STANDARDIZATION OF GEOMETRICAL CHARACTERISTICS IN GESTURE RECOGNITION

Andrew Golik

***Abstract**: This paper covers approach to obtaining geometrical characteristics, in particular defects, for gesture shown in front of web camera. Variant of standardization of defect`s structure is suggested. The paper provides detailed algorithm that allows obtaining standardized feature vector for defect of any size and structure. It is suggested to use obtained feature vectors for recognition of tactile sign language.*

***Keywords**: gesture recognition, defects, standardization, normalization, tactile sign language.*

***ACM Classification Keywords**: I.2 Artificial Intelligence, I.4 Image Processing and Computer Vision, I.5 Pattern Recognition, G.1.3 Numerical Linear Algebra.*

Introduction

Recognition of tactile sign language is an important applied task. There exist people who need such recognition systems not for entertainment, but for everyday life. They are deaf and deaf-and-dumb people. In case when person lost hearing at mature age usually it is possible to continue full-fledged life, but children, who lost hearing in the early childhood, are in much more difficult situation. It is an important social task to help such children. Of cause, availability of money for hearing aids and other equipment partially solves the problem, but children in many orphanages remain without appropriate support. So developing of recognition system, which does not require expensive devices like sensors and can be used for low cost, is an actual task. [1]

Principle of allocation and highlighting of a contour of hand on captured by web camera image is given. Geometrical characteristics, which can be effectively used for gesture recognition, namely defects, are covered. Algorithm of obtaining standardized feature vectors for defects is suggested.

Finding and highlighting of a contour of hand

The majority of approaches to finding of a contour of hand on an image are connected with pixel analysis. Usually two color models, RGB and HSV, are used. RGB is an additive color model in which red, green, and blue light are added together in various ways to reproduce a broad array of colors. HSV is one of most common cylindrical-coordinate representations of points in an RGB color model. HSV stands for hue, saturation, and value. Both models have advantages and shortcomings. However, HSV model is preferable in gesture recognition. This model allows considering hue as separate component. It is important for situations when finding a contour of hand is connected with finding areas that correspond to color of skin. In this case, smoothing and filtration, which helps to get rid of noises, are applied too. This task is not trivial because of environment conditions (lighting, background and so on) which have direct influence on results of recognition. Existing systems require detailed configuration before someone can use them in certain environment. This problem can be partially solved with usage of a red mitten.

The most useful for implementation of mentioned above functionality is OpenCV library. It is written on C++, but there are many wrappers, which allow using it in different programming environments, for example Java (JavaCV) or C# (EmguCV). The library offers great capabilities for image analysis, processing and smoothing, finding of contours and so on. While developing of recognition system for tactile sing language, stage of finding of a contour of hand was implemented. Illustration of corresponding results is shown in Figure 1. [1]

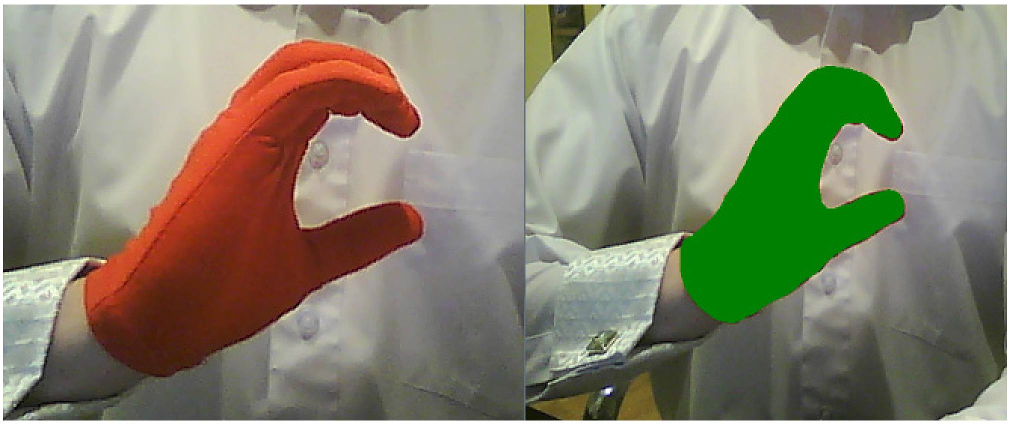

Figure 1. A contour of hand is found and highlighted.

Convex hull and defects

Information about highlighted part of an image is presented as a closed contour that is a formalized description of a captured object. The most useful information is stored in "defects". Defects are geometrical characteristics of gesture, which allow identifying of tactile sign. They are characterized by relative positioning of a contour of hand and convex hull built for this contour. In addition, defects can be considered as "curvilinear triangles" with (blue) parts of convex hull as basis (Figure 2).

Highlighted contour of hand and convex full for this contour is shown in Figure 2. The figure represents four defects, start, end and depth points of each of them. However, there are only "significant" defects on the figure and there exist many small defects most of which even cannot be seen at the first sight.

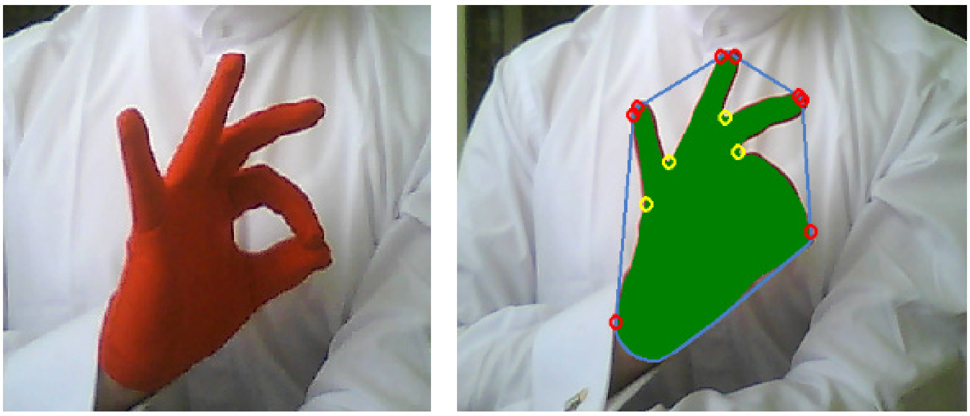

Figure 2. Finding of a contour of hand, convex hull and defects

Defects are sorted in order to find "significant" defects, which can be effectively used for recognition. This stage is rather complex because criteria of importance of defect are not trivial. Following parameters can be used for sorting: length and depth of defect, relation of length to depth and so on. The most efficient approach is to use rules like "if length more than A, depth more than B and relation of length to depth is more than C then defect is significant". [1, 2]

Formation of standardized feature vectors for structure of defect

As mentioned in previous section there are numerical characteristics which can be obtained for defects, namely length, height, area, perimeter and so on. In addition, usage of relative values shows their efficiency on practice. However, structure of defect deserves special attention. The paper provides an approach to formation of feature vectors that represent structure of defect in the most adequate way.

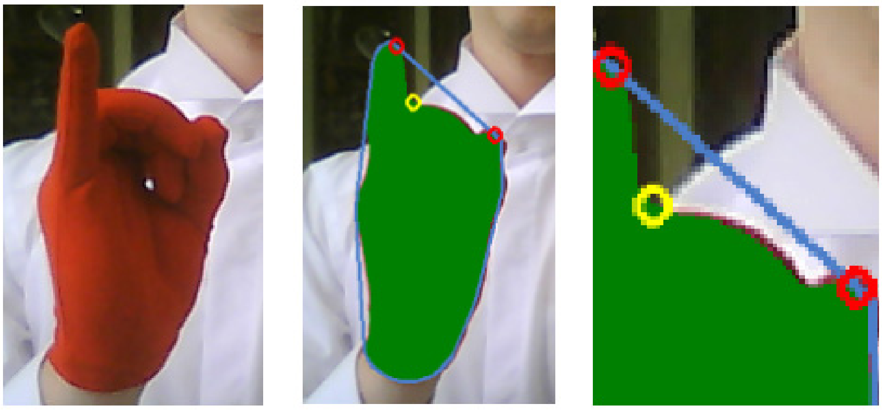

Figure 3. Tactile sign "I" with highlighted defect.

Gesture with one significant defect is shown in Figure 3. Structure of defect is described by positions of points of a contour within it. It should be represented in such way that would allow not only displaying positions of mentioned points, but also comparing structures of different defects and finding similarity degree between them. Algorithm of formation of feature vectors, which meet all the requirements, is given below:

1. Obtain equation of line that passes through start $c_{start}(x_1, y_1)$ and end $c_{end}(x_2, y_2)$ points of defect:

$$(y_2 - y_1)x + (x_1 - x_2)y + (x_2 y_1 - x_1 y_2) = 0;$$

$$\mathrm{A} = y_2 - y_1; \mathrm{B} = x_1 - x_2; \mathrm{C} = x_2 y_1 - x_1 y_2;$$

$$\mathrm{A}\,x + \mathrm{B}\,y + C = 0;$$

2. Having coordinates of points of a contour within defect $c_i(x_{i,} y_i), i = \overline{start, end}$ for each point obtain distance between this point and line from step 1 (Figure 4).

$$d_i = \frac{|\,\mathrm{A}\,x_i + \mathrm{B}y_i + \mathrm{C}\,|}{\sqrt{(\mathrm{A}^2 + \mathrm{B}^2)}}, i = \overline{start, end};$$

Numerical vector is obtained $d = (d_{start}, ..., d_{end})$. Size of this vector d_{size} equals to amount of points of a contour within defect.

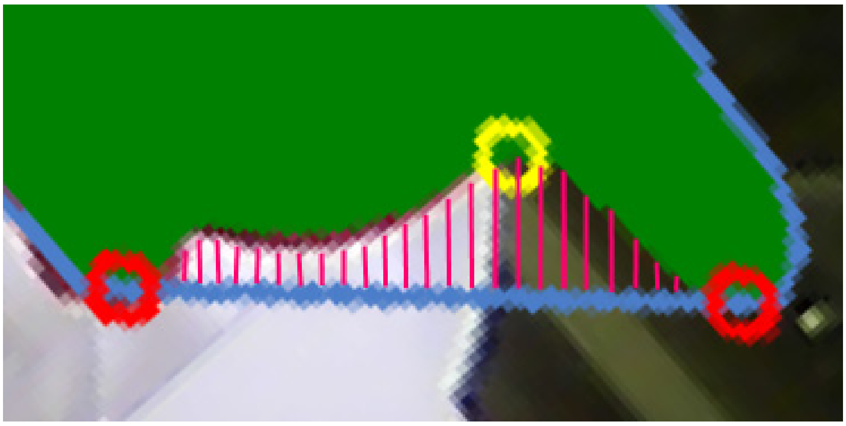

Figure 4. Generalized illustration of second step of the algorithm

3. Standardize vector d .

Third step of the algorithm should be covered more detailed. First, necessity in standardization appears because of different dimensions d_{size} of feature vectors. It is naturally that amount of points of a contour within defect differs for each demonstration of gesture. In order to solve this problem following algorithm of standardization of feature vector is suggested:

1. Set etalon dimension $d_{etalonSize}$ for feature vectors.
2. Compare dimension of current feature vector d_{size} with $d_{etalonSize}$. If $d_{size} \geq d_{etalonSize}$, move to step 3 a), else move to 3 b).
3. a) Divide dimension of current feature vector d_{size} by $d_{etalonSize}$. Round result to bigger number $amount = ceil(d_{size} / d_{etalonSize})$. Starting from the first element of vector d group elements in groups of size $amount$ (or less in case when there are not enough elements).

$$
\begin{aligned}
&\text{set}_1 = \{d_{start}, \ldots, d_{start+amount-1}\}, \\
&\text{set}_2 = \{d_{start+amount}, \ldots, d_{start+j*amount-1}\}, \\
&\ldots \\
&\text{set}_j = \{d_{start+(j-1)*amount}, \ldots, d_{start+j*amount-1}\}, \\
&\ldots \\
&\text{set}_{etalonSize} = \{d_{start+(d_{etalonSize}-1)*amount}, \ldots, d_{start+d_{etalonSize}*amount-1}\};
\end{aligned}
$$

Find average value for elements in each group:

$$
v_j = \frac{(d_{start+(j-1)*amount} + \ldots + d_{start+j*amount-1})}{amount},
$$
$$
j = \overline{1, amount};
$$

b) Increase quantity of feature vector elements on $d_{size} - 1$ by inserting additional elements between existing ones. Values for created elements are obtained by calculating average for two adjacent elements.

$$v_1 = d_{1,}$$
$$v_2 = \frac{(d_1 + d_2)}{2}, v_3 = d_2,$$
$$v_4 = \frac{(d_2 + d_3)}{2}, v_5 = d_3,$$
$$...$$
$$v_k = \frac{(d_{k/2} + d_{k/2+1})}{2}, v_{k+1} = d_{k/2+1},$$
$$...$$
$$v_{2d_{size}-2} = \frac{(d_{d_{size}-1} + d_{d\ size})}{2}, v_{2d_{size}-1} = d_{d\ size};$$
$$(\mathrm{k}/2+1) \le d_{size};$$

If $2d_{size} - 2 \ge d_{etalonSize}$ move to step 3 a), else repeat step 3 b).

4. Normalize feature vector by dividing all its elements by maximal element of the vector:

$$v_{\max} = \max_{i=1,d_{etalonSize}}(v_i),$$
$$v_i = \frac{v_i}{v_{\max}},$$
$$v = (v_1, v_2, ..., v_i, ..., v_{d_{etalonSize}}), v_i \in [0,1];$$

Normalization can solve problems that are connected with different distances between hand and web camera so this step is necessary.

Conclusion

Overall, standardized feature vector has etalon dimension $amount$ and its elements belong to interval from 0 to 1 (Figure 5). The feature vector can be used for finding similarity degree between defects. It is suggested to use Euclidian, ellipsoidal or orthogonal compliance distances with obtained feature vector [3].

Figure 5. Illustration of standardized feature vector

Bibliography

1. Donchenko V.S., Golik A.O. Usage of matrix feature vectors in recognition of tactile sing language. // SAIT 2013 System Analysys and Information Technologies, May 27-31, 2013, Kyiv, Ukraine.

2. Golik A.O. Recognition of tactile sign language: designing the feature vectors based on amount and structure of defects. XIX International Conference “Problems of Decision Making under Uncertainties”(PDMU -2012),– April 23-27, 2012.– Mukachevo, Ukraine, ст. 83.

3. Donchenko V.S., Golik A.O. Usage of ellipsoidal and orthogonal distances in speech and gesture recognition. // “Bulletin of Taras Shevchenko National University of Kyiv. Series: Physics & Mathematics”, №1, 2013.

4. Donchenko V.S., Golik A.O. Matrix feature vectors in speech and gesture recognition // International Journal “Information Theories and Applications”, Vol. 20, Number 2, 2013, p.143-155.

5. Golik A.O. Designing the feature vectors: implementation of speech recognition algorithms using parallel computing capabilities. Visn., Ser. Fiz.-Mat. Nayky, Kyiv Univ. im. Tarasa Shevchenka. – 2012. – N 3. – P. 141-150. (in Ukrainian).

7. Donchenko V.S., Golik A.O. Using of parallel computing technology in speech recognition problems // XVII International Conference “Problems of Decision Making under Uncertainties”(PDMU -2011),– May 23-27, 2011.–Skhidnytsia, Ukraine.

8. .Donchenko V.S., Golik A.O. Implementation of algorithms of speech recognition using parallel computing capabilities // XVI International Conference “Problems of Decision Making under Uncertainties”(PDMU -2010),– October 4-8, 2010.– Yalta, Ukraine.

Authors' Information

Andrew Golik *– post-graduate student, Taras Shevchenko National University of Kyiv, Department of System Analysis and Decision Making Theory, e-mail: analytical@ukr.net*

Major Fields of Scientific Research: computer vision (especially speech and gesture recognition), artificial intelligence, pattern recognition, algorithms, web development.