Abstract—Starting from a biologically inspired framework, Gabor filters were built up from retinal filters via LMSE algorithms. A subset of retinal filter kernels was chosen to form a particular Gabor filter by using a weighted sum. One-dimensional optimization approaches were shown to be inappropriate for the problem. All model parameters were fixed with biological or image processing constraints. Detailed analysis of the optimization procedure led to the introduction of a minimization constraint. Finally, quantization of weighting factors was investigated. This resulted in an optimized cascaded structure of a Gabor filter bank implementation with lower computational cost.

Keywords—Gabor filter, image processing, optimization

I. INTRODUCTION

GABOR filters have desirable properties for picture analysis and feature extraction: They are selective in space, spatial frequency and orientation, achieving the theoretical limit for conjoint resolution in the spatial and spatial frequency domain ([1]). Therefore, they have been widely used in these fields in recent years ([2], [3], [4]). Those filters were also used to describe the behavior of simple cells in area V1 of the human visual cortex, which has turned out to be very successful ([5]). It is thus worth going back to the roots and taking a look on how Gabor filters are composed in nature.

The human retina gathers optical information about the environment via its $10^8$ photoreceptors. Horizontal, bipolar and amacrine cells filter the information, and the $10^6$ ganglion cells code the filter outputs into spike trains, sending them to the optic nerve. In this process, a lot of information compression is needed ([6], [7]) due to the relatively few number of outputs. To achieve this, the overall retinal filtering consists of lowpass and bandpass filtering, emphasizing changes in illumination and suppressing regions with uniform brightness. The next processing stage in the human visual system is area V1 of the visual cortex. There, basic features of the image are extracted by combining retinal outputs, so that overall Gabor-like receptive fields arise ([8]).

The idea of building Gabor filters from retinal responses is not new; it was already stated by Thiem et al. ([9], [10]). We use their principal idea of building Gabor filters by linearly combining outputs of retinal filters, using least mean square error (LMSE) optimization. In contrast to their emphasis on the system’s architecture and the hexagonal sampling, we here investigate the optimization process in more detail. This includes adjusting the retinal filter parameters, analysing and extending the standard LMSE optimization, and testing the influence of quantised weighting factors on the optimization results. In particular, we show that one-dimensional optimization approaches are insufficient for the problem, and introduce a minimization equation leading to a far better weighting factor distribution. The latter is a prerequisite for successfully quantising the weights. From the quantisation results, we propose a simplified cascaded structure for Gabor filter banks compared to that introduced in [10], and compare the computational cost of the approach with standard Gabor filter bank implementations.

In section II, we develop a model for the retinal filters. Section III will introduce a Gabor filter model and describe and analyse the different optimization procedures. Finally, section IV shows how quantization of coefficients influences the results and describes the simplified cascaded structure of a Gabor filter bank.

II. MODEL OF THE RETINA

In a simple approach, the retina as a whole can be seen as performing a discrete convolution of the input image with a retinal filter kernel. In other words, it could be modeled by a discrete, linear filter. Each output of a ganglion cell represents the filter output at a particular position, coded in a spike train.

The retinal filter kernels have a center-surround structure, i.e. center and surround have opposite sign (see [11], for example). As negative filter outputs are hard to code into a spike train by retinal ganglion cells, there exist cell pairs for each filter output: The so-called ON-center cell codes the positive values, and the OFF-center cell codes the negative values using a negated filter kernel ([12]).

A common model for the filter kernels is the Difference-of-Gaussian (DoG) model. It consists of two Gaussians with different variances and can generally be written as

$$d(x, y) = a_1 \cdot e^{-\frac{x^2+y^2}{2\sigma_1^2}} - a_2 \cdot e^{-\frac{x^2+y^2}{2\sigma_2^2}}.$$  

In our model, the filter will be pure bandpass, meaning that uniform brightness will be generating zero filter output. To ensure this, the integral over the filter kernel has to be zero as well. Because the overall amplitude of the DoG has no influence on the analysis and is adjusted later by the coefficients used to combine the retinal filter kernels to Gabor filters, we simply set the maximum of the first Gaussian to 1.

Then we receive the form

$$d(x, y) = e^{-\frac{x^2+y^2}{2\sigma_1^2}} - \frac{1}{a^2} \cdot e^{-\frac{x^2+y^2}{2a^2\sigma_1^2}}, \text{ with } a = \frac{\sigma_2}{\sigma_1}. \quad (2)$$

Whereas $\sigma_1$ fixes the spatial size of the filter kernel, $a$ defines its shape. In contrast to what results of biological experiments
suggest (e.g. [12]), we will not vary \( a \) in our model, because it is not thought to be biologically exact but aims for efficient signal processing. According to [13], we choose \( a = 1, 6 \) to approximate the Laplacian-of-a-Gaussian (LoG) model, which was also used by [10]. The retinal filter according to this value is shown in picture 1A.

To fix the value of \( \sigma_1 \), we consider the Fourier transform of the DoG, which is

\[
D(u, v) = 2\pi\sigma_1^2 \left( e^{-2\sigma_1^2(u^2+v^2)} - e^{-2a^2\sigma_1^2(u^2+v^2)} \right). \tag{3}
\]

As can be seen in figure 1C, the DoG behaves like a bandpass filter with the mid frequency proportional to \( \frac{1}{\sigma_1} \). The Gabor filter is also bandpass, so a constraint on \( \sigma_1 \) is that the frequency range of the DoG filters lies within the frequency range of the Gabor filter. After introducing Gabor filtering in the next section, we will describe this a bit more precisely.

III. SYNTHESIS OF GABOR FILTERS

Before starting to compose Gabor filters, we have to describe them mathematically. Lee ([1]) has derived a Gabor filter model that is based on biological experiments, but also considers constraints from information processing, such as zero mean or \( L^2 \) normalization. It has the form

\[
g(x, y) = \frac{\omega_0}{\sqrt{\pi d \cdot k}} \cdot e^{-\frac{\omega_0}{2\pi}(x^2 + y^2)} \cdot (e^{j\omega_0 x} - e^{-\frac{k^2}{\omega_0} x^2}) \tag{4}
\]

and its spectrum is given by

\[
G(u, v) = 2\sqrt{\pi d \cdot k} \frac{\omega_0}{\sqrt{2\pi}} \cdot e^{-\frac{k^2}{2\omega_0} u^2} \cdot e^{-\frac{\omega_0}{2\pi}(u^2 + v^2)} \cdot (e^{j\omega_0 u} - e^{-\frac{k^2}{\omega_0} u^2} \cdot e^{-\frac{\omega_0}{2\pi} v^2}) \tag{5}
\]

Both the filter and its spectrum are displayed in figure 2.

As the filter function is complex, real and imaginary part have to be treated separately. In fact, simple cells in V1 are arranged in pairs, with each pair approximately having a quadrature characteristic ([11], [14]). The Gabor function consists of a two-dimensional elliptic Gaussian, spreading along the y-axis, and a zero-mean complex oscillation along the x-axis. Its spatial size and spatial mid-frequency can be adjusted by \( \omega_0 \), whereas the parameter \( k \) defines the spatial frequency bandwidth and \( d \) the length-to-width-ratio of the elliptic Gaussian. According to [1], we choose \( d = 2 \) and a bandwidth of 1.5 octaves, which corresponds to \( k = 2.5 \). For different orientations of the Gabor filter, the local filter coordinates \((x, y)\) have to be rotated with respect to the global image coordinates.

Now that we have described the Gabor filter kernels, we can start composing them from retinal DoG filters. Like Thiem et al. ([9]), we use a linear approach:

\[
g(x, y) \approx \sum_{j} a_j \cdot d_{\sigma_j}(x, y), \tag{6}
\]

where \( d_{\sigma_j} \) is the DoG kernel with standard deviation \( \sigma_1 = \sigma_j \). The coefficients \( a_j \) have to be determined by an optimization procedure. This can be done by defining a set of control points \((x_i, y_i)\). For each control point, (6) defines one equation of a linear system of equations. For good optimization results, there should be far more control points \((x_i, y_i)\) than coefficients \( a_j \); we used 10 times more control points than coefficients. The resulting overdetermined system of equations can be solved in a least mean square error (LMSE) sense.

One of the noteworthy properties of the Gabor filter is cartesian separability, i.e. the possibility to separate the function into a product \( g(x, y) = g_1(x) \cdot g_2(y) \), where \( x \) and \( y \) are oriented along the main axes of the Gabor function. This property was also found in receptive fields of V1 simple cells (see [8]). It could be used to simplify the optimization from a 2-dimensional problem into two 1-dimensional problems. A prerequisite for this method is cartesian separability of the basis functions, which is not the case for DoG-functions. So our results would suffer from an additional systematic error. Even when using another, cartesian separable approximation of a retinal filter kernel, there is a further drawback: The solution would lead to a rectangular grid of retinal cells,
oriented along the main axes of the Gabor filter. So one would have to choose whether to take a grid of retinal cells for each desired orientation of Gabors (there is at best a re-use of a grid for the 90°-rotated filters), which would be very inefficient, or whether to take a single grid and then use for each required retinal cell the nearest available position, which would likely be highly inaccurate or (if the grid was fine) again very inefficient. For this reasons, the 1D-approach is unsuitable for the problem.

With a more general 2D-approach, the positions of retinal filter cells can be chosen arbitrarily. For simplicity, we will use quadratic grids of retinal filter cells with different spatial and spatial frequency size, i.e. different \( \sigma_i \). A hexagonal grid would be more efficient in terms of the sampling theorem and is approximately used in the retina of mammals and humans ([9]), but that would be incompatible with common image processing and quadratically sampled digital images. Furthermore, the density of photoreceptors and ganglion cells in the retina is not constant ([13]), so a uniform grid can only be a coarse, local approximation to the arrangement in a human retina.

To determine the \( \sigma_i \), we consider the spectra of DoG and Gabor, shown in figures 1 and 2. As mentioned, the frequency range of the DoG filter has to be within the frequency range of the Gabor. Especially the maximum of the DoG amplitude spectrum has to lie inside the Gabor spectrum. To express this, we use the amplitude of frequency \( r = \sqrt{u^2 + v^2} \). As the DoG spectrum is radially symmetric, it becomes a 1-dimensional function \( D(r) \):

\[
D(r) = 2\pi\sigma_i^2 \left( e^{-2\sigma_i^2 r^2} - e^{-2\sigma_i^2 u^2 \sigma_i^2 v^2} \right) \tag{7}
\]

with maximum at:

\[
r_{\text{max}} = \frac{1}{\sigma_i} \sqrt{\frac{\ln a}{a^2 - 1}} \approx \frac{0.55}{\sigma_i} \quad (a = 1.6) \tag{8}
\]

For describing the extent of the Gabor spectrum, we adopt a typical estimation from statistics for Gaussian distributions:

We calculate the amplitude maximum \( (\bar{u}, \bar{v}) \) in the frequency domain and take the standard deviations \( \sigma_u \) and \( \sigma_v \) as a measure for the expansion of the spectrum. In the following, we will separately derive expressions for the \( u \)- and \( v \)-direction. As the Gabor function is a single Gaussian in the \( v \)-direction, we can simply read off its mean and standard deviation:

\[
\bar{v} = 0, \quad \sigma_v = \frac{\omega_0}{kd} \tag{9}
\]

Deriving the maximum in the \( u \)-direction is a bit more complicated. The part of the spectrum dependent on \( u \) is given by:

\[
G_1(u) = e^{-\frac{u^2}{\omega_0^2}} - e^{-\frac{u^2}{\sigma_u^2}} - e^{-\frac{u^2}{\omega_0^2}} + e^{-\frac{u^2}{\sigma_u^2}} \tag{10}
\]

The derivation of this function has to be set to zero, yielding a nonlinear equation that has to be solved:

\[
\frac{1}{z} + e^{-k^2 z} - 1 = 0, \quad z = \frac{\bar{u}}{\omega_0} \rightarrow z \approx 1 \text{ for } k = 2.5 \tag{11}
\]

The standard deviation is simply taken from the two Gaussian functions in \( G_1(u) \), giving as result for the \( u \)-direction:

\[
\bar{u} = \omega_0, \quad \sigma_u = \frac{\omega_0}{k} \tag{12}
\]
We quantize the amplitudes of the coefficients in order to avoid numerical side effects. This problem could be solved by introducing an additional minimization condition, using a sum of squares:

\[ \sum_{(j)} a^2_j \rightarrow \text{min} . \] (18)

The LMSE-solution of this nonlinear system of equations gives as good results as without minimization, but with far better distribution of coefficients, as Figure 5 shows. Such an additional minimization equation should therefore be included into optimization problems like this one.

### IV. QUANTIZATION

We will now investigate the effects of quantized coefficients on the results of the composition. Those depend on how strong deviations from the optimal solution worsen the filter masks. We quantize the amplitudes of the coefficients in \( K \) equal intervals relative to a maximum value, so the allowed values are:

\[ a_{\pm k} = \pm \frac{k}{K-1} a_{\text{max}} , \quad k = 0, 1, \ldots, K-1 . \] (19)

For optimal results, we use a successive approximation procedure. First, the full overdetermined system of equations as in equation (6) is solved, including minimization of coefficients (equation (18)), since a narrow range of values is necessary for good quantization. Then \( a_{\text{max}} \) is determined slightly smaller than the maximal occurring amplitude \( (a_{\text{max}} = 0.95 a_{\text{max}}) \). Thereafter, all coefficients that are in a range \( \epsilon \) around an allowed value are fixed to that value, and the system of equations is solved for the remaining non-fixed \( a_j \). This procedure is continued, increasing the range \( \epsilon \) subsequently, until all coefficients are fixed to an allowed value.

Figure 6 A-C shows results for the standard case (Figure 3). The masks look slightly different, with some jitter occurring at the outer regions, which, however, does not affect the filtering results. So, even when only switching on and off DoG masks without any weighting factor (except the sign) as shown in C, the results become not significantly worse. This is an indication for the robustness of the synthesis process.

To test the limits of the method, we repeat the quantization
are only 16 coefficients unequal to zero in the last case (see reasonable approximation of a Gabor filter. Notice that there as can be seen from figure 6 D,E, but could still be used as a results become significantly worse for strong quantization, for an arrangement with fewer DoG masks (figure 4 E). Here, the results become significantly worse for strong quantization, as there have to be solved several overdetermined, nonlinear systems of equations. But this can be done once, and the combination of DoG masks is then stored for subsequent filterings.

For showing the advantage of the approach described above, let us compare the computational cost with a standard Gabor filter bank. In the standard approach, a discrete convolution is calculated for each desired Gabor filter. Assuming a number of $N_G$ Gabor filters, an (output) image size of $X \cdot Y$ pixels and a size of the filter kernel of $U \cdot V$ pixels, filtering an image would require approximately $N_G \cdot X \cdot Y \cdot U \cdot V$ multiplications and the same number of additions. The approach with composed Gabor filters requires to calculate discrete convolutions for each of the $N_R$ retinal DoG-layers with approximately $N_R \cdot X \cdot Y \cdot U \cdot V$ multiplications and additions. One could argue that the filter kernel size varies from layer to layer, but this is compensated by the coarser sampling of DoG filters with greater spatial extent. The composition of the Gabor filter approximation requires $M_G$ additions for each position, which leads to $N_G \cdot X \cdot Y \cdot M_G$ additions for the whole sampling process. The difference between standard Gabor filter bank and composed-Gabor filter bank is then

$$\Delta C = N_G \cdot X \cdot Y \cdot \left[ U \cdot V \cdot \left( 1 - \frac{N_R}{N_G} \right) - M_G \right].$$

For a big filter bank, the approximation $N_R \ll N_G$ holds, so the term in brackets simplifies to $[U \cdot V - M_G]$. For practical filter sizes, $U \cdot V$ is much bigger than $M_G$. In the figures presented in this paper, $U = V = 16$ was used, which leads to a reduction of computational cost of factor 5 for the first quantisation (with ca. 50 coefficients unequal to zero) and of factor 16 for the arrangement with fewer DoG masks. Note that we combined multiplication and addition to one operation and thus did not take the reduced computational costs due to the missing multiplications in case of the composed-Gabor filter bank into account.

V. CONCLUSION

In this paper, a model for the retinal filter kernels has been developed inspired by experimental results and image processing methods. In a biology-inspired approach, Gabor filters were composed from the retinal filters using LMSE-algorithms. Therefore, parameters of the DoG filters were fixed
relative to the Gabor masks and methods with separate dimensional composition were proven to be insufficient to the problem. The composed masks are very similar to the original Gabor filters and it was shown that they are relatively insensible to parameter variations. To optimize the amplitudes of the coefficients, an additional minimization equation was used, which dramatically improved the distribution of amplitudes. Finally, the effects of quantization of coefficients were investigated with no significant worsening of results. This leads us to the conclusion, that an "all-or-nothing" approach, despite its coarse quantization can yield accurate Gabor filter masks. Inspired by the results, we propose an efficient implementation of a Gabor filter bank, which could enable the use of a high number of filters with relatively low computational expense.

Composition of Gabor filters could in principle be done with localized filters other than DoG filters, which would be maybe as efficient as with DoG filters or yield even better results. In this paper, we have restricted ourselves to a biological framework, for which filters similar to DoG filters are necessary to reproduce retinal behavior. Testing our approach with other filter kernels would be very interesting in that it helps assessing the results of composition with DoG filters. Furthermore, the composition of a large number of filters from a few basis filters could be a useful strategy for the efficient implementation of filter banks in general.

REFERENCES