Classification of color textures by gabor filtering

Christoph Palm* and Thomas M. Lehmann
Institute of Medical Informatics
Aachen University of Technology (RWTH)
Aachen, Germany

Abstract. A novel approach to Gabor filtering of color textures is introduced. It is based on the complex chromatic Fourier transform. Complex colors are derived from the HSL color space representing intensity independent color textures. Additionally, a novel Gabor texture feature for the grayscale as well as the color domain is proposed. It relies on local phase changes characterizing the homogeneity of a texture in the spatial frequency domain. Several classification experiments on two image databases are performed to study the texture features according to different color spaces and Gabor filter bank variants. The color features show significantly better results than the grayscale features. Although they are completely intensity independent, the features on basis of the complex color space show satisfying results. The RGB based features, where inherently color and intensity work together, performed best. Especially the local phase change measure supplement the known amplitude measure appropriately.

Key words: Gabor filter, color texture, classification, color Fourier transform

1. Introduction

Texture is an important aspect of image analysis for classification and segmentation as well as image generation, computer vision, graphics and image processing. By now, a large number of algorithms for texture feature extraction have been developed. These algorithms are useful especially for surface inspection in industry, [21, 32], as well as tissue characterization in medicine [29].

Although textures occuring in real world are colored, image processing research has focussed on the field of grayscale textures. Reasons for that were high costs of color cameras, high computational costs of color image processing, and the high complexity of color textures. Even now, many researchers negate a close relationship between color and texture, e.g. [18, 24], defining texture as intensity pattern, with color information restricted to histograms. Therefore, both phonomena, color and texture, are tackled independently. Representative for this group of the image processing community, Dubisson-Jolly et al., stated the thesis: "(...) we believe that there is no need to compute color texture features".

*Corresponding author. Christoph Palm is now with the Aixplain AG, Monheimsallee 22, D-52062 Aachen, Germany, Email: Christoph.Palm@web.de

In this paper, we argue with this statement proposing a detailed study of the relationship between color and texture. We concentrate on two related questions:

- Does color enhance grayscale texture features?
- Are there color textures, which are intensity independent?

For this, a large number of color texture images is needed first. We focus on natural color textures. To enable generalization of the conclusions, two large databases are used in our experiments. Both databases are available to the public to enable comparability of the results. Second, a number of grayscale texture features has to be extended to the color domain.

To get an answer to the first question, the corresponding features have to integrate color and intensity. For instance, the RGB color space yields an integrated approach. Each color band is related to the intensity of the incoming light. Nevertheless, the color aspect is integrated by the restriction to a range of wavelength.

The second question needs the separation of color and intensity. The HSL color space splits the lightness from the color features: hue and saturation. Therefore, it is suitable for our purposes.

Grayscale texture features are subdivided into statistical and signal threoretic algorithms. The classes are selected according to an underlying stochastic process or a specific Fourier pattern, respectively. Prominent representatives for statistical texture feature extraction methods are Gauss-Markov random fields [6] and Co-occurrence matrices [3], whereas Gabor and wavelet representations [9, 10] are both based on a windowed Fourier transform. Just a few steps have been made to describe and recognize colored textures. They cover the following methods: Co-occurrence matrices [29], auto-covariance matrices, [16], wavelets, [23] and Gabor filters, [19]. Migration of methods from grayscale to color images is difficult because of the transition from scalars to vectors which involves the loss of order-relation.

Signal theoretic approaches are claimed to outperform statistical methods, [15]. Especially Gabor filters show encouraging results. Hence, we study the behaviour of the Gabor transform on color images. The first approach by Jain, [19] with amplitude energy and correlation features is extended threefold: Firstly, we present novel phase energy Gabor features, which to our best knowledge are not only new in the color but in the grayscale domain as well. Secondly, beside the RGB color space, the HSL color space is used for color texture analysis, providing the separation of color and texture. The complex interpretation of the HSL color space has favorable characteristics especially for the color Fourier transform . Thirdly, we study not only the commonly used Gabor wavelet filter bank, but also the logpolar filter bank, which is claimed to perform better by Smeraldi et al. [31]. All these feature variants are carefully analysed performing a large number of classification experiments. The results are not only encouraging, but provide deeper insight into the structure of colored textures.

The paper is organized as follows. We shortly repeat the formula of grayscale Gabor filters with special emphasis on the two different filter banks in sec. 2. Because it is also relevant for grayscale Gabor filters, we introduce a new feature called phase energy. In sec 3, variants of the color Fourier transform are explained. Within that section, a complex color space is introduced. Two of the color Fourier transforms are used in sec. 4 to define color Gabor filters and the features derived from the Gabor coefficients. In sec. 5, experiments, and in sec. 6 their results, are presented, followed by a discussion in sec. 7.

2. Grayscale Gabor filters

2.1. 2D–Gabor filters

Most commonly, the Gabor transform is introduced as a windowed Fourier transform. A discrete Fourier transform (DFT) $F(u,v)$ decomposes an image $f(m,n)$ with the spatial coordinates $m \in \{0,\ldots,M-1\}$ and $n \in \{0,\ldots,N-1\}$ into their spatial frequencies $u \in \{-\frac{M}{2},\ldots,\frac{M}{2}\}$ and $v \in \{-\frac{N}{2},\ldots,\frac{N}{2}\}$. They represent the vertical and the horizontal decomposition, [4]:

$$F(u,v) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left(f(m,n)\cdot\exp\Big(-\frac{2\pi imu}{M}\Big)\cdot\exp\Big(-\frac{2\pi inv}{N}\Big)\right). \tag{1}$$

In general, the DFT results in complex Fourier coefficients, i.e. $F(u,v) \in \mathbb{C}$. A typical property of the DFT is the Hermitean symmetry, if $f(m,n)$ is real valued

$$F(-u,-v) = F^{\star}(u,v), \tag{2}$$

where $F^{\star}$ denotes the conjugate complex value of F. An undesired property of DFT is the lack of location dependency within the spatial frequency domain. In other words, high amplitudes of Fourier coefficients just denote frequent occurrence of this frequency in the image, but give no hint, where this frequency occurres. This is especially problematic for texture characterization, because texture is often inhomogeneously composed of harmonics with discrete frequencies. Local differences of non-stationary textures may be rather informative for modeling the texture and discerning it from others. Hence, a kind of local DFT is required.

Spatial locality is achieved by applying a gaussian window $w(m_\alpha - m_0, n_\alpha - n_0)$ to the image at a distinct position (m_0,n_0), and rotating it mathematically positivly by the angle α before performing DFT. The coordinates (m_α,n_α) result from the vector $(m,n)^T$ by

$$\begin{pmatrix} m_\alpha \\ n_\alpha \end{pmatrix} = \mathbf{A}_\alpha \begin{pmatrix} m \\ n \end{pmatrix} \qquad \text{with} \qquad \mathbf{A}_\alpha = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} \tag{3}$$

Applying a gaussian window to the windowed Fourier transform is called a Gabor transform ${}^{\mathcal{G}}F_{u_0,v_0}(m_0,n_0)$ [1]:

$${}^{\mathcal{G}}F_{u_0,v_0}(m_0,n_0) = \frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f(m,n)\cdot w(m_\alpha - m_0, n_\alpha - n_0)\cdot\exp\left(-2\pi i\Big(\frac{mu_0}{M}+\frac{nv_0}{N}\Big)\right). \tag{4}$$

Using the symmetry of the gaussian window $w(m,n)$

$$w(m,n) := \frac{1}{2\pi\sigma_m\sigma_n}\cdot\exp\Big(-\frac{1}{2}\frac{m^2}{\sigma_m^2}\Big)\cdot\exp\Big(-\frac{1}{2}\frac{n^2}{\sigma_n^2}\Big), \tag{5}$$

we are able to transform (4) to a convolution equation

$${}^{\mathcal{G}}F_{u_0,v_0}(m_0,n_0) = f(m_0,n_0)\star g_{u_0,v_0,\alpha}(m_0,n_0). \tag{6}$$

The convolution mask $g_{u_0,v_0,\alpha}(m,n)$ is called Gabor filter, which is tuned to the spatial frequency $(u_0,v_0)^T$:

$$g_{u_0,v_0,\alpha}(m,n) = \frac{1}{2\pi\sigma_m\sigma_n}\cdot\exp\left(-\frac{1}{2}\Big(\frac{m_\alpha^2}{\sigma_m^2}+\frac{n_\alpha^2}{\sigma_n^2}\Big)\right)\cdot\exp\left(2\pi i\Big(\frac{mu_0}{M}+\frac{nv_0}{N}\Big)\right). \tag{7}$$

It exhibits the functional form of a complex sinusoidal modulated window function. Fig. 1 visualizes the real part and the imaginary part on the left and right, respectively.

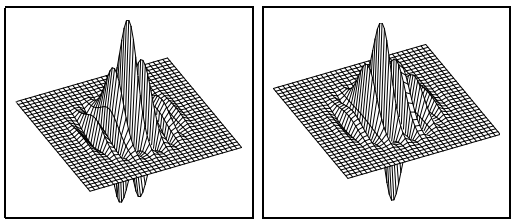

Fig. 1. Gabor filter in the spatial domain with its real part (left) and imaginary part (right).

Following the notation of Bovic et al., [8], we couple the orientation α of the gaussian window to the direction of the complex sinusoidal function, which is determined by (u_0,v_0). Additionally, the variance values σ_m and σ_n are transformed into a direction-independent variance σ and a scaling ratio τ

$$\sigma = \sigma_n \qquad \text{and} \qquad \tau = \frac{\sigma_m}{\sigma_n}. \tag{8}$$

Then, according to (7) and (8) the Gabor filter is given by

$$g_{u_0,v_0}(m,n) = \frac{1}{2\pi\sigma^2\tau} \cdot \exp\left(-\frac{1}{2\sigma^2}\left(\frac{m_\alpha^2}{\tau^2} + n_\alpha^2\right)\right) \cdot \exp\left(2\pi i\left(\frac{mu_0}{M} + \frac{nv_0}{N}\right)\right). \quad (9)$$

The transfer function $G_{u_0,v_0}(u,v)$ of a Gabor filter in the spatial frequency domain is used for implementation of this filter:

$$G_{u_0,v_0}(u,v) = \exp\left(-2\pi^2\sigma^2\left[\left(u_\alpha - (u_0)_\alpha\right)^2\tau^2 + \left(v_\alpha - (v_0)_\alpha\right)^2\right]\right). \quad (10)$$

The rotated spatial frequency coordinates (u_α, v_α) and $((u_0)_\alpha, (v_0)_\alpha)$ are transformed in analogy to (3). The functional form of $G_{u_0,v_0}(u,v)$ is equivalent to a discrete gaussian function with reciprocal variance. The orientation of this Gabor transfer function corresponds to that of the Gabor filter. The frequency of the sinusoid in (7) transforms into a spatial translation in the spatial frequency domain. By the convolution theorem, a convolution in the spatial domain reduces to a multiplication in the frequency domain. For this, both the image and the filter have to be Fourier-transformed

$${}^{\mathcal{G}}F_{u_0,v_0}(m,n) = \mathcal{F}^{-1}\Big\{F(u,v)\cdot G_{u_0,v_0}(u,v)\Big\}. \quad (11)$$

$\mathcal{F}^{-1}$ denotes the inverse DFT. Note that the real part and the imaginary part of the Gabor filter, ${}^{\mathcal{I}}g_{u_0,v_0}$ and ${}^{\mathcal{R}}g_{u_0,v_0}$, respectively, form approximately a quadrature filter pair, [8]. One advantage of a quadrature filter pair is the possiblity to derive the local amplitude and the local phase, ${}^{\mathcal{G}}A_{u_0,v_0}$ and ${}^{\mathcal{G}}\Phi_{u_0,v_0}$, respectively, from the Gabor coefficients at each pixel position (m,n):

$${}^{\mathcal{G}}A_{u_0,v_0} = \sqrt{\mathcal{R}\big({}^{\mathcal{G}}F_{u_0,v_0}(m,n)\big)^2 + \mathcal{I}\big({}^{\mathcal{G}}F_{u_0,v_0}(m,n)\big)^2} \quad (12)$$

$${}^{\mathcal{G}}\Phi_{u_o,v_0} = \arctan\left(\frac{-\mathcal{I}\big({}^{\mathcal{G}}F_{u_0,v_0}(m,n)\big)}{\mathcal{R}\big({}^{\mathcal{G}}F_{u_0,v_0}(m,n)\big)}\right). \quad (13)$$

The real part is noted as $\mathcal{R}(\cdot)$ or ${}^{\mathcal{R}}\cdot$, the imaginary part as $\mathcal{I}(\cdot)$ or ${}^{\mathcal{I}}\cdot$. For the purpose of texture analysis, not just a single Gabor filter is used but a number of filters called the Gabor filter bank. The aim of such a filter bank is to cover the entire spatial frequency domain with small redundancy by the transfer functions of the Gabor filters. Because of (2), covering of one half of the Fourier domain is sufficient for real (grayscale) images.

2.2. Gabor filter banks

We focus on two different design variants of Gabor filter banks: the Gabor wavelets and the logpolar Gabor filter bank. The characteristic of a Gabor filter bank is determined by the appearance of the filter transfer functions and their position in the spatial frequency domain. Hence, the following four parameters are important: σ, τ, u_0, v_0. The dependency between these parameters is especially important. Obviously, the frequency of the modulated sinusoidal function in the spatial domain should be taken into account.

This results in a dependency of filter width, given by σ in the spatial frequency domain, which is inverse proportional to the width in the spatial domain, and location of the filter (u_0, v_0) in the spatial frequency domain. As a colloquial statement: The higher the frequency the broader the filter in the Fourier domain.

2.2.1. Gabor wavelets

Gabor wavelets satisfy this requirements in the filter bank design process using polar coordinates (ζ_0, φ_0) instead of the linear coordinates (u_0, v_0). Therefore, we can rewrite (10) to the Gabor filter transfer function in polar notation, [9]

$$G_{\zeta_0,\varphi_0}(u,v) = \exp\left(-2\pi^2\sigma^2\left[\left(u_\alpha - \zeta_0\right)^2\tau^2 + v_\alpha^2\right]\right). \tag{14}$$

The decomposition of the Fourier domain by polar Gabor filters takes place in octaves and radians for the ζ_0 and φ_0 direction, respectively. To define the Gabor wavelet filter bank, we have to define the smallest possible frequency ζ_{00}, and the number of filters on the unit circle N_φ. The following equations define the position of each filter $G_{\zeta_{0i},\varphi_{0i}}$:

$$\begin{aligned} \zeta_{00} &= \sqrt{2} \\ \zeta_{0i} &= 2^i \cdot \zeta_{00} \\ \varphi_{0i} &= i \cdot \frac{2\pi}{N_\varphi}. \end{aligned} \tag{15}$$

Then, the parameter σ and τ are determined by the functional form of the Gabor filter, the position (ζ_0, φ_0) and the requirement of a dense filter bank. We call a filter bank dense, if each elliptical Gabor filter transfer function touches its direct neighbored filter at their respective half-peak frequency. The formular for σ_W and τ_W result from some mathematical transformations. The index W refer to the Gabor wavelet filter bank:

$$\sigma_W = \frac{1}{\zeta_0 \pi \tan\left(\frac{\pi}{N_\varphi}\right)} \qquad \text{and} \qquad \tau_W = 3\tan\left(\frac{\pi}{N_\varphi}\right). \tag{16}$$

The term *Gabor wavelets* results from properties similar to those of wavelets, especially with respect to their translation and scaling properties. Additionally, the Morlet wavelet has a very similar functional form, [11]. Gabor wavelets are non-orthogonal and show redundancies. In contrast to image compression, where a minimum redundancy is necessary for optimum reconstruction properties, redundant filters for texture analysis show more robustness against noise and other distortions, [22].

2.2.2. Logpolar Gabor filters

Like Gabor wavelets, logpolar Gabor filters are based on polar coordinates. But instead of a frequency dependent definition of σ on a linear graduation of ζ, the dependency of σ on ζ is realized by a logarithmic scale. Therefore, the functional form of $G_{\zeta_0,\varphi_0}(u,v)$

changes again from (14). Instead of (u, v) the polar coordinates (ζ, φ) are used and the logarithm is applied on ζ:

$$G_{\zeta_0,\varphi_0}(\zeta,\varphi) = \exp\left(-2\pi^2\sigma^2\Big[\big(\ln(\zeta)-\ln(\zeta_0)\big)^2\tau^2 + \big(\ln(\zeta)\sin(\varphi-\varphi_0)\big)^2\Big]\right). \qquad (17)$$

The formula for defining the position of the filter in the Fourier domain equals (15). But the change of the functional form of the filter transfer function in (17) results in different σ_L and τ_L values. The index L refers to a logpolar Gabor filter bank:

$$\sigma_L = \frac{1}{\ln(\zeta_0)\pi\sin\left(\frac{\pi}{N_\varphi}\right)}\sqrt{\frac{\ln 2}{2}} \qquad \text{and} \qquad \tau_L = \frac{2\ln(\zeta_0)\sin\left(\frac{\pi}{N_\varphi}\right)}{\ln 2}\sqrt{\frac{\ln 2}{2}}. \qquad (18)$$

Obviously, the logpolar filter definition results just in another filter shape, not in another filter position. In contrast to Gabor wavelets, which are symmetric with respect to their principal axis, logpolar Gabor filters show a translation of the maximum from the center of gravity in direction of a lower frequency and flattening of the high frequency part. The polar notation of the spatial frequency coordinates results in curvature of the filter and, therefore, in a polar sampling of the Fourier domain. Whereas the isocurves of the Gabor wavelets intersect the neighboring curves slightly, in the logpolar definition they just touch each other. Fig. 2 shows the transfer functions of Gabor wavelets (left) and the transfer functions of logpolar Gabor filters (right). To visualize the shape of the filters the isocurves of the half-peak frequency are plotted. Whereas Gabor wavelets represent a kind of standard for the design of Gabor filter banks, logpolar Gabor filters have been introduced recently for the purpose of texture analysis [20, 31].

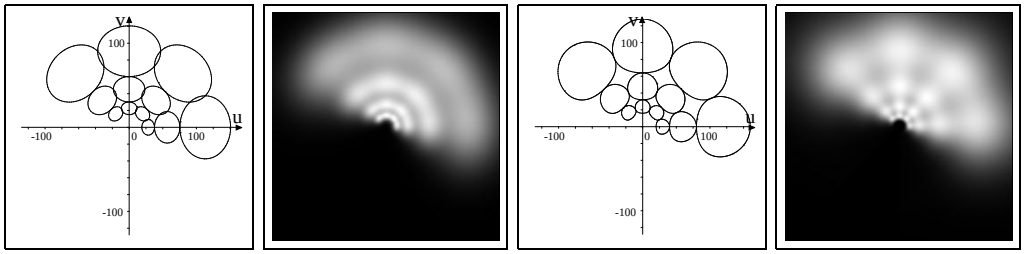

Fig. 2. Gabor filter bank (frequency domain). The shape variety is visualized by isocurves of the half-peak frequency and a gray-coding of the filter values for the Gabor wavelets the logpolar Gabor filters on the left and on the right, respectively.

2.3. Grayscale Gabor features

The Gabor decomposition of an image $f(m, n)$ results in a number of complex images ${}^{\mathcal{G}}F_{u_0,v_0}(m, n)$ of Gabor coefficients (6). The number of coefficient images depends on the

sampling rate of the Fourier domain by the transfer functions of the Gabor filters. Commonly, these Gabor coefficients are not used directly for texture classification because of the high dimensionality of the resulting feature space. For instance, an image with dimensions 256×256 and a filter bank with three and four filters in ζ and φ direction, respectively, results in twelve complex Gabor images and 786.432 complex features. Hence, textural features have to be derived from Gabor coefficients. For the domain of grayscale textures, an energy measure relying on local amplitudes is used by most authors. Beside this, we introduce a novel energy measure that is based on local phases.

2.3.1. Local amplitude energy

By (12), the local amplitude ${}^{\mathcal{G}}A_{\zeta_0,\varphi_0}$ of Gabor coefficients at position (m,n) specifies the intensity of an individual frequency in the area of (m,n). The energy of this amplitude indicates the intensity of that frequency for the entire texture:

$$\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0,\varphi_0}\right) := \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left({}^{\mathcal{G}}A_{\zeta_0,\varphi_0}(m,n)\right)^2 = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1}\left|{}^{\mathcal{G}}F_{\zeta_0,\varphi_0}(m,n)\right|^2. \tag{19}$$

Computation of the energy in the spatial domain allows to exclude some parts of the image, if it consists of more than one texture patch and can be segmented beforehand. To ensure rotation-invariant texture features, the amplitude energy values are averaged according to the φ-dependent coefficient images. This results in a mean and a variance energy amplitude value for a fixed ζ_0, $\mu\left(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0}\right)\right)$ and $\sigma\left(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0}\right)\right)$, respectively:

$$\mu\Big(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0}\right)\Big) = \frac{1}{N_\varphi}\sum_{\varphi_0}\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0,\varphi_0}\right) \tag{20}$$

$$\sigma\Big(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0}\right)\Big) = \frac{1}{N_\varphi - 1}\sum_{\varphi_0}\left(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0,\varphi_0}\right) - \mu\Big(\operatorname{eng}\left({}^{\mathcal{G}}A_{\zeta_0}\right)\Big)\right)^2 \tag{21}$$

.

2.3.2. Local phase energy

Local phases themselves are inappropriate for texture analysis. Beside their rotational invariance, the requirement of translational invariance of the features is self-evident. Phase measures cannot ensure that requirement. Nevertheless, the changing of phases is translation-invariant, and is used for texture segmentation [8]. The idea for segmentation is to detect the change of phases between two textures with constant phase values. In practice, such homogeneous textures with constant phase are rather seldom, especially for natural textures we are interested in. Natural textures are composed of several homogeneous texture patches. Therefore, we propose the use of local phase changes as a feature which characterize the homogeneity of a texture with respect to frequency features.

According to (13), a straight foreward detection of phase changes is not possible, because the arctangens function shows discontinuities at $\frac{\pi}{2}$ and $-\frac{\pi}{2}$. , and it is difficult

to distinguish these discontinuities from phase changes. Therefore, the phase change $\frac{\partial {}^{\mathcal{G}}\Phi_{\zeta_0,\varphi_0}}{\partial \mathbf{r}}$ is not detected using the gradient of the original phase function but using the combined gradient of the real and imaginary component of the complex Gabor coefficients [8]. The direction of the gradient equals the orientation φ_0 of the Gabor filter. The partial derivation of ${}^{\mathcal{G}}\Phi_{\zeta_0,\varphi_0}$ in the direction of $\mathbf{r} = (\cos\varphi_0, \sin\varphi_0)^T$ is computed as

$$\frac{\partial {}^{\mathcal{G}}\Phi_{\zeta_0,\varphi_0}}{\partial \mathbf{r}} = \frac{\frac{\partial \mathcal{R}({}^{\mathcal{G}}F_{\zeta_0,\varphi_0})}{\partial \mathbf{r}} \mathcal{I}\left({}^{\mathcal{G}}F_{\zeta_0,\varphi_0}\right) - \frac{\partial \mathcal{I}({}^{\mathcal{G}}F_{\zeta_0,\varphi_0})}{\partial \mathbf{r}} \mathcal{R}\left({}^{\mathcal{G}}F_{\zeta_0,\varphi_0}\right)}{\mathcal{R}\left({}^{\mathcal{G}}F_{\zeta_0,\varphi_0}\right)^2 + \mathcal{I}\left({}^{\mathcal{G}}F_{\zeta_0,\varphi_0}\right)^2}. \tag{22}$$

The partial derivations in (22) are implemented by simple difference filters in the discrete Gabor coefficient domain. The location of a phase change is detected with help of a local maximum of $\frac{\partial {}^{\mathcal{G}}\Phi_{\zeta_0,\varphi_0}}{\partial \mathbf{r}}$. We propose the energy of these local maxima as novel Gabor texture feature for grayscale texture classification, which indicates the frequency of phase changes within a unique texture. Similar to the energy of local amplitudes, the mean and variance of the phase energy measure $\mathrm{eng}({}^{\mathcal{G}}\Phi_{\zeta_0,\varphi_0})$ are computed to achieve rotational invariance.

3. Chromatic Fourier transforms

Since we use the Fourier domain for filter bank design as well as its implementation, the Fourier transform is the main element of the Gabor transform. Hence, extension of grayscale Gabor filters to colors requires a detailed study of variants of chromatic Fourier transforms. Three approaches are found in literature:

- dimension-extended chromatic Fourier transform, introduced by Nicolás et al., [26],
- vectorial chromatic Fourier transform, [27, 17],
- complex chromatic Fourier transform, [7, 13, 25].

3.1. Dimension-extended chromatic Fourier transform

DFT is known to be separable with respect to the m and n axis. Consequently, its extension to color as a third direction is self-evident as a first idea. Formally, a color image $\mathbf{f}(m,n)$ is defined as a vector of grayscale images $f_k(m,n)$ with $\mathbf{f}(m,n) = \big(f_1(m,n), \ldots, f_K(m,n)\big)$. Within this paper, we restrict ourselves to RGB images and transformations. Therefore, $K = 3$ is assumed. Hence, the dimension extended chromatic Fourier transform (DECFT) is

$$F(u,v,\bar{u}) = \frac{1}{K}\sum_{k=0}^{K-1}\left(\frac{1}{MN}\sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f_k(m,n)\cdot\exp\left(-2\pi i\Big(\frac{mu}{M}+\frac{nv}{N}\Big)\right)\right)\exp\left(-2\pi i\frac{k\bar{u}}{K}\right), \tag{23}$$

where $\bar{u} \in \{-1, 0, 1\}$ represents the frequency components in the k-th direction. Two problems wit DECFT arise. Firstly, interpretation of the sinusoids in the k-th direction is nearly impossible, if we assume just three sampling points for the three color bands. Secondly, the Fourier transform is defined for periodic signals only. Virtually, the signal has to be continued in each direction. Toward discontinuities, equalization of the border values is required, e.g. by windowing [5]. Window functions reduce the border values to zero. Since the signal consists of just three values in k-th direction, the window function leaves only the middle value non-zero. In case of RGB images, this middle value represents the green channel . Therefore, DECFT degrades to DFT on the green channel. For these reasons, DECFT is an inappropriate method for Fourier transforming of color images in general.

3.2. Vectorial chromatic Fourier transform

Two kinds of vectorial approaches have been introduced in the literature: the sequential chromatic Fourier transform and the quaternion chromatic Fourier transform.

3.2.1. Sequential chromatic Fourier transform

The sequential chromatic Fourier transform (SCFT) corresponds to a real-valued DFT on each of the K color bands. Therefore, SCFT results in K independent spectra. At the end, these spectra are combined into a K-dimensional complex valued coefficient vector:

$$F_k(m,n) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_k(m,n) \cdot \exp\Big(-\frac{2\pi imu}{M}\Big) \cdot \exp\Big(-\frac{2\pi inv}{N}\Big). \qquad (24)$$

The advantage of SCFT is the possibility to interpret the transformation result. Knowing which characterics of a signal is measured may help to understand the performance differences of different texture analysis approaches. For visualization purposes, fig. 3 shows sinusoidal functions, for which SCFT can be tuned. The resulting spectrum has just one frequency value, which is not equal to zero:

$$F_k(u,v) = \begin{cases} {}^{\mathcal{R}}F_k(u,v) + i \cdot {}^{\mathcal{I}}F_k(u,v) & \text{if } (u,v) = (2,0) \\ {}^{\mathcal{R}}F_k(u,v) - i \cdot {}^{\mathcal{I}}F_k(u,v) & \text{if } (u,v) = (0,2) \\ 0 & \text{else.} \end{cases} . \qquad (25)$$

Assuming the values of spatial frequencies to be zero for $k = G$ and $k = B$ and $({}^{\mathcal{R}}F_R(u,v), {}^{\mathcal{I}}F_R(u,v)) = (1000, 0)$ for $k = R$, a red intensity and saturation harmonic arises in the spatial domain. The corresponding red, green and blue harmonics are shown in fig. 3. The SCFT is most frequently used to transform color images into the spatial

frequency domain. It has been used for image compression [27] and contrast enhancement [12]. In this paper, we apply the SCFT for color texture analysis on the basis of the Gabor transform.

3.2.2. Quaternion Fourier transform

Another kind of vectorial color Fourier transform was introduced by SANGWINE as the quaternion Fourier transform (QFT) [14, 17]. A quaternion number q consists of one real and three imaginary parts, i_1, i_2 and i_3, which are independent dimensions of an orthonormal system:

$$q = q_1 + i_1 q_2 + i_2 q_3 + i_3 q_4 \qquad \text{with} \qquad i_1^2 = i_2^2 = i_3^2 = i_1 i_2 i_3 = -1. \tag{26}$$

A color image $\mathbf{f}(m,n)$ is interpreted as a quaternion image when the real component is set to zero and the three imaginary components are used to represent the color bands. On this basis, the QFT is defined by

$$Q(u,v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \exp\Big(-\frac{2\pi i_2 m u}{M}\Big) \cdot q(m,n) \cdot \exp\Big(-\frac{2\pi i_3 n v}{N}\Big). \tag{27}$$

When using quaternion values, the non-existing commutativity has to be taken into account. In addition, (27) does not support any interpretation of what is measured. Due to these problems, we do not use the QFT for our purpose of color texture classification. Nevertheless, efficient algorithms for implementation of QFT have been introduced recently [33]. Therefore, this transform might be analyzed for color texture classification in the near future. SANGWINE has already applied the method successful for autocorrelation of color images [30].

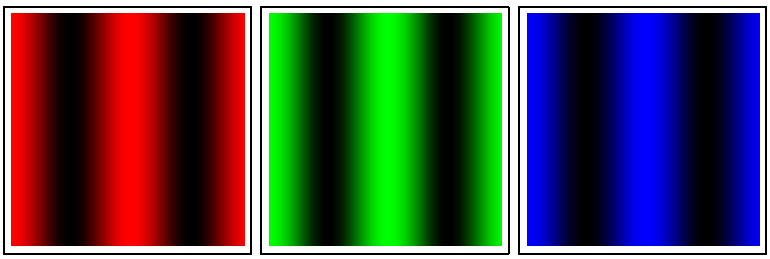

Fig. 3. The components of a spatial frequency domain vector resulting from SCFT are tuned to intensity and saturation harmonics, separated into the different color bands red (left), green (middle) and blue (right).

3.3. Complex chromatic Fourier transform

The complex chromatic Fourier transform (CCFT) is based on a complex representation of color images. Therefore, we introduce the complex color representation first.

3.3.1. Color in the complex representation

The complex representation of colors is not a new color space, but a new interpretation of the well-known HSL color space. The HSL color space is a non-linear transform of the RGB cube:

$$f_H(m,n) = \arctan\left(\frac{\sqrt{3}\big(f_G(m,n) - f_B(m,n)\big)}{2f_R(m,n) - f_G(m,n) - f_B(m,n)}\right), \tag{28}$$

$$f_S(m,n) = 1 - \frac{\min\big\{f_R(m,n), f_G(m,n), f_B(m,n)\big\}}{f_L(m,n)}, \tag{29}$$

$$f_L(m,n) = \frac{f_R(m,n) + f_G(m,n) + f_B(m,n)}{3}. \tag{30}$$

The chromatic components hue, saturation and brightness (intensity) correspond closely to the categories of human color perception. One disadvantage are labile values near the zero saturation, and a singularity at $f_S(m,n) = 0$. Therefore, the HSL space is only suitable for images with saturated colors. Another problem arises if the SCFT is applied on the hue channel. Hue is an angular coordinate that is stored as a scalar value. This implies high frequencies for smooth color changes between red and magenta. The complex representation introduced by Frey overcomes this problem, [7]:

$$\begin{aligned} z(m,n) &= f_S(m,n) \cdot \exp\big(i f_H(m,n)\big) \\ &= f_S(m,n) \cdot \Big(\cos\big(f_H(m,n)\big) + i \cdot \sin\big(f_H(m,n)\big)\Big). \end{aligned} \tag{31}$$

In this representation, the saturation is interpreted as the magnitude and the hue as phase of a complex value z. The intensity channel is not included. Hence, a complex color space integrates the intensity channel as an axis orthogonal to the complex plane. The complex color space has three main advantages. Firstly, it handles the problems of the HSL color space connected with low saturation by low weighting. Secondly, since the Fourier transform is defined for both real and complex values, the complex plane can be used as an input for a complex-valued DFT. This strategy results in a certain method, which is understood in detail, and which is suitable for color texture analysis. Thirdly, the intensity is separated from the color values – therefore, we are able to study pure color textures.

3.3.2. Fourier transform based on complex colors

The definition of CCFT is closely related to DFT. In (1), the real-valued image $f(m,n)$ is substituted by a complex image $z(m,n)$:

$$Z(m,n) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} z(m,n) \cdot \exp\left(-\frac{2\pi i m u}{M}\right) \cdot \exp\left(-\frac{2\pi i n v}{N}\right). \tag{32}$$

The complex input to CCFT results in a loss of symmetry with respect to the amplitude spectrum like in (2). For visualization purposes, the difference image $D(u,v)$ of an

amplitude spectrum and its version rotated by $\mathbf{A}_\pi$ (3) is computed:

$$D(u,v) = \Big|Z(u,v)\Big|^2 - \Big|Z\big((u,v)\cdot\mathbf{A}_\pi\big)\Big|^2. \tag{33}$$

The Fourier transform of a real-valued image results in a constant difference image $D(u,v)$. The difference image $D(u,v)$ of the complex representation of an color image shows noise with values significantly different from zero (fig. 4).

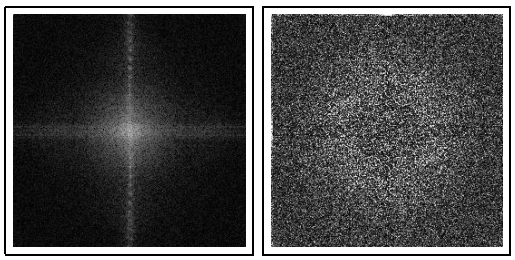

Fig. 4. Left: $|Z(u,v)|^2$ of the complex representation of a color image. Right: $D(u,v)$. The noise results from the loss of symmetry due to the complex input for the DFT.

A phase diagram helps to understand the harmonics to which the CCFT is tuned [25]. It visualizes the Fourier transform as a vector rotating around the center of a complex plane. By assuming one-dimensional images $f(m)$ and a one-dimensional Fourier transform for the positive and negative coordinates, $F(u)$ and $F(-u)$, respectively, different phase diagrams are generated. The rotation of the vector follows a mathematically positive or negative orientation for positive or negative coordinates, respectively. The length of the vector correlates with the amplitude of the frequency. The inverse Fourier transform results from vector addition of the rotation vectors corresponding to $F(u)$ and $F(-u)$. In case of two-dimensional images, we call the coordinates (u,v) positive if and only if

$$0 \le \arctan\left(\frac{v}{u}\right) < \pi \tag{34}$$

The idea of CCFT is clear, if the complex plane of the rotation vector is correlated with the complex color plane. Fig. 5 shows this complex plane and the location of the complex representation of red, green and blue, using the vectors z_R, z_G and z_B, respectively. The starting point of the rotation depends on the real and imaginary parts of the studied frequency component. For a real-valued component, the starting point coincides with red. For an imaginary-valued component, it is located near green. Assuming existence of both the components, the starting point is located in between. The orientation of rotation depends on positive and negative frequencies.

These conditions are simplified in fig. 6, where only one real-valued frequency is present. If this frequency is positive (negative) according to (34), a rainbow oscillation

is characterized, which touches the colors red-green-blue (red-blue-green) during rotation (fig. 6, left and middle). In case of a positive and a negative frequency component, the rainbow degrade to a straight line on the complex plane. Here, the line coincides with the real axis of the complex plane (fig. 6, right).

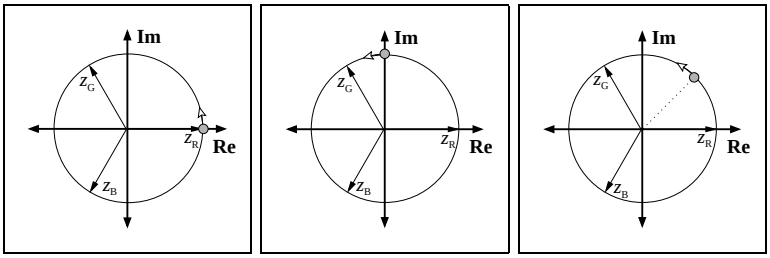

Fig. 5. A phase diagram of starting points (gray point) and direction of rotation (white arrow) in the complex plane: real (left), imaginary (middle) and complex frequency component with equal real and imaginary parts (right).

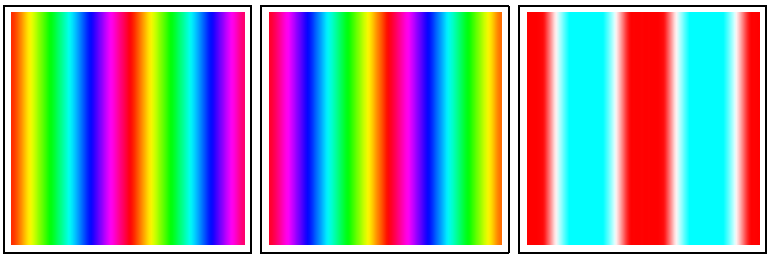

Fig. 6. Harmonics to which CCFT is tuned. All images show real frequency components, which results in a red starting point (the left side of each image). The rainbow harmonic for the positive frequency component (left) follows the red - green - blue - red path, whereas for the negative frequency component (middle) the red-blue-green-red path is given. In presence of positive as well as negative frequency components (right), the path is generated by vector addition. In case of a red starting point the real axis is followed.

4. Color Gabor filters

The approaches to color Gabor transform are based on the concepts of the color Fourier transform. The SCFT processes the color bands in the RGB color space separately. Although it was not explicitly mentioned, JAIN has published the only approach to color Gabor filtering using the SCFT, [19]. Therefore, the features introduced by JAIN are

included in our experimental framework to provide a baseline for the methods derived in this paper. The SCFT-related Gabor filters closely combine the intensity and color aspects, whereas the novel CCFT-based filters enable separate processing of the intensity texture and the color texture.

4.1. Gabor filter in RGB color space

More formally, the SCFT approach separates the RGB color image $\mathbf{f}(m,n)$ into three color bands $f_k(m,n)$ with $k \in \{R,G,B\}$ and applies the Fourier transform separately for each f_k. Extending this to the Gabor transform, for which the image as well as the filter is transformed into the spatial frequency domain, the filter remains the same as for the grayscale approach, but the image is transformed via SCFT. The resulting vector of Fourier coefficient images is multiplied by a vector of the Gabor filter transfer function, where each vector component has the same filter values. According to vector multiplication, the filter vector is transposed. The resulting image vector is retransformed via SCFT. For the sake of simplicity, we give the Gabor transform just for a single component k:

$$ {}^{\mathcal{G}}F^{k}_{\zeta_0,\varphi_0} = \mathcal{F}^{-1}\Big\{ F_k(u,v) \cdot G_{\zeta_0,\varphi_0} \Big\}. \tag{35} $$

Two kinds of features are extracted from these vector-valued Gabor coefficient images. The "within-color" features are closely related to the grayscale features, because the vector elements are again separately processed. The single Gabor image is handled as if it were a grayscale Gabor image. Therefore, in this paper we propose the same features as introduced in the last section: amplitude energy measures and a phase energy measure. Whereas the first was already proposed in [19], the second is new in the grayscale domain and, hence, novel for the color domain. The equations for the band specific local amplitude ${}^{\mathcal{G}}A^{k}_{\zeta_0,\varphi_0}$ and the band specific local phase ${}^{\mathcal{G}}\Phi^{k}_{\zeta_0,\varphi_0}$ are analogous to (12) and (13), respectively.

The main disadvantage of the within-color features is their restriction to one color band regardless the correlation between the color bands. This is expressed by correlation features, which are adopted from the opponent process theory of human color vision [19]. The correlation features can be defined by the energy of the difference of two normalized Gabor filtered color bands k_1 and k_2 with a fixed orientation φ_0 and fixed ($\zeta_0' = \zeta_0$) or neighboring ($\zeta_0' = 2\zeta_0$) center frequencies:

$$ \sum_{m,n} \left(\frac{{}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}(m,n)}{\operatorname{eng}\big({}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}\big)} - \frac{{}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}(m,n)}{\operatorname{eng}\big({}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}\big)} \right) = 2 - 2\sum_{m,n} \frac{{}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}(m,n) \cdot {}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}(m,n)}{\operatorname{eng}\big({}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}\big) \cdot \operatorname{eng}\big({}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}\big)}. \tag{36} $$

By removing the redundant information, we obtain the correlation Gabor features $\operatorname{corr}\Big({}^{\mathcal{G}}A^{k_1,k_2}_{\zeta_0,\zeta_0',\varphi_0}\Big)$ as the normalized cross-correlation between ${}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}(m,n)$ and ${}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}$, $k_1 \neq k_2$

$$\operatorname{corr}\left({}^{\mathcal{G}}A^{k_1,k_2}_{\zeta_0,\zeta_0',\varphi_0}\right) = \sum_{m,n} \frac{{}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}(m,n) \cdot {}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}(m,n)}{\operatorname{eng}\left({}^{\mathcal{G}}A^{k_1}_{\zeta_0,\varphi_0}\right) \cdot \operatorname{eng}\left({}^{\mathcal{G}}A^{k_2}_{\zeta_0',\varphi_0}\right)} \tag{37}$$

Assuming $k_1 \neq k_2$ and a Gabor filter bank consisting of N_ζ scales and N_φ orientations, the number of $\frac{1}{2}(K^2-K)\cdot N_\zeta \cdot N_\varphi$ and $(K^2-K)\cdot(N_\zeta-1)\cdot N_\varphi$ features are extracted for the $\zeta_0' = \zeta_0$ and the $\zeta_0' = 2\zeta_0$ case, respectively. Due to the mean and variance computation to achieve rotational invariance (see (20) and (21), respectively), the number of features reduces to $(K^2-K)\cdot N_\zeta$ and $(K^2-K)\cdot 2\cdot(N_\zeta-1)$, respectively.

4.2. Gabor filter in complex color space

The correlation features handle the connection between color bands as a postprocessing step after the SCFT-based Gabor filtering. The following approach to complex color features combines the color bands before the filtering is applied, using the complex color space and the CCFT approach. The main advantage of this filter is utilization of the HSL color space and, hence, separation of color and the intensity.

An obvious difference between the SCFT-based and the CCFT-based Gabor transforms is the loss of Hermitean symmetry, which results in the necessity of a complete coverage of the Fourier domain by the filter bank. The filters G_{ζ_0,φ_0} corresponding to the real Gabor filter bank cover the positive central frequencies (u_0, v_0) following (34). They are located in the upper half of the Fourier domain (fig. 2). Now, they are supplemented by the $G_{\zeta_0,\varphi_0+\pi}$'s, which sample the part of the Fourier domain with negative spatial frequencies. The main difference between G_{ζ_0,φ_0} and $G_{\zeta_0,\varphi_0+\pi}$ is the kind of rainbow oscillations they are tuned to. Whereas the order of colors for G_{ζ_0,φ_0} is determined by a positive rotation on the complex color plane (red-green-blue-red), a negative rotation for G_{ζ_0,φ_0} results in a reverse order (red-blue-green-red). Hence, both filters measure different kinds of color oscillations. Consequently, we differenciate between the Gabor coefficients ${}^{\mathcal{G}}Z^1_{\zeta_0,\varphi_0}$ and ${}^{\mathcal{G}}Z^2_{\zeta_0,\varphi_0}$ corresponding to G_{ζ_0,φ_0} and $G_{\zeta_0,\varphi_0+\pi}$, respectively:

$${}^{\mathcal{G}}Z^1_{\zeta_0,\varphi_0} = \mathcal{F}^{-1}_{\mathrm{CCFT}}\Big\{Z(u,v)\cdot G_{\zeta_0,\varphi_0}(u,v)\Big\} \tag{38}$$

$${}^{\mathcal{G}}Z^2_{\zeta_0,\varphi_0} = \mathcal{F}^{-1}_{\mathrm{CCFT}}\Big\{Z(u,v)\cdot G_{\zeta_0,\varphi_0+\pi}(u,v)\Big\}. \tag{39}$$

where $\mathcal{F}^{-1}_{\mathrm{CCFT}}$ denotes the inverse CCFT.

To understand the differences between these variants of rainbow oscillations, the local amplitudes ${}^{\mathcal{G}}A^{z_2}_{\zeta_0,\varphi_0}$ and ${}^{\mathcal{G}}A^{z_2}_{\zeta_0,\varphi_0+\pi}$ corresponding to ${}^{\mathcal{G}}Z^1_{\zeta_0,\varphi_0}$ and ${}^{\mathcal{G}}Z^2_{\zeta_0,\varphi_0}$, respectively, are studied in more detail. To compare color results with grayscale results, the color image (Fig. 7, left) is transformed into a grayscale image by weighting the means of the color bands equally using (30). The grayscale images are Gabor transformed in two different scales by $G_{6\sqrt{2},0}$ and $G_{5\sqrt{2},0}$, Figure 7 (middle) and Figure 7 (right), respectively. The

local amplitudes connected with these filters show a high sensivity to intensity changes at the wall splices, resulting in regular peaks in both amplitude images.

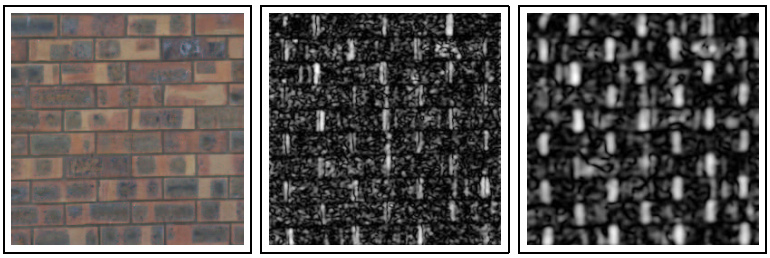

Fig. 7. Left: A color image from VisTex database (ftp://whitechappel.media.mit.edu/pub/VisTex). Middle and right: Local Gabor amplitudes ${}^{\mathcal{G}}A_{6\sqrt{2},0}$ and ${}^{\mathcal{G}}A_{5\sqrt{2},0}$, respectively, obtained from the corresponding grayscale image. Both coefficient images show the focus on intensity changes.

The corresponding amplitude images ${}^{\mathcal{G}}A^{z_1}_{5\sqrt{2},0}$ and ${}^{\mathcal{G}}A^{z_2}_{5\sqrt{2},0}$ (fig. 8, left and middle) of the complex color plane of the image in fig. 7, left, differ noticeably. Obviously, the color changes inside the bricks are modelled rather than the intensity changes. Additionally, they are different among each other. To visualize this effect, the difference image of both amplitude versions is shown in fig. 8, right. The light values are identified with strong differences. The white arrow marks an exemplarily a location of a strong difference. At this location, the colors red-yellow-blue can be noticed in the original image. This color order results in high amplitudes using the filter $G_{5\sqrt{2},0}$, whereas the amplitudes obtained with the reverse order filter $G_{5\sqrt{2},0+\pi}$ are low.

Obviously, the color texture features, which are based on the complex color space, are suitable to model different texture characteristics and, therefore, supplement the RGB-based and the grayscale-based Gabor features.

5. Experimental evaluation

The novel color texture features are not only studied qualitatively but quantitatively using classification experiments. Two image databases, VisTex and BarkTex, are explored. The VisTex images are available to the public at the Massachusetts Institute of Technology (ftp://whitechapel.media.mit.edu). All images show a frontally recorded natural texture. We took the same set of 30 images of natural scenes as in [23] (fig. 9). Each image of size 512×512 was subdivided into 64 disjunct 64×64 subimages to generate 1920 images of 30 classes. Therefore, the within-class variance is related to the within-image variance of the texture. In the following, this dataset is noted as $\underline{DS}_1$.

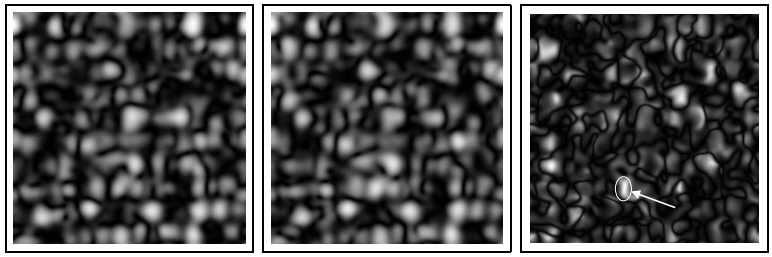

Fig. 8. Left and Middle: Local Gabor amplitudes ${}^{\mathcal{G}}A^{z_1}_{5\sqrt{2},0}$ and ${}^{\mathcal{G}}A^{z_2}_{5\sqrt{2},0}$ according to the complex plane of the color image shown in fig. 7, left. The features focus on image characteristics different than the corresponding grayscale Gabor amplitudes (fig. 7, right). Additionally, they are different among each other. This is clarified by the difference image (right). The white arrow marks exemplary a location of high difference.

The BarkTex data has been prepared by Lakmann [16] at the University of Koblenz-Landau, Germany. It is publicly available at ftp://ftphost.uni-koblenz.de/outgoing/vision/Lakmann/BarkTex. Each image of size 384×256 displays the bark of a certain tree. Six different kinds of trees have been captured, with 68 images per class (fig. 10). Since the image apertures are not restricted to the bark texture and also show background structures, they have been previously masked, with the middle part of 300×200 pixels being selected. The classification task of BarkTex can be stated to be more serious than VisTex, because on the one hand the colors of the barks are very similar and on the other hand, the within-class variance is high according to different trees related to each image. This dataset is denoted as $\underline{DS}_2$.

The following classification experiments make use of the k-Nearest-Neighbor (kNN) classifier (k=5) assuming the Mahalanobis distance measure $d(\mathbf{m}, \widehat{\mathbf{m}})$ between two feature vectors $\mathbf{m} = (m_1, \ldots, m_e, \ldots, m_E)$, $e, E \in \mathbb{N}$, and $\widehat{\mathbf{m}}$ analog:

$$d(\mathbf{m}, \widehat{\mathbf{m}}) = (\mathbf{m} - \widehat{\mathbf{m}})^T \mathbf{\Sigma} (\mathbf{m} - \widehat{\mathbf{m}}), \tag{40}$$

with $\mathbf{\Sigma}$ denoting the diagonal covariance matrix. Therefore, normalization of the dimensions of the feature space spanned by $\mathbf{m}$ is performed within the distance measurement. To ensure both a strict deviation of test and training data as well as a reliable volume of training data, the leaving-one-out procedure is applied for the experiments. That means, one feature vector is split away from the entire set and is classified according to other data. This procedure is iterated for each of the feature vectors. Then, the correctness κ related to a feature space is measured, comparing the original class assignment and the assignment of the classifier for each vector averaged over all vectors.

Eleven feature spaces have been studied for evaluation of different features and their combinations for a fixed filter bank and a fixed database. By varying the Gabor wavelet filter bank (GW) and the logpolar Gabor filter bank (LG) as well as the databases $\underline{DS}_1$ and $\underline{DS}_2$, the total 44 experiments have been performed. For the design of the filter

banks, the number of orientations N_φ has to be determined. Together with the fixed smallest center frequency ζ_{00} and the largest image dimension M the number of possible scales $\widehat{N}_\zeta = \log_2(M) - 1$ is derived. Since for texture analysis high frequencies are the most important ones, we take only the highest N_ζ scales for feature extraction. Within the experimental framework producing the results listed in tab. 1, we have chosen $N_\varphi = 4$ orientations as well as $N_\zeta = 3$ and N_ζ=4 scales for $\underline{DS}_1$ and $\underline{DS}_2$, respectively.

Fig. 9. Images of the VisTex database (left-right, up-down): Bark0, Bark4, Bark6, Bark8, Bark9, Brick1, Brick4, Brick5, Fabric0, Fabric4, Fabric7, Fabric9, Fabric11, Fabric13, Fabric16, Fabric17, Fabric18, Food0, Food2, Food5, Food8, Grass1, Sand0, Stone4, Tile1, Tile3, Tile7, Water6, Wood1, Wood2.

Fig. 10. Images of the BarkTex database, one example for each class (left-right): birch, beech, spruce, pine, oak, robinia

6. Results

Rows 1–3 of tab. 1 are related to grayscale decomposition. For this, the images were transformed to grayscale images using (30) and decomposed by the Gabor wavelet filter bank and the logpolar Gabor filter bank described in detail in (2.2.1) and (2.2.2), respectively. The energy of the local amplitude based on (20) and (21), the energy of the phase changes based on the corresponding mean and variance measures related to (22), and the cross-product of both feature spaces $\mathbf{m}\big(L_{\text{engA}}\big)$, $\mathbf{m}\big(L_{\text{engP}}\big)$, and $\mathbf{m}\big(L_{\text{engA, engP}}\big)$, respectively, are compared. Since we eliminate the rotation dependency (see 2.3.1), the dimension d of the feature spaces $\mathbf{m}\big(L_{\text{engA}}\big)$, $\mathbf{m}\big(L_{\text{engP}}\big)$, and $\mathbf{m}\big(L_{\text{engA, engP}}\big)$ are $2N_\zeta$, $2N_\zeta$ and $4N_\zeta$, respectively.

The correctness of the novel local phase measure with $\kappa = 0.567$ is similar to that of the well-known local amplitude measure with $\kappa = 0.582$ according to $\underline{DS}_1$. Combination of both feature spaces increases the performance of the texture characterization significantly, up to $\kappa = 0.759$. In case of $\underline{DS}_2$ the phase features outperform both the amplitude features and as the combined features, yielding $\kappa = 0.838$. The GW-related results according to $\underline{DS}_1$ outperform the LG-related ones. For this filter bank, the best result of $\kappa = 0.679$ is achieved for $\mathbf{m}\big(L_{\text{engA, engP}}\big)$. Nevertheless, while the overall best result for $\underline{DS}_2$ is again related to GW, the correctness value of $\kappa = 0.777$ is better for LG than for GW with $\kappa = 0.760$.

Rows 4–6 in tab. 1 show the corresponding results for the RGB color space. Consequently, d rises by a factor of 3 in comparison to Rows 1–3. The relationship between the feature spaces remains constant in comparison to the grayscale features. Again, the values for the amplitude energy features and for the phase energies are similar for both datasets and both filter bank variants. For $\underline{DS}_1$ as well as $\underline{DS}_2$, the best results of $\kappa = 0.856$ and $\kappa = 0.826$, respectively, are obtained for the combined feature space. All the correctness values of the RGB-based color texture features are higher than the corresponding values in the grayscale domain.

For comparison with already existing methods of color texture analysis using Gabor

filters, Rows 7–8 show the correlation features $\mathbf{m}(RGB_{\text{corr}})$ and $\mathbf{m}(RGB_{\text{engA,corr}})$ following (37) [19]. In the latter case, they are combined with the amplitude energy. JAIN stated a nearly 100% correctness for his dataset when applying these combined features. However, we cannot reproduce this high performance on our data. While the correctness of the correlation features without combination is rather low, $\kappa < 0.270$, the combined features $\mathbf{m}(RGB_{\text{engA,corr}})$ decrease in performance of single $\mathbf{m}(RGB_{\text{engA}})$ features significantly in all experiments.

Rows 9–11 show the correctness values for the complex color space. Comparing the results of amplitude and phase features, we find that for $\underline{DS}_1$ the amplitude energies $\mathbf{m}(Z_{\text{engA}})$ and for $\underline{DS}_2$ the phase energies $\mathbf{m}(Z_{\text{engP}})$ perform best. As for the RGB color space, in all experiments combination of both feature spaces increase the classification rate. The best results are achieved for the GW filter bank with $\kappa = 0.814$ for $\underline{DS}_1$ and $\kappa = 0.745$ for both GW and LG for $\underline{DS}_1$. Comparing the color spaces, RGB outperforms the complex color space in all experiments. The best correctness using $\underline{DS}_1$ and $\underline{DS}_2$ and the RGB color space are comparable, $\kappa = 0.856$ and $\kappa = 0.826$, respectively. However, they are rather different in the complex space ($\kappa = 0.814$ and $\kappa = 0.745$).

7. Discussion and conclusion

Studying the novel phase energy features, our results indicate at least similar performance to the amplitude energy features in the grayscale as well as the color domain. Even more important is the increase of performance for the combination of both feature spaces in nearly all experiments. This indicates different focal points of these two features, which supplement each other. Therefore, the phase is useful not only for texture segmentation but also for texture classification.

According to the comparison of two Gabor filter bank variants, in individual cases substantial performance differences between GW and LG can be stated. Statistically, these differences do not establish a trend towards one of the filter banks. Therefore, a better performance of the LG filter bank stated in [31] cannot be validated. For classification purposes, both Gabor filter banks are suitable.

The bad performance of the RGB-based correlation features can be explained on the background of different databases and classifiers. Unfortunately, the database of JAIN is not available for the public. Therefore, reliable tests for comparison are not possible.

Referring to the first question posed in the introduction, grayscale-based and integrated color texture analysis have to be compared. A better performance of RGB color features is obvious. The single results of the amplitude and energy features as well as their combination outperform the corresponding grayscale-based results. Our results demonstrate that the intensity textures on the color bands are not completely redundant, because the combination of them increases their discrimination ability and, hence, information. Therefore, color is able to enhance grayscale texture features.

The second question posed in the introduction is related to intensity-independent

			$\underline{DS_1}$		$\underline{DS_2}$	
		d	GW	LG	GW	LG
1	$\mathbf{m}\left(L_{\text{engA}}\right)$	$2N_\zeta$	0.582	0.578	0.760	0.777
2	$\mathbf{m}\left(L_{\text{engP}}\right)$	$2N_\zeta$	0.567	0.459	**0.838**	0.676
3	$\mathbf{m}\left(L_{\text{engA, engP}}\right)$	$4N_\zeta$	**0.759**	0.679	0.821	0.821
4	$\mathbf{m}\left(RGB_{\text{engA}}\right)$	$6N_\zeta$	0.714	0.713	0.801	0.819
5	$\mathbf{m}\left(RGB_{\text{engP}}\right)$	$6N_\zeta$	0.692	0.701	0.810	0.820
6	$\mathbf{m}\left(RGB_{\text{engA, engP}}\right)$	$12N_\zeta$	**0.856**	0.831	0.824	**0.826**
7	$\mathbf{m}\left(RGB_{\text{corr}}\right)$	$18N_\zeta - 12$	0.199	0.190	0.262	0.257
8	$\mathbf{m}\left(RGB_{\text{engA,corr}}\right)$	$24N_\zeta - 12$	0.592	**0.675**	**0.591**	0.583
9	$\mathbf{m}\left(Z_{\text{engA}}\right)$	$4N_\zeta$	0.702	0.701	0.699	0.676
10	$\mathbf{m}\left(Z_{\text{engP}}\right)$	$4N_\zeta$	0.618	0.589	0.730	0.713
11	$\mathbf{m}\left(Z_{\text{engA,engP}}\right)$	$8N_\zeta$	**0.814**	0.777	**0.745**	**0.745**

Tab. 1. Classification results for the VisTex and the BarkTex images, $\underline{DS_1}$ and $\underline{DS_2}$, respectively. Rows 1-3 refer to grayscale results, Rows 4-8 refer to RGB-related results and Rows 9-11 are based on the complex color domain. Three kinds of feature vectors are studied: energy measure based on local amplitudes (engA), energy measure based on local phase changes (engP), and a measure based on the correlation between the local amplitudes of different Gabor coefficient images. The best result for each block and each dataset is marked in bold.

color textures. Assuming a strict separation of texture as an intensity pattern and color as a global feature, texture analysis in the complex space is expected to perform not significantly better than random means. In measuring an intensity-independent pattern, our results are much better than this expection. Therefore, their existence has been proved experimentally. Since the intensity-independent and intensity-dependent pattern are non-redundant, the corresponding features may supplement each other, resulting in a further enhancement of correctness. For this, a significant reduction of the feature space dimensionality has to be obtained first, to get reliable feature space sizes and to disregard the "curse of dimensionality", which means a decrease in classification rates with increasing feature space dimensionality, [2].

Concluding our results, color has been shown to enhance intensity texture features

as well as composing an intensity independent-pattern. Both RGB and HSL color space are suitable for applying Gabor filters. Concerning the Gabor features, the phase energy supplements the amplitude energy, showing similar discrimination capabilities.

References

1946
[1] Gabor D.: Theory of Communication, Journal of IEE, 93, 429-457.

1973
[2] Duda R.O., Hart P.E.: Pattern Classification and Scene Analysis, John Wiley & Sons, NY.

[3] Haralick R.M., Shanmugam K., Dinstein I.: Textural features for image classification, IEEE Trans. SMC, 3, 610-621.

1975
[4] Oppenheim A.V., Schafer R.W.: Digital Signal Processing, Prentice-Hall, Englewood Cliffs, NJ.

1978
[5] Harris F.J.: On the use of windows for harmonic analysis with the discrete Fourier-transform, Proc. IEEE, 66, 51-83.

1986
[6] Derin H., Cole W.S.: Segmentation of textured images using gibbs random fields, CVGIP, 35, 72-98.

1988
[7] Frey H.: Digitale Bildverarbeitung in Farbräumen, PhD thesis, Technical University of Munich, Germany (in German).

1990
[8] Bovic A.C., Clark M., Geisler W.S.: Multichannel texture analysis using localized spatial filters, IEEE Trans. PAMI, 12(1), 55-73.

1991
[9] Jain A.K., Farrokhnia F.: Unsupervised texture segmentation using Gabor filters, PR, 24(12), 1167-1186.

1992
[10] Chang T., Kuo C.C.J.: Tree-structured wavelet transform for textured image segmentation. Proc. of the SPIE, 1770, 394-405.

[11] Daubechies I.: Ten Lectures on Wavelets. CBMS-NSF Regional Conf. Series in Applied Mathematics 62, Society for Industrial and Applied Mathematics, Philadelphia.

1995
[12] Myler H., Weeks A., Voicu L.: RGB Color enhancement using homomorphic filtering, Proc. of the SPIE, 2421, 43-50.

[13] Thornton A., Sangwine S.: Colour object location using complex coding in the frequency domain. IEE 5th Int. Conf. on Image Processing and Applications, Edinburgh, UK, 820-824.

1996
[14] Sangwine S.J.: Fourier transforms of colour images using quaternion or hypercomplex numbers. Electronic Letters, 32(21), 1979-1980.

1997
[15] Fountain S.R., Tan T.N.: Rotation invariant texture features from Gabor filters. 3rd Asian Conf. on Computer Vision (ACCV), Hong Kong, China, Proceedings, Springer, Berlin, 2, 57-64.

[16] Lakmann R., Priese L.: A reduced covariance color texture model for micro-textures. Proc. 10th Scandinavian Conf. on Image Analysis, Lappeenranta, Finland, 947-953.

[17] Sangwine S.J.: The discrete quaternion Fourier transform. 6th Int Conf. on Image Processing and its Applications, Dublin, Ireland, 2, 790-793.

1998

[18] Belongie S., Carson C., Greenspan H., Malik J.: Color- and texture-based image segmentation using EM and its application to content-based image retrieval. 6th Int. Conf. on Computer Vision, Bombay, India, 675-682.

[19] Jain A.K., Healey G.: A multiscale representation including opponent color features for texture recognition. IEEE Trans. on Image Processing, 7(1), 124-128.

1999

[20] Haley G.M., Manjunath B.S.: Rotation-invariant texture classification using a complete space-frequency model. IEEE Trans. on Image Processing, 8(2), 255-269.

[21] Hauta-Kasari M., Parkkinen J., Jaaskelainen T., Lenz R.: Multi-spectral texture segmentation based on the spatial co-occurrence matrix. Pattern Analysis & Applications, 2(4), 275-284.

[22] Randen T., Husøy J.H., Filtering for texture classification: a comparative study. IEEE Trans. on PAMI, 21(4), 291-310.

[23] Van de Wouwer G., Scheunders P., Livens S., Van Dyck D.: Wavelet correlation signatures for color texture characerization. PR, 32, 443-451.

2000

[24] Dubisson-Jolly M.-P., Gupta A.: Color and texture fusion: application to aerial image segmentation and GIS updating. Image & Vision Computing, 18(10), 823-832.

[25] McCabe A., Caelli T., West G., Reeves A.: Theory of spatiochromatic image encoding and feature extraction. J. of the Optical Society of America A, 17(10), 1744-1754.

[26] Nicolás J., Yzuel M.J., Campos J.: Colour information as a third dimension in Fourier transform and correlation. Proc. Int. Conf. on PR (ICPR), Barcelona, Spain, 2, 515-518.

[27] Plataniotis K.N., Venetsanopoulos A.N.: Color Image Processing and Applications. Springer, Berlin, 2000.

[28] Palm C., Keysers D., Lehmann T.M., Spitzer K.: Gabor filtering of complex hue/saturation images for color texture classification. Wang PP (Ed.) Proc. of the 5th Joint Conf. on Information Science, 2, The Association for Intelligent Machinery, Atlantic City, NJ, 45-49.

[29] Palm C., Lehmann T.M., Spitzer K.: Color Texture Analysis of Moving Vocal Cords using Approaches from Statistics and Signal Theory, In: Braunschweig T., Hanson J., Schelhorn-Neise P., Witte H. (Eds.): Advances in Quantitative Laryngoscopy, Voice and Speech Research, 4th Int. Workshop, Friedrich-Schiller University, Jena, Germany, 49-56.

[30] Sangwine S.J.: Colour in image processing. Electronics and Communication Engineering Journal, 12(5), 211-219.

[31] Smeraldi F., Carmona O., Bigün J.: Saccadic search with Gabor features applied to eye detection and real-time head tracking, Image & Vision Computing, 18, 323-329.

2001

[32] Kukkonen S., Kälviäinen H., Parkkinen J.: Color features for quality control in ceramic tile industry. Optical Engineering, 40(2), 170-177.

[33] Pei, Soo-Chang, Ding, Jian-Jiun, Chang, Ja-Han: Efficient implementation of quaternion Fourier transform, convolution, and correlation by 2-D complex FFT, IEEE Trans. on Signal Processing, 49(11).

Christoph Palm studied computer science at the Aachen University of Technology (RWTH). He received his master's degree in computer science in 1997 from the RWTH. Between 1997 and 2001 he was PhD student at the Institute of Medical Informatics, Department of Medical Image Processing, at the University Hospital Aachen. During his research work his main interest covered all aspects of pattern recognition, texture classification, and color image processing. His PhD thesis about integrative analysis of color and texture is now in the reviewing process. He is co-editor of a proceedings volume (Quantitative Laryngoscopy, Voice and Speech Research) and an author of several scientific conference and journal papers. He is peer reviewer of the Optical Society of America.

Christoph Palm joined the AIXPLAIN Inc., Aachen, Germany, in 2001. Due to the focus of the company, his research in pattern recognition is now complemented by the information retrieval and machine translation field.

Thomas M. Lehmann received his master's degree in electrical engineering and the PhD degree (summa cum laude) in computer science from the Aachen University of Technology (RWTH), Aachen, Germany, in 1992 and 1998, respectively.

In 1992, he was research scientist at the Faculty of Electrical Engineering, RWTH Aachen. Since 1992 he has been with the Institute of Medical Informatics, Medical Faculty, RWTH Aachen, where he currently heads the Department of Medical Image Processing at the assistant professor level (Wissenschaftlicher Assistent). He co-authored a textbook on image processing for the medical sciences (Springer-Verlag, Berlin, Germany, 1997) and co-edited the Handbook of Medical Informatics (Hanser Verlag, Munich, Germany, 2002). His research interests are discrete realizations of continuous image transforms, medical image processing applied to quantitative measurements for computer-assisted diagnoses and content-based image retrieval from large medical databases.

Dr. Lehmann received the DAGM-Preis '93. The award from the German Association for Pattern Recognition was given for his work on automatic strabometry using Hough transform and covariance filtering. In 1998, he received the Borcher's Medal from the RWTH Aachen for his work on medical image registration and interpolation. He is the Chairman of the German Workshop on Medical Image Processing and a Vice-President of the working group Medical Image Processing within the German Society of Medical Informatics, Biometry and Epidemiology (GMDS). He is the Chairman of the IEEE Joint Chapter Engineering in Medicine and Biology (IEEE German Section). He is member of the Institute of Electrical and Electronics Engineers (IEEE), the International Association of Dentomaxillofacial Radiology (IADMFR), and the Society of Photo-Optical Instrumentations Engineering (SPIE). He serves on the International Editorial Board of Dentomaxillofacial Radiology (a Nature publishing group).