LOCAL COLOR IMAGE SEGMENTATION USING SINGULAR VALUE DECOMPOSITION

Clifton B. Phillips and Ramesh C. Jain

Visual Computing Laboratory - University of California, San Diego

ABSTRACT

A method was developed to segment images of complex scenes based on color content. The output of an interest operator provides focus toward regions within an image to be sampled for color content. Statistics for each data set sampled are used to cluster and estimate bounded regions within a transformed color space. Each region respectively represents a specific set. Mappings to transformed regions of color space is found using the singular value decomposition. The mean and variance of each color sample in the transformed color space represent characteristic features for their sampled set of points. Color segmentation is accomplished by establishing whether image pixels belong to any subset represented by the characteristic features. This work contributes a method to color-segment targets in images using local color information within an image stream.

I. INTRODUCTION

Segmentation is the first step to extracting information from images. It is the partitioning of images into smaller regions having some common characteristic within each region. Kanade [9] states that it is desirable to accomplish segmentation in such a way to extract semantically meaningful objects or parts of objects and provides an overview of the segmentation problem from the framework of systematically extracting information from images in a hierarchical structure. Several researchers have investigated how to accomplish segmentation using color [2,3,10,11,14]. Two groups appear within the paradigm of color segmentation. The first uses global techniques such as the Karhunen-Loeve (K-L) transform or the scale-space filter scheme to get information for cluster-centering in an appropriate color space. It is well known that the K-L transform allows an image to be reduced to its principal components for coding and compression purposes. The same property is also useful for feature extraction. The idea is that the larger eigenvalues from the K-L transform relate respectively to dominant characteristics within an image. The scale-space filter method [12] analyzes histogram data to determine the best peaks. In most segmentation methodologies, the trick is to get the characteristic feature information from decomposed color components converted to spatial locations for "cluster zones" in the color space of choice. The second group uses local information to predict what value a certain site has. Such methods include Markov Random Fields (MRF's) and iterated conditional modes (ICM).

One difference in the two methodologies mentioned is that the former is applied globally, while the latter is applied locally. The main disadvantage of methods utilizing local information is that cluster information must be known a-priori. This work applies the eigenvalue-based concept locally. An interesting target is first found, then it is sampled for its color content. The covariance color matrix is determined based on the sampled data. A singular value decomposition (SVD) is applied to the color-covariance-matrix resulting in a natural normal transformation to a new color space aligned with the principal components of the data. The variance of the data in the transformed color space are the singular values provided by the SVD operator. In turn the statistics of the sample can be used as basis for color segmentation. Issues such as number of partitions and their respective sizes can be solved locally on a target of interest instead of processing an entire image.

Section II provides an overview of the SVD. Section III discusses the acquisition of an image sample and the formation of a cluster in a transformed space using the SVD. Section IV relates the points bounded by the transformed space to the traditional outcomes of a hypothesis test. Section V summarizes the procedure discussed in this paper. Section VI discusses the experiments and results from applying this method to several color spaces. Some justification is provided for the desirability of working in each color space used in this work. Results are presented from the National Television Systems Committee (NTSC) receiver primary red-green-blue (RGB) system, the [†]CIE XYZ and xyY system, and hue-saturation-intensity (HIS) system of color spaces. Section VII discusses the conclusions and restrictions for this method.

II. DECOMPOSING A MATRIX USING THE SVD

The following arguments are stated without complete and rigorous proof. Details about the singular

[†] Commission Internationale de L'Eclairage - The International Committee on Color Standards.

0-7803-4876-1/98/$10.00 © 1998 IEEE

value decomposition can be found in [1,4,12,13]. The singular value decomposition can be applied to factor the $M \times N$ matrix A into

$$A = U\Sigma V^t \qquad (1)$$

where U and V are respectively $M \times M$ and $N \times N$ orthogonal matrices. Σ is $M \times N$ having non-zero entries only on the main diagonal at locations ii from $i = 1$ through $i = rank(A)$. These non-zero diagonal entries are organized in descending order and called singular values. In general, the SVD can be applied to decompose non-square matrices. The method discussed in this paper uses the SVD to extract a transformation U^t that maps the data to a basis aligned with its principal components.

III. ACQUISITION OF CHARACTERISTICS

A sample can be obtained manually, or by an autonomous interest operator. Consider the goal of tracking an object as it moves through a 3-dimensional space. A simple interest operator can be formed by considering the differences between a three image sequence I_{k-2} through I_k. Let the sequence S_k represent the absolute value of pixels in an image sequence such that $S_{k-1} = |I_{k-1} - I_{k-2}|$ and $S_k = |I_k - I_{k-1}|$ represent returns where changes have occurred in the scene of interest. The interest operator provides output $S_k \cap S_{k-1}$ and is further conditioned by a threshold operator to generate a binary mask. The mask is used to locate the color sample from the original image.

The color analysis method discussed here can be applied to the color space of choice, although selecting an appropriate color space for the problem will enhance the quality of the result. A sample may contain a variety of colors. If the colors are too far apart in the selected color space, it is necessary to partition the sampled colors into appropriately sized clusters. Hartigan [7] provides several strategies for the clustering problem and provides advantages and disadvantages for each. This work applied a partitioning scheme to each sample based on limiting the variance of any classification. If the variance exceeds a user defined threshold, a partition is created and the data is split into two classes along the hyper-plane orthogonal to the direction of maximum variance. Each class is reiterated in the algorithm until all the components of the variance are within the defined threshold.

Consider extracting a sample set of N RGB points from a target, where the RGB color space is defined based on the NTSC receiver primary system definition. Each pixel sampled can be represented at a coordinate in RGB color space. The sampled mean for each set in RGB is determined as

$$\hat{x} = \left[\hat{r} \;\; \hat{g} \;\; \hat{b}\right]^t = \frac{1}{N}\left[\sum_{i=1}^{N} r_i \;\; \sum_{i=1}^{N} g_i \;\; \sum_{i=1}^{N} b_i\right]^t. \qquad (2)$$

First we consider applying this technique to one set of sampled data, then the result shown can be extended to multiple sets where a sample has been partitioned into multiple sub-clusters. The variance can be used to establish a best ellipsoid in RGB space that contains the desired set of RGB pixels. Before any variance can be computed, it is necessary to carry out a transformation to a new basis centered about the sampled mean. The new basis is to be lined up with an estimate of the best principal axis of the sampled data [5]. The transformation extracted from the SVD guarantees the estimated principal component axes are orthogonal. Therefore the statistics are de-coupled when estimated along the principal component axes. The singular value decomposition is useful to establish such a natural orthogonal basis for a sampled set of data. Each RGB point consists of three values corresponding to the red, green, and blue components of color space. Let the set of N RGB points be formed into a matrix A, where A is $M \times N$, with $M = 3$, and $N \geq M$. Let C be the covariance matrix for all the sampled RGB data points represented as the columns of A. For each sampled set of data, the entries for the covariance matrix C are given as

$$C = \begin{bmatrix} C_{rr} & C_{gr} & C_{br} \\ C_{rg} & C_{gg} & C_{bg} \\ C_{rb} & C_{gb} & C_{bb} \end{bmatrix},$$

where the matrix covariance entries are given by:

$$C_{rr} = \frac{1}{N}\sum_{i=1}^{N}(r_i - \hat{r})^2, \qquad C_{rg} = C_{gr} = \frac{1}{N}\sum_{i=1}^{N}(r_i g_i) - \hat{r}\hat{g},$$

$$C_{gg} = \frac{1}{N}\sum_{i=1}^{N}(g_i - \hat{g})^2, \qquad C_{rb} = C_{br} = \frac{1}{N}\sum_{i=1}^{N}(r_i b_i) - \hat{r}\hat{b},$$

$$C_{bb} = \frac{1}{N}\sum_{i=1}^{N}(b_i - \hat{b})^2, \qquad C_{bg} = C_{gb} = \frac{1}{N}\sum_{i=1}^{N}(b_i g_i) - \hat{b}\hat{g}.$$

It is interesting to note that U is the same, whether the SVD is performed on the data itself or upon the covariance matrix. The importance is that the columns of U span the space of A and are the desired orthogonal basis vectors corresponding to the principal values for

the transformed ellipsoid [5]. Consider translating the RGB axis by $T = \hat{x}$ such that the set of sampled data points contained in A is transformed to a zero-mean set along its principal components by equation (3) where the transformation U_i is found from implementing the SVD on set i. A_j and B_j are $M \times 1$ column components of matrices A and B, with $M = 3$.

$$B_j = U_i^t (A_j - T_i) \tag{3}$$

The matrix B is the range space and contains the transformed vectors $[\rho \quad \gamma \quad \beta]^t$. Here the sampled data is assumed to be normally distributed. The transformed data contained in B is aligned with the principal axes and the variance of the data can be found from equation (4). It is easy to show the result of equation (4) can be obtained directly by applying the SVD to the covariance matrix C instead of the data itself. As previously stated, the transformation U will be identical, however the singular values obtained from implementing this method on the matrix C are the variances $\sigma_\rho^2, \sigma_\gamma^2, \sigma_\beta^2$. These variances represent the eigenvalues of the covariance matrix C.

$$\sigma_\rho^2 = \frac{\sum_{k=1}^{N} (\sigma_{\rho_k} - \hat{\sigma}_\rho)^2}{N-1}$$
$$\sigma_\gamma^2 = \frac{\sum_{k=1}^{N} (\sigma_{\gamma_k} - \hat{\sigma}_\gamma)^2}{N-1} \tag{4}$$
$$\sigma_\beta^2 = \frac{\sum_{k=1}^{N} (\sigma_{\beta_k} - \hat{\sigma}_\beta)^2}{N-1}$$

IV. ESTABLISHING PROBABLE POINT SETS

Each of the three values of the variance determined by (4) represents the intersection of the one-sigma ellipsoid with either the semi-major or one of the semi-minor axes. The ellipsoid was selected as the bounding geometry because the location of its bounding surface is well understood when the data is normally distributed. Using these one-sigma values as characteristic parameters describing the boundary of a set of interest means that 63% of the sampled data falls within the set. Scaling the sigma values by two or three respectively results in 86% and 95% of the sampled data falling within the set. There is a trade-off with the selection of a scale value for the sigma value. One would like to hypothesize that an arbitrary pixel mapped into the established boundary from the transformation actually belongs to the set. For higher-scaled values of sigma there is a higher likelihood of misclassification.

The estimated variances should be scaled to maximize the probability of detection, and to minimize the probabilities of misclassification. Let the parameter k be the scaled value for the each σ_i.

The ellipsoid in RGB space bounding the set of points of interest is

$$\frac{\rho^2}{\sigma_\rho^2} + \frac{\gamma^2}{\sigma_\gamma^2} + \frac{\beta^2}{\sigma_\beta^2} = k^2 \tag{5}$$

where κ is the scaled σ_i value. The parameter κ is selected to optimize the outcome of the hypotheses tests as mentioned above. Each point can be tested using equation (5). The bounding ellipsoid can be transformed back into RGB space by the inverse transformation shown by equation (6).

$$\begin{bmatrix} r \\ g \\ b \end{bmatrix}_i = U_i \begin{bmatrix} \rho \\ \gamma \\ \beta \end{bmatrix}_i + T_i \tag{6}$$

V. PROCEDURE SUMMARY

Color segmentation using the SVD method described here can be accomplished as follows.

1. Obtain a sample from the targets of interest.
2. Find the means for each target using (2), then transform samples from each target such that they are zero mean samples.
3. Use the singular value decomposition (1) to find the transformation matrix U for each target data set.
4. Map each data set into a range space where the data is aligned with the principle axes using (3).
5. Use (4) to find the variance in the transformed space or use the eigenvalues of the covariance matrix C.
6. If the variance is above a selected threshold, partitioning the data into sub clusters as required.
7. Establish the boundaries of the sample set as an intersection of clusters found for each target using (5) applied to each cluster.
8. Each set can be visualized by transforming the set boundaries back into the RGB space by the inverse transformation of equation (6).
9. Test each pixel in the entire image, or a smaller area of interest within the image to detect the target.

VI. EXPERIMENTS

Two sets of experiments were conducted to investigate the SVD based method. The objective for the first experiment was to evaluate method in various color spaces. The objective of the second experiment was to evaluate the method as extended to complex scenes with texture and shadows. Color representations of all figures in this paper are available on the world wide web at our visual computing laboratory world wide web site URL "http://www.vision.ucsd.edu".

Figure 1. Test image of scene with Gumby and his friends.

A. Experiment 1

A scene of Gumby and his friends was used as a test image for the first experiment. Figure 1 shows this scene. Statistics from a 30x9 pixel swath on Gumby's belly was used to find the transformation to an appropriate principal component space for this target. Tests were conducted using color models from the receiver primary RGB color space, the C.I.E. XYZ system, xyY system, and the HSI system. The C.I.E. XYZ system, xyY system, and the HSI systems were selected to de-couple the intensity from the chromaticity components. For humans, vision is conveniently satisfied by red, green, and blue color primaries because of the spectral properties of the RGB cones in the human vision system. However, machine vision perception systems have no such limitation to RGB. In order to segment colors in the presence of shadows, color systems which de-couple the intensity from the chromaticity values were evaluated. This method was applied directly to each color space evaluated.

The test sample was taken using camera data reported in the NTSC receiver primary RGB space. Each evaluation was carried out using a 2- sigma ellipsoid in the selected color space by scaling the standard deviations by a factor of 2. Figure 2 shows the 2-sigma ellipsoid found from applying this method to the sampled data in the RGB space. The "o" annotations shown in Figures 2, 3, and 4 represent the sampled raw data points acquired in RGB space. The "x" annotations in Figures 2, 3, and 4 show locations of the mean RGB values projected onto each plane surface. The "*" annotations shown in Figures 2, 3, and 4 represent the bounding ellipsoids found in the respective color space and re-mapped back to the RGB space as presented here. The captions of Figures 2 through 10 describe experiment 1 results. Details about transformation derivations and a general discussion of color spaces can be found in Gonzalez and Woods [6] and Jain [8]. Each result shown represents the negative of the image to minimize a mostly black background.

B. Experiment 2

Application of the SVD based method was extend to texture rich images in experiment 2. The samples were segmented into clusters within the xyY color space. The results of the cluster-splitting combined with the SVD based segmentation are shown in Figures 11 through 19.

VII. CONCLUSIONS

This method was demonstrated in RGB, XYZ, xyY, and HSI color spaces. It is generally believed that it is better to de-couple intensity from the chromaticity values within a color subspace. However, this SVD based segmentation method produced a good result in all the color spaces evaluated. Based on the de-coupling of the intensity argument, the results from the first experiment are also favorable when one unknowingly attempts arbitrary implementation of this method in RGB color spaces. The results show fairly good segmentation of the target sample in all the spaces shown. Figure 8 shows some of the particle board stage as being in the set of desired color points. The result of projecting the XYZ ellipsoid onto the X-Z plane let an apparent significant more amount of colors into the set. The "just noticeable color difference" for the X-Y chromaticity plane in XYZ space is highly spatially dependent. The clustering methodology used in the second experiment eliminated a general restriction requiring targets need be localized within the selected color space. Partitioning was satisfactorily accomplished using the variance tested against a selected threshold. Measures of the need for a partition and an accompanying strategy for splitting the color space into subsets region is the subject of further research. Further work is required to investigate the default selection of the ellipse for the boundaries of the set of color points. Evaluation of super-quadrics as a bounding element may lead to a more promising description of these sets for many applications.

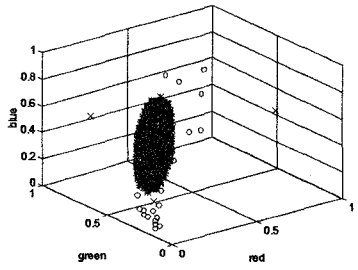

Figure 2. Two-sigma ellipsoid computed and represented in RGB space.

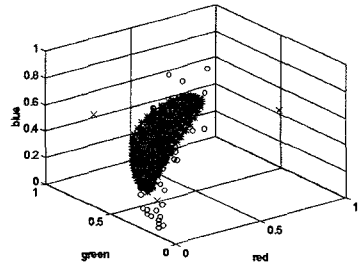

Figure 3. A transformed 2-sigma ellipsoid computed in xyY space and re-mapped into RGB space.

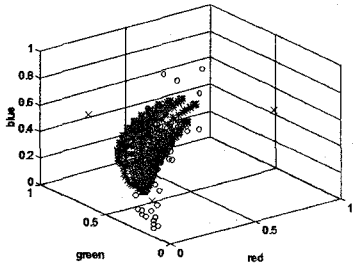

Figure 4. Two-sigma ellipsoid computed in HSI space and re-mapped into RGB space.

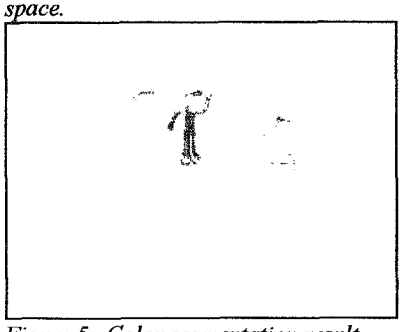

Figure 5. Color segmentation result from applying SVD to localized color sample in RGB space.

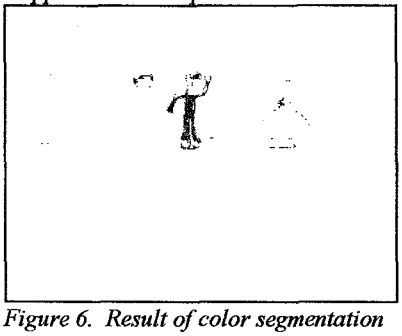

Figure 6. Result of color segmentation from applying SVD method to localized color sample in xyY space.

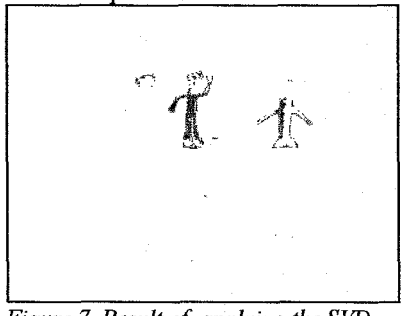

Figure 7. Result of applying the SVD method to localized color sample in xyY space and projecting the bounding ellipsoid onto the x-y chromaticity plane.

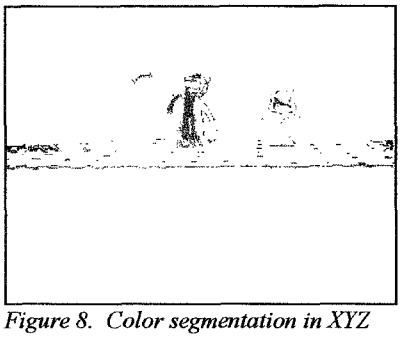

Figure 8. Color segmentation in XYZ space and projecting the bounding ellipsoid onto the X-Z chromaticity plane.

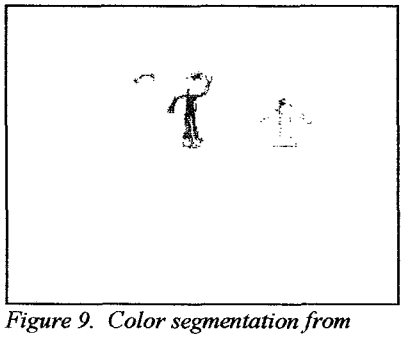

Figure 9. Color segmentation from applying the SVD method to localized color sample in HSI color space.

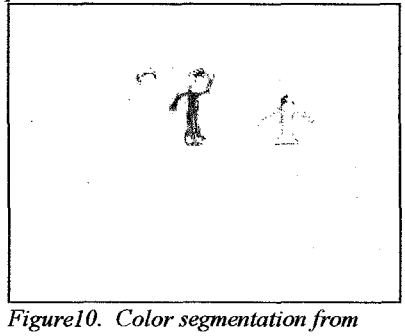

Figure10. Color segmentation from applying the SVD method to localized color sample in HSI color space and projecting the bounding ellipsoid onto the H-S plane.

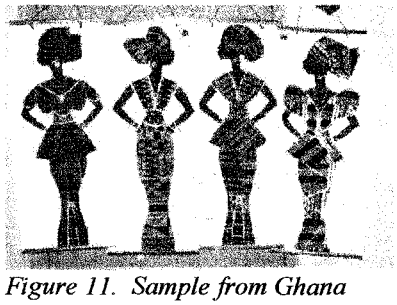

Figure 11. Sample from Ghana scrolls of women in Kente cloth.

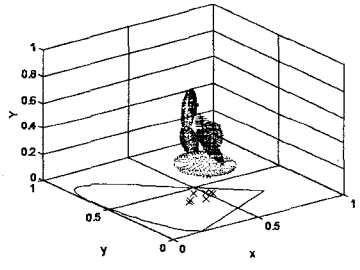

Figure 12. Two-sigma partitions for scrolls formulated in xyY color space.

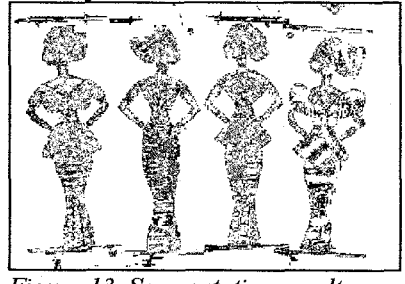

Figure 13. Segmentation result from sampling the dress in the second scroll from the left.

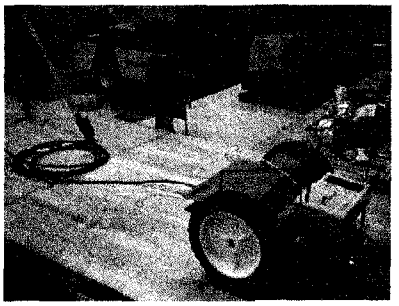

Figure 14. Sample from robot in our visual computing laboratory.

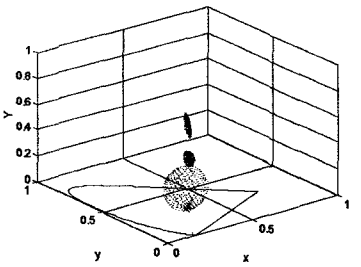

Figure 15. Two-sigma partitions for robot sample formulated in xyY.

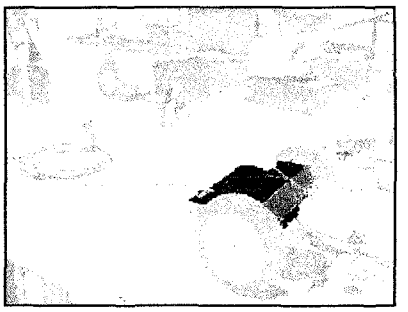

Figure 16. Segmentation result from sampled robot in foreground.

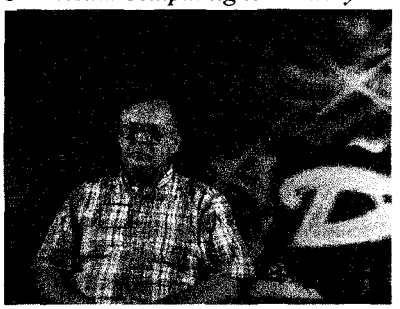

Figure 17. Sample from motion picture professional's shirt posing in cluttered studio scene.

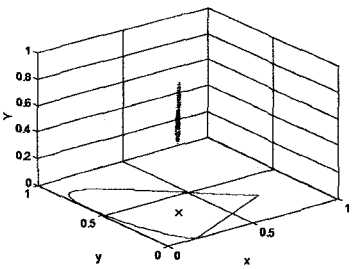

Figure 18. Two-sigma partitions for shirt sample formulated in xyY.

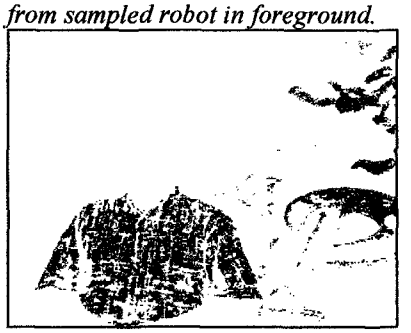

Figure 19. Segmentation result from sampling area on shirt.

ACKNOWLEDGMENTS

We want to thank Mr. Mark Elder for his support and suggestions during data acquisition at the University of California, San Diego motion picture studio.

REFERENCES

[1] W. L. Brogan, Modern Control Theory, 3rd ed., Prentice Hall, 1991.

[2] M. Celenk, "A color clustering technique for image segmentation," Computer Graphics and Image Processing, Vol. 52, 1990, pp. 145-170.

[3] C.-C. Chang, and L.-L. Wang, "Color texture segmentation for clothing in a computer-aided fashion design system," Image and Vision Computing, Vol. 14, 1996, pp. 685-702.

[4] C. -T. Chen, Linear System Theory and Design, Harcourt Brace Jovanovich, 1984.

[5] J. Craig, Introduction to Robotics: Mechanics and Control, 2nd ed., Addison-Wesley, 1989.

[6] R. C. Gonzalez, and R. E. Woods, Digital Image Processing, Addison-Wesley, 1992.

[7] J. A. Hartigan, Clustering Algorithms, Wiley, 1975.

[8] A. K. Jain, Fundamentals of Digital Image Processing, Prentice Hall, 1989.

[9] T. Kanade, "Region Segmentation: Signal vs. Semantics," *Proceedings of the 4th International Joint Conference on Pattern Recognition*, pp. 95-105, Nov. 7-10 1978, Kyoto Japan.

[10] J. Liu, and Y.-H. Yang, "Multiresolution color image segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 16, No. 7, Jul. 1994, pp. 689-700.

[11] Y.-I. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," Computer Graphics and Image Processing, Vol. 13, 1980, pp. 222-241.

[12] W. H. Press, S. A. Teukolosky, W. T. Vetterling, B. P. Flannery, Numerical Recipes in C - The Art of Scientific Computing, 2nd ed., Cambridge University Press, 1995.

[13] G. Strang, Linear Algebra and its Applications, 3rd ed., Harcourt Brace Jovanovich, 1988.

[14] S. E. Umbaugh, R. H. Moss, W. V. Stoecker, and G. A. Hance, "Automatic color segmentation algorithms with application to skin tumor feature identification," IEEE Engineering in Medicine and Biology, Vol. 12, Sep. 1993, pp. 75-82.

[15] A. P. Witkin, "Scale space filtering: a new approach to multiscale descriptor," Image Understanding, eds. S. Ullman, and W. Richards, 1984, pp. 79-85.