Color texture measurement and segmentation

Minh A. Hoang, Jan-Mark Geusebroek*, Arnold W.M. Smeulders

ISLA, Informatics Institute, Faculty of Science, University of Amsterdam, Kruislaan 403, 1098 SJ Amsterdam, Netherlands

Abstract

In computer vision, measurement of image properties such as color or texture is essential. In this paper, we propose a solid framework for the local measurement of texture in color images. We give a physical basis for the integration of the well-known Gabor filters with the measurement of color. Our method implies that the color–texture is measured in the wavelength-Fourier domain. The measurement filter in this domain boils down to a 3D Gaussian, representing a Gabor–Gaussian in the spatial-color domain. In addition, the extended measurements invariant to shadow or shading can be derived directly from the proposed method. We apply a simple segmentation algorithm to illustrate the performance of our proposed color–texture measurement. The derived color–texture filter is demonstrated to be accurate in capturing texture statistics. In addition, the method is compatible with the measurement of gray-value texture.
© 2004 Elsevier B.V. All rights reserved.

Keywords: Color texture; Gaussian probes; Gabor filters; Shadow invariant; Segmentation

1. Introduction

In image processing, the light as it bounces off the scene is the result of many different causes. In image retrieval and in many other tasks of image processing, we are interested in the light response of only one of them: the light-characteristics of the object embedded in the scene. In this paper, we summarize the causes for the purpose of separating the conditions intrinsic to the object's appearance from the accidental scene.

To handle the problem, one could model the influence of the scene on the appearance of the object, or one could try to capture the intrinsic properties of the object in invariant features. At any rate, modelling scene-specific circumstances has to be bootstrapped by the second approach of invariant characteristics. The invariant approach has the advantage of being less complex at the expense of throwing away essential information. For a complete analysis, neither of the two approaches can be missed. One or another basic invariant observations will bootstrap a model which may invoke more detailed invariant

---
*Corresponding author. Tel.: +31 20 525 7552; fax: +31 20 525 7490.

E-mail addresses: minh@science.uva.nl (M.A. Hoang), mark@science.uva.nl (J.-M. Geusebroek), smeulders@science.uva.nl (A.W.M. Smeulders).

descriptions. In turn, it will bootstrap a model of the scene and so on. In view of this dichotomy, we aim to advance object retrieval in broad domains from tight invariant descriptions.

The aim of invariant descriptions is to identify objects at the loss of the smallest amounts of the information content. If two objects or two appearances of the same object t_i are equivalent under a group of transformations G they are said to belong to the same equivalence class

$$t_1 \overset{G}{\sim} t_2 \Longleftrightarrow \exists g \in G : t_2 = g \circ t_1. \tag{1}$$

A property f of t is invariant under G if and only if f_t remains the same regardless of the unwanted condition expressed by G,

$$t_1 \overset{G}{\sim} t_2 \Longrightarrow f_{t_1} = f_{t_2}. \tag{2}$$

In general, a feature with a very wide class of invariance looses the power to discriminate among essential differences. The size of the class of images considered equivalent grows with the dimensionality of G. In the end, the invariance may be so wide that little discrimination among objects is retained. The aim is to select the tightest set of invariants suited for the expected set of non-constant conditions. Hence, in the context of image retrieval, the invariant conditions are to be specified (indirectly) by the user as they reflect the intentions of the user. The oldest work on invariance in computer vision has been done in object recognition. Invariant description in image retrieval is relatively new, but quickly gaining ground. This presentation feeds on the much larger [7,21].

In this paper, we consider the invariant assessment of color and texture in combination. Much work on texture measurement has been done, of which we consider as most important the work on Gabor filters [2] and Gaussian derivative filters [5,17]. These measurements are often referred to as textons [14]. Measurement of color and texture in combination, rather than color or texture alone, would provide better discriminating power. The combination of color and texture has attracted attention in recent literature. In Mirmehdi and Petrou [19], color–textured images are roughly segmented based on a spatial color

model [25]. The assumption underlying their approach implies that texture can be characterized by its color histogram over a region. The drawback here is that the spatial structure of the texture is not considered since only first order statistics, the histogram, is taken into account. Thai and Healey [22] propose measuring color–texture by embedding the Gabor filters into an opponent color representation. Hall et al. [10] apply Gaussian derivative filters to an opponent color space to describe colored texture. Both methods provide a useful structural representation for color–texture. However, the combination of Gabor filtering and opponent color space is put forward as an empirical result. The methods are not well-defined from a physical point of view. We aim at a well-founded integrated color–texture method for the local measurement of color–texture. To achieve this, we adopt the scale-space framework as the theory of local measurement. We give a solid physical basis for the integration of well-known texture measurements and the measurement of color. We demonstrate the performance of the proposed measurement in segmentation of both synthetic and natural images.

The outline of this paper is as follows. Section 2 presents our approach to the measurement of color–texture. Section 3 gives further extension of the measurements, which are invariant to cast shadows. In Section 4, we briefly describe a practical method for estimating color conforming to our measurement scheme. In this section, we put forward a simple scheme for color–texture segmentation to illustrate the method. We conclude in Section 5.

2. Measuring of color–texture

A color image is observed by integrating over some spatial extent and over a spectral bandwidth. Before observation, a color image may be regarded as a three-dimensional energy density function $E(x, y, \lambda)$, where (x, y) denotes the spatial coordinate and λ denotes the wavelength. Observation of the energy density $E(x, y, \lambda)$ boils down to correlation of the incoming signal with a measurement

probe $p(x, y, \lambda)$,

$$\hat{Q}(x, y, \lambda) = \int \int \int E(x, y, \lambda) p(x, y, \lambda) \, dx \, dy \, d\lambda. \tag{3}$$

For a linear spatially shift invariant imaging system, correlation boils down to convolution

$$\hat{Q}(x, y, \lambda) = \int E(x, y, \lambda) * p(x, y, \lambda) \, d\lambda. \tag{4}$$

The yet unknown measurement function $p(x, y, \lambda)$ estimates quantities of the energy density $E(x, y, \lambda)$. In case of texture, we are interested in the local spatial frequency characteristics of $E(x, y, \lambda)$. These properties are better investigated in the domain of spatial frequency. Thus, it is appropriate to represent the joint color–texture properties in a combined *wavelength-Fourier* domain $\mathscr{E}(u, v, \lambda)$, where λ remains the wavelength of the light energy, and (u, v) denotes the spatial frequency,

$$\hat{\mathscr{Q}}(u, v, \lambda) = \int \mathscr{E}(u, v, \lambda) \mathscr{P}(u, v, \lambda) \, d\lambda. \tag{5}$$

Scale space theory suggests that the probe should have a Gaussian shape in order to prevent the probe from adding extra details to the function when observed at a coarser scale [16]. Therefore, the Gaussian function is chosen to probe the signal $\mathscr{E}(u, v, \lambda)$ along its three axes. The measurement of the signal $\mathscr{E}(u, v, \lambda)$ at a given spatial frequency (u_0, v_0) and wavelength λ_0 is obtained by a 3D Gaussian probe centered at (u_0, v_0, λ_0) at a frequency scale σ_f and wavelength scale σ_λ

$$\hat{\mathscr{M}}(u, v, \lambda) = \int \mathscr{E}(u, v, \lambda) G(u - u_0, v - v_0, \\ \lambda - \lambda_0; \sigma_f, \sigma_\lambda) \, d\lambda. \tag{6}$$

Note that the Gaussian probe is a separable function, we can rewrite Eq. (6) as

$$\hat{\mathscr{M}}(u, v, \lambda) = \int \mathscr{E}(u, v, \lambda) G(u - u_0, v - v_0; \sigma_f) \\ \times G(\lambda - \lambda_0; \sigma_\lambda) \, d\lambda. \tag{7}$$

Frequency selection is achieved by tuning the parameters u_0, v_0, and σ_f, and color information is captured by the Gaussian specified by λ_0 and σ_λ.

In physic based vision, the use of filter functions for probing the input spectrum has been presented in many works in literature [1,4,7,8]. In Geusebroek et al. [6], it is shown that color measurements can be achieved by probing with Gaussian derivative apertures, combined with the Gaussian aperture for luminance or intensity measurement. These derived measurements fit well into the scalespace framework, in that no extra detail is created. More importantly, the measurements correspond well with the opponent color theory for human vision [11]. Here, the first-order Gaussian derivative probe compares the blue region at the color spectrum with the yellow part, whereas the secondorder measurement compares the middle green part with the outer (magenta) regions. More formally, we write

$$\hat{E}_{\lambda^{(n)}}(x, y) = \int E(x, y, \lambda) G_n(\lambda - \lambda_0; \sigma_\lambda) \, d\lambda, \tag{8}$$

where $\hat{E}_{\lambda^{(n)}}(x, y)$, $n = 0, 1, 2, \dots$ is the color measurement of $E(x, y, \lambda)$ obtained by sampling with a nth order Gaussian derivative. In practice, it is sufficient to take up to the secondorder derivative since higher derivatives do not affect color as observed by the human visual system [6]. We will use $\hat{E}(x, y)$, $\hat{E}_\lambda(x, y)$, and $\hat{E}_{\lambda\lambda}(x, y)$ for $\hat{E}_{\lambda^{(n)}}(x, y)$ with $n = 0, 1, 2$, respectively.

We now transform Eq. (7) back to the wavelength-spatial domain (x, y, λ), taking the Gaussian derivative probes into account. The multiplication with the shifted Gaussian $G(u - u_0, v - v_0; \sigma_f)$ in Eq. (7) is equivalent to the convolution with a Gabor filter in the spatial domain [2,12]. Therefore, the combined color–texture measurement in the (x, y, λ) domain at wavelength λ_0 is

$$\hat{M}(x, y, \lambda) = h(x, y) * \int E(x, y, \lambda) G_n(\lambda - \lambda_0; \sigma_\lambda) \, d\lambda \tag{9}$$

or for short

$$\hat{M}(x, y, \lambda) = h(x, y) * \hat{E}_{\lambda^{(n)}}(x, y), \tag{10}$$

where

$$h(x, y) = \frac{1}{2\pi\sigma_s^2} e^{-(x^2 + y^2)/2\sigma_s^2} e^{2\pi j(Ux + Vy)} \tag{11}$$

is the 2D Gabor function at the radial center frequency $F = \sqrt{U^2 + V^2}$ (cycles/pixel) and the filter orientation $\tan(\theta) = V/U$, and $j^2 = -1$. Furthermore, the color probes are given by the Gaussian derivatives

$$G_n(\lambda - \lambda_0; \sigma_\lambda) = \frac{\sigma_\lambda^n}{\sqrt{2\pi}\sigma_\lambda} \frac{\partial^n}{\partial \lambda^n} e^{-(\lambda - \lambda_0)^2 / 2\sigma_\lambda^2}. \qquad (12)$$

The probe functions in Eq. (9) are illustrated in Fig. 1.

We have achieved our goal of integrating texture and color measurement. The color–texture measurement is obtained by sampling the wavelength-Fourier domain with a three-dimensional Gaussian probe. This leads to the use of Gabor filter on the opponent color representation, as empirically derived by Jain and Healey [13].

3. Shadow and shading invariant

Within a single texture patch, the value of the Gabor filter response varies proportionally to the local intensity of the texture. The darker region has a smaller response value than the value of the brighter one. Therefore, the shadow effect may compromise the post processing such as segmentation or region interpretation. In this section, we present an approach to correct the shadow and
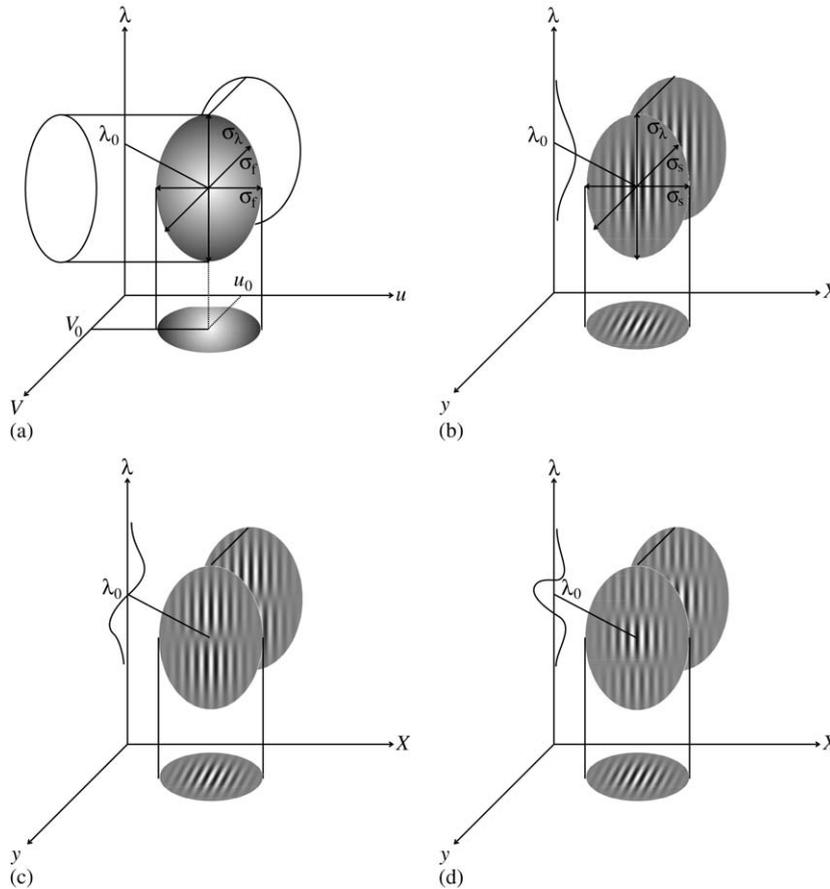

Fig. 1. The 3D probe functions: (a) the color–texture probe in the wavelength-Fourier domain having a 3D Gaussian aperture. The lightness in the aperture indicates the value of the probe function. (b) The probe for color–texture in the wavelength-spatial domain (x, y, λ). This is a product of a 2D Gabor in the spatial domain and a Gaussian in the wavelength domain. (c) and (d) the Gaussian derivatives in the wavelength domain.

shading effect with the assumption that the camera is not saturated.

3.1. Colored illumination and matte dull surfaces

Consider an image before observation $E(x, y, \lambda)$. Under the *Lambertian* reflection model, it may be decomposed into a component $e(x, y)$ denoting variations of the incoming illumination spectrum, and a component $R_\infty(x, y, \lambda)$ representing the material reflectance at each location, resulting in $E(x, y, \lambda) = e(x, y)R_\infty(x, y, \lambda)$. Here, $R_\infty(x, y, \lambda)$ bears the color–texture information of the object. The goal is to remove the undesired varying illumination spectrum $e(x, y)$ from the observed image $\hat{E}(x, y)$.

To factor out the intensity changes across the extent of the image plane, we take the approach similar to homomorphic filtering method [23], in which we first apply the log function to transform the signal from multiplicative form into additive form. Applying the high-pass filtering on the additive form of the signal assures that the low frequency part is factored out, since in this case the form remains additive in the frequency domain. Transforming the multiplicative image formation into the log domain,

$$\log E(x, y, \lambda) = \log e(x, y, \lambda) + \log R_\infty(x, y, \lambda) \tag{13}$$

yields an addition of a smoothly varying illumination spectrum component and a high frequency surface texture component. Since most of the incoming spectra in the real world are smoothly varying functions of the wavelength λ, or in other words $R_\infty(x, y, \lambda)$ is smoothly varying along the λ-axis, sampling $R_\infty(x, y, \lambda)$ with a Gaussian aperture centered at λ_0 provides well approximate measurement of $R_\infty(x, y, \lambda_0)$. We can therefore write

$$\log \hat{E}(x, y) \approx \log e(x, y, \lambda_0) + \log R_\infty(x, y, \lambda_0). \tag{14}$$

High-pass filtering of $\log \hat{E}(x, y)$ suppresses the influence of $e(x, y, \lambda_0)$. Hence, the application of the Gabor filters on the log-transformed image

yields a shadow and shading invariant result,

$$S(x, y) = \text{Gabor}\{\log \hat{E}(x, y)\}$$
$$\approx \text{Gabor}\{\log R_\infty(x, y, \lambda_0)\}. \tag{15}$$

The color channels E_λ and $E_{\lambda\lambda}$ are obtained by differentiating E with respect to λ once and twice. In order to incorporate the color information while avoiding $e(x, y, \lambda)$, we differentiate Eq. (13) with respect to λ once and twice, and apply the Gabor filtering, which results in,

$$S_\lambda(x, y) = \text{Gabor}\left\{\frac{\hat{E}_\lambda}{\hat{E}}\right\}, \tag{16}$$

$$S_{\lambda\lambda}(x, y) = \text{Gabor}\left\{\frac{\hat{E}\hat{E}_{\lambda\lambda} - \hat{E}_\lambda^2}{\hat{E}^2}\right\}. \tag{17}$$

The measurements $\hat{E}_\lambda/\hat{E}$ and $(\hat{E}\hat{E}_{\lambda\lambda} - \hat{E}_\lambda^2)/\hat{E}^2$ conform to the suggested invariant measurements in Geusebroek et al. [7]. From Eqs. (15)–(17) we achieved our goal of measuring color–texture with regard to shadow and shading invariant.

3.2. White but uneven illumination and matte dull surfaces

A more special case can be considered in which the pre-observed color–texture image $E(x, y, \lambda)$ is the multiplication of the incoming white illumination $i(x, y)$ and the material reflectance $R_\infty(x, y, \lambda)$

$$E(x, y, \lambda) = i(x, y)R_\infty(x, y, \lambda). \tag{18}$$

The log domain of Eq. (18) has the form

$$\log E(x, y, \lambda) = \log i(x, y) + \log R_\infty(x, y, \lambda). \tag{19}$$

Again, under the assumption that $i(x, y)R_\infty(x, y, \lambda)$ is smoothly varying along λ-axis, we derive the following approximation:

$$\log \hat{E}(x, y) \approx \log i(x, y) + \log R_\infty(x, y, \lambda_0). \tag{20}$$

As $i(x, y)$ is slowly varying over the scene, shadow and shading invariant may be obtained by applying Gabor filtering,

$$S(x, y) = \text{Gabor}\{\log \hat{E}(x, y)\}$$
$$\approx \text{Gabor}\{\log R_\infty(x, y, \lambda_0)\}. \tag{21}$$

With similar reasoning, and note that $i(x, y)$ is constant over λ-axis, we derive the invariant result

for color channels,

$$S_\lambda(x, y) = \text{Gabor}\{\log \hat{E}_\lambda(x, y)\}$$
$$\approx \text{Gabor}\{\log R_{\infty\lambda}(x, y, \lambda_0)\}, \tag{22}$$

$$S_{\lambda\lambda}(x, y) = \text{Gabor}\{\log \hat{E}_{\lambda\lambda}(x, y)\}$$
$$\approx \text{Gabor}\{\log R_{\infty\lambda\lambda}(x, y, \lambda_0)\}. \tag{23}$$

Therefore, in the special case of white illumination, we obtained the measurements for shadow and shading invariant from Eqs. (21)–(23).

4. Experiments

Up to this point, we have established a unified framework for color–texture measurement. Since the probe functions (Fig. 1) are decomposable, they can be represented as a product of three one-variable functions of λ, x and y. The measurement is carried out in two steps: color measurement with the Gaussian color model and then texture measurement with Gabor filters. We employ a simple segmentation algorithm with a scheme similar to Jain and Farrokhnia [12] to illustrate the performance of our proposed color–texture measurement. The overall scheme is depicted in Fig. 2.

4.1. Implementation of color–texture measurement

A camera is developed to capture the same color space as humans, hence we assume the RGB-sensitivities to span a similar spectral bandwidth and to have a similar central wavelength. When camera response is linearized, a RGB-camera approximates the CIE 1964 XYZ basis for colorimetry by the linear transform [3]. Note that we try to achieve derivative filters in the spectral domain by transforming the spectral responses as given by the RGB-filters. The transformed filters may be imperfect, but are likely to offer accurate estimates of differential measurements. When the spectral responses of the RGB-filters are known, a better transform can be obtained.

Gaussian color measurements are obtained by tuning the parameters λ_0 and σ_λ. Using the expressions describing similarity between different
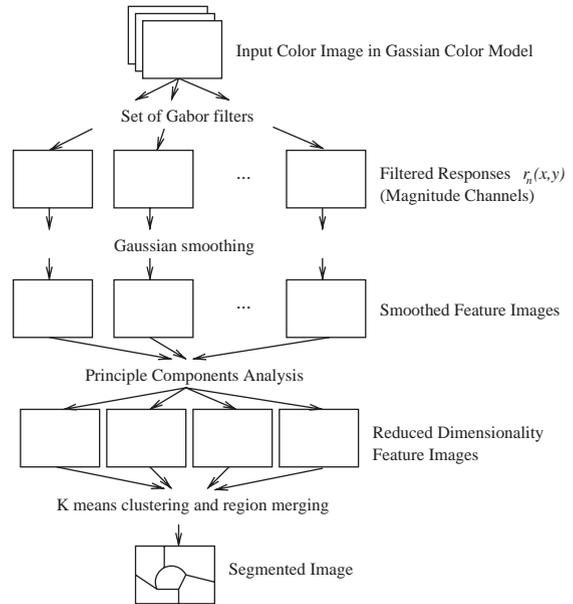

Fig. 2. The color–texture segmentation scheme.

subspaces [9], we compare the subspace of the Gaussian color model with the human visual system by using the XYZ color matching function. Hence, parameters for the Gaussian color model may be optimized to capture a similar spectral subspace as spanned by human vision. Let the Gaussian color matching functions be given by $G(\lambda - \lambda_0; \sigma_\lambda)$. We have two degrees of freedom in positioning the subspace of the Gaussian color model: the mean λ_0 and scale σ_λ of the Gaussian. We wish to find the optimal subspace that minimizes the largest principal angle between the subspaces. An approximate solution is obtained for $\lambda_0 = 520$ nm and $\sigma_\lambda = 55$ nm [6]. For a RGB-camera, we can find the best linear transform from RGB to the Gaussian color model as shown in Geusebroek et al. [6]. The desired implementation of the Gaussian color model in RGB terms is given by

$$\begin{bmatrix} \hat{E} \\ \hat{E}_\lambda \\ \hat{E}_{\lambda\lambda} \end{bmatrix} = \begin{pmatrix} 0.06 & 0.63 & 0.31 \\ 0.19 & 0.18 & -0.37 \\ 0.22 & -0.44 & 0.06 \end{pmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}.$$

Color–texture measurements are obtained by applying a set of Gabor filters on each channel

$\hat{E}$, $\hat{E}_\lambda$ and $\hat{E}_{\lambda\lambda}$ (Eq. (9)). Each measurement results in two terms: the magnitude response $r(x, y)$ and the phase response $p(x, y)$. The magnitude is computed from the real and imaginary part of $\hat{E}$ by

$$r(x, y) = \sqrt{\hat{E}_{re}^2 + \hat{E}_{im}^2}. \tag{24}$$

The magnitude response emphasizes texture regions, which are in tune with the chosen frequencies of the filter. While the phase response describes the texture transitions. In this paper, we are only interested in the magnitude responses and use them as the output of measurement. Methods for designing an efficient set of Gabor filters can be found in Jain and Farrokhnia [12], Weldon et al. [24], Manjunath and Ma [18]. In our experiment, we use 20 Gabor filters built from five scales $\sigma_s = 4$, 3.5, 2.95, 2.35, 1.75, corresponding to five center frequencies $F = 0.05$, 0.08, 0.14, 0.22, 0.33 (cycles/pixel), and four orientations $\theta = 0, \pm\pi/4, \pi/2$. These values of scale and center frequency are calculated based on the method proposed by Manjunath and Ma [18]. We therefore obtain 60 filtered response images from which we consider the magnitude, $r_n(x, y)$, $n = 1, \ldots, 60$. Each image pixel (x_i, y_j) is now represented by a 60-dimensional feature vector whose nth component is denoted by $r_n(x_i, y_j)$. Pixels in one color–texture homogeneous region will form a cluster in the feature space, which is compact and may be discriminated from clusters corresponding to other regions.

4.2. Segmentation

The segmentation algorithm is based on clustering pixels using their associated feature vectors. For preprocessing, every filtered magnitude image $r_n(x, y)$ is smoothed by a Gaussian kernel to suppress the variation of the feature vectors within the same color–texture region. Since the feature vectors are highly correlated, we apply the principal components analysis (PCA) [15] to reduce the feature space dimensionality down to four. The 4-dimensional feature vectors are used as the input for clustering. The clustering algorithm has two steps. In the first step, the k-means algorithm with large number of k is applied on the feature space to derive the initial clusters. In the second step, a region merging method is used to combine adjacent clusters which are statistically similar (Fig. 2).

The region merging is done in an agglomerative manner where in each iteration the two most similar regions are merged. We employ a region similarity measure analogous to the one proposed in Nguyen et al. [20]. The similarity between regions R_i and R_j is given by

$$\mathscr{S}_{i,j} = (\mu_i - \mu_j)^\top [\Sigma_i + \Sigma_j]^{-1} (\mu_i - \mu_j), \tag{25}$$

where μ_i, μ_j are the mean vectors and Σ_i, Σ_j are the covariance matrices computed from feature vectors of regions R_i and R_j, respectively. The formula recalls the Mahalanobis metric with an extension. Here, $S_{i,j}$ measures the distance between two sets. If one of the two reduces to a single point, $S_{i,j}$ becomes the Mahalanobis distance. The advantage of this measure is that the uncertainty of the vectors μ_i and μ_j as expressed by their respective covariances $\Sigma_{i,j}$ is taken into account. The two regions R_i and R_j are merged if the value of $\mathscr{S}_{i,j}$ is under a threshold. In our experiment, the similarity threshold t in the range of [6..9] produces almost the same result for every test image. Therefore, we fix the similarity threshold at $t = 7.5$ for all our experiments. Finally, a simple post-processing technique is utilized to remove small-sized isolated regions.

The segmentation results are illustrated in Fig. 3. The input image is created by combining five sub-images of natural and artificial color–texture. In this image, two patches on top are chosen to be similar in texture but different in color. The two patches on the left are chosen to be similar in color but different in texture. The results in Fig. 3 show that five regions are correctly discriminated when using the proposed measurement.

Results from the Corel database achieved by using the method on over 40,000 examples showed a remarkable good segmentation. Segmentations of real images using the proposed method are illustrated in Fig. 4. Furthermore, segmentation results obtained by using invariant features are shown in Fig. 5. In general cases, the results using
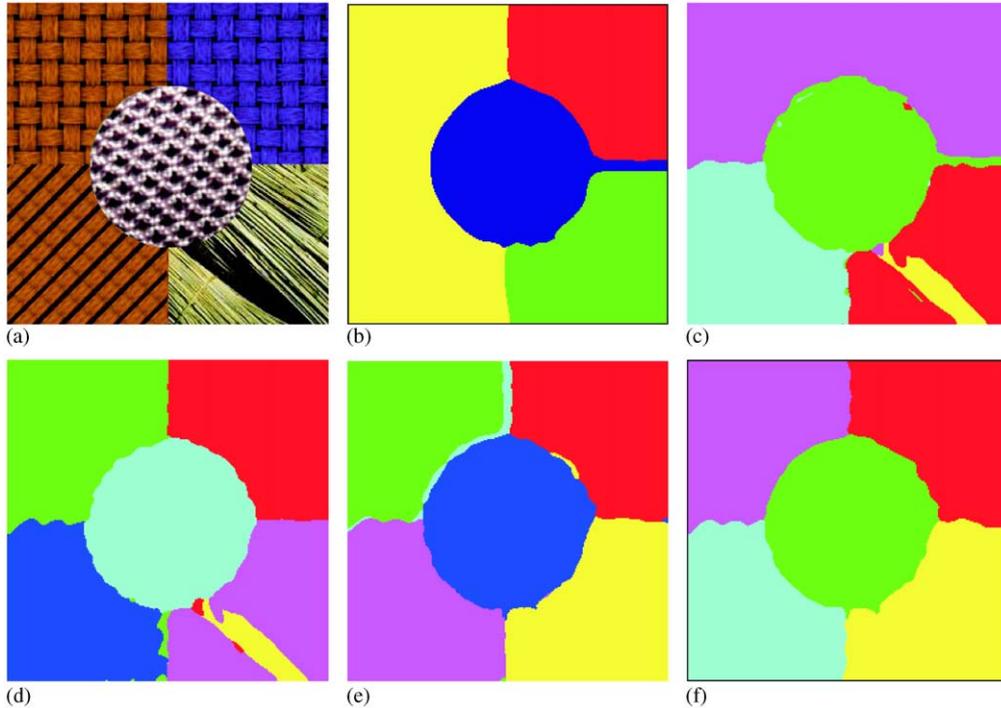
Fig. 3. (a) Synthetic color–texture image with five different regions. (b) The segmentation result using only color features. The original color image is smoothed by a set of Gaussian filters at different scales as in Mirmehdi and Petrou [19]. Here, two regions with identical color are merged. (c) Segmentation result using only gray-value texture. Note that the regions with identical texture but different color are merged. (d) The segmentation result using the proposed color–texture features without shadow invariance. The regions are correctly segmented, but affected by shadow. (e) The segmentation result using the shadow invariant color–texture feature. In this case, all regions are correctly segmented. (f) Post-processing of the invariant segmentation result to remove small isolated regions.

invariant features are worse since it has larger distorted boundary between segmented regions. The proposed algorithm takes about 8 s to segment an image from the Corel database on a state-of-the-art PC (Pentium III, 1 GHz CPU).

5. Conclusion

We have proposed a framework for the local measurement of texture in color images. Color–texture is analyzed in wavelength-Fourier space. We measure the spatial frequency by sampling the incoming image with a shifted Gaussian in the spatial frequency domain, and measure color by sampling the signal with Gaussian in wavelength domain. Therefore, color–texture measurement implies the sampling with a 3D Gaussian in a wavelength-Fourier space. This yields the decomposition of the measurement into a spectral Gaussian and a spatial Gabor filter. Hence, we have derived a solid physical basis for the integration of the Gabor texture measurement with opponent color measurement.

The proposed filters are the extension of the existing intensity Gabor filters to the combined color–texture filtering. Hence, the methods can be applied to both full color images and monochromatic images. We showed that the incorporation of the physics of light reflection results in shadow invariant texture segmentation. Therefore, robustness against various illumination conditions and cast shadows is achieved. When applying invariance for content-based retrieval, the degree of invariance should be tailored to the recording circumstances. Clearly, a feature with a very wide
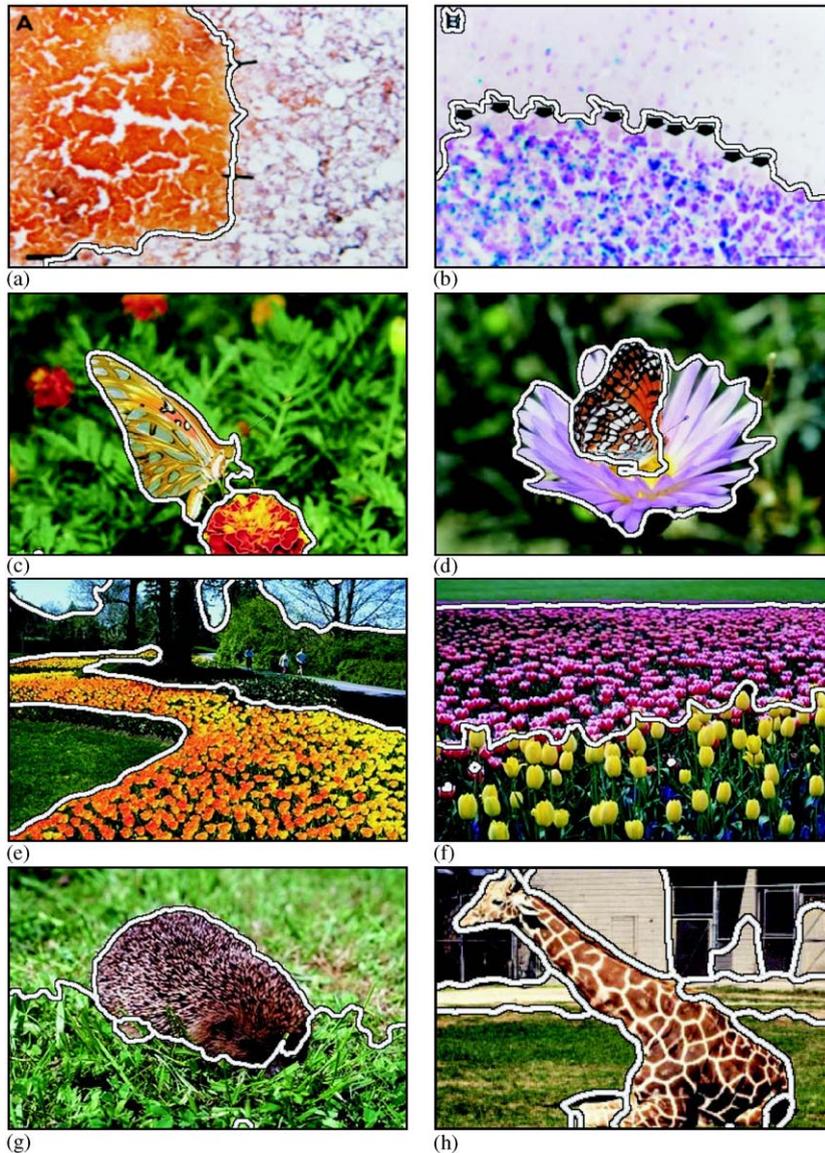
Fig. 4. Segmentations of real images. (a), (b) Segmentations of bright-field images from Elsevier database. (c), (d), (e), (f), (g), (h) Segmentations of natural images from Corel database. Note that all important objects are well identified.

class of invariants looses the power to discriminate among essential differences. The aim is to select the tightest set of invariants. What is needed is a complete set of image properties with well-described variant conditions that they are capable of handling, see Geusebroek et al. [7].

In practice, the results for non-invariants can be obtained by applying the Gabor filters directly to RGB values. However, RGB values are strongly correlated, hence smoothing or filtering would lead to mixing artifacts. In this paper, the proposed transformation of RGB to opponent derivative sensitivities directly links physics–reflectance model-based approaches to the domain of color imaging. Furthermore, invariant measures can be explicitly derived from the model.
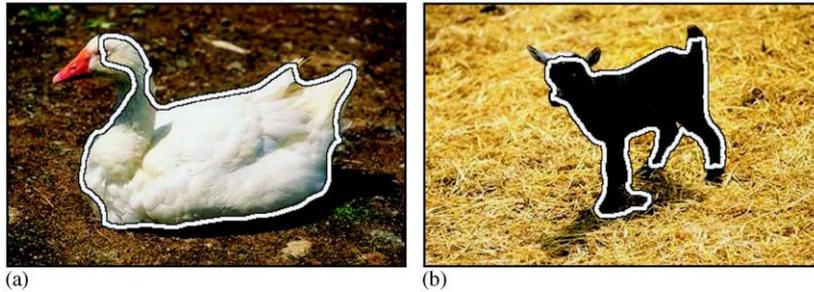
Fig. 5. Segmentations of real images using invariant features. Note that the backgrounds with casting shadows are well segmented.

Experimental results have shown that our color–texture measurements provide robust discriminating power. The algorithm has only one tunable parameter, the merging threshold. The choice for the value of this threshold was not critical, and only one value was used for all our experiments. Good segmentations have been achieved for the various natural images.

Acknowledgements

References

[1] E. Angelopoulou, S. Lee, R. Bajcsy, Spectral gradient: a material descriptor invariant to geometry and incident illumination, in: The Seventh IEEE International Conference on Computer Vision, IEEE Computer Society, September 1999, pp. 861–867.

[2] A. Bovik, M. Clark, W. Geisler, Multichannel texture analysis using localized spatial filters, IEEE Trans. Pattern Anal. Machine Intell. 12 (1) (1990) 55–73.

[3] ITU-R Recommendation BT.709, Basic parameter values for the HDTV standard for the studio and for international programme exchange, Technical Report BT.709, ITU, Switzerland, 1990.

[4] G. Finlayson, Color in perspective, IEEE Trans. Pattern Anal. Machine Intell. 18 (10) (1996) 1034–1038.

[5] J. Gårding, T. Lindeberg, Direct computation of shape cues using scale-adapted spatial derivative operators, Internat. J. Comput. Vision 17 (2) (1996) 163–191.

[6] J. Geusebroek, R. van den Boomgaard, A. Smeulders, A. Dev, Color and scale: the spatial structure of color images, in: The Sixth Europian Conference on Computer Vision, Springer, Berlin, July 2000, pp. 331–341.

[7] J. Geusebroek, R. van den Boomgaard, A. Smeulders, H. Geerts, Color invariance, IEEE Trans. Pattern Anal. Machine Intell. 23 (12) (December 2001) 1338–1350.

[8] T. Gevers, A. Smeulders, Color based object recognition, Pattern Recognition 32 (3) (1999) 453–464.

[9] G. Golub, C.V. Loan, Matrix Computations, The Johns Hopkins Press, London, 1996.

[10] D. Hall, V. deVerdiere, J. Crowley, Object recognition using coloured receptive fields, in: The Sixth Europian Conference on Computer Vision, Springer, Berlin, July 2000, pp. 164–177.

[11] E. Hering, Outlines of a Theory of the Light Sense, Harvard University Press, Cambridge, MA, 1964.

[12] A. Jain, F. Farrokhnia, Unsupervised texture segmentation using gabor filters, Pattern Recognition 24 (12) (1991) 1167–1186.

[13] A. Jain, G. Healey, A multiscale representation including opponent color features for texture recognition, IEEE Trans. Image Processing 7 (1) (January 1998) 124–128.

[14] B. Julesz, Textons, the elements of texture perception and their interactions, Nature 290 (1981) 91–97.

[15] M. Kirby, F. Weisser, G. Dangelmayr, A model problem in the representation of digital image sequences, Pattern Recognition 26 (1) (1993) 63–73.

[16] J. Koenderink, The structure of images, Biol. Cybern. 50 (1984) 363–370.

[17] J. Koenderink, A. van Doorn, Generic neighborhood operators, IEEE Trans. Pattern Anal. Machine Intell. 14 (6) (1992) 597–605.

[18] B. Manjunath, W. Ma, Texture features for browsing and retrieval of image data, IEEE Trans. on Pattern Anal. Machine Intell. 18 (8) (1996) 837–842.

[19] M. Mirmehdi, M. Petrou, Segmentation of color textures, IEEE Trans. Pattern Anal. Machine Intell. 22 (2) (2000) 142–159.

[20] H. Nguyen, M. Worring, A. Dev, Detection of moving objects in video using a robust motion similarity measure, IEEE Trans. Image Processing 9 (1) (2000) 137–141.

[21] A. Smeulders, M. Worring, S. Santini, A. Gupta, R. Jain, Content-based image retrieval at the end of the early years, IEEE Trans. Pattern Anal. Machine Intell. 22 (12) (2001) 1349–1380.

[22] B. Thai, G. Healey, Modeling and classifying symmetries using a multiscale opponent color representation, IEEE Trans. Pattern Anal. Machine Intell. 20 (11) (1998) 1224–1235.

[23] S. Umbaugh, Computer Vision and Image Processing, Prentice Hall PTR, Upper Saddle River, NJ, 1998.

[24] T. Weldon, W. Higgins, D. Dunn, Efficient gabor-filter design for texture segmentation, Pattern Recognition 29 (12) (1996) 2005–2016.

[25] X. Zhang, B. Wandell, A spatial extension of cielab for digital color image reproduction, in: Proceedings of the Society for Information Display Symposium, 1996.