

Perceptual Organization of Directional Primitives Using a Pseudocolor Hough Transform

Marta Penas¹, María J. Carreira², and Manuel G. Penedo¹

¹ Dpto. Computación. Fac. Informática, Universidade da Coruña. 15071 A Coruña. SPAIN. {infmpc00,cipenedo}@dc.fi.udc.es

² Dpto. Electrónica e Computación. Universidade de Santiago de Compostela. 15782 Santiago de Compostela. SPAIN. mjose@dec.usc.es

Abstract. This paper describes a computational framework developed for the extraction of low-level directional primitives present in an image, and subsequent organization using the laws of perceptual grouping. The system is divided in three stages. Extraction of the directional features in the image, through an efficient implementation of Gabor wavelet decomposition. Reduction of these high dimensionality results by means of growing cell structures. And extraction of the segments from the image by means of a Fuzzy Hough Transform.

Keywords: Perceptual primitives, Gabor wavelets, growing cell structures, chromaticity diagram, Hough transform.

1 Introduction

In human vision, perceptual organization refers to the human visual system basic capability to derive relevant groupings and structures from an scene without prior knowledge of its contents. In computer vision, perceptual organization is the study of how features are clustered prior to object recognition.

The main goal of perceptual organization is object recognition, which is basically a searching problem. Perceptual organization will be crucial for reducing the size of this search. In this frame, the most important functions of perceptual organization are [1]: segmentation, or division of the image into sets of related features; three-space inference, as perceptual organization results in the formation of two-dimensional relations between image features that can lead to specific three dimensional interpretation; and indexing of world knowledge, as in large databases a very important factor determining the size of the searching space is the selection of the appropriate object from the set of possibilities.

The final goal of our work is developing a computational framework that integrates the relationships offered by Gestalt psychology (parallelism, continuity, similarity, symmetry, common region and closure) among extended tokens to form larger groups, as the significance of large organizations is higher than a small organized form. All the approaches to this goal [2,3] are based on the

results extracted from traditional edge detectors. As stated in a previous paper [4], despite the introduction of the computationally expensive process of Gabor wavelet decomposition, the perceptual organization system's performance can improve: the results obtained by this process provide global information about the orientation of the pixels in the image, instead of local information provided by traditional edge detectors. They are also independent of the kind of images processed, as opposed to classical edge detectors which need parameter tuning depending on the kind and quality of images. Beside these advantages, we have used an approximation to Gabor wavelets [4] in order to reduce the computation time by using a pyramidal multiscale Gabor wavelet transform in the spatial domain, which is faster than conventional FFT implementations. Thus, an optimized spatial-domain implementation was developed, using 1D, 11-tap filter masks to approximate the original filters. This is a very efficient strategy and reduces computational costs while keeping high fidelity.

We have modified the scheme in [4] by including the important property of Gabor filters of good edge localization in the computation of each component of the gaborjet. Thus, Gabor wavelet decomposition results in a set of gaborjets, one for each frequency channel, composed of 9 components. Reduction of the dimensionality of this input space is necessary. To this end, auto-organized structures based on artificial neural networks will be employed. At first [4], we employed self-organized maps, but later [5], three different auto-organized structures were studied and compared: self-organized maps, growing cell structures (GCS) and growing neural gas. GCS was selected, as it was the structure that produce the best results.

The results are RGB images where each color represents one of the possible orientations considered and intermediate colors represent the distance to each of the 8 main orientations

In this paper, we develop the following step in our perceptual grouping process, the detection of segments in the image by means of a Fuzzy Hough Transform in which each pixel votes only for the directions similar to its own direction (defined by its color).

2 Extraction of Perceptual Primitives

In this section, we will give a brief introduction of the first stages of the whole process, that is, computation of the gaborjets and reduction of the dimensionality of the input space by means of a hierarchy of GCSs organized in two-layers. These steps construct the RGB images that the Hough transform developed in this work for performing the perceptual grouping process will receive as input.

The main drawback of Gabor wavelet decomposition is its high complexity both in memory and computational time. This is the reason why a multiresolution spatial domain of the 2D Gabor wavelet decomposition [4] was applied. This implementation has an important drawback, the thickness of the edges detected. In this work, we have modified the previous extraction of edges as, for each main orientation considered, the exact position of the edges will be detected as the

conjunction between a maximum in the modulus and a zero-crossing in the even or the odd part of the Gabor result. Only these positions and their direct neighbors have been considered edgels, and the rest of pixels were assigned a zero value. This way, we are using the important property of Gabor filters of good localization. The application of Gabor filters centered at two main frequencies ($\frac{1}{4}, \frac{1}{8}$) and eight main orientations ($\frac{k\pi}{8}, k = 0..7$) mapped the original image into a high dimensional space. An eight component feature vector or gaborjet has been assigned to each pixel in each frequency level, containing the results from Gabor decomposition centered in each orientation. A ninth component, called response [4], was added to the gaborjet in order to emphasize the presence of a directional feature and to reduce inter-image variability. As previously mentioned, a reduction of the dimensionality of this input space was achieved using a self-organized structure. Several of such structures were analyzed previously [5] and, as a conclusion, GCS was the one selected, as results were significantly better. Each GCS will be organized in a two-layer hierarchy. First layer organizes information from different orientations and second layer will receive the output of the first layer for each frequency in order to organize information from different frequencies.

Growing Cell Structures (GCS) [6] are neural networks based on self-organized maps, but they try to eliminate the restrictions of the a priori network size definition, incorporating a mechanism to add new processing elements when needed, but maintaining the network topology. Adjustable parameters that define these structures are: weight vector assigned to each processing element and number of processing elements in the final structure. GCS incorporates a parameter called *resource value* associated to each processing element, which indicates where new processing elements must be inserted to satisfy a determined merit function.

In order to represent the results graphically, we use a colormap, where each processing element was assigned a different color that represented in some way the distance between processing elements. Thus, if a main cluster is assigned a determined color, then the elements around must have similar colors. The objective was the selection of 8 equidistant colors representing the 8 main orientations. For this purpose, we have resorted to the chromaticity diagram [7]. Results of applying this hierarchy of GCSs can be seen in second row of fig. 1.

3 Segment Extraction Using the Hough Transform

Hough transform is widely used in artificial vision and pattern recognition for the detection of geometrical shapes that can be defined through parametric equations. Traditional Hough transform implementations are based on the results obtained by classical edge detectors, like Canny or Sobel. This paper describes the design and implementation of a Hough transform based on the pseudo-color images obtained from previous processes, this is, color images where each color represents a specific orientation.

Hough transform is based on the normal equation of a line, which states that, if the normal to the line makes an angle θ with x axis and the length of this normal is ρ , then the equation of the line is given by:

$$\rho = x \cdot \cos \theta + y \cdot \sin \theta \quad (1)$$

Classically, the continuous $\rho - \theta$ space is quantized into suitable sized $(\Delta\theta \times \Delta\rho)$ rectangles and each rectangle is associated with an element of an accumulator array A with size $N_\rho \times N_\theta$. However, this parameterization needs to be enhanced to deal with the lines in our color-labelled images. We need that all points with similar colors (close orientations) and neighboring positions vote for the same line. Our accumulator will have 2 dimensions $p = 0..N_\rho - 1$ and $q = 0..N_\theta - 1$, and we must search for the maxima in this space.

The process begins with the quantization of the Hough space in $N_\rho \times N_\theta$ cells, where $N_\rho = \sqrt{I_x^2 + I_y^2} / \Delta\rho$ and $N_\theta = \pi / \Delta\theta$, depending on the size of the image $(I_x \times I_y)$ and the quantization in ρ ($\Delta\rho$) and θ ($\Delta\theta$). Then, we compute the contribution of each pixel P in the labelled image, from its angle q_P (determined by the color it has assigned). For this angle, we compute the corresponding θ_P from the quantization in θ ($\theta_P = q \Delta\theta$), and then ρ_P from eq. 1. Thus, p_P is obtained from the quantization in ρ ($p_P = \rho_P / \Delta\rho$). A pixel P contributes to all the neighbors $A(p, q)$ of $A(p_P, q_P)$ incrementing their value with $\Delta A = C_p C_q$ where:

$$C_p = e^{-\beta_p d^2(p, p_P)} \text{ and } C_q = e^{-\beta_q d^2(q, q_P)} \quad (2)$$

p being the p-neighbor of p_P if $C_p < \varepsilon_p$ and q the q-neighbor of q_P if $C_q < \varepsilon_q$. β_p and β_q are the parameters of the gaussians which define the contributions to the accumulator A and d is the distance between the two points in p-space and q-space, respectively. These parameters control how smooth the decay in the Gaussian is.

Parameters β_p and ε_p were fixed from $\Delta\rho$. As we have chosen $\Delta\rho = 1$, we have fixed $\beta_p = 0.03$ in order that the nearest neighbors in p-space have high contributions to the accumulator array. With this β_p , when the distance between p and p_P is greater than about 10, the contribution is too small to be considered, and so the threshold of C_p will be set to $\varepsilon_p = 0.05$. Parameters β_q and ε_q were selected from $\Delta\theta$ in a similar way as those for index p . First, we have fixed $\Delta\theta = \pi/24$. As C_q must have a higher contribution to its nearest angles not overlapping the next of our 8 main orientations, we have fixed $\beta_q = 0.33$ and $\varepsilon_q = 0.05$ in order that the contributions of angle indices q further than 3 from q_P are very small.

Once the voting process has finished, the following step is the maxima detection. Each maximum detected in the accumulator array corresponds to a line in the image that can contain one or more segments. For each maximum detected over a predefined threshold, an inverse Hough transform takes place removing all the contributions to the accumulator array of the pixels belonging to the line detected. A detection of segments takes place simultaneously to the inverse

Hough transform. When a pixel belonging to a line is removed from the accumulator array, its orthogonal projection to the line it belongs to is determined. Once all the pixels involved have been analyzed, the line is sequentially searched determining which pixels have been colored and though belong to a segment.

The final result of this process is an array of segments. Each segment is defined by the polar coordinates of the line it belongs to, and its edgels. From these, all the defining characteristics of the segment, like length or slope, can be computed.

4 Results

Fig. 1 shows the final results obtained by the system with a set of different kinds of images. It can be observed that GCS results contain the most important directional features in the image. Each orientation has the same color assigned in each of the images, which will be crucial for the detection of the segments that the image contains. All the results have been obtained from the same set of parameters, a priori fixed without intervention of the final user.

Hough transform results contain the segments in the image. As Hough transform is a global instead of a local operation, the information contained in the whole image can influence in the deviation of the segments and the detection of spurious segments. These problems will be solved in higher level analysis.

5 Discussion and Conclusions

In this paper, the directional primitives of an image have been extracted by means of an efficient multi-resolution spatial-domain implementation of Gabor wavelet decomposition, centered at 2 frequency channels and 8 orientations.

Later on, these primitives have been organized and grouped by means of growing-cell structures (GCS) in a two layer hierarchy. The first layer organizes the information from different orientations in each frequency channel. The second layer integrates the organizations from different frequency channels so as to have an unique result where information from different frequency channels complement each other. This result has been a RGB image containing the most important directional features in the image.

The next step has been the analysis of this RGB image by means of a fuzzy Hough transform, that operates on color images where each color represents an orientation instead of on binary images. The objective is the detection of the segments contained in the image.

The result of all these processes is a list of segments. The next step will be the organization of these segments into larger groups that will be later organized in order to select those low-level features that belong to a single object. This is, the directional features extracted from Gabor decomposition will be used as the primitives of a perceptual organization process.

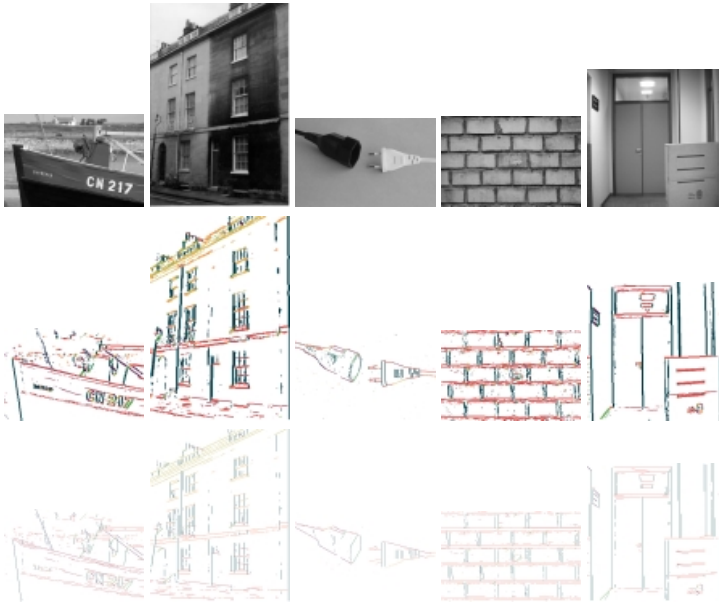


Fig. 1. First row shows original images, second row results from GCS analysis, and third row segments obtained after applying the Hough transform.

Acknowledgements. This paper has been partly supported by Xunta de Galicia through grant PGiDT01PXI10502PR.

References

1. D. G. Lowe. Perceptual Organisation in Computer Vision: A Review and a Proposal for a Classificatory Structure. *IEEE Trans. Syst. Man Cybern.* **23(2)** (1993) 382–399.
2. S. Sarkar, K. L. Boyer: *Computing Perceptual Organization in Computer Vision*. Machine Perception and Artificial Intelligence, Volume 12. World Scientific.
3. S. Sarkar, P. Soundararajan: Supervised Learning of Large Perceptual Organization: Graph Spectral Partitioning and Learning Automata. *IEEE Trans. Pattern Anal. Mach. Intell.*, **22(3)** (2000) 504–517.
4. M.J. Carreira, M. Mirmehdi, B.T. Thomas, M. Penas. Perceptual primitives from an extended 4D Hough transform. *Image Vis. Comput.* **20(13-14)** (2002) 969–980.
5. M. Penas, M. J. Carreira, M. G. Penedo. Autoorganised Structures for the Extraction of Perceptual Primitives. *Lect. Notes Comp. Science* **2085** (2001) 628–636.
6. B. Frizke: Growing Self-organising Networks – Why?. 4th European Symposium on Artificial Neural Networks (1996) 61–72.
7. J. D. Foley: *Computer Graphics. Principles and Practice*. Addison-Wesley.