to be published in Proc. of the IEEE Int. Symposium on Ciruits and Systems, ISCAS'98

THE SVD APPROACH FOR STEERABLE FILTER DESIGN*

Gerald Sommer[1], Markus Michaelis[2], and Rainer Herpers[1,3]

[1] Cognitive Systems Group, Computer Science Institute,
University Kiel, D-24105 Kiel, Germany Email: [gs, rhe]@informatik.uni-kiel.de
[2] Plettac Electronics, D-90766 Fürth, Germany
[3] GSF – Institute of Medical Informatics and Health Services Research, MEDIS
D-85764 Neuherberg, Germany

ABSTRACT

The first processing step in computational early vision usually consists of convolutions with a number of kernels. These kernels often are derived from a mother kernel that is rotated, scaled, or deformed with respect to other degrees of freedom. This paper presents an efficient computational approach to calculate the responses of arbitrary mother kernels with arbitrary deformations. Analytical solutions to this problem in most cases are difficult or not possible. Therefore, we present in this paper a numerical approach that emphasizes an algebraical point of view.

1. INTRODUCTION

It seems to be a natural demand on the early vision part of a general vision system to provide explicit information of images about orientation, scale, and other degrees of freedom. Since a few years there has been a growing interest in generating continuous responses of filters with respect to these degrees of freedom. Instead of resulting in a tremendous computational burden the application of the steerable filter approach scales only linearly with the number N of used basis functions.

The idea of steerable filtering [2, 11, 13] is the following. Let $F(\vec{x})$ denote the mother kernel with $\vec{x} \in \mathbb{R}^n$ or $\vec{x} \in \mathbb{Z}^n$ and $F_\alpha(\vec{x})$ the deformed kernels with α denoting one or several deformations of the mother kernel. These deformations correspond to the chosen degrees of freedom. 'Steerability' then refers to a reconstruction formula for the kernels:

$$F_\alpha(\vec{x}) = \sum_{k=1}^{N} b_k(\alpha) A_k(\vec{x}) \tag{1}$$

The kernels $A_k(\vec{x})$ are called basis functions, the $b_k(\alpha)$ are called interpolation functions. By (1) the response of an image I to a set of kernels $\{F_\alpha, \alpha\}$ in a (quasi) continuum of the parameter α is calculated by N projections: $\langle I|F_\alpha\rangle = \sum_{k=1}^{N} b_k(\alpha)\langle I|A_k\rangle$.

There are different aspects of steerability which are of interest and which have been addressed in the literature. One may ask for kernels F that are exactly steerable with a small number of basis functions. Another problem is the approximated reconstruction for a given kernel. The basis functions may be predefined or one asks for optimal basis functions for a given kernel and/or deformation.

A general approach for exact steerability is the Lie group approach introduced by Michaelis & Sommer in [8, 9]. There, the basis functions are defined by the invariant subspaces of the deformation Lie groups. A Lie group approach to steerability has also been suggested by Hel-Or & Teo [3].

In practice, the Lie group approach turns out to have several drawbacks [10]. The main drawback is that in the Lie group approach basis functions are predefined by the deformation. They are independent of the kernel. In practice, however, often the kernels are given and one is looking for an appropriate steering equation.

Perona [12] proposed an approach to steerability based on the singular value decomposition (SVD) which does not have the drawbacks of the Lie group approach. The basis functions and interpolation functions for a given kernel and deformation are obtained as the left and right singular vectors of the SVD of a linear operator $\mathcal{L}$. This operator is defined by mapping images I to the continuous responses of the deformed kernels: $\mathcal{L}I(\alpha) := \langle F_\alpha | I\rangle$. For kernels and deformations where the Lie group approach needs an infinite number of basis functions this approach has an infinite number of singular components (i.e. basis functions) as well, if exact steerability is required. If a certain L^2-error is tolerated, however, it guarantees the minimal number of basis functions.

Although the concept of steerability could improve many early vision methods, it is far from being a standard tool in computer vision. From our point of view the best approach for most applications is the one of Perona [12]. However, this approach is hardly used by others. One reason might be that Perona addresses the case of infinite dimensional function spaces. In practice, however, finite dimensional approximations are needed that allow for a numerical solution. Moreover, in practice the images and kernels are discrete and hence they are restricted to finite dimensional vector spaces anyway. The main distinguishing feature of our paper (see also [10]) in contrast to the original Perona's formulation is that we explicitly discuss the numerical implementation of the method in finite dimensional function spaces.

On that base the steering of no more than two simul-

*This work is partially supported by the DFG, Grants So 320/1-2, Ei 322/1-2, and He 2967/1-1.

taneous deformations of 2D kernels turns out to be simple using a standard procedure that we will present in section 2. We will discuss the general calculation of optimal basis functions to approximate steerability in section 3 and also the handling of higher dimensional kernels together with more deformations in section 4. In addition, we will discuss the benefit of using orthogonal basis functions instead of deformed copies of the mother kernels. Finally, we will present in section 5 some consequences of steerable filters for the design of computational early vision strategies. For a more complete and by formal proofs enhanced version of this contribution we refer the reader to [10].

2. A STANDARD PROCEDURE

The intention of this section is to demonstrate that the application of the theory of steerability by SVD actually reduces to a few formulas that are straight forward and simple to implement. This example, however, will suffice in most cases to steer arbitrary 2D-kernels with no more than two deformations.

Let F_l denote the deformed filters, where 'l' samples all deformations together (e.g. orientation and scale). The continuous (multi-) parameter is denoted by α. The matrix $(\langle F_k|F_l\rangle)$ is called the Gramsian matrix. It is real and symmetric and therefore, it has a complete set of eigenvectors. The eigen decomposition of the Gramsian matrix is denoted by

$$\langle F_{l'}|F_l\rangle \;=\; \sum_k u_{k,l'}\sigma_k u_{k,l} \tag{2}$$

where u_k denotes the k'th eigenvector with eigenvalue σ_k. The optimal (and orthogonal) basis functions $A_k(x,y)$ are given by

$$A_k = \sum_l u_{k,l}F_l \tag{3}$$

'Optimality' here means that a minimal number of basis functions is necessary for a given L^2 error. The interpolation functions for arbitrary deformations (α) are given by

$$b_k(\alpha) = \frac{\langle F_\alpha|A_k\rangle}{||A_k||^2} \tag{4}$$

The sampling of α in (4) is independent of the sampling $F_l \equiv F_{\alpha_l}$ for the Gramsian matrix in (2). The reconstruction equation now reads:

$$F_\alpha \approx \sum_k b_k(\alpha)A_k \tag{5}$$

It should be emphasized that the calculation of the b_k and A_k by (2), (3), and (4) is done once off-line and thus, it is not time critical for real time applications.

For most practical cases when no more than two deformations are steered this will be the easiest approach. In other cases, however, the Gramsian matrix can become too large or too singular. In the following we will discuss the theory in more detail. This will not only give a deeper understanding of the method but provides also some more ideas to calculate the optimal basis functions in more complex situations.

3. OPTIMAL STEERING EQUATIONS IN FINITE DIMENSIONAL SPACE

3.1. Calculation of the interpolation functions

Given a set of basis functions, the interpolation functions are calculated by (4) or its generalization for nonorthogonal basis functions.

$$\begin{pmatrix} b_1(\alpha) \\ \vdots \\ b_N(\alpha) \end{pmatrix} \;=\; G^{-1}\begin{pmatrix} \langle F_\alpha|A_1\rangle \\ \vdots \\ \langle F_\alpha|A_N\rangle \end{pmatrix} \tag{6}$$

Here, G denotes the Gramsian matrix for the basis functions $G = \langle A_k|A_{k'}\rangle$. The calculation of the interpolation functions is done off-line and only once for a given set of basis functions.

3.2. Calculation of optimal basis functions

The number M of basis functions for the exact reconstruction of all kernels F_α is clearly the dimension of the vector space spanned by all F_α:

$$M \;=\; \dim(\mathrm{span}\{F_\alpha,\alpha\}) \tag{7}$$

Unfortunately this number will be infinite in general and approximated solutions are needed in practice. Therefore, we investigate the special problem where we admit approximated reconstructions of the deformed kernels F_α using a small number N of basis functions ($N \ll M$).

$$F_\alpha(\vec{x}) \;\approx\; \sum_{k=1}^{N} b_k(\alpha)A_k(\vec{x}) \tag{8}$$

The task is then to find optimal basis functions, so that for a given approximation error a minimal number of basis functions is needed or vice versa that for a given number of basis functions a minimal error is guaranteed.

We investigate numerical solutions what means embedding the problem in a finite dimensional space. Let $\{Z_m\}$ be an orthonormal basis that spans such a finite dimensional space. All functions then are given in Z-representation, i.e. by the real numbers $\langle F_\alpha|Z_m\rangle$ and $\langle A_k|Z_m\rangle$. For $\{Z_m\}$ one can choose e.g. the Dirac base, the canonical basis functions of the Lie group approach [9], or the Fourier base. There is no standard optimal choise because it depends on the filters and deformations. But any basis that approximately spans all deformed kernels works and usually it is not difficult to find a worthwhile one. In addition to the finite basis $\{Z_m\}$ we discretize the deformation (multi-) parameter, $\alpha \to \alpha_l \equiv l$.

As our optimality criterium for a set of basis functions $\{A_k\}$ we introduce the following distance that sums up the least square reconstruction errors of all deformed kernels embedded in the finite dimensional space spanned by $\{Z_m\}$.

$$\mathrm{d}(\{F_l\},\{A_k\}) = \sum_{l,m}\Big(\langle F_l|Z_m\rangle - \sum_k \langle F_l|A_k\rangle\langle A_k|Z_m\rangle\Big)^2 \tag{9}$$

The r.h.s. of (9) is a dyadic approximation of the Z-expansion matrix $\Lambda := \langle F_l|Z_m\rangle$. Least square optimal

dyadic matrix approximations are obtained by the SVD. This motivates the following central statement that is derived in more detail in [10].
Let $\{F_l\}$ be any set of kernels and $\langle F_l|Z_m\rangle = \sum_k u_{k,l}\sigma_k v_{k,m}$ the SVD of its Z-expansion matrix. Then, the optimal orthonormal basis functions are the right singular functions $A_k = \sum_m v_{k,m} Z_m$.
Here, u_k, v_k, and σ_k denote the left singular vectors, right singular vectors, and singular values for the matrix Λ. The corresponding interpolation functions are calculated by (6).

3.3. Reconstruction errors

Qualitatively two types of errors are possible. First, the reconstructed kernel differs from the original mother kernel. However, this error could be independent of the deformation, i.e. the approximated kernel itself could be steered exactly. The second type of error depends on the deformation so that the reconstructed deformed kernels have varying shapes. If these variations are too strong they result in disturbing structures in the filter response that are caused by the steering scheme and not by the local image structure.

The total discretized distance measure (9) sums up the error of all deformations and is a measure for the first type error:

$$d(\{F_l\}, \{A_k, k = 1 \dots N\}) \; = \; \sum_{k=N+1}^{L} \sigma_k^2 \qquad (10)$$

The deformation-dependent error reads:

$$d_l(F_l, \{A_k, k = 1 \dots N\}) \approx \sum_{k=N+1}^{L} \sigma_k^2 u_{k,l}^2 \widehat{=} \sum_{k=N+1}^{L} b_k^2(\alpha_l) \qquad (11)$$

Depending on the sigular vector u_k the maximum distance with respect to α can differ significantly from the average distance. The SVD basis functions are not optimal to minimize this maximum distance or to guarantee that it will be below a certain threshold.

4. ALTERNATIVE MATRIX REPRESENTATIONS

The Z-expansion matrices $\Lambda = (\langle F_l|Z_m\rangle)$ have as many rows as there are samples for the deformations and as many columns as auxiliary basis functions Z_m are needed to span all kernels. In the case of 3D kernels or if several deformations are steered simultaneously the SVD can become computationally too expensive. Therefore, we discuss different matrix representations which are advantageous to handle these cases.

4.1. The Gramsian matrix ($\Lambda\Lambda^T$)

A matrix that avoids the auxiliary basis $\{Z_m\}$ is the Gramsian matrix $(\langle F_{l'}|F_l\rangle)$ of the deformed kernels. We have $(\langle F_{l'}|F_l\rangle) = (\sum_m \langle F_{l'}|Z_m\rangle\langle Z_m|F_l\rangle) = \Lambda\Lambda^T$. If the SVD is denoted $\Lambda = U\Sigma V^T$, we obtain $\Lambda\Lambda^T = U\Sigma V^T V\Sigma U^T = U\Sigma^2 U^T$, i.e. any set of kernels $\{F_l\}$ the Gramsian matrix has the SVD $\langle F_{l'}|F_l\rangle = \sum_k u_{k,l'}\sigma_k^2 u_{k,l}$. The optimal basis functions are then $A_k = \sum_l u_{k,l} F_l$. In contrast to the Z-representation the basis functions here have absorbed the weights (singular values), $A_k = \sum_l u_{k,l} F_l = \sigma_k \sum_m v_{k,m} Z_m$.

4.2. The matrix $\Lambda^T\Lambda$

When steering more than two deformations simultaneously, we encounter the same problem as for steering nD filters. The concatenated sampling of all parameters results in very large columns of the matrix Λ. In this case the matrix $\Lambda^T\Lambda = (\sum_l \langle Z_{m'}|F_l\rangle\langle F_l|Z_m\rangle)$ can be used to calculate the basis functions.
Let $\{F_l\}$ be any set of kernels and $\langle F_l|Z_m\rangle = \sum_k u_{k,l}\sigma_k v_{k,m}$ the SVD of its Z-expansion matrix Λ. Then the SVD of $\Lambda^T\Lambda$ is given by $(\Lambda^T\Lambda)_{m,m'} = \sum_k v_{k,m}\sigma_k^2 v_{k,m'}$. The optimal (normalized) basis functions are $A_k = \sum_m v_{k,m} Z_m$ which is the same as in Z-representation. Like for the Gramsian matrix, $\Lambda^T\Lambda$ has squared singular values compared to Λ and therefore, it is more singular which may cause numerical problems.

4.3. Auxiliary functions

In case of steering several deformations it can be advantageous to give up the Dirac sampling of the (multi-) parameter α and to use an orthonormal and complete base $\{c_l(\alpha)\}$ of the parameter space instead (or $\{c_{l,l'}\}$ in the discrete case). We define auxiliary functions B_l by

$$B_l \;\; := \;\; \sum_{l'} c_{l,l'} F_{l'} \qquad (12)$$

For an appropriate choice of $\{c_l\}$, there may be less B_l significantly different from zero than there are F_l. If $\{c_l(\alpha)\}$ is an orthonormal basis, e.g. the Fourier basis, the two matrices $\langle F_l|Z_m\rangle$ and $\langle B_l|Z_m\rangle$ have the same right singular vectors and singular values and therefore, both sets of kernels have the same optimal basis functions. Both sets span the same vector space with the same L^2-weights of its components. In this case we will call them 'steering equivalent'.

4.4. Splitting large matrices

If the matrices Λ, $\Lambda\Lambda^T$, or $\Lambda^T\Lambda$ are still too large to be handled numerically we can split the matrix according to the following scheme.

A special case of steering equivalent sets is the original set of kernels and the optimal basis functions. Hence, the SVD for a large matrix with M rows can be performed by splitting it in N smaller matrices, each containing approximately M/N rows. Then put the weighted right singular vectors $\sigma_k v_k$ from all submatrices in a new large matrix and perform the SVD for this matrix. The following statement expresses that the resulting matrix will provide the correct basis functions.

Let $\{B^n_{l'}, l'\}$, $\{F^n_l, l\}, n = 1, \dots, N$ be N sets of functions, where for each n, $\{B^n_{l'}, l'\}$ and $\{F^n_l, l\}$ are steering equivalent. Then, the two sets $\{B^n_{l'}, l', n\}$ and $\{F^n_l, l, n\}$, where each is the union of all the N 'small' sets, are steering equivalent too.

So far nothing is gained because the new large matrix has the same size as the original one. To obtain a smaller matrix we sort all singular values of all partial matrices by magnitude (one queue for all) and drop all singular components below a threshold. This procedure will introduce some unknown errors. However, usually the error is small and it has the order of magnitude of the threshold.

5. STEERABLE FILTERS FOR EARLY VISION

Usually steerable filters are only used as a black box tool for the efficient calculation of signatures with respect to certain deformations. However, the following ideas suggest a deeper impact on computational early vision methods. Especially the design of visual architectures following behaviour based approaches requires to support attentive, purposive, and progressive recognition [14]. All three aspects are intrinsic features of steerable filters if they are interpreted as projection or matching operators.

5.1. Progressive filtering scheme

Corresponding to the principle of economy of time a progressive scheme of recognition allows to abandon the process if the measure of the confidence or evidence is sufficiently high. The template in this frame is the equivalence class which is a purposively constrained approximation of the ideal one. That necessitates a gradual decomposable template to allow for a gradual increase of matching results.

The principle of progression is inherent to the steering filter approach of recognition, especially to the SVD approach. This results from the orthogonality and ordinality of basis functions. Indeed, any number of basis functions reconstruct all deformed filters. Only the quality of the reconstruction changes (fig. 1, middle box). Therefore, the whole parameter range of the deformations is covered with any number of basis functions. This allows for an adaptation of the reconstruction of the deformed mother kernel to the required quality of the equivalence class with respect to the purpose. The purpose of the task allows for a modulation of the reconstruction within a feedback loop [14].

5.2. Hierarchical filtering scheme

Inherent to the proposed architecture of early vision is a generic scheme of grouping, realized as a layered architecture with feedback. At the bottom of this architecture a set of basis functions may constitute a set of irreducible, invariant mother kernels. At higher levels aggregations of these primitive kernels synthesize more complex kernels with any useful invariance properties [1]. May be that there is the one for the grandmothers face. Figure 1 (bottom box) shows a simple example of constructing several filters from the one in figure 1 (middle box). This approach of synthesizing successfully complex steerable templates from primitive ones (fig. 1, middle to bottom box) has been used for face processing applications [4].

5.3. Sequential filtering scheme

A set of several steerable filters embedded in a purposive controlled feedback loop allows for a local selection (with respect the structural features) and adaptation (with respect to the degrees of freedom) of the mother kernels to respond optimally to the local signal structure. With other words, non stationary filtering with all the above mentioned features of progression and hierarchy may be organized as a very efficient tool. Edge tracking (e.g. fig. 1, bottom right) with translation, rotation, and scale as degrees of freedom has been successfully implemented for the detailed analysis of facial regions [7]. Figure 2 illustrates several examples of recognizing facial structures and anatomical landmark positions.

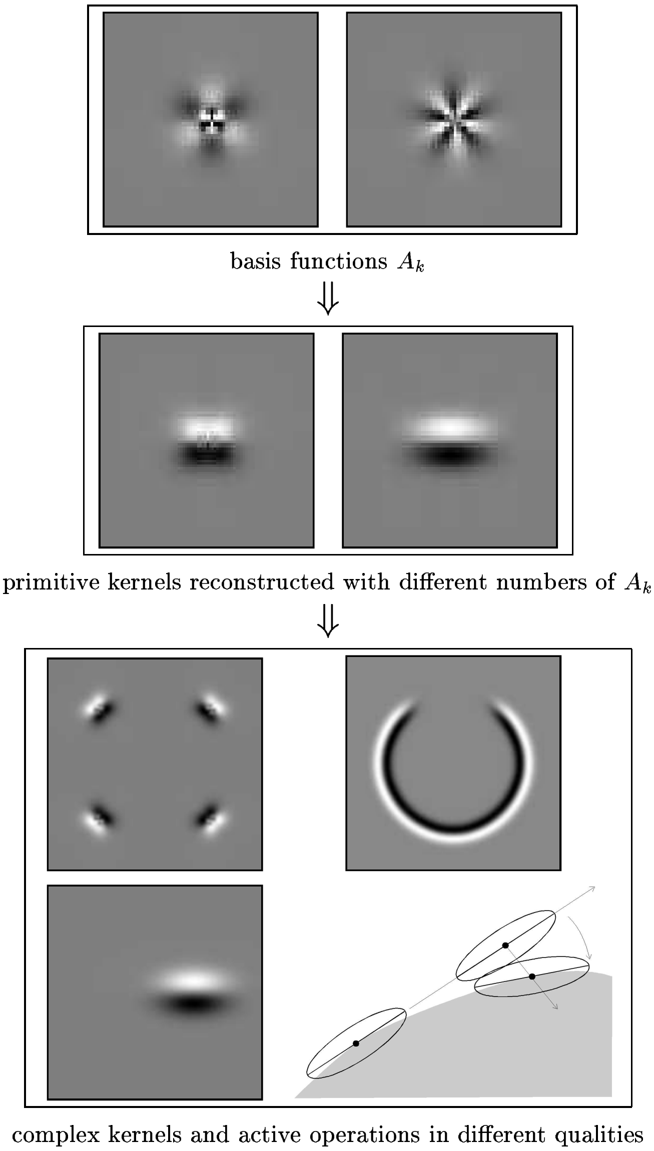

Figure 1: Fundamental idea of the steerable filtering scheme. First step (top box): examples of basis functions. Second step (middle box): primitive kernels are reconstructed with different numbers of basis functions. Third step (bottom box): circular kernel with 4 primitive kernels and 10 basis functions for each (top left). Circular kernel with 28 primitive kernels and 30 basis functions for each (top right). The gap in the circular kernel is motivated by its use as an iris-detector. An one-sided kernel that is rotated around a shifted center (bottom left). An active edge tracking operation that searches for the next edge element (bottom right).

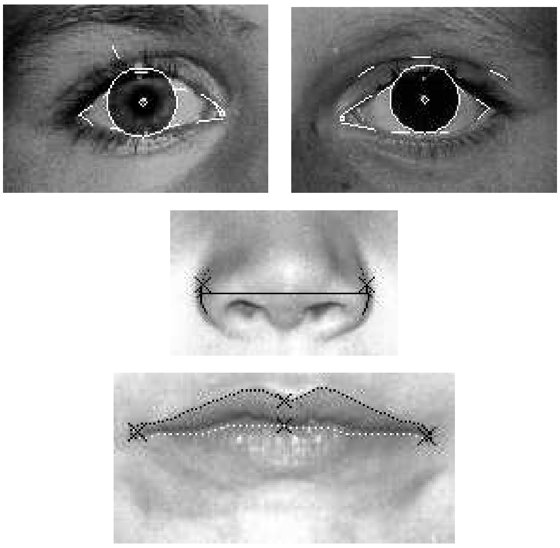

Figure 2: Results of the sequential filtering scheme applied to several facial regions to detect anatomical landmark positions. An initial detection of prominent edge or line features is performed applying model knowledge, followed by a sequential edge and line tracking. During all these processing steps the filters used are adapted optimally to the underlying image structure to ensure optimal filter responses.

6. CONCLUSIONS

In this paper we discussed the numerical approach of steerable filter design based on SVD. The benefits of the SVD method in comparison to the algebraic (Lie group) approach [9] are:

- It steers any set of filters, not only continuous deformations of a mother filter.
- A minimal number of basis functions is required (with respect to the overall L^2-error). This number is significantly smaller than for the Lie group approach.
- For rotations and periodic translations it is identical to the Lie group approach. Hence, (partial) analytical solutions are possible and the steering is exact in these cases.
- The numerical calculation of the basis functions (versus analytical in the case of the Lie group approach) allows to apply simple black box tools to any new kernel and deformations without difficult analytical calculations.

We want to mention that steerable filters have no intrinsic drawbacks and they are not intrinsically expensive. The off-line and on-line extra costs and the complexity are 'scalable'. On that base we presented some adaptive processing schemes which apply the embedding of steerable filters in purposive early vision. Our experience with such implementations allows us to recommend steerable filters also in time-critical applications as we have shown in a face processing example [4, 5, 6].

7. ACKNOWLEDGEMENTS

This work has been partially supported by the DFG, Grants So 320/1-2, Ei 322/1-2, and He 2967/1-1. We are also grateful to our graduate students L. Witta, D. Müller and K.-H. Lichtenauer.

8. REFERENCES

[1] W. Beil, Steerable filters and invariance theory, Pattern Recognition Letters, Vol.15, 1994, 453-460.

[2] W.T. Freeman and E.H. Adelson, The design and use of steerable filters for image analysis, IEEE Trans. PAMI, Vol.13, 1991, 891-906.

[3] Y. Hel-Or, P.C. Teo, Canonical decomposition of steerable functions, Proc. IEEE CVPR, 1996, 809-816.

[4] R. Herpers, M. Michaelis, G. Sommer, and L. Witta, Context based detection of keypoints and features in eye regions, Proc. IEEE ICPR, Vol. B, 1996, 23-28.

[5] R. Herpers, H. Kattner, H. Rodax, and G. Sommer, GAZE: An attentional processing strategy to detect and analyze the prominent facial regions, Proc. of the Int. Workshop on Autom. Face- and Gesture-Rec., M. Bichsel (ed.), Zurich, Switzerland, 1995, 214-220.

[6] R. Herpers, M. Michaelis, K.H. Lichtenauer and G. Sommer, Edge and keypoint detection in facial regions, Proc. IEEE 2. int. Conf. on Autom. Face- and Gesture Rec., 1996, 212-217.

[7] R. Herpers, GAZE: A common attentive processing strategy for the detection and investigation of salient image regions, PhD thesis, also Tech. Rep. No. 9714, Christian-Albrechts-Universität, Inst. für Informatik, D-24105 Kiel, Germany, 1997.

[8] M. Michaelis, Low level image processing using steerable filters, PhD thesis, Christian-Albrechts-Universität Kiel, Germany, 1995, also Tech. Rep. No. 9716, Christian-Albrechts-Universität, Inst. für Informatik, D-24105 Kiel, Germany, 1997.

[9] M. Michaelis and G. Sommer, A Lie group approach to steerable filters, Pattern Recognition Letters, Vol.16, 1995, 1165-1174.

[10] M. Michaelis and G. Sommer, Steerable filters in finite dimensional function spaces, Tech. Rep. No. 9715, Christian-Albrechts-Universität, Inst. für Informatik, ftp://ftp.informatik.uni-kiel.de/pub/kiel/publications /CognitiveSystems/, D-24105 Kiel, Germany, 1997.

[11] P. Perona, Steerable-scalable kernels for edge detection and junction analysis, Proc. ECCV, LNCS Vol. 588, Springer-Verlag, Berlin, 1992, 3-18.

[12] P. Perona, Deformable kernels for early vision, IEEE Trans. PAMI, Vol.17, 1995, 488-499.

[13] E. Simoncelli, W.T. Freeman, E.H. Adelson, and D.J. Heeger, Shiftable multi-scale transforms, IEEE Trans. IT, Vol.38, 1992, 587-607.

[14] G. Sommer, Algebraic aspects of designing behaviour based systems. Proc. Algebraic Frames for the Perception-Action Cycle, LNCS Vol. 1315, Springer-Verlag, Berlin, 1997, 1-28.