A novel pixon-representation for image segmentation based on Markov random field
Lei Lin\textsuperscript{a,b,1}, Litao Zhu\textsuperscript{b,1}, Faguo Yang\textsuperscript{b}, Tianzi Jiang\textsuperscript{b,∗}
\textsuperscript{a} Department of Mathematics, Zhejiang University, Hangzhou 310027, PR China
\textsuperscript{b} National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, PR China

\textbf{A R T I C L E  I N F O}

Article history:
Received 6 June 2004
Revised 27 March 2008
Accepted 24 April 2008

Keywords:
Image segmentation
Pixon-representation
Markov random field
Region labeling

\textbf{A B S T R A C T}

In this paper, a pixon-based image representation is proposed, which is a set of disjoint regions with variable shape and size, named pixon. These pixons combined with their attributes and adjacencies construct a graph, which represents the observed image. A Markov random field (MRF) model-based image segmentation approach using pixon-representation is then proposed. Compared with previous work on region-based and pixon-based segmentation methods, the present method has some remarkable improvements over them. Firstly, a set of significant attributes of pixons and edges are introduced into the pixon-representation. These attributes are integrated into the MRF model and the Bayesian framework to obtain a weighted pixon-based algorithm. Secondly, a criterion of GOOD pixon-representation is presented and a fast QuadTree combination (FQTC) algorithm is proposed to extract the good pixon-representation. The experimental results demonstrate that our pixon-based algorithm performs fairly well while reduces the computational cost sharply compared with the pixel-based method.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{In this paper, our approach belongs to clustering-based approach under Bayesian framework is then described on the pixon-representation approach under Bayesian framework.}
\end{figure}

\section{1. Introduction}

Image segmentation is the process of segmenting an image into a set of disjoint regions which have similar characteristics such as intensity, color, texture, etc. Image segmentation is a crucial low-level processing and a common problem in many fields like computer vision. It has been studied for decades and several approaches had been developed. These approaches can be roughly classified into three classes: (i) edge-based approaches, which rely on edges found in an image by edge detecting operators, such as Canny method\cite{26}, (ii) region-based approaches, which group pixels into homogeneous regions and segment the image to some major areas (Fig. 1b), such as normalized cut\cite{20} and region growing\cite{21}, (iii) clustering-based methods, which segment the feature space of image to several clusters and get a sketch of the original image\cite{23} (Fig. 1c), such as mean shift\cite{22} and fuzzy C-means (FCM) clustering. The differences of these methods are showed in Fig. 1. In this paper, our approach belongs to clustering-based method.

In general, there are two main criteria to be considered in image segmentation: one is the homogeneity of the region and the other is the discontinuity between adjacent regions. However, due to the low resolution, the lack of contrast, the presence of various image noises, etc., it is difficult to get the satisfactory segmentation result without the prior knowledge. For example, the noise model and the probability distribution of real image are adopted in the Bayesian segmentation framework\cite{3,5,24,25}; the continuity of objects’ boundary is introduced in the deformable model-based methods\cite{6,7}; the continuity of the image characteristics is utilized in the Markov random field (MRF) model-based approaches\cite{11,12,24,25}.

In this paper, we propose a pixon-representation for image segmentation, which is a set of disjoint regions with variable shape and size, named pixon. These pixons combines with their attributes and adjacencies construct a graph, which represents the observed image. Since the characteristics such as intensity, color, texture, etc., of pixels in each pixon are similar, an assumption is taken into account that the pixels in one pixon belong to the same class in the segmented image. A MRF model-based image segmentation approach under Bayesian framework is then described on pixon-representation.

The pixon concept has been introduced by Pina and Puetter\cite{1–3} to restore astronomy images. The pixon-based method draws on the idea that images can be concisely described using fewer degrees-of-freedoms where the characteristics of image are smooth and using more degrees-of-freedoms where there is greater detail. The size of pixon defines locally the scale of underlying image information. Pina and Puetter adopted fuzzy pixons in image restoration, which model the image by the local convolution of a pseudo-image with a blurring function with a local scale. This pixon model was also used by X. Desombes\cite{4} for low-level image description. In this paper, we focus on image segmentation with the assumption that the observed image is not blurred because blurring will smooth the image and then decrease the degrees-of-freedoms in the detail. So we adopt the pixon definition...
introduced by F. Yang et al. [5] rather than fuzzy pixon. The pixon-representation of image is then systematically in this paper. A Fast QuadTree Combination (FQTC) algorithm is also proposed to extract the GOOD pixon-representation.

Markov random field introduced by Besag [8] is widely used in image segmentation, restoration and interpretation because of its powerful capability of describing the continuity of image characteristics [10–19]. Besides the pixon-based MFR model in [5], the region-based MRF model is introduced in [10–12] to reduce the computational cost. Compared with previous work, an image representation based on pixon is proposed systematically and a Fast QuadTree Combination algorithm is described to extract the good representation with respect to a discriminant. As a result, the size and shape of pixons can be controlled by adjusting the discriminant and arranging the order of the pixon combinations in FQTC algorithm. Moreover, the attributes of pixons and edges, such as the size of pixon, the mean and variance of pixon intensity, the length of boundaries between adjacent pixons, are integrated into the pixon-based MRF model and the Bayesian framework. These improvements remarkably reduce the computational cost while guarantee the robust segmented result.

This paper is organized as follows. In Section 2, we introduce the pixon-representation of image and the shortest pixon-representation with respect to a discriminant. The fast QuadTree combination algorithm to extract the good pixon-representation with respect to a discriminant will be described in Section 3. Then Section 4 is devoted to the MRF model-based segmentation approach under Bayesian framework, which is performed on pixon-representation. Finally, the experimental results and the conclusion are given in Section 5.

2. Definition of pixon-representation

Pina and Puetter [1–3] introduced the fuzzy pixon,

\[
I(\tilde{x}) = (K \ast I_{\text{pseudo}})(\tilde{x}) = \int d\tilde{y} K(\tilde{x}, \tilde{y}) \tilde{y}
\]

where the image is the local convolution of a pseudo-image, \(I_{\text{pseudo}}\), with a blurring function, \(K\), with a local scale, \(\tilde{y}\).

Focusing on image segmentation with the assumption that the observed image is not blurred, we introduce the pixon-representation and propose a fast approach to extract a good pixon-representation for an observed image.

2.1. Definition of pixon-representation

**Definition 1.** Let \(X = \{x_i\}_{i=1}^N\) be the set of all the image pixels. A subset of \(X\) is a pixon if and only if all the pixels in it are connected. A pixon is then denoted by \(P_i = \{x_i\}_{j=1}^N\).

An attribute vector of the pixon is extracted from the observed image,

\[
\tilde{p}_i = (n_i, b_i, \max_{i} \min_{i} \mu_i, \sigma_i^2),
\]

where \(n_i\) is the number of pixels in \(P_i\), \(b_i\) is the perimeter of \(P_i\), namely the length of the boundary between \(P_i\) and the other part of the observed image, \(\max_i, \min_i, \mu_i\) and \(\sigma_i^2\) are the maximum, minimum, mean and variance of the observed image intensities in \(P_i\), respectively. Let \(I(x)\) denote the image intensity on the pixel \(x\). The attributes of the pixon intensity can be obtained by

\[
\begin{align*}
\max & = \max_i (I(x_i)|x_i \in P_i), \\
\min & = \min_i (I(x_i)|x_i \in P_i), \\
\mu_i & = \sum_{i=1}^N I(x_i)/n_i, \\
\sigma_i^2 & = \sum_{i=1}^N (I(x_i))^2/n_i - \mu_i^2.
\end{align*}
\]

**Definition 2.** A set of pixons, \(P = \{P_i\}_{i=1}^N\), is a pixon-representation if and only if

\[
\begin{align*}
(1) & \quad P_i \neq \emptyset, \\
(2) & \quad P_i \cap P_j = \emptyset, \text{ if } i \neq j, \\
(3) & \quad \bigcup_{i=1}^N P_i = X.
\end{align*}
\]

The above definition shows that the pixon-representation segments the image into a set of disjoint regions. A set of edges, \(E\), can be acquired from these regions,

\(E = \{e_{ij} | P_i \in P \text{ and } P_i \cap P_j \text{ are adjacent}\}\)

where \(P_i\) and \(P_j\) are adjacent if \(3x_i \in P_i\) and \(x_j \in P_j\), which are neighboring pixels to each other in the image.

The strength of an edge can be defined as the length of the boundary between the two adjacent pixons, which is denoted by \(b_{ij}\), so \(b_{ij} = \sum b_{ij}\). The attributes similar to Boundary Process in [11] can also be defined to describe the image intensity on the boundary. We use a attribute vector, \(e_{ij}\), to denote all the attributes of an edge.

The pixons and edges, combined with their attribute vectors, construct a graph, \(G = (P, E)\), which represents the observed image, as shown in Fig. 2.

2.2. Shortest pixon-representation with respect to a discriminant

There are two trivial pixon-representations, \(P_0 = \{X\}\) and \(P_0 = \{\{x_i\}|x_i \in X\}\). The former takes all the image pixels as one pixon, the latter takes each pixel as a pixon, which is a lossless representation. In order to represent the image using as few pixons as possible while limiting the representation error, the shortest pixon-representation with respect to a discriminant is defined.
Definition 3. A function $f(P) \geq 0$ of pixons is a pixon error function if and only if

1. $f(P) = 0$, if $P = \{x_i\}$,
2. $f(P_i) \leq f(P_j)$, if $P_i \supset P_j$.

Definition 4. For a given pixon error function, $f(\cdot)$, and a non-negative constant, $T$, the inequality, $f(\cdot) \leq T$, defines a pixon discriminant.

Definition 5. A pixon-representation is called the shortest pixon-representation with respect to a given discriminant, $f(\cdot) \leq T$, if its number of pixons is least among all the pixon-representation satisfying $\forall P_i \in P, f(P_i) \leq T$.

In general, using the pixon attribute vector to describe the region of the observed image will lose some information, so we use the pixon error function to denote the error between the pixon and the region of the observed image. In this paper we define error function to denote the error between the pixon and the region of the observed image.

3. Extraction of pixon-representation

The shortest pixon-representation with respect to a discriminant is not unique, as shown in Fig. 3. And it is hard to extract the shortest one from a large and complex image. In this section, an approach to extract a GOOD pixon-representation is proposed, which combines the adjacent pixons of the lossless pixon-representation, $P_i = \{x_i | x_i \in X\}$, iteratively, until no pixons can be combined considering the given discriminant. The obtained good pixon-representation is dependent on the order of combination besides the discriminant.

3.1. Combination of adjacent pixons

The adjacent pixons in a pixon-representation, $G = \{P, E\}$, can be combined to form a new pixon, denoted by $P_{\text{new}} = P_i \oplus P_j$, whose attribute vector, $p_{\text{new}}$, can be obtained from $p_i$ and $p_j$.

\[ n_{\text{new}} = n_i + n_j, \]
\[ b_{\text{new}} = b_i + b_j - 2b_{ij}, \]
\[ \max_{\text{new}} = \max(\max_{i}, \max_{j}), \]
\[ \min_{\text{new}} = \min(\min_{i}, \min_{j}), \]
\[ \rho_{\text{new}} = (n_i\rho_i + n_j\rho_j)/n_{\text{new}}, \]
\[ \sigma^2_{\text{new}} = [n_i(\sigma^2_i + \mu^2_i) + n_j(\sigma^2_j + \mu^2_j)]/n_{\text{new}} - \rho^2_{\text{new}}, \]

where $b_{ij}$ is the edge strength, i.e. the length of the boundary between $P_i$ and $P_j$.

It can be proved that $P - \{P_i, P_j\} + \{P_{\text{new}}\}$ is still a pixon-representation. And the edge set of the new pixon-representation can be obtained from $E$ by combining the edges connecting the same two pixons after the pixon combination.

3.2. Combination-based extraction of pixon-representation

Given a discriminant, $f(\cdot) \leq T$, we define the edge error function as $f(E) = f(P_i \oplus P_j)$. Since $P_i = \{x_i | x_i \in X\}$ satisfies all the discriminants, the shortest pixon-representation with respect to $f(\cdot) \leq T$ can be extracted by combining the pixons of $P_i$, the lossless representation, until all error function values of the edges are larger than $T$.

In fact, the pixon-representation obtained by combination scheme may not always be the shortest, which is dependent on the order of combinations. However, it is a substitute to the shortest, for the number of pixons has been sharply cut down.

3.3. Fast QuadTree combination algorithm

A fast QuadTree combination algorithm is proposed to extract the shortest pixon-representation here. Firstly, a QuadTree-based multi-resolution pixon-representation is constructed, as shown in Fig. 4. Then a initial pixon-representation with respect to $f(\cdot) \leq T_0$, $T_0 \in [0, T]$, is extracted by coarse-to-fine selecting a set of disjoint squares from the multi-resolution pixon-representation, which satisfy $f(\cdot) \leq T_0$. Finally, the pixons connected by the edge with the minimal edge error are combined iteratively, until the minimal edge error is larger than $T$.

If the image region is not a square whose edge length is the power of 2, the multi-resolution pixon-representation can be constructed as follows. Firstly, the image is put into a large enough square like (a) in Fig. 4. For each scale, the pixon is then defined as the set of pixels falling into a square of this scale; the squares including no pixel are ignored. An example using the fast QuadTree combination algorithm is given in Fig. 5, where the error function is defined as $f(P) = \max - \min$.

4. Image segmentation based on pixon-representation

A Markov random field model-based image segmentation approach under Bayesian framework is proposed based on pixon-representation. Compared with Yang’s method [5], the noise model of the Bayesian framework in our approach is based on the pixel intensity rather than the mean of pixel intensity. As a result, the size of pixon and the variance of the pixon intensity are integrated into the computation of the posterior probability. Further more, the normalized strength of edges and the size of pixon are adopted to form a weighted MRF model.
4.1. Bayesian framework

Let \( I \) be the observed image and \( S \) be the segmented image. In the Bayesian segmentation framework, the segmented image is obtained by maximizing the posterior probability,

\[
S^* = \arg \max_S P(S|I)
\]

where

\[
P(S|I) \propto P(I|S)P(S).
\]

We assume \( I = S + N \), where \( N \) is independent Gaussian white noise. Then the conditional probability is

\[
P(I|S) = \prod_{k=1}^{K} \prod_{x \in I_k} \frac{1}{\sqrt{2\pi\sigma_k}} \exp \left( -\frac{(l(x_k) - u_k)^2}{2\sigma_k^2} \right)
\]

where \( K \) is the number of classes, \( I_k \) is the set of pixels segmented into the \( k \)-th class, and \( u_k \) is the intensity mean of pixels in \( I_k \). Let \( \mathbf{G} = \{ \mathbf{P}, \mathbf{E} \} \) be a pixon-representation of \( I \). Since the characteristics of pixels in each pixon are similar, we assume that the pixels in one pixon will be segmented into the same class. So using (2) and (6), we get

\[
P(I|S) = \prod_{k=1}^{K} \prod_{x \in I_k} \prod_{y \in I_k} \frac{1}{\sqrt{2\pi\sigma_k}} \exp \left( -\frac{(l(x_k) - u_k)^2}{2\sigma_k^2} \right)
\]

The computation of \( P(I|S) \) is simplified since the number of pixons is far less than that of pixels.

\( P(S) \) is the prior probability. In this paper, we adopt the MRF model based on the pixon-representation to define the prior probability distribution as follows.

4.2. MRF model based on pixon-representation

A neighborhood system of the graph, \( \mathbf{G} = \{ \mathbf{P}, \mathbf{E} \} \), is defined as

\[
N(\mathbf{P}) = \{N(P_i)|P_i \in \mathbf{P}\}
\]

where

\[
N(P_i) = \{P_j|x_k \in \mathbf{E}\}, \quad 1 \leq i \leq N
\]

is the neighborhood of each pixon.

Let \( \Lambda = \{\lambda_1, \ldots, \lambda_k\} \) be the set of possible labels denoting the classes in the segmented image and \( L = \{l_1, \ldots, l_k\} \) be a family of random variables where \( l_i \in \Lambda \) denotes the label of \( i \)-th pixon and \( N \) is the number of pixons. The segmented image \( S \) can then be described by the event, \( L = \omega \), since we assume that the pixels in one pixon will be segmented into the same class.

Let \( \Omega \) be the set of all possible configurations, \( \Omega = \{\omega = (\omega_1, \ldots, \omega_N)|\omega_i \in \Lambda, 1 \leq i \leq N\} \). \( L \) is a MRF with respect to the neighborhood, \( N(\mathbf{P}) \), if

\[
\begin{align*}
(1) & \quad P(L = \omega) > 0, \quad \forall \omega \in \Omega \\
(2) & \quad P(l_i = \omega_i|l_j = \omega_j, P_j \neq P_i) = P(l_i = \omega_i|l_j = \omega_j, P_j \in N(P_i)), \\
& \quad \forall P_i \in \mathbf{P} \text{ and } \omega \in \Omega
\end{align*}
\]

where \( P(\cdot|\cdot) \) and \( P(\cdot) \) are the joint and conditional probability density functions, respectively.

It has been proved in [9] that the configurations of MRF obey a Gibbs distribution

\[
P(\omega) = \frac{1}{Z} \exp(-U(\omega)/T)
\]

where \( Z \) is a normalizing constant and \( T \) is a constant called temperature. \( U(\omega) \) is the energy function, which is a sum of clique potentials \( V_c(\omega) \) on all possible cliques, i.e.

\[
U(\omega) = \sum_{c \in \mathbf{C}} V_c(\omega).
\]

In this paper, the set of cliques is defined as \( \mathbf{C} = \{\mathbf{c}_i|\mathbf{c}_i = \{P_i\} \cup N(P_i), P_i \in \mathbf{P}\} \), where each pixon in \( \mathbf{G} = \{\mathbf{P}, \mathbf{E}\} \) defines one clique. The clique potential is defined by

\[
V_c(\omega) = w_c(\overline{\mathbf{p}}) \sum_{P_i \in N(P)} \frac{1}{b_i} \sum_{b_j \in b_i} \frac{1}{b_j} \eta_{ij} \frac{n_{ij}}{|\mu_i - \mu_j|} \exp \left( -\frac{1}{T} \sum_{P_k \notin N(P)} n_{ik} \frac{b_i}{b_k} \frac{n_{ik}}{|\mu_i - \mu_k|} \right)
\]

where \( \eta_{ij} \) is a binary variable which has the value 1 if \( P_i \) and \( P_j \) have the same label and the value 0 otherwise; \( w_c(\overline{\mathbf{p}}) = n_i \) is the clique weight; \( w_c(b_i, b_j, b_k) = b_i/b_j \) is the normalized edge weight; and \( w_c(\overline{\mathbf{p}}, \overline{\mathbf{p}}) = 1/|\mu_i - \mu_j| \) is the pixon distance weight that denotes the difference of image characteristics between two pixons.

In all, the prior probability is defined as

\[
P(S) = P(\omega) = \frac{1}{Z} \exp \left( -\frac{1}{T} \sum_{P_j \notin N(P)} n_{ij} \frac{b_i}{b_k} \frac{n_{ij}}{|\mu_i - \mu_k|} \right)
\]

4.3. Optimization

From (4) and (5), the optimal segmented image can be written as

\[
S^* = \arg \min_S (-\ln P(I|S) - \ln P(S)).
\]

Using (7) and (13), the objective function is then obtained,

\[
F(S) = F(\omega) = \sum_{k=1}^{K} \sum_{x \in I_k} n_{ik} \ln \left( \frac{b_i}{b_k} \frac{n_{ik}}{|\mu_i - \mu_k|} \right)
\]
where $\alpha = 1/T$ is a weight of MRF model, which denotes the tradeoff between the fidelity to the observed image and the smoothness of the segmented image. The constant term has been removed from the objective function.

The class number $K$ and the weight $\alpha$ are given before optimization. The initial segmented image is obtained using Fuzzy C-Means (FCM) clustering, and the initial parameters of each class are estimated from the initial segmented image, i.e. the means $u_k$ and variances $\sigma_k$. Then we compute the threshold $T$, the value of $T$ should not be too large, otherwise the pixon will contain many pixels which actually belong to two different classes. So we using follow empirical function:

$$T = \min_{0 \leq k < K, x \in \mathcal{I}} |u_k - u_i| - \sigma_i - \sigma_j.$$

Finally, the segmented image and the parameters are optimized, simultaneously.

Let $F(\omega, \omega_{\text{new}})$ denote the objective function value when the ith label of $\omega$ is changed into $\omega_{\text{new}}$ and $\Delta F(\omega, \omega_{\text{new}})$ denote $F(\omega, \omega_{\text{new}}) - F(\omega)$. The optimization is described as follows:

1. Initialize the number of classes $K$; the total number of iteration $\text{NUM}$; $u_1, u_2, \ldots, u_K$ and $\sigma_1, \sigma_2, \ldots, \sigma_K$ according to an initial segmentation, which is obtained using FCM method; compute the threshold $T$; and the iteration index $j = 0$.
2. Extraction of pixon-representation, then initialize the pixon-based image model: assign a label $\hat{\omega}_i$ to each pixon $P$, which minimizes the expression $|u_P - u_i|$.
3. Find the best label for each pixon, $\hat{\omega}_{\text{best}}$, $1 \leq i \leq \text{NUM}$, which minimizes $F(\omega, \omega_{\text{best}})$.
4. Find the $\Delta F(\omega, \omega_{\text{best}})$, satisfying
   $$\Delta F(\omega, \omega_{\text{best}}) \leq \Delta F(\omega, \omega_{\text{new}}), 1 \leq i \leq \text{NUM}.$$
5. If $\Delta F(\omega, \omega_{\text{best}}) < 0$ and $j < \text{NUM}$, go to step 4, otherwise stop iteration.
6. Update the best label of each pixon and re-estimate $u_i$, $\sigma_i$ using new $\omega_i$.
7. $j = j + 1$, Go to step 3.

In fact, $\Delta F(\omega, \omega_{\text{best}})$ can be calculated using the correlative terms with the ith label in $F(\omega)$, i.e.

$$F_j(\omega, \omega_{\text{new}}) = n_j\left(\frac{(u_j - u_{\text{new}, j})^2}{\sigma_{\text{new}, j}^2} + \ln \sigma_{\text{new}, j}\right) + \alpha \sum_{k \in (j)} \left(\frac{\bar{n}_{j\bar{k}}}{b_{j\bar{k}}^2} + \frac{\bar{n}_{j\bar{k}}}{b_{j\bar{k}}^2}\right) \frac{\bar{n}_{\text{new}, j\bar{k}}}{|\mu_i - \mu_j|}.$$

5. Experimental results and conclusions

5.1. Synthetic images

Because of the noise model of the Bayesian framework in our approach is based on the pixel intensity rather than the mean intensity of pixon intensity, our approach would theoretically more robust to noise than Yang’s method. Our approach also has the same time complexity as Yang’s method, so it should have less computational time than traditional MRF method. We have set up experiments using synthetic images to qualitatively evaluate our method.

We create a synthetic image whose size is $256 \times 256$ and have three classes with different intensity values (Fig. 6a). We add the Gauss noise with signal-to-noise ratio (SNR) from 0 to 8 dB to the synthetic image get nine observes images (Fig. 6b). Then we apply our segmentation method to these images, as shown in Fig. 6f. Here, the pixon error function is defined as $f(P) = \max_{L} - \min_{L}$. The number of class, $K = 3$, and the weight of MRF model, $\alpha = 200$, NUM = 100 are used for all the images.

The true segmentation result is the synthetic image, and for evaluation, we define the errors of a result is the number of the pixels which are assigned wrong labels. The errors of each method’s results are show in Fig. 8a.

Fig. 6. The experimental results based on Pixon-Representation in synthetic images. (a) The synthetic image. (b) The observe image with SNR = 4dB. (c) The segmented result of FCM method. (d) The segmented result of traditional MRF method. (e) The segmented result of Yang’s method. (f) The segmented result of our method.
5.2. Real images

We apply our MRF segmentation method based on pixon-representation to three classes of images, as shown in Fig. 7. Here, the pixon error function is defined as \( f(P_i) = \max \frac{e_i}{c_0} - \min \frac{e_i}{c_0} \). The number of class, \( K = 3 \), and the weight of MRF model, \( \alpha = 200 \), \( \text{NUM} = 100 \) are used for all the images.

Since the lack of the gold standard for quantitatively evaluation of our experimental results, we used the function proposed in [27], defined as:

\[
Q(S) = \frac{1}{10000(N \times M)} \sqrt{R} \times \sum_{i=1}^{R} \left[ \frac{e_i^2}{1 + \log A_i + \left( \frac{R(A_i)}{A_i} \right)^2} \right]
\]

Fig. 7. The experimental results based on Pixon-Representation. The first column is the observed images; the second column is their Pixon-Representations extracted by the FQTC algorithm, where the green lines denote the boundaries between Pixons; the third column is the initial segmented image obtained by FCM clustering on the Pixon-Representation; the last column is the final segmented results after optimization.

Fig. 8. The experimental results of synthetic images. (a) Errors of each experimental results, and (b) is the computational cost.
where S is the result to be evaluated, N × M is the image size, R is the number of regions, A_i is the size of the i-th region, e_i is the average error of the i-th region and R(A_i) is the number of regions having exactly area A_i. e_i is defined as the sum of Euclidean distances between the features extracted from each pixel and those for the corresponding cluster. The smaller value of Q(S), the better segmentation result should be. More details can be found in [27] [28].

5.3. Results and conclusions

The experimental results show that our algorithm is robust to noise as shown in Fig. 8a and Table 2. At the same time, the computational cost is sharply reduced by using pixon-representation as shown in Fig. 8b and Table 3. The number of pixons is 8.20–23.31% of that of pixels and the mean number of each pixon’s neighbors is less than 5 when 4-neighbor connectivity of pixels is used (Table 1).

The results show our approach get similar errors and smaller errors and Q(S) than Yang’s method (Fig. 8a and Table 2), because we use more information of image such as the Pixel intensity the normalized strength of edges and the size of pixon rather than the Yang’s method. Since we use the same model and the relatively small number of pixons to replace the pixels, the computational cost is sharply reduced by using our approach as well as Yang’s method (Fig. 8b Table 3). But because the neighbor of each pixons is always more than 4, and extraction of pixon-representation needs some time, the final ratio between the cost of our approach and that of the traditional MRF method is a little more than the ratio between the number of pixons and pixels.

The results also indicate the weakness of our algorithm, i.e. it is hard to differ the image details from the noises. For example, the face of Baboon covered by hair is not correctly segmented in the second row of Fig. 7. This problem can be solved by integrating the texture of image into our method. Generally, more characteristics and prior knowledge are considered, more satisfactory results can be obtained.

The FQTC algorithm and MRF model-based approach proposed in this paper are both based on the graph, so they can be extended to 3-D directly. And we can choose various pixel connectivity systems to extract pixon-representation, besides the 4-neighbor connectivity system.

In summary, we have proposed a novel image representation based on pixon. A criterion of GOOD pixon-representation has also been presented and a fast QuadTree combination (FQTC) algorithm has been proposed to extract the good pixon-representation. Then the proposed image representation has been applied to image segmentation based on MRF model. The experimental results have demonstrated that our pixon-based algorithm has good performance while reduces the computational cost sharply compared with the pixel-based method.

Acknowledgements

This work was partially supported by the Natural Science Foundation of China, Grant Nos. 30425004, 30730035, and the National Key Basic Research and Development Program (973) Grant No. 2003CB716100.

Table 1

<table>
<thead>
<tr>
<th>Image</th>
<th>The size of image</th>
<th>The number of pixons</th>
<th>The mean number of each pixon’s neighbors</th>
<th>The ratio between the number of pixons and pixels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepper image</td>
<td>512 × 512 = 262144</td>
<td>21494</td>
<td>4.671</td>
<td>8.20%</td>
</tr>
<tr>
<td>Baboon image</td>
<td>512 × 512 = 262144</td>
<td>55853</td>
<td>4.953</td>
<td>23.31%</td>
</tr>
<tr>
<td>Cortex image</td>
<td>128 × 128 = 16384</td>
<td>1691</td>
<td>4.704</td>
<td>10.32%</td>
</tr>
</tbody>
</table>

Table 2

<table>
<thead>
<tr>
<th>Image</th>
<th>Cost of FCM (ms)</th>
<th>Cost of traditional MRF (ms)</th>
<th>Cost of Yang’s method (ms)</th>
<th>Cost of our method (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepper image</td>
<td>9987</td>
<td>60156</td>
<td>16143</td>
<td>17034</td>
</tr>
<tr>
<td>Baboon image</td>
<td>9596</td>
<td>111385</td>
<td>18549</td>
<td>19924</td>
</tr>
<tr>
<td>Cortex image</td>
<td>306</td>
<td>4714</td>
<td>712</td>
<td>697</td>
</tr>
</tbody>
</table>

Table 3

<table>
<thead>
<tr>
<th>Image</th>
<th>FCM method</th>
<th>Traditional MRF method</th>
<th>Yang’s method</th>
<th>Our method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pepper image</td>
<td>4067</td>
<td>2216</td>
<td>2395</td>
<td>2099</td>
</tr>
<tr>
<td>Baboon image</td>
<td>30887</td>
<td>4668</td>
<td>3979</td>
<td>1782</td>
</tr>
<tr>
<td>Cortex image</td>
<td>51</td>
<td>48</td>
<td>57</td>
<td>29</td>
</tr>
</tbody>
</table>

References


