Correspondence

Histogram-Based Segmentation in a Perceptually Uniform Color Space

L. Shafarenko, M. Petrou, and J. Kittler

***Abstract*—In this work, we present a segmentation algorithm for color images that uses the watershed algorithm to segment either the two-dimensional (2-D) or the three-dimensional (3-D) color histogram of an image. For compliance with the way humans perceive color, this segmentation has to take place in a perceptually uniform color space like the Luv space. To avoid oversegmentation, the watershed algorithm has to be applied to a smoothed histogram.**

***Index Terms*—Adaptive filters, image color analysis, image segmentation, morphological operations, noise.**

I. Introduction

The spectral properties of the surfaces of objects play a very important role in their recognition and classification. In problems of surface industrial inspection and color grading, the spectral classes recognized in an image by a computer vision system have to correspond to chromatic classes perceived as distinct by the human vision system. For this purpose, the Luv color space is used, in which the Euclidean distance between two points is approximately proportional to the perceptual difference between the two colors represented by these points.

Using the Luv color space has, however, a drawback arising from the fact that the transformation from the RGB to the Luv space is highly nonlinear. Nonlinearity transforms the homogeneous noise in the RGB space to inhomogeneous noise. This means that even if we smooth the RGB data before the transformation, any small residual amount of noise may be significantly amplified by the nonlinear transformation, depending on the actual RGB values it refers to. The implication is that the Luv data are bound to contain nonuniform noise, which could be significant at places.

Although several people have reported work on color segmentation, in most cases the effects of noise and in particular nonhomogeneous noise are neglected. Quite often, the segmentation is performed in nonperceptually uniform color spaces, treating color as spectral information and not as a property of surfaces that is meaningful only with respect to the human vision system. For example, Healey [2] presented a color segmentation algorithm based on the physical properties of the reflected spectrum from various materials and assuming negligible sensor noise. Another example is the work of Huang *et al.* [3] who used a combination of scale space filters and projection of the three-dimensional (3-D) color histogram onto three one-dimensional (1-D) histograms to perform color segmentation. In order to compensate for the errors introduced by this projection, and in order to reduce oversegmentation, the segmented image was postprocessed by global optimization to force spatial consistency in the results. In their model, no sensor noise was considered, and color was not treated as a human sensor-based attribute.

In this correspondence, we present an algorithm for color segmentation that imitates human perception. To this end, an adaptive filter is used that effectively removes noise from a 3-D color histogram in the Luv color space, with subsequent perceptual coarsening. A color clustering method based on the morphological watershed transform [1] is then applied to the color histogram. The segmentation results are presented for a variety of color images.

Manuscript received July 23, 1996; revised October 14, 1997. This work was supported by the EPSRC under Grant GR/K44534. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. John Goutsias.

The authors are with the School of Electronic Engineering, Information Technology, and Mathematics, University of Surrey, Guildford GU2 5XH, U.K. (e-mail: m.petrou@ee.surrey.ac.uk).

Publisher Item Identifier S 1057-7149(98)06460-4.

II. Calculating the Covariance Matrix of the Noise in Luv

Let us assume that the noise probability density function in the RGB space is given by $f(R,G,B)$. In case of Gaussian noise, this function usually takes significant value near its maximum at (R_0,G_0,B_0) and inside a volume of linear size of the order of its standard deviation σ.

We express the expected value of another function $g(R,G,B)$ as

$$E\{g(R,G,B)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(R,G,B) f(R,G,B)\, dR\, dG\, dB. \tag{1}$$

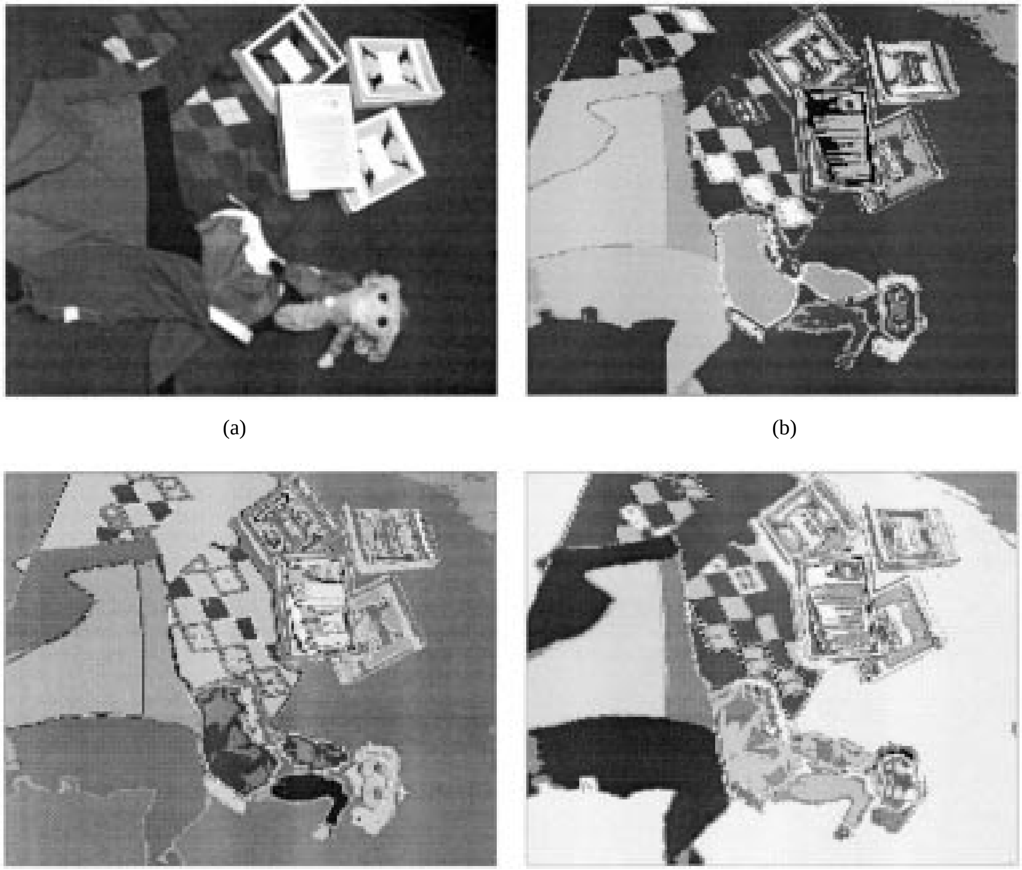
If inside the region of significant values of $f(R,G,B), g(R,G,B)$ is slowly varying, we can expand it about point (R_0,G_0,B_0) and replace it in the right-hand side of (1) by this series approximation:

$$\begin{aligned} g(R,G,B) \approx{}& g(R_0,G_0,B_0) + \frac{\partial g}{\partial R}(R-R_0) \\ &+ \frac{\partial g}{\partial G}(G-G_0) + \frac{\partial g}{\partial B}(B-B_0) \\ &+ \frac{1}{2}\left(\frac{\partial^2 g}{\partial R^2}(R-R_0)^2 + \frac{\partial^2 g}{\partial G^2}(G-G_0)^2 \right.\\ &\left. + \frac{\partial^2 g}{\partial B^2}(B-B_0)^2\right) \\ &+ \frac{\partial^2 g}{\partial R \partial G}(R-R_0)(G-G_0) \\ &+ \frac{\partial^2 g}{\partial R \partial B}(R-R_0)(B-B_0) \\ &+ \frac{\partial^2 g}{\partial B \partial G}(B-B_0)(G-G_0) + \cdots. \end{aligned}$$

Note that all the derivatives of g that appear in the expansion take constant values, as they are calculated at point (R_0,G_0,B_0). It is also assumed that $f(R,G,B)$ is an even function with respect to each of its arguments and that the range of integration over each of the variables is symmetric. This means that all integrals involving odd powers will vanish. Therefore the first nonvanishing term in the approximate calculation of (1) is

$$\begin{aligned} E\{g(R,G,B)\} &= g(R_0,G_0,B_0) \\ &\quad \cdot \left(1 + \frac{1}{2}\, \frac{\sigma_R^2 g''_{RR} + \sigma_G^2 g''_{GG} + \sigma_B^2 g''_{BB}}{g(R_0,G_0,B_0)}\right). \end{aligned} \tag{2}$$

In this expression, double primes indicate second derivatives with respect to the subscripts calculated at point (R_0,G_0,B_0) and σ_R^2, and σ_G^2 and σ_B^2 are the variances of the noise distribution along the $R, G,$ and B axes, respectively.



(a) (b) (c) (d)

Fig. 1. (a) Original image. (b) Segmentation using both noise filtering and perceptual coarsening. (c) Segmentation without perceptual coarsening (82 clusters). (d) Segmentation without noise filtering.

It has been established experimentally [4] that not only the noise in each of the three channels $R, G,$ and B is statistically independent from the noise in the other channels, but also that it is Gaussian with zero mean and diagonal covariance matrix with approximately the same standard deviations σ in all bands. Using the above information, we can apply (1) to find the approximate expectation value of any function of $R, G,$ and B. We denote the calculated expectation values by the same name as the function with an over-line. For $\overline{L}$, therefore, we find

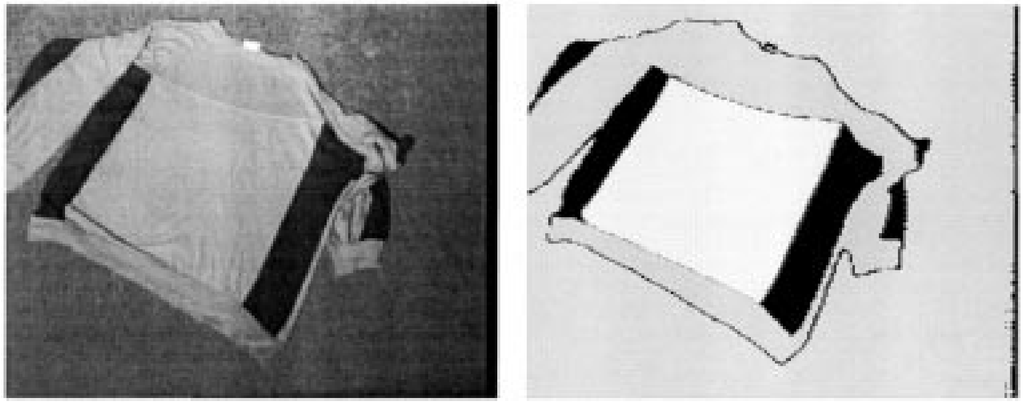
$$\overline{L} = \frac{1}{\sqrt{8\pi^3}\sigma^3} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} L(R,G,B) \cdot e^{-((R-R_0)^2+(G-G_0)^2+(B-B_0)^2/2\sigma^2)}\, dR\, dG\, dB$$

with similar formulas holding for $\overline{u}, \overline{v}$, as well as for $\overline{L-\overline{L}}, \overline{u-\overline{u}}$ and $\overline{v-\overline{v}}$.

This way, we can calculate the elements of the covariance matrix $\boldsymbol{C}$ of the noise distribution in the Luv space as (3), shown at the bottom of the page. This covariance matrix is different for different parts of the color space and thus leads to an adaptive filter appropriate for the noise reduction in the 3-D color histogram in the Luv space before any clustering takes place.

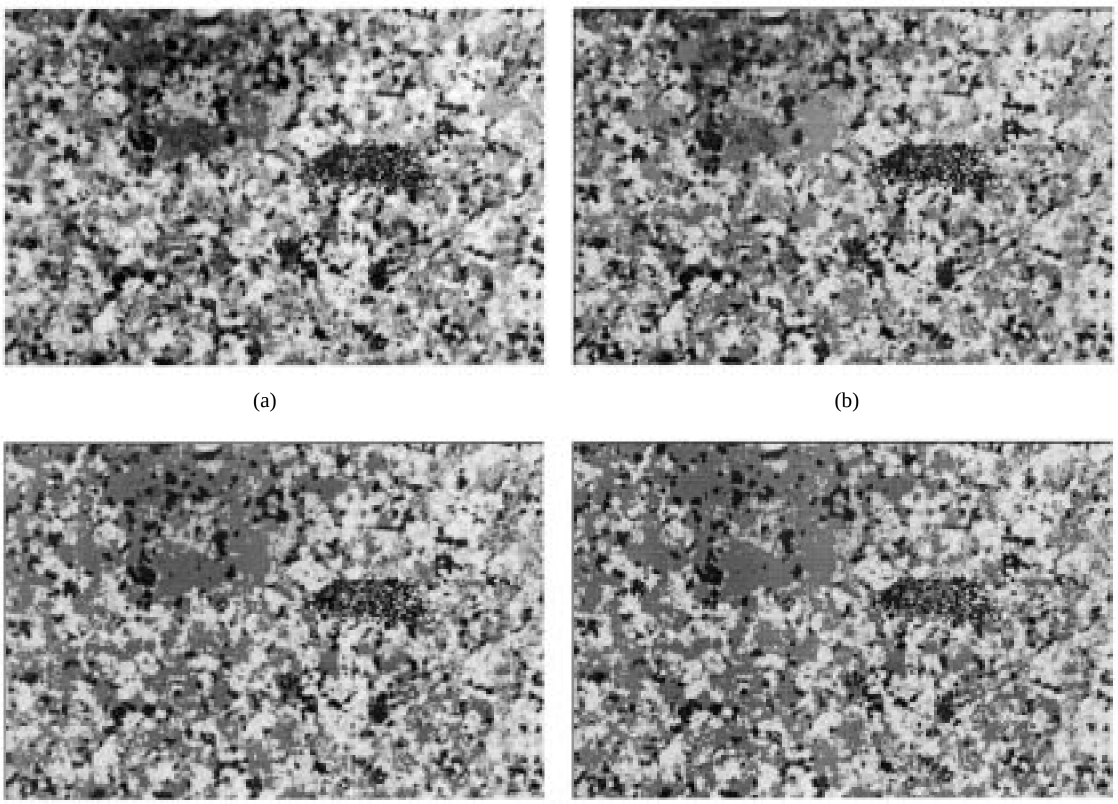
Once we have estimated the probability density function of the noise, we must use it to improve the color histogram of the image. Noise, however, is a stochastic process, and it is not possible to reverse its effect, unless we assume a certain prior model for the signal we are trying to recover. For example, in image restoration, noise can be removed by the method of simulated annealing, provided that a *prior* model for the signal has been adopted. In our case, no such prior model is available, as our signal is the color histogram of the image, with unknown number of clusters and unknown cluster profiles. Even in the simplest of cases, when the noise is additive and uniform, it is well known that the histogram of a grey image is *convolved* with the histogram of the noise field. No deconvolution can be attempted as that would be equivalent to subtracting two random fields, which is known to double the noise power. Instead, we view each histogram value as a measurement that carries with it a certain degree of uncertainty. That is, each measurement we are having, could have arisen with varied degrees of probability from a whole range of possible true measurements. As we cannot possibly know from which of these measurements exactly it arose, all we can do is to replace this measurement (which can be thought of as a delta function in the absence of any noise) by the finite width Gaussian. The width of this Gaussian is measurement dependent. Our histogram then looks like a

$$\boldsymbol{C} = \sigma^2 \begin{pmatrix} (L'_R)^2 + (L'_G)^2 + (L'_B)^2 & L'_R u'_R + L'_G u'_G + L'_B u'_B & L'_R v'_R + L'_G v'_G + L'_B v'_B \\ L'_R u'_R + L'_G u'_G + L'_B u'_B & (u'_R)^2 + (u'_G)^2 + (u'_B)^2 & u'_R v'_R + u'_G u'_G + u'_B v'_B \\ L'_R v'_R + L'_G v'_G + L'_B v'_B & u'_R v'_R + u'_G u'_G + u'_B v'_B & (v'_R)^2 + (v'_G)^2 + (v'_B)^2 \end{pmatrix} \tag{3}$$



(a) (b)

Fig. 2. Chromaticity-based segmentation. (a) Original image. (b) Segmentation results.



(a) (b)

(c) (d)

Fig. 3. Three-dimensional segmentation of a granite image. (a) Original image. (b) Segmentation with $\rho = 1.5$ (nine clusters). (c) Segmentation with $\rho = 2.$ (seven clusters) (d) Segmentation with $\rho = 4.$ (six clusters). $\sigma = 3.6$ for all cases.

whole lot of shifted and overlapping Gaussians and to cope with the uncertainty they manifest, we must integrate them out. In other words, we must convolve the histogram with this variable width Gaussian.

III. Histogram Preprocessing

Prior to clustering we need to "smooth out" the color histogram in order to get rid of the noise. This is achieved by applying filtering as described above. In addition to this, another step is required before clustering could take place. Namely, the color histogram needs to be "coarsened" to correspond to human perception; in other words, we do not want to distinguish between clusters that the human eye is not able to recognize as different, even though they might be separated by the algorithm. Thus, we smooth the color histogram using a spherical window of radius ρ. The exact value of ρ chosen depends on the application we are interested in.

As an illustration that both noise filtering and perceptual coarsening are essential for good segmentation, we present Fig. 1, which shows the output with either of these important stages being omitted. One can observe that the quality of the results is very poor in terms of oversegmentation, when compared with the results obtained when both preprocessing stages are used.

IV. Clustering

We are aiming at automatic clustering, where all information should be extracted from the image itself. We are therefore restricted in our choice of the clustering technique. For the purpose, therefore, of

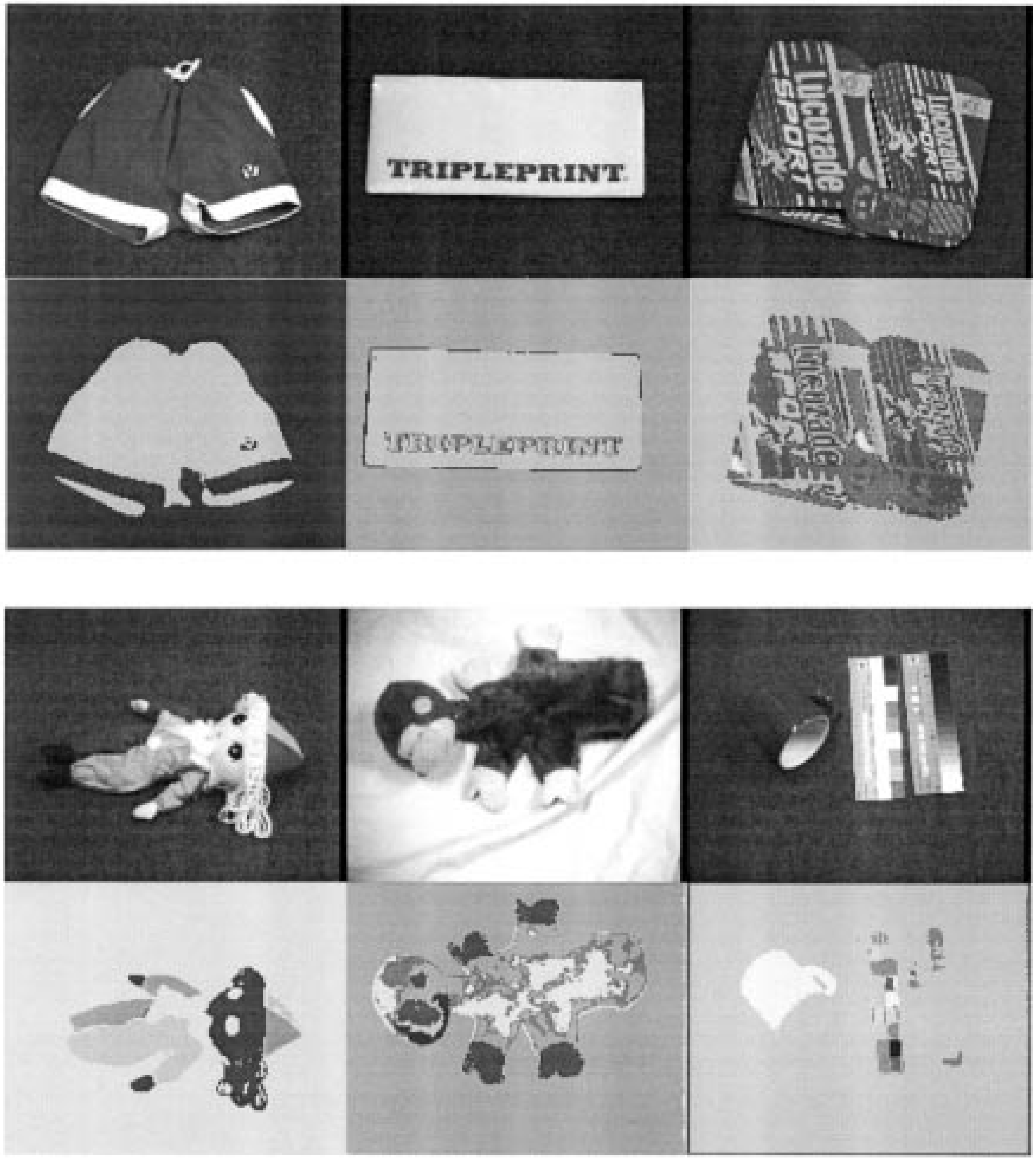


Fig. 4. Segmentation on a variety of images using a fixed set of parameters for the algorithm.

identifying the valleys of the color histogram, we used a well-known morphological algorithm, the watershed transform. We shall give its brief description now; for details, the reader is referred to extensive literature on the subject (e.g., [1]).

The idea of watershed is drawn from a topographic analogy. Consider a 2-D histogram of features as a topographic relief. Find all local minima and "pierce" them. Immerse the whole relief into water. As the relief goes deeper into the water, the regions surrounding the seeds become flooded. Eventually two or more such regions expand to a point at which they would come into contact unless the waters are separated. This is the moment that a dam is raised. In the watershed method, the dams are all infinitely tall and are arbitrary complex sets of pixels depending on the line of contact. This is a very informal definition though, since in the situation of discrete altitude of the relief, which is laid out on a discrete grid, there is no way of gradually bringing the flooding water up to the point of contact. However, it helps to visualize the procedure.

Technically, for every bin with an identifier, the neighboring bins are checked on being "under water," and if any of them are, they receive the same identifier, provided that they have not been identified with a different flood area already. We assume four-connectivity for the 2-D grid and 18-connectivity for the 3-D grid (used for clustering the 3-D color histogram).

V. Segmentation

We are addressing two different segmentation procedures, the first being chromatically based, and the second taking into consideration both chromaticity and intensity of the image. In both of them, we use the watershed algorithm to segment the *color histogram* of the image.

A. Segmentation Based on Chromaticity Alone

The need for this type of segmentation arises when we are presented with a problem of segmenting the image according to color information alone, ignoring the intensity. A good example of such a task would be segmentation of the image in Fig. 2, where the creases of the fabric and the shadows due to illumination changes across the scene should not prevent us from segmenting out the region in question as having uniform color. Having adopted this approach, we should not expect to be able to distinguish between points in color space that differ only in intensity.

It might appear that for chromaticity-based segmentation, we should consider only a 2-D color histogram, summing up votes for all intensities occurring at each point of the chromaticity plane (which is the uv plane in the case of the Luv color space). This is not the case, however, due to the fact that noise "mixes up" color coordinates, that is, each point in color space contributes into several points on the chromaticity plane, rather than solely into its projection onto the chromaticity plane. The noise filtering we propose is essentially 3-D, and summation over intensity should be done only after such filtering has been performed. The situation is different for the operation of perceptual coarsening which is commutative with the operation of summation over intensity and could therefore be performed in the chromaticity plane rather than in the 3-D color space.

Thus, the algorithm for chromaticity-based segmentation is as follows.

- Calculate the color histogram of the image.
- Project it onto the chromaticity plane by summing over the intensity coordinate.
- Perform perceptual coarsening.
- Perform clustering using the watershed algorithm in two dimensions.

The segmentation results using this algorithm are shown in Fig. 2 and also in Fig. 1(b).

B. Segmentation Based on Both Intensity and Chromaticity

A very different type of segmentation occurs when we are facing the problem of segmenting the image "as seen," that is, when it is necessary to recognize as different those colors that differ in their luminance value. In this case, we have the ability to distinguish between black and white, but we lose the tolerance to shades and creases. An example of a thus posed segmentation problem is that of segmenting a typical granite image as presented in Fig. 3.

In this case, we have to do clustering on the 3-D color histogram, and the algorithm is as follows:

- calculate the color histogram of the image;
- filter it for noise reduction;
- perform perceptual coarsening;
- perform clustering using the watershed algorithm in the 3-D Luv space.

Segmentation results using this algorithm with several different values of parameters are shown in Fig. 3. To give the reader an idea of the algorithm performance, color values for each pixel were replaced by those of the cluster the pixel belongs to. The color values of the cluster were calculated as its mean $L, u,$ and v. Note that this was done to visualize the result rather than for any further use of the image. If such use is intended (for compression purposes, for example) the ways to represent the cluster should be researched, which is beyond the scope of this work.

VI. Conclusions

A new algorithm is proposed for segmentation of color images, which takes into account the noise that is inevitably present during the image acquisition. Such noise affects human perception of the image due to the nonlinear nature of the human perception. This leads to a situation when even a low absolute value of the noise is noticeable to the human eye in certain areas of the color space.

The Luv color space was used for perceptual coarsening of the color histogram, as well as for the resolution gain it can offer compared to the RGB space.

The clustering method was based on the morphological watershed transform performed on the 3-D color histogram.

The resulting algorithm is highly suitable for automatic color segmentation. Indeed, there are only two parameters involved: the width of the noise distribution and the size of the window for perceptual coarsening. The former describes the hardware setup, while the latter reflects the desired degree of coarseness in the segmentation. When a combination of these two parameters is found that results in a good segmentation, the algorithm performs well for a wide range of images acquired using the same hardware. To illustrate this, we present segmentation results obtained with these parameters set constant. The results on a variety of images are shown in Fig. 4.

References

[1] S. Beucher and F. Meyer, "The morphological approach to segmentation: The watershed transformation," in *Mathematical Morphology in Image Processing*. New York: Marcel Dekker, 1993, pp. 443–481.

[2] G. Healey, "Segmenting images using normalized color," *IEEE Trans. Syst., Man, Cybern.*, vol. 22, pp. 64–73, 1992.

[3] C. L. Huang, T. Y. Cheng, and C. C. Chen, "Color image segmentation using scale space filter and Markov random fields," *Pattern Recognit.*, vol. 25, pp. 1217–1229, 1992.

[4] L. Shafarenko, "Perception-driven automatic segmentation of color images using mathematical morphology," Ph.D. dissertation, Univ. Surrey, Guildford, U.K., 1996.