

Extended Pixel Representation for Image Segmentation

Deeptha Girish, Vineeta Singh, Anca Ralescu

EECS Department

University of Cincinnati

Cincinnati, OH 45221-0030, USA

girishde@mail.uc.edu, singhvi@mail.uc.edu, anca.alescu@uc.edu

Abstract

We explore the use of extended pixel representation for color based image segmentation using the K-means clustering algorithm. Various extended pixel representations have been implemented in this paper and their results have been compared. By extending the representation of pixels an image is mapped to a higher dimensional space. Unlike other approaches, where data is mapped into an implicit features space of higher dimension (kernel methods), in the approach considered here, the higher dimensions are defined explicitly. Preliminary experimental results which illustrate the proposed approach are promising.

Introduction

Image segmentation is a key processes in image analysis. This step is done irrespective of the goal of the analysis. It is one of the most critical tasks in this process. The aim of segmentation is to divide the image into non overlapping areas such that all pixels in one segment have similar features. This representation makes the image more meaningful and makes its analysis and understanding easier. For example, in medical images each color represents a particular stain that is essentially a protein binder that binds to a certain type of molecule. Therefore the different clusters obtained after segmenting such an image according to color can help us identify different biological components in that image. This can help us to do further analysis like finding meta data about each of the structures, finding one structures relative position to the other etc.

One way to achieve this is to cluster similar pixels into one cluster. Therefore, what defines similarity and how it is calculated becomes very important. A lot of research has been done on finding similarity measures for images and videos (Wang et al.(2005)Wang, Zhang, and Feng), (Huttenlocher et al.(1993)Huttenlocher, Klanderman, and Rucklidge), (Gualtieri et al.(1992)Gualtieri, Le Moigne, and Packer). It is important that the measure we use captures the slightest difference in pixels. This plays an essential role in segmentation. K-means (MacQueen et al.(1967)) is a popular algorithm that works well for image segmentation. The aim of the K-means algorithm as stated by Hartigan ET. al. (Hartigan and Wong(1979)) is to divide M points

in N dimensions into K clusters so that the within-cluster sum of squares distances is minimized. Thus, distance is an important factor. One of the major difficulties of image segmentation is to differentiate two similar regions which actually belong to different segments. This is because the existing features and the standard distances used to measure the dissimilarity do not capture the small differences very well. Increasing the sensitivity to small differences is the motivation for using the extended pixel representation. Image data is displayed on a computer as a bitmap which is a rectangular arrangement of pixels. The number of bits used to represent individual pixel has increased over the recent years, as computers have become more powerful. Today's computer systems often use a 24 bit color system. The most common color system in use is the RGB space (Red, Green and Blue).

Selecting the right color space for segmentation can be very important. Different parts of the image get highlighted better in different color spaces. It is also dependent on the purpose of the segmentation. To simplify this choice of color space, it is better to know in advance the features that represent maximum variation. This can be done by finding the coefficient of variation for each of the features of different color spaces. For a given data set, the coefficient of variation, defined as the ratio of standard deviation to the mean, represents the variability within the data set with respect to the mean. When selecting features/dimensions, it then makes sense to take into consideration the coefficient of variation, more precisely, to select the features that have the highest variability. In selecting a color space, it is usual to adopt one (e.g., RGB, HSV) space. This study departs from this policy by selecting more than one color spaces, which further helps in better differentiate between pixels.

Related Work

The idea of extended pixel representation has been used previously to detect and repair halo in images (Ohshima et al.(2006)Ohshima, Ralescu, Hirota, and Ralescu), to calculate the degree of difference between pixels in color images in order to identify the halo region. The extended pixel representation used in (Ohshima et al.(2006)Ohshima, Ralescu, Hirota, and Ralescu) is

Table 1: The original and extended color spaces

Original Color Space	Extended Color Space
(RGB)	$(RGB \rho^{RG} \theta^{RG} \rho^{GB} \theta^{GB} \rho^{BR} \theta^{BR})$
(HSV)	$(HSV \rho^{HS} \theta^{HS} \rho^{SV} \theta^{SV} \rho^{VS} \theta^{VS})$
$(YCbCr)$	$(YCbCr \rho^{YCb} \theta^{YCb} \rho^{CbCr} \theta^{CbCr} \rho^{CrY} \theta^{CrY})$
(RGB) (HSV) $(YCbCr)$	$(RGBHSVYCbCr)$

described next. Given a pixel p in the image I , its extended pixel representation of p , is the real-valued vector $(R, G, B, \rho^{RG}, \theta^{RG}, \rho^{GB}, \theta^{GB}, \rho^{BR}, \theta^{BR})$ where RGB are pixel values in the RGB color space, and ρ^{XY} and θ^{XY} are the polar coordinates in the two dimensional color space, (X, Y) , with $X, Y \in R, G, B, X \neq Y$. This representation better support image segmentation because it captures the difference between pixels more accurately. In particular, it gives much better results for images with low contrast.

The same extended pixel representation can be implemented in other color spaces as well, as our experimental results for HSV, which uses H for Hue, synonymous to color, S for Saturation, the colorfulness of a color relative to its own brightness, and V for Value which refers to the lightness or darkness of a color. It is known that for most images, HSV color space is more suitable for segmentation and we show that the extended pixel representation for HSV improves the results even for this color space. The YCbCr space uses Y for Luminance, the brightness of light, C for Chrominance, the color information of the signal, which is stored as two color-difference components (Cb and Cr). Table 1 shows the features for each color space used in the experiments.

Current Approach

To illustrate the effect of the extended pixel representation, we consider two experiments as follows.

Sensitivity of distance to the extended space

Two pixels which are almost of the same color and having very similar R, G and B values are selected and their corresponding values in each of the extended pixel representations considered are calculated, using both the Euclidean distance and Manhattan distance. It is seen that the extended pixel representations are more sensitive to the dissimilarity between the pixels irrespective of the distance measure used. This is the idea behind using extended pixel representation. Especially for tasks like segmentation, this is a very efficient way to separate two samples which have similar feature values into different clusters.

Table 2 shows the results for computing the distance between the two fixed pixels, in three standard color spaces, their extensions, and in an extended representation based on three color spaces.

Table 2: Euclidean distance (ED) and Manhattan distance (MD) for two fixed pixels in the original and extended color spaces, and the percentage of change due to the extended representation.

Pixel Representation	ED	MD
(RGB)	54.08	75.00
$(RGB \rho^{RG} \theta^{RG} \rho^{GB} \theta^{GB} \rho^{BR} \theta^{BR})$	80.47	175.40
% of change	48.78	133.86
(HSV)	5.00	5.13
$(HSV \rho^{HS} \theta^{HS} \rho^{SV} \theta^{SV} \rho^{VS} \theta^{VS})$	8.59	15.17
% of change	71.69	195.62
$(YCbCr)$	29.15	32.81
$(YCbCr \rho^{YCb} \theta^{YCb} \rho^{CbCr} \theta^{CbCr} \rho^{CrY} \theta^{CrY})$	48.24	92.10
% of change	65.52	180.72
$(RGBHSVYCbCr)$	61.64	112.94
% of change with respect to average distances for the original color spaces	109.58	200.00

Obviously, all distances in the extended space should be larger than those in the original space, because adding more features to the pixel representation in effects adds positive quantities to each of the distances considered and therefore this experiment merely confirms that theoretical fact.

Sensitivity of the distance in the original and extended space when the pixel changes

A second small experiment was run as follows. Starting with two pixels p , and p' , ED and MD were computed just as in the previous section. Then p_2 was altered slightly, to obtain p'' , and the same distances were computed for p and p'' , and compared with those computed for p and p' . Tables 3 and 4 show the results.

Table 3: Changes in the Euclidean Distance (ED) when the pixel p' changes to p'' . RGB pixel values are $p = (10, 100, 150)$, $p' = (25, 120, 135)$, $p'' = (30, 125, 130)$.

Pixel Representation	(p_1, p')	(p_1, p'')	diff %
(RGB)	29.15	37.75	29.48
$(RGB \rho^{RG} \theta^{RG} \rho^{GB} \theta^{GB} \rho^{BR} \theta^{BR})$	54.77	71.24	30.06
(HSV)	15.00	20	33.29
$(HSV \rho^{HS} \theta^{HS} \rho^{SV} \theta^{SV} \rho^{VS} \theta^{VS})$	25.89	34.52	33.32
$(YCbCr)$	19.22	24.74	28.68
$(YCbCr \rho^{YCb} \theta^{YCb} \rho^{CbCr} \theta^{CbCr} \rho^{CrY} \theta^{CrY})$	38.40	49.43	28.70
$(RGBHSVYCbCr)$	38.01	49.36	29.88

Experiments and Results

Segmentation using the k-means algorithm was implemented for three different images. Each image is in a different color space. Segmentation using K-means was done on these images using the standard pixel representation and the extended pixel representation. The Euclidean distance

Table 4: Changes in the Manhattan Distance (ED) when the pixel p' changes to p'' . RGB pixel values are $p = (10, 100, 150)$, $p' = (25, 120, 135)$, $p'' = (30, 125, 130)$.

Pixel Representation	(p_1, p')	(p_1, p'')	diff %
(RGB)	50	65	30
$(RGB \rho^{RG} \theta^{RG} \rho^{GB} \theta^{GB} \rho^{BR} \theta^{BR})$	120.3	155.4	29.17
(HSV)	15.15	20.22	33.39
$(HSV \rho^{HS} \theta^{HS} \rho^{SV} \theta^{SV} \rho^{VS} \theta^{VS})$	45.11	60.16	33.36
$(YCbCr)$	27.40	35.82	30.73
$(YCbCr \rho^{YCb} \theta^{YCb} \rho^{CbCr} \theta^{CbCr} \rho^{CrY} \theta^{CrY})$	82.13	106.1	29.22
$(RGBHSVYCbCr)$	92.55	121.0	30.77

between pixels (in original color space and extended space) was used. For every image, the number of clusters was decided visually. The same number of clusters was used for all the extended pixel representations. The mean square error, equation (1) and the signal to noise ratio, equation (2) are calculated.

$$MSE(r, t) = \frac{1}{M \times N} \sum_{n=1}^M \sum_{m=1}^N [t(n, m) - r(n, m)]^2 \quad (1)$$

$$Snr(t, r) = 10 \log_{10} \left(\frac{\sum_{m=1}^M \sum_{n=1}^N [r(n, m)]^2}{\sum_{m=1}^M \sum_{n=1}^N [r(n, m) - t(n, m)]^2} \right) \quad (2)$$

where $r(n, m)$ represents the original image and $t(n, m)$ represents the segmented image, both of size $[N, M]$.

Table 5 shows the results of the segmentation and MSE and Snr for each of these segmentations. Inspecting Table 5 it is observed that for all images, even though the distance measure and the number of clusters were same, the mean square error was lower and correspondingly signal to noise ratio was higher when the extended pixel representation was used. It can be seen that the results look very similar to the original image because the clusters are colored with the mean color of all the pixels in that cluster. It can also be observed that the difference in results when the extended pixel representation is used is higher for images of low contrast. The extended pixel representation also performs better for pictures with high texture content.

The smaller changes in color values of the pixels are separated and recognized better in extended pixel representations. Irrespective of the color space used, the extended pixel representation gives lower mean squared error and higher signal to noise ratio than the standard pixel representation for all images. All pixels that are similar in color get clustered into one segment which might represent a particular structure in the image. Thus, this representation adds more meaning to the image.

Conclusion

The segmentation problem considered in this paper can be addressed by using the extended pixel representations.

The idea of representing each pixel with more information proved to be successful for image segmentation. It was observed that extended pixel representation can more effectively distinguish similar pixels. Although in the second experiment, it was seen that the percentage change of the euclidean distance between the extended pixel representation and the standard pixel representation when p was changed to p' is very small, the effect turns out to be significant in the clustering step. With the same number of clusters and the same distance measure used, the mean square of the segmentation done using the extended pixel representation is lower for all images.

The promising experimental results of this idea encourages exploration of other extended pixel representations in the future. It will be interesting to include texture information like edges or frequency domain information as part of the pixel representation. Tailoring the extended pixel representations for the task in hand and automatic learning of the most important features and optimal number of features to represent a pixel for a particular task is an important idea to be considered for future work.

References

JA Gualtieri, J Le Moigne, and CV Packer. Distance between images. In *Frontiers of Massively Parallel Computation, 1992., Fourth Symposium on the*, pages 216–223. IEEE, 1992.

John A Hartigan and Manchek A Wong. Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28(1):100–108, 1979.


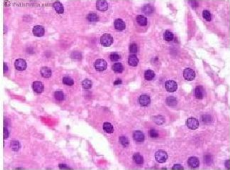
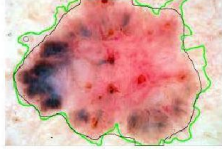

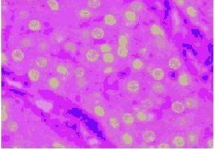


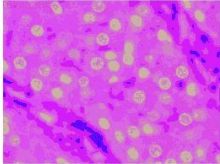
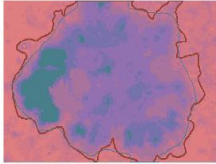

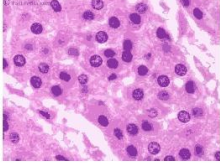

Daniel P Huttenlocher, Gregory A Klanderman, and William J Rucklidge. Comparing images using the Hausdorff distance. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 15(9):850–863, 1993.

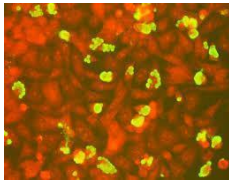
James MacQueen et al. Some methods for classification and analysis of multivariate observations. In *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, volume 1, pages 281–297. Oakland, CA, USA., 1967.

Chihiro Ohshima, Anca Ralescu, Kaoru Hirota, and Dan Ralescu. Processing of halo region around a light source in color images of night scenes. In *Proc. of the Conference IPMU*, volume 6, pages 2–7, 2006.

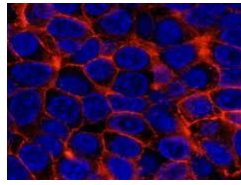
Liwei Wang, Yan Zhang, and Jufu Feng. On the Euclidean distance of images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 27(8):1334–1339, 2005.

Table 5: Results on segmentation of six images, using each of the three color spaces separately (2nd row), extended pixel representation (3rd row), and all three color spaces (4th row). Bold font indicates the smallest MSE.

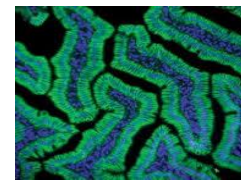
		
a) Original Image	b) Original Image	c) Original Image
		
a) RGB Snr=35.3260db MSE=0.0050	b) HSV Snr=54.8959db MSE = 0.0035	c) YCbCr Snr=55.8133db MSE=0.0031
		
a) $RGB\rho^{\text{RG}}\theta^{\text{RG}}\rho^{\text{GB}}\theta^{\text{GB}}\rho^{\text{BR}}\theta^{\text{BR}}$ Snr=39.3334db MSE=0.0033	b) $HSV\rho^{\text{HS}}\theta^{\text{HS}}\rho^{\text{SV}}\theta^{\text{SV}}\rho^{\text{VH}}\theta^{\text{VH}}$ Snr=56.1241db MSE=0.0026	c) $YCbCr\rho^{\text{YCb}}\theta^{\text{YCb}}\rho^{\text{CbCr}}\theta^{\text{CbCr}}\rho^{\text{CrY}}\theta^{\text{CrY}}$ Snr=58.8801db MSE=0.0015
		
a) RGB HSV YCbCr Snr=37.4722db MSE=0.0173	b) RGB HSV YCbCr Snr=57.064db MSE=0.0021	c) RGB HSV YCbCr Snr=35.4811db MSE=0.0173



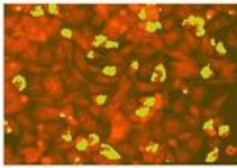
d) Original Image



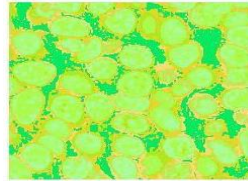
e) Original Image



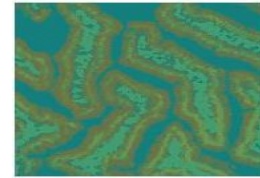
f) Original Image



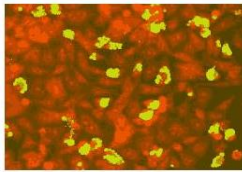
d) RGB
 Snr=47.4465db
 MSE=0.0014



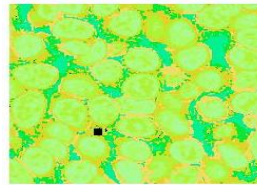
e) HSV
 Snr=44.0748db
 MSE=0.0062



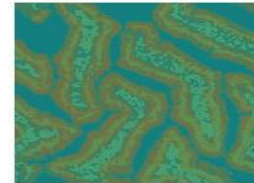
f) YCbCr
 Snr=54.3042db
 MSE=0.00075



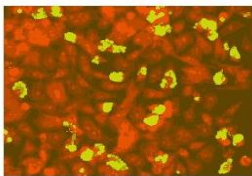
d) RGB $\rho^{RG} \theta^{RG} \rho^{GB} \theta^{GB} \rho^{BR} \theta^{BR}$
 Snr=49.2535db
MSE=0.0012



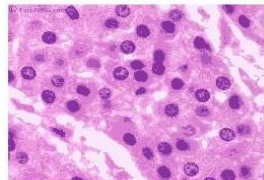
e) HSV $\rho^{HS} \theta^{HS} \rho^{SV} \theta^{SV} \rho^{VH} \theta^{VH}$
 Snr=47.3483db
 MSE=0.0044



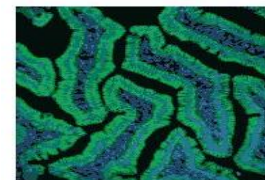
f) YCbCr $\rho^{YCb} \theta^{YCb} \rho^{CbCr} \theta^{CbCr} \rho^{CrY} \theta^{CrY}$
 Snr=54.9753db
MSE=0.0007



d) RGB HSV YCbCr
 Snr=48.1598db
 MSE=0.0013



e) RGB HSV YCbCr
 Snr=26.9429db
MSE=0.0038



e) RGB HSV YCbCr
 Snr=32.4912db
 MSE=0.0037