Research Article
An Automatic Image Inpainting Algorithm Based on FCM

Jiansheng Liu, Hui Liu, Shangping Qiao, and Guangxue Yue

1 College of Science, Jiangxi University of Science and Technology, Ganzhou 341000, China
2 Graduate School, Jiangxi University of Science and Technology, Ganzhou 341000, China
3 College of Mathematics and Information Engineering, Jiaxing University, Jiaxing 314000, China

Correspondence should be addressed to Jiansheng Liu; jxgjscn@126.com

Received 10 October 2013; Accepted 5 December 2013; Published 2 January 2014

Academic Editors: J. Shu and F. Yu

Copyright © 2014 Jiansheng Liu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

There are many existing image inpainting algorithms in which the repaired area should be manually determined by users. Aiming at this drawback of the traditional image inpainting algorithms, this paper proposes an automatic image inpainting algorithm which automatically identifies the repaired area by fuzzy C-mean (FCM) algorithm. FCM algorithm classifies the image pixels into a number of categories according to the similarity principle, making the similar pixels clustering into the same category as possible. According to the provided gray value of the pixels to be inpainted, we calculate the category whose distance is the nearest to the inpainting area and this category is to be inpainting area, and then the inpainting area is restored by the TV model to realize image automatic inpainting.

1. Introduction

Inpainting is a trimming process for the defects and cracks of the art, originated from the art field which was originally the heritage restoration of experts on the basis of being faithful to the original. Image inpainting is a technology for restoring the damaged parts of an image by referring to the information from the undamaged parts to make the restored image look “complete”, “continuous” and “natural”.

The terminology “digital image inpainting” was firstly put forward on the international conference in Singapore in 2000. There are many typical image inpainting algorithms proposed by researchers during the past decade. The BSCB [1] model was presented by Bertalmio, Sapiro, Caselles, and Ballester. The TV (total variation) model [2, 3] was proposed by Chan and his team. The CDD (curvature driven diffusions) model [4, 5] was introduced by Chan T. and Shen J. An isotropic diffusion of fast image inpainting was proposed by Oliverira. A model based on RBF was proposed by Zhou et al. [6]. The p-Laplace inpainting model was calculated from p-Laplace operator by Zhang et al. [7–9]. Wang et al. proposed a new spread function to improve the TV model [10], algorithm based on wavelet decomposition was proposed by Zhang and Dai [11] and so forth. However, these models belong to semi-automatic image restoration algorithms that need to manually determine the repaired area. It is a difficult problem for automatically identifying the repaired area of the image and achieving automatically image inpainting function in the field of image inpainting.

This paper attempts to overcome this drawback of the traditional image inpainting techniques to a certain degree. The proposed algorithm utilizes the fuzzy C-means (FCM) clustering algorithm to automatically identify the damaged area and also combines the TV model to realize image automatic inpainting.

2. Theoretical Basis of FCM Identify Damaged Areas

2.1. Basic Knowledge of the Fuzzy Sets. Fuzzy set theory and fuzzy logic [12] are presented by Zadeh, that is, to use precise methods, formulas, and models to measure and handle phenomena and rules which are vague, incomplete, or incorrect. Membership function is defined as the degree to which an object \( x \) belongs to a fuzzy set \( A \), usually recorded as \( \mu_A(x) \); the scopes of its argument are the all possible objects which belongs to the set \( A \) (that all the points of the space
where the set $A$ in) in the range of $[0,1]$; namely, $0 \leq u_A(x) \leq 1$. $u_A(x) = 1$ means that $x$ completely belongs to set $A$, which is equivalent to the traditional set concept $x \in A$. A membership function defined in the space of $X = \{x\}$ is defined as a fuzzy set $A$, or defined as fuzzy subset $A$ in the domain of $X = \{x\}$. For the fuzzy set $A$ which has finite objects $x_1, x_2, \ldots, x_n$ can be expressed as

$$A = \{ (u_A(x_i), x_i) | x_i \in X \}.$$  

(1)

Fuzzy set theory has played an important role in many applications, such as fuzzy clustering analysis, fuzzy pattern recognition [13], fuzzy synthetic judgments [14], fuzzy decision and forecast [15, 16], fuzzy programming, fuzzy probability [17], and fuzzy statistics [18]. In the problem of clustering, the clusters generated by clustering can be regarded as fuzzy sets; therefore, the value of membership of each sample point belongs to a cluster that is in the interval $[0,1]$.

2.2. Theory of FCM Algorithm. Fuzzy C-mean clustering (FCM) [19, 20] is a clustering algorithm that uses membership to determine the degree of each data point which belongs to a cluster. In 1973, Bezdek proposed the algorithm that is an improved algorithm of the early hard C-means clustering (HCM). The fuzzy clustering algorithm uses the initialization method to determine several initial clustering centers and then performs a repeated iteration which continuously adjusts and optimizes the clustering centers to achieve the minimal variance within clusters. It is a method based on the principle of least squares using iterative methods to optimize objective function to get the fuzzy partition of the data sets. Its objective function is defined as

$$\min I_{FCM}(U,V) = \sum_{k=1}^{c} \sum_{i=1}^{n} u_{ik}^m \| x_i - v_k \|^2,$$  

(2)

where $X = \{x_1, x_2, \ldots, x_n\} \in \mathbb{R}^{n \times p}$ is a data set that composed of $n$ samples of dimension $p$, which is divided into $c$ classes. $x_i$ is a $p$-dimensional vector; $m \in [1, \infty)$ is the weighted index; when $m = 1$, the fuzzy C-means clustering is degraded as the classical C-means clustering. Nikhil et al. studies show that the best choice for $m$ is in range of $[1.5, 2.5]$ and its ideal value is usually set to $m = 2$; $d_{ik}$ is generally defined as the Euclidean distance, it means the distance between sample $x_i$ and center $v_k$, that is, $d_{ik} = \| x_i - v_k \|_2$; $\| \cdot \|_2$ defines as norm of $A$, usually $A = 2$; that is, $d_{ik}$ is 2-norm; $V = \{v_1, v_2, \ldots, v_c\} \in \mathbb{R}^{c \times p}$ is the set of $c$ cluster centers. $v_i$ is $p$-dimensional vector and denotes the ith cluster center; $u_{ik}$ is the membership of the $k$th pixel that belongs to $i$th class. Therefore, the classification results can be expressed as a $c \times n$-order fuzzy classification matrix $U$:

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{c1} & u_{c2} & \cdots & u_{cn} \end{bmatrix},$$  

(3)

where $u_{ik}$ need to satisfy the following constraints:

(1) $u_{ik} \in [0,1], \forall i,k$;

(2) $\sum_{i=1}^{c} u_{ik} = 1, \forall k$;

(3) $0 < \sum_{i=1}^{c} u_{ik} < n, \forall i$.

FCM is to search the optimum number of $(U,V)$ for minimizing the $I_{FCM}(U,V)$. Due to the constraint $\sum_{i=1}^{c} u_{ik} = 1$, it introduces a Lagrange multiplier $\lambda$ to solve the minimum of the following formula:

$$J(U,V) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^m \| x_i - x_k \|^2 + \lambda \left( \sum_{i=1}^{c} u_{ik} - 1 \right),$$  

(5)

where

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left( \| x_i - v_j \|^2 / \| x_k - v_j \|^2 \right)^{1/(m-1)}}, \forall i, k,$$

(6)

So the implementation of FCM algorithm is as follows:

Step 1. Select $\varepsilon > 0$, set the initial clustering centers $V^{(0)} = \{v_1, v_2, \ldots, v_c\}$, and set the number of iteration $l = 1$.

Step 2. The elements of membership matrix $U^{(l)}$ are calculated as

$$u_{ik}^{(l)} = \left\{ \sum_{j=1}^{c} \left( \| x_i - v_j \|^2 / \| x_k - v_j \|^2 \right)^{2/(m-1)} \right\}^{-1}, \forall i, k.$$  

(7)

If $d_{kj}^{(l)} = \| x_i - v_j \|^2 = 0$, then $u_{ik}^{(l)} = 1$, $u_{kj}^{(l)} = 0, k \neq j$.

Step 3. Calculate the new cluster center set $V^{(l+1)}$, whose elements are

$$V^{(l+1)} = \frac{\sum_{k=1}^{n} \left( u_{ik}^{(l)} x_i \right) x_k}{\sum_{k=1}^{n} \left( u_{ik}^{(l)} x_k \right)^m}, \forall k.$$  

(8)

Step 4. If $\| V^{(l+1)} - V^{(l)} \| < \varepsilon$, then stop; otherwise $l = l + 1$ and go to Step 2.
The length $h = 1$ and then solve the approach values of $v^1_e$, $v^1_w$, $v^2_n$, and $v^2_s$:

\[
\begin{align*}
  v^1_e &= \frac{1}{\|\nabla u_{\|e}\|} \left[ \frac{\partial u}{\partial x} \right]_e = \frac{1}{\|\nabla u_{\|e}\|} \frac{u_E - u_O}{h}, \\
  |\nabla u|_e &= \frac{1}{h} \sqrt{e^2 + (u_E - u_O)^2} + \left( \frac{u_{NE} - u_{SE} + u_N - u_S}{4} \right).
\end{align*}
\]  

(14)

Similarly we can get

\[
\begin{align*}
  v^1_w &= \frac{1}{\|\nabla u_{\|w}\|} \left[ \frac{\partial u}{\partial x} \right]_w = \frac{1}{\|\nabla u_{\|w}\|} \frac{u_W - u_O}{h}, \\
  v^2_n &= \frac{1}{\|\nabla u_{\|n}\|} \left[ \frac{\partial u}{\partial y} \right]_n = \frac{1}{\|\nabla u_{\|n}\|} \frac{u_N - u_O}{h}, \\
  v^2_s &= \frac{1}{\|\nabla u_{\|s}\|} \left[ \frac{\partial u}{\partial y} \right]_s = \frac{1}{\|\nabla u_{\|s}\|} \frac{u_S - u_O}{h}.
\end{align*}
\]  

(15)

So

\[
\begin{align*}
  -\nabla \cdot \left[ \frac{\nabla u}{\|\nabla u\|} \right] &= -\nabla \cdot v = (v^1_w - v^1_e) + (v^2_s - v^2_n) \\
  &= \sum_{Q \in \Lambda \cap q \in \Lambda} \frac{1}{\|\nabla u_Q\|} \left[ u_Q - u (i, j) \right],
\end{align*}
\]  

(16)

where $Q \in \Lambda \cap O$, $q \in \Lambda$ means that when $Q$ is a pixel of $\Lambda \cap O$, then $q$ is the corresponding half pixel of $\Lambda$; as $Q$ take $E$, correspondingly $q$ take $e$. Formula (12) can be remarked as

\[
\begin{align*}
  u (i, j) &= \sum_{Q \in \Lambda \cap q \in \Lambda} \frac{1}{\|\nabla u_Q\|} \left[ u_Q - u (i, j) \right] \\
  &\quad + \lambda_D (i, j) (u_0 (i, j) - u (i, j)) = 0.
\end{align*}
\]  

(17)

So

\[
\begin{align*}
  u (i, j) &= \sum_{Q \in \Lambda \cap q \in \Lambda} \frac{1}{\|\nabla u_Q\|} \lambda_D (i, j) \\
  &= \frac{\lambda_D (i, j) u_0 (i, j)}{\sum_{Q \in \Lambda \cap q \in \Lambda} \lambda_D (i, j)}.
\end{align*}
\]  

(18)

where

\[
\begin{align*}
  \omega_q &= \frac{1}{\|\nabla u_Q\|}, \\
  h_Q &= \lambda_D (O) + \sum_{Q \in \Lambda \cap q \in \Lambda} \omega_q, \\
  h_O &= \lambda_D (O) + \sum_{Q \in \Lambda \cap q \in \Lambda} \omega_q.
\end{align*}
\]  

(19)

So formula (18) can be remarked as:

\[
\begin{align*}
  u (i, j) &= \sum_{Q \in \Lambda \cap q \in \Lambda} h_q u_Q + h_O u_0 (i, j).
\end{align*}
\]  

(20)
where \( \sum_{Q \in \Lambda, O \in \Lambda} h_Q + h_O = 1 \); apparently, the inpainting objective pixel value \( u(i, j) \) adds the weight \( h_Q \) to be repaired through neighbor field point \( u_{OQ} \).

Use Gauss-Jacobi iteration method to search the optimum:

\[
u^{(n)}(i, j) = \sum_{Q \in \Lambda, O \in \Lambda} h_Q^{(n-1)} u_Q^{(n-1)} + h_O^{(n-1)} u_O(i, j), \tag{21}\]

where \( n \) is the number of iterations. According to the discretization process of TV model, the process of TV inpainting model mainly is shown in Algorithm 1.

### 4. Computer Experiment

#### 4.1. FCM Experimental Analysis

In the FCM experiment in this paper given gray values of the pixel in the repaired areas, we can calculate the Euclidean distance of it with each clustering center concluded by formula (8), the cluster which has the minimum distance is the repaired area \( D \).

In this experiment, it is crucial to select the appropriate values of these parameters, namely, the number of image clustering (C value), the initial value of the clustering center (V), C, and the clustering center initial value selected which directly influence whether the defect area can be accurately identified. Generally apply the gray value distribution of the gray histogram of the image to be repaired to determine the value of C and V.

We discussed the experimental results of FCM to identify scratches, text, and big block inpainting areas

#### 4.1.1. Scratch Inpainting

In Figure 2(a) there is identification effect of areas to be repaired; Figure 2(b) is gray histogram of Figure 2(a).

As Table 1 has shown in the experiment, we set \( C = 6 \); to initialize clustering centers \( V = [0, 20, 50, 75, 100, 125, 175, 255] \) and update the clustering centers, obtained clustering centers are 21.6442, and the clustering center 2 is 69.3399, the clustering center 3 is 108.8676, the clustering center 4 is 154.7014, the clustering center 5 is 192.2138, and the clustering center 6 is 217.2253. Apparently the area to be repaired in cluster 1 identifying the scratches of the repaired area was shown in Figure 2(a), the pixel in the image to be repaired where it close to black gray value will impact identification of the repaired areas from the figure.

When set \( C = 8 \), identification effect of areas to be repaired was shown in Figure 2(c).

Apparently, the effect of Figure 2(c) is better than that of Figure 2(a); at this time, we initialize the clustering center \( V = [0, 20, 50, 75, 100, 125, 175, 255] \) and the clustering center 1 is 9.6953, the clustering center 2 is 45.918, the clustering center 3 is 78.5321, the clustering center 4 is 107.6317, the clustering center 5 is 139.0719, the clustering center 6 is 169.5539, the clustering center 7 is 196.3286, the clustering center 8 is 218.8029. Area to be repaired which has been in Cluster 1, as shown in Tables 1 and 2, the cluster center value of 1 is closer to the color of the area to be repaired (gray value is 0) than the cluster center value of 1 when \( C = 6 \). Therefore, for the greater value of C, the identification is more accurate.

#### 4.1.2. Text Inpainting

In Figure 3(a), there is the identification result of the areas to be repaired; Figure 3(b) is gray histogram of Figure 3(a).

As Table 3 shows, in the experiment, we set \( C = 5 \) initialize clustering center \( V = [0, 75, 150, 200, 255] \), obtained the clustering center 1 is 52.0827, the clustering center 2 is 85.3504, the clustering center 3 is 119.696, the clustering center 4 is 171.9693, and the clustering center 5 is 254.6589. Apparently, the area to be repaired in cluster 5 identifying the scratches of the repaired area was shown in Figure 3(a).
4.1.3. The Big Block Area Inpainting. In Figure 4(a), there is the identification result of areas to be repaired; Figure 4(b) presents gray histogram of Figure 4(a).

As Table 4 has shown, in the experiment, we set \( C = 7 \), to initialize clustering center \( V = [10, 20, 50, 100, 125, 175, 255] \), update the clustering center, finally obtained the clustering center 1 is 22.732, the clustering center 2 is 60.23, the clustering center 3 is 93.2463, the clustering center 4 is 124.5873, the clustering center 5 is 162.8123, the clustering center 6 is 194.4558, and the clustering center 7 is 217.9387.

Apparentely, area to be repaired in cluster 1 identifying the scratches of the repaired area was shown in Figure 4(a); the pixels in the image to be repaired nearly black will be misclassified as noise, so if the grayscale values of the area to be repaired are similar to the gray values of the known area, they will be misclassified.

When we set \( C = 10 \), the extracted area to be repaired was shown in Figure 4(c).

Apparently, the result of Figure 4(c) is better than Figure 4(a); at this time, to initialize clustering center...
Figure 3: (a) Identification effect of areas to be repaired with $C = 5$. (b) The gray histogram of image to be repaired.

### Table 2: Clustering centers table.

<table>
<thead>
<tr>
<th>Initial clustering centers</th>
<th>Obtained clustering centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering center 1</td>
<td>0</td>
</tr>
<tr>
<td>Clustering center 2</td>
<td>20</td>
</tr>
<tr>
<td>Clustering center 3</td>
<td>50</td>
</tr>
<tr>
<td>Clustering center 4</td>
<td>75</td>
</tr>
<tr>
<td>Clustering center 5</td>
<td>100</td>
</tr>
<tr>
<td>Clustering center 6</td>
<td>125</td>
</tr>
<tr>
<td>Clustering center 7</td>
<td>175</td>
</tr>
<tr>
<td>Clustering center 8</td>
<td>255</td>
</tr>
</tbody>
</table>

### Table 3: Clustering centers table.

<table>
<thead>
<tr>
<th>Initial clustering centers</th>
<th>Obtained clustering centers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clustering center 1</td>
<td>0</td>
</tr>
<tr>
<td>Clustering center 2</td>
<td>75</td>
</tr>
<tr>
<td>Clustering center 3</td>
<td>150</td>
</tr>
<tr>
<td>Clustering center 4</td>
<td>200</td>
</tr>
<tr>
<td>Clustering center 5</td>
<td>255</td>
</tr>
</tbody>
</table>

$V = [0, 10, 20, 30, 40, 50, 100, 125, 175, 255]$, update the clustering center, finally obtained the clustering center 1 is 12.2282, the clustering center 2 is 39.6536, the clustering center 3 is 63.6337, the clustering center 4 is 85.1273, the clustering center 5 is 106.6592, the clustering center 6 is 129.8529, the clustering center 7 is 156.3771, the clustering center 8 is 179.4131, the clustering center 9 is 198.9847, and the clustering center 10 is 219.6205. At this time, Area to be repaired which has been in Cluster 1, as shown in Tables 4 and 5, the cluster center value of 1 is closer to the color of the area to be repaired (gray value is 0) than the cluster center value of 1 when $C = 7$. Therefore, the greater value of $C$ and the initial value of clustering center are more appropriate, and the more accurate identification results can be obtained.

4.2. Realization of the Image Automatic Inpainting. Let $\lambda = 0$, $\varepsilon = 0.0001$ with improved signal-to-noise (ISNR) as...
Figure 4: (a) Identification effect of areas to be repaired with $C = 7$. (b) The gray histogram of image to be repaired. (c) Identification effect of areas to be repaired with $C = 10$.

the objective metric standard of the quantity of the inpainting image, the definition of ISNR as

$$\text{ISNR} = 10 \cdot \log_{10} \left\{ \frac{\sum_{x=1}^{M} \sum_{y=1}^{N} \tilde{I}(x, y) - I^0(x, y)}{\sum_{x=1}^{M} \sum_{y=1}^{N} \tilde{I}(x, y) - I(x, y)} \right\}^2,$$

(22)

where $\tilde{I}(x, y)$ is the original gray image, $I^0(x, y)$ is the damaged image, and $I(x, y)$ is the recovered image.

Combining TV model inpainting algorithm with FCM identifying the damaged areas, we have discussed the automatic inpainting effect of scratches and big block damaged areas as follows.

4.2.1. Scratch Inpainting. In the following experiments, $n$ is the number of iteration, $T$ is the inpainting time, and ISNR is improved signal-to-noise.
4.2.2. The Big Block Area Inpainting. Comparing Figures 5 and 6, it shows that the automatic image inpainting model may generate blurred edges for big block damaged area, not as good as small inpainting areas.

5. Conclusions

FCM clustering is an unsupervised clustering technique applied to classify images into clusters with similar properties. It utilizes the distance between pixels and cluster centers in the repaired area to compute the membership function. Experiments showed that the inpainting effect of automatic image inpainting model was decided by the inpainting model, identifying the repaired area that depends on FCM algorithm.
When the selected clustering number $C$ and the selected clustering center initial value are appropriate, FCM can accurately identify the repaired area. As combined with the inpainting model, we can realize automatic digital image inpainting function.

**Conflict of Interests**

The authors declare that there is no conflict of interests regarding the publication of this paper.

**References**


