Hand Geometry Analysis by Continuous Skeletons

Leonid Mestetskiy, Irina Bakina, and Alexey Kurakin

Moscow State University, Moscow, Russia
Moscow State University, Moscow, Russia
Moscow Institute of Physics and Technology, Moscow, Russia
l.mest@ru.net, irina_msu@mail.ru, alekseyvk@yandex.ru

Abstract. The paper considers new approach to palm shape analysis that is based on continuous skeletons of binary images. The approach includes polygonal approximation of binary image, skeleton construction for the polygonal approximation and skeleton regularization by pruning. Skeleton of polygonal shape is a locus of centers of maximum inscribed circles. Both internal and external skeletons of palm shape are used for analysis. Segmentation of initial image, palm orientation and structure identification, fingers segmentation and characteristic points detection are performed based on image skeleton. Algorithm of color palm images binarization and computational experiments with large database of such images are described in the paper.

Keywords: Skeleton, binarization, segmentation, external skeleton, hand geometric points detection

1 Introduction

Skeleton (or medial axis representation) is a powerful and widely used tool for image shape analysis [1]. Originally, the notion of skeleton was defined for continuous objects [2]: the skeleton of a closed region in the Euclidean plane is a locus of centers of maximum empty circles in this region. While the circle is considered to be empty if all its internal points are internal points of the region.

To use skeletons for image analysis it is necessary to adapt this notion for discrete images. Many papers consider "discrete skeleton" for the purpose of shape analysis. Discrete skeleton of a binary image is an analogue of continuous one. Skeleton on this image consists of one pixel width lines, and all these lines are approximately equidistant from the edges of the source shape. There are several approaches for discrete skeleton construction: topological thinning, morphological erosion, Euclidian distance transform [3].

But discrete skeletons have significant drawbacks in comparison with continuous skeletons. First of all, the values of radial function for discrete skeleton could be calculated only approximately. In addition, pruning of noisy branches (regularization) of such skeletons is a hard problem and requires different heuristics in each particular case. And there is a problem of computational efficiency for discrete skeleton construction. At the present moment the algorithms of discrete skeleton construction are speeded up by parallelization [4]. But such speedup is limited due to the fact that number of steps of discrete skeletonization increases with the increase of image resolution.

Continuous approach for binary image skeleton construction is considered in papers [5, 6, 8]. It outperforms discrete methods in many aspects.

This paper demonstrates application of continuous skeleton for palm shape analysis. Experiments are held on Hand Geometric Points Detection Competition database [11]. Advantages of the proposed method are especially noticeable in case of poor palm images when fingers are stuck. Usage of external skeleton of the palm helps to segment such images successfully.

The paper is organized in the following way. Notion of continuous skeleton is described briefly in Section 2. The basic algorithm for palm characteristic point detection is presented in Section 3. It works in the case of "good" palm image, i.e. image in which the fingers are separable and could be easily segmented. The algorithm for "poor" image segmentation with stuck fingers is considered in Section 4. And Section 5 concludes the paper and demonstrates some experiments on HGC-2011 database.

2 Continuous Skeleton of Binary Image

Polygonal shape is a closed connected region in the Euclidean plane such that its boundary consists of finite union of closed polygons without self-intersections.

Maximum inscribed circle is a circle C that is completely lying inside the polygonal shape P and any other circle $C' \subset P$ does not contain C.

Skeleton of polygonal shape is a locus of centers of maximum inscribed circles. Notion of skeleton could be defined not only for polygons, but this is beyond the scope of the paper.

Radius of maximum inscribed circle is associated with any point of skeleton. Function $R(x)$ that maps points of skeleton to radiuses of maximum inscribed circles is called the *radial function*.

It can be proved [8] that skeleton of a polygonal shape consists of finite union of lines and arcs of the parabolas. Thus, skeleton of a polygonal shape has a dual nature. First of all, skeleton is a union of curves. In addition, it could be treated as a planar graph.

In applications it is usually required to construct a skeleton of a shape that is described by the raster binary image. In such a situation silhouette of the binary image should be approximated as a polygon. Then continuous skeleton of the polygon could be constructed. And, finally, the constructed skeleton should be pruned to get rid of unimportant branches.

Detailed algorithm of skeleton construction is described in [6], but its brief description is the following:

1. **Boundary corridor construction.**
   For initial binary image (Fig. 1a) boundary corridor is calculated based on

boundary tracing. Boundary corridor consists of two sequences of pixels: black sequence and white sequence. White sequence corresponds to internal boundary of the corridor and black to external. On Fig. 1b corridor is drawn in gray color.

2. **Construction of minimal perimeter polygon inside the boundary corridor.**
   Closed path of minimal perimeter is constructed inside the boundary corridor [5]. Such path is a closed contour without self-intersections. Physical model of the path is a rubber thread stretched along the boundary corridor. Minimal perimeter polygon is shown on 1b.
3. **Skeleton construction.**
   Skeleton is constructed for the minimal perimeter polygon. Fig. 1c shows skeleton and maximum inscribed circles at the vertices of the axial graph.

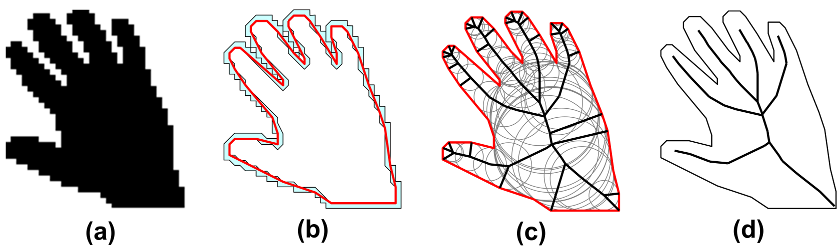

Fig. 1. The process of skeleton construction: (a) source binary image; (b) boundary corridor and minimal perimeter polygon; (c) initial skeleton of polygon, and (d) skeleton after pruning.

4. **Skeleton regularization.**
   When skeleton is constructed it is subjected to pruning. Pruning is performed by sequential cutting of certain terminal edges of the skeleton. Cutting criteria is based upon the following principle. Initial polygon could be represented as a union of all inscribed circles with the centers at the skeleton points. When skeleton edge is removed, associated circles are removed too. Union of the remaining circles forms a figure which is called silhouette of the remaining skeleton part. This silhouette is a subset of the initial polygon and it is situated inside it. If Hausdorff distance of this silhouette and initial polygon is less than the threshold then skeleton edge is cut, otherwise—not. In the example on Fig. 1d threshold is equal to 2 pixels.

Example on Fig. 1 is an illustration for the skeleton construction process. Palm size is only 33×36 pixels. Sample skeleton for real palm from Fig. 2a is presented on Fig. 2b, where palm image size is 540×420 pixels.

Internal skeleton of the palm is considered above. But skeleton of the external area of the palm could be constructed and used for shape analysis. Let S be an initial shape, and B_S be its external boundary. Consider rectangle R such

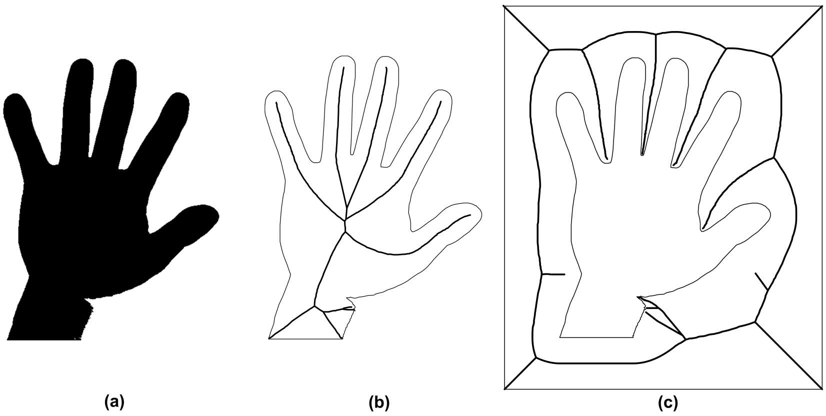

Fig. 2. (a) Binary image of palm; (b) its boundary polygon of minimal perimeter and internal skeleton, and (c) external skeleton.

that R contains entire image of the shape S. Rectangle R and curve B_S form a polygonal shape S_{ext} which is lying between R and B_S. Skeleton of the shape S_{ext} is called the *external skeleton* of S. An example of an external skeleton of a palm is depicted on Fig. 2c.

3 Characteristic Points Detection

The procedure described in the previous section allows us to obtain regularized skeleton of any binary image. Assume that we have a binary image of a palm and its skeleton. Let's consider how it can be used to locate hand characteristic points—finger tips and valleys. The process consists of the following steps:

1. Extraction of palm center
2. Extraction of fingers
3. Validation of palm
4. Detection of tips and valleys

At first, we determine the position of palm center. Then, extraction of fingers is performed. It is obvious that each finger corresponds to one of the skeleton branches. So, we extract all skeleton branches, validate them and leave only those which correspond to potential fingers. On the next step we check the correctness of the whole palm segmentation. And, finally, calculate finger's tips and valleys.

3.1 Extraction of hand center

As it is described earlier, there is an association between each skeleton point and radius of maximum inscribed circle with the center at this point. It is called radial function. Thus, we define the *center of palm* as the skeleton node with the maximal value of radial function. The illustration is given on Fig. 3a: point O is the extracted center of hand, while the circle with center at O represents the inscribed circle of maximal radius.

3.2 Extraction of Fingers

Axial graph of any binary image contains only vertices of degree one, two or three [7]. For each vertex of degree one we extract skeleton branch starting from it, going through the vertices of degree two and ending at the vertex of degree three. As a result, the list of all skeleton branches is produced. Some of these branches are structural (like those going to hand wrist) or fake (those which were not removed during pruning procedure), while others correspond to potential fingers.

The originating node of a branch is called the *top* node. The end point of a branch is called the *root* node. For example, nodes A_5 and O on Fig. 3a are top and root nodes for the branch A_5B_5O.

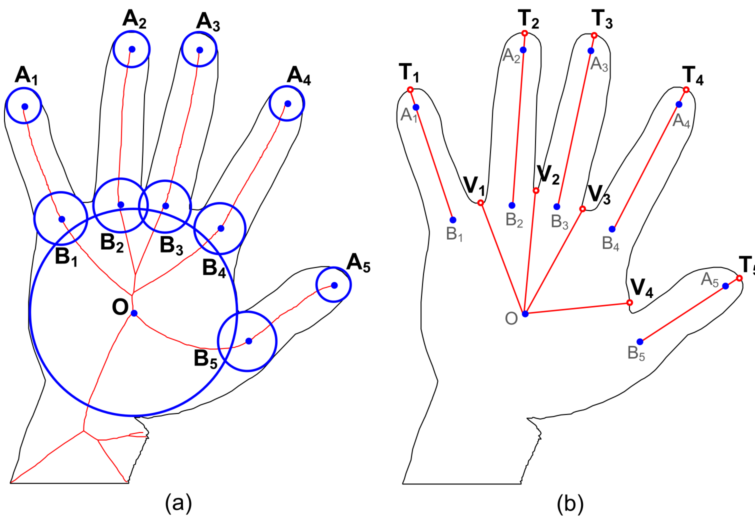

Fig. 3. Detection of (a) tops, roots and bottoms of fingers, and (b) their tips and valleys.

Now, let's consider the sequence of branch nodes from root to top node. For each node we calculate three characteristics: r is the value of radial function for this node, r_p—the value of radial function for the previous node, r_t—the value of radial function for the top node, R_{max}—the maximal value of radial function for the whole skeleton and α—the angle between two segments connecting the center of circle (associated with the node) with its tangency points. The first node in the sequence to fulfill the following conditions is considered to be the *bottom* node of the branch (as usual, the braces designate "and" condition, square brackets—"or" condition):

$$\left[\begin{array}{l} \left\{\begin{array}{l} r < 0.3R_{max} \\ \left[\begin{array}{l} r < r_p \\ \alpha > \alpha_0 \end{array}\right. \end{array}\right. \\ r < r_t \end{array}\right. \quad (1)$$

Variable α_0 is the method parameter. In this work we used $\alpha_0 = 2.7$ radians. An example of bottom node is given on Fig. 3a—it is point B_5 for the branch A_5B_5O.

So, for each skeleton branch top, root and bottom nodes are extracted. The line connecting top and bottom nodes is called the *axis* of branch. Next, we check all the branches to leave only those of potential fingers.

Let's consider one of the skeleton branches. Firstly, we calculate its length l—the total length of branch edges between top and bottom nodes and r_m—the maximal value of radial function in the neighborhood of the top node. The skeleton branch is considered to be the branch of a potential finger if:

$$\left\{\begin{array}{l} l \in [l_0, l_1] \\ r_m < r_{max} \end{array}\right. \quad (2)$$

The variables l_0, l_1 and r_{max} are method parameters. They can be set heuristically or estimated by learning. In our work they were set as $l_0 = 60$, $l_1 = 250$ and $r_{max} = 35$ pixels.

3.3 Validation of Hand

As a result of the previous step we have the list of the branches which correspond to potential fingers. If their number is less than 5 (total number of hand fingers) we admit that hand was segmented incorrectly, so tips and valleys cannot be extracted. If the number is equal to 5 we proceed to the next step. However, there are situations when the total number of branches is greater than 5. In such a case special procedure is applied to remove extra skeleton branches. It is based on the analysis of hand structure.

Firstly, it should be noted that all skeleton branches in the list are arranged according to the traversal of axial graph. Due to this fact we can analyze only the sequences formed of 5 successive branches. The sequence with the minimal angle between axes of its first and last branches is considered to be the sequence of fingers branches.

So, fingers branches are determined. The last thing is to establish correspondence between each finger (little, ring, etc.) and its branch. For this purpose we calculate the angles between axis of each finger and axis of its previous finger in the sequence. And then we select the branch with the maximal angle and announce it to be the branch of thumb.

Below in the text, top, root, bottom and axis of finger branch are called top, root, bottom and axis of finger respectively.

3.4 Detection of Tips and Valleys

Let's consider a palm with extracted fingers. Denote by A_i the top nodes and by B_i the bottom nodes of fingers, $i = 1, \ldots, 5$ (see Fig.3b). The lines $A_i B_i$ are fingers axes.

For each finger its *tip* T_i is the point of intersection between half-line $B_i A_i$ and hand boundary. The *valley* between two fingers i and $i + 1$ is the point V_i that is the nearest point of hand boundary between T_i and T_{i+1} to the hand center O, $i = 1, 2, 3$. The initial position V_4' of valley V_4 between thumb and index finger is calculated by the same rule, but then additional correction is applied. It is totally heuristic.

We go over the sequence of hand boundary points from V_4' to T_5 and calculate $d_4 = |PB_4|$ and $d_5 = |PB_5|$, where P is one of the considered points. If r_5 is the value of radial function for the node B_5, then the first point P that fulfills the conditions $d_4 > d_5$ and $d_5 < 2.1r_5$ is called the valley V_4. Fig. 3b shows extracted tips T_i and valleys V_i for hand.

4 Hand Image Segmentation

The tips and valleys detection algorithm based on the analysis of internal skeleton of palm binary image produces good results only if the fingers are separable, like on Fig. 2a. But in a case of fingers touching each other (Fig. 4a) the obtained binary image has indistinguishable fingers. So, direct construction of skeleton would not work. However, the problem can be solved if continuous skeleton approach is combined with other image processing methods. The idea is to enhance image on Fig. 4d by reducing it to the image on Fig.4h. Let's consider basic operations that are used in our algorithm.

Denote by

V_{rgb} the space of RGB color images;
V_{gs} the space of grey-scaled images;
V_b the space of binary images;
V_{skel} the space of continuous skeletons.

Firstly, we introduce the operations

Red: $V_{rgb} \to V_{gs}$ that extracts red component from RGB color image;
Sob: $V_{gs} \to V_{gs}$ that is Sobel operator with kernels
$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{pmatrix}$;
Bin: $V_{gs} \to V_b$ that produces binary image from gray-scaled imaged based on thresholding;
Neg: $V_b \to V_b$ that produces negative image for binary image;
Skel: $V_b \to V_{skel}$ that constructs continuous skeleton from binary image;
Prun: $V_{skel} \to V_{skel}$ that performs pruning of skeleton;
Silh: $V_{skel} \to V_b$ produces binary image that is silhouette of skeleton.

Let H (Fig. 4a) be the initial color palm image: $H \in V_{rgb}$. The segmentation algorithm can be defined as follows.

1. $H_1 = Red(H)$, $H_1 \in V_{gs}$. Red component (Fig. 4b) is extracted from color image H.
2. $H_2 = H_1 - Sob(H_1)$, $H_2 \in V_{gs}$. The grey-scaled image H_1 (Fig. 4b) is subjected to Sobel operator and subtracted from itself. The resultant image H_2 is shown on Fig. 4c.
3. $H_3 = Bin(H_2)$, $H_3 \in V_b$. Binarization of H_2 (Fig. 4c) produces H_3 (Fig. 4d).
4. $G = Skel(H_3)$, $G \in V_{skel}$. Continuous skeleton G (Fig. 4e) is constructed for binary image H_3.
5. $G_1 = Prun(G)$, $G_1 \in V_{skel}$. Skeleton G is subjected to pruning of two types. Firstly, all the terminal edges having value of radial function less than 3 pixels are removed. Secondly, regularization (see Section 2) with threshold equals to 6 pixels is performed. The result of this step is shown on Fig. 4f.
6. $H_4 = Silh(G_1)$, $H_4 \in V_b$. The silhouette H_4 (Fig. 4g) of pruned skeleton G_1 is constructed.
7. $H_5 = Neg(H_4)$, $H_5 \in V_b$. The negative image H_5 (Fig. 4h) for the image H_4 is produced.

Thereby, we obtain binary image of palm from its initial color image. Generally, the image has separable fingers and can be subjected to characteristic points detection process described in Section 3. The results of segmentation for the considered color palm image on Fig. 4a are shown on Fig. 4i.

It should be noted, that the proposed algorithm could be applied to any palm image (either "good" with separable fingers as on Fig. 2a or "bad" as on Fig. 4a). So this algorithm along with algorithm from section 3 gives us entire procedure for fingers' tips and valleys detection.

5 Implementation and Experiments

Experiment was carried out for HGC-2011 competition [11]. In the competition it was required to detect 5 tips and 4 valleys for each provided palm image. Source data included training set with 300 palm images. It was used for manual tuning of algorithm parameters (binarization and pruning thresholds). Quality of point's detection was calculated on testing set (160 palm images), which was unavailable before the publication of results.

According to the competition rules point is detected correctly if the distance between the ground truth point position and the detected point position is less than 20 pixels. Each incorrect detection results in 1 point penalty for participating algorithm, and each refuse of classification results in 0.7 points penalty (i.e. refuse of entire palm classification results in 6.3 points penalty).

In HGC-2011 competition, our algorithm showed the best result among the 15 registered participants. Second, third and forth algorithms got 57.3, 77.9 and 98.1 penalty points on testing set. Detailed results of our algorithm are presented in the table 1.

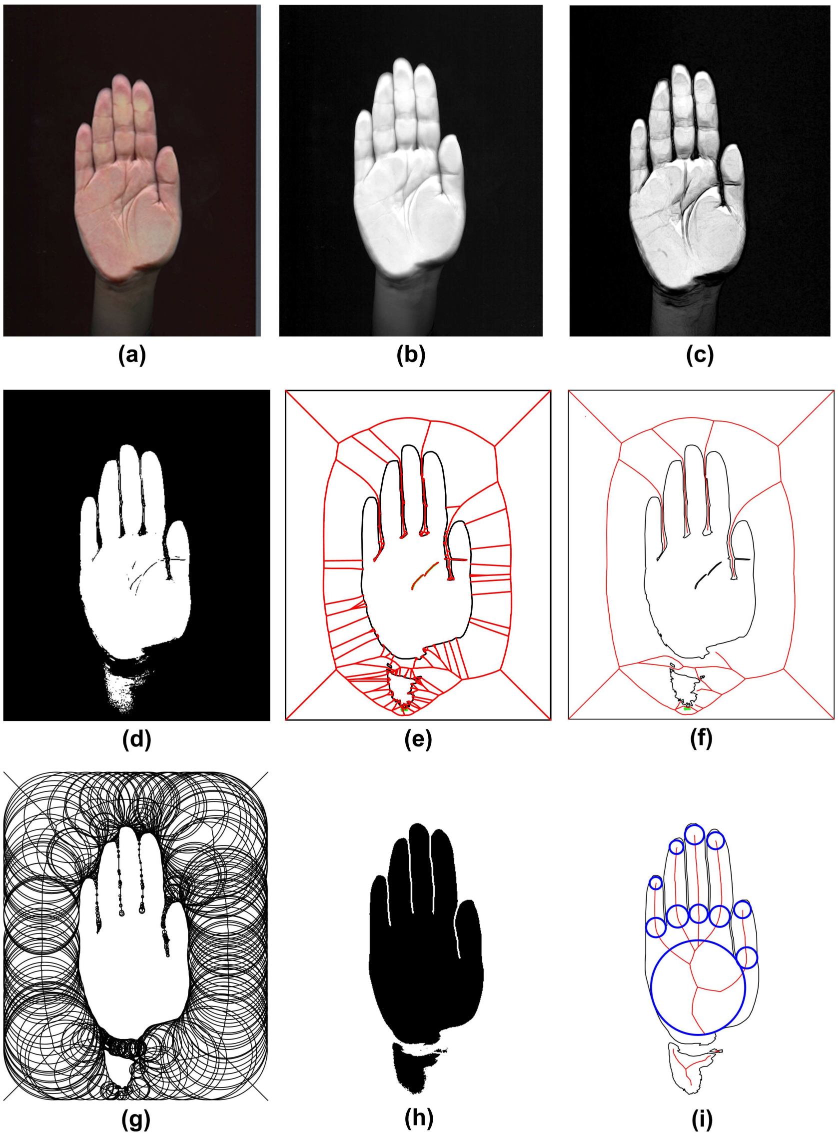

Fig. 4. Palm image segmentation steps.

	Detection rate	Running time (sec)	Penalty
Training set (300 palms, 2700 points)	2682 of 2700 (99.33%)	58	18
Testing set (160 palms, 1440 points)	1415 of 1440 (98.26%)	34	22.3

Table 1. Point detection results on HGC database

Moreover similar approach for palm shape analysis was successfully used in systems for biometric person identification [9] and hand gesture recognition [10], which demonstrates rich capabilities of skeleton based shape analysis.

Acknowledgments. The authors thank the Russian Foundation for Basic Research for the support on this study (grants $10-07-00609$ and $11-01-00783$), and the organizing committee of Hand Geometric Points Detection Competition (HGC-2011) [11] for the provided database of hand images used in our experiments.

References

1. Siddiqi K., Pizer S. M.: Medial Representations: Mathematics, Algorithms and Applications. Springer (2008)
2. Blum, H.: A Transformation for Extracting New Descriptors of Shape. In Proc. Symposium Models for the Perception of Speech and Visual Form, MIT Press Cambridge MA, 1967
3. Costa, L., Cesar, R.: Shape Analysis and Classification. CRC Press (2001)
4. Strzodka, R., Telea, A.: Generalized Distance Transforms and Skeletons in Graphics Hardware. In: Proc. Eurographics IEEE TCVG Symp. Visualization, pp. 221–230 (2004)
5. Mestetskiy, L.: Continuous Skeleton of Binary Raster Bitmap (in Russian). In: Graphicon'98, International Conference on Computer Graphics, Moscow (1998)
6. Mestetskiy, L. and Semenov, A.: Binary Image Skeleton—Continuous Approach. In: Proc. of 3rd Int. Conf. on Computer Vision Theory and Applications, vol. 1, pp. 251–258. INSTICC Press (2008)
7. Mestetskiy, L.M.: Continuous Morphology of Binary Images: Figures, Skeletons and Circulas (in Russian). FIZMATLIT, Moscow (2009)
8. Mestetskiy, L.: Skeleton Representation Based on Compound Bezier Curves. In: Proc. of 5th Int. Conf. on Computer Vision Theory and Applications, vol. 1, pp. 44-51. INSTICC Press (2010)
9. Bakina, I.: Palm Shape Comparison for Person Recognition. In: Proc. of 6th Int. Conf. on Computer Vision Theory and Applications. Portugal (2011)
10. Kurakin, A., Mestetskiy, L.: Hand Gesture Recognition through On-line Skeletonization. Application of Continuous Skeleton to Real-Time Shape Analysis. In: Proc. of 6th Int. Conf. on Computer Vision Theory and Applications. Portugal (2011)
11. Magalhães, F., Oliveira, H. P., Matos, H. and Campilho, A. (Published: 2010, Dez 23). HGC2011—Hand Geometric Points Detection Competition Database (Available at: `http://paginas.fe.up.pt/~hgc2011/`)