ISSN No:-2456-2165

Image Reconstruction using its Spatial and Geometrical Information

K. B. Ranushka Pasindu Dharmaranga
Typefi systems (pvt) Ltd,
Colombo, SriLanka

Ligitha Sakthymayuran
Trincomalee campus,
Eastern University, Srilanka

Abstract:- Image reconstruction is currently often used in a wide range of technological and medical applications. The local image feature descriptor is the most critical factor influencing the performance of object reconstruction or image retrieval systems. This study provides and demonstrates a strategy for replicating images. In this approach, Training photos are used to extract local feature descriptors; at first images are recreated using local feature descriptors and geometric information. Scale invariant feature transform (SIFT) descriptors are used to characterize images, and the feature extraction method is similar to how descriptors are used in the training phase. The unknown image closest neighbor descriptor built by using pairwise matching. For each of the regions of interest, visually equivalent patches may be in the external image database.To detect patch overlapping regions between the new patch and the patch already present in the query image, the Mean Squared Error (MSE) is used. To eliminate overlapping patches, the highest MSE threshold value is chosen as the default threshold (DT) in this experimental technique. Based on the experimental results, an image may be approximated and rebuilt using image local feature descriptors.

Keywords:- *Image reconstruction, image retrieval, image feature descriptor, geometric information, partial information.*

I. INTRODUCTION

Researchers are becoming increasingly interested in reconstructed images based on local visual features. Local feature descriptors are used in several cutting-edge image reconstruction research investigations (Vondrick, Khosla, Pirsiavash, Malisiewicz, & Torralba, 2013) (Weinzaepfel, Jgou, & Prez, 2011) (d'Angelo, Jacques, Alahi, & Vanderg, 2014) to get improved results. Pixel information is becoming increasingly important while looking for images in today's digitalized environment. Local feature descriptors can also give a wealth of information about an image or areas of an image. Furthermore, cutting-edge image reconstruction approaches such as the Scale Invariant Feature Transform (SIFT) (Lowe, 2004) and Speeded Up Robust Features (SURF) (Amato, Falchi, & Gennaro, 2013) rely on local feature descriptors to increase performance (Vondrick, Khosla, Pirsiavash, Malisiewicz & Torralba, 2013) (Weinzaepfel, & Torralba, 2013) (d'Angelo, Jacques, Alahi, & Vanderg, 2014). To define a picture uniquely, a set of local descriptors with matching geometrical metadata can be employed. As a result, the amount of storage required to store these descriptors is significantly reduced. Using local descriptors significantly enhances the efficiency of visual search (Amato, Falchi, & Gennaro, 2013), (Xiao, Zhiwei, Lei, Wei-Ying, & He, 2009). Picture patches that are very similar to query descriptors and may be retrieved from a massive collection of picture patches and descriptor information. Image feature quality are the most essential factor in influencing object recognition performance. Reconstructing original images from their local features is one of the most natural approaches to examine feature characteristics. Picture reconstruction, for example, is used to analyse classification problems of the HOG feature (Vondrick, Khosla, Pirsiavash, Malisiewicz, & Torralba, 2013). Another important use of image reconstruction is image synthesis, which makes advantage of image properties. Image attributes are used to express semantic information in images. SIFT features are used to characterize the pictures in this investigation. SIFT is a feature descriptor that is picture size and orientation invariant. This descriptor is also robust to perspective distortions and affine Image Reconstruction Using Local Feature Descriptors.

A. *Objective:*

The purpose of this study is to focus on image reconstruction using geometrical information and local feature descriptors. The geometrical information in the local descriptor includes the (x; y) coordinates, orientation, and scale of the image patch.

B. *Contribution*

The major contribution of this research is the creation of a unique method for reproducing digital photos. The most challenging component about reproducing digital photographs is a lack of information. To address this issue, in this approach a revolutionary method used for locating lost metadata. It implemented this new technique for data collecting and saw a substantial increase in speed.

II. LITERATURE REVIEW

(Weinzaepfel, Jgou, & Prez, 2011)This research was conducted to illustrate how a picture may be approximately recreated using the output of a black box local description tool, such as those previously used for image indexing. As a first step in this research, they created a database of image patches from which they can easily match and discover image patches from locations of interest.

These patches are then deformed into the original image and neatly stitched together based on the original image's interest zones, which were used for research testing reasons.

At the end of the study, gaps in still-missing texture-free areas are filled using smooth interpolation. In this investigation, only local feature descriptors such as SIFT were used to recreate visually meaningful images. The INRIA holiday dataset and the Copyday dataset, both of which contain holiday photographs, were used in this study.

This work presented a technique for recreating a digital image utilizing local descriptors such as SIFT and geometrical information.

(Mortensen, Deng, & Shapiro, 2005) In this work, they looked at an approach for matching numerous pictures based on local feature descriptors like SIFT. While modern descriptors such as SIFT may find matches between features with unique local neighbors, they frequently neglect the global context when resolving ambiguities that may emerge when an image has several related locations.

In this work, a feature descriptor was presented to complement SIFT with a global context vector. When picture key points are more successfully matched individually, a more robust way of hand modification is provided. The study was conducted in three parts. For each interest key point detected in the original image that was used as the testing image, a two component vector consisting of a SIFT descriptor that reflects local characteristics and a global context vector to disambiguate locally similar features was produced.

SIFT was used for this investigation since it outperforms the other known key-point descriptors.

This is the case because the SIFT descriptor computes the gradient vector for each pixel in the feature point's vicinity and generates a normalized histogram of gradient directions.

Given two photographs and a collection of feature points that can be easily identified in each image, as well as robust descriptions for those characteristics, we then match feature points between them.

They did not undertake the extremely expensive group-wise consistency checks when matching two or more photographs using the proposed approach. To compare descriptors, we employ a simple nearest neighbour distance or nearest neighbour.

Furthermore, the authors of (Hiroharu & Tatsuya, 2014) proposed altering BREAF (Calonder, Lepetit, Strecha, & Fua, 2010) and FREAK (Alahi, Ortiz, & Vandergheynst, 2012) descriptors to generate image patches. This method generates the best possible image patch for a given input description. According to their testing results, the generated images are clear, and this method is only applicable to certain types of descriptors.

(Vondrick, Khosla, Pirsiavash, Malisiewicz, & Torralba, 2013) recreates images using HOG properties. Four strategies for image reconstruction are proposed in (Vondrick, Khosla, Pirsiavash, Malisiewicz, & Torralba, 2013). According to the data in (Vondrick, Khosla, Pirsiavash, Malisiewicz, & Torralba, 2013) one of the strategies outperforms the others. The proposed technique is based on a feature vocabulary and its corresponding images. In theory, this method may be used for any arbitrary feature.

The SIFT feature descriptor is used to reconstruct images in (Weinzaepfel, Jgou, & Prez, 2011). This technique is more suited to addressing numerous privacy problems that emerge when image descriptors are simply sent over an unsecured network, such as when used for image recognition or classification. In this image reconstruction approach, similar picture patches are found using local attributes.

A local binary descriptor is used to recreate the images (d'Angelo, Jacques, Alahi, & Vanderg, 2014). This approach (d'Angelo, Jacques, Alahi, & Vanderg, 2014) is used to rebuild image patches in certain wavelet frames by solving an optimization problem with a regularization term that quantifies the sparsity of the reconstructed patch.

III. DATA AND METHODOLOGY

In this section, we'll look at how to reconstruct images utilizing local feature descriptors and geometrical data. Although recovering a single patch from its descriptor is not possible, local feature descriptors and their associated geometrical information are extremely valuable in identifying the connected patch in the original picture database. This method also makes use of an external image database to recreate picture patches. Consequently, the proposed reconstruction approach creates intensity pictures by utilizing external local feature descriptors and an external image database.

In the beginning, training images are utilized to extract local feature descriptors. The descriptor database is built by integrating the extracted local feature descriptors with the geometrical information of the region of interest and the source picture index. Each extracted descriptor, for example, is labeled D(i) = f(i), x(i), y(i), O(i), S(i), index(i) f(i) [Rd] is the d dimension of the feature descriptor. The region of interest's spatial coordinates are x(i),y (i). The orientation and scale of the obtained feature descriptor are O(i),S. (i). index(i) is the index of the source image from which the feature descriptor was extracted.

In this method, SIFT descriptors are utilized to characterize images. Because SIFT is a well-known feature detection approach. The SIFT method is also used to identify prominent and stable feature points.

Consequently, during the testing time, an external image database and a database of feature descriptors are utilized to reconstruct every picture patch.

During the testing step, local feature descriptors T(j) are extracted from the query image. Consequently, this feature extraction approach is analogous to how descriptors are employed during the training phase.

The unknown image closest neighbour descriptor for T(j) is then constructed by utilizing pairwise matching to

ISSN No:-2456-2165

derive the unknown image closest neighbour descriptor for T(j) from D. The nearest neighbour descriptor is then used to derive the appropriate image patch using the original picture database.

Furthermore, Mean Squared Error (MSE) is used to find patch overlapping regions between the new patch and the patch already existing in the query image. In this experimental approach, the highest MSE threshold value is used as the default threshold (DT) to reduce overlapping patches.

If the new patch's threshold value is greater than DT, it is considered a non-overlapping patch with the existing patches in the query image. Consequently, the query image has received a fresh patch. Otherwise, the new patch will be considered an overlapping patch and will not need to be fixed into the query image.

If the MSE between the new patch and the existing patch vary, the patches are regarded to be two separate patches. If the MSE value is less than the threshold, the two patches will be handled as a single patch, and no more patches will be performed.

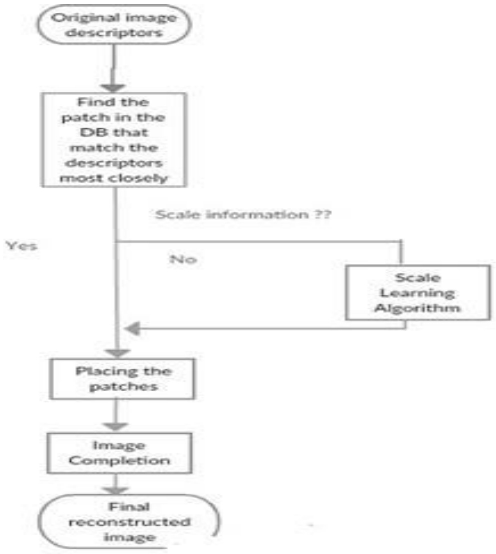

Fig. 1: workflow

A. *Background Feature Extraction*

The input combinations are referred to as features. Feature extraction is the process of extracting characteristics from image data. Color, shape, texture, local invariant traits, and local binary patterns are all characteristics. The scale-invariant feature transform (SIFT) method turns picture data into scale-invariant coordinates associated with local features. This approach generates many features that cover the image densely at all sizes and locations, which is an important property. The ability to distinguish tiny things in noisy backgrounds needs the matching of at least three qualities from each object for reliable identification, hence feature quality is very important for object recognition. The following are the primary computing processes necessary to generate SIFT features:

- Scale-space extrema detection: The first stage in computing seeks out extrema at all scales and places in the image. It is simple to do by utilizing the Difference-of-Gaussian (DoG) function to find future interest locations that are scale invariant.
- Key point localization: To determine the position and scale of each prospective location, a thorough model is employed. The important points are chosen based on their stability.
- Orientation assignment: Depending on the local image gradient directions, each key point position is assigned one or more orientations. All following procedures are performed on image data that has been altered in relation to the assigned orientation, scale, and location for each feature, resulting in transformation invariance.
- Creating key point descriptors: At the selected scale, the local image gradients in the region around each key point are measured. These are translated to an 8 representation, allowing for significant local shape distortion and illumination alterations.

B. *Patch Matching*

It is customary to use comparison or matching in image editing or building operations to make the process easier and more successful. Many contemporary methods, such as Mean Square Error (MSE) and Root Mean Square Error (RMSE), can be used to do this (RMSE).

Mean Square Error (MSE) The mean square error provides us with a single value that contains extra information about the degree of fit of the regression line. In other words, it provides us with a broad concept of how to compare two matrices. Matrix values move closer to one another as the Mean Square Error value declines, since lower values indicate decreasing magnitudes of error. Consider the case where all the data points are perfectly aligned with the regression line. When two identical matrices are used to calculate the Mean Square Error, the result is zero. This would result in residual errors of 0 for all locations, as well as an MSE calculation of 0, which is the minimum MSE number imaginable. A measure of picture quality is required so that Mean Square Error may be used to compare restoration outcomes.

C. *Dataset*

For this investigation, the Zurich building data collection is used. This collection contains 1005 photographs in Portable Network Graphic (PNG) format. Each of the 201 courses has 5 images. Each image has a resolution of 640x480 pixels.

Fig. 2: Image Sample

D. Euclidean Distance

In most circumstances, we use the Euclidean distance calculation based on Pythagoras' theorem in geometry, particularly when computing the distance between two points in the plane.

In this research, utilized the Euclidean distance measure to compute the distances between two significant spots on the image plane, and used the SIFT descriptor to calculate all of the necessary coordinate values.

E. Experimental Setting

The ZuBuD Image Database (Svoboda, Shao, & Gool, 2003) is used to evaluate the performance of the suggested technique. This collection has 201 building courses, each with five photographs. Figure 1 depicts a sample of images from the ZuBuD Image Database. To assess the effectiveness of the proposed approach, 20 building classes are chosen at random from the ZuBuD Image Database. In each class, three images are chosen at random for training and the other two are chosen at random for assessment. SIFT feature descriptors are likewise regarded as local feature descriptors in this experiment. A variety of criteria, including the default threshold (DT) and the size of the image patch, are changed with different values to yield better reconstruction images. In addition, the nearest neighbor descriptor is used to select an acceptable picture patch from the original image database. 3X3, 6X6, and 11X11 sized patches are selected from the source image to determine the approximate size of the image patch that is used to reconstruct pictures. Furthermore, based on testing findings, 11X11 sized patches produce better reconstructed images than other sized patches. An 11X11 sized photo patch is used in this experiment.

IV. RESULTS

Figures represent some of the testing outcomes of the proposed experimental design. Based on our testing results, we propose an approach that develops an approximation of the unknown image progressively by creating its region of interest one by one.

Fig. 3: Test image and expected output

Fig. 4: Input images and output images

Figure 4 displays the test image as Input and the created image as Output after 100 iterations of the algorithm and increasing the number of iterations yields higher quality photographs. Completing the task, on the other hand, will need extra time and processing resources.

When applied the proposed technique to a test image from the data set, as shown in Figure 3, the model correctly predicted the outcome. The actual image can be viewed on the left, while the anticipated result can be found on the right.

V. CONCLUSION

The purpose of this research was to explore if it was feasible to reproduce digital images using just their local feature descriptors. In this study, the Zurich building dataset was used to test the built model. In this work, the programming language used for implementation was MATLAB. This proposed model only uses its SIFT-derived local feature descriptors to reconstruct an image (Scale Invariant Feature Transform). A set of images of that item taken from four different viewpoints is utilized to collect image patches that will be stitched together to produce the image. picked SIFT for this project because it allows me to use the same image from different angles and scale settings. As a result, SIFT is a better fit for these requirements. Because it is unaffected by either the scaling factor or the rotation angle.

Dealing with overlapping visual patches was the most difficult component of this investigation. The image will be too fuzzy to discern since overlapping picture patches play havoc with the rebuilt pictures. Consequently, an algorithm called the Scale learning algorithm was developed and used in this work. In this work, image patch matching and insertion in the correct location remains a tough task. If we can overcome that barrier, we will be able to recreate photos with greater quality while also reducing the time it takes to execute the algorithm.

REFERENCES

[1.] Alahi, A., Ortiz, R., Vandergheynst.(2012).FREAK: Fast Retina Keypoint. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* (pp. 510-517).
[2.] Amato, G., Falchi, F., & Gennaro, C. (2013). On Reducing the Number of Visual Words in the Bag-ofFeatures Representation. *International Conference on Computer Vision Theory and Applications*, (pp. 657-662).
[3.] Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). Brief: Binary robust independent elementary features. *11th European conference on Computer vision*, (pp. 778-792).
[4.] d'Angelo, E., Jacques, L., Alahi, A., & Vanderg. (2014). From Bits to Images: Inversion of Local Binary Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 874-887.
[5.] Hiroharu, K., & Tatsuya, H. (2014). Image Reconstruction from Bag-of- Visual-Words. *Conference on Computer Vision and Pattern Recognition*, (pp. 955-962).
[6.] Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 91-110.
[7.] Mortensen, E. N., Deng, H., & Shapiro,
[8.] L. (2005). A SIFT Descriptor with Global Context. *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005),* (pp. 20-26).
[9.] Svoboda, T., Shao, H., & Gool, L. V. (2003). *ZuBuD ± Zurich Buildings Database for Image Based Recognition,Technique Report No.*
[10.] *260.* Swiss Federal Institute of Technology.
[11.] Vondrick, C., Khosla, A., Pirsiavash, H., Malisiewicz, T., & Torralba, A. (2013). HOGgles: Visualizing Object Detection Features. *International Conference on Computer Vision ICCV*, 1-8.
[12.] Weinzaepfel, P., Jgou, H., & Prez, P. (2011). Reconstructing an image from its local descriptors. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, (pp. 337-344).
[13.] Xiao, Z., Zhiwei, L,Lei, Z., Wei-Ying, M., & He. (2009). Efficient indexing for large scale visual search. IEEE 12th International Conference on Computer Vision, (pp. 1103-1110).