Scientia Iranica, Vol. 6, No. 1, pp 9–17

SCIENTIA
IRANICA

A New Smoothing Approach for Pattern-Color Separability

M. Mirmehdi* and M. Petrou[1]

In this paper, a set of masks is introduced which are used in combination for simulating pattern-color separability function of the human visual system. These kernels provide a perceptual degree of smoothing, corresponding to measurements estimated from human psychophysical experiments. It has been found that these kernels yield a better approach for smoothing color images than the traditional widely practiced Gaussian masks. The masks are applied to color texture segmentation and some results are presented. The masks are also useful for other vision tasks where smoothing is a major step.

INTRODUCTION

Segmentation of texture images is a major field of research in computer vision [1-3]. In this paper, the preprocessing smoothing stage of an image, critical for a good segmentation, not the segmentation itself, is considered. Image smoothing within a multiscale or multiresolution framework using Gaussian (or other) masks is generally practiced as a precursor to texture segmentation. This is outlined in some detail by Watt [4]. Malik and Perona [5] used Gaussian masks to construct a computational model of human texture perception. Vinken et al.[6] devised smoothed versions of an image by convolution of the original image with sampled Gaussian kernels of increasing width.

Many researchers have performed a preprocessing smoothing stage by simply generating a multiresolution pyramid representation of the image. In this regard, an important work for texture segmentation is [7], while many others can be found in [2,3,8]. Matalas et al. [9] used a B-spline function, as a fast transform, in order to obtain images at several smoothing levels and calculate vector dispersion and gradient orientation at different scales. A small disparity function was then applied to segment grey-level textures.

*. *Corresponding Author, Department of Computer Science, University of Bristol, Bristol BS8 1UB, England.*
1. *School of Electrical Engineering, Information Theory and Mathematics, University of Surrey, Guildford GU2 5XH, England.*

Other works, related to smoothing as a precursor to segmentation, are based on multi-channel filtering theory, using for instance Gabor filters. These filters take the form of a sinusoidal grating oriented in particular directions and modulated by a two-dimensional Gaussian. For example, Jain and Farrokhnia [10] presented a grey-level texture segmentation algorithm which used a bank of Gabor filters as its first stage to produce a bank of filtered images. More recently, Campbell et al. [11] examined a combination of Gabor filters and low-pass color filters for segmentation of natural scenes. Their system was implemented as a self-organising feature map trained on a large number of hand-segmented outdoor scenes. The Wigner distribution [12] is another Gaussian-based smoothing technique used to derive features for texture segmentation [13].

In other approaches, features are first derived from the image and then undergo Gaussian smoothing, e.g. [14-17]. For example, Huang et al. [16] performed Gaussian smoothing on the histograms of each band of a color image prior to applying Markov Random Fields (MRFs) for color image segmentation. Gaussian smoothing for scale space analysis of other features such as edges or junctions has also been reported [18,19].

Psychophysical and neurophysiological evidence indicate that the human visual system performs multiscale analysis [20,21]. Here, work by Zhang and Wandell [22] is extended and smoothing, in a way that imitates human vision, is investigated and compared with the traditional method of Gaussian smoothing for

color images. Then, the smoothing task is applied as a preprocessing step to a color segmentation technique. Finally, the results are compared with work by Ma and Manjunath [23].

PERCEPTUAL APPROACH

The human vision system is able to extract color textures as single entities without difficulty; color is a segmentation cue and an integral part of the preattentive vision. Also, image features, which characterize a texture at a certain resolution, may be entirely different from image features that characterize the same texture at another resolution. In fact, the criteria by which a segmentation is judged are subjective. In the absence of specific application requirements, it is expected that the segmentation of an image matches with that performed by the human vision system. Thus, here the interpretation of color texture perception is based on how texture can be perceived as separate homogeneous regions in the preattentive stage of vision. One piece of work with a close nature is that of Roan et al. [24] who described a method for classification of grey-level textured surfaces viewed at different resolutions, i.e., viewed at different scales or distances, while the image size remained constant. They used grey-level cooccurrence matrices and the Fourier power spectrum of an unknown texture image, taken at one of several resolutions (by physically increasing the distance between the sensor and the textures), to classify it as one of the six known textures. On the other hand, Zhang and Wandell [22] (hereafter ZW) systematically studied the color perception of human subjects for different frequencies of spatial color variation. Hence, they proposed an extension of the CIE-Lab space based on a spatially filtered pattern-color separable transformation to simulate the spatial blurring of the human visual system. This new color space, S-CIE-Lab, was then applied for measuring color reproduction errors of digital images.

They proposed that their filters should be applied to image data in the opponent color space. Their filters depend on the distance from which the smoothed scene is supposed to be viewed and are expressed as sums of different Gaussian kernels for the different opponent color planes. JPEG-DCT images, half-toned images, and some test patterns were employed to compare reproduction errors in both CIE-Lab and S-CIE-Lab spaces. The images for the S-CIE-Lab were smoothed using filters calculated for a user viewing a computer monitor from a fixed distance.

Here, the ZW kernels are extended to formulate a new set of masks and are used within a multi-scale segmentation framework for smoothing the image to a level where in the coarsest version, different textured areas become patches of almost uniform color.

The major question that arises here is whether one has to use the new masks, or any smoothing performed in the opponent color space would work equally well. This issue has been examined and a series of experiments using either the new masks or a Gaussian mask of the same size has been performed in both cases applying them in the opponent color space. The blurred images were used in a multi-level relaxation scheme [25,26] to obtain the final color segmentation.

PERCEPTUAL SMOOTHING

Recent experiments by Zhang and Wandell [22] demonstrated that the human eye perceives high spatial frequencies of color as a uniform color instead of separating them. An algorithm which takes this into account must smooth the image in luminance and chrominance color planes separately with different filter matrices for the planes. ZW advocated the use of the opponent color space, which consists of three different color planes, O_1, O_2, O_3, representing the luminance, the red-green and the blue-yellow planes, respectively. Each of these planes in the $O_1O_2O_3$ color space is smoothed separately with two-dimensional spatial kernels, defined as sums of Gaussian functions with different values of standard deviation σ. The result of this operation is that the luminance plane is blurred lightly, whereas the red-green and the blue-yellow planes are blurred more strongly. This spatial processing technique is pattern-color separable. ZW filtered representation was then transformed back to CIE-XYZ and then to CIE-Lab resulting in their Spatial CIE-Lab space, namely S-CIE-Lab. The CIE (Commission Internationale de L'Éclairage) Colorimetry system (including XYZ, Lab and Luv spaces) is described in [27].

In this application, smoothing is performed in the $O_1O_2O_3$ space using the new masks. Then the image data is transformed from the $O_1O_2O_3$ to the CIE-Luv space [28] and is used as the input in the ensuing segmentation steps. The reason for using the CIE-Luv space is that Euclidean distances between two points in this space correspond to a measure of the perceived color differences between them. This is not true in the RGB space, as it is not a perceptually uniform space.

IMPLEMENTATION

It has been assumed that the input data is available in the RGB format. The input RGB image can be transformed into the $O_1O_2O_3$ opponent color space in one step but, for clarity, this transformation is shown

via the XYZ space:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.431 & 0.342 & 0.178 \\ 0.222 & 0.707 & 0.0714 \\ 0.020 & 0.130 & 0.939 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} = \begin{pmatrix} 0.279 & 0.720 & -0.107 \\ -0.449 & 0.290 & -0.077 \\ 0.086 & -0.590 & 0.501 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \quad (2)$$

Once smoothing has taken place in the $O_1O_2O_3$ opponent color space, the data must be transformed back to the XYZ space before it can be transformed to the CIE-Luv space, in readiness for the segmentation process. The transformation from $O_1O_2O_3$ to XYZ is as follows:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.626 & -1.867 & -0.153 \\ 1.369 & 0.934 & 0.436 \\ 1.505 & 1.421 & 2.536 \end{pmatrix} \begin{pmatrix} O_1 \\ O_2 \\ O_3 \end{pmatrix} \quad (3)$$

The process of transforming the XYZ to CIE-Luv space, as described in [28], can be summarized as follows:

$$L = \begin{cases} 116(Y/Y_0)^{1/3} - 16 & \text{if } Y/Y_0 > 0.008856 \\ 903.3Y/Y_0 & \text{if } Y/Y_0 \leq 0.008856 \end{cases},$$

$$u = 13L\left(\frac{4X}{X + 15Y + 3Z} - u_0\right),$$

$$v = 13L\left(\frac{9Y}{X + 15Y + 3Z} - v_0\right) \quad (4)$$

where,

$$u_0 = \frac{4X_0}{X_0 + 15Y_0 + 3Z_0}, \quad v_0 = \frac{9Y_0}{X_0 + 15Y_0 + 3Z_0} \quad (5)$$

and (X_0, Y_0, Z_0) is the reference white. The L and the Y components are linked to the luminosity, while u, v and X, Z are chromatic components. For smoothing, after the RGB to $O_1O_2O_3$ transformation, three convolution matrices $(P_{lum}, P_{rg}, P_{by})$ are set up for the luminance plane, red-green and blue-yellow color planes. Each of these consists of a weighted sum of Gaussian kernels which achieve a different level of smoothing for each plane (as mentioned earlier). For example, the convolution matrix for P_{lum} as computed according to ZW is:

$$P_{lum} = \frac{1}{m} \sum_i \frac{w_i}{n_i} e^{-\frac{x^2+y^2}{\sigma_i^2}}, \quad (6)$$

for $(i = 1, 2, 3)$. For P_{rg} and P_{by}, the index i ranges from 1 to 2. The number of summations can be verified against Table 1. The values for (w_i, σ_i) which have been determined from psychovisual measurements of color appearance on human subjects are given in Table 1 [22]. Divisor n_i in Equation 6 is introduced to normalize the sum of the matrix elements of each individual Gaussian kernel before the weighted sum is applied. Divisor m normalizes the sum of the final matrix to 1.

In their subsequent coding error analysis in [22], Zhang and Wandell experimented for an output devise resolution of 90 dpi ($\sim 3 - 4$ pixels/mm) at a viewing distance of 18in for 1° of visual angle. The sizes of the convolution matrices are found for different distances as follows. Consider Figure 1. For a visual angle of $\theta°$ and a particular distance S away from the scene being viewed, a viewing area $\hat{x}$ units wide is obtained. Given the output device resolution of r pixels/mm, then the value of x (and similarly y) can be computed in pixels to be implemented in Equation 6, using $x = r\hat{x}$, i.e.:

$$x = rS \tan\frac{\theta\pi}{180}. \quad (7)$$

Some typical convolution kernels are illustrated in Figures 2a-c. For ease of reference, these are named the perceptual masks or kernels. For comparison, Figure 2d shows the corresponding Gaussian kernel, computed so that it has the same size as the new perceptual masks with σ being such that at the cut-off size its value is 1% of its central value.

Once the kernels are applied to the image in the opponent color space, the image can be converted to CIE-Luv which, as described earlier, is a perceptually uniform space and, therefore, more suitable for carrying out color measurements. Figure 3 demonstrates a real

Table 1. Values for weight and spread of the convolution kernels for a distance of 1m from the scene.

Plane	Weight	σ (degrees)	σ (pixels)
Luminance	0.921	0.0212	1.11
	0.105	0.0956	5.00
	-0.108	2.973	155.80
Red-Green	0.531	0.0291	1.52
	0.330	0.351	18.38
Blue-Yellow	0.488	0.0391	2.047
	0.371	0.274	14.35

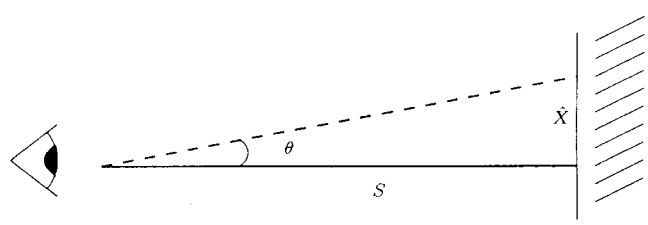
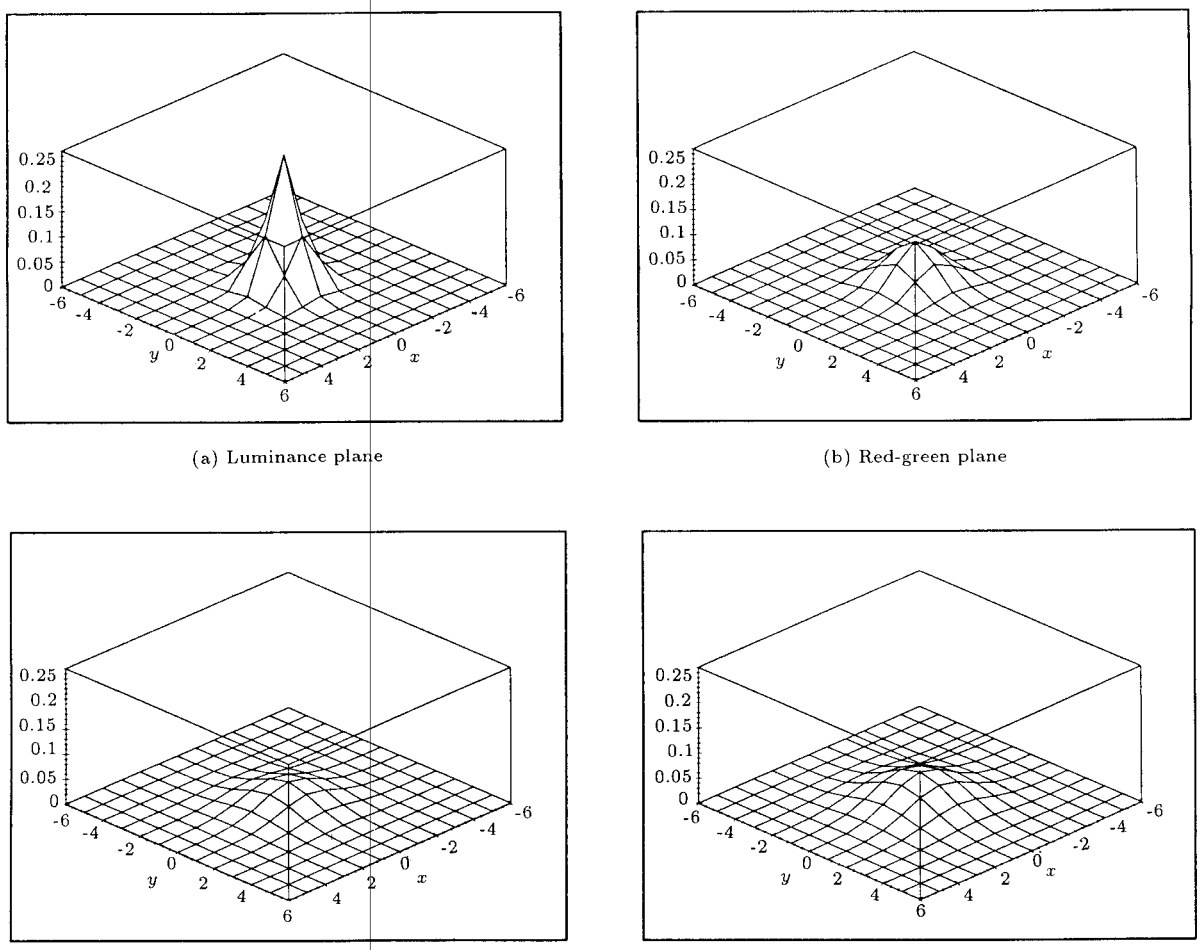


Figure 1. The visible width x for a particular distance S at $\theta°$of visual angle.



(a) Luminance plane

(b) Red-green plane

(c) Blue-yellow plane

(d) Gaussian

Figure 2. Convolution kernels for (a-c) perceptual smoothing and (d) Gaussian smoothing.

texture collage and its associated smoothed images at various distances for both perceptual and Gaussian smoothings. As the masks used are immense for large distances and in order to avoid an ever increasing border, the image is extended in two ways; it is assumed that the image is embedded in a much larger image of either black pixels, or white pixels. These two cases correspond to the physical situations where the viewed object is seen against a black or a white background, respectively. The overall color impression that results in the two cases, for both perceptual and Gaussian smoothing is very different as illustrated in Figure 3. However, the obtained segmentation results did not differ significantly for either backgrounds. Therefore, in all remaining experiments the option of padding the image with zeros is adopted to deal with the border effects. It is clear from Figure 3 that the perceptually smoothed images provide a more realistic representation and blurring of an "object" viewed at varying distances than the Gaussian. Most particularly, the Gaussian has mixed and smoothed the color values when convolved with each of the three color channels.

SEGMENTATION

The use of perceptual smoothing is advocated as a preprocessing step to any segmentation process that requires smoothed images as its input data. Thus, the segmentation process itself is not of concern here. However, to test and demonstrate the use of perceptual smoothing, a segmentation technique reported elsewhere [25,26] is now briefly mentioned. A multi-level relaxation scheme is developed [25,26] in which a mechanism of segmenting color textures is proposed, by constructing a causal multiscale tower of image versions based on perceptual smoothing of the color texture image. The reason it is called "tower" and not a "pyramid" was the abscence of a subsampling performance leading to the preservation of the same number of pixels at all levels. The levels of the tower were constructed with the help of different blurring masks by assuming that the same color-textured object was seen at 1,2,3,... meters distance. Hence, each coarser version of the image imitated the blurred version that the human vision system would have "seen" at the corresponding distance. The analysis of the image started at the

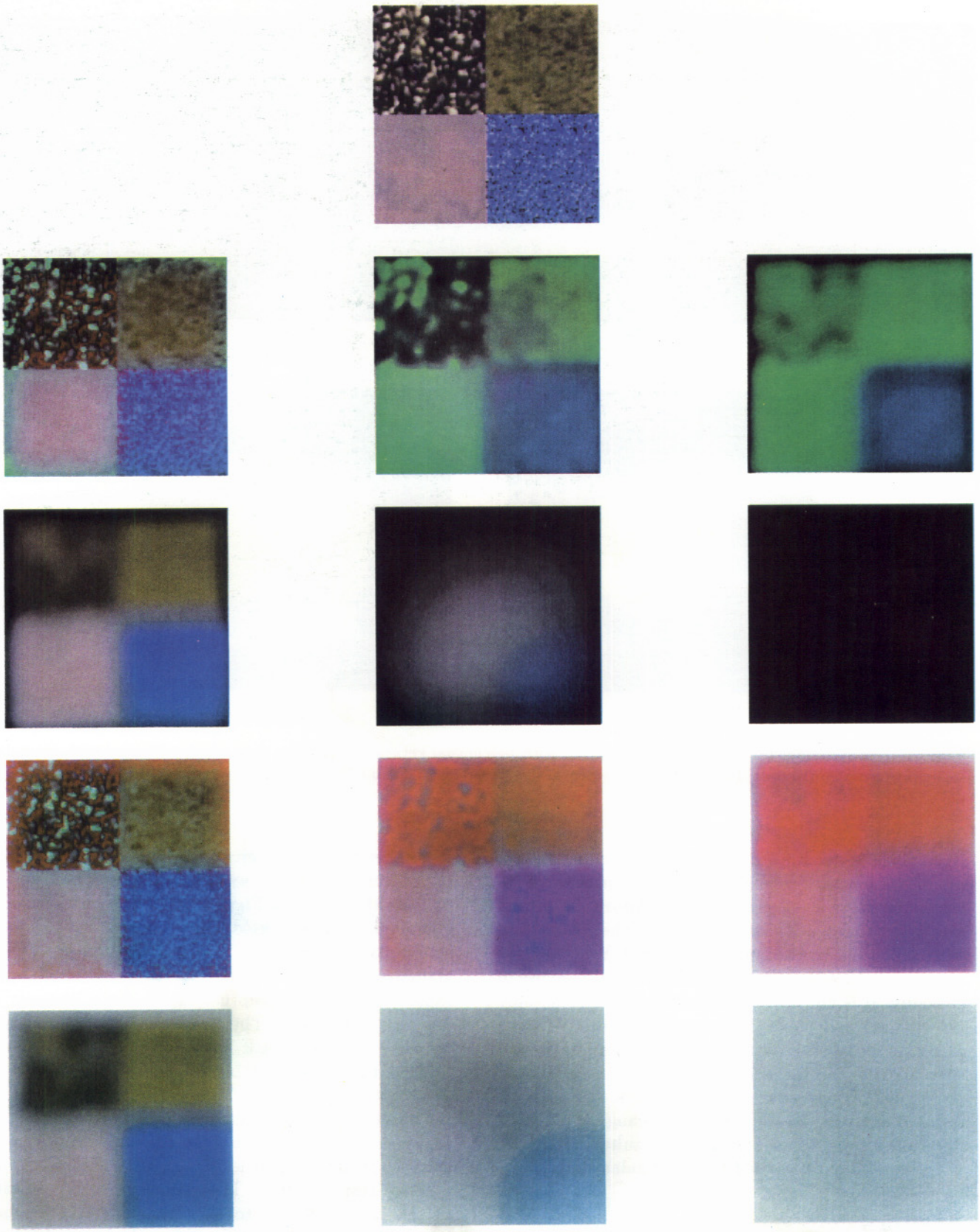


Figure 3. Convolution kernels for (a-c) perceptual smoothing and (d) Gaussian smoothing. Real texture collage image with (first row) perceptual and (second row) Gaussian smoothed transformations corresponding to viewing distances of 1, 5, and 10m against a "black background", and (third and fourth rows) as before but against a "white background".

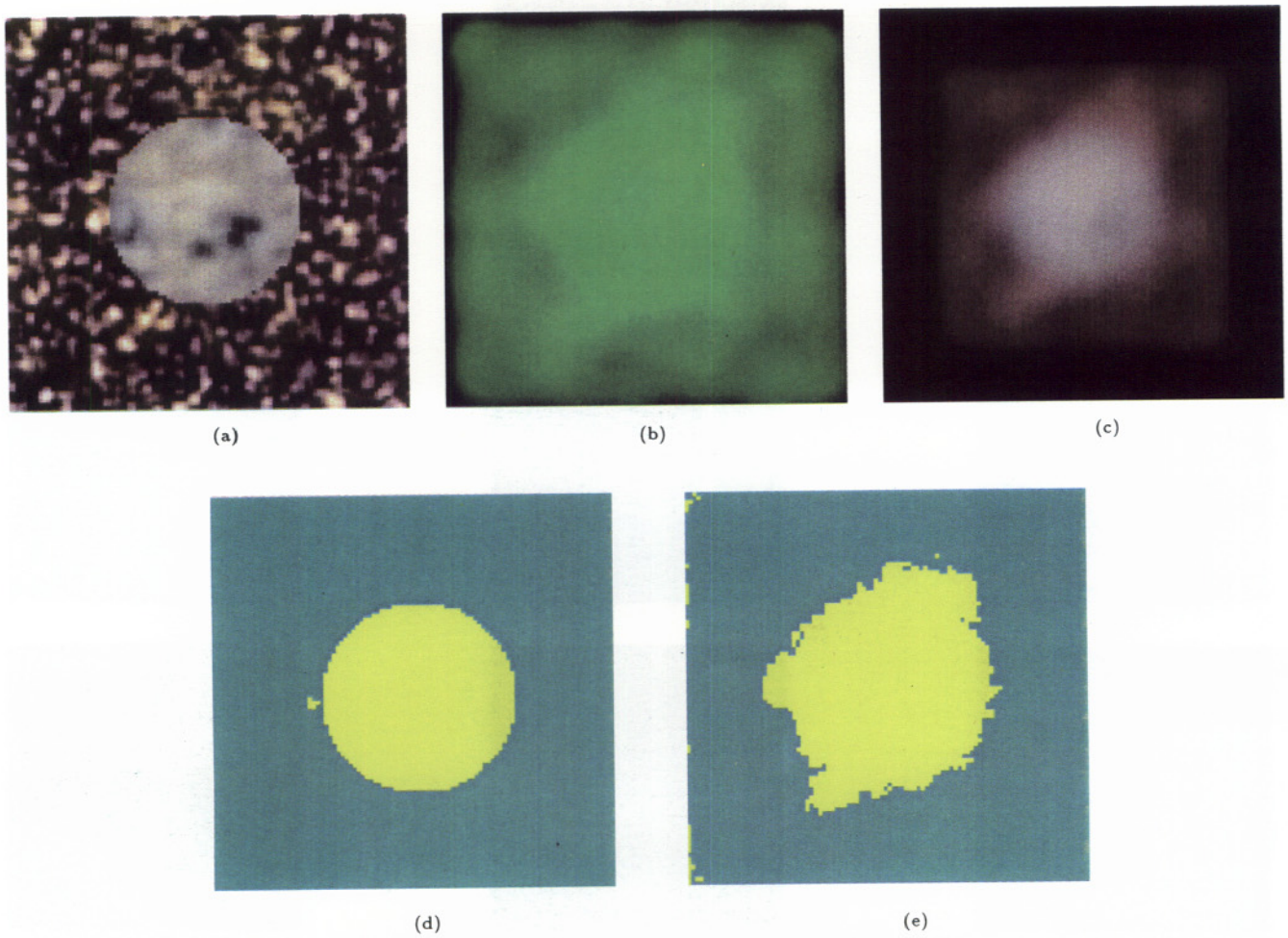
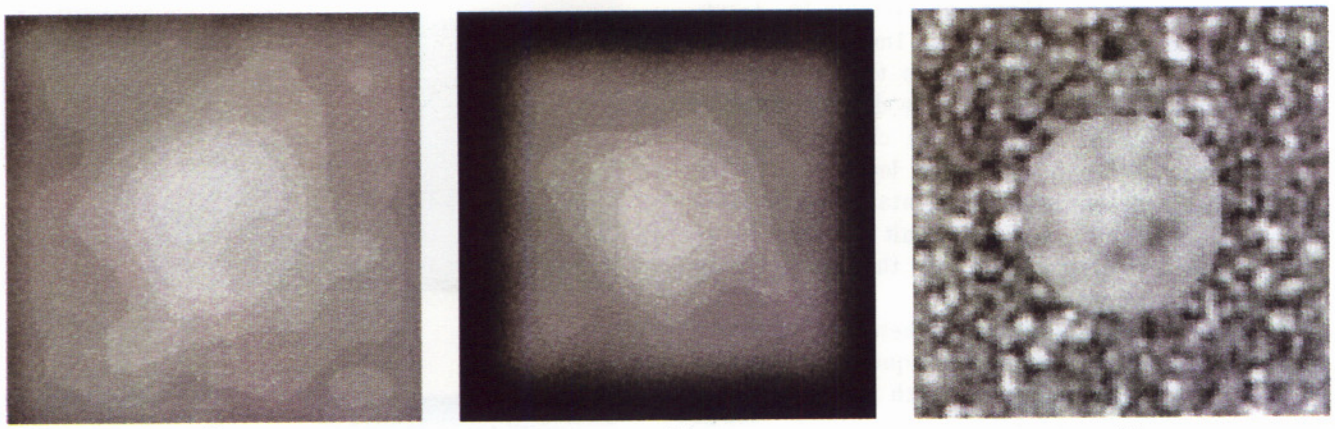


Figure 4. (a) Original texture image, (b) Blurred with perceptual kernels for viewing distance, (c) Blurred with Gaussian kernel for viewing distance, (d) Multilevel segmentation starting with image (b), (e) Multilevel segmentation starting with image (c).

coarsest level where the image is segmented into a number of clusters. The processing then proceeded towards the finest level, just as would happen if a person slowly approached a distant object. The mechanism with which information was transferred from a coarse to a finer level was probability theory that made use of causality. It is not advocated here that this is actually the mechanism deployed by humans; probability theory was used because it was a sound mathematical tool that allowed the incorporation of both features and preliminary conclusions that refer to many different levels of analysis. At the finest level image or the final iteration step, a segmented image resulted with each pixel labelled and assigned to a particular cluster.

RESULTS

Now, the perceptual smoothing approach is illustrated along with some smoothing and segmentation results. The original image in Figure 4a was created by patching together two different color textures. This image was blurred with the perceptual masks, assuming that the object it represents was viewed from 10m distance, to form the image in Figure 4b. If the corresponding Gaussian mask was used, all structure in the image would have been lost. Instead, in Figure 4c, the Gaussianly blurred image is shown, with a mask the same size as the perceptual mask, for a distance of 1m from the object. This was the blurring obtained by the Gaussian mask that seemed to correspond best to the perceptual blurring. Figures 4d and 4e demonstrate the results of segmenting the image within the same multiscale framework, with the parameters tuned in each case to produce the best results (each starting at the coarsest level, i.e., at configurations 4b and 4c, respectively). The difference in the quality of the results is due to the fundamental difference between the two blurring masks which can be best observed if the luminance of the two images is plotted. Figures 5a and 5b show only the L components of the images in Figures 4b and 4c. It can be seen that the perceptual masks smooth the luminance very little in

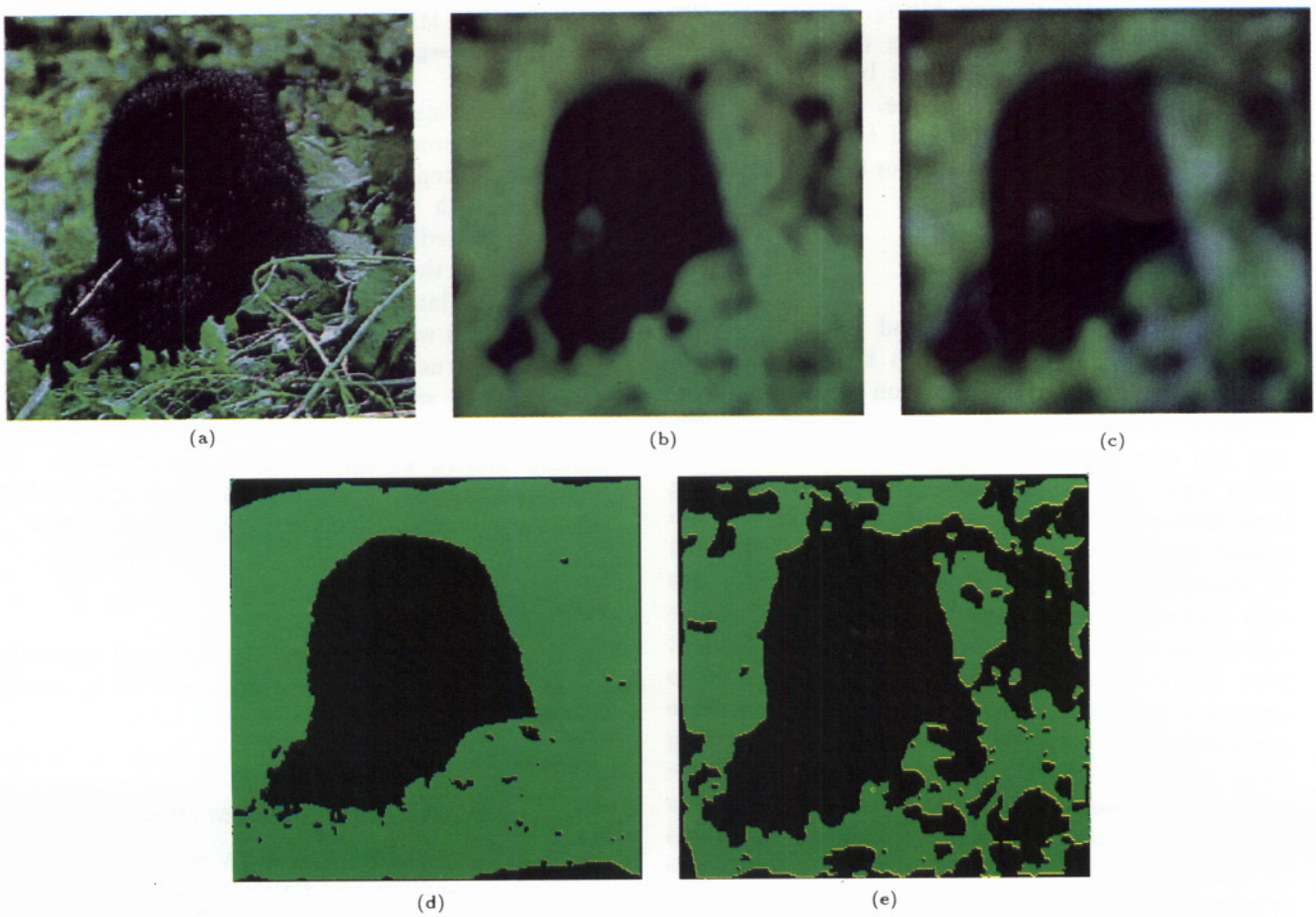


(a) (b) (c)

Figure 5. (a) Luminance band for 4(b), (b) Luminance band for 4(c), (c) Luminance band for perceptually smoothed image from a distance of 1m.

comparison to the Gaussian masks, with the desirable result of preserving the luminance edges, which is significant in the creation of a segmentation result that seems perceptually correct. For direct comparison with Figure 5b, the L component of the perceptual smoothed image at 1m distance is shown in Figure 5c.

In Figure 6, another example is illustrated. Figure 6b shows the perceptually smoothed image viewed



(a) (b) (c)

(d) (e)

Figure 6. (a) Original monkey image, (b) Blurred with perceptual kernels for viewing distance, (c) Blurred with Gaussian kernel for 1m viewing distance, (d) Multilevel segmentation starting with image (b), (e) Multilevel segmentation starting with image (c).

6-1

from a 10m distance. Figure 6c presents the Gaussianly blurred image viewed from a 1m distance which again seems to best correspond to the perceptually blurred masks. Figures 6d and 6e display the final segmentation results. It is quite clear again that the perceptually-based smoothing leads to a much more subjectively appealing segmentation, while with Gaussian smoothing, only the result from 1m is remotely close. At greater distances, the results do not correspond in any way.

Next, the results presented here are compared with two recent segmentation techniques, by Matas and Kittler [29] and Ma and Manjunath [23], referred to hereafter as MK95 and MM97, respectively. MK95 grouped color pixels by simultaneously taking into account both their feature space similarity and spatial coherence. Their approach uses color histograms and incorporates neighbourhood connectivity. MM97 combined color intensity and Gabor texture features (including magnitude and phase information) characterized by a predictive coding method to detect and separate color texture regions. Figure 7 shows a color texture image collage made using real images of ceramic tiles and granite stone, as well as the segmentation obtained using MK95, MM97 and the approach based on the smoothing process proposed here. It is subjectively clear that better results are obtained using the new technique. More precisely, the error rate against the ground truth is 1.3% for MK95, 9.7% for MM97 and 0.04% for the new proposed method.

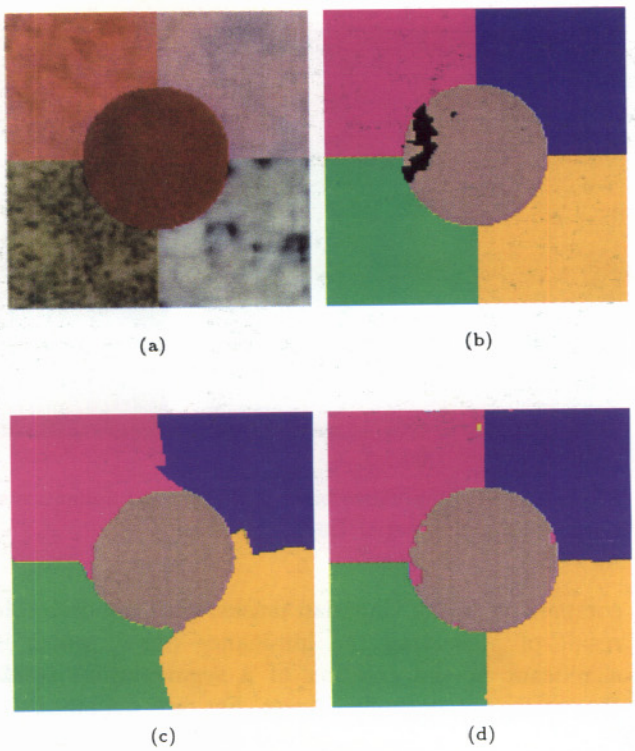


Figure 7. (a) Real texture collage and its segmentation by applying (b) MK95, (c) MM97 and (d) our complete smoothing and segmentation approach.

CONCLUSION

Recent experiments by Zhang and Wandell [22] for perceptual smoothing, applied to the evaluation of image coding schemes, are based on measurements of psychophysical studies which showed that discrimination and appearance of small-field or fine patterned colors differ from similar measurements made using large uniform fields. The human eye perceives high spatial frequencies of color as a uniform color instead of separating them. An algorithm which takes this into account must smooth the image in luminance and chrominance color planes separately with different filter matrices for each plane. This has been used to propose a perception based approach for color texture smoothing and segmentation. For the initial smoothing stage, masks for smoothing the image in a progressive manner have been developed to represent the viewing of the image from increasing distances. This led to the construction of a multi-scale tower of images which were inputs for the second stage of a texture segmentation process. This second stage is an iterative probabilistic relaxation-based pixel classification process [25,26]. The smoothing stage can act as a preprocessing step for any other technique which is based on the processing of coarsened images.

The perceptually derived blurring masks were compared with the classical Gaussian mask. The results presented in this paper clearly demonstrate the superiority of the former for color texture segmentation over the latter which is often used in multiscale schemes. This was verified using segmentation results based on the new proposed method when compared against other segmentation techniques. This new human-perception approach to smoothing is not necessarily limited to color texture segmentation. Its intrinsic pattern-color separability is also applicable to color edge detection and color region segmentation.

REFERENCES

1. Haralick, R.M. "Statistical and structural approaches to texture", *Proceedings of the IEEE*, **67**(5), pp 786-804 (1979).
2. Gool, L.V., Dewaele, P. and Oosterlinck, A. "Texture analysis anno 1983", *Computer Vision, Graphics and Image Processing*, **29**, pp 336-357 (1985).
3. Reed, T.R. and Du Buf, J. "A review of recent texture segmentation and feature extraction techniques", *CVGIP: Image Understanding*, **57**, pp 359-372 (1993).
4. Watt, R., *Understanding Vision*, Academic Press (1991).

5. Malik, J. and Perona, P. "A computational model of texture perception", Technical Report CSD-89-491, University of California Berkeley, CS (1989).
6. Vincken, K.L., Koster, A.S.E. and Viergever, M.A. "Probabilistic multiscale image segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **19**(2), pp 109-120 (1997).
7. Bouman, C. and Liu, B. "Multiple resolution segmentation of textured images", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **13**(2), pp 99-113 (1991).
8. Rosenfeld, A., *Multiresolution Image Processing and Analysis*, Springer (1984).
9. Matalas, I., Roberts, S. and Hatzakis, H. "A set of multiresolution texture features suitable for unsupervised image segmentation", in *Proceedings of Signal Processing VIII, Theories and Applications*, **III**, pp 1495-1498 (1996).
10. Jain, A.K. and Farrokhnia, F. "Unsupervised texture segmentation using gabor filters", *Pattern Recognition*, **24**(12), pp 167-1186 (1991).
11. Campbell, N.W. and Thomas, B.T. "Segmentation of natural images using self-organising feature maps", in *Proceedings of the 7th British Machine Vision Conference*, **I**, pp 222-232 (1996).
12. Reed, T.J. and Wechsler, H. "Segmentation of textured images and gestalt organisation using spatial/spatial-frequency represenations", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **12**(1), pp 1-11 (Jan. 1990).
13. Kittler, J., Marik, R., Mirmehdi, M., Petrou, M. and Song, J. "Detection of defects in color texture surfaces", in *IAPR Proc. of Machine Vision Applications 94*, pp 558-567 (Dec. 1994).
14. Unser, M. and Eden, M. "Multiresolution feature extraction and selection for texture segmentation", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **11**(7), pp 717-728 (July 1989).
15. Hsiao, J.Y. and Sawchuk, A. "Unspervised texture image segmentation using feature smoothing and probabilistic relaxation techniques", *Computer Vision, Graphics and Image Processing*, **48**, pp 1-21 (1989).
16. Huang, C.L., Cheng, T.Y. and Chen, C.C. "Color images' segmentation using scale space filter and markov random field", *Pattern Recognition*, **25**, pp 1217-1229 (1992).
17. Xie, Z.Y. and Brady, M. "Texture segmentation using local energy in wavelet scale space", in *ECCV96*, **I**, pp 304-313 (1996).
18. Perona, P. and Malik, J. "Scale space and edge detection using anisotropic diffusion", *IEEE Trans. Pattern Analysis and Machine Intelligence*, **12**(7), pp 629-639 (1990).
19. Lindeberg, T. "Discrete scale-space theory and the scale-space primal sketch", PhD Thesis, KTH (Royal Institute of Technology), Sweden (1991).
20. Hubel, D.H., *Eye, Brain, and Vision*, Scientific American Library (1988).
21. ter Haar Romeny, B.M. "Introduction to scale-space theory: Multiscale geometric image analysis", Technical Report ICU-96-21, Utrecht University, Holland (1996).
22. Zhang, X. and Wandell, B.A. "A spatial extension of CIELAB for digital color image reproduction", in *Society for Information Display Symposium, San Diego* (1996), WWW address: ftp://white.stanford.edu/scielab/spie97.ps.gz.
23. Ma, W.Y. and Manjunath, B.S. "Edge flow: A framework of boundary detection and image segmentation", in *Conference on Computer Vision and Pattern Recognition*, pp 744-749 (1997).
24. Roan, S.J., Aggarwal, J.K. and Martin, W.N. "Multiple resolution imagery and texture analysis", *Pattern Recognition*, **20**(1), pp 17-31 (1987).
25. Petrou, M., Mirmehdi, M. and Coors, M. "Segmentation of color textures", *Submitted to IEEE Transactions on Pattern Analysis and Machine Intelligence* (1997).
26. Petrou, M., Mirmehdi, M. and Coors, M. "Perceptual smoothing and segmentation of color textures", in *Proceedings of the 5th European Conference on Computer Vision 98*, Freiburg, Germany, pp 623-639 (1998).
27. Publication CIE No 15.2. in *Colorimetry*, Commission Internationale de L'Eclairage, Second Edition (1986).
28. Wyszecki, G. and Stiles, W.S., *Color Science: Concepts and Methods, Quantitative Data and Formulea*, John Wiley (1982).
29. Matas, J. and Kittler, J. "Spatial and feature based clustering: Applications in image analysis", in *CAIP95*, pp 162-173 (1995).