

Image processing is
applied to large
numbers cultural
studies essay



**ASSIGN
BUSTER**

ABSTRACT

Image processing is applied to large number of practical applications in a majority of sectors like Engineering, Medical Science, Remote Sensing etc. basically speaking Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or a set of characteristics or parameters related to the image. Most image-processing techniques involve treating the image as a two-dimensional matrix standard processing techniques to it. Image processing usually refers to digital image processing, but optical and analog image processing also are possible. This report is about general techniques that apply to template matching process and manipulating it. As the Image Processing topic is very vast and so is the Template matching Process having a large number of methodologies presently available and also proposed in large number of research papers, this report focuses on the major techniques used in template matching process with some practical implemented in MATLAB. **KEYWORDS:** Template Matching, Image registration, Feature detection, Feature matching, Mapping function, Resampling, Fourier Transform.

Chapter 1

Introduction,

An digital image is defined as two dimensional function $f(x, y)$ where x, y represent spatial coordinates and the amplitude of function f at any pair of coordinates x, y is called the intensity or grey level of the image at that point. Image Processing is used in a wide variety of applications, for two

somewhat different purposes. a. Improving the visual appearance of images to a human observer, including their printing and transmission, and b.

Preparing images for the measurement of the features and structures which they reveal. Images are used in wide variety of applications depending on the type of application and its functionality. A focused researched area is based on the template matching inside the images where we use to identify the location of an object inside any image using the credentials that is to be known by the user that is referred to as template image itself. An digital image is defined as two dimensional function $f(x, y)$ where x, y represent spatial coordinates and the amplitude of function f at any pair of coordinates x, y is called the intensity or grey level of the image at that point. Image

Processing is used in a wide variety of applications, for two somewhat different purposes. a. Improving the visual appearance of images to a human observer, including their printing and transmission, and b. Preparing images for the measurement of the features and structures which they reveal.

Images are used in wide variety of applications depending on the type of application and its functionality. A focused researched area is based on the template matching inside the images where we use to identify the location of an object inside any image using the credentials that is to be known by the user that is referred to as template image itself.

Template matching:

Template matching is a technique in digital image processing for finding small parts of an image which match a template image. It can be used as a way to detect edges in images. Template matching can be subdivided between two approaches: feature-based and template-based matching. The

feature-based approach uses the features of the search and template image, such as edges or corners, as the primary match-measuring metrics to find the best matching location of the template in the source image. The template-based, or global, approach, uses the entire template, with generally a sum-comparing metric (using SAD, SSD, cross-correlation, etc.) that determines the best location by testing all or a sample of the viable test locations within the search image that the template image may match up to.

[2]

1. 2. 1 Feature-based approach

If the template image has strong features, a feature-based approach may be considered; the approach may prove further useful if the match in the search image might be transformed in some fashion. Since this approach does not consider the entirety of the template image, it can be more computationally efficient when working with source images of larger resolution, as the alternative approach, template-based, may require searching potentially large amounts of points in order to determine the best matching location.

1. 2. 2 Template-based approach

For templates without strong features, or for when the bulk of the template image constitutes the matching image, a template-based approach may be effective. As aforementioned, since template-based template matching may potentially require sampling of a large number of points, it is possible to reduce the number of sampling points by reducing the resolution of the search and template images by the same factor and performing the operation on the resultant downsized images (multiresolution, or pyramid,

image processing), providing a search window of data points within the <https://assignbuster.com/image-processing-is-applied-to-large-numbers-cultural-studies-essay/>

search image so that the template does not have to search every viable data point, or a combination of both.

1. 2. 3 Motion tracking and occlusion handling

In instances where the template may not provide a direct match, it may be useful to implement the use of Eigen spaces - templates that detail the matching object under a number of different conditions, such as varying perspectives, illuminations, color contrasts, or acceptable matching object "poses".[6] For example, if the user was looking for a face, the Eigen spaces may consist of images (templates) of faces in different positions to the camera, in different lighting conditions, or with different expressions. It is also possible for the matching image to be obscured, or occluded by an object; in these cases, it is unreasonable to provide a multitude of templates to cover each possible occlusion. For example, the search image may be a playing card, and in some of the search images, the card is obscured by the fingers of someone holding the card, or by another card on top of it, or any object in front of the camera for that matter. In cases where the object is malleable or possible, motion also becomes a problem, and problems involving both motion and occlusion become ambiguous. In these cases, one possible solution is to divide the template image into multiple sub-images and perform matching on each subdivision.

1. 2. 4 Template-based matching and convolution

A basic method of template matching uses a convolution mask (template), tailored to a specific feature of the search image, which we want to detect.

This technique can be easily performed on grey images or edge images. The

convolution output will be highest at places where the image structure
<https://assignbuster.com/image-processing-is-applied-to-large-numbers-cultural-studies-essay/>

matches the mask structure, where large image values get multiplied by large mask values. This method is normally implemented by first picking out a part of the search image to use as a template: We will call the search image $S(x, y)$, where (x, y) represent the coordinates of each pixel in the search image. We will call the template $T(x_t, y_t)$, where (x_t, y_t) represent the coordinates of each pixel in the template. We then simply move the center (or the origin) of the template $T(x_t, y_t)$ over each (x, y) point in the search image and calculate the sum of products between the coefficients in $S(x, y)$ and $T(x_t, y_t)$ over the whole area spanned by the template. As all possible positions of the template with respect to the search image are considered, the position with the highest score is the best position. This method is sometimes referred to as 'Linear Spatial Filtering' and the template is called a filter mask. For example, one way to handle translation problems on images, using template matching is to compare the intensities of the pixels, using the SAD (Sum of absolute differences) measure. The mathematical representation of the idea about looping through the pixels in the search image as we translate the origin of the template at every pixel and take the SAD measure is the following: S_{rows} and S_{cols} denote the rows and the columns of the search image and T_{rows} and T_{cols} denote the rows and the columns of the template image, respectively. In this method the lowest SAD score gives the estimate for the best position of template within the search image. The method is simple to implement and understand, but it is one of the slowest methods.

Background

The majority of the template matching methods consist of the following four steps: Feature detection. Salient and distinctive objects (closed-boundary regions, edges, contours, line intersections, corners, etc.) are manually or, preferably, automatically detected. For further processing, these features can be represented by their point representatives (centers of gravity, line endings, distinctive points), which are called control points (CPs) in the literature. Feature matching. In this step, the correspondence between the features detected in the sensed image and those detected in the reference image is established. Various feature descriptors and similarity measures along with spatial relationships among the features are used for that purpose. Transform model estimation. The type and parameters of the so-called mapping functions, aligning the sensed image with the reference image, are estimated. The parameters of the mapping functions are computed by means of the established feature correspondence. Image resampling and transformation. The sensed image is transformed by means of the mapping functions. Image values in non-integer coordinates are computed by the appropriate interpolation technique. Here in this case we will focus on the feature detection and matching methods. Fig. 1. Four steps of image registration: top row—feature detection (corners were used as the features in this case). Middle row—feature matching by invariant descriptors (the corresponding pairs are marked by numbers). Bottom left—transform model estimation exploiting the established correspondence. Bottom right—image resampling and transformation using appropriate interpolation technique.[8]

Feature Detection

Formerly, the features were objects manually selected by an expert. During an automation of this registration step, two main approaches to feature understanding have been formed. 3. 1 Area-based methods Area-based methods put emphasis rather on the feature matching step than on their detection. No features are detected in these approaches so the first step of image registration is omitted. The methods belonging to this class will be covered in sections corresponding to the other registration steps. 3. 2. Feature-based methods The second approach is based on the extraction of salient structures—features—in the images. Significant regions (forests, lakes, fields), lines (region boundaries, coastlines, roads, rivers) or points (region corners, line intersections, points on curves with high curvature) are understood as features here. They should be distinct, spread all over the image and efficiently detectable in both images. They are expected to be stable in time to stay at fixed positions during the whole experiment. The comparability of feature sets in the sensed and reference images is assured by the invariance and accuracy of the feature detector and by the overlap criterion. In other words, the number of common elements of the detected sets of features should be sufficiently high, regardless of the change of image geometry, radiometric conditions, presence of additive noise, and of changes in the scanned scene. The ‘remarkableness’ of the features is implied by their definition. In contrast to the area-based methods, the feature-based ones do not work directly with image intensity values. The features represent information on higher level. This property makes feature-based methods suitable for situations when illumination changes are expected or

multisensor analysis is demanded. Region features. The region-like features
<https://assignbuster.com/image-processing-is-applied-to-large-numbers-cultural-studies-essay/>

can be the projections of general high contrast closed-boundary regions of an appropriate size, water reservoirs, and lakes, buildings, forests, urban areas or shadows. The general criterion of closed boundary regions is prevalent. The regions are often represented by their centers of gravity, which are invariant with respect to rotation, scaling, and skewing and stable under random noise and gray level variation. Region features are detected by means of segmentation methods. The accuracy of the segmentation can significantly influence the resulting registration. Goshtasby et al. proposed a refinement of the segmentation process to improve the registration quality. The segmentation of the image was done iteratively together with the registration; in every iteration, the rough estimation of the object correspondence was used to tune the segmentation parameters. They claimed the subpixel accuracy of registration could be achieved. Recently, selection of region features invariant with respect to change of scale caught attention. Alhichri and Kamel proposed the idea of virtual circles, using distance transform. Affinely invariant neighborhoods were described in [1], based on Harris corner detector and edges (curved or straight) going through detected corners. Different approach to this problem using Maximally Stable Extremal Regions based on homogeneity of image intensities was presented by Matas et al. Line features. The line features can be the representations of general line segments, object contours, coastal lines, roads or elongated anatomic structures in medical imaging. Line correspondence is usually expressed by pairs of line ends or middle points. Standard edge detection methods, like Canny detector or a detector based on the Laplacian of Gaussian [2], are employed for the line feature detection. The survey of existing edge detection method together with their evaluation can be found (SAR <https://assignbuster.com/image-processing-is-applied-to-large-numbers-cultural-studies-essay/>)

images with speckle noise, which is a typical degradation present in this type of data). They applied elastic contour extraction. Point features. The point features group consists of methods working with line intersections road crossings, centroids of water regions, oil and gas pads , high variance points, local curvature discontinuities detected using the Gabor wavelets, inflection points of curves , local extrema of wavelet transform the most distinctive points with respect to a specified measure of similarity , and corners . The core algorithms of feature detectors in most cases follow the definitions of the ' point' as line intersection, centroid of closed-boundary region or local modulus maxima of the wavelet transform. Corners form specific class of features, because ' to-be-a-corner' property is hard to define mathematically (intuitively, corners are understood as points of high curvature on the region boundaries). Much effort has been spent in developing precise, robust, and fast method for corner detection. A survey of corner detectors can be found in Refs. [9] and the most up-to-date and exhaustive in Ref. [10]. The latter also analyzes localization properties of the detectors. Corners are widely used as CPs mainly because of their invariance to imaging geometry and because they are well perceived by a human observer.

Problem Statement encountered

The template is of the same spectral band as the reference image and of different spectral band (the graphs on the right demonstrate red-blue channel matching). In a general case the normalized cross correlation could fail in case of multimodal data here in this project we are dealing with a part of dissertation so we are considering only single band image and performing Pattern matching in it.

Chapter 2

Technique description

5. 1 Feature matching The detected features in the reference and sensed images can be matched by means of the image intensity values in their close neighborhoods, the feature spatial distribution, or the feature symbolic description. Some methods, while looking for the feature correspondence, simultaneously estimate the parameters of mapping functions and thus merge the second and third registration steps. In the following paragraphs, the two major categories (area-based and feature-based methods, respectively), are retained and further classified into subcategories according to the basic ideas of the matching methods.

4. 1. Area-based methods Area-based methods, sometimes called correlation-like methods or template matching merge the feature detection step with the matching part. These methods deal with the images without attempting to detect salient objects. Windows of predefined size or even entire images are used for the correspondence estimation during the second registration step. The limitations of the area-based methods originate in their basic idea. Firstly, the rectangular window, which is most often used, suits the registration of images which locally differ only by a translation. If images are deformed by more complex transformations, this type of the window is not able to cover the same parts of the scene in the reference and sensed images (the rectangle can be transformed to some other shape). Several authors proposed to use circular shape of the window for mutually rotated images. However, the comparability of such simple-shaped windows is violated too if more complicated geometric deformations (similarity, perspective

transforms, etc.) are present between images. Another disadvantage of the area-based methods refers to the 'remarkableness' of the window content. There is high probability that a window containing a smooth area without any prominent details will be matched incorrectly with other smooth areas in the reference image due to its non-saliency. The features for registration should be preferably detected in distinctive parts of the image. Windows, whose selection is often not based on their content evaluation, may not have this property. Classical area-based methods like cross-correlation (CC) exploit for matching directly image intensities, without any structural analysis. Consequently, they are sensitive to the intensity changes, introduced for instance by noise, varying illumination, and/or by using different sensor types.

5. 1. 1. Correlation-like methods

The classical representative of the area-based methods is the normalized CC and its modifications. This measure of similarity is computed for window pairs from the sensed and reference images and its maximum is searched. The window pairs for which the maximum is achieved are set as the corresponding ones (see Fig. 2). If the subpixel accuracy of the registration is demanded, the interpolation of the CC measure values needs to be used. Although the CC based registration can exactly align mutually translated images only, it can also be successfully applied when slight rotation and scaling are present. There are generalized versions of CC for geometrically more deformed images. They compute the CC for each assumed geometric transformation of the sensed image window and are able to handle even more complicated geometric deformations than the translation-usually the

similarity transform. Similar to the CC methods is the sequential similarity detection algorithm (SSDA). It uses the sequential search approach and a computationally simpler distance measure than the CC. It accumulates the sum of absolute differences of the image intensity values and applies the threshold criterion—if the accumulated sum exceeds the given threshold, the candidate pair of windows from the reference and sensed images is rejected and the next pair is tested. The method is likely to be less accurate than the CC but it is faster. Sum of squared differences similarity measure was used in for iterative estimation of perspective deformation using piecewise affine estimates for image decomposed to small patches. Recently big interest in the area of multimodal registration has been paid to the correlation ratio based methods. In opposite to classical CC, this similarity measure can handle intensity differences between images due to the usage of different sensors—multimodal images. It supposes that intensity dependence can be represented by some function. Two main drawbacks of the correlation-like methods are the flatness of the similarity measure maxima (due to the self-similarity of the images) and high computational complexity. The maximum can be sharpened by preprocessing or by using the edge or vector correlation. the edge-based correlation, which is computed on the edges extracted from the images rather than on the original images themselves. In this way, the method is less sensitive to intensity differences between the reference and sensed images, too. Extension of this approach, called vector-based correlation, computes the similarity measures using various representations of the window. Despite the limitations mentioned above, the correlation like registration methods are still often in use, particularly thanks

to their easy hardware implementation, which makes them useful for real-time applications.

5. 1. 2. Fourier methods

If an acceleration of the computational speed is needed or if the images were acquired under varying conditions or they are corrupted by frequency-dependent noise, then Fourier methods are preferred rather than the correlation like methods. They exploit the Fourier representation of the images in the frequency domain. The phase correlation method is based on the Fourier Shift Theorem and was originally proposed for the registration of translated images. It computes the cross-power spectrum of the sensed and reference images and looks for the location of the peak in its inverse (see Fig. 2). Fig. 2. Area-based matching methods: registration of small template to the whole image using normalized cross-correlation (middle row) and phase correlation (bottom row). The maxima identify the matching positions. The template is of the same spectral band as the reference image (the graphs on the left depict red red channel matching) and of different spectral band (the graphs on the right demonstrate red-blue channel matching). In a general case the normalized crosscorrelation could fail in case of multimodal data.

Chapter 3

Description of the tool

The MATLAB high-performance language for technical computing integrates computation, visualization, and programming in an easy-to-use environment where problems and solutions are expressed in familiar mathematical

notation. The Image Processing Toolbox is a collection of functions that extend the capability of the MATLAB numeric computing environment. The toolbox supports a wide range of image processing operations, including:

- Geometric operations
- Neighborhood and block operations
- Linear filtering and filter design
- Transforms
- Image analysis and enhancement
- Binary image operations
- Region of interest operations

The Image Processing Toolbox in matlab supports four basic types of images:

- Indexed Images
- Intensity Images
- Binary Images
- RGB Images

Chapter 4 Experiments and Results
Chapter 5 How the problem is being extended for dissertation
Chapter 6 Timeline Chart of work done and work to be completed