

Gabor Parameter Selection for Local Feature Detection [★]

Plinio Moreno, Alexandre Bernardino, and José Santos-Victor

Instituto Superior Técnico & Instituto de Sistemas e Robótica
1049-001 Lisboa - Portugal
{plinio, alex, jasv}@isr.ist.utl.pt

Abstract. Some recent works have addressed the object recognition problem by representing objects as the composition of independent image parts, where each part is modeled with “low-level” features. One of the problems to address is the choice of the low-level features to appropriately describe the individual image parts. Several feature types have been proposed, like edges, corners, ridges, Gaussian derivatives, Gabor features, etc. Often features are selected independently of the object to represent and have fixed parameters. In this work we use Gabor features and describe a method to select feature parameters suited to the particular object considered. We propose a method based on the Information Diagram concept, where “good” parameters are the ones that optimize the filter’s response in the filter parameter space. We propose and compare some concrete methodologies to choose the Gabor feature parameters, and illustrate the performance of the method in the detection of facial parts like eyes, noses and mouths. We show also the rotation invariance and robustness to small scale changes of the proposed Gabor feature.

1 Introduction

The object recognition problem has been tackled recently with several successful results [1–4]. All of these works exploit the idea of selecting various points in the object and building up a local neighborhood representation for each one of the selected points. Two related problems are involved in this process: (i) which points in the object should be used and (ii) how to represent the information contained in their neighborhood. In the present work, we address the latter problem, assuming that interest points are obtained by some methodology [1–3]. In the experiments we present later, interest points are selected manually.

Regarding the problem of local neighborhood representation, there are several types of features being proposed in the literature: gradient magnitude and orientation maps [1], Gaussian derivatives [2, 3], rectangular features [5] and Gabor features [6], amongst others. However, the parameters when using Gabor features are often fixed [7, 8] or chosen manually [6]. In this

[★] Research partly funded by European project IST 2001 37540(CAVIAR), the FCT Programa Operacional Sociedade de Informação(POSI) in the frame of QCA III, and Portuguese Foundation for Science and Technology PhD Grant FCT SFRH\BD\10573\2002

work we also use Gabor features to represent a local image neighborhood but select their parameters according to the particular image pattern to detect.

The adaptation of feature parameters to particular object parts was first exploited in [9]. They propose to select Gabor function parameters in a semi-automatic fashion, using the “Information Diagram” concept. The Information Diagram is the representation of Gabor feature magnitude, at a certain image point, as a function of the Gabor filter orientation and frequency parameters. The scale and wavelength (inverse of frequency) have a fixed ratio. In our work, we extend the “Information Diagram” concept to consider scale and wavelength as independent parameters, thus resulting in a 3-dimensional function. We show different methodologies to select “good” feature parameters from this Extended Information Diagram.

In order to evaluate different methodologies for parameter selection, we have set-up a facial feature learning and detection experiment. The evaluation of results will be based on the detection rates achieved. Since the focus of the work is on the selection of feature parameters, we will employ very straightforward techniques for the learning and detection steps. In the learning step we compute the object model, consisting in the average and covariance matrix of vectors containing the response of selected Gabor features in a large training set. In the detection step, we compute the distance (Euclidean and Mahalanobis) between novel image points and the acquired models. We have performed experiments in the identification of facial points like eyes, mouths and noses, and obtain high success rates with the proposed features. Then we evaluate the robustness of the method to pattern variations in scale and orientation.

2 Gabor Functions as Local Image Descriptors

The motivation to use Gabor functions is mostly biological, since Gabor-like receptive fields have been found in the visual cortex of primates [10]. Gabor functions act as low-level oriented edge and texture discriminators and are sensitive to different frequencies and scale information. These facts raised considerable interest and motivated researchers to extensively exploit the properties of Gabor functions.

2.1 The Gabor Function

Mathematically, a 2D Gabor function, g , is the product of a 2D Gaussian and a complex exponential function. The general expression is given by:

$$g_{\theta, \lambda, \sigma_1, \sigma_2}(x, y) = \exp \left\{ -1/2 (x \ y) M (x \ y)^T \right\} \exp \left\{ \frac{j\pi}{\lambda} (x \cos \theta + y \sin \theta) \right\}$$

where $M = \text{diag}(\sigma_1^{-2}, \sigma_2^{-2})$. Some examples of Gabor functions are shown in Fig.1. The parameter θ represents the orientation, λ is the wavelength, and σ_1 and σ_2 represent scale at orthogonal directions. When the Gaussian part is symmetric, we obtain the isotropic Gabor function:

$$g_{\theta, \lambda, \sigma}(x, y) = \exp \left\{ -\frac{x^2 + y^2}{2\sigma^2} \right\} \exp \left\{ \frac{j\pi}{\lambda} (x \cos \theta + y \sin \theta) \right\} \quad (1)$$

However, with this parameterization the Gabor function does not scale uniformly, when σ changes. It is preferable to use a parameter $\gamma = \lambda/\sigma$ instead of λ so that a change in σ corresponds to a true scale change in the Gabor function. Also, it is convenient to apply a 90 degrees counterclockwise rotation to Eq. (1), such that θ expresses the orthogonal direction to the Gabor function edges. Therefore, in the remainder of the paper we will use the following definition for the Gabor functions:

$$g_{\theta,\gamma,\sigma}(x,y) = \exp\left\{-\frac{x^2+y^2}{2\sigma^2}\right\} \exp\left\{\frac{j\pi}{\gamma\sigma}(x\sin\theta - y\cos\theta)\right\}$$

By selectively changing each of the Gabor function parameters, we can “tune” the filter to particular patterns arising in the images. In Fig. 1 we illustrate the variation of parameters(γ, θ, σ) in the shape of the Gabor function.

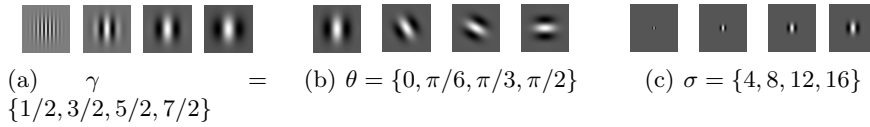


Fig. 1. Examples of Gabor functions. Each sub-figure shows the real part of Gabor function for different values of γ, θ , and σ

2.2 Gabor Response

By convolving a Gabor function with image patterns $I(x,y)$, we can evaluate their similarity. We define the Gabor response at point (x_0, y_0) as:

$$G_{\theta,\gamma,\sigma}(x_0, y_0) = (I * g_{\theta,\gamma,\sigma})(x_0, y_0) = \int I(x,y)g_{\theta,\gamma,\sigma}(x_0 - x, y_0 - y)dx dy \quad (2)$$

where $*$ represents convolution. The Gabor response obtained from Eq. (2) can emphasize basically three types of characteristics in the image: edge-oriented characteristics, texture-oriented characteristics and a combination of both. In order to emphasize different types of image characteristics, we must vary the parameters σ, θ and γ of the Gabor function.

The variation of θ changes the sensitivity to edge and texture orientations. The variation of σ will change the “scale” at which we are viewing the world, and the variation of γ the sensitivity to high/low frequencies. We would like to find the most adequate combinations of σ, θ and γ to represent particular parts of objects for recognition/detection tasks.

3 Object Part Model

As mentioned in the introduction, we consider objects composed of parts, like eyes, noses and mouths in human faces. Each part is modeled as random vector containing (a) the absolute value of Gabor responses, and (b) the real and imaginary parts of Gabor responses with different parameters. In the case of (a), the feature vector is:

$$v_{(x,y)} = (v_{(x,y)}^1, \dots, v_{(x,y)}^i, \dots, v_{(x,y)}^m)^T; \quad v_{(x,y)}^i = |G_{\theta_i,\gamma_i,\sigma_i}(x,y)| \quad (3)$$

and (x, y) represents the coordinate of the object part center. In the case of (b), the feature vector is:

$$v = (v^1, \dots, v^i, \dots, v^{2m})^T; \quad v^{2i} = \text{Re}(G_{\theta_i, \gamma_i, \sigma_i}); \quad v^{2i-1} = \text{Im}(G_{\theta_i, \gamma_i, \sigma_i}) \quad (4)$$

The rationale is to model image parts by analyzing their contents in terms of edges and textures of different scales, orientations and frequencies. We assume that the random feature vector follows a normal distribution with average \bar{v} and covariance matrix Σ , $v_{(x,y)} \sim \mathcal{N}(\bar{v}_{(x,y)}, \Sigma_{(x,y)})$.

For the detection of parts, we will compute the distance between the obtained model and the novel patterns. We consider both the Euclidean and Mahalanobis distances. The decision of whether a part feature is present or not in a certain image pixel will depend on the computed distance values.

4 Parameter Selection

In this section we focus on selecting the parameters (orientation, scale and frequency) for each of the Gabor functions used in the feature model. We assume a limited (constant) number of Gabor filters to constrain the computational cost of the methods. A straightforward approach to define the parameters would be to sample the parameter space uniformly. However, this strategy does not exploit the particular characteristics of the object part under test. Instead, we could analyse the Gabor response function in the full parameter space (σ, γ, θ) and select those parameters that best describe the particular object characteristics. However, this strategy could bias the parameter distribution to a narrow range and reduce the capability to discriminate the modeled object from others. To enforce some variability in the parameter space and still be able to adapt the representation to the particular object under test, we will sample uniformly one of the parameters and perform a 2D search in the remaining dimensions. This strategy extends the concept of **Information Diagram**[9].

4.1 Information Diagram

The ‘‘Information Diagram’’ (ID) concept proposed in [9] selects the Gabor filter parameters semi-automatically. The ID represents the magnitude of the Gabor response at a certain interest point (x, y) , as a function of θ and σ , keeping the γ parameter constant. The ID function is defined as:

$$\text{ID}_{x,y}(\theta, \sigma) = |G_{\theta, \gamma=1, \sigma}(x, y)|$$

Then, local maxima coordinates of ID are chosen as ‘‘good’’ Gabor function’s parameters because they represent the object’s characteristic orientations, scales and frequencies, thus being considered good descriptors of the local image content.

In this work, we extend the ID concept to consider variability also in the γ parameter. We define the Extended Information Diagram as the 3D function:

$$\text{EID}_{x,y}(\theta, \sigma, \gamma) = |G_{\theta, \gamma, \sigma}(x, y)|$$

Then we denote θ -ID, γ -ID, and σ -ID as slices of the EID function, keeping constant one of the parameters, $\theta = \theta_0$, $\gamma = \gamma_0$ or $\sigma = \sigma_0$:

$$\theta\text{-ID}_{x,y}^{\theta_0}(\sigma, \gamma) = \text{EID}_{x,y}(\theta_0, \gamma, \sigma); \quad \sigma\text{-ID}_{x,y}^{\sigma_0}(\theta, \gamma) = \text{EID}_{x,y}(\theta, \gamma, \sigma_0);$$

$$\gamma\text{-ID}_{x,y}^{\gamma_0}(\theta, \sigma) = \text{EID}_{x,y}(\theta, \gamma_0, \sigma)$$

According to our notation, the work in [9] uses a γ -ID with $\gamma_0 = 1$. In Fig. 2 we show some examples of the θ -ID, σ -ID and γ -ID for the left eye's center point of the image used in Fig. ??.

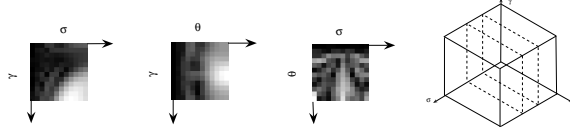


Fig. 2. Examples of θ -ID, σ -ID, γ -ID, and σ slices in the parameter space from left to right

4.2 Searching Multiple Information Diagrams

Our strategy to find good parameters for the object part's model features is based on uniformly discretizing one of the parameters (say θ), and search local extrema in the resulting set of θ -ID's. For example, a set of θ -IDs for $\mathcal{T} = \{\theta_1, \dots, \theta_n\}$, at point (x, y) is given by:

$$\Theta\text{-ID}_{x,y}^{\mathcal{T}} = \{\theta\text{-ID}_{x,y}^{\theta_1}, \dots, \theta\text{-ID}_{x,y}^{\theta_n}\}$$

The several $\theta_i \in \mathcal{T}$ are uniformly spaced in the range $[0, \pi)$. Then we compute the parameters of the highest local maximum and smallest local minimum:

$$(\hat{\sigma}_i^{\max}, \hat{\gamma}_i^{\max}) = \arg \max_{\sigma, \gamma} \theta\text{-ID}_{x,y}^{\theta_i}; \quad (\hat{\sigma}_i^{\min}, \hat{\gamma}_i^{\min}) = \arg \min_{\sigma, \gamma} \theta\text{-ID}_{x,y}^{\theta_i}$$

Then, the set of chosen Gabor function parameters in Eq.(3) and Eq.(4), are such that $(\gamma_i, \theta_i, \sigma_i)$ belongs to $\{(\hat{\sigma}_1^{\min}, \hat{\gamma}_1^{\min}, \theta_1), \dots, (\hat{\sigma}_n^{\min}, \hat{\gamma}_n^{\min}, \theta_n)\}$ and/or $\{(\hat{\sigma}_1^{\max}, \hat{\gamma}_1^{\max}, \theta_1), \dots, (\hat{\sigma}_n^{\max}, \hat{\gamma}_n^{\max}, \theta_n)\}$.

5 Experimental Results

In this section we present the results of the tests done for the various approaches to object modeling and feature parameter selection. Then we select the most successful approach and perform tests in order to verify the rotation invariance and robustness to scale changes of the selected feature vector.

The experimental tests performed in this work use 90 subjects from the AR face database [11], all without glasses, where 60 of them are used for training (object part modeling) and 30 for testing (object part detection). We represent four different parts: left eye, right eye, nose and mouth.

5.1 Selection of the object model and the modified ID

Experiments are set-up for evaluating the discretized parameters (σ , γ or θ), the number and type of the extrema computed at each ID, the distance metrics (Euclidean and Mahalanobis), and the feature model type (magnitude *vs* real-imaginary parts). A list of the experiments and related configurations is shown in Table 1.

Table 1. List of the performed tests. Performance in last two columns(%)

Test ID	type	# local max	# local min	distance	mag	re+im
1	θ	1	1	Mah	68.49	78.33
2	θ	2	0	Mah	85.92	95.83
3	γ	2	0	Mah	58.19	74.16
4	γ	1	1	Mah	54.41	75.83
5	σ	2	0	Mah	58.19	72.50
6	σ	1	1	Mah	50.21	72.50
7	θ	1	1	Euc	31.93	85
8	θ	2	0	Euc	38.87	87.5
9	γ	2	0	Euc	17.86	53.33
10	γ	1	1	Euc	15.55	45
11	σ	2	0	Euc	24.79	74.17
12	σ	1	1	Euc	15.97	75.83

In every experiment performed we use $n = 12$ IDs, and choose either one local maxima and one local minima or two local maxima, so the number of filters is kept constant ($m = 24$). The sets of values for the θ , γ and σ -IDs are, respectively, $\mathcal{T} = \{0, \pi/12, \dots, 11\pi/12\}$, $\mathcal{G} = \{0.5, 0.8, \dots, 4\}$, and $\mathcal{S} = \{4, 7, \dots, 39\}$.

All IDs are calculated from the mean images \bar{I}_{part} in the training set, centered at each object part (left eye, right eye, nose, mouth). To evaluate the performance of each experiment we count the number of hits (successful detections) in the test set. Given an object part model, a distance function and an image point, a hit exists if the global minima of the distance is located inside a circle of radius r around the center of the object part. Considering only the tests using real and imaginary parts of the Gabor response, we can see, in Fig. 3 that mean performance is better when using θ -IDs, Mahalanobis distance, and 2 local maxima. In this case, success rates are as high as 95%.

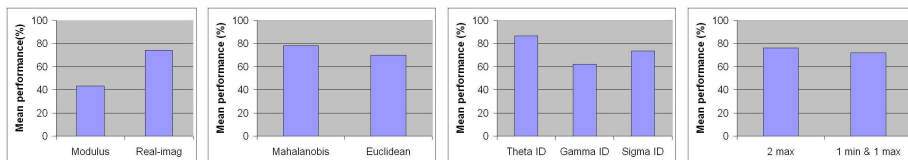


Fig. 3. Mean detection rate of marginalized tests of table 1

5.2 Rotation Invariance

We test the rotation invariance of the Gabor filters on a synthetic image, and evaluate, in the face data set, the effects of Gabor response variations to rotated patterns on the correct detection rate. Due to discretization effects and imperfect filter symmetry, Gabor response presents small variations with the amount of rotation. In Fig.4 we show the effect of image α -rotation in the response of a Gabor α -rotated to a synthetic image at the image's center point. We can observe that there are some errors in the magnitude and phase that, not being dramatic, can change the performance of the detection algorithm. The variation in the error change the success rate in the object part model when using rotated images. If we shift the angles in every component of the feature vector in Eq.(4), the rotated model is: $v_{(x,y)} = (v^1, \dots, v^i, \dots, v^{2m})^T$; $v_{(x,y)}^{2i} = \text{Re}(G_{\theta_i+\alpha, \lambda_i, \sigma_i}(x, y))$; $v_{(x,y)}^{2i-1} = \text{Im}(G_{\theta_i+\alpha, \lambda_i, \sigma_i}(x, y))$. In Fig. 4 we observe the variation of the success rate when rotating the image parts and the model. In our tests, for simplicity, we rotate the image regions every $\pi/4$, because it does not involve a recomputation of the filters response. It is important to remark that we use the object model learned when $\alpha = 0$, computed in the previous section for test 2 in Table 1. We observe a very good behaviour of the learned model in the rotated images, with a performance above 91%.

5.3 Scale robustness

To check the robustness to scale variations, we compute the success rate in rescaled images maintaining the object model learned in the original images (θ -IDs, Mahalanobis distance, and two local maxima). In Fig. 4 we observe that the performance is above 90% for image rescaling upto $\pm 20\%$, corresponding to a range of about 0.6 octaves. To cope with larger scale variations, one should cover the scale dimension with additional object part models. If we sample the scale space every 0.6 octaves we should be able to keep performance above 90%, provided that an adequate scale selection method is available.

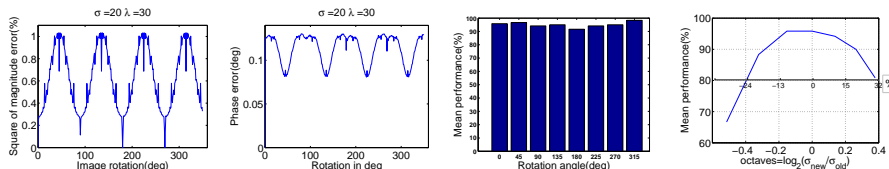


Fig. 4. Gabor filter rotation invariance tests(magnitude error, phase error, and success rate variation in rotated images) and scale robustness test, from left to right

6 Conclusions and future work

In this work we present an automatic feature selection method that can be applied to different image regions successfully. The representation is

based on Gabor features and our methodology selects automatically a set of parameters that are good descriptors for a particular image pattern, representing a part of an object. The technique is based on the Information Diagram concept [9], that is extended, in this work, to consider optimization along all dimensions of the Gabor function parameters. We illustrate the richness of the descriptor and parameter selection methods in a facial feature detection task.

The face detection tests allowed us to evaluate certain design criteria:

- a representation using the full Gabor response (real and imaginary parts) is more powerful than using the magnitude alone;
- using θ -ID's provided significantly better performance;
- the Mahalanobis distance outperforms the Euclidean distance in the detection success;

We also show some tests illustrating the rotation and scale robustness characteristics of the method. The detection method is based on simple distance metrics to stress the feature capability in representing image patterns, independently of sophisticated learning algorithms. Even though the learning algorithm is very simple, results are promising and should further improve with more powerful techniques.

References

1. Lowe, D.: Object recognition from local scale-invariant features. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (1999) 1150–1157
2. Schmid, C., Mohr, R.: Local grayvalue invariants for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence **19** (1997) 530–534
3. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: European Conference on Computer Vision, Springer (2002) 128–142
4. Weber, M., Welling, M., Perona, P.: Towards automatic discovery of object categories. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2000)
5. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. IEEE Conf. on Computer Vision and Pattern Recognition. (2001)
6. Smeraldi, F., Bigun, J.: Retinal vision applied to facial features detection and face authentication. Pattern Recognition Letters **23** (2002)
7. Jain, A., Ratha, N., Lakshmanan, S.: Object detection using gabor filters. Pattern Recognition **30** (1997) 295–309
8. Wu, H., Yoshida, Y., Shioyama, T.: Optimal gabor filters for high speed face identification. In: 16th International Conference on Pattern Recognition. Volume 1. (2002) 11–15
9. Kamarainen, J.K., Kyrki, V., Kälviäinen, H.: Fundamental frequency gabor filters for object recognition. In: Proc. of the 16th International Conference on Pattern Recognition. (2002)
10. Daniel, P., Whitteridge, D.: The representation of the visual field on the cerebral cortex in monkeys. Journal of Physiology **159** (1961) 203–221
11. Martinez, A., Benavente, R.: The ar face database. Technical report, CVC (1998)