A Machine Learning Approach on Image Matching in the Modern Environment

S. Vidya Sagar Appaji, Prof P.V. Lakshmi, Dr P. Srinivasa Rao

1 Research Scholar, CSE Department, GITAM [Deemed to be University], Visakhapatnam
2 Professor, IT Department, GITAM [Deemed to be University], Visakhapatnam
3 Associate Professor, MVGRCE(A), Chintalavalasa, Vizianagaram.

Email: sagarsetti4u@gmail.com

Received: 26th October 2018, Accepted: 29th January 2019, Published: 28th February 2019

Abstract

In recent trends, based on the environment, the quality of the digital photo may degrade by making it blurry, low contrast and we may get some situations where a person need to be recognized for a specific reason. To provide the quality and identity of photograph, this paper explores Image Enhancement and Image Matching algorithms using Hadoop Mapreduce Framework. The proposed algorithms can enhance the image quality and can give the approximate matching result from the dataset. Here, we need a framework that is an abstraction that provides generic functionality that is build to concentrate on image processing tasks. Map reduce paradigm framework allows the image to divide into clusters in order to facilitate distributed processing. We chose hadoop as it is known for its scalability, cost effective, flexibility, fast, security and authentication, parallel processing, and MapReduce is noted for its data processing across different computing nodes.

Keywords
Blurry, Enhancement, Matching, Framework, Hadoop, Computing Nodes

Introduction

Image Enhancement:
The main objective is to process or alter an image so that result is more suitable than that of the original image for further image analysis or for specific application. It is the process that uses software to manipulate the stored image digitally. Filters, image editors and other tools are used for image enhancement for manipulating different properties of parts of an image or an entire image. These techniques provide more number of choices in order to improve the quality of images visually. These tools will simply change the contrast or brightness of an image or manipulate the gray scale or the red-green-blue color patterns of an image. Image enhancement includes enhancing images, sharpening and softening images, contrast change and brightening, color adjustments. Image enhancement focuses on operations of image processing which will increase the quality of image which in turn will overcome the human visual system weakness. The methods of enhancement techniques are:

- Fuzzy Set Theory
- Histogram Equalization
- Histogram Matching
- Equalized Histogram Equalization

Histogram Equalization:
It is a method which uses image’s histogram for contrast adjustment in image processing which will be very helpful in increasing image’s global contrast. Let us assume a discrete gray level scale image \( \{x\} \) and let \( n_i \) be assumed as the occurrences of gray level \( i \). The pixel occurrence of level \( i \) probability in the image is

\[ p_x(i) = \frac{n_i}{n}, \quad 0 \leq i < L \]

Here \( L \) is the total gray levels count in the image, \( n \) is the total pixels count in the image, and \( p_x(i) \) is the pixel value \( I \) of image's histogram, which is normalized to \([0,1]\). Here we are defining a function for cumulative distribution which corresponds to \( p_x \) as

\[ cd f_x(i) = \sum_{j=0}^{i} p_x(j) \]

which is also known as the image's overall equalized histogram. Here we had created a transformation for the form \( y = T(x) \) which in turn a new image \( \{y\} \), which contains a flat histogram is produced. The produced image will have a cumulative distribution function(CDF) which is linearized across the range of values, i.e.
for some K constant. The CDF properties helps us in performing such a transform; which is defined as

\[ cd f_y(i) = iK \]

Here the range of K is \([0, L]\). Here what we have to notice is that \([0, 1]\) is the range where \(T\) points the levels into, since we had used a equalized histogram of image \(x\). The simple transformation which is needed to be applied is:

\[ y' = y \cdot (\max\{x\} - \min\{x\}) + \min\{x\} \]

**Image Matching:**

**SURF Feature Detection:**

The integral image concept’s implementation is facilitated by Hessian-matrix approximation use which will speed up the detection and is noted to result in a better performance in accuracy. The advantage of providing a complementary information is the best reason for implementing this method of feature extraction method. This function takes the values of \(x\) and \(y\) from the feature detected coordinates. The following relation is for the computation:

\[ \text{int } S = 2 \cdot \text{Convert.ToInt32}(2.5f \cdot \text{ip. scale}); \]

let us consider a two variables continuous function which will be like \(f(x; y)\). The Hessian matrix, \(H\), is function partial derivates matrix \(f\)

\[
H(f(x, y)) = \begin{bmatrix}
\frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\
\frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2}
\end{bmatrix}
\]

The discriminant is the determinant of above matrix, and is calculated by

\[
\text{det}(H) = \left( \frac{\partial^2 f}{\partial x \partial y} \right)^2 - \left( \frac{\partial^2 f}{\partial x^2} \right) \left( \frac{\partial^2 f}{\partial y^2} \right)
\]

The extreme values of the function will be classified by discriminant value.

**SURF Feature Matching:**

The indexing process is used by SURF which accelerates the matching stage. The features which are having same type of contrast will be compared in matching stage which allows us for faster matching.

### Related Work

In [1] a comparison was made on various histogram equalization methods. In this paper, they have observed that the brightness preservation handled well by RMSHE, RSIHE, and RSWHE-M when checked with other methods. In [2] they implemented a working set that involved Blob detection, template matching, SURF feature extraction as the three image matching techniques for performance comparison. [3] Studied and compared the three algorithms: SIFT, SURF, FAST for Image Matching and further dealt with accuracy levels achieved in determining each object in the image. [4] presented a paper in which they have shown a technique direct matching, that deal with differences of pixel values for the gray scale images. [5] surveyed some of the areas where image enhancement is done and presented the most important techniques for image enhancement in digital image processing. [6] analyzed four

<table>
<thead>
<tr>
<th>Key points Matched</th>
<th>Match Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SIFT Library</td>
<td>858 of 3705</td>
</tr>
<tr>
<td>Lowe’s executable</td>
<td>1087 of 4635</td>
</tr>
</tbody>
</table>
important image enhancement methods. The algorithms proposed in this paper used software techniques to improve the contrast levels of fingerprints, medical, and selenography images. [7] presented an analysis of spatial domain processing techniques and in specific, this paper categorized processing methods based on image enhancement representative techniques. [8] shown a similarity measure to obtain image detection, and the best precision is noted to be achieved by CVA. In [9] they have presented a procedure for the evaluation of segmented images. The methodology proposed consists in comparing the segmentation image obtained by any algorithm with the ideal (hand-made) segmentation. In [10] they proposed an algorithm that enhances the foreground and background image part separately and stretch the contrast of an image at inter-object level and intra-object level and then combines it to an enhanced image. In [11] they demonstrated the performance of some algorithms of image enhancement which is evaluated both qualitatively and quantitatively and gives a brief description of techniques of spatial domain. [12] discussed about Big Data, its storage aspect and its challenges and working of mapreduce and then the techniques of image enhancement that are useful in performing mapreduce job for statistical analysis. In [13] it was explained about four most used image enhancement techniques and also determines the best parameters and transformations used.

**Environmental Setup**

Based on cluster equipped with hadoop this experimentation is being implemented. This project has been executed with one Name Node and four Data Nodes. Later on, a gigabit switch connects the computing nodes. Hadoop 2.7.1, Ubuntu 14.04 LTS, and Python were installed on both the Name Node and the Data Nodes.

<table>
<thead>
<tr>
<th>Table 1. Master and Slave Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Master</td>
</tr>
<tr>
<td>--------</td>
</tr>
</tbody>
</table>
|        | 1      | 2x2.5 GHz Central processing unit(
|        |        | CPU) Node, 4 GB RAM, 100 GB Hard disk space |
| Slaves | 2      | 2x2.5 GHz Central processing unit(
|        |        | CPU) Node, 4 GB RAM, 100 GB Hard disk space |

<table>
<thead>
<tr>
<th>Table 2. Hardware Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Name Node</td>
</tr>
<tr>
<td>Data Node</td>
</tr>
<tr>
<td>Network</td>
</tr>
<tr>
<td>Storage</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3. Software Specifications</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
</tr>
<tr>
<td>---------</td>
</tr>
<tr>
<td>Ubuntu</td>
</tr>
<tr>
<td>Python</td>
</tr>
</tbody>
</table>

**Results**

Figure 3: A Part of Image was Visible but the Proposed Methodology will Check for Match
Figure 4 shows that the given input is matched with one of the images in given dataset. Figure 4 shows that image cannot be found.

**Conclusion and Future Scope**
This paper presents the working of histogram equalization technique which is used for image enhancement. And further, also explains the SURF algorithm to determine similarity from the given image with the available data set of images, and this image matching is computed using feature detection, followed by feature extraction. The features obtained from the detection are matched more likely.

**References**
8. Comparison of Near-Duplicate Image Matching, Li Chen, F. W. M. Stentiford, University College London, Adastral Park Campus, IP5 3RE, UK. {l.chen, f.stentiford}@adastral.ucl.ac.uk
9. A new algorithm for color image comparison based on similarity measures, Daniel Paternain, Mikel Galar, Aranzazu Jurio, Edurne Barrenechea, 8th Conference of the European Society for Fuzzy Logic and Technology (EUSFLAT 2013)
