A new approach to machine contour perception based on the human visual system

Vassilios Vonikakis[1], Ioannis Andreadis[1] and Antonios Gasteratos[2]

1. Laboratory of Electronics,
Section of Electronics and Information Systems Technology,
Department of Electrical and Computer Engineering,
2. Laboratory of Robotics and Automation,
Section of Production Systems,
Department of Production and Management Engineering,

Democritus University of Thrace,
GR-67100 Xanthi,
Greece.

This paper presents a new shape descriptor suitable for machine contour perception, that mimics the orientation selective cells of the human visual cortex. It extracts a description for any 2-dimensional shape, regardless of its size, rotation and position, with affordable computational cost. Furthermore, it can identify different shapes of the same object, distorted shapes and heavily contaminated shapes with up to 20% spike noise. Special queries are also possible, for the detection of certain shape patterns in their contours. The extracted shape description can be used in many applications, such as image understanding, pattern recognition, autonomous systems and shape-based image retrieval systems.

The earliest stage of visual processing in the Human Visual System (HVS) is located in the primary visual cortex in an area known as V1. V1 cells are biological orientation filters, meaning that they respond only to edges with certain orientation. There are 12-24 different groups of such cells in every point of the visual field, each one specialized on the detection of a particular orientation. A stimulated V1 cell carries, among others, two basic pieces of information; its location in the visual field and the orientation that detects. Every point of the visual field contains all 24 groups of V1 cells, in order to detect every possible edge direction. The role of these cells is of great importance for the perception of shape, since they are directional filters, identifying changes in intensity of any direction and at any point of the visual field. Cells from higher levels of the HVS, receive whole visual maps of stimulated V1 cells, and identify certain combinations among them, resulting to the gradual perception of more complex forms.

The algorithm

Similarly to the HVS, the proposed shape descriptor uses 12 groups of artificial V1 cells each one specialized in detecting any edge with a particular orientation within its receptive field. Every cell is a 10×10 kernel containing an oriented straight line segment. Every group comprises 5 cells with the same orientation and different position of the oriented segment, resulting to a set of 60 artificial cells (Table 1). The orientations of the 12 groups are 0º, 15º, 30º, 45º, 60º, 75º, 90º, 105º, 120º, 135º, 150º and 165º. Every cell has a slight tolerance, meaning that can be also stimulated by edges with orientations differing by some degrees from its main orientation. The main stages of the algorithm are:

1. The image is divided to 10×10 non-overlapping regions. In each region all 60 cells are tested to find the one with the higher stimulation. The stimulation process is similar to the biological V1 cells and counts the number of edge pixels of the particular region that fall only inside the straight segment of each cell. If not appropriate stimulation is found in any cell, no edges are contained in the particular region. The whole process aims in finding which cell (orientation filter) describes better the edges of every image region, and thus, knowing the inclination of the particular edge at that particular section of the image. Previous techniques convolved series of orientation filters with the image which was computationally intensive. The proposed approach does not use

convolution, since not all pixels of the kernel are used in the calculation of its stimulation (only the pixels of the straight line segment) and the kernels are not shifted to every possible position in the image (only to non-overlapping regions). This process demands less computational effort and higher parallelism due to the fact that there are no dependences in the analysis of every region.

2. The main objective of the shape descriptor is to record the relative angles that are formed from the orientations of the stimulated kernels moving in a clockwise fashion on the contour of the shape. In order to do so, the algorithm searches for a kernel which will be the starting point. It then moves clockwise and calculates the angle that is formed between the orientation of the current kernel and the orientation of the following one. The process continuous until a loop has been completed and the starting kernel is reencountered.

3. The proposed shape description is a chain of the angles formed by the orientations of the stimulated kernels of a certain shape, when moving clockwise. Its novelty is that no proportions in straight line segments are recorded, but only relative angles. This is an important advantage for many image understanding algorithms, because it provides scale and rotation invariance. Additionally, an appropriate normalization of the chain provides interpretation invariance (invariance of the starting kernel). A special metric, examining the dissimilarity of two descriptions, provides an appropriate measure for the comparison of two shapes. For the purposes of shape-based image retrieval, a comparison rule is employed, based on the aforementioned metric, which provides a crisp decision on whether two shapes are similar. Generally, if the result of the metric is less than the length of the chain, then the two shapes match.

Advantages

The method was extensively tested with binary edge images up to 1,300×1,300 pixels, containing various shapes, at different scales and rotations. Experimental results confirm that the method achieves scale, rotation and interpretation invariance. The fact that no proportions are recorded, and only angles, makes the shape descriptor quite flexible for the retrieval of different shapes that derive from the same object. Particularly, the method is capable of identifying as similar, different shapes of the same object. This is depicted in Fig. 1, where all shapes have dissimilarity less than 7 (which is the length of their description) and for that reason they are classified by the comparison rule as being similar. Furthermore, for the same reasons, stretch distortions on both axes of the shape, do not almost affect the relative angles and result to small changes in dissimilarity metric, thus not affecting the final result. Another advantage that derives from the form of the shape description is the fact that the proposed shape descriptor can search for certain patterns in the contour of a shape or partial similarities, rather than searching for a match with the whole contour. An important capability of the new shape descriptor is its ability not to be affected by noise, and additionally to reconstruct the original noisy image to a new filtered one. Noisy images with up to 20% spike noise can be handled successfully. The reconstruction can be achieved by tiling the stimulated kernels of every image region, making them part of the actual image and not useful only for the computation of edge orientations. The computational demands of the proposed method are quite affordable considering its advantages. Implemented in C code and executed by an Intel Celeron Processor, running at 1 GHz, the typical execution time for an image of 700×700 pixels, is approximately 0.5 seconds and for images less than 400×400 pixels the execution time is unnoticeable to the human observer.

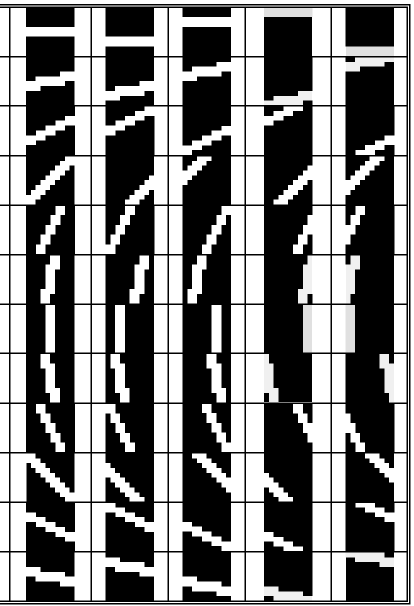
Group 1	0º	
Group 2	15º	
Group 3	30º	
Group 4	45º	
Group 5	60º	
Group 6	75º	
Group 7	90º	
Group 8	105º	
Group 9	120º	
Group 10	135º	
Group 11	150º	
Group 12	165º	

Table 1

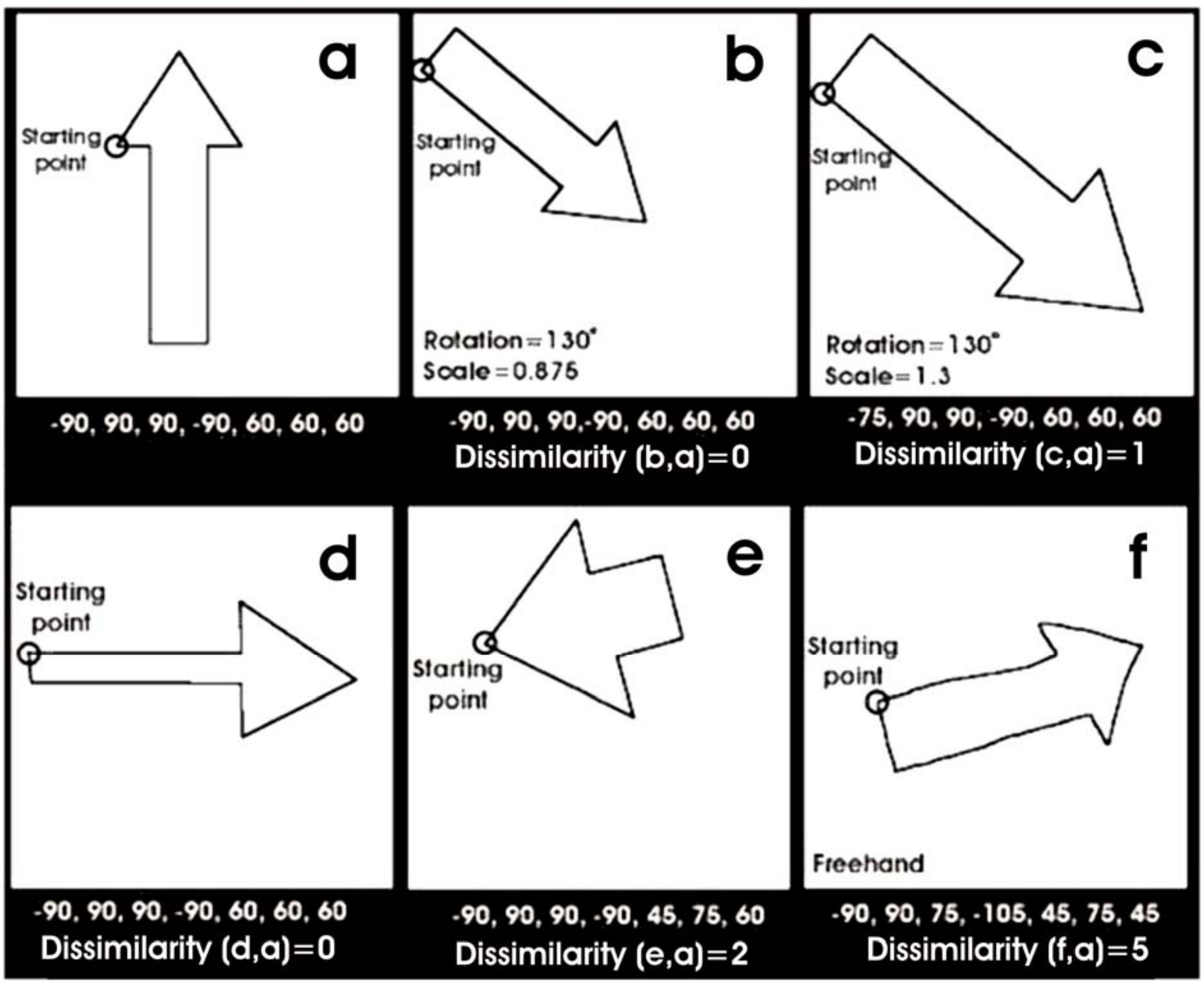


Fig. 1