Image Registration Techniques

Shivam Pandey (17807671)
Department of Civil Engineering
Indian Institute of Technology Kanpur
April 25, 2021

Abstract: Image registration process is used in aligning two images captured in different conditions or viewpoints. The method is hepful in tracking the changes in the scene, and find its usage in the field of defence, remote sensing, computer vision, robotic perception, medical science, and almost everywhere the images are an output of the system and need to be processed to get information. This work will discuss the usage, and different methods of Image registration. Also the work provides a demo of feature descriptor based image registrations with a simple GUI.

1. Introduction

Image registration is the process of aligning two images over each other such that the common part of the two images can't be distinguished. Figure 1 shows a clear illustration of the image registration concept. The image sources may be different, time of clicking may differ, or even images may be taken from different view angles [1] [2]. In the problem of image registration one of image is taken as the target image and the considered as the source image. We perform some kind of manipulation on the source image in order to get the source image aligned with the target image. The whole essence of the image registration is to get the source image co-ordinates in the reference co-ordinate frame defined by the target image [3].

The image registration is an important and significant part of any image processing pipeline. The process is helpful in getting higher resolution images, fine refinement of images, motion and target tracking, and map generation. The application of image processing also includes the fields of environmental monitoring, image stitching, making high resolution images, target tracking and automatic target recognition, GIS, medical imaging analysis (MRI & CT scan stitching) and many more [4] [5] [6] [7] [4] [8].

2. Image registration techniques Classification

The registration techniques are classified based on many criterion's, including type of computations, image acquisition method, portion of images considered for matching etc.

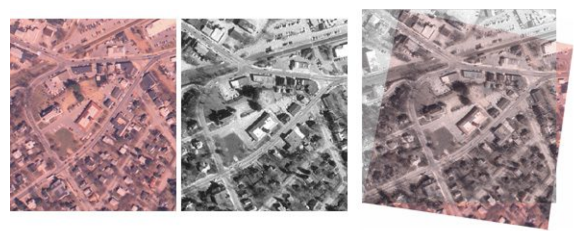

Figure 1: Illustration of image registration

On the basis of Image acquisition methods the image registration techniques are divided into 3 main segments:

- Multi-view registration
- Multi-temporal registration
- Multi-modal registration

Multi-view registration In these type of image registration the images are taken from different view-points for the same environment. The method is helpful in determining grater details of the scene. Image stitching, and 3D-multi-view reconstruction are main examples in this category.

Multi-temporal registration The changes in the environment with the time lapse causes problems in image registration, and thus contains more information, and needs better methods to get the registered images. This type of image registration finds its usage in much interesting task like target tracking and classification etc.

Multi-modal registration Sometimes the information comes from different sources, either from different cameras or from completely different sensors like MRI & CT scan. These problems falls in the category of multi-modal registration process.

Some other criteria of the image registration classification are given here: [9]

1. Sources involved: *The type of sensors involved in the image acquisition task.*
2. Segment undergoing in transformation: *Sometime only a part of an image goes under transformation at a time, and on contrary to that if a complete image is undergoing in transformation.*
3. Dimensionality of transformation: *Dimensions of different registration may be different, depending how the data is there. The transformation may be 2D-2D(RGB-RGB), 2D-3D(RGB-RGBD), 3D-3D(RGBD-RGBD).*
4. Transformation type: *Thr transformation model applied could be rigid, affien, or more flexible Homography transform containing 8-DoF.*
5. Human intervention: *The human intervention may be complete, or partial or may be nil in fully automatic methods.*

3. Image Registration Procedure Details

The image registration can divided into 4 main steps as shown in figure 2, and these are **feature detection**, **feature matching**, **transformation model**, **Image registration with resampling**.

3.1. Feature Detection

The most important and initiation task of the image registration pipeline is the feature point identification, and descriptor generation. As we have already stated this point if divided in two main parts and described here:

3.1.1. Feature point identification

The process of feature point identification involves in finding out locally distinct points that could be identified uniquely. The main target is, they should be globally unique and identifiable but with a world of symmetry this can't be possible thus we take a sub-condition and reduce down our process of interest point detection at a local level. These points have been one of the core area of research for the image processing since the starting. Thus there are lot of conventional ways to get the interest point out of the image. Some of the famous are Harris corner detector, ShiTomasi edge detector, Difference of Gaussian etc. Some of the deep learning based have also been tried recently and proved to be way more effective and ganarilize well than the conventional algorithms. We will talk about DNN based methods later. [10]

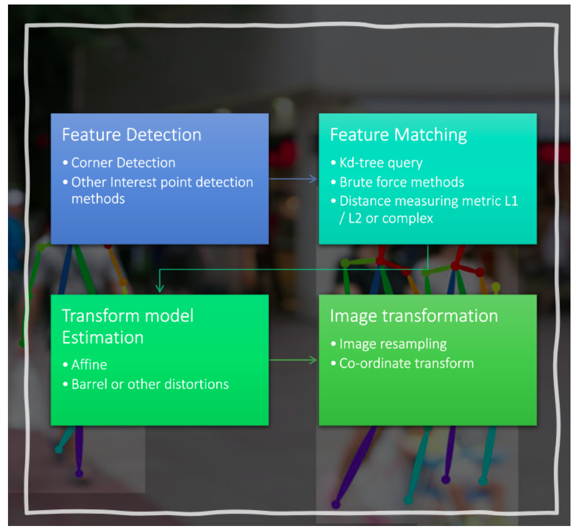

Figure 2: Image registration pipeline

3.1.2. Feature point descriptor

Once we have obtained the interest points we need some kind of identification string in order to distinguish them numerically. The feature descriptors works on the local region of the interest point to get some information about the locality variation. Conventional feature descriptors like SIFT, SURF etc. [10] has used a specified window sized around the feature points. Some of the conventional feature descriptor methods are listed below:

1. SIFT: Scale-Invariant Feature Transform
2. SURF: Speeded-Up Robust Features- This was a faster version of SIFT.
3. BRIEF: Binary Robust Independent Elementary Features- Memory efficient and fast than SIFT and SURF
4. ORB: Oriented FAST and Rotated BRIEF-Faster version of BRIEF

Some method of descriptors have also used the concepts from the signal theory, those were based on wavelets etc.

The descriptor being the crucial part in identifying each of the interest point uniquely without any kind of aliasing. The conventional descriptors are not enough to generalize over multiple scene conditions with enough degrees of in-variances. Here comes the data hungry deep learning algorithms thriving on their experience based embedding/descriptors. We will talk about this later in ANN section.

3.2. Feature Matching

Once we have got the interest points and their respective unique identity vector, we can then match the feature points of both the images on the basis of their feature vector and if some of them commutes to a certain threshold we are then able to call them as a match. Although the used of $L2$ and $L1$ norm for these metric may become irrelevant as the feature descriptors are multidimensional and doesn't follow the euclidean like norms any more. The methods like *Brute force* and *FLANN kd-tree* based search with $L1$or $L2$ norms as a distance b/w correspondences are still highly used due to their simplicity. The modern methods applied the DNN in order to learn the new disparity metric based on which the correspondences can be evaluated. [9]

3.3. Transformation model

The transformation model is the mapping function that will be involved in the final transformation of the images. The transformation model is chosen on the basis of a priori knowledge of the data acquisition. There are lot of options including translation, rotation, scaling, shearing. Also there is more freedom available through holography matrix H upto 8-DoF.

3.4. Image transformation and re-sampling

Once we are done with the model estimation, we can then work on the optimization step to minimize the distance metric defined by correspondences in order to estimate our model parameters. After the estimation of the parameters we can then transform all the points of the source image and then try to re-sample the partially overlapping pixels and filling the values of each pixel.

4. Image Registration Techniques

Based on how we handle all of the 4 steps given previously, the registration techniques of the images gets evolved out with amny of researches. We will list down the description of some of the methods here:

4.1. Extrinsic Registration Technique

The method involved usage of synthetic easily identifiable objects put down with the body. Since these objects are easily identifiable thus they could be used a control point for the registration. The marker attached to the target body, or stereo-tactic frame rigidly attached to the system are the 2 main example usage of these category. [9]

4.2. Moments and Principle Axes Methods

The method is able to register even similar objects with the help of principal axis determination by analysing the physical property of the structure. There is a need to pre-segment things from the complete image in order to get required results.

4.3. Surface based Image registration

This method is majorly applicable in the medical sciences as the surfaces are distinct. multimodal data for the brain specific scans are used to be registered with this method. The input images are taken in a peaceful environment, thus we don't need the homographs in there, so these methods use rigid transformation. [9]

4.4. Correlation based registration

The methods if used for mono-modal and also in case of similar objects, found it's immense usage in medical sciences for analysing he treatment. The crosss corelation methods uses Fourier transforms to get the cross correlation of images. [9]

4.5. Mutual Information Based Methods

The mutual information methods are probabilistic in nature. The method uses the voxelised form of normal distribution fit to each voxel, and tries to maximize the likelihood of overlap b/w the two scans. The cross correlation methods were inefficient for such problems but the mutual information based methods don't suffer from these problems. [9]

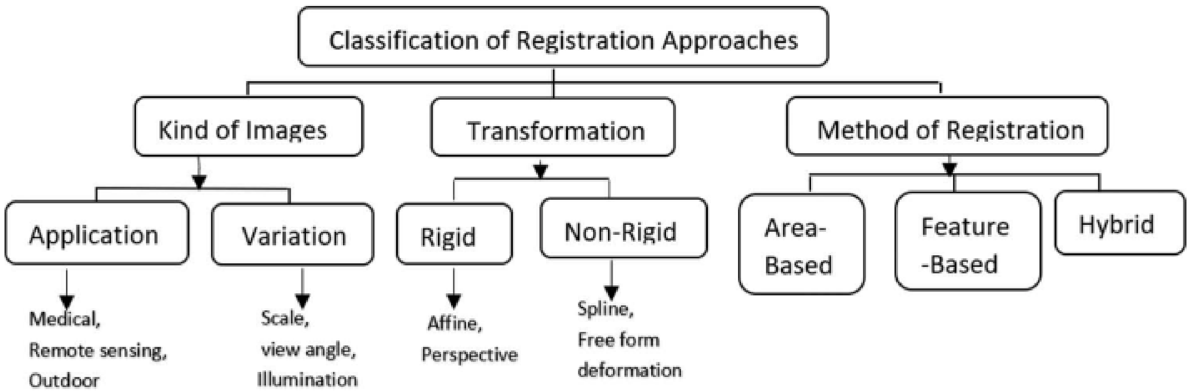

Figure 3: Classification different methods of image registration

4.6. Wavelets based Method

Similar to mutual information the Wavelet Transform was introduced to get an idea of the time instant at which a particular frequency exists. From this method we get both the selectivity in temporal as well as spatial resolution. [9]

4.7. Neural networks based method

With the advent of deep learning based algorithms the model estimation with very large amount of data experience is possible, that to with the high non-linearity in the model. These method os deep neural network can boost the performance of image registration significantly. Here we will discuss the possible problems and how they have been solved with the DNN based algorithms.For more information on deep learning based approaches the reader can refer to [11] [12] [8] [13] [14].

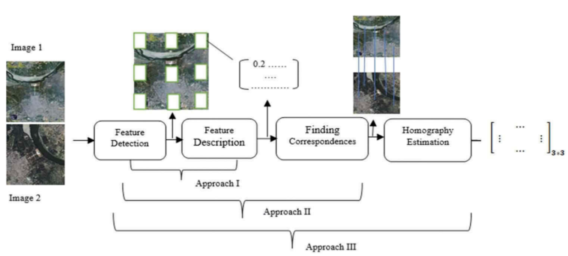

Figure 4: Deep Learning based approaches divided at the level of details we extract from DCNN.

Figure 3 shows the classification of the image registration methods based on their operative characteristics. The main registration method based classification creates 3 pool of algorithms known as Area based method, Feature based methods, and Hybrid method.

Here we won't be considering area based methods as those methods are effective with smooth featureless patches, but in the application of our field i.e robotics, remote sensing, stereo reconstruction, the surface provides us lot of interest points and thus we will be discussing about DNN boost towards feature based methods.

From the figure 5 shows the pipeline of image registration on the basis of images we can work on the each step of this pipeline with the DNN based methods, the whole structure is divided in a hierarchical manner of registration based on deep learning methods, as shown in figure 4 shows the 3 approaches of registration:

1. Approach1: DCNN is used in interest point detection and feature generation.
2. Approach2: DCNN is used in simultaneous estimation of distance metric for correspondence establishing,with the feature descriptor.
3. Approach3: DCNN is directly used in a end-to-end manner to determine direct Homography.

Let's talk about Approach1 in more depth as that is more physically legit.

The feature extraction and interest point determination is the most crucial part of the complete pipeline because without that the possibility of success becomes null. The most reliable type of neural-network is encoder-decoder based model, as they by default don't need any kind of labels that could have messed up the complete training

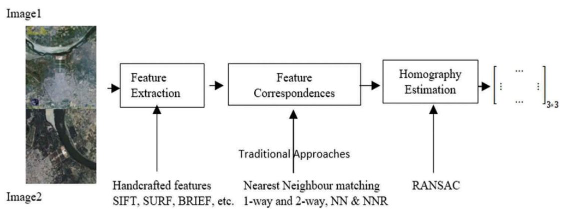

Figure 5: Image registration pipeline for feature based methods.

with some discrepancies. The encoder-decoder model doesn't needs any kind of label, they generate labels by their own. Here are some interest point detectors and descriptors using encoder-decoder model [9].

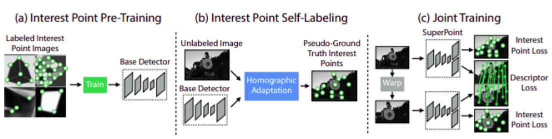

Figure 6: SuperPoint algorithm training process.

SuperPoint: *Self-Supervised Interest Point Detection and Description* The SuperPoint utilizes the encoder-decoder structure with synthetic data augumented with a Homography adaption method. The Homography adaption method is mainly for increasing the reliability of what we are giving to the network. Figure 6 shows the complete training process of the superpoint algorithm [15].

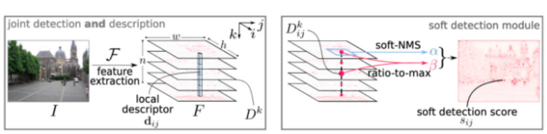

Figure 7: D2-net algorithm schematics for a single CNN.

D2-Net: *A Trainable CNN for Joint Description and Detection of Local Features* The authors suggested a single convolution neural net that does both the tasks, i.e extraction of features, interest points and descriptors as well. Figure 7 shows the schematics of the D2-net being a single DCNN generating both the interest points, and descriptors [16].

LF-Net: *Learning Local Features from Images* The method uses sparse-matching deep architecture and use an end-to-end training approach on image pairs having relative pose and depth maps.

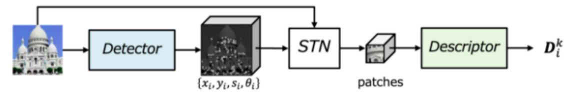

Figure 8: LF-net algorithm schematics for sparse matching based deep CNN for detection and description.

They run their detector on the first image, find the maxima and then optimize the weights so that when run on the second image, produces a clean response map with sharp maxima at the right locations. Figure 8 shows the single sided inference model for the LF-net based feature extraction [17].

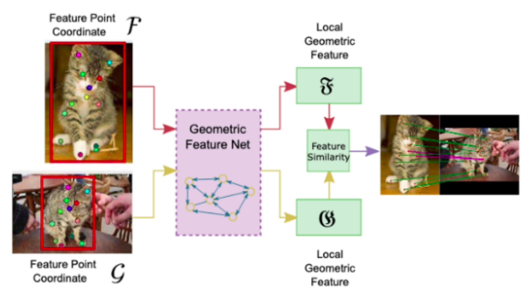

Figure 9: Deep Graphical Feature Learning for the Feature Matching Problem

Deep Graphical Feature Learning for the Feature Matching Problem The method proposed use of graph neural network in the interest point detection and descriptor. The claim by the method is that it is able to solve the traditional np-hard assignment problem into simple assignment problem. The schematic for the method is given in the figure 9, showing the use of graphical neural network in place of CNN. [18].

5. GUI based Demonstration Application

In order to demonstrate the working of feature based image registration methods we have created a GUI based application with python backend. The application implements the conventional methods based on descriptors like SIFT, SURF, etc. For more infromation and detailed Documentation about the application please refer to out Github Repository for ImageRegistration-ToolBox

6. Conclusion

Concluding the work, we have discussed various kind of registration methods varying from old classical one to modern latest one. There is alot more to discuss especially with the Deep Learning approach as they have potential to even search more robust correspondence comparison metric that could be used in the registration proceed via optimization steps. The Deep learning methods claim to be more robust and generalise well but the system needs to though off well as there the approach2 of encoder-decoder is the best in performance if contrasted against approach1 & approach3. This implies that the system must be though of well before any kind of implementation. Finally, the image registration algorithms found their use case from robotics to medical science we need some method that can generalize this much of wast extent.

References

[1] Barbara Zitová and Jan Flusser. Image registration methods: A survey. *Image Vis. Comput.*, 21(11):977–1000, 2003.

[2] Sayan Nag. Image Registration techniques: A survey, 2017.

[3] Medha V Wyawahare, Pradeep M Patil, and Hemant K Abhyankar. Image Registration Techniques : An overview. *Int. J. Signal Process. Image Process. Pattern Recognit.*, 2(3):11–28, 2009.

[4] Denis P. Shamonin, Esther E. Bron, Boudewijn P.F. Lelieveldt, Marion Smits, Stefan Klein, and Marius Staring. Fast parallel image registration on CPU and GPU for diagnostic classification of Alzheimer's disease. *Front. Neuroinform.*, 7(JAN), 2014.

[5] Frank G. Zöllner, Amira Šerifović-Trbalić, Gordian Kabelitz, Marek Kociński, Andrzej Materka, and Peter Rogelj. Image registration in dynamic renal MRI—current status and prospects, 2020.

[6] Rong Ji. Research on Basketball Shooting Action Based on Image Feature Extraction and Machine Learning. *IEEE Access*, 8, 2020.

[7] Iyad Abu Doush and Sahar AL-Btoush. Currency recognition using a smartphone: Comparison between color SIFT and gray scale SIFT algorithms. *J. King Saud Univ. - Comput. Inf. Sci.*, 29(4):484–492, 2017.

[8] Grant Haskins, Uwe Kruger, and Pingkun Yan. Deep learning in medical image registration: a survey. *Mach. Vis. Appl.*, 31(1), 2020.

[9] Sombir Singh Bisht, Bhumika Gupta, and Parvez Rahi. Image Registration Concept and Techniques: A Review. *Int. J. Eng. Res. Appl.*, 4(4):30–35, 2014.

[10] Yujuan Ding, Wai Kueng Wong, Zhihui Lai, and Zheng Zhang. Bilinear Supervised Hashing Based on 2D Image Features. *IEEE Trans. Circuits Syst. Video Technol.*, 30(2):590–602, 2020.

[11] Kavitha Kuppala, Sandhya Banda, and Thirumala Rao Barige. An overview of deep learning methods for image registration with focus on feature-based approaches, 2020.

[12] Bob D. de Vos, Floris F. Berendsen, Max A. Viergever, Hessam Sokooti, Marius Staring, and Ivana Išgum. A deep learning framework for unsupervised affine and deformable image registration. *Med. Image Anal.*, 52:128–143, 2019.

[13] Yipeng Hu, Marc Modat, Eli Gibson, Wenqi Li, Nooshin Ghavami, Ester Bonmati, Guotai Wang, Steven Bandula, Caroline M. Moore, Mark Emberton, Sébastien Ourselin, J. Alison Noble, Dean C. Barratt, and Tom Vercauteren. Weakly-supervised convolutional neural networks for multimodal image registration. *Med. Image Anal.*, 49:1–13, 2018.

[14] Xiao Yang, Roland Kwitt, Martin Styner, and Marc Niethammer. Quicksilver: Fast predictive image registration – A deep learning approach. *Neuroimage*, 158:378–396, 2017.

[15] Daniel Detone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperPoint: Self-supervised interest point detection and description. In *IEEE Comput. Soc. Conf. Com-*

put. Vis. Pattern Recognit. Work., volume 2018-June, pages 337–349, 2018.

[16] Mihai Dusmanu, Ignacio Rocco, Tomas Pajdla, Marc Pollefeys, Josef Sivic, Akihiko Torii, and Torsten Sattler. D2-net: A trainable CNN for joint description and detection of local features. In *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, volume 2019-June, pages 8084–8093, 2019.

[17] Yuki Ono, Pascal Fua, Eduard Trulls, and Kwang Moo Yi. LF-Net: Learning local features from images. In *Adv. Neural Inf. Process. Syst.*, volume 2018-Decem, pages 6234–6244, 2018.

[18] Zhen Zhang and Wee Sun Lee. Deep graphical feature learning for the feature matching problem. In *Proc. IEEE Int. Conf. Comput. Vis.*, volume 2019-Octob, pages 5086–5095, 2019.