

A Modified Fuzzy C-Means Algorithm for Document Image Segmentation

Mohamed Nooman Ahmed, Ph.D.

Applied Software Research, Lexmark International Inc., Lexington, Kentucky

Email: mahmed@lexmark.com

Abstract

In this paper, we present a new system to segment and label document images into text, images, and background using a modified Fuzzy C-means Algorithm. Each pixel is assigned a feature pattern extracted from the gray level distribution and computed at different scales. The invariant feature pattern is then assigned to a specific region using Fuzzy logic. Our algorithm is formulated by modifying the objective function of the standard fuzzy c-means (FCM) algorithm to allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood. The neighborhood effect acts as a regularizer and biases the solution towards piecewise-homogeneous labelings. Such a regularization is useful in segmenting scans corrupted by scanner noise.

1 Introduction

Several approaches for document segmentation have been proposed.^{1,6} These techniques can be broadly classified as bottom-up and top-down. Bottom-up methods start from pixel level and merge regions together into larger and larger components. Top-down techniques apply a priori knowledge about the page to hypothesize and split the page into blocks which are subsequently identified and further subdivided. Top-down approaches work well with pre-specified layouts such as technical papers. However, the performance of these techniques degrades significantly when different components are touching or overlapping.

Among bottom-up approaches, texture-based schemes have attracted much attention.^{4,6} These methods treat different components of a document image as different textures. The scanned document images are convolved with a set of masks to generate feature vectors. Each feature vector is then classified into different classes using a pre-trained classifier.

The results of these techniques, however, are highly sensitive to noise. Another problem associated with these approaches is the mask size for extracting local features. If the mask size is too small, it is difficult to detect large scale textures such as large fonts. On the contrary, if a large mask

is chosen, the computational complexity will increase dramatically.⁶

To solve the problem of noise sensitivity and to exploit both local features and image structure, we present in this paper a different approach for fuzzy segmentation of document images. Each pixel is assigned a feature pattern consisting of gray level distribution computed at different scales. The invariant feature pattern is then assigned to a specific region using Fuzzy logic.⁸⁻¹²

Our algorithm is formulated by modifying the objective function of the standard fuzzy c-means (FCM) algorithm to allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood. The neighborhood effect acts as a regularizer and biases the solution towards piecewise-homogeneous labelings. Such a regularization is useful in segmenting scans corrupted by scanner noise.

2 Feature Extraction

The first step in our approach is to extract a feature pattern for each pixel in a gray level image. A feature vector X is a set of measurements $\{x_1, x_2, \dots, x_d\}$ which condenses the description of relevant properties of the image into a small, Euclidean feature space of N dimension. The number of needed features depends on the complexity of the image. The components of the feature vector may include gray values, gray values through different filters, texture measures, Markov random field features, fractal dimension measures, and gradient magnitudes and directions.⁵

In this paper, we will construct our feature vector from measurements obtained from gray level distribution in a window, W , of dimension $w \times w$ centered around each pixel. These features describe the first order gray level distribution without considering the spatial interdependence. Two features were selected:

1. The mean gray level, μ ,

$$\mu = \frac{1}{w^2} \sum_{(x,y) \in W} I(x,y) \quad (1)$$

where $I(x,y)$ is the gray level at location (x,y) .

2. The variance σ^2 of the gray level variation

$$\sigma^2 = \frac{1}{w^2 - 1} \sum_{(x,y) \in \mathcal{W}} (I(x,y) - \mu)^2. \quad (2)$$

3 Standard Fuzzy-C-Means

The standard FCM objective function for partitioning $\{x_k\}_{k=1}^N$ into c clusters is given by

$$J = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|x_k - v_i\|^2 \quad (3)$$

where $\{x_k\}_{k=1}^N$ are the feature vectors for each pixel, $\{v_i\}_{i=1}^c$ are the prototypes of the clusters and the array $[u_{ik}] = U$ represents a partition matrix, $U \in \mathcal{U}$, namely

$$\mathcal{U} \{ u_{ik} \in [0, 1] \mid \sum_{i=1}^c u_{ik} = 1 \quad \forall k \text{ and} \\ 0 < \sum_{k=1}^N u_{ik} < N \quad \forall i \} \quad (4)$$

The parameter p is a weighting exponent on each fuzzy membership and determines the amount of fuzziness of the resulting classification. The FCM objective function is minimized when high membership values are assigned to voxels whose intensities are close to the centroid of its particular class, and low membership values are assigned when the pixel data is far from the centroid.

4 Modified Fuzzy C-Means Objective Function

We propose a modification to Eq.(3) by introducing a term that allow the labeling of a pixel to be influenced by the labels in its immediate neighborhood. As mentioned before, the neighborhood effect acts as a regularizer and biases the solution towards piecewise-homogeneous labeling. Such a regularization is useful in segmenting scans corrupted by salt and pepper noise. The modified objective function is given by

$$J_m = \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \|x_k - v_i\|^2 \\ + \frac{\alpha}{N_R} \sum_{i=1}^c \sum_{k=1}^N u_{ik}^p \left(\sum_{x_r \in \mathcal{N}_k} \|x_r - v_i\|^2 \right) \quad (5)$$

where \mathcal{N}_k stands for the set of neighbors that exist in a window around x_k and N_R is the cardinality of \mathcal{N}_k . The effect of the neighbors term is controlled by the parameter α . The

relative importance of the regularizing term is inversely proportional to the signal to noise ratio (SNR) of the MRI signal. Lower SNR would require a higher value of the parameter α .

Formally, the optimization problem comes in the form

$$\min_{U, \{v_i\}_{i=1}^c} J_m \quad \text{subject to} \quad U \in \mathcal{U} \quad (6)$$

5 Parameter Estimation

The objective function J_m can be minimized in a fashion similar to the standard FCM algorithm. Taking the first derivatives of J_m with respect to u_{ik} , v_i , and setting them to zero results in two necessary but not sufficient conditions for J_m to be at a local extrema. In the following subsections, we will derive these three conditions.

5.1 Membership Evaluation

The constrained optimization in equation (6) will be solved using one Lagrange multiplier

$$F_m = \sum_{i=1}^c \sum_{k=1}^N \left(u_{ik}^p D_{ik} + \frac{\alpha}{N_R} u_{ik}^p \gamma_i \right) + \lambda \left(1 - \sum_{i=1}^c u_{ik} \right) \quad (7)$$

where $D_{ik} = \|x_k - v_i\|^2$ and $\gamma_i = \left(\sum_{x_r \in \mathcal{N}_k} \|x_r - v_i\|^2 \right)$

Taking the derivative of F_m w.r.t u_{ik} and setting the result to zero, we have, for $p > 1$

$$\left[\frac{\delta F_m}{\delta u_{ik}} = p u_{ik}^{p-1} D_{ik} + \frac{\alpha p}{N_R} u_{ik}^{p-1} \gamma_i - \lambda \right]_{u_{ik} = u_{ik}^*} = 0 \quad (8)$$

Solving for u_{ik}^* we have

$$u_{ik}^* = \left(\frac{\lambda}{p(D_{ik} + \frac{\alpha}{N_R} \gamma_i)} \right)^{\frac{1}{p-1}} \quad (9)$$

Since $\sum_{j=1}^c u_{jk} = 1 \quad \forall k$, then

$$\sum_{j=1}^c \left(\frac{\lambda}{p(D_{jk} + \frac{\alpha}{N_R} \gamma_j)} \right)^{\frac{1}{p-1}} = 1 \quad (10)$$

or

$$\lambda = \frac{p}{\left(\sum_{j=1}^c \left(\frac{1}{(D_{jk} + \frac{\alpha}{N_R} \gamma_j)} \right)^{\frac{1}{p-1}} \right)^{p-1}} \quad (11)$$

Substituting into equation (9), the zero-gradient condition for the membership estimator can be rewritten as,

$$u_{ik}^* = \frac{1}{\sum_{j=1}^c \left(\frac{D_{ik} + \frac{\alpha}{N_R} \gamma_i}{D_{jk} + \frac{\alpha}{N_R} \gamma_j} \right)^{\frac{1}{p-1}}} \quad (12)$$

5.2 Cluster Prototype Updating

Using the standard Euclidian distance and taking the derivative of F_m w.r.t v_i and setting the result to zero we have;

$$\left[\begin{array}{l} \sum_{k=1}^N u_{ik}^p (x_k - v_i) + \\ \sum_{k=1}^N u_{ik}^p \frac{\alpha}{N_R} \sum_{y_r \in \mathcal{N}_k} (x_r - v_i) \end{array} \right]_{v_i=v_i^*} = 0. \quad (13)$$

Solving for v_i we have

$$v_i^* = \frac{\sum_{k=1}^N u_{ik}^p \left((x_k) + \frac{\alpha}{N_R} \sum_{x_r \in \mathcal{N}_k} (x_r) \right)}{(1 + \alpha) \sum_{k=1}^N u_{ik}^p} \quad (14)$$

6 Results

We tested our algorithm with several images scanned at 600 dpi. Six different features were extracted for every pixel. These features included first order statistics evaluated at three different scales. The images were segmented into text, images, and background. Results are shown in Figures 1 and 2.

References

1. Y. Y. Tang, C. Y. Suen, and M. Cheriet, "Document Analysis and Understanding: A brief Survey" *Proc. Int. Conf. on Document Analysis and Recognition*, pp.17-31, 1991.
2. H. Cheng and C. Bouman, "Trainable Context Model for Multiscale Segmentation" *Proc. Int. Conf. on Image Processing (ICIP)*, October 1998.

3. L. O'Gorman, "The Document Spectrum for Page Layout Analysis" *IEEE Trans. on Pattern Matching and Machine Intelligence (PAMI)*, Vol. 15, No. 11, November 1993.
4. K. Etemad, D. Doermann, and R. Chellapa, "Page segmentation using decision integration and wavelet packets," *Proc. Int. Conf. of Pattern Recognition*, Vol. 2, pp. 345-349, 1994.
5. A. K. Jain and Y. Zhong, "Page segmentation using texture analysis," *Pattern Recognition*, Vol. 29, no 5, pp. 343-770, 1996.
6. H. Cheng, C. Bouman, and J. Allebach, "Multiscale Document Segmentation," *Proc. of IS&T 50th Annual Conference*, pp. 417-425, May 1997.
7. S. Dellepiane et al., "Model generation and model matching of real images by fuzzy approach," *Pattern Recognition*, Vol. 25, No.2, pp. 115-137, 1992.
8. J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, Vol. 3, pp. 32-57, 1973.
9. J. Bezdek, "A convergence theorem for the fuzzy ISO-DATA clustering algorithms," *IEEE Trans. on Pattern Ana. Machine Intell. (PAMI)*, 1980.
10. C. Xu, D. Pham, and J. Prince, "Finding the brain cortex using fuzzy segmentation, isosurfaces, and deformable surfaces," *Proc. of the XVth Int. Conf. on Information Processing in Medical Imaging (IPMI 97)*, pp. 399-404, 1997.
11. D. L. Pham, and J. L. Prince, "Adaptive fuzzy segmentation of magnetic resonance images," *IEEE Trans. Medical Imaging*, Vol. 18, No. 9, pp. 737-752, 1999.
12. M. N. Ahmed, S. M. Yamany, N. A. Mohamed, A. A. Farag, and T. Moriarty "Bias field estimation and adaptive segmentation of MRI data using modified fuzzy C-means algorithm," *Proc. IEEE Int. Conf. Computer Vision and Pattern Recogn.*, Fort Collins, Colorado, Vol. 1, pp. 250-255, June 1999.

Biography

Dr. Mohamed Nooman Ahmed is a research scientist at Lexmark international. He obtained his Ph.D. in Computer Science and Engineering from the University of Louisville, KY, in 1998.

His research interests include image processing, enhancement, neural networks, pattern recognition and medical imaging.

MICHAEL CRICHTON

THE JURASSIC PARK AND SPHERE WRITER'S LOS ANGELES RESIDENCE



My sense of a home," says Michael Crichton, "is that when you move there you never know how it will be." Considering what he writes, you might picture him as someone who can't wait to find out—to the point of searching the shrubbery for dinosaur eggs or checking the filter of the air-conditioning unit for toxins. But that's to confuse the public with the private Michael Crichton. And in his Los Angeles house, there's only one in residence: the disarmingly difficult husband and father who has taken Taubert's advice—"Be regular and orderly in your life so that you may be violent and original in your work"—to heart.

Crichton is so busy, it's a wonder that he even has time for a home. He publishes a novel almost every year (*Rising Sun*, *Disclosure*, *Airframe*, *Jurassic Park*), and just about as often, a movie made from one of his books (like the recent *Sphere*) is released. His novels dominate the best-seller lists and are coveted by film studios. *Jurassic Park* (1995) is one of the five highest-grossing films of all time. Then there is *ER*, the television show he conceived and executive produces, which is consistently among the top rated. There's a formula at work: Crichton takes science, medicine and social issues and, without pandering, turns them into gripping drama.

His domestic story is far less dramatic. In the early 1980s Crichton was living in a Richard Neutra house in the Hollywood Hills. He met Anne-Marie Martin and soon decided that instead of a "nice, demanding, formal" residence, he wanted something more like a home. In west Los Angeles they found a sprawling, informal house that dated from the late 1920s and might have been designed by architect-to-the-stars Paul Williams. But its legacy was less important than its look. In a city of instant antiques, it seemed venerable—in this most sober-

ON THE INTERIOR DESIGNER Judith Kanner incorporated a mix of period ("I've got a fashion vision," she says in the living room, which is framed by Georgeanna's door round. Seating includes a French settee near a Bretonian chair and a pair of 1960s West German chairs.

INTERIOR DESIGN BY Judith Kanner
RENOVATION DESIGN BY Winston Chappell
TEXT BY Joyce Karshbath
PHOTOGRAPHY BY Mary El. Nichols

"It was one of those constant works in progress," recalls Michael Crichton of his Los Angeles house he shares with his wife, Anne-Marie, and daughter, Taylor. "It evolved from something better rather than formal ideas." Above: The entrance hall features an antique Knoll carpet.

292

MICHAEL CRICHTON

THE JURASSIC PARK AND SPHERE WRITER'S LOS ANGELES RESIDENCE



My sense of a home," says Michael Crichton, "is that when you move there you never know how it will be." Considering what he writes, you might picture him as someone who can't wait to find out—to the point of searching the shrubbery for dinosaur eggs or checking the filter of the air-conditioning unit for toxins. But that's to confuse the public with the private Michael Crichton. And in his Los Angeles house, there's only one in residence: the disarmingly difficult husband and father who has taken Taubert's advice—"Be regular and orderly in your life so that you may be violent and original in your work"—to heart.

Crichton is so busy, it's a wonder that he even has time for a home. He publishes a novel almost every year (*Rising Sun*, *Disclosure*, *Airframe*, *Jurassic Park*), and just about as often, a movie made from one of his books (like the recent *Sphere*) is released. His novels dominate the best-seller lists and are coveted by film studios. *Jurassic Park* (1995) is one of the five highest-grossing films of all time. Then there is *ER*, the television show he conceived and executive produces, which is consistently among the top rated. There's a formula at work: Crichton takes science, medicine and social issues and, without pandering, turns them into gripping drama.

His domestic story is far less dramatic. In the early 1980s Crichton was living in a Richard Neutra house in the Hollywood Hills. He met Anne-Marie Martin and soon decided that instead of a "nice, demanding, formal" residence, he wanted something more like a home. In west Los Angeles they found a sprawling, informal house that dated from the late 1920s and might have been designed by architect-to-the-stars Paul Williams. But its legacy was less important than its look. In a city of instant antiques, it seemed venerable—in this most sober-

ON THE INTERIOR DESIGNER Judith Kanner incorporated a mix of period ("I've got a fashion vision," she says in the living room, which is framed by Georgeanna's door round. Seating includes a French settee near a Bretonian chair and a pair of 1960s West German chairs.

INTERIOR DESIGN BY Judith Kanner
RENOVATION DESIGN BY Winston Chappell
TEXT BY Joyce Karshbath

"It was one of those constant works in progress," recalls Michael Crichton of his Los Angeles house he shares with his wife, Anne-Marie, and daughter, Taylor. "It evolved from something better rather than formal ideas." Above: The entrance hall features an antique Knoll carpet.

Figure 1: Original and segmented document into text (black), image (gray), and background (white)


bliss road trip



with a large selection of out-of-door garden fountains and outdoor furniture. Hints on House Furnishings by W. Shaw Sparrow (855) 1400 n.m. On a larger end table at Peninsula Fountain (866) Emerson St. 650-323-3151, which was opened by the Peninsula Dairy in 1923. 240 n.m. Still in the midday sun as the white-haired ladies go at the Elizabeth F. Gamble Garden Center (1431 Wilshire Street, 650-328-1268), a serene 2.3-acre community horticultural foundation. 230 n.m. Stanford's I. B. & B. Gerald Center Center for Visual Arts (Museum Way, 650-723-4177), boasts a Rodin sculpture garden as well as galleries for artfairs from Africa, the ancient Mediterranean, and Oceania. While on campus, swing by the Track House Sport Shop (Campus Drive, 650-327-8877) for logo T-shirts and sweatshirts. 330 n.m. As you drive from Palo Alto to Menlo Park, stop at Bleu Citron (205 El Camino Real, 650-323-7488), a canary-yellow shop filled with understated modern housewares by Philippe Starck and Ron Arad. 400 n.m. Craftsmen work in Spanish-style bungalows at the Allied Arts Guild (75 Arroyo Road, 650-322-2405). The artisans here make everything from dog topiaries to labrettop fountains. 430 n.m. Menlo Park could be Pleasanton but for exceptional stores like Folk Art International (971 Santa Cruz Avenue, 650-328-9989), which is filled with Day of the Dead candeholders from Puebla, Mexico, tin rattlers from Bolivia, and Henry Moore sculptures (not for sale) from the owners' collection. Fume! (719 Santa Cruz Avenue, 650-327-8900) specializes in Asian antiques. 800 n.m. Twilight is the perfect time to drive around Atherton (population: 7,500, which is 100% of the richest town in America. Most houses (median price is \$11.5 million) are custom well-furnished gems, making the town seem like a laid-back Bel-Air. For serious shopping, check out the real estate listings for open houses in the low Palo Alto Valley. 7:00 n.m. Set in what feels like a modernist Latin Tulum lodge, Wild Here (1102 El Camino Real, 650-327-4273) is the epitome of California casual: locals show up in shorts to eat \$20 wood-grilled veg chips. 9:00 n.m. Menlo Park doesn't shut down until a week. Regler's (1010 El Camino Real, 650-324-4231), an independent bookstore, stays open till 11:00 during the week. Take your bedtime reading next door to Cafe Borrone (650-327-0930) and have a beer under the stars before logging off for the day—03.

Guy Chaddock & Co.
Country French and English Antique Reproductions
Brochure in the Public • Catalog to the Trade
(661) 395-2940 • Fax: (661) 395-9970
2201 East Brundage Lane, Burbank, California 91507
Visit our website at www.guychaddock.com

bliss road trip



with a large selection of out-of-door garden fountains and outdoor furniture. Hints on House Furnishings by W. Shaw Sparrow (855) 1400 n.m. On a larger end table at Peninsula Fountain (866) Emerson St. 650-323-3151, which was opened by the Peninsula Dairy in 1923. 240 n.m. Still in the midday sun as the white-haired ladies go at the Elizabeth F. Gamble Garden Center (1431 Wilshire Street, 650-328-1268), a serene 2.3-acre community horticultural foundation. 230 n.m. Stanford's I. B. & B. Gerald Center Center for Visual Arts (Museum Way, 650-723-4177), boasts a Rodin sculpture garden as well as galleries for artfairs from Africa, the ancient Mediterranean, and Oceania. While on campus, swing by the Track House Sport Shop (Campus Drive, 650-327-8877) for logo T-shirts and sweatshirts. 330 n.m. As you drive from Palo Alto to Menlo Park, stop at Bleu Citron (205 El Camino Real, 650-323-7488), a canary-yellow shop filled with understated modern housewares by Philippe Starck and Ron Arad. 400 n.m. Craftsmen work in Spanish-style bungalows at the Allied Arts Guild (75 Arroyo Road, 650-322-2405). The artisans here make everything from dog topiaries to labrettop fountains. 430 n.m. Menlo Park could be Pleasanton but for exceptional stores like Folk Art International (971 Santa Cruz Avenue, 650-328-9989), which is filled with Day of the Dead candeholders from Puebla, Mexico, tin rattlers from Bolivia, and Henry Moore sculptures (not for sale) from the owners' collection. Fume! (719 Santa Cruz Avenue, 650-327-8900) specializes in Asian antiques. 800 n.m. Twilight is the perfect time to drive around Atherton (population: 7,500, which is 100% of the richest town in America. Most houses (median price is \$11.5 million) are custom well-furnished gems, making the town seem like a laid-back Bel-Air. For serious shopping, check out the real estate listings for open houses in the low Palo Alto Valley. 7:00 n.m. Set in what feels like a modernist Latin Tulum lodge, Wild Here (1102 El Camino Real, 650-327-4273) is the epitome of California casual: locals show up in shorts to eat \$20 wood-grilled veg chips. 9:00 n.m. Menlo Park doesn't shut down until a week. Regler's (1010 El Camino Real, 650-324-4231), an independent bookstore, stays open till 11:00 during the week. Take your bedtime reading next door to Cafe Borrone (650-327-0930) and have a beer under the stars before logging off for the day—03.

Guy Chaddock & Co.
Country French and English Antique Reproductions
Brochure in the Public • Catalog to the Trade
(661) 395-2940 • Fax: (661) 395-9970
2201 East Brundage Lane, Burbank, California 91507
Visit our website at www.guychaddock.com

Figure 2: Original and segmented document into text (black), image (gray), and background (white)