massachusetts institute of technology — computer science and artificial intelligence laboratory

Evaluation of Sets of Oriented and Non-Oriented Receptive Fields as Local Descriptors

Jerry Jun Yokono and Tomaso Poggio

AI Memo 2004-007 March 2004
CBCL Memo 237

Abstract

Local descriptors are increasingly used for the task of object recognition because of their perceived robustness with respect to occlusions and to global geometrical deformations. We propose a performance criterion for a local descriptor based on the tradeoff between selectivity and invariance. In this paper, we evaluate several local descriptors with respect to selectivity and invariance. The descriptors that we evaluated are Gaussian derivatives up to the third order, gray image patches, and Laplacian-based descriptors with either three scales or one scale filters. We compare selectivity and invariance to several affine changes such as rotation, scale, brightness, and viewpoint. Comparisons have been made keeping the dimensionality of the descriptors roughly constant. The overall results indicate a good performance by the descriptor based on a set of oriented Gaussian filters. It is interesting that oriented receptive fields similar to the Gaussian derivatives as well as receptive fields similar to the Laplacian are found in primate visual cortex.

1. Introduction

One of the most challenging tasks in computer vision is to recognize general 3D objects in various environments. Much research has been done on object recognition both at the level of categorization and of identification of specific objects. Schneiderman [23] showed that a histogram-based categorization algorithm works well for detection of faces, cars, telephones and other general objects. Early work of Poggio [25] treated face detection as object categorization task with statistical learning by using gray global template images. The system learns face models to construct a mixture of Gaussian model. Osuna [27] used the Support Vector Machine (SVM) classifiers to learn hyperplanes which separate face and non-face. Recently, Heisele [21][22] presented a novel SVM-based method using facial components and their geometrical configuration. Components are selected automatically using a statistical error bound; a SVM classifier is then generated for each component., followed by a higher-level combination classifier. Viola and Jones [28] described a real-time face detection system using features similar to Haar-like wavelets introduced in object recognition by Poggio [29]. Their system is fast because of the cascade processing they use to implement an AdaBoost learning algorithm. This learning procedure not only tunes the thresholds for the wavelet-like features but also effectively performs feature selection.

On the other hand, the task of a 3D object identification system is to recognize specific objects. Schiele and Crowly [17][18][19] proposed a multi-dimensional histogram approach which can recognize many 3D objects (they used the COIL database as a benchmark). In their system, Gaussian derivatives are used to create a multi-dimensional feature space. Feature representation and distance metrics for histogram matching are thoroughly studied in their work. Mel [20] also developed a histogram-based object recognition system that is inspired by properties of the human visual cortex model and uses multiple low-level attributes such as color, local shape and texture. Although these systems are robust to changes in rotation, position, and deformation, they cannot cope with recognition in a cluttered scene. It is clear that in general significant invariance to rotation and deformation can be achieved at the expense of specificity.

Recently, Lowe [9] developed an object recognition system that works well under a variety of illuminations and achieves rotational and scale invariance by using a unique local descriptor. The descriptor is constructed from several collections of orientation histograms; it is normalized with the canonical orientation of that local region. Matching is performed by efficient nearest neighbor search and by voting for affine parameters. The system can learn an object model from a single image. Schmit [11] also developed an image retrieval system using a rotational invariant local descriptor.

In most of the learning approaches reported in the literature, a large number of training images are needed off-line to solve problems such as face recognition. Our ultimate goal in the work reported in this paper is to develelop an on-line learning scheme than can be effective after just one example but can also improve incrementally when more examples become available.

As a first step in this long-term program, we evaluated the performance of a local descriptor based on sets of oriented Gaussian derivatives. Gaussian derivative filters are known for their selectivity to specific orientation and frequency. Freeman [6] developed an efficient way to compute an arbitrarily oriented response by combining linearly several basis functions, and called the technique *"steerable filter"*. Although local descriptors have become popular in object recognition tasks, "steerable filters" have not been used yet in local descriptors. In this paper, we compare a local descriptor using various combinations of Gaussian derivatives with different orientations and scales with the simplest local descriptor: gray value patches. We performed the comparison in terms of criterion of sectivity and invariance. In particular, we show that our Gaussian descriptor can compute pointwise correspondences in novel images. To deal with scale changes, neighboring descriptors can be combined; the best correspondence is then found searching in scale-space.

In section 2, we will show brief overview of steerable filter.

2. **Steerable Filter**

Gaussian derivative filters are widely used filters which have spatial orientation selectivity as well as frequency selectivity. Those filters, which show robustness to illumination changes and view point changes, have many interesting properties. The first and second derivatives correspond to the edge feature and the bar feature, respectively.
Given the Gaussian distribution G, x and y directional first order derivatives are given by the followings:

$$G = e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

$$Gx = -\frac{x}{\sigma^2}G \quad (2)$$

$$Gy = -\frac{y}{\sigma^2}G \quad (3)$$

The second order derivatives are given by the followings:

$$Gxx = (\frac{x^2}{\sigma^4} - \frac{1}{\sigma^2})G \quad (4)$$

$$Gyy = (\frac{y^2}{\sigma^4} - \frac{1}{\sigma^2})G \quad (5)$$

$$Laplacian = Gxx + Gyy \quad (6)$$

Finally, we get the followings :

$$Gxxx = (\frac{3x}{\sigma^4} - \frac{x^3}{\sigma^6})G \quad (7)$$

$$Gyyy = (\frac{3y}{\sigma^4} - \frac{y^3}{\sigma^6})G \quad (8)$$

Freeman [6] introduced “steerable filters”, an efficient way for computing arbitrarily oriented response instead of computing derivatives for all the necessary orientations. An arbitrarily oriented filter response is computed using several basis functions by linearly combining those responses.

Filter responses of nth order Gaussian derivative $G_n(\theta)$ to an arbitrary orientation θ is denoted as follows:

$$\begin{aligned}
&G_n(\theta) = \sum_{i=1}^{n+1} k_{in}(\theta) \cdot G_n((i-1)\pi/(n+1)) \\
&k_{i1}(\theta) = \cos(\theta - (i-1)\pi/2), i = 1,2 \\
&k_{i2}(\theta) = \frac{1}{3}(1 + 2\cos(2(\theta - (i-1)\pi/3))), i = 1,2,3 \\
&k_{i3}(\theta) = \frac{1}{4}(2\cos(\theta - (i-1)\pi/4) + 2\cos(3(\theta - (i-1)\pi/4))), i = 1,2,3,4
\end{aligned} \quad (9)$$

where $k_{in}(\theta)$ is the coefficient of the basis.

For *n=1,2,3*, we get the following arbitrarily oriented responses using the basis,

$$G_1(\theta) = \cos(\theta)G_1(0^o) + \sin(\theta)G_1(90^o)$$

$$G_2(\theta) = k_{21}(\theta)G_2(0^o) + k_{22}(\theta)G_2(60^o) + k_{23}(\theta)G_2(120^o) \qquad (10)$$
$$k_{2i}(\theta) = \frac{1}{3}\{1 + 2\cos(2(\theta - \theta_i))\}$$

$$G_3(\theta) = k_{31}(\theta)G_3(0^o) + k_{32}(\theta)G_3(45^o) + k_{33}(\theta)G_3(90^o) + k_{34}(\theta)G_3(135^o)$$
$$k_{3i}(\theta) = \frac{1}{4}\{2\cos(\theta - \theta_i) + 2\cos(3(\theta - \theta_i))\}$$

Those basis functions in spatial space are shown in Fig.1.(σ =1)
Note that Laplacian function, not a basis, is added in order to evaluate Laplacian-based descriptor.
To make our filters redundant in frequency space, we use Gaussian widths of 1, 2 and 4.
Fig.2 shows spatial and frequency responses of filters. (0 degree orientation)
As we can see from Fig.3, filters are constructed covering enough orientations and frequencies using three scales of Gaussian derivatives. At each pixel location, the local information captured by the filter depends on the width of Gaussian function. As shown in Fig.2, when the Gaussian is wider, the filter captures information from further locations, responds to lower frequency and operates with more "blurred" representations.

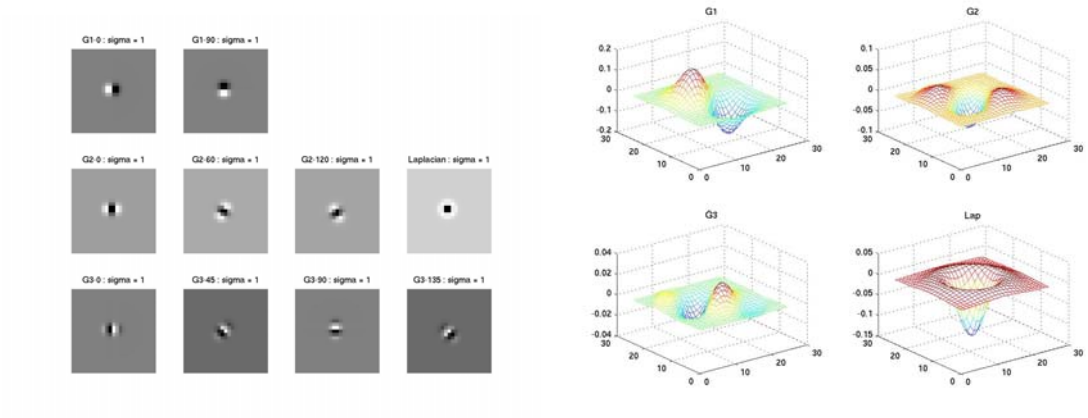


Fig.1 Steerable Filter Basis Functions (Up to 3rd order) and Laplacian (σ =1)

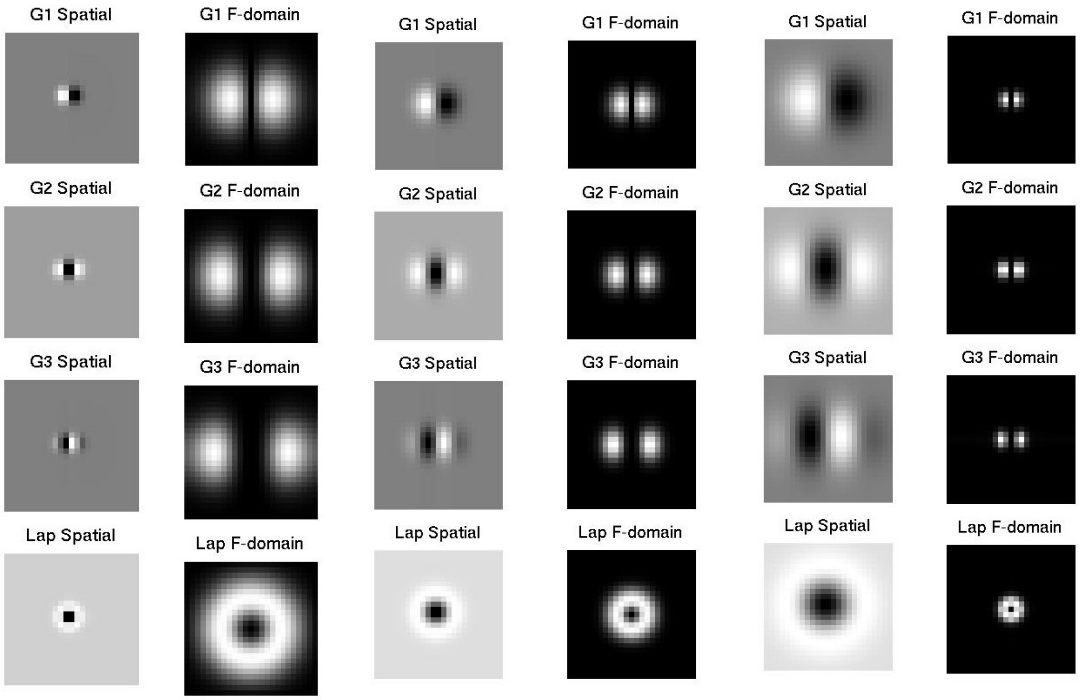


Fig. 2 Frequency responses of basis filters (σ =1,2,4 : 0 deg orientation)

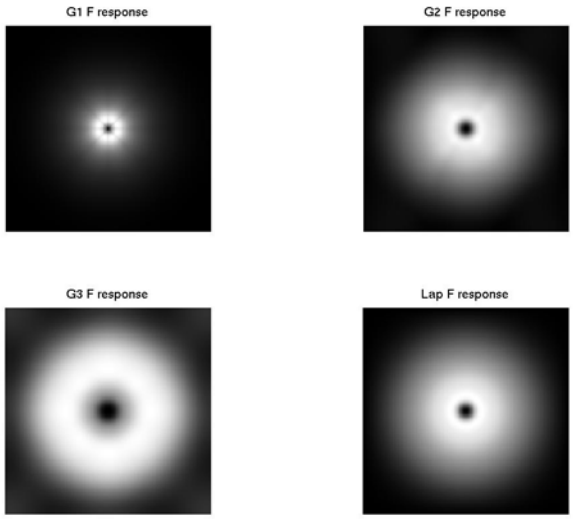


Fig. 3 Frequency responses with all orientations and scales

3. Local Descriptor

Local descriptors should have selectivity to specific local patterns as well as invariance to affine changes such as rotation, scale, and position. If selectivity is poor, many incorrect matches will appear, making recognition difficult. We evaluate the selectivity of the descriptor by computing the empirical correlation distribution associated with the filter as a function of the width of the Gaussian, the number of widths, and the orders of the derivatives.

3.1 Correlation Distribution

Each pixel represents its local information with Gaussian derivative jets, also called steerable jets.
We will show that the Gaussian derivative-based local descriptor is robust regarding selectivity.
If we use Gaussian derivatives up to the third order, four orientations, three widths, the vector length of the jet descriptor associated with a location in the image is 3x3x4=36. An empirical correlation distribution is created by selecting the descriptor associated with one location in the image and computing correlation with all the descriptors associated with the remaining pixels in the image. Normally the distribution changes depending on which pixel is selected; however, we find that the right end of the distribution is usually almost unchanged if a corner-like point is selected. Fig.5 and Fig.6 show the empirical distributions computed from the image shown, when the red-crossed pixel is selected. Fig.5 shows the distribution corresponding to a descriptor with Gaussian derivatives up to the first order is used (24 dimension jets) and Fig.6 shows the distribution for a descriptor with derivatives up to the third order (36 dimension jets). As we can see from these distributions, using high order derivatives makes the descriptor more selective to the specific pattern. In the distribution corresponding to the first order derivative descriptor, there are many other locations in the image which have high correlation to the selected pixel. Of course, there is a trade-off between selectivity and invariance. A descriptor might be very specific but show poor invariance to affine changes such as rotation, view point change, scale and brightness. We later show that the third order descriptor has good performance under both criteria of selectivity and invariance.

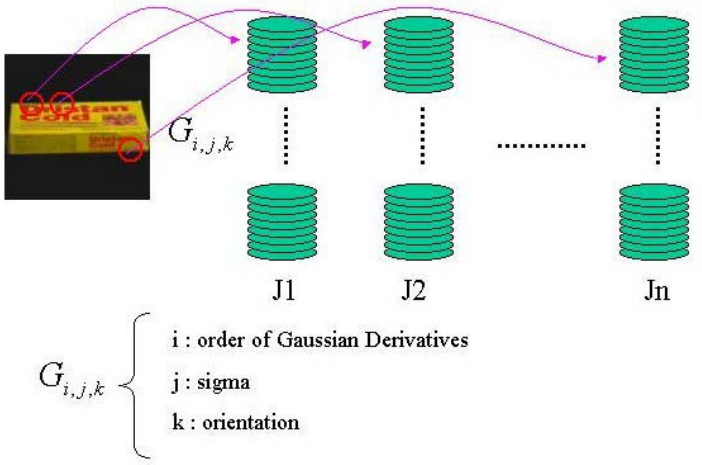
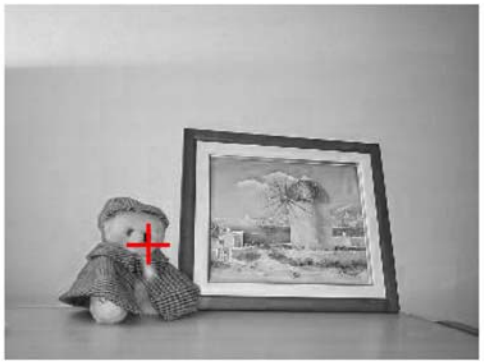


Fig. 4 Steerable Jets



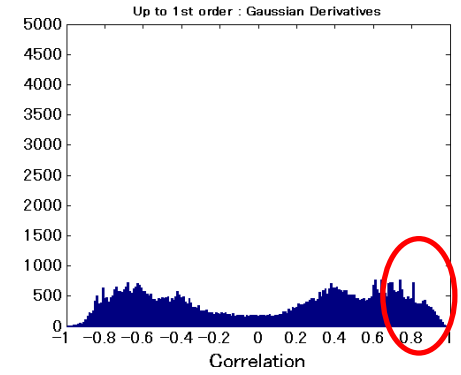


Fig.5 First Order Derivative Only

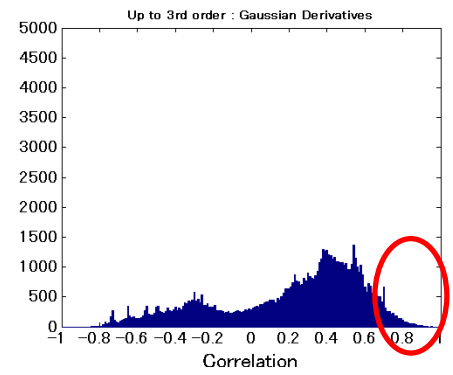


Fig.6 Up to Third Order Derivatives

3.2 Changing width of Gaussian

We evaluated the effect of the width of the Gaussian in the same way using the distribution of the correlation. Fig.7 shows the distribution when the width of the Gaussian filters is changed to 1, 2 and 4, respectively. When the width is large, local patterns are more blurred and the distribution becomes less sharp. When the width is small, the descriptor has good selectivity, but it will have in general less invariance.

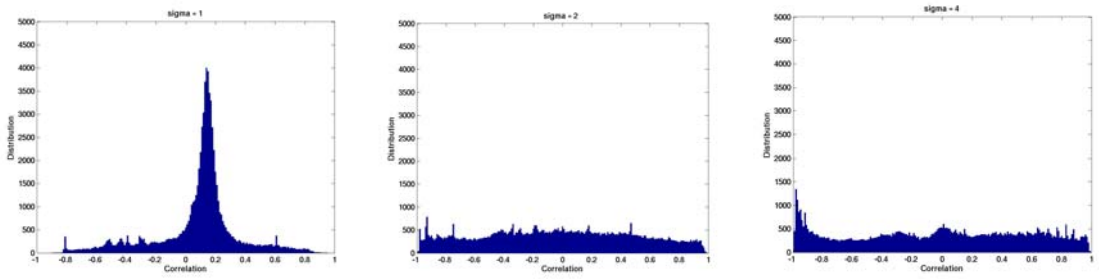


Fig.7 Effect of changing the width of the Gaussian (σ =1,2,4)

3.3 Changing number of width of the Gaussian

Fig.8 shows the distribution when the number of widths is changed. The left-hand distribution corresponds to one width (σ =1); the middle is to 2 widths (σ =1,2); the right-hand to 3 widths (σ = 1,2,4). Again, we obtain distributions that are less sharp with increasing number of widths.

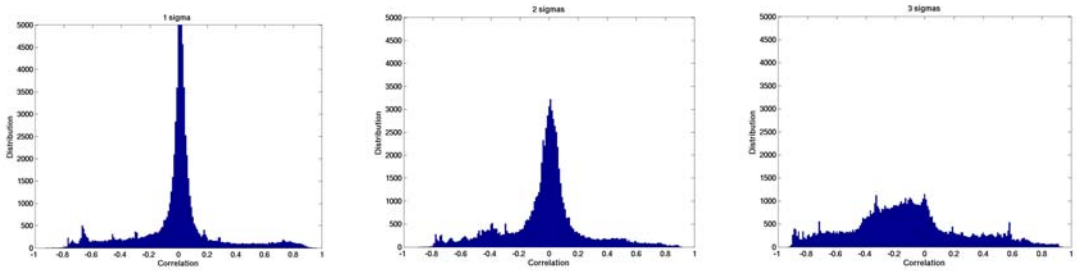


Fig.8 Changing number of width of Gaussian (1, 2, 3)

3.4 Gabor Jets

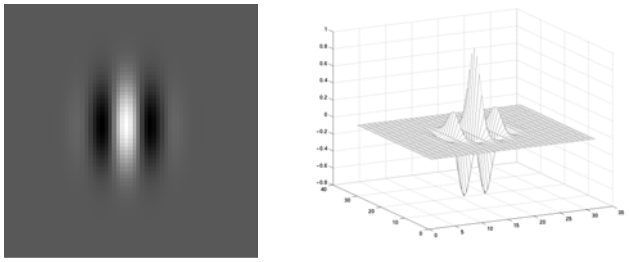


Fig.9 Gabor wavelet (real part)

Gabor functions can be regarded as approximations of high order Gaussian derivatives. They are also a reasonable model of the receptive fields of simple cells in the visual cortex (V1) [13][14].

We use the following formula for complex Gabor function:

$$G(x,y) = K\exp(-\pi(a^2(x-x_0)_r^{\ 2} + b^2(y-y_0)_r^{\ 2}))\exp(j(2\pi(u_0x_0 + v_0y_0) + P))(\exp(2\pi j(u_0(x-x_0) + v_0(y-y_0))) - \exp(-\pi(\frac{u_{0r}^{\ 2}}{a^2} + \frac{v_{0r}^{\ 2}}{b^2})))$$

P : phase of the sinusoidal carrier

x_0, y_0 : peak location of Gaussian envelop

u_0, v_0 : spatial frequency of sinusoidal carrier

K : magnitude of Gaussian envelop

The Gabor functions are products of a Gaussian envelop and a complex sinusoidal carrier and are selective both in the frequency and in the spatial domain. The last term on the right hand side compensates for the DC component. Fig.9 shows the real part of a Gabor filter. We use the following parameters for the Gabor filter:

$$P = 0$$
$$x_0 = y_0 = 0$$
$$\lambda = \frac{a}{b} = 1.24$$
$$\beta = 1.4$$

λ and β represent aspect ratio and half-magnitude frequency bandwidth respectively. Subscript r represents rotation operation. In terms of biological perspective, these values are quite appropriate.
Three frequencies and eight orientations are used to create Gabor Jets. We should say it is not appropriate to compare the result with this 24 dimensional vector since dimensionality is different. However, it is useful to see how the biological parameter model works as a local descriptor. The results indicates the distribution sometimes tends to have poor selectivity as shown in Fig. 10.

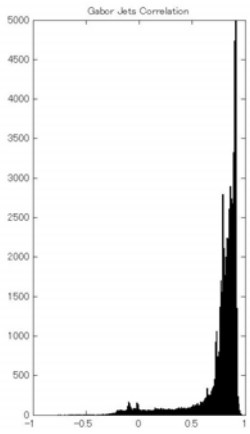


Fig. 10 Correlation Histogram of Gabor Jets

4. Computing Correspondence

To evaluate the descriminant power of the descriptor, we compute the correspondence between a descriptor for each location in one image and all other locations in a second image (Fig. 11). We use as a distance metric normalized correlation denoted as below.

$$Correlation(v_1, v_2) = \frac{(v_1 - \overline{v_1})(v_2 - \overline{v_2})}{\sqrt{(v_1 - \overline{v_1})^2 (v_2 - \overline{v_2})^2}} \qquad (12)$$

The left-hand image is used as a model image. When a pixel is selected, the best corresponding pixel in the right-hand image is found. Note here that only one pixel is used as a descriptor. Steerable jets with only the first order derivatives, do not yield correct correspondence in our tests, as shown in Fig. 12. Gabor Jets descriptor shows a better correspondence; however, as we mentioned, sometimes correspondence fails. If we use the descriptor using up to the third order, correspondence is much better as shown in Fig.14.

Fig. 11 Images used to compute correspondence

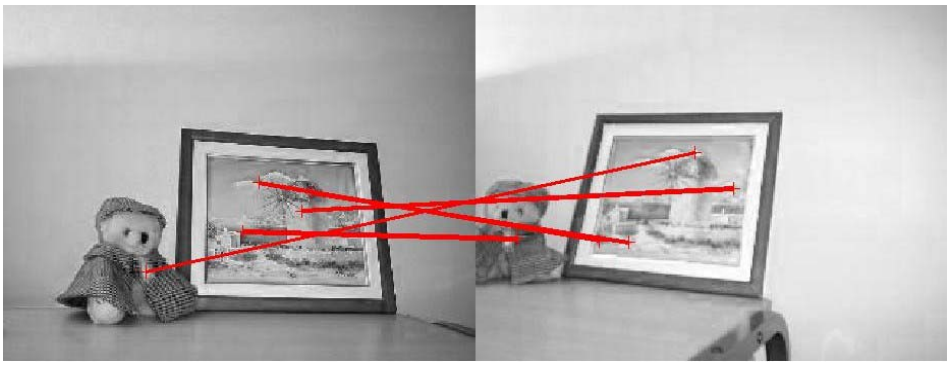


Fig. 12 Correspondence with first order Gaussian Derivative Jets descriptor

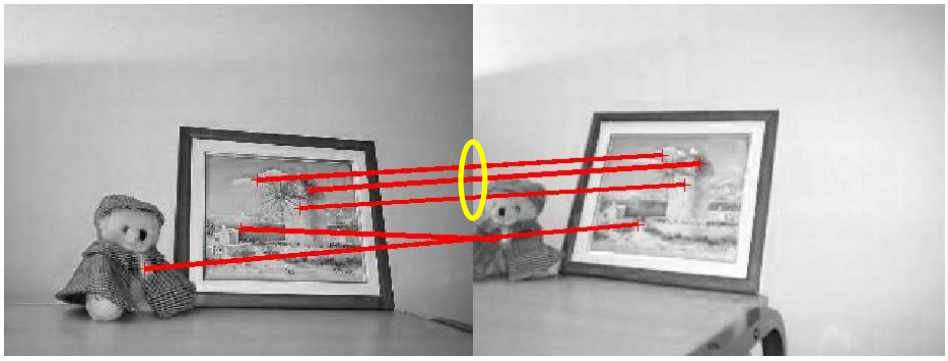


Fig. 13 Correspondence with Gabor Jets descriptor

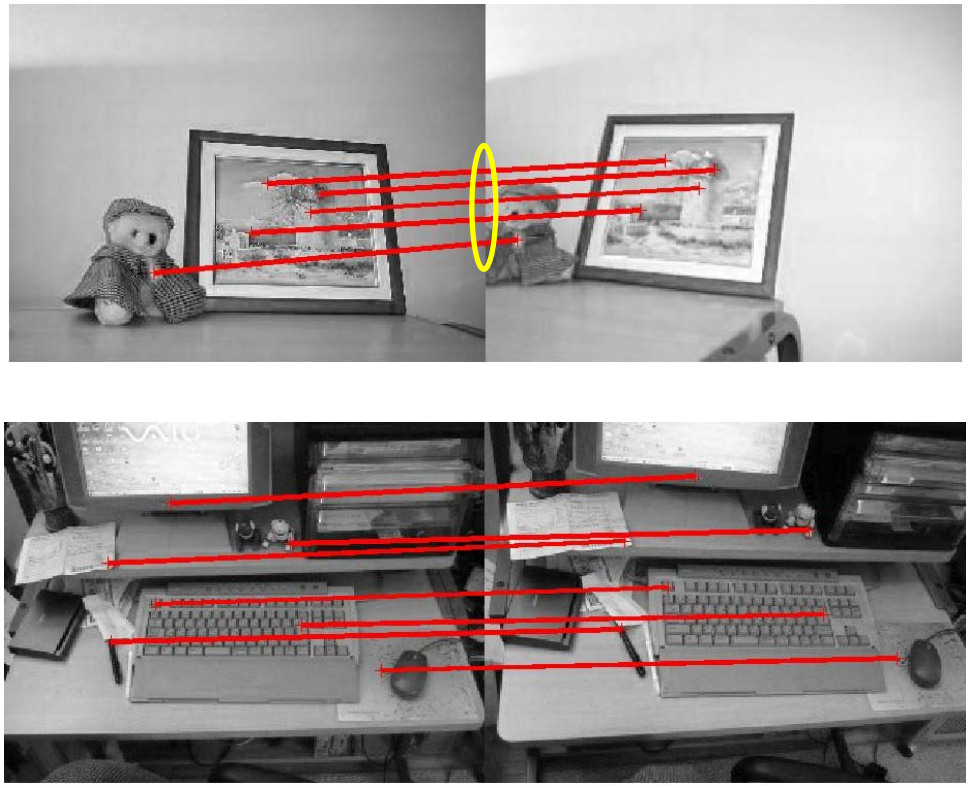


Fig. 14 Correspondence with the third order descriptor

5. Invariance and Selectivity

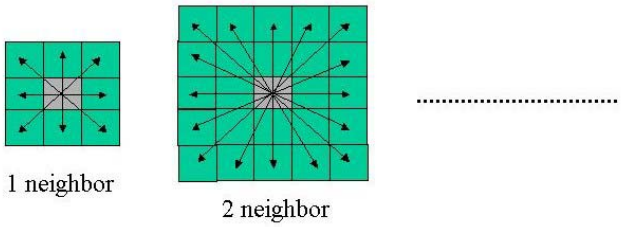
In this chapter, we show further evaluation of the descriptor regarding selectivity and invariance.

No research has been done using these criteria. We compare the descriptor to other classical descriptors such as the gray value patches and the Laplacian-based descriptor.

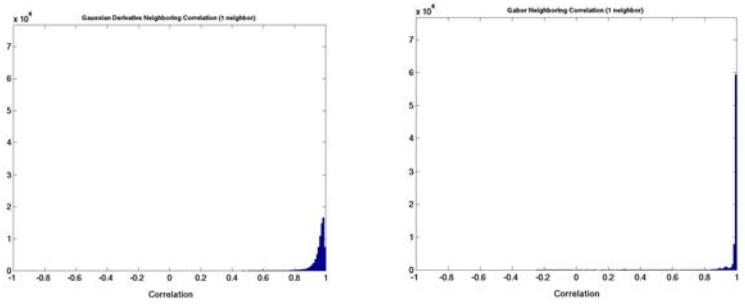
5.1 Invariance to shift changes

To examine the invariance to translations, we compute the correlation histogram with the neighboring pixels. The correlations are computed at all pixel locations in the image and the distribution is created. The distribution is computed for both the Gabor jets and for the up-to-third order descriptor. As we can see from the distributions, the Gabor jets descriptor has a high correlation even for five pixels shift. The Gabor descriptor shows invariance to significant shifts because of low-pass filtering effects: in this case we expect that it may have poor selectivity.

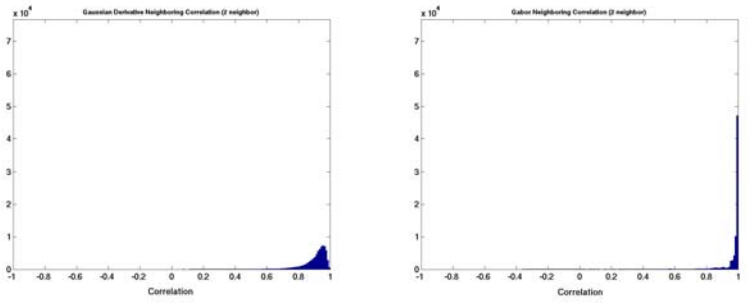
To deal with matching in a low resolution image, the descriptor can be expanded, combining the neighboring four jets, which are five pixels away from the center pixel.



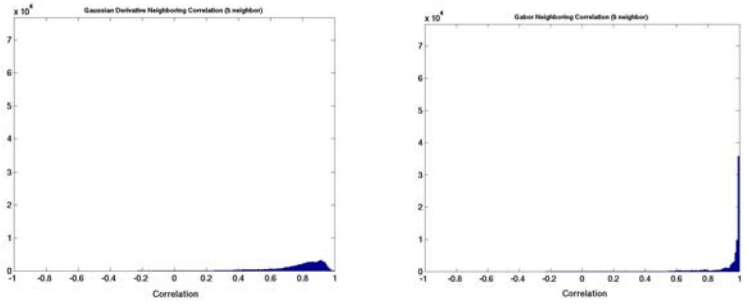
correlation with neighboring pixels



1 neighbor (Up to 3rd order Gaussian derivative and Gabor)



2 neighbor (Up to 3rd order Gaussian derivative and Gabor)



5 neighbor (Up to 3rd order Gaussian derivative and Gabor)

Fig. 15 Correlation distribution of neighboring pixels

5.2 Selectivity

As far as we know there is no formal comparison in the literature between the steerable jets descriptor and the gray value descriptor with respect to selectivity and invariance. A comparison should be performed by using similar dimension vectors. The Gaussian descriptor with derivatives up to the third order vector has a dimensionality of 180. A circular masked gray value patch (15x15) of 193 dimension vector is thus used for the comparison. In addition, we used Laplacian-based descriptors (a descriptor with one scale and 193dims and a descriptor with three scales and 207dims) in the comparison. Correlation distributions of all descriptors are shown in Fig. 16.
With a threshold 0.7 for the steerable jets descriptor, 99.9% of false matches are eliminated in the first match. In the same way, we can take a thresholds of 0.6 for the gray value descriptor, 0.75 for the 3-scale-Laplacian, and 0.4 for the 1-scale-Laplacian. These thresholds measure selectivity. As we will show later, by using these thresholds, invariance to affine changes is examined.

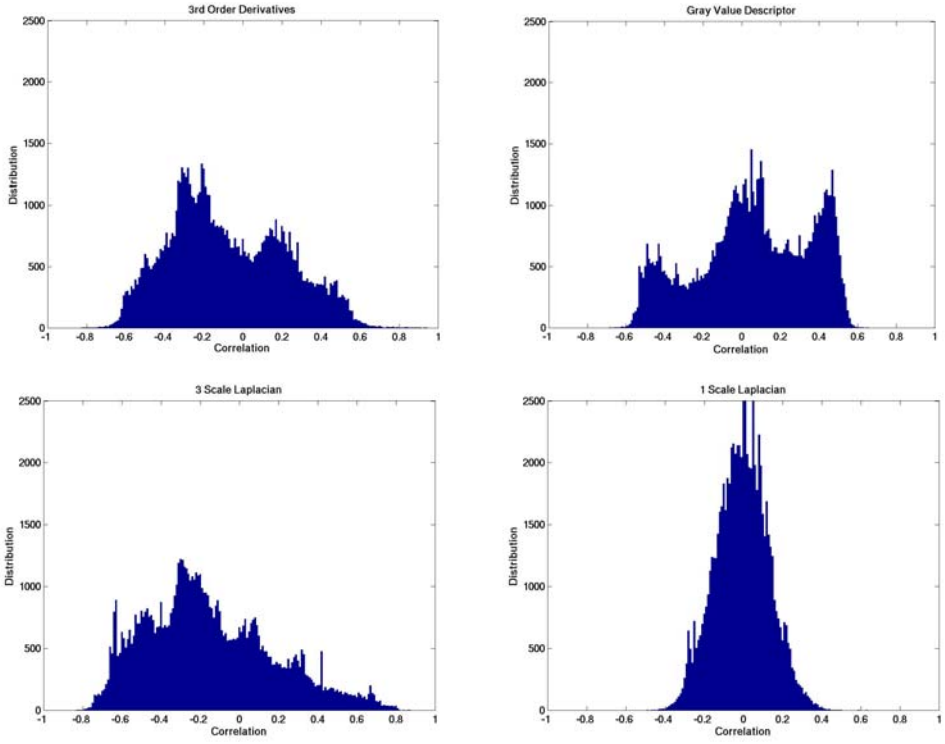


Fig. 16 Correlation distribution of descriptors

5.3 Interesting points

To evaluate invariance to affine changes, the COIL 3D object image database [24] is used.
First, interest points are selected using the Harris corner detector [8] as it is not necessary to see non-texture points.

$$M = \begin{bmatrix} (\frac{\partial I}{\partial x})^2 & (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) \\ (\frac{\partial I}{\partial x})(\frac{\partial I}{\partial y}) & (\frac{\partial I}{\partial y})^2 \end{bmatrix} \quad (12)$$

$$H = \det(M) - k(traceM)^2 \quad (13)$$

Around each interesting point, local descriptors (based on Gaussian derivatives up to the third order, based on the gray value patch, based on the Laplacian with three scales, based on the Laplacian with one scale) are computed.

5.4 Invariance to rotation

Since we know the affine parameter, we can compute the correlation with the corresponding pixel. If we use the thresholds as mentioned in 5.2 as a selectivity measurement, Fig.17 shows that even when the image is rotated 30 degrees, the descriptor based on the Gaussian derivatives up to the third order shows high correlation. The gray patch descriptor on the other hand only allows 20 degrees of rotational invariance. The descriptor with three scales of Laplacian function also has good performance, while the descriptor based on one scale of the Laplacian function works poorly.

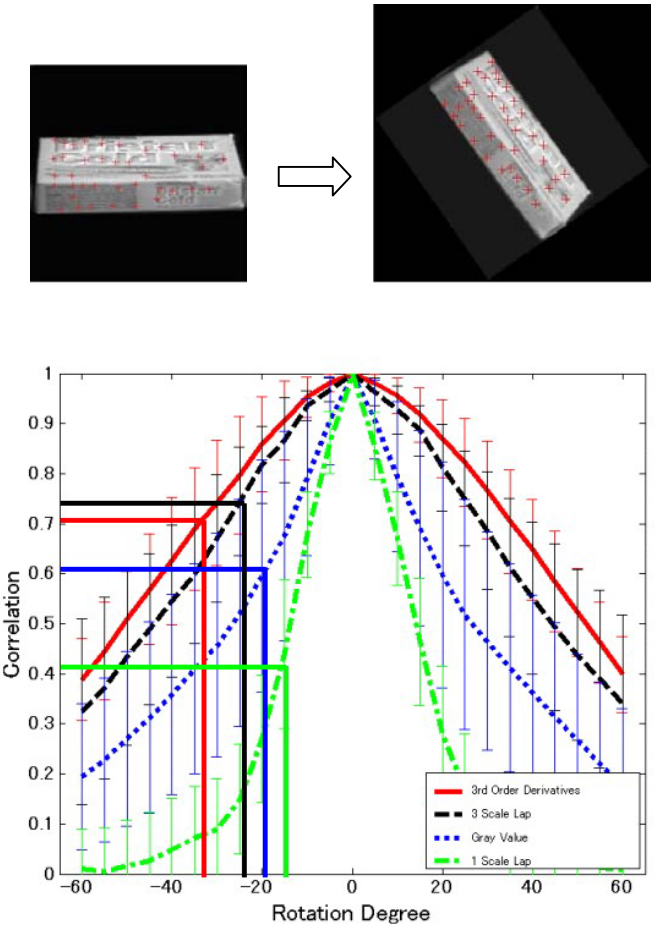


Fig. 17 Invariance to Rotation

5.5 Invariance to scale changes

Invariance to scale changes is examined by linearly rescaling the original image.
Again, as we can see in Fig.18, the descriptor with the Gaussian derivatives up to the third order shows a good performance even with an approximately half-sized image of the original. Here, the gray value patch descriptor and the Laplacian descriptor with three scales show almost the same result.

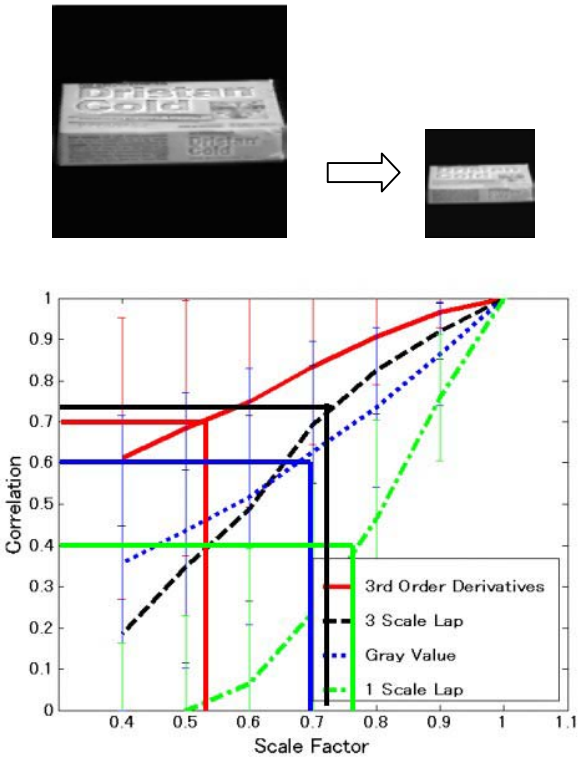
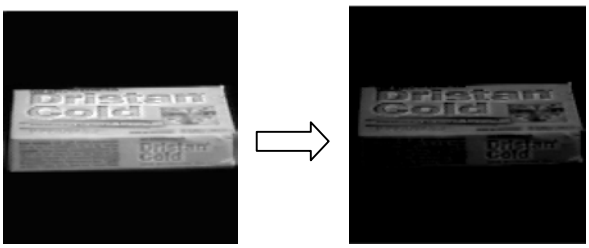


Fig. 18 Invariance to Scale Changes

5.6 Invariance to brightness

Intensity of images are shifted incrementally/decrementally to examine a simple measure of invariance to brightness. As we are computing the normalized correlation for the matching, brightness changes do not effect any of the descriptors. There is no significant difference between these descriptors as shown in Fig.19. Note that image intensity is saturated when the brightness factor is too large.



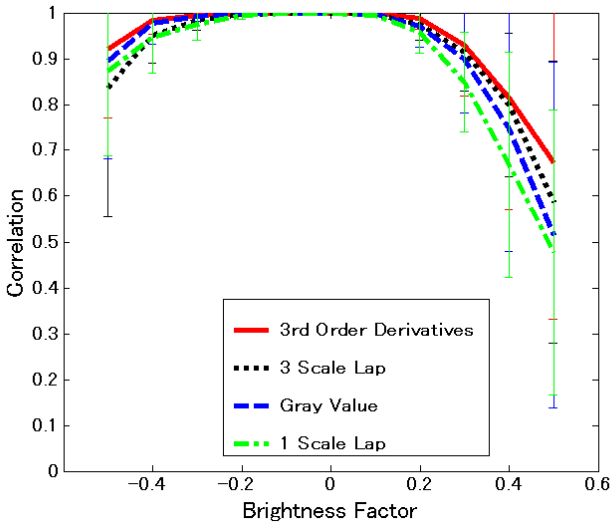


Fig. 19 Invariance to Brightness Changes

5.7 Invariance to View Point Changes

Since the COIL database has images of rotation in depth, we can examine invariance to view point changes. As we can see from Fig. 20, if the object is rotated in depth 25 degrees, the descriptor still shows a high correlation.

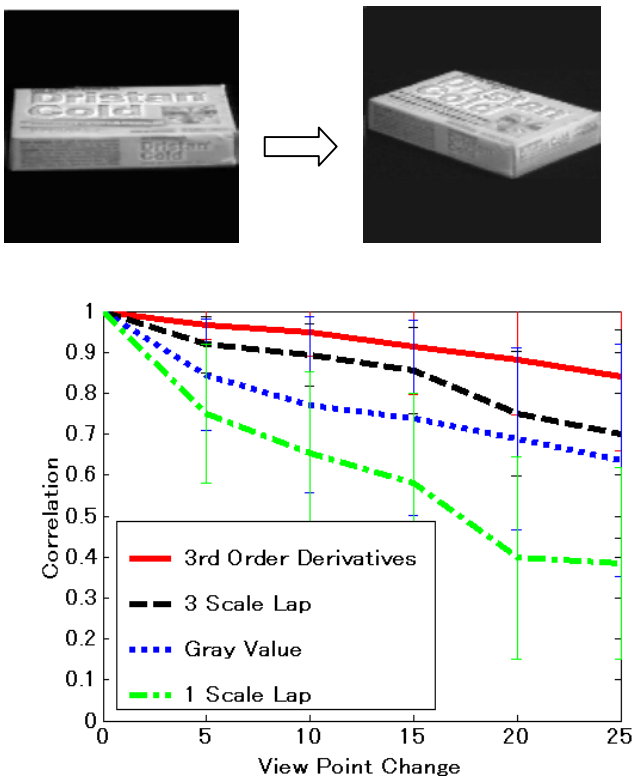


Fig. 20 Invariance to View Point Changes

6. Scale invariant matching

To realize scale invariant matching, a multi-scale representation of an object model is constructed. Yuille and Poggio [1][2], Lindeberg [3] suggested that a multi-scale image can be represented as the solution to the diffusion formulation. The Gaussian kernel is found out to be unique to generate the scale-space:

$$\partial_t L = \frac{1}{2}\nabla^2 L = \frac{1}{2}(\partial_{xx} + \partial_{yy})L \qquad (14)$$

$$g(x, y; t) = \frac{1}{2\pi t} e^{-(x^2+y^2)/2t} \qquad .(15)$$

We tried pointwise matching in the scale-space; however, performance is poor as similar patterns can be found at low resolution levels. The solution to this problem is combining four neighboring jets into the descriptor. The results of scale invariant matching are shown in Fig.15.

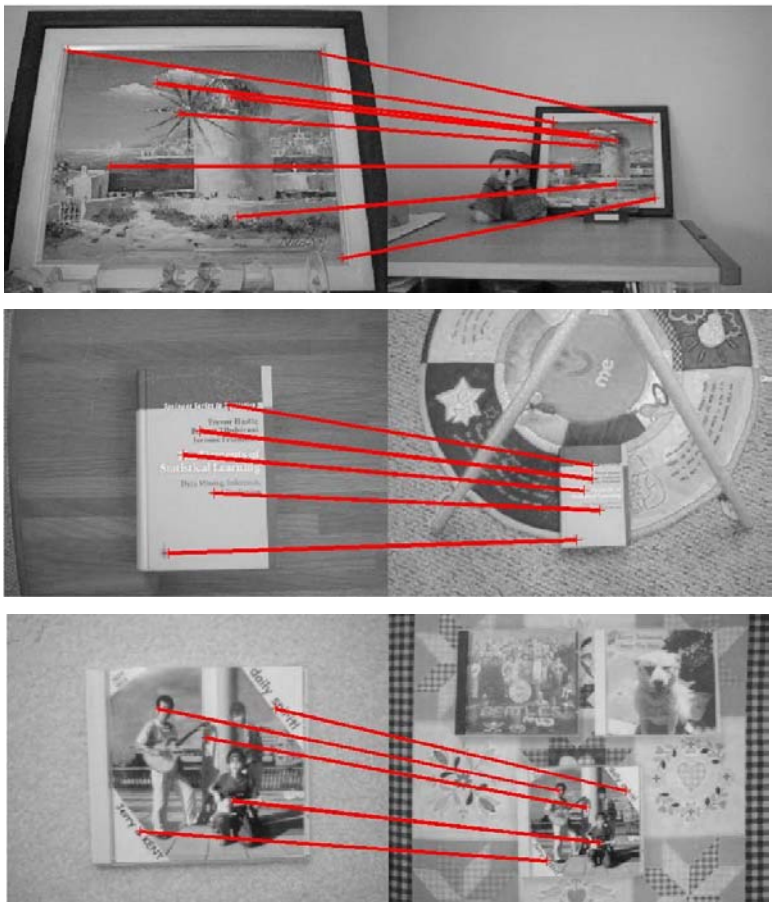


Fig. 15 Scale invariant correspondence with the descriptor

7. Conclusion and Future Work

In this paper we consider several descriptors: the first consists of Gaussian derivatives up to the third order, the second of the gray image patch, and the third of Laplacian filters with either three scales or one scale. We compare their performance in terms of selectivity and invariance to several affine changes such as rotation, scale changes, brightness changes, and viewpoint changes. The results show a good performance by the Gaussian derivative descriptor. It is interesting to point out that oriented receptive fields similar to the Gaussian derivatives as well as receptive fields similar to the Laplacian are found in human visual cortex.

We are now working on an object recognition system using the descriptor with the oriented Gaussian derivatives.

Acknowledgements

This research has been done during the visit to CBCL with the support of Sony Corporation.

References

[1] A.Yuille and T.Poggio. “Scaling Theorems for Zero Crossings”. IEEE Trans. Pattern Analysis and Machine Intell., vol.8, pp.15-25, 1986

[2] A.Yuille and T.Poggio. “Scaling and Fingerprint Theorems for Zero-Crossings”. Advances in Computer Vision, pp.47-78, Lawrence Erlbaum, 1988

[3] T.Lindeberg. “Scale-Space Theory in Computer Vision”. The Kluwer International Series in Engineering and Computer Science, Kluwer Academic Publishers, 1994

[4] G. Carneiro and A. D. Jepson. “Phase-based local features”. ECCV, volume I, pp. 282–296, 2002.

[5] Gustavo Carneiro and Allan D. Jepson. “Multi-scale Local Phase-based Features”. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Madison, WI, USA. 2003.

[6] W. Freeman and E. Adelson. “The design and use of steerable filters”. PAMI, 13(9):891–906, 1991.

[7] R. Rao and D. Ballard. “Object indexing using an iconic sparse distributed memory”. ICCV, pp.24-31, 1995

[8] C. Harris and M. Stephens. “A combined corner and edge detector”. In Alvey Vision Conference, pp. 147–151, 1988.

[9] D. G. Lowe. “Object recognition from local scale-invariant features”. In ICCV, pp. 1150–1157, 1999.

[10] K. Mikolajczyk and C. Schmid. “Indexing based on scale invariant interest points”. In ICCV, pp. 525–531, 2001.

[11] C. Schmid and R. Mohr. “Local gray value invariants for image retrieval”. PAMI, 19(5):530–534, 1997.

[12] C. Schmid, R. Mohr, and C. Bauckhage. “Evaluation of interest point detectors”. IJCV, 37(2):151–172, 2000.

[13] J.G.Daugman. “Two-dimensional Spectral Analysis of Cortical Receptive Field Profiles”. Vision Research, vol.20, pages 847-856, 1980

[14] J.G.Daugman. “Complete Discrete 2-D Gabor Transforms by Neural Networks for Image Analysis and Compression”. IEEE Transactions on Acoustics, Speech and Signal Processing, vol.36, pages 1169-1179, 1988

[15] M.Lades, J.C.Volbruggen, J.Buhmann, J.Lange, C.Malsburg, R.P.Wurtz and W.Konen. “Distortion Invariant Object Recognition in the Dynamic Link Architecture”. IEEE Transactions on Computers. vol.42, No.3, pages 300-310, 1993

[16] M.Potzsch, N.Kruger and C.Malsburg. “Improving Object Recognition by Transforming Gabor Filter Response”. Computation in Neural Systems, Vol.7, No.2, pages 341-347, 1996

[17] B. Schiele and J. L. Crowley. “Probabilistic object recognition using multidimensional receptive field histogram”. In Proc. Int. Conf. on Patt. Recog., vol. B, pp 50-54, Vienna, August 1996.

[18] Bernt Schiele and James L. Crowley. “Object recognition using multidimensional receptive field histograms”. In ECCV (1), pages 610--619, 1996

[19] B. Schiele and J. L. Crowley. “Recognition without correspondence using multidimensional receptive field histograms”. International Journal of Computer Vision, 36(1):31--50, January 2000. 6

[20] B.W. Mel, "SEEMORE: Combining Color, Shape and Texture Histogramming in a Neurally Inspired Approach to Visual Object Recognition", Neural Computation, 9, pp. 777-804, 1997.

[21] Heisele, B., P. Ho, J. Wu and T. Poggio. “Face Recognition: Component-based versus Global Approaches”. Computer Vision and Image Understanding, Vol. 91, No. 1/2, 6-21, 2003.

[22] Heisele, B., P. Ho and T. Poggio. “Face Recognition with Support Vector Machines: Global Versus Component-based Approach”. International Conference on Computer Vision (ICCV'01), Vancouver, Canada, Vol. 2, 688-694, 2001.

[23] Schneiderman, H. and Kanade, T. “A statistical approach to 3d object detection applied to faces and cars”. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), pages 746--751, Hilton Head Island, South Carolina, June 2000.

[24] S. A. Nene, S. K. Nayar and H. Murase, “Columbia Object Image Library (COIL-100)”. Technical Report CUCS-006-96, February 1996.

[25] K. Sung and T. Poggio, "Example-based Learning for View-based Human Face Detection," IEEE TPAMI, Vol. 20(1), pp 39-51, January 1998.

[26] B. Heisele, T. Poggio, and M. Pontil, “Face detection in still gray images”. Technical Report A.I. Memo No. 2001-010, C.B.C.L. Memo No. 197, MIT Center for Biological and Computational Learning (2000)

.

[27] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," IEEE CVPR, 1997.

[28] Paul Viola and Michael J. Jones. “Robust real-time object detection”. Technical Report CRL 2001.

[29] Papageorgiou, C., M. Oren, and T. Poggio. “A General Framework for Object Detection”. In: Proceedings of the Sixth International Conference on Computer Vision (ICCV'98), Bombay, India, 555-562, January 1998.