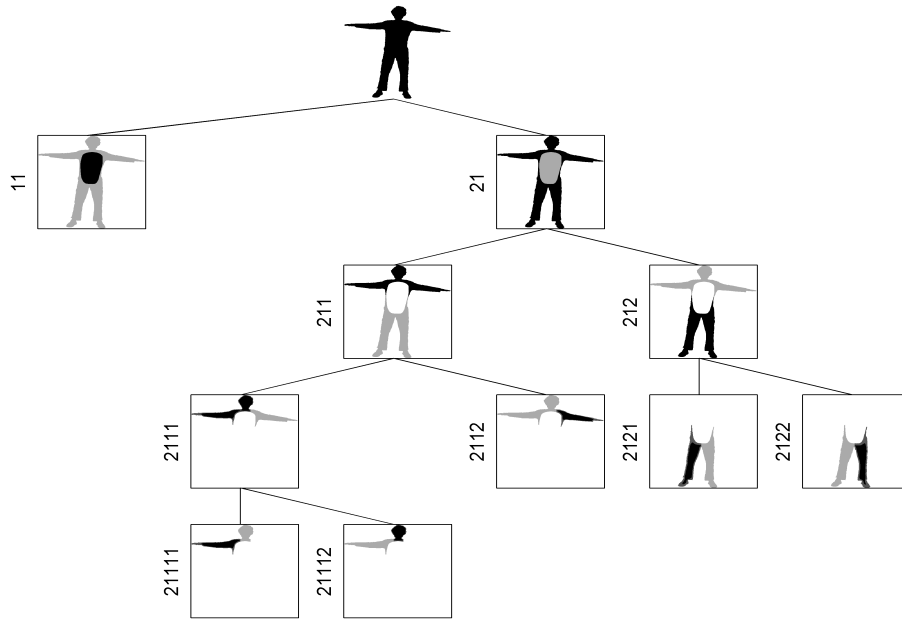


Figure 5.7: The preliminary part hierarchy tree of a human shape.

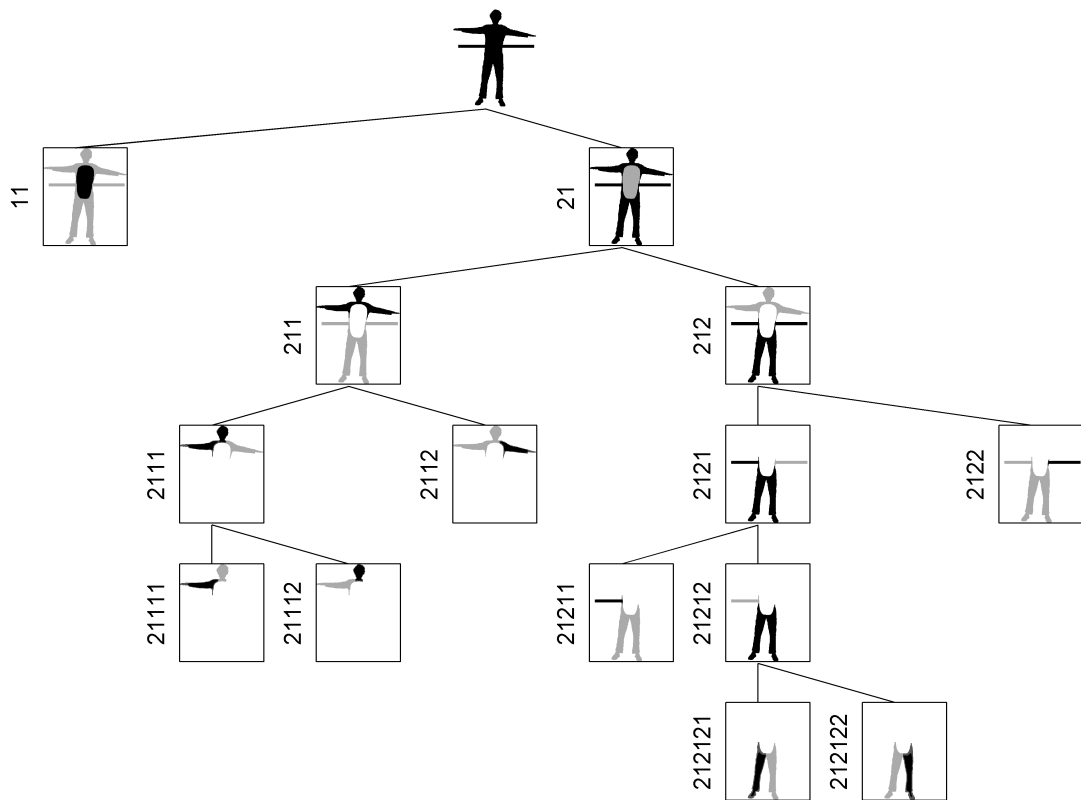


Figure 5.8: The preliminary part hierarchy tree of an occluded human shape.

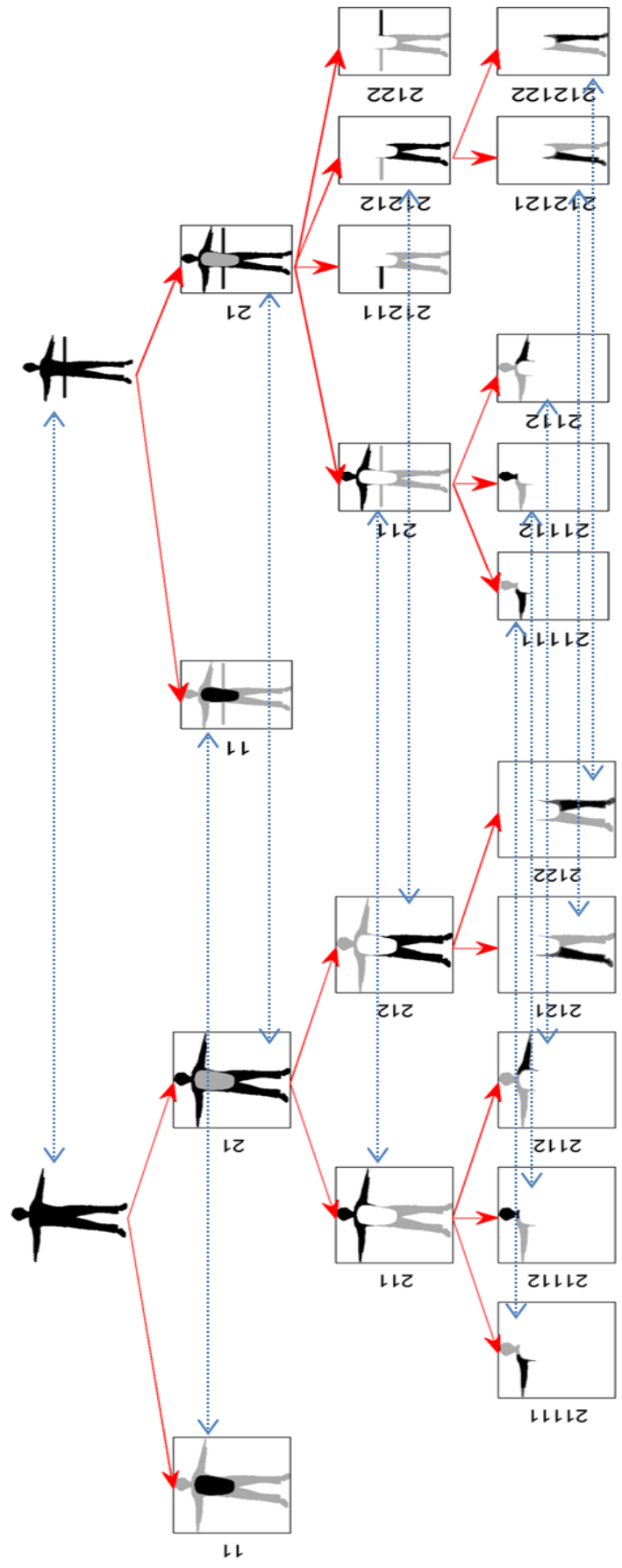


Figure 5.9: Matching human shapes (in Figures 5.7 and 5.8): The winning randomized hierarchy tree pair with node associations.

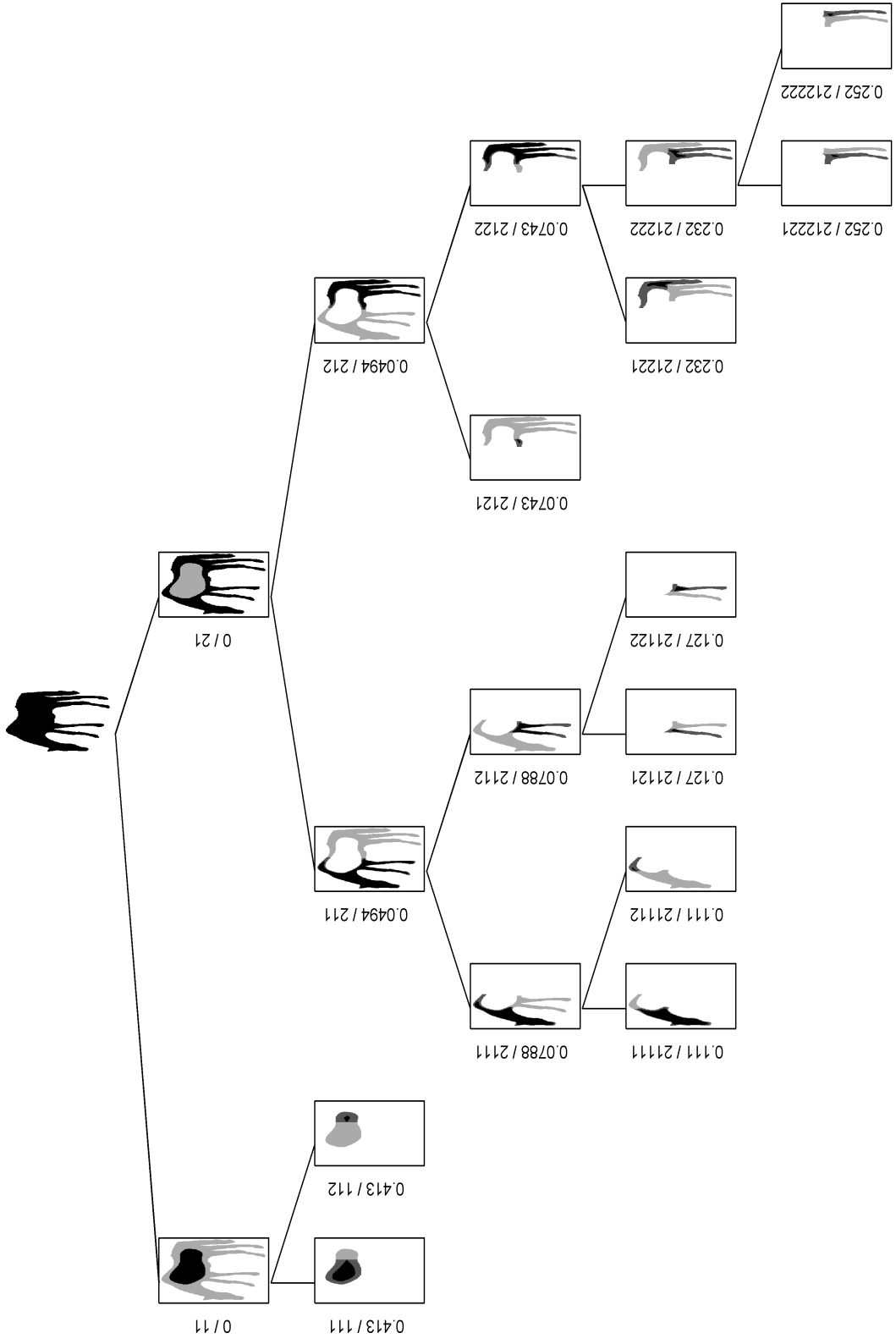


Figure 5.10: The preliminary part hierarchy tree of a horse shape.



Figure 5.11: (Third example) Matchings between leaf parts of a human shape and its occluded version, whose preliminary part hierarchy trees were given in Figures 5.7 and 5.8.

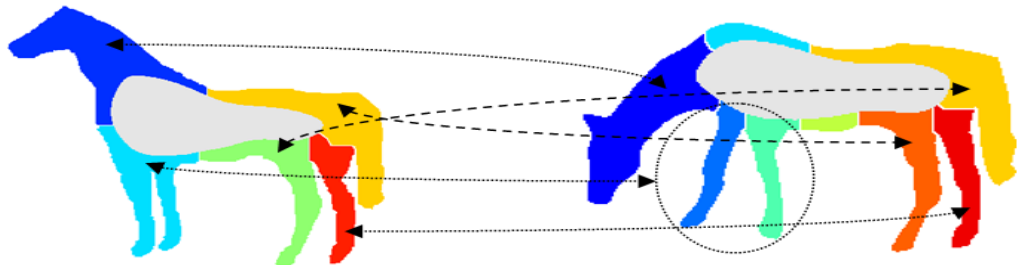


Figure 5.12: (Fourth example, first trial) Optimal matchings between leaf parts of horse shapes, whose preliminary part hierarchy trees were given in Figures 5.3 and 5.10. Confusion appears in matching third legs and tails.

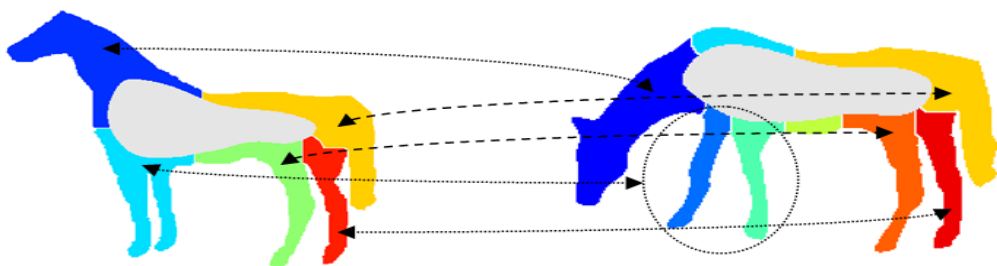


Figure 5.13: (Fourth example, second trial) Sub-optimal matchings between leaf parts of horse shapes, whose preliminary part hierarchy trees were given in Figures 5.3 and 5.10. No confusion appears in this case.

weight of 9.86, whereas the total weight for the sub-optimal solution with 10 associations is 9.33. Here, there are two related indicators in favor of the sub-optimal and fully correct solution. First, it has a greater average weight ($0.93 > 0.89$). Second, the weights of the individual associations reveal incorrect matchings. The tails match with the score 0.95 and the third legs with the score 0.91, whereas the third leg of the first horse matches to the tail of the second horse with the score 0.77 and the tail of the first horse matches to the third leg of the second horse with the score 0.80. This suggest that either the average weight or the individual weights can be examined to determine the correct solution instead of just selecting the one that has the maximum total weight among 100 maximum weighted cliques.

5.2 Matching Global Skeletons

Different from conventional medial descriptors, global skeletons combine axis based (skeleton branches) and region based (coarse structure) descriptions. Considering this, matching of two shapes is formulated as matching the main bodies with their contexts. We call a main body with its surrounding peripheral parts (its context) with a special name "main-context". In global skeleton scheme, the main body is represented by the coarse structure that encloses the central region associated with the main body. Then the surrounding peripheral parts are represented by the skeleton branches.

In the remaining part of this section, firstly the context tree is defined as a data structure to represent main-contexts. Secondly, the tree editing distance algorithm is mentioned briefly that is used for matching leaves of two context trees. Thirdly, application of the method to match different types of shapes is explained. Finally, experimental results are discussed.

5.2.1 The Context Tree

The context tree is a tree structure constructed for representation of main-contexts. The root of the tree is designed for description of the main body (coarse structure), while the leaves are designed for description of the peripheral parts (skeleton branches) in the context of the main body. In terms of context trees, a single-body shape, which has a single main body, is represented by a single context tree, while a multiple-body shape, which has more than one main body, is represented by a forest of context trees. Intermediate-Form shapes are represented in

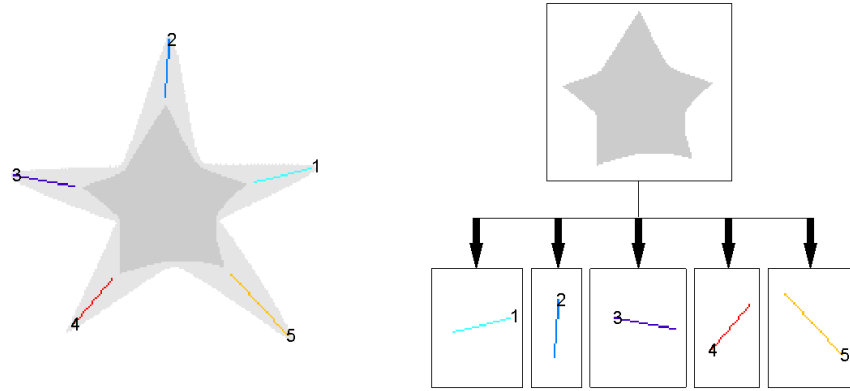


Figure 5.14: The symbolic context tree representation for a single-body shape.

either ways as they yield two different global skeleton representations. The leaves are ordered according to the counterclockwise order of the termination points of skeleton branches on the shape boundary. Symbolic context tree representations for a single and a multiple-body shape are shown in Figure 5.14 and 5.15 respectively.

As the context tree includes two types of nodes (root, leaf) that describe different aspects of the global skeleton (coarse structure and skeleton branches), the matching problem involves two sub-problems which have to be handled distinctly. Description and matching issue of the coarse structure is planned to be handled in a future work, since it is thought to be a problem on its own. Skeleton branches (leaves) are matched using tree editing distance algorithm by Zhang and Shasha [12].

Description of Skeleton Branches

A skeleton branch is described by its following features:

1. Normalized length l
2. Normalized polar coordinate r
3. Polar coordinate θ

Features are normalized in a way similar to the method of computing saliency measures. The normalization method for a feature f is as follows:

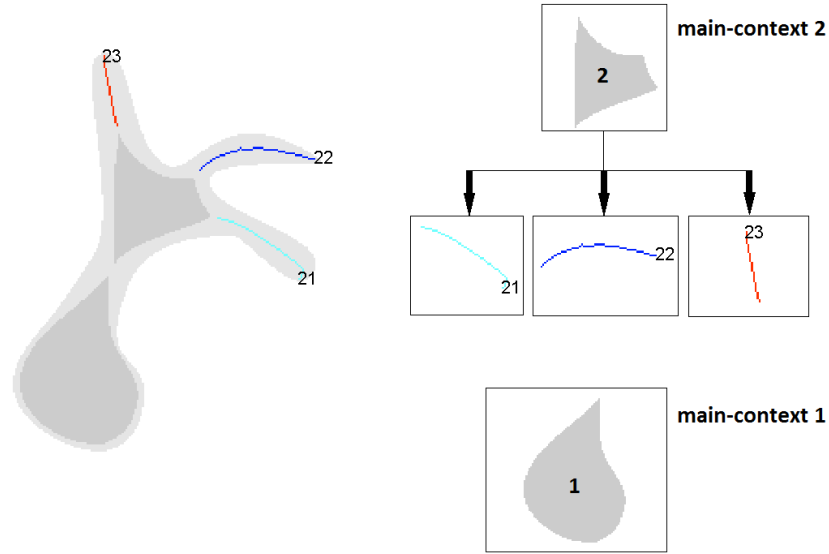


Figure 5.15: The symbolic context tree representation for a multiple-body shape.

Let \mathbf{f} be the vector that stores the observations of the feature f in a shape, \mathbf{f}^n be the vector to store the normalized values and \hat{f} be the mean value of the observations. Then, the normalized feature values are computed using the following equation:

$$\mathbf{f}^n = \mathbf{1} - e^{-c\mathbf{f}} \quad (5.1)$$

where

$$c = \frac{-\ln 0.5}{\hat{f}}. \quad (5.2)$$

5.2.2 Tree Editing Distance Algorithm by Zhang and Shasha [12]

A dynamic programming algorithm is developed by Zhang and Shasha for determination of the editing distance between two ordered labeled trees. Distinctive property of such trees is that left-to-right order among siblings is significant. The distance is considered to be the total cost of transforming one tree to another through edit operations, which are insert, delete and change. Here, changing a node means transforming it to another node. Deleting a node means making its children the children of its parent. Inserting is the complement operation of deleting.

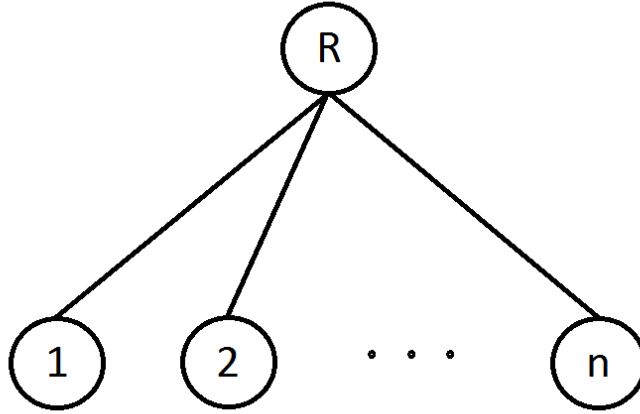


Figure 5.16: Topology of a context tree.

Let S be a sequence s_1, \dots, s_n of edit operations, which transforms a tree T_1 to another tree T_2 . Then, the cost of this particular transformation is as follows:

$$cost(S) = \sum_{i=1}^n cost(s_i) \quad (5.3)$$

Finally, the distance between T_1 and T_2 is the minimum cost that can be achieved by a sequence S that transforms T_1 to T_2 .

$$dist(T_1, T_2) = \min\{cost(S) | S \text{ is an edit operation sequence transforming } T_1 \text{ to } T_2.\} \quad (5.4)$$

Reader may refer to the paper [12] for the details of the general algorithm. Here, we only focus on the simpler algorithm (derived from the general algorithm) for matching leaves of the context tree as a special case (which resembles the case of string matching).

Before going into details, it is convenient to mention their notation. Let $T[i]$ be the i^{th} node in the tree according to the left-to-right postorder numbering. $T[i..j]$ is denoted as the ordered sub forest of T and if $i > j$, then $T[i..j] = \emptyset$.

Topology of a sample context tree looks like the one that is shown in Figure 5.16. Consider, we take the root of the context tree to match the leaves only. The remaining nodes (leaves) becomes a forest of subtrees which have roots only. Then, the simpler algorithm that measures the distance between two forests $T_1[1..m]$ and $T_2[1..n]$ which are derived from two context trees T_1 and T_2 with number of leaves m and n respectively is as follows:

```

forestdist( $\emptyset$ ,  $\emptyset$ ) = 0
for  $i = 1$  to  $m$  do
    forestdist( $T_1[1..i]$ ,  $\emptyset$ ) = forestdist( $T_1[1..i - 1]$ ,  $\emptyset$ ) + delete( $T_1[i]$ )
end for
for  $j = 1$  to  $n$  do
    forestdist( $\emptyset$ ,  $T_2[1..j]$ ) = forestdist( $\emptyset$ ,  $T_2[1..j - 1]$ ) + insert( $T_2[j]$ )
end for
for  $i = 1$  to  $m$  do
    for  $j = 1$  to  $n$  do
        forestdist( $T_1[1..i]$ ,  $T_2[1..j]$ ) = min {
            forestdist( $T_1[1..i - 1]$ ,  $T_2[1..j]$ ) + delete( $T_1[i]$ ),
            forestdist( $T_1[1..i]$ ,  $T_2[1..j - 1]$ ) + insert( $T_2[j]$ ),
            forestdist( $T_1[1..i - 1]$ ,  $T_2[1..j - 1]$ ) + change( $T_1[i]$ ,  $T_2[j]$ )
        }
    end for
end for

```

The distance between the forests $T_1[1..m]$ and $T_2[1..n]$ is then equal to $\text{forestdist}(T_1[1..m], T_2[1..n])$.

Mapping Skeleton Branches

The original work [12] doesn't provide a method that reveals the mapping which yields the distance computed. Nevertheless, it is easy to find the mapping by backtracking the edit operations. The skeleton branches that are operands of change operations are mapped to each other, while the others, which are operands of insert and delete operations, are not mapped.

Editing Costs

$\text{delete}(T_1[i])$ is the cost of deleting i^{th} node of the tree T_1 . The cost is equal to the normalized saliency weighted length (see Chapter 4 for details) of the node to be deleted.

$$\text{delete}(T_1[i]) = l_i^{\text{sw}} \quad (5.5)$$

$\text{insert}(T_2[j])$ is the cost of inserting j^{th} node of the tree T_2 . Similarly, the cost is equal to the normalized saliency weighted length of the node to be inserted.

$$\text{insert}(T_2[j]) = l_j^{\text{sw}} \quad (5.6)$$

$change(T_1[i], T_2[j])$ is the cost of modifying the feature vector α^i of the node $T_1[i]$ into the feature vector α^j of the node $T_2[j]$. Different features contribute with different weights to the changing cost. Weights are stored in the weight vector β , where $\sum \beta_i = 1$.

$$change(T_1[i], T_2[j]) = \beta \cdot |\alpha^i - \alpha^j| \quad (5.7)$$

$$\beta = \begin{bmatrix} \beta_l \\ \beta_r \\ \beta_\theta \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.3 \\ 0.4 \end{bmatrix}, \alpha = \begin{bmatrix} l \\ r \\ \theta \end{bmatrix} \quad (5.8)$$

5.2.3 Matching Strategies for Different Types

Different representations require different strategies to match shapes from different types. Even, some pairs can not be matched against, if they are from different main types, i.e. one is a single-body and the other is a multiple-body shape. In such cases, editing distance is specified as infinity. Remaining cases are matching shapes of same types, and matching a shape of a main type with a shape of intermediate-form.

Matching Two Single-Body Shapes: Each single-body shape is represented by one context tree. Matching them is equivalent to matching these context trees and the computed distance is the dissimilarity of the two shapes. Since the first leaf in the order may vary in each tree, matching is repeated with shifted versions of the second tree, where the order of the leaves is changed by a circular shift operation. At the end, the match with minimum cost is selected.

Matching Two Multiple-Body Shapes: Each multiple-body shape has two context trees. This requires, additionally, finding the correct match between the context trees of each shape, which is achieved by switching the tree pairs to be matched against. Again, the match with minimum cost is selected.

Matching an Intermediate-Form Shape to a Shape of a Main Type: Intermediate-form shapes are represented in such a way that they can be compared with shapes of either main types. It is sufficient to select the appropriate representation and match using the appropriate one of the above methods, according to the type of the shape to be matched against.

Matching Two Intermediate-Form Shapes: An intermediate-form shape can be safely considered as a single-body shape, when it is matched against another intermediate-form shape.

5.2.4 Experimental Results

Shapes in the database are matched to each other using the method mentioned above. Matchings between skeleton branches of shapes from same categories are demonstrated in Appendix B. Matching branch pairs are depicted with the same color and also annotated with connecting arrows, whereas branches with black color have no corresponding branch in the other shape. No confusion appears in matching skeleton branches of the shapes within the same category in the database.

Retrieval results are given in Appendix C, where, for all shapes, first 10 nearest shapes are depicted with tree editing distances. Additionally, the distance matrix is shown in Figure 5.17. Here, dark levels indicate small distance, whereas light levels indicate large distance. The results seem to be promising, although coarse information is not introduced yet. It is observed that intra-category distances are generally small, but inter-category distances are not large as it should be. As a result, confusions appear between similarly articulated shape categories. Inter-category distances are thought to be increased, when main-bodies are compared by introducing coarse information. In the following paragraphs, we examine some problematic cases that appear in the current matching scheme.

Consider the category of star shapes for example, which are confused with some turtle shapes. In fact, placements and lengths of the skeleton branches are very similar in some pairs from these two categories, which then results in a small editing distance between their context trees. An example case is demonstrated in Figure 5.18. Coarse structures seem to be more discriminative in separating two categories.

The category of turtles in the database consists of two different groups, one with a prominent tail and one without a prominent tail. These two groups are distant from each other as a result of deleting (inserting) operations required for extra branches (Figure 5.19).

Another challenging issue is discrimination of cat and horse categories, which are structurally very similar. Even coarse structures are not much discriminative. We observe that retrieval

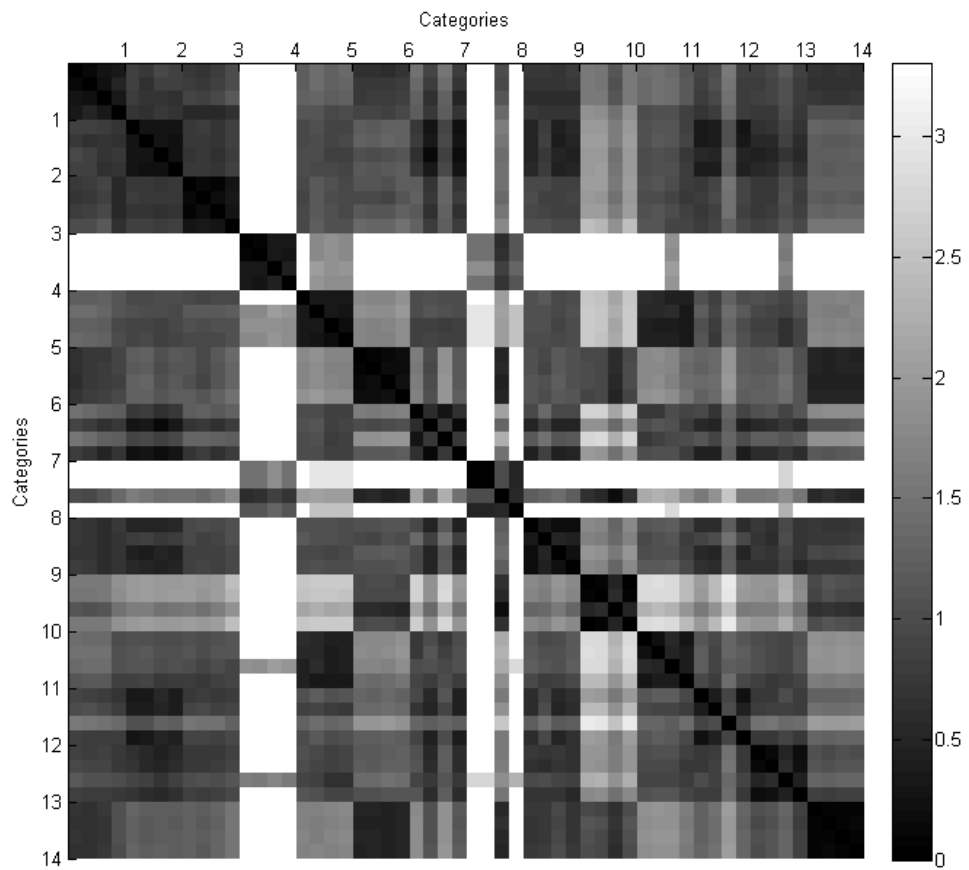


Figure 5.17: Distance matrix.

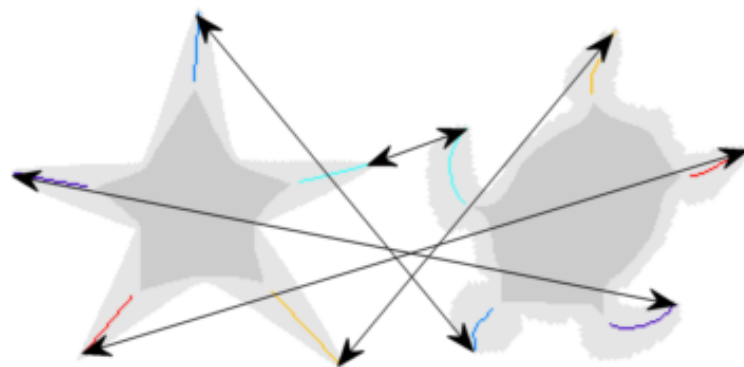


Figure 5.18: Matching a star to a turtle. Placements and lengths of the skeleton branches are very similar. Coarse structures may provide a clear distinction.

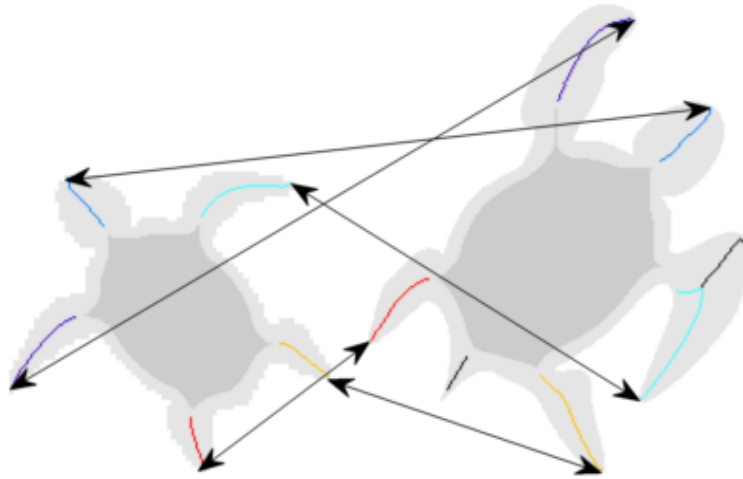


Figure 5.19: Matching turtles. The turtle on the right has extra branches which increase the editing distance.

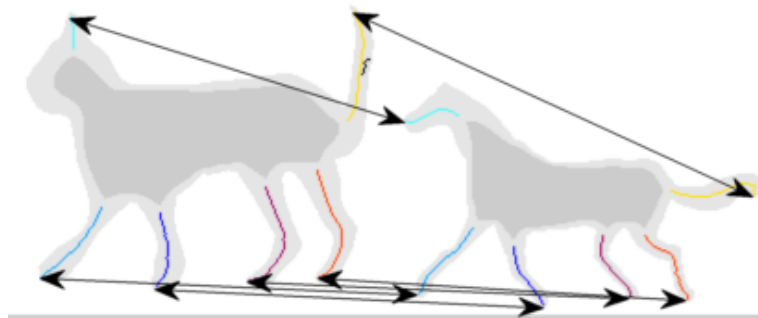


Figure 5.20: Matching a cat to a horse.

of horse shapes are successful in most of the cases, while cat shapes are confused with horse shapes many times. This can be attributed to consistent global skeleton representations of horse shapes and some inconsistencies in global skeletons of cat shapes especially in head and tail parts. Consistency results in small distances between horse shapes. However, inconsistencies keep cat shapes distant from each other causing some horse shapes to appear closer than other cat shapes. For example, the closest shape (after itself) to the third shape in the cat category is the third shape from horse category. Matching these two shapes is illustrated in Figure 5.20.

Finally, there are some shapes such as dog-bones (Category 8) and hearts (Category 10), which have no prominent ribbon-like peripheral parts. Instead, their peripheral parts are blob-

like and as stated before, blob-like parts are subject to be included mostly by the coarse structure. This fact entails introduction of coarse structures into matching scheme to be able to match this kind of shapes robustly, since they are mostly represented by their main bodies.

CHAPTER 6

CONCLUSION

This thesis proposes a new scheme for shape representation that captures the part hierarchy and medial descriptors. The scheme utilizes the ω function defined over the shape domain recently proposed by Tari [9, 10] as an unconventional distance transform. The function is computed as the minimizer of an energy that captures both local and global interactions among shape points.

Global interactions endows the ω function with the interesting property of attaining both positive and negative values which partitions the shape robustly into central and peripheral regions at the zero level curve. Starting with this initial step, the partitioning hierarchy of the parts is extracted via level sets and watersheds of ω according to the order of separating saddle points. The resulting part hierarchy is organized into a proper binary tree named as the preliminary part hierarchy tree.

In addition to the zero level curve, a second closed curve emerges as another interesting result of the oscillation in the ω function. The new region enclosed by this curve can be obtained easily as a watershed region of the $-\omega$ function. Since this region robustly represents a coarse version of the shape, we call it the coarse structure. The coarse structure tends to exclude ribbon-like structures (limbs), whereas blob-like structures are mostly included. This property enables us to easily separate stable skeleton branches that represent limbs from unstable ones that fall inside the coarse structure.

In addition to the part hierarchy, medial descriptors are extracted. This is done in a complementary fashion so that medial descriptors and parts simultaneously support each other. Firstly ridge points are extracted which constitute the raw skeleton following the method proposed in [11]. Then, ridge points that fall inside the peripheral leaf parts of the preliminary

part hierarchy tree and outside of the coarse structure are grouped into skeleton branches. Among the extracted branches, only the ones that represent salient protrusions are selected.

Polar coordinate frames are constructed to describe the spatial organization of the extracted skeleton branches. Different constructions are proposed for shapes of different topologies. Firstly, a rough topological categorization is made based on the connectivity of central region. A single-body shape has a single central region (main body) that holds a single maximum, whereas multiple-body shapes have multiple separated central regions. In addition to these main categories, a third category emerges representing the transition category between them. Such a categorization is important, because spatial organization in multiple-body shapes can not be robustly described using a single global coordinate frame. Considering this, descriptions in such shapes are localized at the neighborhoods of the separate main bodies reducing the adverse effect of deformations in bounding necks. Finally, each main body and surrounding peripheral parts are described separately using depth-1 trees via the coarse structure (root) and skeleton branches (leaves) respectively. Shapes from transition category are represented multiply to be able to match them to shapes from both main categories. Emergent medial descriptor is named as global skeleton.

Experimental results are presented for matching both the part hierarchy trees and global skeletons. Matching two part hierarchy trees is formulated as finding the maximum weighted clique in a weighted association graph derived from the tree pair using the algorithm in [8]. Here, an important contribution is made by employing a randomization procedure before the derivation of the weighted association graph. The procedure yields a novel tree structure (named randomized hierarchy tree), where the depth of a node is determined probabilistically in order to overcome the depth instabilities in the preliminary part hierarchy trees. After trials with hundred randomized hierarchy tree samples, the best match is chosen as the one with the maximum score.

Global skeleton representations of two shapes are compared using the algorithm in [12] that finds the editing distances between depth-1 tree pairs. Although matching the coarse structures is left as a future work (i.e. only the skeleton branches that constitute the leaves of the tree are matched), the matching results seem promising. We expect that objects will be more distinctive after introduction of the coarse structure into matching.

REFERENCES

- [1] C. Aslan. Disconnected skeletons for shape recognition. Master's thesis, Middle East Technical University, 2005.
- [2] C. Aslan and S. Tari. An axis-based representation for recognition. In *ICCV '05: Proceedings of the Tenth IEEE International Conference on Computer Vision*, pages 1339–1346, Washington, DC, USA, 2005. IEEE Computer Society.
- [3] H. Blum. A transformation for extracting new descriptors of shape. *Models for the perception of speech and visual form*, 19(5):362–380, 1967.
- [4] H. Blum. Biological shape and visual science (Part I). *Journal of theoretical Biology*, 38(2):205–287, 1973.
- [5] Benjamin B. Kimia, Allen R. Tannenbaum, and Steven W. Zucker. Shapes, shocks, and deformations i: the components of two-dimensional shape and the reaction-diffusion space. *Int. J. Comput. Vision*, 15:189–224, July 1995.
- [6] D. Macrini, K. Siddiqi, and S. Dickinson. From skeletons to bone graphs: Medial abstraction for object recognition. In *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8, June 2008.
- [7] Stanley Osher and James A. Sethian. Fronts propagating with curvature-dependent speed: algorithms based on hamilton-jacobi formulations. *J. Comput. Phys.*, 79:12–49, November 1988.
- [8] M. Pelillo, K. Siddiqi, and S.W. Zucker. Matching hierarchical structures using association graphs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(11):1105–1120, 1999.
- [9] S. Tari. Hierarchical shape decomposition via level sets. *Mathematical Morphology and Its Application to Signal and Image Processing*, pages 215–225, 2009.
- [10] S. Tari. Extracting parts of 2d shapes using local and global interactions simultaneously. In C. Chen, editor, *Handbook of Pattern Recognition and Computer Vision*. World Scientific, 2010.
- [11] Z.S.G. Tari, J. Shah, and H. Pien. Extraction of shape skeletons from grayscale images. *Computer Vision and Image Understanding*, 66(2):133–146, 1997.
- [12] K. Zhang and D. Shasha. Simple fast algorithms for the editing distance between trees and related problems. *SIAM J. Comput.*, 18:1245–1262, December 1989.
- [13] Song Chun Zhu and Alan L. Yuille. Forms: a flexible object recognition and modeling system. *Int. J. Comput. Vision*, 20:187–212, December 1996.

Appendix A

GLOBAL SKELETONS OF THE SHAPES IN THE DATABASE

In this appendix, global skeletons of the shapes in the database are illustrated. Skeleton branches appear as black lines. A purple curve indicates the boundary of a coarse structure. Red lines are polar axes with black squares representing their origin.

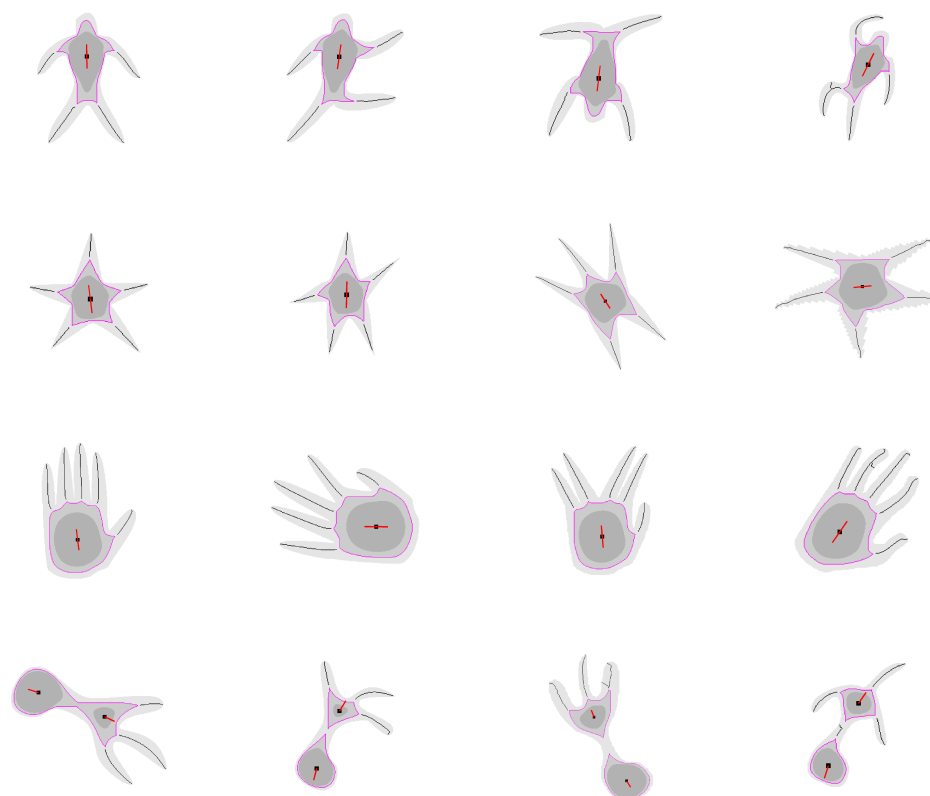


Figure A.1: Global skeletons.



Figure A.2: Global skeletons.

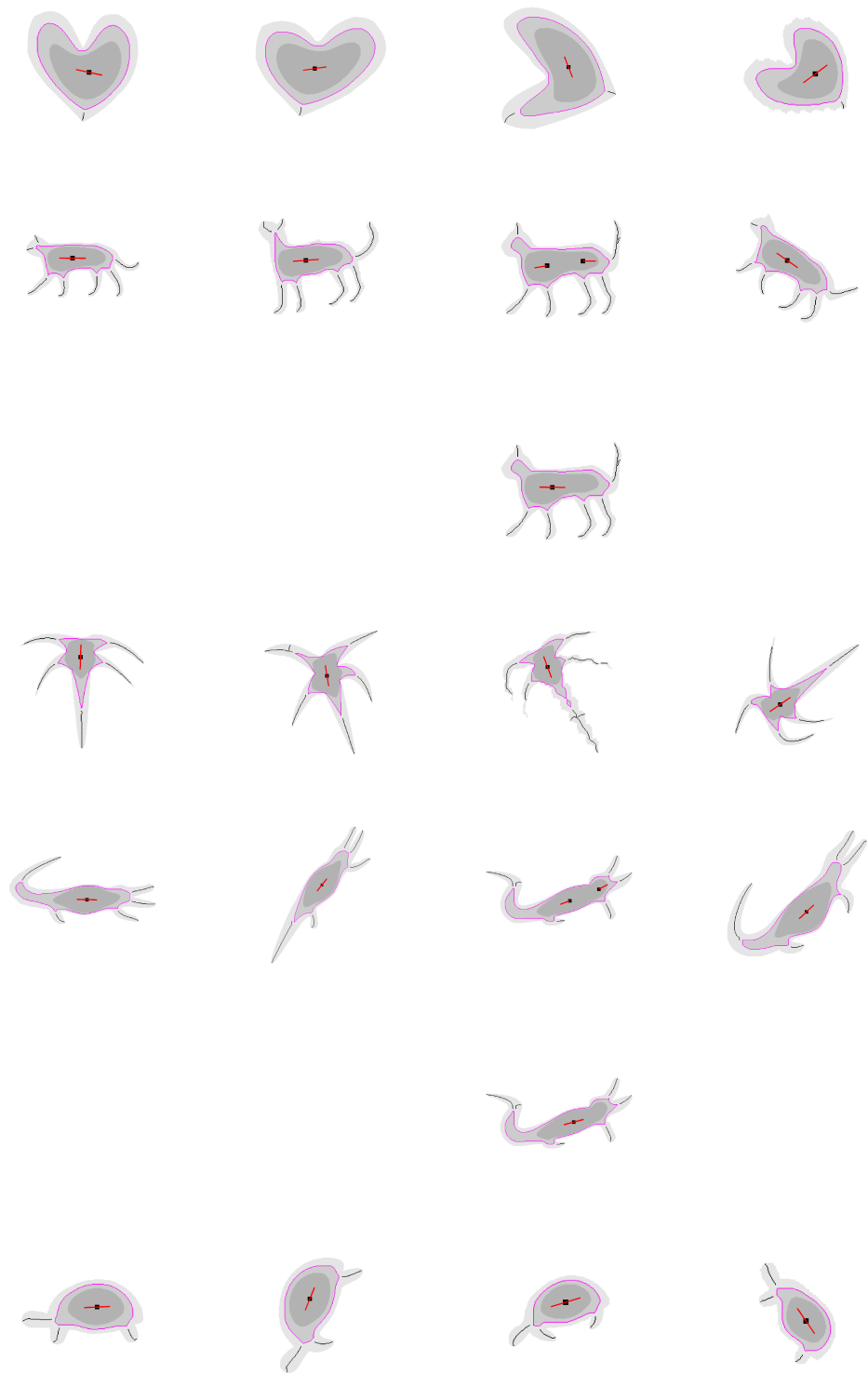


Figure A.3: Global skeletons.

Appendix B

MATCHING GLOBAL SKELETONS: INTRA-CATEGORY DEMONSTRATIONS

In this appendix, within category global skeleton matching demonstrations are given. Matching skeleton branch pairs are indicated with arrows and also with same colors. Black colored branches with no arrows indicating them are not matched.

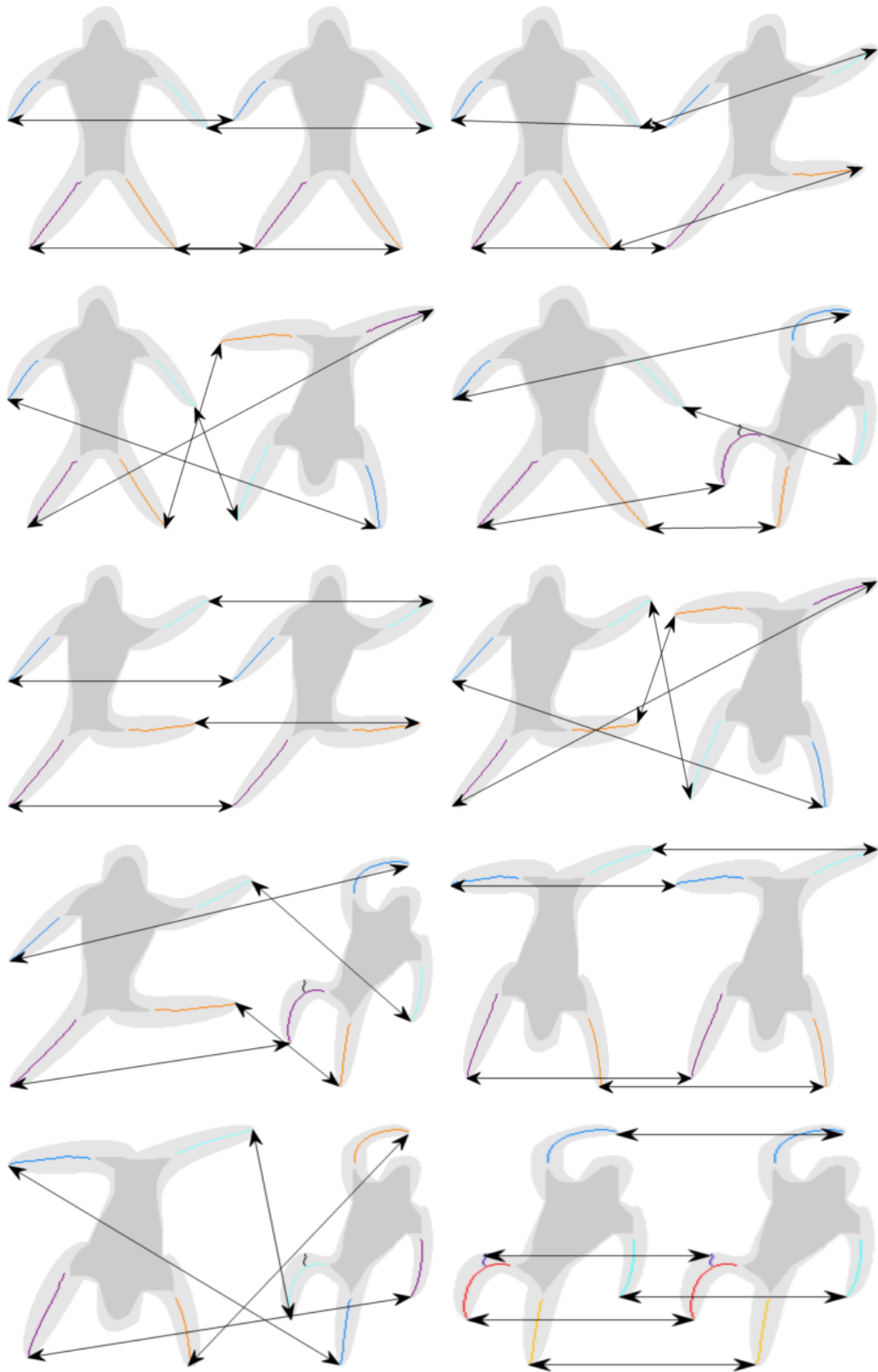


Figure B.1: Matching: Category 1

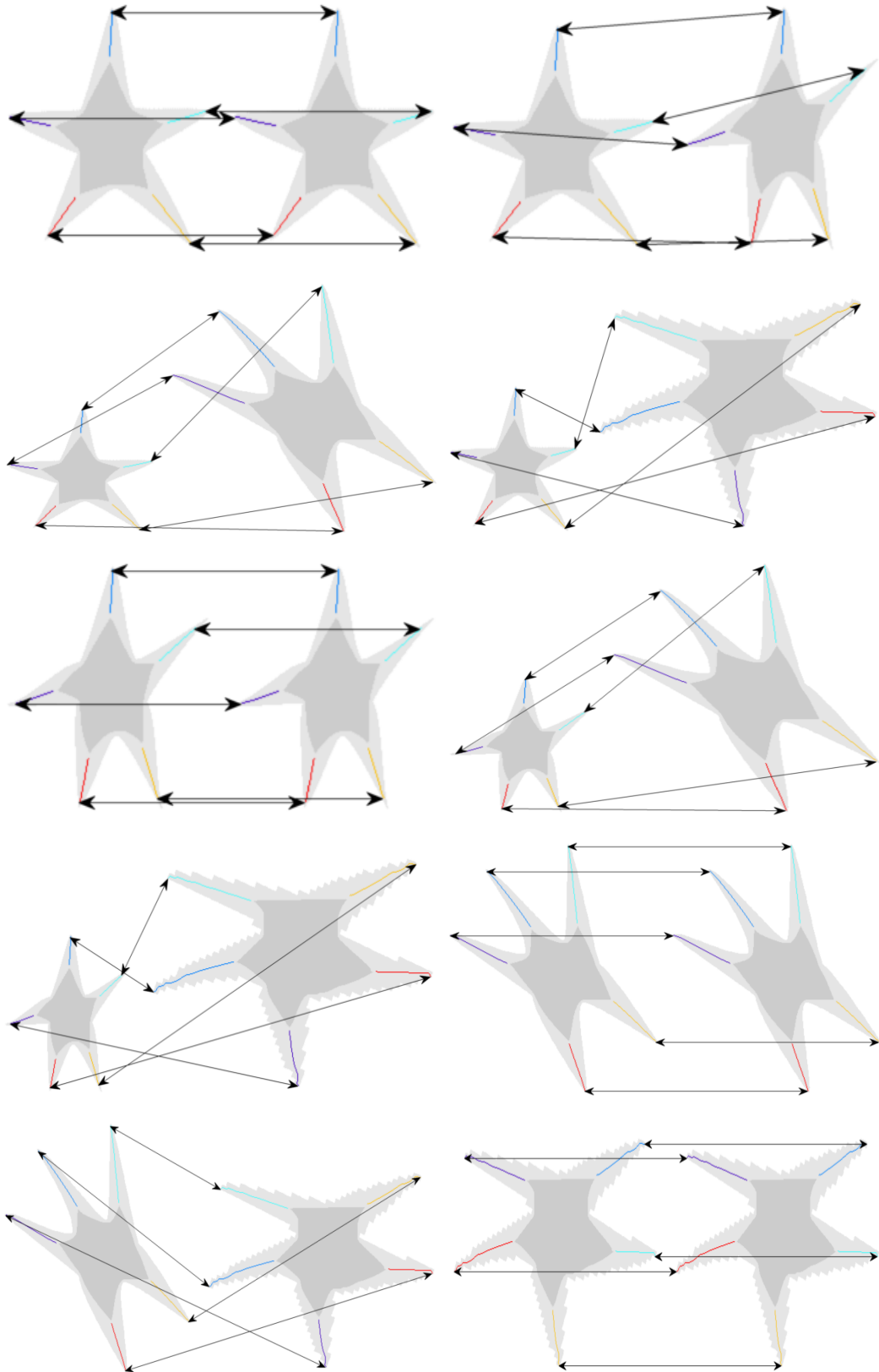


Figure B.2: Matching: Category 2

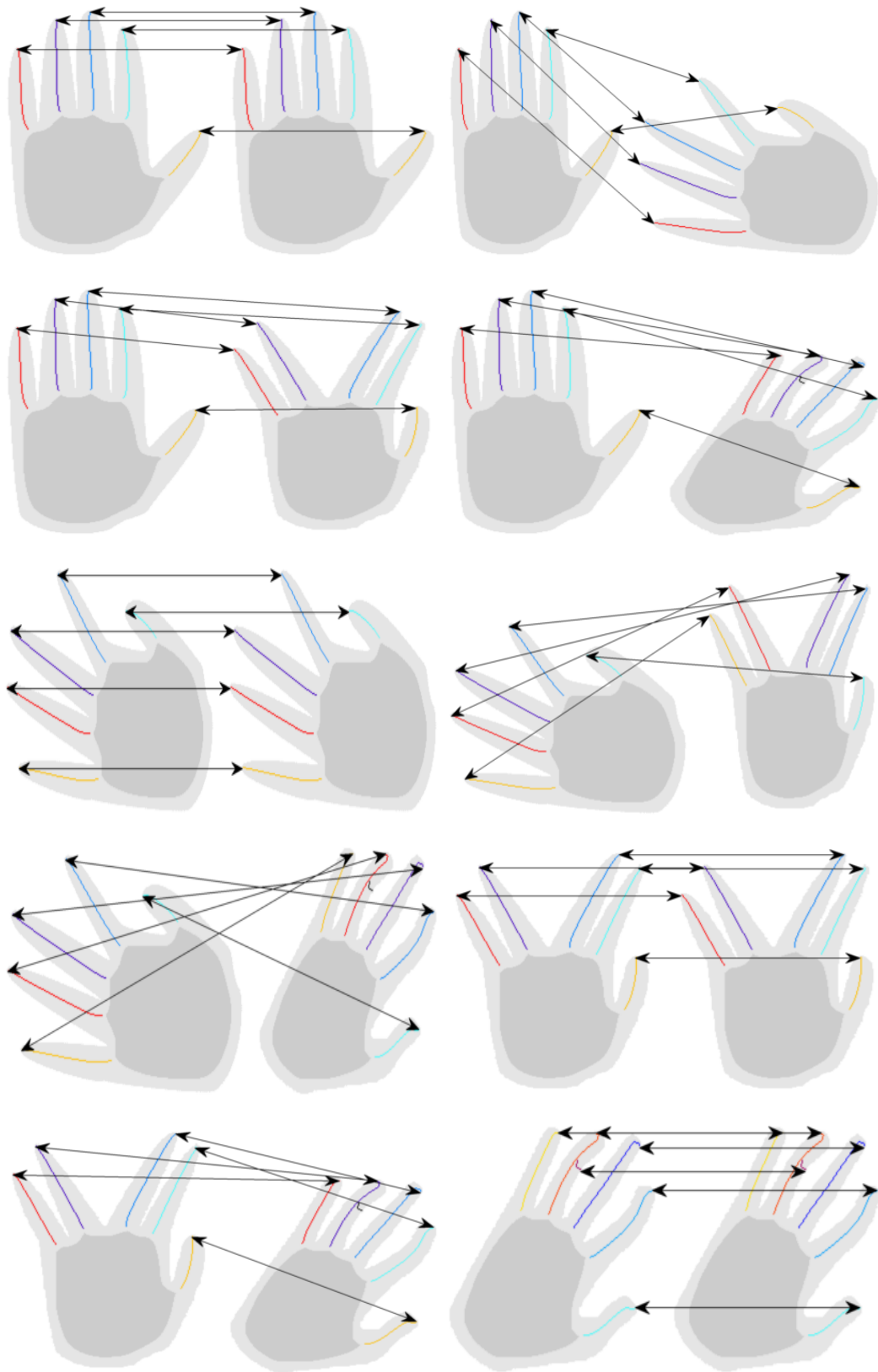


Figure B.3: Matching: Category 3

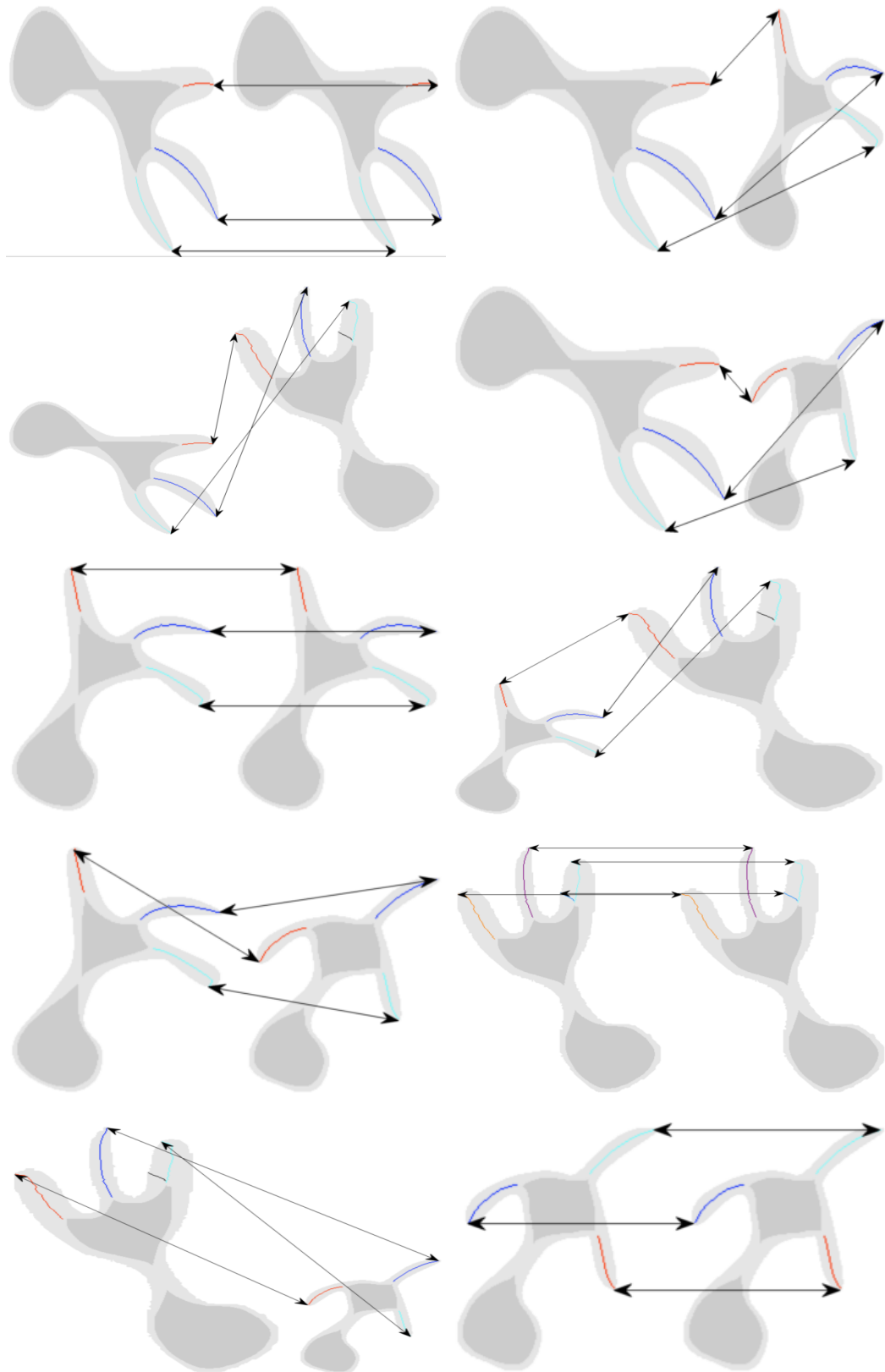


Figure B.4: Matching: Category 4

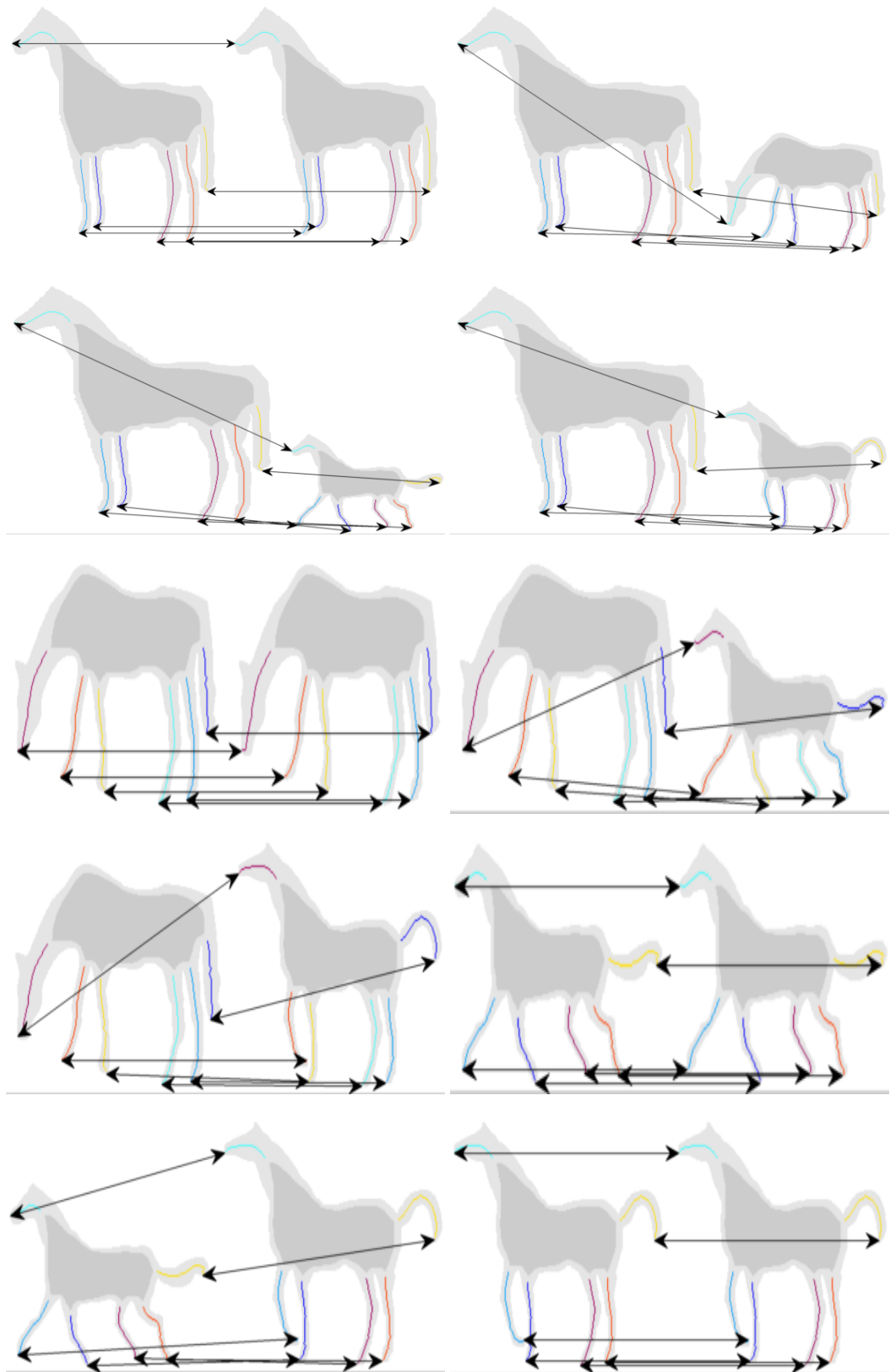


Figure B.5: Matching: Category 5

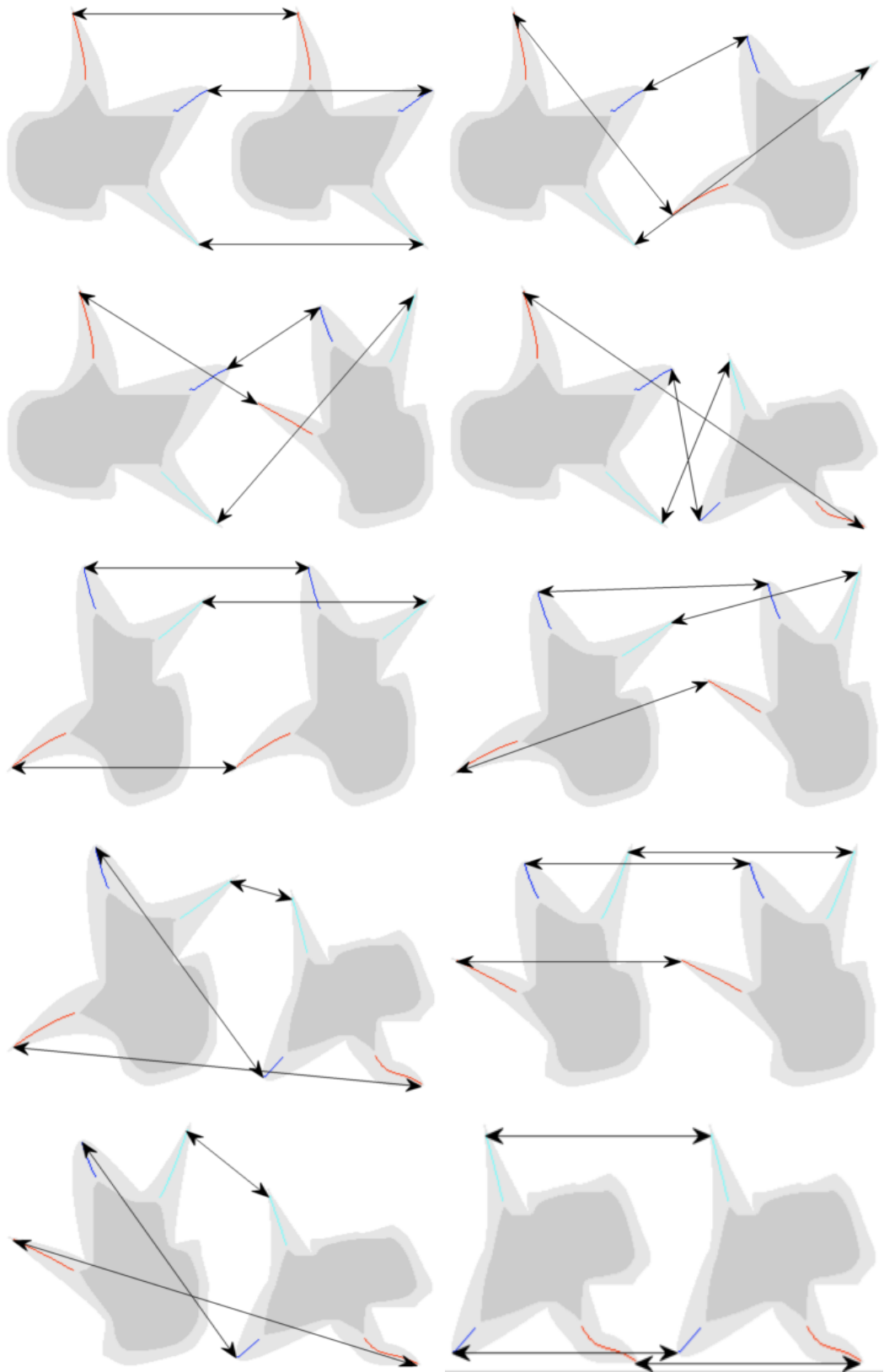


Figure B.6: Matching: Category 6

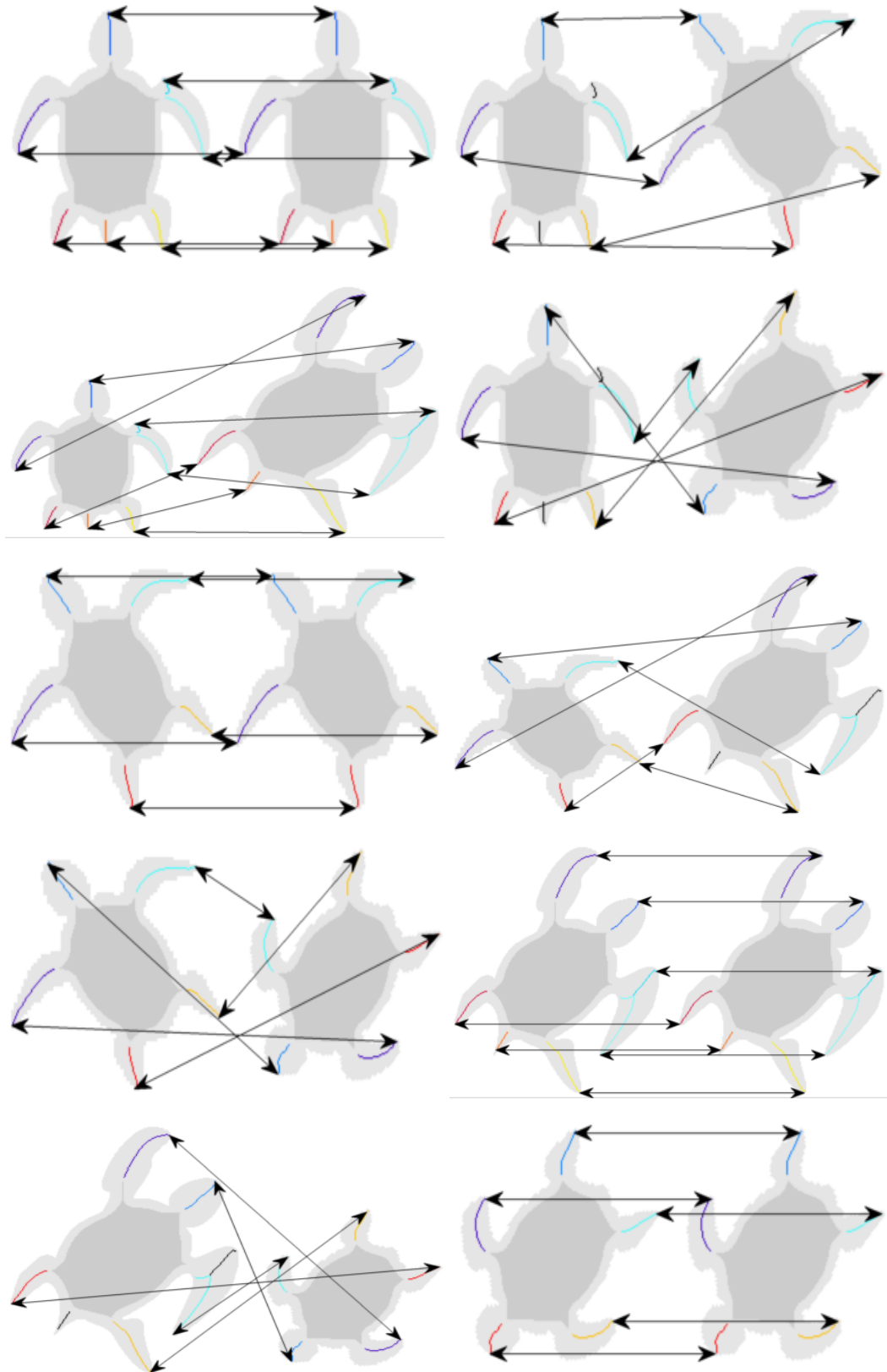


Figure B.7: Matching: Category 7

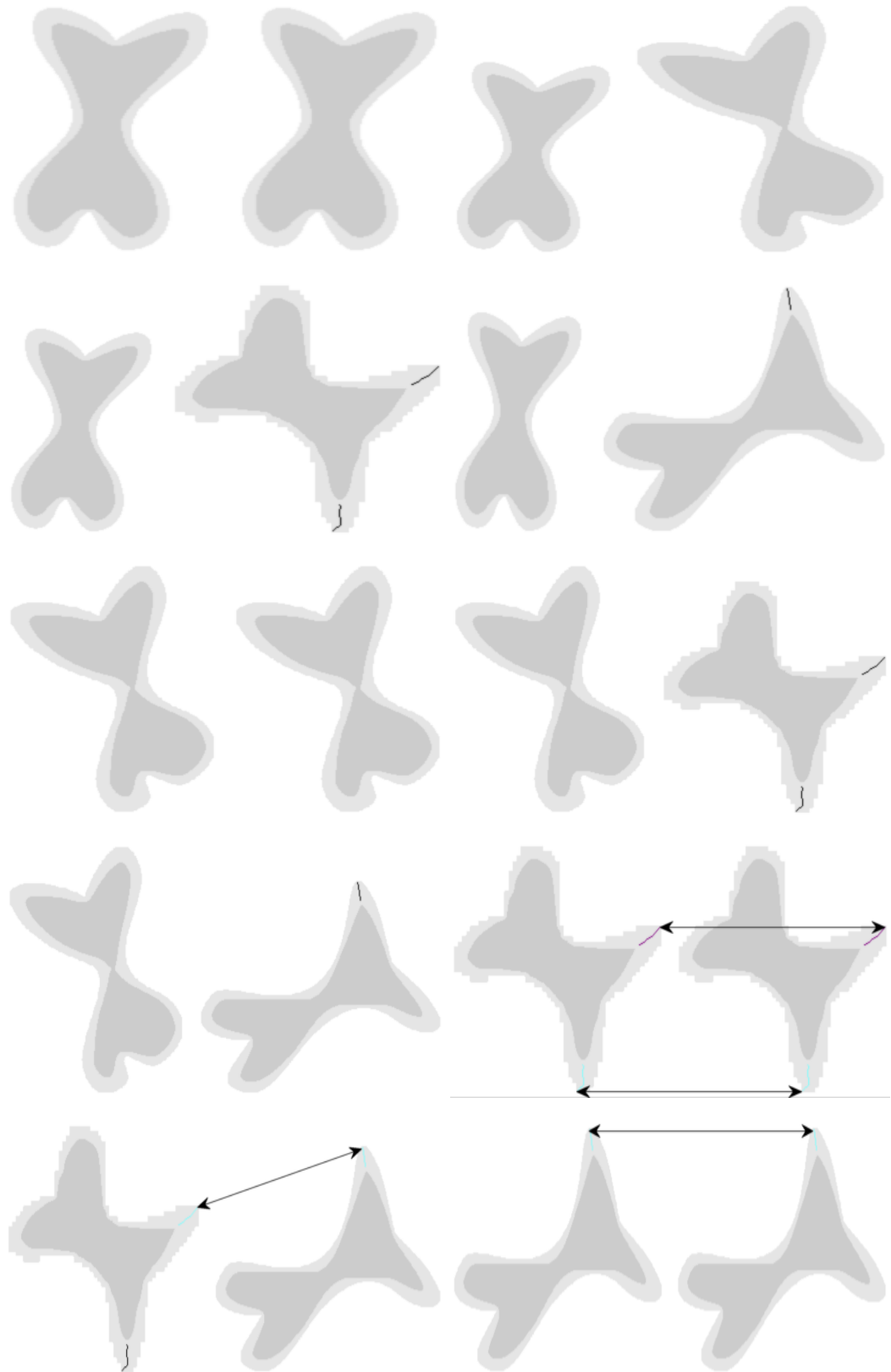


Figure B.8: Matching: Category 8

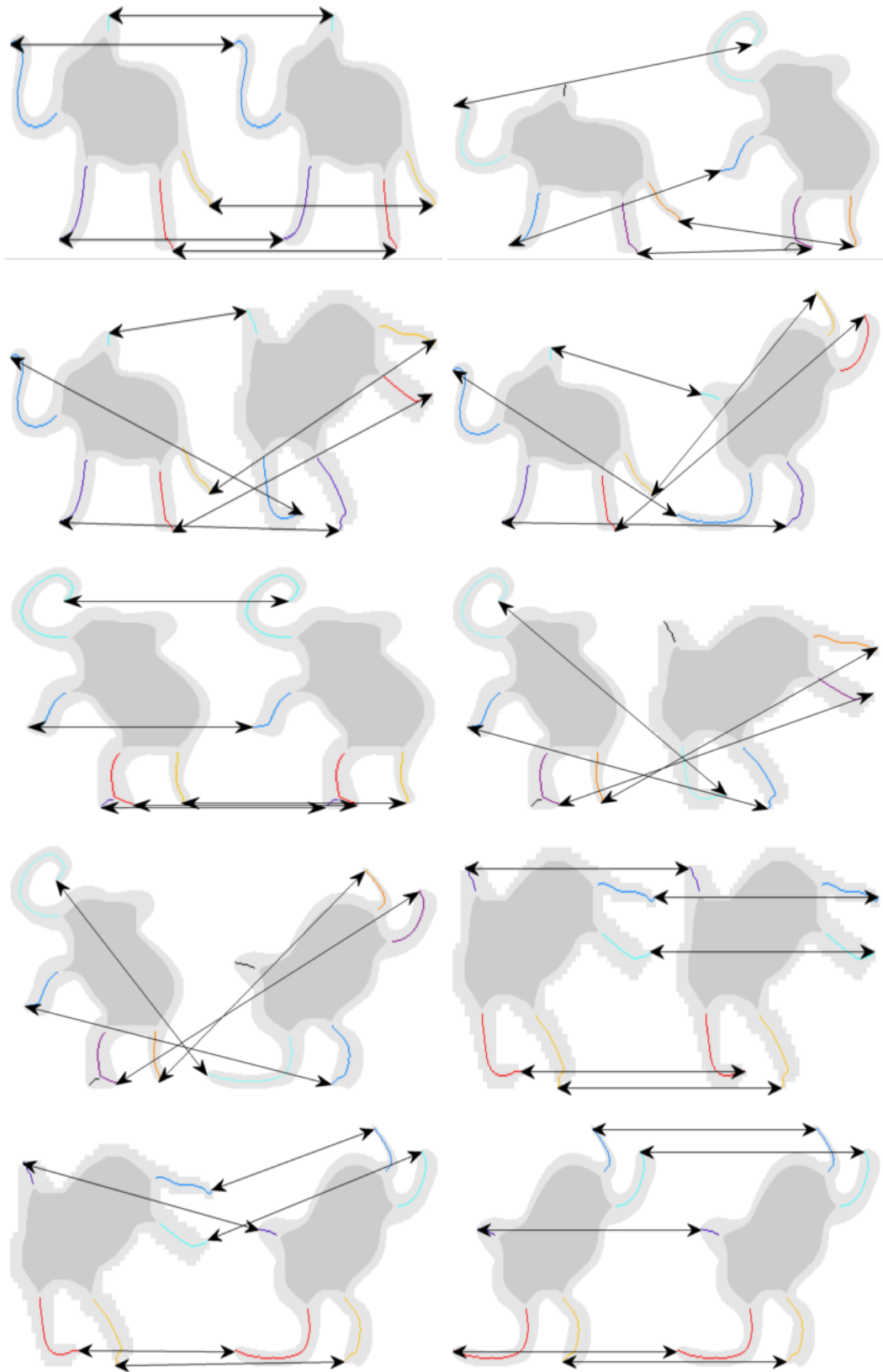


Figure B.9: Matching: Category 9

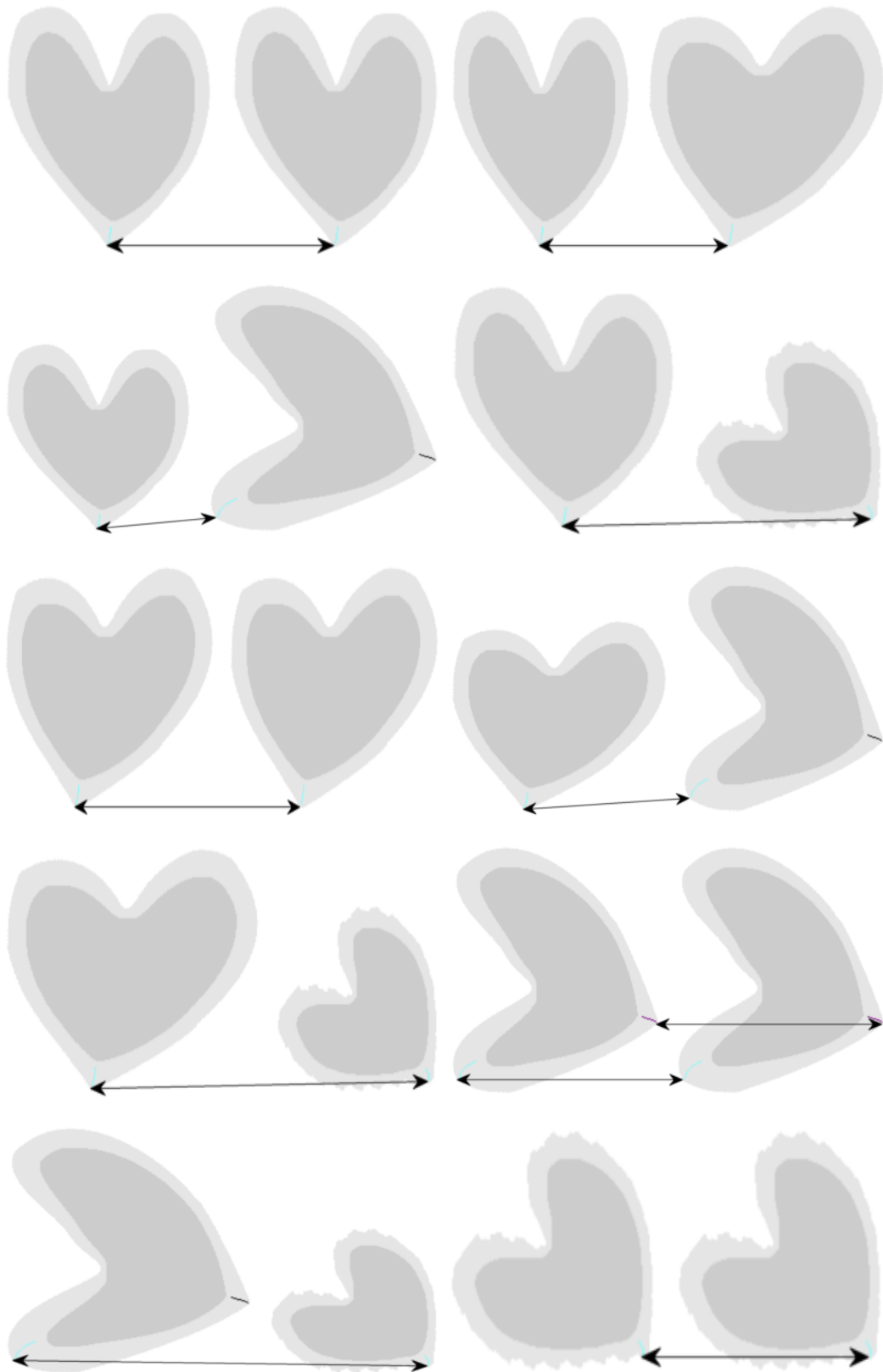


Figure B.10: Matching: Category 10

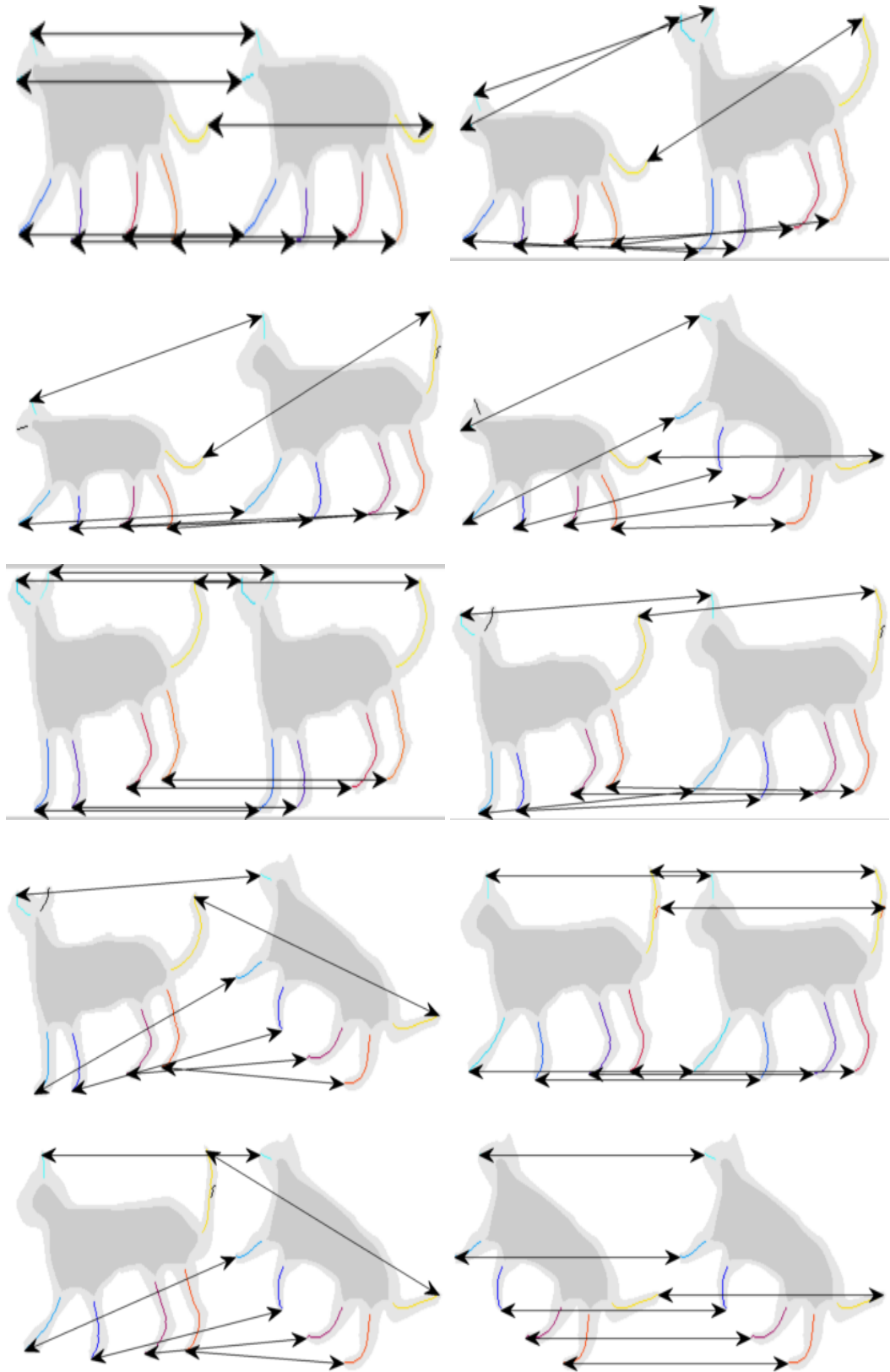


Figure B.11: Matching: Category 11

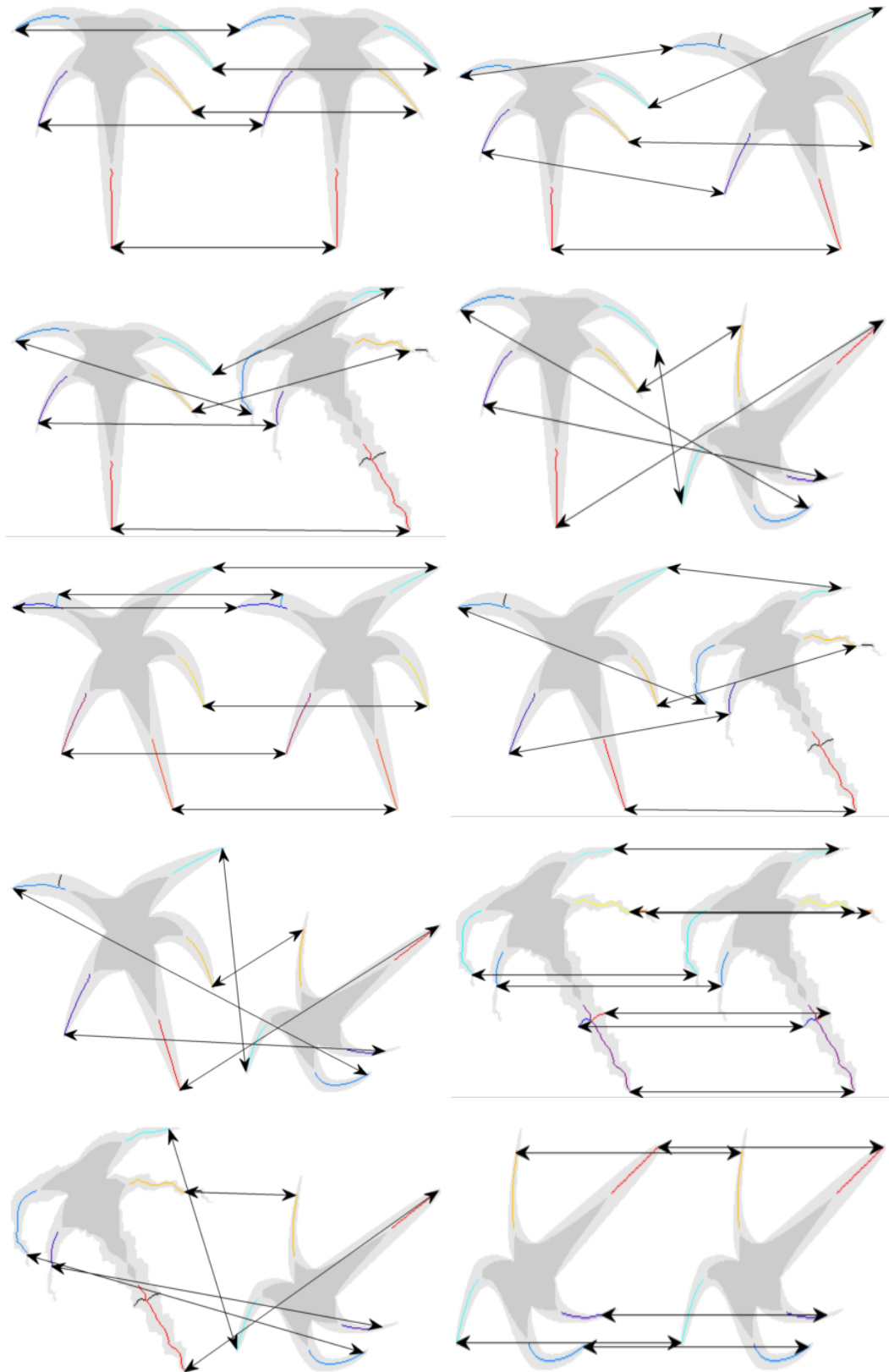


Figure B.12: Matching: Category 12

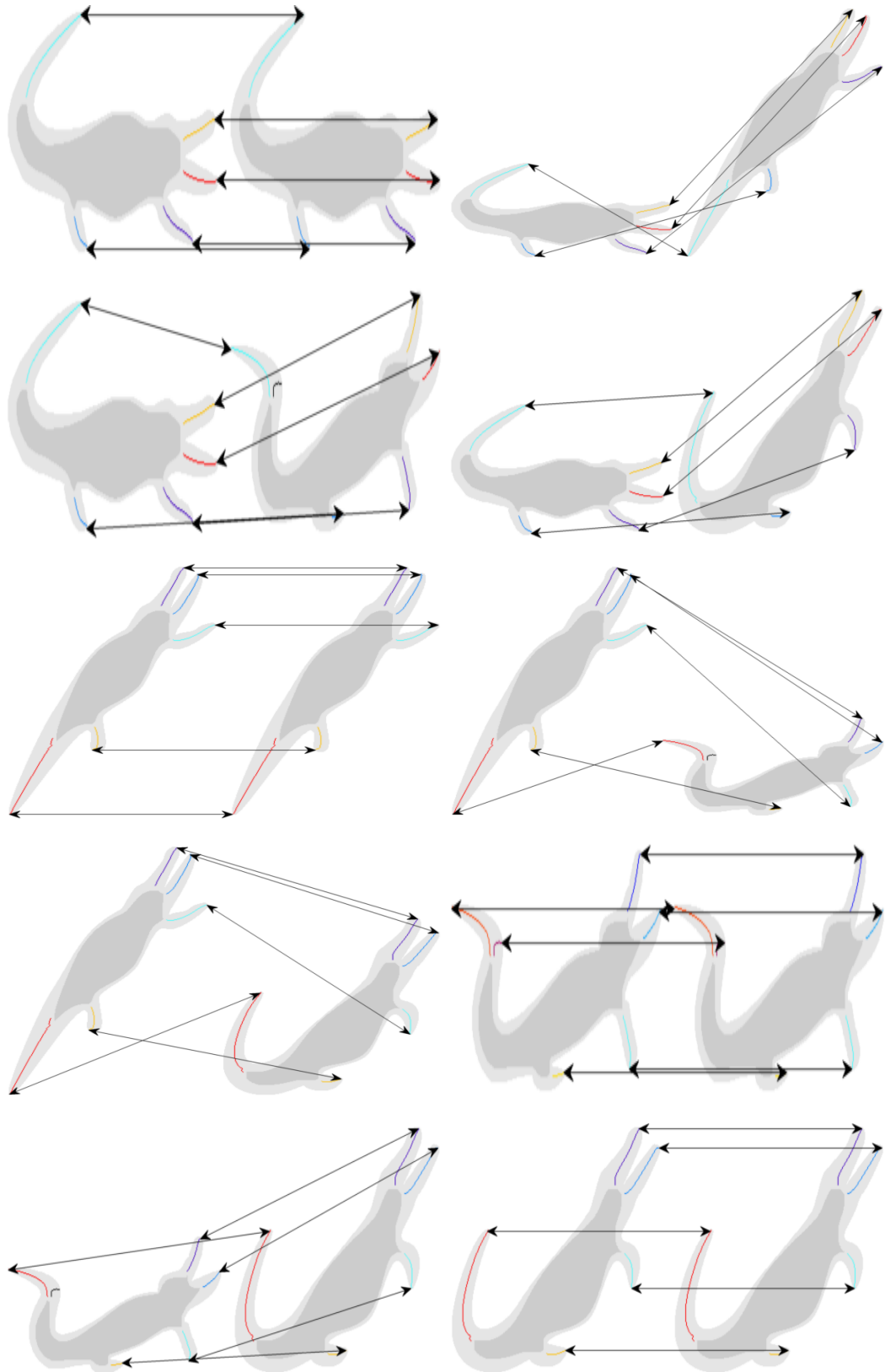


Figure B.13: Matching: Category 13

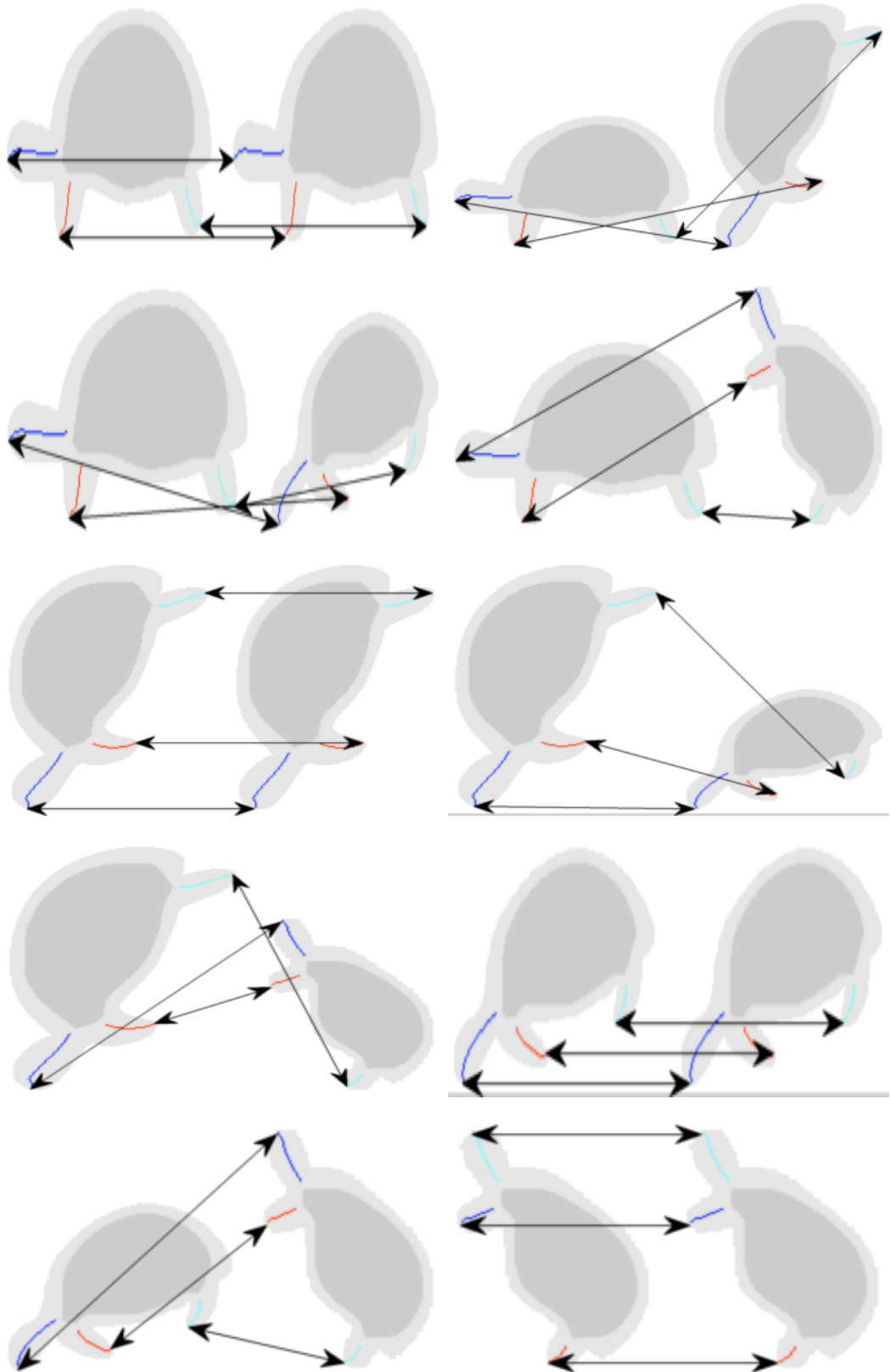


Figure B.14: Matching: Category 14

Appendix C

MATCHING GLOBAL SKELETONS: RETRIEVAL RESULTS

In this appendix, retrieval results are given for each shape in the database. Best 10 matches are shown with dissimilarities below each matching shape.

Table C.1: Retrieval results for category 1.













































										
	0.000	0.154	0.288	0.354	0.608	0.647	0.653	0.662	0.669	0.682
										
	0.000	0.154	0.248	0.491	0.624	0.626	0.646	0.655	0.659	0.686
										
	0.000	0.248	0.288	0.360	0.517	0.576	0.582	0.585	0.605	0.625
										
	0.000	0.354	0.360	0.491	0.547	0.550	0.554	0.629	0.646	0.689

Table C.2: Retrieval results for category 2.













































										
	0.000	0.223	0.264	0.272	0.293	0.297	0.309	0.324	0.400	0.424
										
	0.000	0.228	0.264	0.283	0.299	0.326	0.352	0.354	0.374	0.387
										
	0.000	0.135	0.177	0.218	0.297	0.299	0.420	0.451	0.467	0.473
										
	0.000	0.218	0.228	0.272	0.293	0.311	0.397	0.405	0.435	0.461

Table C.3: Retrieval results for category 3.

















































											
	0.000	0.096	0.150	0.281	0.554	0.675	0.689	0.695	0.726	0.734	
											
	0.000	0.150	0.201	0.359	0.547	0.700	0.732	0.736	0.742	0.746	
											
	0.000	0.096	0.201	0.272	0.550	0.629	0.663	0.669	0.685	0.710	
											
	0.000	0.272	0.281	0.359	0.646	0.781	0.794	0.800	0.819	0.834	

Table C.4: Retrieval results for category 4.

















































											
	0.000	0.098	0.346	0.347	0.588	1.085	1.456	1.456	1.637	1.786	
											
	0.000	0.098	0.266	0.332	0.639	1.125	1.483	1.483	1.591	1.763	
											
	0.000	0.332	0.346	0.439	0.837	1.341	1.737	1.852	1.852	1.902	
											
	0.000	0.266	0.347	0.439	0.708	1.109	1.496	1.496	1.561	1.822	

Table C.5: Retrieval results for category 5.

















































											
	0.000	0.315	0.334	0.350	0.442	0.464	0.524	0.610	0.865	0.883	
											
	0.000	0.298	0.342	0.350	0.353	0.545	0.556	0.571	0.839	0.885	
											
	0.000	0.156	0.316	0.334	0.342	0.382	0.412	0.456	0.669	0.790	
											
	0.000	0.156	0.298	0.315	0.321	0.417	0.419	0.455	0.626	0.730	

Table C.6: Retrieval results for category 6.

















































											
	0.000	0.052	0.145	0.158	0.443	0.461	0.471	0.485	0.489	0.588	
											
	0.000	0.052	0.157	0.180	0.455	0.476	0.482	0.501	0.517	0.618	
											
	0.000	0.145	0.157	0.242	0.426	0.427	0.438	0.443	0.457	0.546	
											
	0.000	0.158	0.180	0.242	0.466	0.475	0.478	0.486	0.487	0.505	

Table C.7: Retrieval results for category 7.













































										
	0.000	0.299	0.607	0.642	0.652	0.721	0.732	0.743	0.748	0.760
										
	0.000	0.105	0.135	0.272	0.311	0.352	0.447	0.465	0.468	0.495
										
	0.000	0.299	0.738	0.746	0.756	0.824	0.832	0.834	0.845	0.893
										
	0.000	0.105	0.177	0.223	0.272	0.283	0.385	0.423	0.432	0.449

Table C.8: Retrieval results for category 8.













































										
	0.000	0.000	0.500	1.000	1.456	1.483	1.496	1.852	2.780	2.983
										
	0.000	0.000	0.500	1.000	1.456	1.483	1.496	1.852	2.780	2.983
										
	0.000	0.168	0.443	0.485	0.487	0.499	0.517	0.532	0.537	0.551
										
	0.000	0.500	0.500	0.537	1.085	1.109	1.125	1.341	2.303	2.484

Table C.9: Retrieval results for category 9.

















































											
	0.000	0.167	0.167	0.333	0.449	0.468	0.478	0.496	0.508	0.542	
											
	0.000	0.333	0.435	0.470	0.582	0.594	0.677	0.698	0.711	0.719	
											
	0.000	0.167	0.251	0.374	0.423	0.424	0.461	0.467	0.470	0.480	
											
	0.000	0.167	0.251	0.420	0.435	0.465	0.468	0.509	0.516	0.566	

Table C.10: Retrieval results for category 10.

















































											
	0.000	0.003	0.071	0.512	0.623	0.902	0.977	0.979	0.983	1.013	
											
	0.000	0.071	0.075	0.565	0.583	0.908	0.933	0.938	0.942	0.950	
											
	0.000	0.168	0.475	0.508	0.512	0.546	0.583	0.584	0.588	0.618	
											
	0.000	0.003	0.075	0.508	0.620	0.906	0.980	0.983	0.986	1.013	

Table C.11: Retrieval results for category 11.

















































											
	0.000	0.247	0.279	0.455	0.456	0.524	0.545	0.610	0.738	0.760	
											
	0.000	0.247	0.376	0.412	0.419	0.466	0.524	0.556	0.748	0.832	
											
	0.000	0.382	0.403	0.417	0.442	0.466	0.524	0.571	0.887	0.918	
											
	0.000	0.279	0.316	0.321	0.353	0.376	0.403	0.464	0.732	0.753	

Table C.12: Retrieval results for category 12.













































										
	0.000	0.209	0.324	0.354	0.374	0.397	0.454	0.520	0.528	0.529
										
	0.000	0.374	0.387	0.400	0.405	0.423	0.432	0.480	0.490	0.499
										
	0.000	0.914	0.915	1.028	1.056	1.071	1.155	1.189	1.189	1.191
										
	0.000	0.209	0.309	0.326	0.385	0.423	0.435	0.447	0.473	0.485

Table C.13: Retrieval results for category 13.

























































































										
	0.000	0.151	0.159	0.451	0.490	0.544	0.560	0.563	0.572	0.639
										
	0.000	0.151	0.213	0.485	0.556	0.569	0.572	0.574	0.586	0.594
										
	0.000	0.460	0.490	0.574	0.626	0.669	0.697	0.733	0.737	0.805
										
	0.000	0.159	0.213	0.460	0.517	0.566	0.598	0.600	0.610	0.637

Table C.14: Retrieval results for category 14.

										
	0.000	0.063	0.075	0.081	0.438	0.471	0.482	0.486	0.551	0.625
										
	0.000	0.075	0.090	0.131	0.426	0.443	0.455	0.466	0.624	0.640
										
	0.000	0.063	0.084	0.090	0.457	0.489	0.501	0.505	0.532	0.626
										
	0.000	0.081	0.084	0.131	0.427	0.461	0.476	0.478	0.499	0.584