March 2023

DIGITAL IMAGE PROCESSING

EDITED BY
DR. ABHIRUP BHAWAL

CIIR Research Publications

Contents

	Title of Chapters	Page (s)
Chapter 1	IMAGE IMPROVEMENT Mr. Manikandan T.	1
Chapter 2	BASED ON CONTENT IMAGE RETRIEVAL Mr. Manikandan T.	4
Chapter 3	VECTOR EXTRACTION IN IMAGE ENHANCEMENT Dr. Abhirup Bhawal	7
Chapter 4	VECTOR MATCHING FEATURE Dr. Abhirup Bhawal	10
Chapter 5	SPECIFICATION FOR HISTOGRAM Dr. Abhirup Bhawal	13
Chapter 6	OBJECTIVE EVALUATION METRICS REQUIRING A REFERENCE IMAGE Ms. Ashwini B.	16
Chapter 7	CLASSIFICATION OF MAJOR APPLICATIONS Ms. Natya S.	19
Chapter 8	PHOTOGRAPHY APPLICATION Mr. Syed Abrar Ahmed	22
Chapter 9	APPROACHES FOR PIXEL-LEVEL IMAGE FUSION Ms. Natya S.	25
Chapter 10	FUSION TECHNIQUE IN SIGNAL TRANSFORM SCHEME Dr. Bhanurekha K.	28
Chapter 11	INTRODUCTION TO IMAGE DEFINITION Ms. Akshatha K.	31
Chapter 12	BENEFITS OF IMAGE PROCESSING Dr. Azra Jeelani	34
Chapter 13	IMAGE FUSION Dr. Muthupandi G.	37
Chapter 14	APPLICATIONS OF IMAGE FUSION Dr. Muthupandi G.	40
Chapter 15	INTRODUCTION TO DWT, PCA AND NSCT METHODS Dr. Sumantra Chaudhuri	43

Preface

The term "digital image processing" refers to the use of a digital computer to process digital images. In order to get an improved picture or to extract some important information, we can also state that it is the usage of computer algorithms. The employment of algorithms and mathematical models in digital image processing allows for the processing and analysis of digital pictures. Digital image processing aims to improve picture quality, extract useful information from photos, and automate operations using images.

The fundamental procedures in digital image processing can be many like picture acquisition is either taking a digital photo, scanning a document, or transferring an already-existing picture into a computer. Image improvement is enhancing an image's aesthetic appeal by boosting contrast, cutting down on noise, and eliminating artifacts, for example. Removal of visual deterioration, such as blurring, noise, and distortion, is known as image restoration. Segmenting a picture entail breaking it into sections or regions, each of which corresponds to a particular item or characteristic in the image. Image representation and description—This entails displaying an image in a manner that permits computer analysis and manipulation, as well as succinctly and meaningfully expressing an image's characteristics. The process of extracting information from a picture, such as identifying objects, spotting patterns, and measuring characteristics, is known as image analysis. Image synthesis and compression: These techniques create new pictures or scale down already-existing images to demand less space for storage and transmission.

Many different applications, such as medical imaging, remote sensing, computer vision, and multimedia, heavily rely on digital image processing. Digital image processing benefits—Image quality may be enhanced using digital image processing techniques, which can make pictures clearer, sharper, and more informative. Object identification, pattern detection, and measurement are just a few of the picture-based jobs that may be automated thanks to digital image processing. Enhanced efficiency—Digital image processing algorithms can process pictures considerably more quickly than people can, allowing for the quick analysis of vast volumes of data. Enhanced precision—When doing jobs that need for exact measurements or quantitative analysis, digital image processing algorithms often outperform humans in terms of accuracy.

Digital image processing drawbacks can also be a matter of concern. High computational cost—Some algorithms used in digital image processing are very resource-intensive and computationally expensive. Limited interpretability—Particularly for complicated or advanced algorithms, certain digital image processing methods may give findings that are difficult for humans to understand. Dependence on input quality—The quality of the photos used as inputs determines much how well digital image processing algorithms perform. Images with low quality input might produce poor quality output. Algorithm limitations—Digital image processing algorithms have several drawbacks, such as the inability to recognize objects with large deformations or occlusions, or the difficulty in identifying objects in crowded or dimly light situations. Dependence on quality training data—A lot of digital image processing algorithms rely on the accuracy of the training data that was used to create them. Poor training data might cause the algorithm to perform poorly.

Dr. Abhirup Bhawal
Editor

CHAPTER 1

IMAGE IMPROVEMENT

Mr. Manikandan T.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- manikandan.t@presidencyuniversity.in

The main goal of image enhancement is to process a picture in such a manner that it is more suited for the intended use than the original image. There is meaning in the term "particular." It hints that the outcomes of such an operation will vary greatly depending on the application. In other words, an image-enhancing method that is effective for topographic X-ray pictures may not be effective for MR images [1]. Based on the domain in which they are used, the approach may be divided into two types. These are the realms of frequency and space. The image's Fourier Transforms are used using the frequency domain techniques. The totality of pixels that make up a picture is referred to as the spatial domain. Techniques that work directly on pixels are referred to as spatial domain methods. The formula for the process is $g(x, y) = T[f(x, y)]$. Where T is an operator on f defined across some region of the input picture, $f(x, y)$, the processed image, and (x, y). There are several improvement methods available for the spatial domain. These include histogram processing, arithmetic-based augmentation, and logical operations and filters [2].

The current system only uses one of the three possible strategies, such as bit plane slicing, histogram equalization, or contrast modification, to improve colour images. When they are used separately, choosing the best improvement method becomes challenging. Choosing the next degree of improvement for individual development is challenging. The suggested method combines all three improvement strategies into a single algorithm, and we can additionally boost the final image's resolution. A single procedure may be used to retrieve both the colour and intensity characteristics. The current system only uses one of the three possible approaches, such as bit plane slicing, histogram equalisation, or contrast modification, to improve colour images [3]. When they are carried out separately, it becomes challenging to choose the right enhancement procedure and to choose the subsequent degree of enhancement. The proposed approach combines all three enhancement strategies into a single algorithm; we can also boost the output image's resolution; and we can derive both the intensity and colour parameters from the same algorithm.

Advantages

Enhance the quality of the picture after enhancement; Switch from one parameter to another by just altering the available choices in the output window; When compared to one of the individual enhancement approaches, images collected in low light may be improved beautifully and minute detail can be achieved. This technology is commonly used in public security systems.

Slicing a Bit Plane

Modern technology has created a requirement for the effective storage of a lot of digital photos. When T. Kato described investigations into automated retrieval of pictures from a database based

ISBN 978-81-962236-5-6

on the colours and forms present, the term CBIR seems to have first been used. Since then, it has come to mean the method of selecting desirable photos from a big collection based on syntactical image characteristics. Statistics, pattern recognition, signal processing, and computer vision are only a few examples of the disciplines from which the methods, tools, and algorithms are drawn.

Visual information sources are used by historians from a range of fields, including art, sociology, medicine, etc., to aid in their study. Images are also used extensively by archaeologists. In certain cases, the visual evidence may be the only one that is accessible. Researchers must use substitutes in the form of photographs, slides, or other images of the objects, which may be amassed within a specific library, museum, or art gallery, when access to the original works of art is constrained or impossible, perhaps due to their geographic distance, ownership restrictions, or factors related to their physical condition b[4].

Many different institutions, including academic and public libraries, have photographic and slide collections. Some people consider CBIR to be a subset of the area of image computing and computer vision since it borrows many of other fields' methodologies. It varies from these disciplines primarily in that it places a strong focus on the recovery of pictures with required properties from a large collection. A significantly broader range of activities is included in image processing, such as picture enhancements, compressing, transmission, and interpretation.

Although there are some grey areas such as object identification via feature analysis, the line separating CBIR from mainstream image analysis is often rather obvious. This might be shown with an example. Automatic facial recognition technologies are now widely used by police departments. One of two methods may be used for such systems. First, to confirm a person's identification, the picture in front of the camera may be compared with that person's database information. Few observers would refer to the matching of these two photos, which only involves two images, as CBIR.

The second option is to scan the whole database for the most similar photographs. This is a true instance of CBIR. Numerous concerns are common to both traditional image processing and information retrieval and the research and development challenges in CBIR. Understanding the demands and information-seeking behaviours of picture users; finding appropriate methods to describe image content; extracting such characteristics from raw photos; and offering compact storage for huge image databases are some of the most crucial [5].

- Comparing stored and requested photos in a manner that takes human similarity assessments into account
- Having effective human interfaces for CBIR systems; efficiently retrieving stored pictures according to content; and.

Image Retrieval Based On Text

The majority of text descriptors or classification codes now used in picture indexing are supported, in some instances, by text retrieval packages that have been specifically created or modified to work with images.

Again, relatively little information on the efficacy of such systems has been made public. The degree of user satisfaction with these systems seems to be quite variable. Manual annotation was necessary for each picture when using text-based image retrieval. It takes a long time to use the text string to describe each picture. A text string that depends on the user explains the content of the picture. Text strings for the same picture by multiple users will change because different users' perceptions of an image vary. Therefore, content-based image retrieval is necessary. The discipline of content-based image retrieval (CBIR), which uses computer vision to retrieve pictures from digital libraries, is a rapidly expanding one [6].

Bibliography

[1] R. Grohs, "Drivers of brand image improvement in sports-event sponsorship," *Int. J. Advert.*, vol. 35, no. 3, pp. 391–420, May 2016, doi: 10.1080/02650487.2015.1083070.

[2] N. V. Slavine, S. J. Seiler, R. W. McColl, and R. E. Lenkinski, "Image improvement method for positron emission mammography," *Phys. Medica*, 2017, doi: 10.1016/j.ejmp.2017.06.025.

[3] L. Liu, Y. Sang, X. Sun, and B. Wang, "A hybrid remote sensing image improvement algorithm in view of NSST," 2021. doi: 10.1088/1742-6596/2026/1/012048.

[4] J. Tsuchiya, K. Yokoyama, K. Yamagiwa, R. Watanabe, K. Kimura, M. Kishino, C. Chan, E. Asma, and U. Tateishi, "Deep learning-based image quality improvement of 18F-fluorodeoxyglucose positron emission tomography: a retrospective observational study," *EJNMMI Phys.*, 2021, doi: 10.1186/s40658-021-00377-4.

[5] M. Mozaffarzadeh, B. Makkiabadi, M. Basij, and M. Mehrmohammadi, "Image improvement in linear-array photoacoustic imaging using high resolution coherence factor weighting technique," *BMC Biomed. Eng.*, 2019, doi: 10.1186/s42490-019-0009-9.

[6] S. Kida, T. Nakamoto, M. Nakano, K. Nawa, A. Haga, J. Kotoku, H. Yamashita, and K. Nakagawa, "Cone Beam Computed Tomography Image Quality Improvement Using a Deep Convolutional Neural Network," *Cureus*, 2018, doi: 10.7759/cureus.2548.

CHAPTER 2

BASED ON CONTENT IMAGE RETRIEVAL

Mr. Manikandan T.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- manikandan.t@presidencyuniversity.in

The application of computer vision to the image retrieval issue, that is, the challenge of finding digital pictures in huge databases, is known as content-based image retrieval (CBIR), sometimes referred to as query by image content (QBIC) and content-based visual information retrieval (CBVIR). "Content-based" refers to the search process' analysis of the image's actual contents. Colours, forms, textures, and any other information that can be inferred from the picture itself are all examples of content in this context. Searches must depend on information like captions or keywords, which may be time-consuming or costly to prepare, in the absence of the capacity to study the content of the images. Journalists, design engineers, and art historians are just a few of the professions that need to discover a certain picture from a collection [1].

The population phase and retrieval phase are the first two tasks in content-based image retrieval. Every picture in the database has a population phase feature vector (signature) taken from it. The picture is faithfully represented by this feature vector. The feature vector should have a lower file size than the picture. For all of the photos, create a feature vector database. Take a query picture in the second task, calculate the feature vector, and determine the Euclidean distance (similarity). It is possible to extract the photos with the smallest Euclidean distance [2]. Calculate the interband average picture for a given colour image using the equation Calculate the single-bit plane using this interband average picture. This bit plane has the same size as the provided colour picture in both the "1" and "0" forms. Determine the row and column pixel distributions for a given bit plane. The binary row vector's number of ones is called a row vector. The column vector consists of all 1s in the appropriate columns.

The representation of a colour image or a grayscale picture is an M x N intensity matrix, where each element has a distinct intensity value, ranging from a grayscale image to a colour image. The binary image, commonly referred to as an 8bit, is a representation of an image's intensity values for each pixel as represented by a single bit, either "1" or "0."

Various methods are used to transform a given picture into a bit plane image (binary image). Using a threshold, a single-bit planar picture is created. The picture intensity value is the mean of this threshold. We start by thinking about the inter-band average picture. Red, Green, and Blue colour plane averages are used to create this interband average picture [3].

BIT PLANE SLICING

Low contrast picture that was improved using image processing software. However, this approach brightens every pixel in the provided picture, resolving the issue with bitplane slicing. A picture is divided into eight binary planes via bit-plane slicing.

ISBN 978-81-962236-5-6

It is obvious that an 8-bit binary vector (b7, b6, b5, b4, b3, b2, b1, b0), where I is from 0 to 7, and each bi is either "0" or "1," may represent the intensity value of each pixel. A picture in this situation may be thought of as an overlay of eight bit-planes. Each bit-plane may be thought of as a binary matrix that represents a two-tone picture.

Given the bit plane's formation where R=Remainder and round, I bp (i,j)= Bit-Plane information

For each grayscale picture, we are obtaining 8-bit planes if we round the components to the closest integer. Therefore, we may generate a total of twenty-four-bit planes for the colour picture if we split out R, G, and B colour components and create 8-bit planes for each colour component [4].

FEATURE VECTOR EXTRACTION

What qualifies as a characteristic for sign recognition has no clear-cut, universal criteria. It often depends on the issue or the language. An "interesting" portion of an image is referred to as a feature, and it serves as the basic building block for all subsequent algorithms. In many cases, the total algorithm is only as good as its feature detector. The repeatability of a feature detector, or whether it will be identified in two or more distinct photographs of the same scene, is hence its desired attribute.

When attempting to identify the indications, the spatial, temporal, and textural elements are the most crucial categories that may be taken into account. The feature extraction step was developed and engineered to work with actual photos. These systems' algorithms are often broken down into three tasks: extraction, selection, and classification. It is the most important assignment since the specific characteristics made available for discrimination directly affect the effectiveness of the classification job. Proper classification requires a reasonable nexus between the features [5].

The definition of the feature is that it is a function of one or more measurements, each of which defines a measurable quality of an item, and that it is calculated such that it quantifies certain important qualities of the object. The many characteristics that have been categorized and are in use include general traits: independent characteristics like form, texture, and colour they may also be further classified into the following categories:

- **Features at the pixel level:** Features, such as colour and position, are determined at each pixel.
- **Local traits:** Features derived from the outcomes of the picture band's subdivision for edge detection.
- **Image Segmentation General Traits:** Features that are estimated across an image's whole surface or simply a regular subarea.

Domain-specific characteristics use reliant traits such as conceptual ones, fingerprints, and human faces. Low-level features and high-level features may be used to broadly categorise all features. While high-level feature extraction relies on low-level feature extraction, low-level feature extraction may be directly retrieved from the source pictures. The following considerations should be used as a reference for selecting the features from the retrieved vector:

ISBN 978-81-962236-5-6

- The extracted features should not need any domain-specific expertise and should provide sufficient information about the picture.
- They should connect well to the human perceptual features as users ultimately decide if the recovered photos are appropriate.

They should be simple to calculate to approach the practicality of a big Image collection and speedy retrieval.

- One of the most prevalent visual elements in images is the colour component.
- Classification Images with distinctive colour characteristics have various benefits.
- Robustness: The colour histogram changes gradually as the picture is rotated or scaled and is invariant to the rotation of the image around the view axis.
- Effectiveness: There is a high proportion of relevancy between the query picture and the retrieved matched photos.

Implementation simplicity is a direct process used to create the colour histogram, which involves scanning the picture, assigning colour values to the histogram's resolution, and creating the histogram using colour components as indices. Simple computations: For photos of size x y, the complexity of the histogram calculation is O(x, y). A single picture match has a linear complexity of O (n), where n is the resolution of the histogram or the number of distinct colours [6].

Bibliography

[1] S. R. Dubey, “A Decade Survey of Content Based Image Retrieval Using Deep Learning,” *IEEE Trans. Circuits Syst. Video Technol.*, 2022, doi: 10.1109/TCSVT.2021.3080920.

[2] L. Piras and G. Giacinto, “Information fusion in content based image retrieval: A comprehensive overview,” *Inf. Fusion*, 2017, doi: 10.1016/j.inffus.2017.01.003.

[3] H. KOYUNCU, M. DİXİT, and B. KOYUNCU, “An analysis of content-based image retrieval,” *Int. Adv. Res. Eng. J.*, 2021, doi: 10.35860/iarej.811927.

[4] K. Iida and H. Kiya, “Privacy-Preserving Content-Based Image Retrieval Using Compressible Encrypted Images,” *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3035563.

[5] C. Burak Akgül, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar, “Content-based image retrieval in radiology: Current status and future directions,” *J. Digit. Imaging*, 2011, doi: 10.1007/s10278-010-9290-9.

[6] A. Ahmed and S. Mohamed, “Implementation of early and late fusion methods for content-based image retrieval,” *Int. J. Adv. Appl. Sci.*, 2021, doi: 10.21833/ijaas.2021.07.012.

CHAPTER 3

VECTOR EXTRACTION IN IMAGE ENHANCEMENT

Dr. Abhirup Bhawal
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- abhirup.bhawal@presidencyuniversity.in

Minimal storage needs Assuming colour quantization, the size of the colour histogram is substantially lower than the picture itself. The shape is a crucial visual element as well as one basic element of a picture. Shape content description is tough to describe precisely because it is challenging to compare forms' similarities. Therefore, feature extraction and similarities assessment between the different classifiers are two phases that are crucial in shape-based picture retrieval. Region-based and contour-based shape descriptors, which employ local characteristics as boundary segments to describe an object's form respectively, may be classified into two primary types [1].

A representation, which is a collection of features that serve as an image representation, is what emerges after feature extraction. The following list includes the four methods utilised in this article for feature vector extraction:

- Feature vector computation using an inter-band average image's single-bit plane.
- Calculation of feature vectors using a three-bit colour picture plane.
- Calculation of feature vectors using an eight-bit grayscale picture plane.
- Feature vector computation using a 24-bit colour picture plane.

Because of this, the feature vector size for each approach varies. For the single-bit plane, it is little, while for the sixteen-bit planes, it is big. By constructing the similarity matrix using these four approaches, we are comparing the performance of the image retrieval system [2].

Average Image for Feature Vector Extraction

Calculate the interband average picture for the provided colour image using the equation. Using this interband average image, calculate the single-bit plane as described in section IV (A).

- This bit plane has the same size as the provided colour picture in both the "1" and "0" forms.
- Determine the row and column pixel distributions for a given bit plane.
- The binary row vector's number of ones is called a row vector. The column vector consists of all 1s in the appropriate columns.

One row and one column of coefficients are obtained for the bit pane after the computation of the column- and row-wise pixel distribution. For example, SBP column (1, j) is the columnist-wise pixel distribution for single bit plane and SBP row (i,1) is the row-wise pixel distribution. It is very important to segment a picture and estimate the attributes of individual image sections [3].

The measure known as "region properties" is a logical array with any number of dimensions that specifies the attributes for each linked component (object) in the binary picture.

Feature Vector Extraction

Using Three Bit Color Image Planes

In this instance, a feature vector was taken from a distinct colour component. Section IV describes the separate single-bit planes for the red, green, and blue components (B). Therefore, each colour picture created using this approach has three-bit planes.

Calculate the column and row pixel distributions for the specified bit planes. The binary row vector's number of ones is called a row vector. The column vector consists of all 1s in the appropriate columns. We get a vector of three rows and three columns of coefficients after computing the pixel distribution row- and column-wise for each bit plane. Determine the mean, standard deviation, and cube root of the third moment for each row and column vector for each bit plane feature vector with a size of 18 (six mean values, six standard deviation values, and six third moment values) using the equations provided in section V (A) [4].

As a result, there are eight-row vectors with eight means, eight standard deviations, and eight cube roots of the third-moment value, and eight columns with eight means, eight standard deviations, and eight cube roots of the third-moment value. Therefore, the size of the feature vector when employing this approach for an eight-bit grayscale picture.

We calculated the picture bit planes' mean, standard deviation, and cube root of pixel distribution. For each picture pixel, this bit plane creation differs from others by employing a binary vector instead of the mean. When the mean is used as a threshold, there is only yet another bit plane for interband average images and one bit plane for colour images. When each pixel in the picture is represented by an 8-bit binary vector, an 8-bit plane for grayscale images and a 24-bit plane for colour images must be created [5].

Each intensity value has an 8-bit binary array, as we are aware. There are eight-bit regions for each image's intensity. As stated in section IV(C), the majority of colour pictures are thus captured using a reference RGB colour space. There are eight-bit planes for each colour plane. To examine a colour picture, each RGB colour plane has a total of 24-bit planes. The same process should be applied to each of the 24-bit planes that we currently have. In this instance, the feature vector consists of 48 rows and column mean values, 48 columns mean and standard deviation values, and 48 cube roots of the third-moment value. Each colour picture has a row and column vector. The feature vector is 144 bytes long overall [6].

Bibliography

[1] R. Scheibler and N. Ono, “Fast Independent Vector Extraction by Iterative SINR Maximization,” 2020. doi: 10.1109/ICASSP40776.2020.9053066.

[2] X. Zhang, X. Zou, M. Sun, T. F. Zheng, C. Jia, and Y. Wang, “Noise Robust Speaker Recognition Based on Adaptive Frame Weighting in GMM for i-Vector Extraction,” *IEEE Access*, 2019, doi: 10.1109/ACCESS.2019.2901812.

[3] W. Li, T. Fu, and J. Zhu, "An improved i-vector extraction algorithm for speaker verification," *Eurasip J. Audio, Speech, Music Process.*, 2015, doi: 10.1186/s13636-015-0061-x.

[4] O. Glembek, L. Burget, P. Matějka, M. Karafiát, and P. Kenny, "Simplification and optimization of i-vector extraction," 2011. doi: 10.1109/ICASSP.2011.5947358.

[5] S. Cumani and P. Laface, "Memory and computation trade-offs for efficient i-vector extraction," *IEEE Trans. Audio, Speech Lang. Process.*, 2013, doi: 10.1109/TASL.2013.2239291.

[6] P. Liang, W. Shi, Y. Ding, Z. Liu, and H. Shang, "Road extraction from high resolution remote sensing images based on vector field learning," *Sensors*, 2021, doi: 10.3390/s21093152.

CHAPTER 4

VECTOR MATCHING FEATURE

Dr. Abhirup Bhawal
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- abhirup.bhawal@presidencyuniversity.in

When a user submits a query picture, we must calculate the feature vector again and compare it to the feature vector that has already been computed and stored in the database. In Fig. 8, this is shown. The morphological operation, feature vector storage process, and similarity measure procedure are shown in the retrieval process block diagram. The following principles form the basis of the feature extraction procedure. Covers the following phases of the batch feature extraction and storing procedure [1].

- Pictures pulled one at a time from the database.
- The feature extraction procedure is used to calculate the feature.
- For the database photos provided, create a feature vector database.

CBIR'S Performance

Precision and recall are two metrics that may be used to evaluate an image retrieval system's performance. How effectively the system avoids returning results that are completely unrelated to the source from the user's point of view and how well the system can retrieve photos that are comparable to the query image are the two main factors in determining an image search engine's efficacy.

A representative query picture must be chosen from one of the database's image categories. The user must count the number of photos returned by the search engine once it returns the results, as well as the proportion of those images that are similar to the query image. The first test is known as Recall. The database recalls all of the pertinent photos. The following equation may be used to calculate recall. Precision is the name of the second metric. The following mathematical formula describes how accurately a retrieval system presents both relevant and irrelevant photos from the database [2].

Using MATLAB 7.0 and a computer with an Intel Core 2 Duo Processor T8100 (2.1GHz) and 2 GB RAM, the CBIR approach is implemented. The CBIR method is evaluated on a library of 800 photos of different sizes, including 8 categories of animals, buses, flowers, bicycles, beaches, historical sites, and mountains, among others.

For 40 query photos (five from each category in the database), sample images from each category. The precision and recall for the suggested approaches are computed, and the average recall precision is shown versus category. The average recall and accuracy of this CBIR approach serve as significant performance indicators [3].

When we take into account the total average recall and accuracy for all approaches, as indicated in table I, charts are shown. It was shown that while bit plane creation keeps rising, both general average accuracy and recall rise.

An example of a histogram that visualizes the tonal distribution in a digital picture is an image histogram. Each tonal value's number of pixels is plotted. A frequency distribution is illustrated by a histogram [4].

The discrete function is used to construct the histogram of a digital picture with G total potential intensity levels in the range where rk is the original raw image's intensity level. The number of pixels in the picture with an intensity level of nk is equal to the total number of pixels in the image.

Histogram Equalization Description

Enhancing the contrast of the picture using histogram equalisation distributes the intensity values throughout the whole range. Because the histogram equalisation approach only adds additional pixels to the bright portions of the picture and subtracts extra pixels from the dark sections of the image, the produced image has a large dynamic range and cannot be utilised for photographs with backgrounds that are illuminated unevenly. Histogram equalisation aims to distribute an image's contrast uniformly over the whole dynamic range that is accessible.

It is the probability density function (pdf) that is being changed in the histogram equalisation procedure. To put it simply, the histogram equalisation approach converts a particular image's probability density function (pdf) into a uniform distribution that spreads out from the lowest pixel value (0 in this example) to the highest pixel value (L 1). If the pdf is a continuous function, then this may be accomplished relatively quickly [5].

The pdf, however, will be a discontinuous function since we are working with a digital picture. Let's say we have a picture x and the intensity's dynamic range is from 0 (black) to L1 (white). The probability p(rk) based on the histogram may be used to approximate this pdf as follows: One of the most popular techniques for processing coloured images in this scheme. The RGB space's channels are treated separately using histogram equalisation. After equating the R, G, and B components, we combine all three and get a better picture than the original image.

This method requires converting the picture to the hue, saturation, and luminance (HSV) colour space before processing it in RGB. In this case, brightness and intensity are synonyms. The perception associated with the predominant wavelength of the colour stimulus is represented by hue. The degree of colour purity is shown by saturation (the amount of white light in the colour).

The chromaticity coordinates are defined as the combination of hue and saturation (polar system). This approach uses the HSV colour space with the V component of the histogram equalized. We merge the V with H and S after adjusting the V. The following picture is superior to the supplied image [6].

Bibliography

[1] S. Walia and K. Kumar, "Calibrating thresholds based on trade-offs between detection accuracy and fpr for copy move forgery detection," *Int. J. Recent Technol. Eng.*, 2019, doi:

10.35940/ijrte.B2083.078219.

[2] J. Liu, C. Zhang, R. Zhang, Y. Li, and J. Cheng, "A video steganalysis method based on coding cost variation," *Int. J. Distrib. Sens. Networks*, 2021, doi: 10.1177/1550147721992730.

[3] Y. Zhou, Y. Song, S. Cui, H. Zhu, J. Sun, and W. Qin, "A Novel Change Detection Framework in Urban Area Using Multilevel Matching Feature and Automatic Sample Extraction Strategy," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2021, doi: 10.1109/JSTARS.2021.3064311.

[4] G. Kalakoti and G. Prabakaran, "Key-frame detection and video retrieval based on dc coefficient-based cosine orthogonality and multivariate statistical tests," *Trait. du Signal*, 2020, doi: 10.18280/ts.370509.

[5] Y. Li, P. Hao, S. Zhang, and Y. Li, "Feature-matching pattern-based support vector machines for robust peptide mass fingerprinting," *Mol. Cell. Proteomics*, 2011, doi: 10.1074/mcp.M110.005785.

[6] W. K. Tam and H. J. Lee, "Accurate shade image matching by using a smartphone camera," *J. Prosthodont. Res.*, 2017, doi: 10.1016/j.jpor.2016.07.004.

CHAPTER 5

SPECIFICATION FOR HISTOGRAM

Dr. Abhirup Bhawal
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- abhirup.bhawal@presidencyuniversity.in

The histogram equalisation approach is expanded upon by histogram specification. The goal of the histogram equalisation approach is to get the output histogram to follow a uniform PDF. However, when specifying a histogram, we want the resulting histogram to match that histogram exactly. To do this, the input picture is first histogram equalised, and the resultant equalised image's PDF is then matched to the required histogram's PDF. Let z be the pixel value for the desired histogram, and let r and s be the pixel values for the input and equalised images, respectively [1].

He is often the approach that is employed to increase contrast in digital photos, however, the results it produces are artificial artefacts including intensity saturation, over-enhancement, and noise amplification. The picture histogram had to be divided into two sections to solve these issues. Subsequently, various transformation functions were applied to each component of the split image histogram. Following that, the picture histogram was divided into several divisions and the same procedure was used with a few more characteristics. AHE is an extension of DHE, which is a multi-histogram approach. When these approaches are compared against HE, it is discovered that not only do both methods outperform HE, but the DHE method also outperforms the CLAHE method [2].

Image enhancement is a procedure used to identify the hidden information included in photographs and eliminate the undesired distortion caused by degradation, in contrast, unwanted noise, inappropriate intensity saturation, blurring effect, etc. Additionally, image enhancement produces superior visual pictures that are used as input by numerous programmes for processing digital photos. The ratio between an image's greatest and lowest pixel intensities is one way to describe the picture element of contrast. In an image processing system, improving contrast is a taxing activity that must be completed, and obtaining a good contrast picture is highly challenging. A picture may have poor contrast for several reasons, including the use of low-quality imaging technology. The lack of centred pixel intensity, poor lighting, specimen spotting, and other characteristics that may be changed by image enhancement techniques often result in poor picture quality, a decline in contrast, and the appearance of shading and artefacts [3].

The look of a picture worldwide is described by its histogram. Techniques for restoring the decreased picture contrast are often utilised, including histogram equalisation and linear contrast modification. However, none of these strategies is appropriate for preserving the natural contrast; instead, they may introduce undesired artefacts and artificial contrast. The histogram provides a wealth of information about an image's properties, hence histogram modelling using a spatial domain approach is crucial to DIP and has been implemented on many systems. The many forms of histogram equalisation techniques are researched and contrasted in this study.

ISBN 978-81-962236-5-6

Better than the bi-histogram equalisation approach is the multi-histogram equalisation technique known as "dynamic histogram equalisation" (DHE). A clip histogram equalisation method that is superior to traditional AHE is contrast limited adaptive histogram equalisation (CLAHE). The CLAHE method extends the adaptive histogram equalisation technique by adding a few more parameters that help to solve the adaptive histogram equalisation issue [4].

Summary of Histogram Equalization

The histogram of the output picture is systematically flattened and stretched using the spatial domain technique known as histogram equalisation, which results in an output image with a uniform distribution of pixel intensity.

Due to its simplicity and relative superiority over other conventional approaches, this strategy often produces positive results for image improvement paradigms. The input picture histogram allows us to get the probability density function (PDF) and cumulative density function (CDF). Create the processed picture and histogram for the final image by using these two functions, PDF and CDF, to replace the grey levels in the input image with the new grey levels. And we discovered that the grey level intensities are consistently extended and lowered when we contrast the input picture histogram with the modified image histogram. As a result, we discover that the output image's histogram is predictably distributed. However, this permits over-enhancing pictures over the range of real grayscale [5].

The middle grey level is always the processed image's mean brightness during the histogram equalisation procedure, regardless of the input mean. Because this technology tends to generate unnecessary visual degradation like the concentration effect, it is not particularly practical to use it in consumer electronics like television.

The specific solution for this problem is to keep the average brightness of the input picture indoors in the output image to overcome this limitation.

Dynamic histogram equalisation

By applying local minima separation of the histogram, Dynamic Histogram Equalization (DHE), which Wadudetal popularised in 2007, aims to reduce the effect of higher histogram components on lower histogram components in the picture histogram and control the amount of grey level spreading. DHE exhibits continuous and superior picture improvement compared to the conventional paradigm. Additionally, the DHE controls mean brightness persistence and affect intensity saturation artefacts.

The shortcomings of histogram equalisation have been solved by the DHE approach, which has shown superior brightness preservation and contrast improvement versus HE. DHE does not alter the image in any way; it just strengthens it. If the user is unsatisfied, they may change simply one parameter to limit the degree of the improvement. Additionally, DHE is transparent and computationally sound, making it simple to develop and suitable for use in real-time systems.

The histogram is separated using DHE based on local minima. Formally, it smooths out the histogram in one dimension to get rid of pointless minima. After that, it creates sub-histograms using the histogram's section that is between two local minima [6].

Bibliography

[1] K. Hussain, S. Rahman, M. M. Rahman, S. M. Khaled, M. Abdullah-Al Wadud, M. A. Hossain Khan, and M. Shoyaib, “A histogram specification technique for dark image enhancement using a local transformation method,” *IPSJ Trans. Comput. Vis. Appl.*, 2018, doi: 10.1186/s41074-018-0040-0.

[2] D. Coltuc, P. Bolon, and J. M. Chassery, “Exact histogram specification,” *IEEE Trans. Image Process.*, 2006, doi: 10.1109/TIP.2005.864170.

[3] C. C. Sun, S. J. Ruan, M. C. Shie, and T. W. Pai, “Dynamic contrast enhancement based on histogram specification,” *IEEE Trans. Consum. Electron.*, 2005, doi: 10.1109/TCE.2005.1561859.

[4] H. D. Liu, M. Yang, Y. Gao, and C. Cui, “Local histogram specification for face recognition under varying lighting conditions,” *Image Vis. Comput.*, 2014, doi: 10.1016/j.imavis.2014.02.010.

[5] Y. Wan and D. Shi, “Joint exact histogram specification and image enhancement through the wavelet transform,” *IEEE Trans. Image Process.*, 2007, doi: 10.1109/TIP.2007.902332.

[6] G. Thomas, D. Flores-Tapia, and S. Pistorius, “Histogram Specification: A Fast and Flexible Method to Process Digital Images,” *IEEE Trans. Instrum. Meas.*, vol. 60, no. 5, pp. 1565–1578, May 2011, doi: 10.1109/TIM.2010.2089110.

CHAPTER 6

OBJECTIVE EVALUATION METRICS REQUIRING A REFERENCE IMAGE

Ms. Ashwini B.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- ashwini.b@presidencyuniversity.in

In certain situations, an ideal fused photograph may already exist or be manually created. This image could serve as a baseline to assess very well how image fusion works. For illustration, the input multi-spectral (MS) and panchromatic (PAN) pictures may first deteriorate in remote sensing picture fusion. The deteriorated photographs are then combined and shown side by side with the original. In certain specific scenarios of multi-focus face recognition, a reference fused picture may very well be created by manually dividing the input image [1]. Various objective nuclear fission metrics may be employed because when reference fused imagery is provided it is known as full-reference quality metrics. Two famous ones are the maximal signal-to-noise ratio (PSNR) and the root-mean-square error (RMSE). The different methods, however, have been discovered to be poorly associated with sensory consciousness in a few unique situations. Overall manual segmentation conclusions of the fused photographs are therefore assessed using the generated background model [2].

The algorithm is the second iteration of a wavelet-based fusion technique. Through the assessment of two quality indices a spectral index (ERGAS index) and a spatial one it enables personalization of the trade-off between the spatial and temporal quality of the fused picture. Comparative between the performance of the hypothesized fusion approach and others based on wavelet Mallet and Fourier domain filtering have been made. Its algorithm is quite uncomplicated and just uses basic algebraic operations like products, variances, and convolutions. Especially in comparison to other fusion approaches that need more complicated processing, it produces data with excellent spatial and spectral quality [3]. The suggested procedure requires three essential steps: To begin started, a crude segmentation conclusion is created using the focus characteristics from each different component that was collected by morphological processing. After that, the picture embossing method is utilized to acquire the precise creation of conducive correlations between neighbouring pixels. The fused picture would then be created by combining the specific areas.

Although many complete picture fusion measures have been put proposed, the following issues have remained unresolved:

- Different image fusion algorithms may not be able to incorporate various metrics because they may be excellent for detecting certain types of picture instability caused during the fission reaction but inferior for everyone else.
- There are no effective methods for determining the quality of cross-resolution images. For instance, the full-reference image integration metrics in software for remote sensing can only be used at a medium speed.

- The majority of today's full-reference image fusion measurements cannot be employed directly for photographs with many channels, which include colour, multi-spectroscopy, hyper-spectrum, and four different volumes [4].

OBJECTIVE EVALUATION METRICS WITHOUT REQUIRING A REFERENCE IMAGE:

The ideal fused frame isn't always accessible in the preponderance of image fusion applications. Therefore, objective image reconstruction quality evaluation performance indicators described as no-reference excellence metrics that do not assume the maintenance of a fused picture that seems to be ground truth are extremely desired. The recommended pooling approach is ubiquitous and may be used with any fusion quality and effectiveness measure based on local discrepancies between the incoming and fused picture and assumptions of structure preservation. We establish that for all the structural nuclear fission measures studied, performance may indeed be improved by regionally integrating local performance results spanning 3° to 6° of visual angle and designating the lowest performing section as the global score. The frequency and severity are then applied to these local pictures. A human central visual function is used to weigh the frequency domain discrepancy between the local and metropolitan pictures (CSF). The MSE of the weighted discrepancy pictures derived from the combined regional image and primary regional images is used to measure the quality of a local regional image [5]. The weighted combination of the local and regional image quality indicators, in the end, determines the quality of a fused picture. A metric based on mutual information, whereas mutual information (MI) needs to be measured of the separation seen between joint summary statistics of two variables. A corrected MI measure that additionally takes the permeability variance between the many source pictures into consideration when assessing results. A more comprehensive definition of the correlation-based feature measure for the information transfer from either the input pictures to the fused conclusion using Tsallis entropy. For fusion evaluation criteria, it presented a nonlinear regression equation that is comparable here to the idea of information gain. This links the assessment of the educational ecosystem [6].

In recent years, measures for determining the retention of structural properties in fused images have been constructed, based on the structural similarity measures (SSIM). To be more precise, the SSIM maps between each various component and fused image are first created on a pixel-by-pixel basis. The performance of the fusion process may be measured globally using a spatial pooling methodology. A visual information fidelity (VIF)-based image fusion measurement that has shown strong performance for image resolution. A comparison of several image reconstruction metrics for applications that require night vision information fusion. Theoretical experiments were carried out by Wei and Blum for variable quality indicators used in the system that will allow the fusion technique. The usefulness of several fusion measurements on a database of multiple distribution images [7]. For the currently available non-reference objective fusion metrics, there are still various difficult issues that need to be handled. The majority of current quantifiable fusion measures are limited to monochromatic pictures. The photos some of which are taken as input in certain apps are dynamic. Using the given test dataset, the proposed fusion refutes hypotheses about a specific is contrasted with something like several current fusion measurements. Each block evaluates input visual data with and without distortion data. Each sub-visual band's informational fidelity for fusion (VIFF) is estimated. There are two principal categories for the no-reference fusion measurements now in use:

- Metrics based on knowledge theory only take into account the total amount of data that is communicated from the inputs to the combined picture [8].
- Fusion metrics are influenced by local features. This assessment determines the proportion of properties from the input pictures that are communicated to the fused image and are susceptible to the human visual system.

Bibliography

[1] H. M and S. M.N, "A Review on Evaluation Metrics for Data Classification Evaluations," *Int. J. Data Min. Knowl. Manag. Process*, 2015, doi: 10.5121/ijdkp.2015.5201.

[2] Y. Benny, T. Galanti, S. Benaim, and L. Wolf, "Evaluation Metrics for Conditional Image Generation," *Int. J. Comput. Vis.*, 2021, doi: 10.1007/s11263-020-01424-w.

[3] M. R. M. Habibi, F. Mohammadabadi, H. Tabesh, H. Vakili-Arki, A. Abu-Hanna, and S. Eslami, "Effect of an Online Appointment Scheduling System on Evaluation Metrics of Outpatient Scheduling System: a before-after MulticenterStudy," *J. Med. Syst.*, 2019, doi: 10.1007/s10916-019-1383-5.

[4] J. Luiten, A. Osep, P. Dendorfer, P. Torr, A. Geiger, L. Leal-Taixe, and B. Leibe, "HOTA: A Higher Order Metric for Evaluating Multi-Object Tracking," *Int. J. Comput. Vis.*, Sep. 2020, doi: 10.1007/s11263-020-01375-2.

[5] L. Chen, P. Xu, and H. Wang, "Photocatalytic membrane reactors for produced water treatment and reuse: Fundamentals, affecting factors, rational design, and evaluation metrics," *Journal of Hazardous Materials*. 2022. doi: 10.1016/j.jhazmat.2021.127493.

[6] D. Deutsch, R. Dror, and D. Roth, "A statistical analysis of summarization evaluation metrics using resampling methods," *Trans. Assoc. Comput. Linguist.*, 2021, doi: 10.1162/tacl_a_00417.

[7] R. Takami, H. Shibata, and Y. Takama, "A Visual Analytics Interface for Formulating Evaluation Metrics of Multi-Dimensional Time-Series Data," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3098621.

[8] R. F. Wunderlich, Y. P. Lin, J. Anthony, and J. R. Petway, "Two alternative evaluation metrics to replace the true skill statistic in the assessment of species distribution models," *Nat. Conserv.*, 2019, doi: 10.3897/natureconservation.35.33918.

CHAPTER 7

CLASSIFICATION OF MAJOR APPLICATIONS

Ms. Natya S.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id-natyas@presidencyuniversity.in

This subject contains the following applications, each of whom is described in depth below:

- Remote Sensing,
- Medical Diagnosis,
- Surveillance,
- Photography Applications

i. REMOTE SENSING

Two versions of image fusion in applications for remote sensing are shown in remote sensing techniques. The first example is the synthesis of multi-spectral (MS) and panchromatic (PAN) photographs [1].

a. Application:

To develop the imagery used within the well-known Google Maps/Earth products, it had been effectively deployed. The fusion of multi-spectral and hyper-spectral photographs comes in second. In more modern years, spectral immixing and land-cover classifications have shown to be tremendously aided by hyperspectral (HS) imaging, which photographs a landscape containing hundreds of continuous wavelength bands [2].

b. Disadvantages:

Due to the complexity of something like the sensors and economic constraints, they typically have lower positioning accuracy than MS images. Compared to pan-sharpening, this is a more challenging issue because the surround sound audio MS image contains both spatial information, making the existing pan-sharp-cutting methods completely irrelevant or ineffective for the combining of HS and MS images.

The following are the most complicated questions in the area of images from remote sensing fusion:

- Spectral and Spatial Distortions,
- Mis-registration

The source pictures utilized in remotely sensed data are often compiled from various spectral bands, captures, or timeframes. The two instruments do not perfectly point in the same direction and their reception times are not similar, even for the PAN and MS data sets acquired by the same platform [3].

To model flaws, observation noise, environmental elements, end member variability, and original dataset size, the inverse issue of stationery is difficult and poorly presented. The methodology focuses on the SAR texture, which is derived by reducing the SAR picture to its low-pass approximate. It modifies the MS image's broad sweeping intensity, which is determined by a continuous function that stretches the intensity-hue-saturation transforms to any number of bands. The GI originating from the original MS data is replaced with the chewiness panchromatic-sharpened GI. The technique uses SAR texture that was recovered using a toits-low-pass interpretation of the SAR picture [4].

ii. MEDICAL DIAGNOSIS

Fusion of magnetic resonance imaging (MRI) and computed tomography (PET) pictures and fusion of MRI and computed tomography (CT) images are two examples of face recognition for medical diagnostic purposes. The list of data sources that are often exploited in medical picture fusion techniques. For instance, the Harvard Medical School has made downloadable a brain imaging data collection that incorporates registered MRI, PET, and CT images. Matting combines the correlation between neighbourhood nodes and focuses information, which tends to provide more consistent fusion results. Many related image fusion studies have taken the extensive application of this data collection.

a. Advantage:

The soft tissue characteristics of organs including the brain, heart, and eyes may be captured by MRI in contrast to MRI imaging. High spatial resolution photographs of the human body's bone structures may be obtained by CT imaging. While the photographs that are generated typically have a limited spatial resolution, PET imaging is a significant kind of nuclear medicine imaging. Image fusion technologies may be utilized in association with MRI, CT, and PET scans to improve imaging performance and practical clinical relevance [5].

The following constitute the most difficult issues in this area of medical image fusion:

- The absence of clinical signs fusion techniques.
- Assessment of the effectiveness of fusion.
- Mis-registration.

iii. SURVEILLANCE APPLICATIONS

Uses for image fusion in eavesdropping The picture matting-based procedure is used in this figure to integrate visible and infrared pictures as well as daytime and nighttime images. In contemporary memory, region-based photogrammetry has received a lot of payments charged to its alleged benefits, which would include:

- Instead of using pixels, the merging rules combine areas. As a result, more effective tests for selecting the proper areas from the source photographs based on a region's multiple factors may be applied preceding fusion.
- By processing semantic regions then instead of individual pixels, gradient blending-based procedures can capture real-world spectral response fields while addressing some of the challenges with pixel-fusion methods, such as responsivity to noise, blurring effects, and miss certification. This new method has been demonstrated as symmetric photography. The infrared image is capable of seeing dark without light because it is hypersensitive to items that are brighter than the area surrounding.

b. Disadvantage

The low sensitivity and specificity of IR images is a drawback and the Fusion of IR and visible pictures, or the multistate directional bilateral filter (MDBF), a unique multistage mathematical analysis that adds the non-sub sampling bidirectional filter bank into the systematic sampling bilateral filter is developed, may easily solve this issue. Experimental findings encompassing pictures in the visible and infrared spectrum combined with medical images show that our methodology is superior to traditional methods when it comes to visual examination and measurement systems [6].

Bibliography

[1] D. Gautam and V. Pagay, "A review of current and potential applications of remote sensing to study thewater status of horticultural crops," *Agronomy*. 2020. doi: 10.3390/agronomy10010140.

[2] M. M. Bennett, "Is a pixel worth 1000 words? Critical remote sensing and China's Belt and Road Initiative," *Polit. Geogr.*, 2020, doi: 10.1016/j.polgeo.2019.102127.

[3] F. W. LEBERL, " A review of: ' Microwave Remote Sensing—Active and Passive '. By F. T. Ulaby. R. K. Moore and A. K. Fung. (Reading, Massachusetts: Addison-Wesley, 1981 and 1982.) Volume I: Microwave Remote Sensing Fundamentals and Radiometry. [Pp. 473.] Price U.S. $46.50. Volume II: Radar Remote Sensing and Surface Scattering and Emission Theory. [Pp. 628.] Price U.S. $49.50. ," *Int. J. Remote Sens.*, 1984, doi: 10.1080/01431168408948820.

[4] W. Ha, P. H. Gowda, and T. A. Howell, "A review of downscaling methods for remote sensing-based irrigation management: Part I," *Irrigation Science*. 2013. doi: 10.1007/s00271-012-0331-7.

[5] F. Paul, A. Käär, M. Maisch, T. Kellenberger, and W. Haeberli, "The new remote-sensing-derived Swiss glacier inventory: I. Methods," *Ann. Glaciol.*, 2002, doi: 10.3189/172756402781817941.

[6] J. M. Chen, S. C. Thomas, Y. Yin, V. Maclaren, J. Liu, J. Pan, G. Liu, Q. Tian, Q. Zhu, J. J. Pan, X. Shi, J. Xue, and E. Kang, "Enhancing forest carbon sequestration in China: Toward an integration of scientific and socio-economic perspectives," *J. Environ. Manage.*, 2007, doi: 10.1016/j.jenvman.2006.08.019.

ISBN 978-81-962236-5-6

CHAPTER 8

PHOTOGRAPHY APPLICATION

Mr. Syed Abrar Ahmed
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- syedabrarahmed@presidencyuniversity.in

Examples of feature extraction are used in photography. In this, the fusion of multiple focus and simultaneous exposure photos is conducted using a guided filtering-based manner and a recursive filtering-based manner, respectively. The use of the multi-focus information fusion method is a crucial component of creating a series of photos with every item in focus. The development of a useful activity-level assessment to determine the clarity of source pictures is the main goal of multi-focus image fusion. The unique image fusion strategy for multi-focus pictures with dense feature scale invariant transform is presented in this paper (SIFT). This work's major improvement is that it demonstrates the enormous potential of image local characteristics like the dense SIFT employed in picture fusion [1].

In particular, the local classification technique may be utilized to better the quality of the fused picture by matching the misregistered pixels from different reference images in addition to being applied to quantify the activity level. Using the sliding window approach, our program first creates an initial decision map by estimating the activity level of source image patches using the dense SIFT descriptor. The decision map is then enhanced using feature matching and a local focus way of measuring comparison [2]. Using two examples of image fusion in photography applications the fusion of multi-focus images and indeed the fusion of multi-exposure images experimental findings demonstrate that the proposed method can be successful in the market with or even outmanoeuvre the state-of-the-art fusion methods in terms of both subjective sensory perception and objective evaluation metrics. It is difficult for all components in a photograph to remain in focus at slightly varied distances from the camera because of the small depths of the field of cameras.

Multi-focus image fusion is a process that may be used to combine many photographs of the same scene, each with a different centre point, to generate a composite image with every item in focus. The development of a useful activity-level evaluation to assess the clarity of source photographs is the main goal of multi-focus image fusion. The unique image reconstruction approach for multi-focus pictures with dense feature scale invariant transform is proposed in this work (SIFT) [3]. This work's major improvement is that it demonstrates the immense potential of imaging local characteristics like that of the dense SIFT employed in picture fusion. The proposed scheme may compete with or even surpass state-of-the-art fusion methods, according to experimental findings, in terms including both subjective sensory perception and objective assessment criteria. In many machine imaging and image processing jobs, the fused picture should effectively maintain the pertinent information contained in the original data, making it very desired. Multi-exposure picture fusion, like multi-focus image fusion, tries to create an image that is evenly supplied by integrating many photos which have been captured at various exposures. Applying flash in low light is another photographing method in addition to adjusting the exposure level. Flash photography does,

ISBN 978-81-962236-5-6

however, also provide several undesirable effects, such as red eyes, harsh lighting that is flat, and distractingly sharp shadowing of silhouettes [4].

Flash and non-flash photogrammetry is an intriguing research topic in this area as a potential answer to this problem. Digital photography has made it quick and simple to take two images of low-light situations: one with a flash to capture detail and the other without a flash to capture soundscape illumination. We provide many tools that evaluate as well as combine the benefits of such flash/no-flash picture combinations. The theoretical study of these correlation-based quality metrics when they are applied to evaluate the effectiveness of weighted averaging image fusion technologies is the main topic of this work [5]. This paper's goal is to firmly establish that, in certain scenarios, correlation-based performance indicators exhibit some inappropriate behavior. We use a statistical framework to evaluate the sensor pictures that were seen, and also look at the characteristics of these causal connection quality indicators. Our examination found that these correlation-based performance metrics can behave poorly when the target signal's power or the input pictures' noise are changed, giving the impression of greater quality when worse performance is present. Studies using actual pictures also show the theoretical analyses' value by highlighting it's predicting the future powers. Furthermore, there are various pairings of multi-focus photos and multi-exposure photographs in the well-established Patriotic' image fusion database [6]. The most difficult questions in the area of photography information fusion are finally resolved in the photograph application:

- The impact of shifting attractions on the multiple focus and simultaneous exposure photos are always captured at various periods in photography applications. In this case, moving particles may appear in several positions while being captured. Moving objects may contribute ghost artefacts to the fused view as a result of fusion [7].
- Use for consumer electronics and the world of photography, the imaging process entails obtaining several pictures using multiple camera settings, which takes a lot of time. Consequently, another significant topic is how to incorporate the multi-focus and non-linear and non-image fusion algorithms into consumer gadgets to produce high-quality fused photographs in real-time.

Bibliography

[1] K. D. Wyatt, B. N. Willaert, C. M. Lohse, P. J. Pallagi, J. A. Yiannias, and T. R. Hellmich, "Experiences of Health Care Providers Using a Mobile Medical Photography Application," *Applied Clinical Informatics*. 2020. doi: 10.1055/s-0040-1701254.

[2] H. U. Hua, D. J. Pieramici, and N. Bagheri, "Understanding seizure risk with wide field fundus photography: Implications for screening guidelines in the era of COVID-19 and telemedicine," *Am. J. Ophthalmol. Case Reports*, 2020, doi: 10.1016/j.ajoc.2020.100844.

[3] L. Ray, "Social Theory, Photography and the Visual Aesthetic of Cultural Modernity," *Cult. Sociol.*, 2020, doi: 10.1177/1749975520910589.

[4] R. Dermawi and H. Tolle, "What factors influence users to install a photography-category application to their android smartphone," *Int. J. Interact. Mob. Technol.*, 2019, doi: 10.3991/ijim.v13i02.6726.

[5] I. Mykytka, "Noun Compounds In Photography," *Atlantis*, 2020, doi: 10.28914/ATLANTIS-2020-42.2.04.

[6] D. O. Dumestre and F. Fraulin, "Avoiding Breach of Patient Confidentiality: Trial of a Smartphone Application That Enables Secure Clinical Photography and Communication," *Plast. Surg.*, 2020, doi: 10.1177/2292550319880910.

[7] F. Kokkinos and S. Lefkimmiatis, "Iterative residual CNNS for burst photography applications," 2019. doi: 10.1109/CVPR.2019.00608.

CHAPTER 9

APPROACHES FOR PIXEL-LEVEL IMAGE FUSION

Ms. Natya S.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id-natyas@presidencyuniversity.in

As previously stated, pixel-level image fusion is frequently active in:

i. **Remote Sensing**

By measuring the reflected and transmitted radiation of an area from either a distance, remote sensing is the method for recognizing and keeping track of the physical features of a locality typically from satellite or aircraft. Remotely sensed photos are captured by sophisticated cameras and are used by astronomers to "feel" the Earth. diagnostic imaging Health care image fusion, which is the process of registering and consolidating multiple images from one or much more imaging modalities, aims to increase the enables individual of medical images for the assessment and diagnosis of medical issues by improving electron microscopy quality and reducing random noise and redundancy. Clinical decision-making based on medical photographs is now more accurate because of multi-modal brain tumour segmentation algorithms and tools [1].

ii. **Computer Vision**

Multisensory image fusion in video processing is the technique of combining pertinent data from two or maybe more pictures into one image. The three fundamental processes below may be used to characterize the bulk of picture segmentation methods:

- ***a.*** Picture transform an image may be submitted to an image transform to change its domain. Features that might be less readily seen in the spatial domain may indeed be identified by viewing a photograph in domains like frequency or Hough space.
- ***b.*** Combining the parameters of the transform.
- ***c.*** To produce the alternatives under the required probabilities, inverse transform sampling is done over a designated time. With the help of this sampler, it is possible to control the percentages of various situations, among them the most and least plausible.
- ***d.*** The contour let transform, which in itself is superior for handling two-dimensional data, seems to be another transform in addition to curvature let that may capture the inherent architectural architecture of pictures [2].

Depending on the reactions that occur used, existing background subtraction approaches may be classified into four distinct groups:

- ***a.*** Multi-scale anaerobic fermentation methods,
- ***b.*** Methods using a non-linear model,
- ***c.*** Techniques that perform integration on image pixels or in other transformed contexts directly,

d. Techniques that combine principal component analysis, multi-scale decomposition, local binary patterns, and other transformations.

In addition to the signal understanding of the relevant, the fusion technique also has a significant impact on the outcomes of the fusion. The fusion approach is a technique for figuring out how to generate the fused image from the coefficients or pixels of the source image [3]. Two face photos from distinct spectrums are combined somewhere at the image level using a DWT-based fusion technique. At the document level, a 2D log-polar Gabor wavelet is utilized to extract the amplitude and phase components from the fused picture. The best face image recognition results were obtained using a short-wave IR and visible light combine, with an identical error rate of 2.86%. The suggested image feature fusion good recommendation for other fusion algorithms based on performance [4]. The data sources which are frequently utilized in image fusion applications for espionage are outlined. Several noteworthy infrared and visible picture pairings thoughts come to mind. The Equinox face repository provides a collection including both visible and Long Wave Thermal imaging facial photographs for computer vision applications. In recent years, the investigation of combining images from several sensors has become more well-known throughout the scientific world. The operator functionality models that were created using the contemporaneous protocol's findings were employed to develop the first values obtained for the fusion of different sensor image data [5].

The cognitive benefits of different sensors picture fusion, however, have so far not gotten as much emphasis. A concurrent protocol approach had been utilized in this research to examine how people combine data from visible and infrared imaging throughout low- and high-stress conditions. The operator functionality models that were created using the contemporaneous protocol's findings were utilized to create the very first design points for the fusion of different sensor image data. Then, in combat target identification simulations, fused image data were applied, and operator proficiency, accuracy, and speed were examined with the outcomes produced from injected data. The findings indicate that the model accurately captures how individuals interpret information from various senses in this specific situation, and the program design considerations are prospective for supporting fighter pilots in even more rapid and precise object recognition [6].

Effective image nuclear fission algorithms should merge the information gathered from the original photos properly and provide a clear fused image for remote monitoring. More crucially, steady monthly monitoring is a major element of surveillance systems. Consequently, increasing the speed of information fusion in remote sensing applications is an intriguing approach. Environmental stress circumstances. The possibility of erroneous picture acquisition is a serious problem in the realm of monitoring image fusion .

Bibliography

[1] Z. Shao and J. Cai, "Remote Sensing Image Fusion with Deep Convolutional Neural Network," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, 2018, doi: 10.1109/JSTARS.2018.2805923.

[2] C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, and L. He, "Local Feature Descriptor for Image Matching: A Survey," *IEEE Access*, 2019, doi: 10.1109/ACCESS.2018.2888856.

[3] M. A. Hasan, N. A. Abdullah, M. M. Rahman, M. Y. I. Bin Idris, and O. F. Tawfiq, "Dental Impression Tray Selection from Maxillary Arch Images Using Multi-Feature Fusion and Ensemble Classifier," *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3059785.

[4] M. Kumar and S. Dass, "A total variation-based algorithm for pixel-level image fusion," *IEEE Trans. Image Process.*, 2009, doi: 10.1109/TIP.2009.2025006.

[5] K. Elaiyaraja and M. Senthil Kumar, "Fusion imaging in Pixel Level image processing technique - A literature review," *Int. J. Eng. Technol.*, 2018, doi: 10.14419/ijet.v7i3.12.15913.

[6] X. Jia, C. Zhu, M. Li, W. Tang, and W. Zhou, "LLVIP: A Visible-infrared Paired Dataset for Low-light Vision," 2021. doi: 10.1109/ICCVW54120.2021.00389.

CHAPTER 10

FUSION TECHNIQUE IN SIGNAL TRANSFORM SCHEME

Dr. Bhanurekha K.
Associate Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- bhanurekha@presidencyuniversity.in

The fusion technique, in addition to the signal transform scheme, is another important component that influences the fusion results. The fusion strategy is the method for determining the generation of the fused picture from the source image's coefficients or pixels [1].

i. Multi-scale Decomposition Methods:

Some of the most commonly heard multi-scale decomposition methodologies for image fusion in earlier research are the Laplacian pyramid, wavelet packet decomposition (DWT), and constant discrete wavelet transform. In a lecture on wavelet-based photo clustering algorithms, they provide a thorough comparison of the several pyramid merging strategies, different appearance levels, and various wavelet families. The bulk of pyramid-based and typical wavelet-based image fusion approaches are based on multi-scale segmentation, as was first proposed. Alternative multi-scale techniques for feature extraction, such as dual-tree complex basis function, curve let, contour let, among shear let, have been introduced in current history in addition to the wavelet as well as pyramid transforms [2].

The discrete transform of wavelets has well-known flaws such as shift variance, aliasing, and lack of directionality. These issues were successfully addressed by image fusion that used the dual complex wavelet tree complex technique. The dual-tree complex wavelet has two key advantages over the discrete wavelet transform, shift invariance and unidirectional selectivity, which may assist to eliminate artefacts produced by the discrete wavelet transform. The contour let transform, which is superior for handling two-dimensional communications, is another transmogrify in addition to curve let that may encapsulate the inherent geometrical architecture of pictures. In the contourlet transform, a moment in time discontinuity is first acquired by the Laplacian pyramid, followed by being connected to linear structures by a unidirectional filter bank [3].

Due to the contourlet transform's capability to accurately depict spatial structures, it has been effectively used throughout remote sensing, reconnaissance, and medical imaging applications. However, contourlet lacks a shift-invariant characteristic that's because the transform process also uses down-sampling techniques. The non-subsampled form of the contourlet offers a prospective remedy, however, it will necessitate more time. Also because the directional filter bank of the contourlet is permanent, it is unable to accurately depict complex spatial structures with several orientations. Recently, image fusion has been accomplished with the help of the shearlet. There are no boundaries on the number of shearing positions or support size, making the shearlet more computationally economical than the contourlet [4].

Recent years have witnessed the development of multi-scale representations of images using edge-preserving filtering for the integration of several-exposure images. For multi-scale decomposition-based fusion, other potent computer vision technologies such as anisotropic temperature diffusion, log-Gabor transform, and critical to organizational success incorporation have also been effectively used. The capacity to properly discern fine-scale variables incorporating, middle-scale margins and large-scale spatial structures of a photograph are generally considered to be some very approaches' main benefit. By decreasing halo and aliasing artefacts even during the fusion process, this characteristic makes it possible to produce nuclear fission outcomes that are suitable for human visual processing. Finally, the number of convolution layers has a significant impact on the appearance of the fused pictures in addition to that same choice of the multi-scale disintegration approach. The fused pictures' spatial resolution could suffer if fewer deconstruction stages are implemented [5].

ii. Strategies Used for Fusion of Multi-scale Representations:

The use of efficient segmentation techniques is another area in multi-scale decomposition-based merging that might be researched to optimize fusion quality. A neighbourhood geometrical processing step is utilized following the choose-max approach at each transformed scale, which may increase the uniformity of coefficient selection. In summation, the bulk of research conducted in this field is on creating new approaches for multi-scale decomposition, a problem that is still in demand. Fusion rules didn't garner nearly as much attention throughout the beginning as decomposition process development did. Recent research has shown that many of the shortcomings of multi-scale transforms may very well be addressed by sophisticated fusion rules, leading to state-of-the-art fusion performances. But even though the choose-max and averaging procedures, which are often employed in conventional fusion rules, are straightforward in just about all cases, they may cause eye abnormalities when the pictures are also not correctly registered which includes noise. These methods thoroughly use the dependencies here between coefficients of multiple scales and the significant connection amongst neighbouring pixels [6].

iii. Sparse Representation-based Methods

By restricting the dimension reduction mechanism of human vision systems, the revolutionary respect to the applied theory known as background subtraction, which deals with fragmentary solutions for a linear system of equations, has successfully been applied to several image processing problems like interpolation and recognition. The pictures or image patches in the background subtraction may be described by a sparse linear function of a set of atoms drawn from the over-complete dictionary. Only a few non-zero variables in the sparse coefficients may be able to accurately represent the saliency information contained in the original pictures, according to the incompleteness of the weighted coefficients generated. The sparse coefficients' peculiarities were initially used to apply classical pattern-matching theory to image fusion. To safeguard shift invariance and capture local descriptive format, the input pictures from numerous sources are first divided into spanning multiple patches. Decomposing patchwork from distinct pictures with the same over-complete dictionary yields the equivalent sparse coefficients. Then, an increasing network is used to the parameters from numerous sources [7].

Bibliography

[1] R. K. Tripathy, M. R. A. Paternina, J. G. Arrieta, A. Zamora-Méndez, and G. R. Naik, "Automated detection of congestive heart failure from electrocardiogram signal using Stockwell transform and hybrid classification scheme," *Comput. Methods Programs Biomed.*, 2019, doi: 10.1016/j.cmpb.2019.03.008.

[2] A. M. Cheema, S. M. Adnan, and Z. Mehmood, "A novel optimized semi-blind scheme for color image watermarking," *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3024181.

[3] J. Liu, P. Wang, H. Chen, and J. Zhu, "A combination forecasting model based on hybrid interval multi-scale decomposition: Application to interval-valued carbon price forecasting," *Expert Syst. Appl.*, 2022, doi: 10.1016/j.eswa.2021.116267.

[4] P. Song, Y. Tan, X. Geng, and T. Zhao, "Noise Reduction on Received Signals in Wireless Ultraviolet Communications Using Wavelet Transform," *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.3009944.

[5] M. YANG, W.-W. LI, and S.-Z. ZHU, "Comparison and Application of Four Multi-scale Decomposition Methods for the Price of Different Carbon Markets," *DEStech Trans. Environ. Energy Earth Sci.*, 2020, doi: 10.12783/dteees/peems2019/33942.

[6] A. Mobasher Amini, D. Dureisseix, and P. Cartraud, "Multi-scale domain decomposition method for large-scale structural analysis with a zooming technique: Application to plate assembly," *Int. J. Numer. Methods Eng.*, 2009, doi: 10.1002/nme.2565.

[7] G. Chen, L. Li, W. Jin, S. Qiu, and H. Guo, "Weighted sparse representation and gradient domain guided filter pyramid image fusion based on low-light-level dual-channel camera," *IEEE Photonics J.*, 2019, doi: 10.1109/JPHOT.2019.2935134.

CHAPTER 11

INTRODUCTION TO IMAGE DEFINITION

Ms. Akshatha K.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- akshatha.k@presidencyuniversity.in

The amplitude of F at each given pair of coordinates (x, y) is referred to as the intensity of that picture at that location. An image is characterized as a two-dimensional function, F (x, y), where x and y are spatial coordinates. We refer to F as a digital picture when it's x, y, and amplitude values are all finite. Based on the number of pixels, an image's dimensions (height and breadth) serve as a representation [1]. For instance, if a picture is 500 × 400 (width x height), then 200000 pixels make up the whole image. This pixel is a point higher in the picture that assumes a certain hue, level of transparency, or colour and typically, it appears as one of the following:

- **Grayscale:**

Pixels are integers with values ranging from 0 to 255. (0 is completely black whereas 255 is completely white).

- **RGB:**

Three numbers, ranging from 0 to 255, make up a pixel. (These numbers show how intense the colours red, green, and blue are). Each pixel of a picture is subject to specified procedures during image processing. The image processor implements the first set of operations on the information pixel by pixel. After that is accomplished, it will go on to the next operation, and so forth. Any pixel in the picture might have its output value estimated using these techniques [2].

Image Processing

The process of transforming a physical picture into a digital representation and then performing operations on it to extract pertinent information is referred to as image processing. The image processing system typically treats all pictures as 2D signals when implementing certain particular signal processing techniques. Image processing may be divided into five categories:

- Visualization: Finding items that are hidden in the picture requires skill.
- Recognize: Locate or identify things in the picture.
- Sharpening and restoration: From the original picture, produce an improved image.
- Pattern recognition: Count the different patterns that surround the image's objects.
- Retrieval: Look through and perform searches on photographs from a vast library of digital pictures that resemble the original picture.

Fundamental Image Processing Steps

The fundamental steps in image processing are displayed in Figure 1.

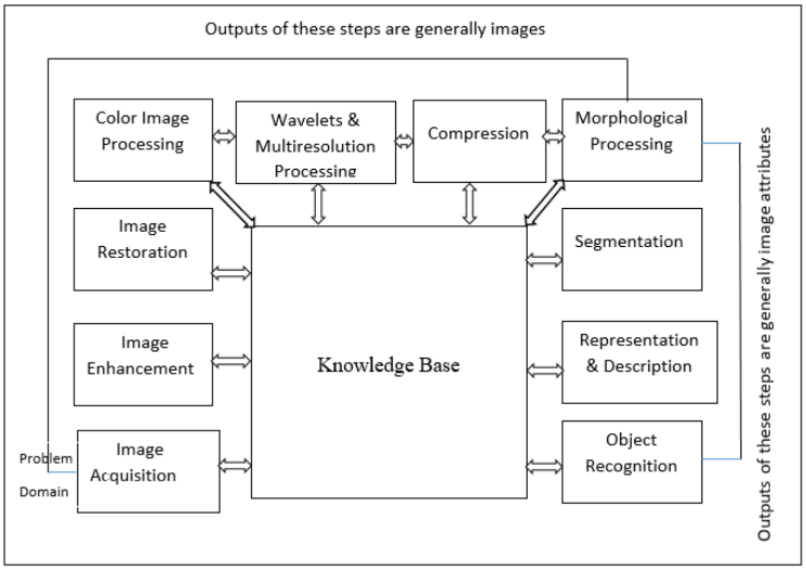

Figure 1: Represented the Fundamental Steps in Image Processing.

i. **Image Acquisition:**

The first stage of image processing is picture acquisition and in image processing, this process is frequently referred to as pretreatment. The information must be retrieved from a source, a hardware-based one [3].

ii. **Image Enhancement:**

The proposed method is the technique of bringing out too and underlining certain interesting characteristics in an obscured image. Changing the brightness, contrast, etc., may do this.

iii. **Image Restoration:**

Image restoration refers to the method of enhancing an image's look. Picture restoration, as compared to image augmentation, is implemented by applying specific mathematical or probability theory models.

iv. **Colour Image Processing:**

A variety of digital color forecasting models are used in color picture processing. The widespread consumption of digital photos on the internet had also given rise to the predominance of this stage.

v. **Wavelets and Multiresolution Processing:**

Images of varying dimensions are presented by wavelets. For the reasons of pyramidal representation and data compression, the pictures are fragmented into wavelets as well as smaller sections [4].

vi. **Compression:**

Compression is a method used to reduce the amount of space needed to record or transmit a picture. This is done exceptionally when the photograph will be published online.

vii. **Morphological Processing:**

An assortment of processing technology known as "procedures and methodologies" are used to change the appearance of photographs.

ISBN 978-81-962236-5-6

viii. **Segmentation:**

One of the most significant challenges of image recognition is segmentation. It entails dividing a photograph into its items as well as components.

ix. **Representation and Description:**

The segmentation process separates a picture into areas, and each region is presented and documented in a method that is appropriate for some further computer processing. The qualities and geographical aspects of the picture are represented via representation. It is the job of descriptive to extract measurable data that helps distinguish one class of items from that other [5].

x. **Recognition:**

A label is given to objects using recognition based on their characteristics.

Bibliography

[1] A. W. Moskovic, "On the definition of image as a representation of reality," *Doxa Comun.*, 2021, doi: 10.31921/doxacom.n33a924.

[2] H. Zhou, J. Peng, C. Liao, and J. Li, "Application of deep learning model based on image definition in real-time digital image fusion," 2020. doi: 10.1007/s11554-020-00956-1.

[3] G. Cui, K. Zhang, L. Mao, Z. Xu, and H. Feng, "Micro-image definition evaluation using multi-scale decomposition and gradient absolute value," *Guangdian Gongcheng/Opto-Electronic Eng.*, 2019, doi: 10.12086/oee.2019.180531.

[4] K. Nie, W. Liu, and J. Wang, "Image definition evaluation method based on image moments function," *Chinese J. Sensors Actuators*, 2013, doi: 10.3969/j.issn.1004-1699.2013.10.016.

[5] E. K. Moru and M. Qhobela, "Social science students' concept images and concept definitions of anti-derivatives," *Pythagoras*, 2019, doi: 10.4102/PYTHAGORAS.V40I1.484.

CHAPTER 12

BENEFITS OF IMAGE PROCESSING

Dr. Azra Jeelani
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- azra.jeelani@presidencyuniversity.in

The use of image-processing technologies has had a serious influence on many IT organisations. The following are only just some of the most important advantages of image processing, independently of the use:

- The digital picture may be sent in whatsoever specified format (improved image, X-Ray, photo negative, etc.).
- It helps with the transformation of images for comprehending scientific concepts.
- Information from processed images may be utilized in automated interpretations.
- The pixel density and hardness of the picture may be tweaked to any ideal consistency.
- Images may be easily achieved and stored [1].

APPLICATIONS OF IMAGE PROCESSING

i. Medical Image Retrieval:

Image processing has been increasingly used in medical research, in addition to having greatly improved the efficiency and accuracy of treatment plans. For example, but use a high-tech nodule produced through brain scans, as illustrated in Figure 1, may be used for the early recognition of brain tumours (MRI, CT). These systems must be subjected to extensive research and testing without first being approved for usage because pharmacological treatments need professionally educated picture interpreters.

Figure 1: Represented the Medical Image Retrieval Processing [2]**.**

ii. Traffic Sensing Technologies:

Specifically, a traffic sensor video image processing system (VIPS) consists of a method for communication, a system for synthesizing pictures, and a system for exchanging photos. A VIPS has several screening zones that produce an "on" signal when a vehicle enters this same zone and

ISBN 978-81-962236-5-6

an "off" signal when it leaves while simultaneously capturing video. These recognition zones in Figure 2 may be tailored for a variety of directions and are capable of identifying traffic in the right place. It can also distinguish the kind of vehicle, photograph the registration number automatically, monitor the driver's motions on the interstate, and do a lot more.

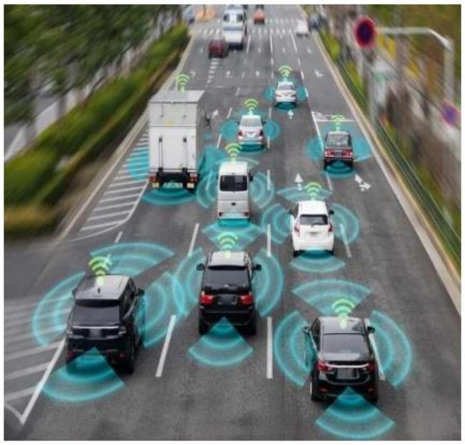

Figure 2: Represented the Traffic Sensing Technology [3]**.**

iii. Image Reconstruction:

Image processing could be employed to reconstruct and repair missing or destroyed parts of a photograph. Image processing algorithms that have been properly trained to employ existing photo datasets are utilized to synthesize newer copies of old and damaged images [4].

iv. Face Detection:

Face recognition has become one of the most frequently used image processing techniques in use today also shown in Figure 3.

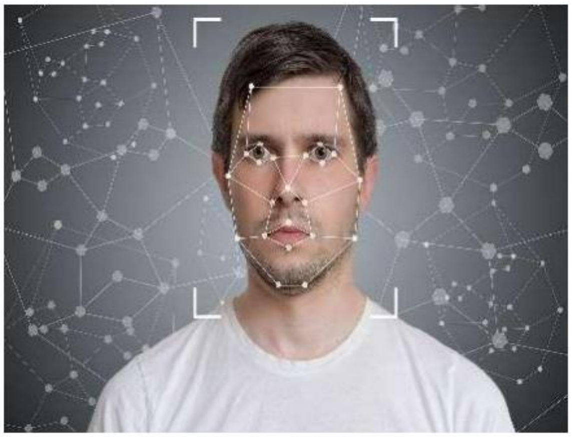

Figure 3: Illustrated the Face Detection.

The computer is trained and tested with the distinctive traits that relatively autonomous faces, such as the curvature of the face, the distance seen between eyes, etc., using techniques called deep learning. The computer should begin to identify any things in an image that approximate a human face shortly after it has learned these properties of a human face. Face recognition is a powerful mechanism used in security, fingerprinting, and even filters featured on the vast majority of social media websites nowadays [5].

Bibliography

[1] S. Saxena, S. Sharma, and N. Sharma, "Parallel Image Processing Techniques, Benefits and Limitations," Res. J. Appl. Sci. Eng. Technol., 2016, doi: 10.19026/rjaset.12.2324.

[2] P. Shamna, V. K. Govindan, and K. A. Abdul Nazeer, "Content based medical image retrieval using topic and location model," J. Biomed. Inform., 2019, doi: 10.1016/j.jbi.2019.103112.

[3] V. Lücken, N. Voss, J. Schreier, T. Baag, M. Gehring, M. Raschen, C. Lanius, R. Leupers, and G. Ascheid, "Density-Based Statistical Clustering: Enabling Sidefire Ultrasonic Traffic Sensing in Smart Cities," J. Adv. Transp., 2018, doi: 10.1155/2018/9317291.

[4] C. Arndt, F. Güttler, A. Heinrich, F. Bürckenmeyer, I. Diamantis, and U. Teichgräber, "Deep Learning CT Image Reconstruction in Clinical Practice," RoFo Fortschritte auf dem Gebiet der Rontgenstrahlen und der Bildgebenden Verfahren. 2021. doi: 10.1055/a-1248-2556.

[5] A. Kumar, A. Kaur, and M. Kumar, "Face detection techniques: a review," Artif. Intell. Rev., 2019, doi: 10.1007/s10462-018-9650-2.

CHAPTER 13

IMAGE FUSION

Dr. Muthupandi G.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- muthupandi@presidencyuniversity.in

Image fusion is a process that combines several input photos into something like a single output image that adequately captures the scene compared to any of the individual source image could. Image fusion is implemented to improve image quality and reduce randomization and redundancy because then diagnosing medical issues is made simpler. Multi-sensor image compression in computer form is the process of fusing essential data from two or more photographs into one image. Any of the input photographs will be less instructional than the final image. The fundamental objective of image fusion is to create images that can be more helpful for extra vision processing and contain additional data necessary for human visual processing [1].

Need for image fusion:

Different sensors data fusion is just becoming a regulation, with many application scenarios demanding highly generalized formal findings. Only high spatial and high spectral information can be extracted from a single picture when processing it under varying environments. It is essential in the field of remote sensing. However, either by design or as a condition of observational limitations, the sensors are also not suitable for providing such types of data. Data fusion is the only remedy, consequently [2].

Types of image fusion:

Low level, expected ability, and high level, or pixel fragment, feature, and determination ranges, are the three ways this same image fusion method may be used.

- Fusing just particular pixel values somewhere at the pixel level.
- Feature level fusion of the incoming image's segmented segments taking into account their peculiarities.

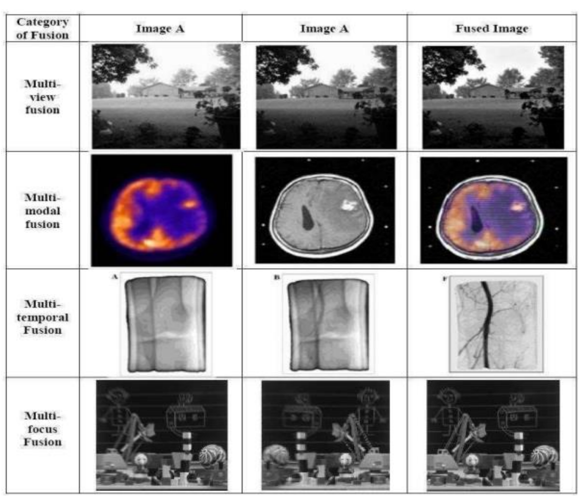

Figure 1: Represented the Different Image Fusion Methods [3].

ISBN 978-81-962236-5-6

Segmentation areas of something like the input pictures are fused somewhere at the decision level while taking initial object detection and classification into account. Researchers have recently shown that combining objects or locations rather than image pieces is far more important. The region-based approach offers several benefits over the pixel-based methodology, including the fact that it is lower sensitive to perturbation, has brighter discrimination, and is less affected by mix registrants, but at the price of increased complexity. In Figure 1, several picture fusion algorithms are shown with demonstrations.

Categories of Image Fusion

According to the nature of the images to be fused, fusion can be categorized as follows:

i. **Multi-Temporal Fusion:**

Although the photos to be amalgamated are of the same modalities, they were captured at particular phases. The major goal of the fission reaction in this instance is to examine differences in the scene at various intervals. It is performed by comes two or more photos.

ii. **Multi-Focus Fusion:**

The portions of the photographs that need to be fused are removed, and the fusion is then leveraged to create the merged photograph. Multi-modal fusion: In this category, numerous modalities are the source photographs, and the major objective is to create segmentation results that retain the greatest amount of information from the different modalities as possible without destroying the image's overall meaning.

iii. **Multi-Modality Fusion:**

To increase the practical application of medical pictures, the procedure of medical image fusion means integrating several images from various types of imaging into a single, fused image containing a large amount of data.

iv. **Multi-View Fusion:**

The primary objective of the fusion process throughout this category is to have all the extra information under the different situations in the fused picture. The images to be mashed are of the same sense modality and transmitted at the same instant in time, although under varying conditions [4].

Advantages of Image Fusion

- It is the quickest to understand.
- The fused picture has up to a better.
- It works best for recognition and identification.
- It is inexpensive.
- Its great resolution is suitable for photographs of various scales.
- Information in fused pictures is strengthened by image fusion.
- Image fusion preserves its capacity to decipher signals throughout all domains.
- Because image fusion offers very many benefits in terms of sharpness, it should improve the images overall [5].

Bibliography

[1] J. Ma, Y. Ma, and C. Li, “Infrared and visible image fusion methods and applications: A survey,” Inf. Fusion, 2019, doi: 10.1016/j.inffus.2018.02.004.

[2] N. Ogrinc, A. Kruszewski, P. Chaillou, P. Saudemont, C. Lagadec, M. Salzet, C. Duriez, and I. Fournier, “Robot-Assisted SpiderMass forIn VivoReal-Time Topography Mass Spectrometry Imaging,” Anal. Chem., 2021, doi: 10.1021/acs.analchem.1c01692.

[3] H. Zhang, H. Xu, X. Tian, J. Jiang, and J. Ma, “Image fusion meets deep learning: A survey and perspective,” Information Fusion. 2021. doi: 10.1016/j.inffus.2021.06.008.

[4] P. Shamsolmoali, M. Zareapoor, E. Granger, H. Zhou, R. Wang, M. E. Celebi, and J. Yang, “Image synthesis with adversarial networks: A comprehensive survey and case studies,” Information Fusion. 2021. doi: 10.1016/j.inffus.2021.02.014.

[5] D. Li, Z. Song, C. Quan, X. Xu, and C. Liu, “Recent advances in image fusion technology in agriculture,” Computers and Electronics in Agriculture. 2021. doi: 10.1016/j.compag.2021.106491.

CHAPTER 14

APPLICATIONS OF IMAGE FUSION

Dr. Muthupandi G.
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- muthupandi@presidencyuniversity.in

1. Intelligent Robots:

- Require motion control based on input from the environment through force/torque, force/tactile, vision, and other kinds of sensors, as shown in Figure 1.
- Control for intelligent watching.
- Automatic target tracking and identification.
- Automated intelligence in production lines.

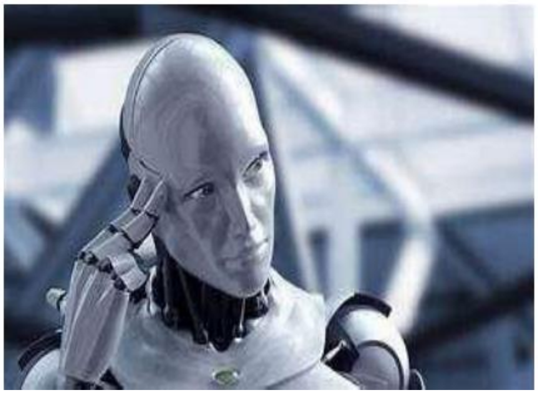

Figure 1: Represented the Intelligent Robot [1]**.**

2. Manufacturing:

- Examining electronic circuits and parts, as illustrated in Figure 2.
- Inspection and measuring of the product's surface.
- Inspection of non-destructive materials.
- Monitoring the manufacturing process.
- Complex device or machine diagnostics.
- Automated intelligence in production lines.

Figure 2: Illustrated the Manufacturing Process [2]**.**

3. **Military and Law Enforcement:**
   - Detection, tracking, and identification of ocean (air, ground) target/event.
   - Concealed weapon detection.
   - Battle-field monitoring.
   - Night pilot guidance [3].
4. **Remote Sensing:**
   - Making use of several electromagnetic spectrum bands.
   - Sensors range from multi-spectral active microwave ovens to aerial black-and-white photos.

Spatiotemporal and non-phenomenological data photograph fusion approaches are the two categories, alternately. The approach for fusing data is constructed by phenomenological algorithms, which use knowledge of the basic natural laws as a base. Various specialists are already pursuing such strategies. However, developing and implementing such techniques must take a lot of time. Contrarily, as shown in Figure 3, non-phenomenological strategies often ignore the physical process in preference of attempting to integrate information by using the data connected to specific information segments. Intending to effectively use the data gathered by the many sensors as well as from the same sensor in different assessment settings, it is important in this context to construct an efficient data fusion system that is capable of leveraging such multi-sensor properties [4].

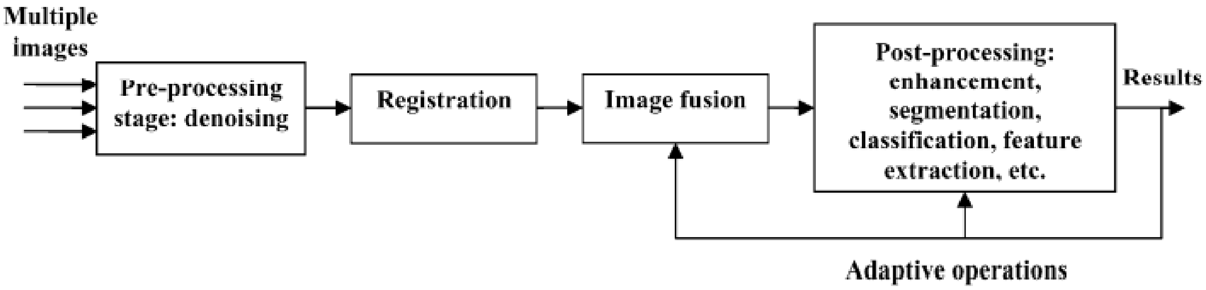

Figure 3: Represented the Block Schema of a General Image Fusion Procedure [5]**.**

Data fusion aims to combine bits of information from several sensors to provide a more thorough picture of the system being evaluated. Technological breakthroughs in this field, particularly concerning computed tomography, have made it possible to synthesize data through electromagnetic waves via the fusing of non-linear and non-pictures, or segmentation techniques.

Bibliography

[1] F. Ren and Y. Bao, "A review on human-computer interaction and intelligent robots," *International Journal of Information Technology and Decision Making*. 2020. doi: 10.1142/S0219622019300052.

[2] T. M. Wang, Y. Tao, and H. Liu, "Current Researches and Future Development Trend of Intelligent Robot: A Review," *International Journal of Automation and Computing*. 2018. doi: 10.1007/s11633-018-1115-1.

[3] A. M. Boutté, B. Thangavelu, J. Nemes, C. R. Lavalle, M. Egnoto, W. Carr, and G. H. Kamimori, "Neurotrauma Biomarker Levels and Adverse Symptoms among Military and Law Enforcement Personnel Exposed to Occupational Overpressure without Diagnosed Traumatic Brain Injury," *JAMA Netw. Open*, 2021, doi: 10.1001/jamanetworkopen.2021.6445.

[4] B. Meher, S. Agrawal, R. Panda, and A. Abraham, "A survey on region based image fusion methods," *Inf. Fusion*, 2019, doi: 10.1016/j.inffus.2018.07.010.

[5] M. Guo, J. Li, C. Sheng, J. Xu, and L. Wu, "A review of wetland remote sensing," *Sensors (Switzerland)*. 2017. doi: 10.3390/s17040777.

CHAPTER 15

INTRODUCTION TO DWT, PCA AND NSCT METHODS

Dr. Sumantra Chaudhuri
Assistant Professor, Department of Electronics and Communication Engineering,
Presidency University, Bangalore, India
Email Id- sumantra.c@presidencyuniversity.in

Spatial and frequency domain classifications have been applied to group image fusion technologies.

- **Spatial Domain:**

Each element in a 2D matrix that describes a pixel's intensity may be utilized to characterize a picture. The state of 2D matrices that characterize the intensity distribution of a photograph is referred to as the spatial domain [1].

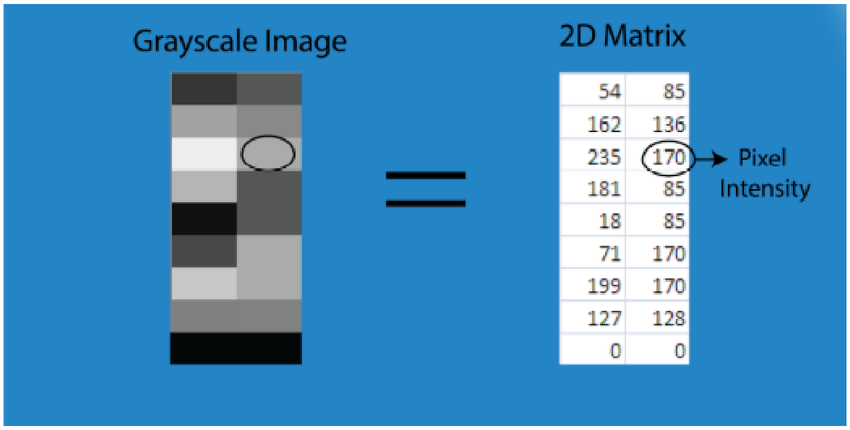

Figure 1: Illustrated the Spatial Domain.

A 3D vector made up of 2D matrices is used to represent the spatial domain of an RGB picture. Each 2D matrix contains the intensities for a certain colour, as seen below.

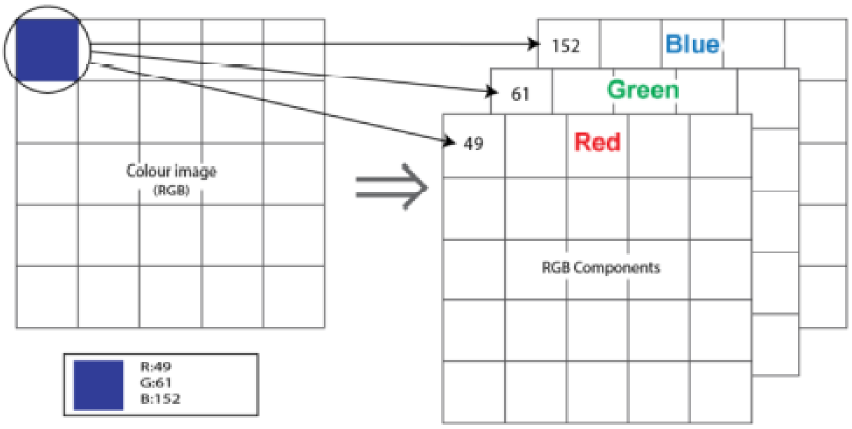

Figure 2: Represented the spatial domain for colour image (RGB).

ISBN 978-81-962236-5-6

- **Frequency Domain:**

Frequency domain authors made use of an image's Fourier Transform. Frequency in a photograph refers to the frequency of change in pixel values [2]. The graphic below shows how information is transformed from the spatial domain into the frequency domain:

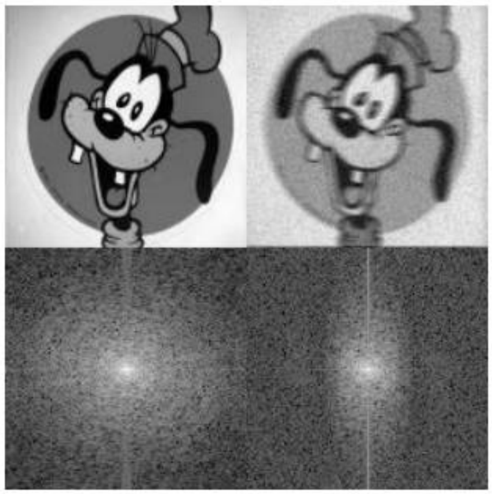

Figure 3: Represented the Conversion of the Image from Spatial Domain to Frequency Domain.

Among all the methods of image fusion, two existing methods are chosen which are Discrete Wavelet Transform (DWT) from the frequency domain and Principal Component Analysis (PCA) from the spatial domain.

- **DWT-PCAv Fusion**

The input source photographs are first fragmented using DWT into various multi-scale dimensions and orientations following the MRR algorithm. With each deconstructed component containing a varied volume of information, this technique is used to analyze input pictures at multiple configurations. The main components are instead computed on each grayscale picture level after multi-scale decomposition. The average of the major components is then assessed at each level of the fragmented picture, and weights are applied towards each coefficient element of the image to determine how to apply unification rules [3].

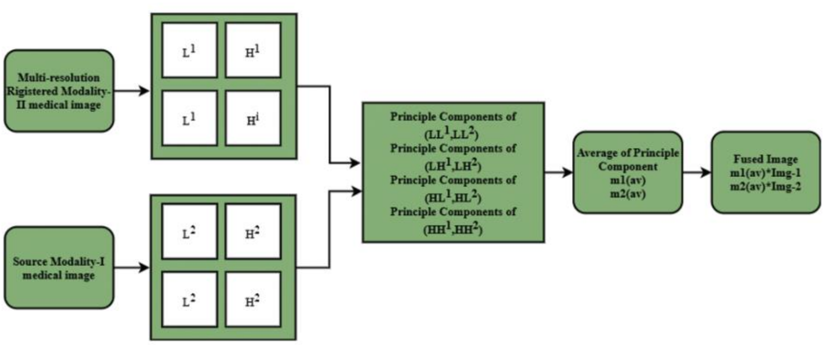

Figure 4: Represented the Block diagram of the DWT-PCAv Fusion Process.

ISBN 978-81-962236-5-6

Different coefficient levels, such as Low-Low (LL), Low-High (LH), High-Low (HL), and High-High, are decomposed from the input pictures (HH). The levels of the detailed coefficient scale are LH, HL, and HH, whereas the element reflecting approximation coefficients is LL [4]. The PCA receives as input the LL coefficient values extracted from the two source photographs. The greatest basic components from the LL coefficients element define the two new coefficient elements m1 and m2. The primary components are produced like the ones used to analyse other detailed parameter elements. Then, to produce the m1 and m2 average components, the important parts of the approximation and elaboration coefficients are averaged. The final picture is synthesized using these two average main components. Figure 4 shows's entire step-by-step fusion procedure in greater detail [5].

Bibliography

[1] X. Zhang, “Benchmarking and comparing multi-exposure image fusion algorithms,” *Inf. Fusion*, 2021, doi: 10.1016/j.inffus.2021.02.005.

[2] Y. Liu, X. Chen, H. Peng, and Z. Wang, “Multi-focus image fusion with a deep convolutional neural network,” *Inf. Fusion*, 2017, doi: 10.1016/j.inffus.2016.12.001.

[3] H. Ghassemian, “A review of remote sensing image fusion methods,” *Information Fusion*. 2016. doi: 10.1016/j.inffus.2016.03.003.

[4] L. Tang, J. Yuan, and J. Ma, “Image fusion in the loop of high-level vision tasks: A semantic-aware real-time infrared and visible image fusion network,” *Inf. Fusion*, 2022, doi: 10.1016/j.inffus.2021.12.004.

[5] A. Song, H. Duan, H. Pei, and L. Ding, “Triple-discriminator generative adversarial network for infrared and visible image fusion,” *Neurocomputing*, 2022, doi: 10.1016/j.neucom.2022.02.025.

PUBLISHER
M/S CIIR RESEARCH PUBLICATIONS
B-17, SECTOR-6, NOIDA,
UTTAR PRADESH, INDIA.
201301
Email: info@ciir.in

March 2023
ISBN 978-81-962236-5-6