

Automatic Classification of Image Registration Problems

Steve Oldridge, Gregor Miller, and Sidney Fels*

University of British Columbia, Electrical and Computer Engineering
{steveo,gregor,ssfels}@ece.ubc.ca
<http://www.ece.ubc.ca/~hct>

Abstract. This paper introduces a system that automatically classifies registration problems based on the type of registration required. Rather than rely on a single “best” algorithm, the proposed system is made up of a suite of image registration techniques. Image pairs are analyzed according to the types of variation that occur between them, and appropriate algorithms are selected to solve for the alignment. In the case where multiple forms of variation are detected all potentially appropriate algorithms are run, and a normalized cross correlation (NCC) of the results in their respective error spaces is performed to select which alignment is best. In 87% of the test cases the system selected the transform of the expected corresponding algorithm, either through elimination or through NCC, while in the final 13% a better transform (as calculated by NCC) was proposed by one of the other methods. By classifying the type of registration problem and choosing an appropriate method the system significantly improves the flexibility and accuracy of automatic registration techniques.

Key words: Image Registration, Computational Photography, Panorama Stitching, Focal Stacking, High-Dynamic Range Imaging, Super-Resolution

1 Introduction

Image registration is the process of calculating spatial transforms which align a set of images to a common observational frame of reference, often one of the images in the set. Registration is a crucial step in any image analysis or understanding task where multiple sources of data are combined. It is commonly used in computational photography [1], remote sensing [2, 3], medical image processing [4, 5], and other computer vision tasks.

Image registration methods vary significantly depending on the type of registration being performed. This paper introduces a system that attempts to automatically classify registration problems based on the variation between image pairs. Rather than rely on a single “best” algorithm to solve all types of registration, the proposed system is made up of a suite of techniques. Image

* Support for this project was provided by Bell Research and the National Science and Engineering Research Council.

pairs are analyzed according to the types of variation that occur between them, and an algorithm designed to handle that variation solves for the misalignment. In instances where multiple types of variation are detected all potentially appropriate algorithms are run, and a normalized cross correlation of the results in their respective error spaces is performed to select which alignment is best.

This approach has significant advantages over single algorithmic approaches which perform poorly outside their respective domains. Conversely, when compared to simply running all available registration algorithms on the pair, the system is both more efficient and also easier to interpret. No single metric exists to compare the quality of a registration, particularly when error spaces are tuned to particular types of registration, so determining which solution is best from amongst several candidates is difficult. Using normalized cross correlation is possible, however it is less effective at comparing across all algorithms because the normalized combination of all error spaces often doesn't represent the best alignment if one or more of the error spaces is inappropriate.

Our system significantly improves the flexibility and accuracy of automatic registration, approaching what Zitova and Flusser [6] term '*the ultimate registration method,*' which is 'able to recognize the type of given task and to decide by itself about the most appropriate solution.'

Section 2 looks at traditional taxonomies of registration and how these can be used to guide in the classification of registration problems. Section 3 examines the basis by which image pairs are classified by the system. Section 4 outlines the system created to automatically register images with different types of variation. The results of the classification of a large set of registration problems by our prototype system is shared in Section 5. Finally Section 6 concludes.

2 Related Work

Systems for automatic image registration exist, however they are most often limited to a single application such as stitching panoramas [7], super-resolution [8, 9], high dynamic range (HDR) imaging [10, 11], or focal stacking [1]. These techniques can be used on a limited subset of problems from other domains, however no single algorithm exists that will solve all types of registration. Yang et. al [12] extend the flexibility of their algorithm within other problem domains by analyzing the input image pairs and setting parameters accordingly, however the single underlying algorithm still fails in a number of their test cases.

Drozd et. al [13] propose the creation of an expert system based tool for autonomous registration of remote sensing data, and outline a plan to use information derived from image metadata and user tags to select from amongst correlation based, mutual information based, feature based, and wavelet based methods. Unfortunately their description is more of a preliminary proposal and doesn't provide results of the performance of their expert system or of how appropriate the registration techniques selected were at solving the problems they were chosen for. To our knowledge no other attempts at classifying registration have been made, either by rule based systems or by learning methods.

Image registration survey papers provide a methodological taxonomy for understanding the different algorithms used to solve the registration problem. Brown [14] divides registration into four components: feature space, search space, search strategy, and similarity metric. Within Brown's framework, knowledge of the types of variation that occur in image sets is used to guide selection of the most suitable components for a specific problem. Variations are divided into three classes; variations due to differences in acquisition that cause the images to be misaligned, variations due to differences in acquisition that cannot be easily modeled such as lighting or camera extrinsics, and finally variations due to movement of objects within the scene. These are labeled by Brown 'corrected distortions', 'uncorrected distortions', and 'variations of interest' respectively. This paper outlines a system that attempts to automatically detect these variations and use them as a basis for classification of the type of registration problem to be solved, which in turn guides in the selection of suitable algorithms. We focus on Brown's 'corrected distortions,' which are easier to detect and provide significant guidance in the selection process.

More recently, Zitová and Flusser [6] have differentiated the field of registration into area and feature based methods. The four basic steps of image registration under their model are: feature detection, feature matching, mapping function design, and image transformation and resampling. While they do not provide a model of variation equivalent to Brown's, they discuss in detail the advantages and drawbacks of each method, allowing a similar mapping of methodology from situation. In the conclusion of their survey of registration techniques they propose the creation of '*the ultimate registration method*,' which is 'able to recognize the type of given task and to decide by itself about the most appropriate solution.' This paper is an attempt at just that.

Examining these surveys reveals a number of common forms of variation: purely spatial variation, significant variations in intensity, significant variations in focus, variations in sensor type, and finally variations in structure. We have implemented three representative algorithms as a part of our system: a feature based method [7] for purely spatially varying problems, a median thresholding method [10] for intensity varying problems, and finally an area based method [15] for focus varying problems. We do not classify or solve for sensor or structure variations, however these are logical additions to the system.

3 Problem Classification

Image registration methods vary significantly depending on the type of registration being performed. Within our system image pairs are organized into the categories: spatially varying, intensity varying, focus varying, and unrelated, based on their primary form of variation. Examining the types of variations that occur in a pair or sequence of images allows photographers to select an appropriate application, or programmers to select an appropriate algorithm, in order to find the best alignment.

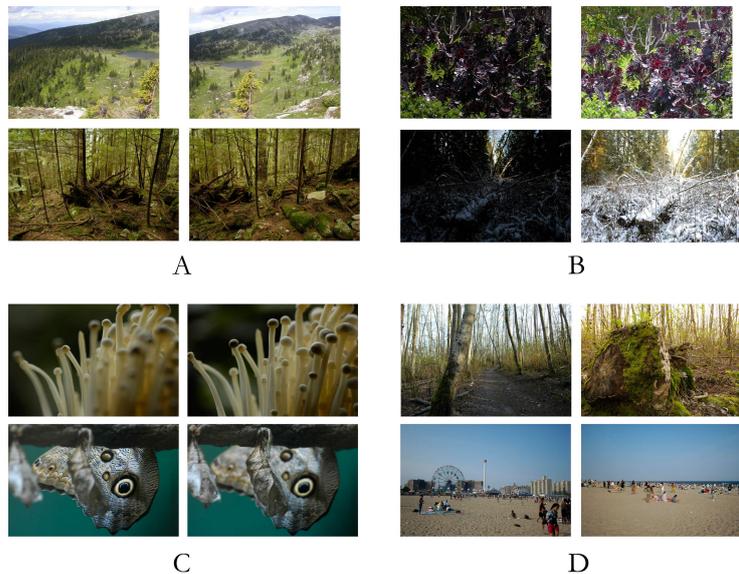


Fig. 1. Image pairs representative of the different types of variation that occur in registration problems. A: Spatially Varying B: Intensity Varying C: Focus Varying D: Unrelated

Similarly in our system each image pair is analyzed to determine the differences in their intensity histograms and hue/saturation histograms, the normalized power of each image, the number of matching features between the images, and the centroid of those matches. Differences between histograms are measured by their intersection. Figure 1 shows two representative image pairs of each type of registration, and Table 1 presents their corresponding values. These values are used by the system to classify what type of variations occur through the application of simple heuristic rules that utilize these operators. Our system is capable of running many algorithms and comparing the results to find the best solution, therefore it is much more important to make true positive classifications than it is to prevent false positives. The basis for these rules within each application domain is examined in detail below.

3.1 Purely Spatial Variations

Image pairs that differ purely spatially, as shown in Figure 1A, are the most common type of image registration problem. Applications that require registration of images that vary spatially include panorama stitching, super resolution, and remote sensing. Although area-based methods derivative of Lucas and Kanade [15] are capable of solving these types of registration problems, feature based methods like Autostitch [7] and Autopano are the most common technique applied

Image Pair	N Feat	Centroid	I	HS	Power
A1 (Spatial)	967	(0.76,0.51)(0.22,0.61)	0.765	0.787	3.30%, 3.10%
A2 (Spatial)	605	(0.71,0.65)(0.30,0.65)	0.923	0.878	3.30%, 3.05%
B1 (Intensity)	944	(0.53,0.62)(0.53,0.62)	0.161	0.303	3.79%, 1.80%
B2 (Intensity)	1483	(0.53,0.60)(0.43,0.50)	0.425	0.651	3.60%, 3.24%
C1 (Focus)	139	(0.52,0.53)(0.52,0.53)	0.834	0.812	1.40%, 0.96%
C2 (Focus)	50	(0.52,0.36)(0.58,0.32)	0.862	0.834	0.16%, 0.43%
D1 (Unrelated)	24	(0.49,0.25)(0.44,0.28)	0.898	0.819	3.18%, 2.76%
D2 (Unrelated)	10	(0.47,0.6)(0.58,0.73)	0.746	0.704	1.88%, 1.73%

Table 1. Values used in the classification of image pairs corresponding to images from Figure 1. For each pair the number of features (N Feat), feature centroid (Centroid), overlap of intensity histogram (I), overlap of hue saturation histogram (HS), and power of each image (Power) is calculated.

and are generally considered much more accurate unless the image pairs contain little high-frequency information from which to find and match features.

Without first aligning the images, calculating the amount of *overlapping* high frequency content in the image pairs is difficult, so instead we calculate the number of matched features [16] directly. Image pairs with on average more than one matched feature per 75x75 pixel patch are classified as ‘purely spatial’ because methods unconcerned with other forms of variation (i.e. feature based methods) are likely to be capable of solving for their alignment. Stitching problems with low overlap are likely to contain a low number of matching features, so we also calculate the centroid of the features detected, allowing us to distinguish these cases. Pairs with feature centroids greater than 30% translation from the origin are considered purely spatial require 1/5 as many matches. Section 5 shows how the results of the combination of these two rules allows us to positively classify purely spatially varying image pairs within our test set.

3.2 Intensity Variations

Significant intensity variations are common amongst high dynamic range image registration problems, and can also appear in panorama image pairs where there is a powerful light source in one of the frames. HDR techniques are predominantly area based; interest points required by feature based methods are most often detected at edges or corners, and are not consistent across large differences in intensity. For those image pairs where image intensity varies significantly, such as those shown in Figure 1B, median thresholding [10] can be used to find a more accurate registration.

Intensity varying image pairs can be easily detected by examining the differences in intensity histograms, providing a simple basis for their classification. Pairs with histograms that differ by more than 30% are classified as intensity varying. Section 5 demonstrates the effectiveness of this rule at finding intensity varying image pairs within our test set.

3.3 Focus Variations

Focus variations are found in image pairs used for focus stacking, and in pairs with motion or gaussian blur, shown above in Figure 1C,. Techniques are predominantly area based for the same reason as HDR techniques; the same edges and corners are not detected across images with different focal planes. Instead intensity based area methods like those derivative of Lucas and Kanade [15] must be used to find the correct alignment.

Focus stacking is used to combine images with limited depth of field, so images are likely to have a low amount of high frequency information. Image pairs are detected by examining the normalized power of each of the images, a measure proportional to the number of in focus pixels in the image. In a number of problems, particularly those relating to registering blurred images, only one of the images is lacking in focus. Image pairs where either image has a normalized power less the 2.5% are classified as focus varying. As we will see in Section 5 this rule is useful for positively classifying focus varying pairs, however it also classifies a number of other pairs which are not considered as primarily focus varying in our ground truth.

4 System

Using the rules described in Section 3 our system is able to identify the types of variation occurring between the image pair. If only a single type of variation is identified then the corresponding algorithm is run to solve for the transform that aligns the pair. As mentioned in Section 2, we have implemented three representative algorithms as a part of our system: a feature based method [7] for purely spatially varying problems, a median thresholding method [10] for intensity varying problems, and finally an area based method [15] for focus varying problems.

When multiple forms of variation are classified for an image pair the system uses all appropriate algorithms, solving for each transform. Normalized cross correlation of the proposed transforms is then performed, calculating the error of each transform across all appropriate error spaces to pick the best. For both our area based and median thresholding methods this value can be calculated directly using the same error function that they use to align the images. Unfortunately this is not directly possible in our feature based method, which uses the number of unmatched features as its metric for finding a transform. Instead an error space based on joint intensity of image patches is used.

5 Evaluation

Image registration methods vary significantly depending on the type of registration being performed. Knowing the types of variations that occur in a pair or sequence of images allows photographers to select an appropriate application or programmers to select an appropriate algorithm in order to find the

best alignment. To test our algorithm we created a set of 64 image pairs from the categories: spatially varying, intensity varying, focus varying, and unrelated, based on their primary form of variation. These images were then classified in a user study by six independent photographers. For each pair we considered the classification to be valid if five of the six photographers classified the image pairs exactly the same, a process which eliminated four pairs. This set of uniformly classified pairs was then used as a ground truth for evaluating the system. For the remaining 60 images the photographers were on average 96% successful at correctly classifying the main form of variation. This allows us to compare how well our system is able to classify registration problems.

Using our two rules outlined in 3.1 we can positively identify 100% of the purely spatial varying problems within the data set. 38% of pairs classified as primarily spatially varying were also proposed as being intensity or focus varying. Once normalized cross correlation has been applied 76% of the purely spatial (according to our ground truth) pairs find the best alignment using the correspond spatial method. Examination of the remaining 24% of spatial pairs shows that in 60% of cases all error spaces agreed the solution chosen was the best, while 40% produced conflicting recommendations.

As expected, intensity varying image pairs can be easily detected by examining the differences in intensity histograms as proposed in 3.2. Using this rule we are able to find 100% of the intensity varying image pairs within the data set. 91% of ground truth intensity varying pairs were also indicated by either the spatial and/or focus varying rules. After NCC however 81% of the selected solutions were from the intensity varying method. The remaining 19% were selected from the spatially varying method.

Similarly, using the rule proposed in 3.3 we are able to classify 94% of the focus varying problems in our test set. 16% of ground truth pairs were also classified as spatially varying, however after NCC all of the solutions were selected from the focus varying method.

Unclassified image pairs are considered to be unrelated by the system. 38% of the unrelated image pairs were correctly identified by the system. A single focus varying problem was also indicated as being unrelated. This poor rate of classification of unrelated image pairs derives from nature of our rules, which were chosen to identify differences in intensity and focus between images, a common occurrence in unrelated images.

Overall the system is able to positively classify 90% of the problems correctly. Removing unrelated image pairs from the set increases the correct classification rate to 98%. 55% of the problems were correctly classified with no alternative variation suggested and were solved using their appropriate method. A further 32% selected the solution by the corresponding method for their classification through normalized cross correlation. Finally, for the remaining 13% of image pairs, 71% of the solutions selected by the system were lowest in all error spaces being considered, suggesting that they represent a better alignment than that proposed by the 'correct' method. Figure 2 shows this set of images. Table 2 summarizes the results of our system's performance classifying the test set.



Fig. 2. Image pairs that were aligned using a method other than that suggested by their main form of variation.

Further verification of the resulting transforms created by the system is unfortunately a difficult prospect. The unbiased evaluation of registration techniques is a significant problem within the field, particularly when error surfaces or representations used by different methods don't agree upon a decisive 'best' solution. Zitová and Flusser [6] identify three measures of registration accuracy in their survey: localization error, matching error, and alignment error. Localization error represents mistakes in the location of feature based methods' interest regions. Matching error is measured as the number of false matches between features. Finally, alignment error measures the difference between the proposed alignment and the actual between-image geometric distortion. Localization and matching error are specific to feature based methods and measures of common problems with the steps of those methods, and are difficult to apply to area based approaches. Alignment error is a much more desirable measure, however this requires a ground truth.

Azzari et al. recently propose the use of a set of synthetic data with a known ground truth which they have made available online [17]. Their error measure consists of a combination of sum of square difference of image intensity and distance metrics that measure the displacement of known points within the image. Unfortunately their image sets are low resolution (320x240), limited to translation and rotation, and contain no variation in intensity or focus, sensor, limiting their usefulness. A much more robust and high resolution set is necessary for evaluation of modern registrations, and would be a critical contribution to the field of registration. Proof that the 13% of ground truth pairs that were aligned with the other methods selected by normalized cross correlation remains impos-

sible to provide without such a data set. A visual inspection seems to suggest that in most cases they are, however this type of evaluation is empirical at best.

Ground Truth Classification	Identified	W No Alt	After NCC
Spatial	100%	62%	76%
Intensity	100%	9%	81%
Focus	94%	78%	94%
Unrelated	38%	38%	38%
Total Related	98%	56%	87%

Table 2. Summary of the system's classification rate.

6 Conclusions and Future Work

We have introduced a novel automatic registration system that attempts to automatically classify registration problems based on the variation between image pairs. The system was validated using a test set of 60 pre-classified image pairs verified by an independent user study of photographers. The system was able to identify 98% of the related ground truth pairs' main form of variation. 55% of pairs were correctly identified by a single form of variation allowing immediate selection of an algorithm. A further 32% of pairs proposed transforms were correctly selected using normalized cross correlation on the solution space of the proposed algorithms. Visual inspection of the final 13% of pairs suggests that the alignments proposed are superior to the solution found by the 'correct' algorithm, however verification of this is hypothetical at best. Empirical evaluation of registration is impossible without knowledge of the ground truth alignments.

Although our system focused on spatial, intensity, and focus variations, extension into automatic registration for sensor and structure variations would greatly benefit researchers, particularly those within the medical imaging community. Such a system would require a greater degree of differentiation between problem types and would likely rely more heavily on image metadata to distinguish the variations between image pairs. Additionally learning based methods such as support vector machines, PCA analysis or Bayesian networks represent an avenue of improvement upon our simple heuristic rule based system. Our initial test set of 60 images is sufficient to prove the concept of the system and provide a frame of reference for establishing a set of rules that work for the given images, however a more robust approach would require a substantially larger set of classified images, particularly if using learning based methods. Ideally this set would also contain ground truth translations as proposed in [17], allowing for direct evaluation of performance post-classification.

The correspondence of a set of images using a system that selects which pairs to align would increase the flexibility of our system further. Autostitch

[7] provides an excellent example of how this can be done for spatially varying problems.

Finally, our system deals with registration only. The inclusion of image re-sampling and transformation techniques appropriate to the type of variation detected, such as tone mapping or focal stacking, would greatly enhance the usefulness of this system as a tool for photographers and researchers alike.

References

1. Agarwala, A., Dontcheva, M., Agrawala, M., Drucker, S., Colburn, A., Curless, B., Salesin, D., Cohen, M.: Interactive digital photomontage. In: SIGGRAPH '04: ACM SIGGRAPH 2004 Papers, New York, NY, USA, ACM Press (2004) 294–302
2. Lillesand, T.M., Kiefer, R.W.: Remote Sensing and Image Interpretation. 6th edn. Wiley (2007)
3. Campbell, J.B.: Introduction to remote sensing. 4th edn. Guildford Press (2008)
4. Maintz, J., Viergever, M.: A survey of medical image registration. *Medical Image Analysis* **2**(1) (1998) 1–36
5. Pluim, J., Maintz, J., Viergever, M.: Mutual-information-based registration of medical images: a survey. *Medical Imaging, IEEE Transactions on* **22**(8) (Aug. 2003) 986–1004
6. Zitov, B., Flusser, J.: Image registration methods: a survey. *Image and Vision Computing* **21** (2003) 977–1000
7. M.Brown, Lowe, D.G.: Recognising panoramas. *Computer Vision, 2003.Proceedings. Ninth IEEE International Conference on* **2** (16-16 Oct. 2003) 1218–1225
8. Flusser, J., Zitov, B., Suk, T.: Invariant-based registration of rotated and blurred images. In: *IEEE 1999 International Geoscience and Remote Sensing Symposium. Proceedings, IEEE Computer Society* (1999) 1262–1264
9. Zitov, B., Kautsky, J., Peters, G., Flusser, J.: Robust detection of significant points in multiframe images. *Pattern Recogn. Lett.* **20**(2) (1999) 199–206
10. Ward, G.: robust image registration for compositing high dynamic range photographs from handheld exposures. *Journal of Graphics Tools* **8** (2003) 17–30
11. Schechner, Y.Y., Nayar, S.K.: Generalized mosaicing: High dynamic range in a wide field of view. *Int. J. Comput. Vision* **53**(3) (2003) 245–267
12. Yang, G., Stewart, C., Sofka, M., Tsai, C.L.: Registration of challenging image pairs: Initialization, estimation, and decision. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **29**(11) (Nov. 2007) 1973–1989
13. Drozd, A.L., Blackburn, A.C., Kasperovich, I.P., Varshney, P.K., Xu, M., Kumar, B.: A preprocessing and automated algorithm selection system for image registration. Volume 6242., *SPIE* (2006) 62420T
14. Brown, L.G.: A survey of image registration techniques. *ACM Computing Surveys* **24** (1992) 325–376
15. Lucas, B.D., Kanade, T.: An iterative image registration technique with an application to stereo vision (darpa). In: *Proceedings of the 1981 DARPA Image Understanding Workshop*. (April 1981) 121–130
16. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60** (2004) 91–110
17. Azzari, P., Stefano, L., Mattoccia, S.: An evaluation methodology for image mosaicing algorithms. In: *ACIVS '08: Proceedings of the 10th International Conference on Advanced Concepts for Intelligent Vision Systems, Berlin, Heidelberg, Springer-Verlag* (2008) 89–100