Abstract — The SIFT descriptor is one of the most widely used descriptors which has considerable stability against changes such as rotation, scale, affine of the image, and illumination. However, because of the greater emphasis on its insensitivity to geometric changes, this descriptor is weak in various illuminations. Therefore, in this article an attempt has been made to boost the SIFT descriptor against changes in illuminations through the use of techniques of creating pictures in various illumination conditions and by extracting the desired features of these conditions. For this purpose, we have used the Power-Law Transform, and the results of the implementation testing have been successful. The efficiency of the proposed algorithm and of the base algorithm of SIFT with regard to the data set ALOI have been investigated, and it has been found that by adding this method to the base SIFT descriptor, the rate of recognition improves by five percent. Moreover, there will be a better response to changes in illumination.

Keywords — Illumination Variance, Object Recognition, Power-Law Transform, SIFT Descriptor

I. INTRODUCTION

Extracting the key points from the image of an object (i.e., points which can function as good representatives for describing the object) which are stable in various views and make the realization of good recognition possible, is one of the main challenges in the area of machine vision, camera calibration, 3D reconstruction, image registration, robot navigation, and object recognition are only a few of the applications of these features. For example, in object recognition these key points can be used in three stages:

1- Finding the key points: this can be achieved through a general strategy by searching in the length of the image and finding unique points which have specific features. These points can be found by searching for the corners, bubbles, and T-junctions.

2- Describing the key points: these points should be described in a way that we will have the same display of the key points in the presence of noise in the environment and when changes occur in geometry and illumination, and that these points will be distinctive and insensitive.

3- The last stage concerns the matching of these points among different images. Normally, methods of calculating distances, such as Euclidean and Mahalanobis methods are used for eigenvectors obtained in the last stage. In a comparison carried out among different methods of describing features [1], it was found that the insensitivity to scale transform in SIFT [2], [3] offers the most distinctive description of the object. The SIFT descriptor emphasizes on key insensitive points extracted through Gaussian Differences (DoG). In the description stage, the magnitude and orientation of the images, gradients based on histogram, and gradient orientations around the key points are also obtained.

This descriptor yields good results in changes such as rotation, scale, and affine of the images, but it is weak against changes in illumination. Therefore, in this article a method with the name MGS\texttrademark is presented to overcome this problem through the use of techniques of constructing new samples of an image in different illumination conditions. The innovation introduced in this article is the collection of more points from the image in different illumination conditions (Fig. 1) in order to help the image become independent of lighting. Adjustment of parameters has been carried out through the use of genetic algorithms. One of the advantages of the proposed method is that most of the processing is performed offline. Therefore, there will be little overhead when the testing is done; and the insensitive recognition of illumination will be performed online. The practical tests carried out in this article show that the proposed method had a considerable effect in increasing the number of matched points and it also increased the accuracy of classification.

II. RELATED WORK

As was stated before, strong local descriptors, which are obtained through extracting key insensitive points, have had

\[1\] Multi Gray Scale
III. BASE SIFT DESCRIPTOR

Since the innovation we present is the improvement of the SIFT descriptor under conditions of different MSGs, in this section we discuss this descriptor. Today, the SIFT descriptor is considered as one of the best and most powerful tools for extracting key points insensitive to different conditions such as rotation, scale, change in viewing orientation, noise, MSG, and the affine transform. As was stated before, the SIFT descriptor has many advantages over other descriptors, and for this very reason it has received a lot of attention in the area of object recognition. In this application, through matching the key points extracted from the original image with their equivalent points in the final image, and by considering a given number of matched points, recognition is performed. In other words, it can be said that this descriptor does not learn the general features of an object in order to classify it. In one of the applications, nowadays SIFT is implemented in some of the FPGA boards. In general, this descriptor is used in three stages which are as follows:

A. Finding the Key Point

The first stage in all methods in which work is carried out on special (key) points of the picture is to find these points.

1. Finding extreme points in scale space

In this method, Gaussian Differences (DoG) are used to find the key points of the image. The process of finding these points starts with constructing a pyramid of images, and with convolution of the image I(x,y) with a Gaussian filter G(x,y,σ). Therefore, the scale-space is represented as follows:

\[ I(x,y, \sigma) = I(x,y)^*G(x,y, \sigma) \]  

(1)

"*" stands for the convolution operator in (x, y), and

\[ G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{x^2+y^2}{2\sigma^2}\right)} \]  

(2)

The degree of blurring is controlled by the standard deviation parameter σ in the Gaussian function. The scale-space DoG can also be obtained by subtracting adjacent scale levels:

\[ D(x,y, \sigma) = [G(x,y,\sigma) - G(x,y, k\sigma)]^*I(x,y) \]  

(3)

Using the Eq. (1) we will get:

\[ D(x,y, \sigma) = I(x,y, \sigma) - L(x,y, \sigma) \]  

(4)

In Fig. 2 the stages of constructing the DoG space are shown.

![Figure 2](image)

Figure 2. For each octave of scale space, the initial image is repeatedly convolved with Gaussians to produce the set of scale space images shown on the left. Adjacent Gaussian images are subtracted to produce the...
difference-of-Gaussian images on the right. After each octave, the Gaussian image is Down-sampled by a factor of 2, and the process repeated.

In the next stage, the maximum and the minimum points in each octave are found. This is achieved by comparing each pixel with its 26 neighbors in region 3*3 of all adjacent DOG levels present in the same octave (Fig. 3). If the point in question is bigger or smaller than all its neighbors, it is chosen as the desired point [3].

2. Locating Key Points

In this stage, we omit some of the extracted key points in two phases so that we can have key points which are less sensitive to noise, or have points which are not located on the edges. To do this, in phase one we use the Taylor expansion to omit extreme points which are unstable and have a low Power-Law contrast [13]. Since DoG has high sensitivity to edges, it is possible that some extracted points will be along weak edges and hence will not be stable in the presence of noise. Therefore, in phase two we use the Hessian matrix to omit points which have the above feature.

3. Orientation Assignment

In this stage, preparations are made for constructing the eigenvector. To each key point, an orientation is assigned based on the local features of the image. For each sample image \( L(x, y) \) in this scale, the gradient range \( m(x, y) \) and the gradient orientation \( \theta(x, y) \) are calculated using differences in pixels and based on the following formulae:

\[
m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \tag{5}
\]

\[
\theta(x, y) = \tan^{-1} \left( \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \tag{6}
\]

The orientation histogram is built by using the gradient range for a key point together with an area around the point (Fig. 4).

B. Display the Key Point Description

In this stage the eigenvector will be developed. First the gradient range and the orientation around the key point are sampled. In his experiment, David Lowe used the 4*4 array with eight orientations in each histogram instead of employing a 2*2 array for orientation histograms. Therefore, the eigenvector length will be 4*4*8=128 members for each key point.

C. Feature Vector Matching

The matching phase of the recognition stage is carried out by comparing each key point extracted from the test image with key points related to the trained image. The best candidate points for matching are found through recognizing the nearest neighbor in the set of key points of the train image. The nearest neighbor has the least distance from its matched point.

IV. PROPOSED ALGORITHM

The following flowchart shows the procedure of carrying out the processes of the proposed algorithm.

A. Power-Law Transform

The Power-Law Transform is defined as follows:

\[
s = cr^\gamma \tag{7}
\]

Where \( c \) and \( \gamma \) are positive constants. Sometimes equation 7 is written as \( s = c(r + e)^\gamma \) when relocation is considered (with the condition the entry be zero). However, in equation 7 we have omitted the relocation value, because we have assumed that the conditions are normal and that calibration has been carried out. Changes in the two variables \( r \) and \( c \), with respect to each other, are shown in Fig. 6 for various values of \( \gamma \).
C. Combination of SIFT and Power-Law Transform

A closer look at the proposal algorithm will show that it consists of the following stages:

1. Constructing samples insensitive to illumination (MGS) using the train samples (S).
2. Extracting key points and extracting feature vectors of these key points for the new sample (Sij)
3. Transforming the new samples from the MGS space of each class (si) to a single sample (Sj)
4. Continuing the algorithm based on the base SIFT descriptor.

Figure 6. Power-Law transform for different value of $\gamma$

Power-Law curves having a small $\gamma$, map an area of dark input value in an extensive area of output values, the reverse is also true for levels with higher input values. Different and possible states of the transform curves can be easily obtained by changing the value of omega.

Based on Fig. 6, it can be seen that the curve relating to values from $\gamma>1$ are the converse of the curves produced by values from $\gamma<1$. The last point to find out from equation 7 is the simultaneous transform with $c = \gamma = 1$. Various image capturing, printing, and displaying equipment operate based on this transform. The defining parameter in this equation is gamma, and the process which makes the use of this transform suitable for a particular application is called gamma correction. Since this transform has a variety of applications, the gamma correction has a very important role in improving the efficiency of the output. It is for this very reason that the optimum values of gamma have been calculated using genetic algorithms.

B. Using SIFT to Key Point Extraction

Assuming that there is a set of objects $S$ in the form of $S=\{s_1,s_2,s_3,\ldots,s_n\}$, for each $s_i$ form the set $S$, there is a set of key points with their eigenvectors that have been extracted by SIFT(Fig. 7). In other words, each row in this table represents one of the key points in our train image, $m$ shows the number of key points related to an image, $(x,y)$ are the coordinates of the key point.

In the testing phase, which is usually in real-time, first the key points together with eigenvectors are extracted from the test image. Therefore, we will have two sets, one representing the train sample and the other related to the test set. We compare the key points of the test set with each and every row in the $S$ set, and if the number of matched points for each $s_i$ is more than the others, that member is determined as the matched object (Fig. 8).

D. Gamma Correction with Genetic Algorithm

To obtain optimal amounts of gamma-standard genetic algorithms are used. Gens of each chromosome contain different doses of gamma. The fitness of each chromosome has $K$ length. Based on two factors and classification
accuracy in a limited amount of permitted distribution is calculated (Eq. (8)). Other parameters according to Table 1 are set.

<table>
<thead>
<tr>
<th>TrainSet</th>
<th>α</th>
<th>K</th>
<th>High Range</th>
<th>Low Range</th>
<th>Pm</th>
</tr>
</thead>
<tbody>
<tr>
<td>20-70%</td>
<td>80%</td>
<td>5</td>
<td>10</td>
<td>0</td>
<td>0.5</td>
</tr>
</tbody>
</table>

Pm mutation rate, Low/High range Gens Confine parameters (admissible confine of γ), K number of scales of MGS space, α parameters division as fitness impact factor (Eq. (8)) and the size of train set appointed with TrainSet.

\[
\text{TotalFit} = α \ast \text{ClassificationFit} + (1 - α)\text{DisperalFit}
\]

\[
\text{DisperalFit} = \sum_{i=1}^{N} \text{dist}(x_i + x_{i-1})
\]

In fitness computation, α in recognition rate and 1-α in rate of dispersal are multiplied. Scattering rates in order to maintain the distance between various γ.

V. EXPERIMENTAL RESULT

The data set used in the tests ALOI [14] included 27 samples with different Illumination for 1000 different objects. (Fig 1 and 9 are show that)

In the first test, adaption points compared with two methods. Points that adapted via MSG-SIFT algorithms are more than of this point that obtained with SIFT base. (Fig. 10) Therefore, when the illumination conditions in different data sets we tested, the proposed algorithm will be more reliability.

![Comparison of SIFT vs. MGS-SIFT](image)

In the second experiment the number of different training and test samples was considered. As the result of experiments in Table 2 shows, in the absence if the number of samples is still small given the proposed algorithm has higher accuracy will be. This is because the distinctive features in the construction of new samples with different illumination conditions.

<table>
<thead>
<tr>
<th>Train and test set</th>
<th>SFIT</th>
<th>MSG-SIFT</th>
</tr>
</thead>
<tbody>
<tr>
<td>20% Train,80% Test</td>
<td>85.77</td>
<td>89.64</td>
</tr>
<tr>
<td>70% Train,30% Test</td>
<td>89</td>
<td>95</td>
</tr>
</tbody>
</table>

So you can conclude that the proposed algorithm on a data set containing examples of the lighting is different is a better result. Therefore, in such matters, the algorithm has provided will be better performance than based SIFT.

VI. CONCLUSION

Two Parameters that should be satisfied by the descriptor is stability and distinction. (Repeatability against changes and having the minimum information to describe objects)

As observed, the proposed method in this article, one of the most popular descriptors in the field of machine vision stability and distinction in two aspects examined.

With regards to the conversion descriptor SIFT rotation, scaling and image stretching and strengthening is invariant against changes in illumination certainly more issues will be used. But in later work on this issue can be reviewed and the color images in color space problem can be solved.

ACKNOWLEDGMENT

This project is supported with Mechatronic Research Centre of Qazvin Azad University.

REFERENCES