Anisotropic Methods for Image Registration

James Murphy

November 24, 2014

Goals of Talk

1. Explain the problem of *image registration*.
2. Discuss existing methods for image registration, many of which use harmonic analysis.
3. Detail an algorithm based on *anisotropic features* extracted from the images. This work is joint with Jacqueline Le Moigne (NASA) and David J. Harding (NASA).

Introduction to image registration

- The process of image registration seeks to align two or more images of approximately the same scene, acquired at different times or with different sensors.
- Images can differ in many ways:
  1. Geometrically: rotated, translated, warped, dilated.
  2. Modally: different sensors, different conditions at time of image capture.
- Noise may be present.
- This problem is relevant to, among other fields, microscopy, biomedical imaging and remote sensing.

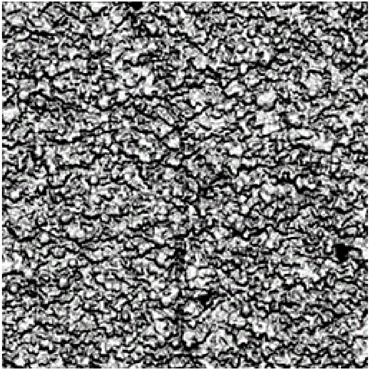
Importance of image registration

- Need to be able to know exact location of a newly captured image; this requires registration against a known image.
- Registration is the first step in image fusion.
- Related to more general problems in computer vision.

Difficulties in registration

- Any registration algorithm uses the content of the image; how it does so varies substantially.
- Difficult images to register include those with few dominant features, and images of different modes.
- We are particularly interested in the second case, of *multimodal registration*.
- Harmonic analytic techniques are well-suited for these types of problems, when compared to other methods.

A motivational example



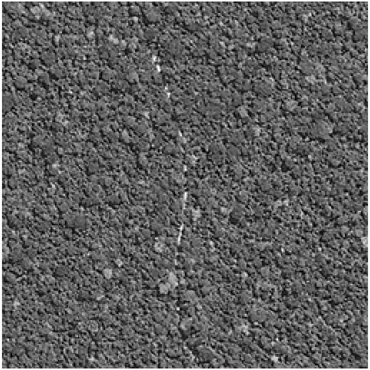


Figure: A LIDAR and optical image of the Amazon rainforest. These images are very homogeneous. How can we register them? It is hard enough for the human eye to do. Can we use mathematical tools to efficiently extract features to be used for matching?

Stages of image registration

Image registration may be viewed as the combination of four separate processes:

1. Selecting an appropriate **search space** of admissible transformations. This will depend on whether the images are at the same resolution, and what type of transformations will carry the input image to the reference image, i.e. rotation-scale-translation (RST), polynomial warping, etc.
2. Extracting relevant **features** to be used for matching. These could be individual pixels that are known to be in correspondence between the two images, or could be global structures in the images, such as roads, buildings, rivers, and textures.
3. Selecting a **similarity metric**, in order to decide if a transformed input image closely matches the reference image. This metric should make use of the features which are extracted from the image, be they specific pixels or global structures.
4. Selecting a **search strategy**, which is used to match the images based on maximizing or minimizing the similarity metric.

1. **Manual Registration**: A human selects matching pixels in the two images, and the transformation that registers them is computed by minimizing the distance between these pixel pairs.
2. **Algorithmic Pixel Matching**: Same as above, but with an algorithm executed automatically. The SIFT algorithm is popular and effective for images *of the same modality*
3. **Global Feature Matching**: Algorithmically determine robust, sparse features in the images, then compute registration based on these features.

The third class has a strong connection with harmonic analysis.

Problem with pixel-based methods

- Finding pairs of corresponding pixels is very difficult in the case of multimodal images. Even SIFT (state of the art) struggles.
- The stumbling block is that although there might be *global correspondences* stemming from large-scale features, these are not induced by *local correspondences*.
- Consider two images of a rural scene in WA: one LIDAR, one optical.

WA multimodal images

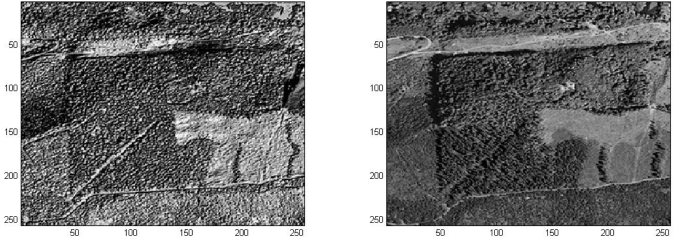


Figure: LIDAR and optical images of a scene in WA state.

SIFT fails for multimodal images

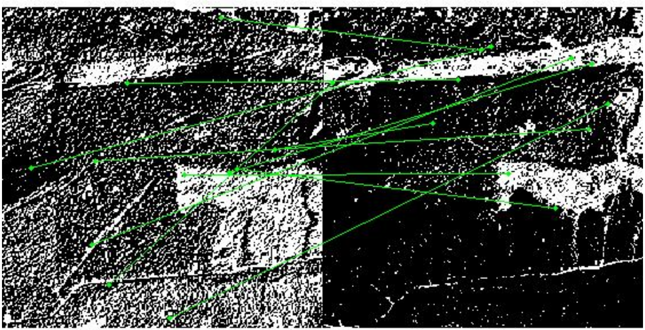
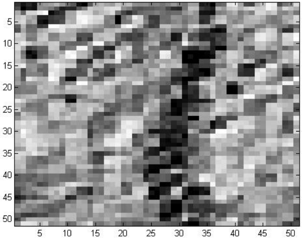


Figure: The “matching” pixels computed in the LIDAR and optical images of WA using the SIFT algorithm. Note the lack of correspondence; such points are unusable for a registration algorithm.

Local dissimilarity in multimodal images

- In multimodal images, the similarities between the images are only manifested on the *global* scale.
- Locally, the images appear dissimilar.



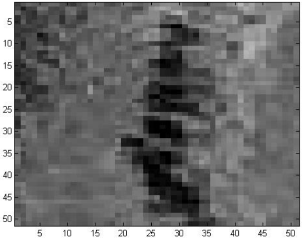


Figure: The same hedge in the LIDAR and optical image. Although there is clear correspondence at the macroscopic level, it is difficult to find pixel-to-pixel correspondences.

Global features for registration

- In order to efficiently register multimodal images, we must use *global features*.
- Substantial work in registration using global wavelet features has been done.
- We are interested in whether generalizations of wavelets, which have anisotropic character, can improve upon these methods.

Shearlets for registration

- We considered the anisotropic representation system of *shearlets*.
- There is mathematical theory explaining that shearlets represent certain images classes with optimal sparsity.
- This is good for registration algorithms, because sparse features increase the robustness of the optimization algorithm that computes the registration transformation.

Background on wavelets

- Mathematically, the continuous wavelet transform decomposes a signal according to *scale* and *translation*.
- For a signal $f \in L^2(\mathbb{R}^2)$ and an appropriately chosen wavelet function ψ, f may be decomposed as

$$f \sim \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \langle f, \psi_{m,n} \rangle \psi_{m,n}, \tag{1}$$

  where $\psi_{m,n}(x) = 2^{\frac{m}{2}} \psi(A^m x - n)$ and $A = \begin{pmatrix} 2 & 0 \\ 0 & 2^{\frac{1}{2}} \end{pmatrix}$.
- The *wavelet coefficients* $\langle f, \psi_{m,n} \rangle$ describe the behavior of f at different scales; $m >> 0$ large gives information at local scales, $m << 0$ gives global information.

Background on shearlets

- Shearlets generalize wavelets by decomposing with respect not just to *scale* and *translation*, but also *direction*.
- Mathematically, given a signal $f \in L^2(\mathbb{R}^2)$ and an appropriate base function ψ, we may decompose f as

$$f \sim \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{k \in \mathbb{Z}^2} \langle f, \psi_{i,j,k} \rangle \psi_{i,j,k}, \tag{2}$$

where:

- $\psi_{i,j,k}(x) = |\det A|^{\frac{i}{2}} \psi(B^j A^i x - k)$.
- $A = \begin{pmatrix} 2 & 0 \\ 0 & 2^{\frac{1}{2}} \end{pmatrix}$, $B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.

An SAR image

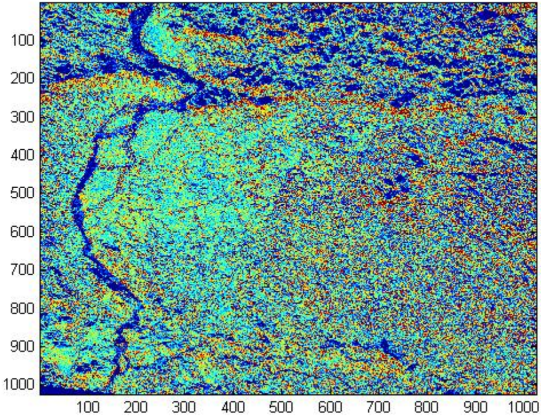


Figure: A synthetic aperture radar image containing several edge-like features.

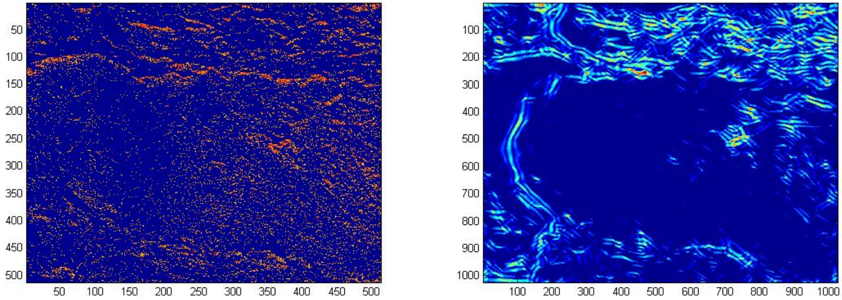


Figure: Wavelet-like and shearlet features extracted from the original SAR image.

Evaluating registration quality

- When registering images, there are two significant criteria of registration algorithm quality:
  1. **Accuracy** of computed registration when compared to the true registration.
  2. **Robustness** of algorithm to initial distance between the two image.
- The accuracy of the registration is obviously important.
- The robustness of the algorithm is important because the initial closeness of the two images depends greatly on the GPS technology in the sensors and the distance of the sensing device to the location being imaged.

Importance of robustness

- If the images to be aligned start far apart, the registration algorithm could fail to converge.
- This situation occurs if the initial registration between the images is poor, which can happen for a variety of reasons: poor GPS, lost data, satellite orbit far from earth, etc.
- Our hypothesis was that using shearlet features would increase registration robustness, due to their sparsity in representing anisotropic features.

Experimental Overview

- As we prototyped, we realized that using shearlets did increase robustness, but at a slight cost in accuracy, usually a few pixels.
- Consequently, we devised a two-stage registration algorithm: first, use shearlets to get an *approximate registration*, then refine this with another iteration of the algorithm, using wavelet features.
- We compared this algorithm to using wavelets alone.
- We performed experiments on *synthetic* images, as well as *multimodal* image pairs.

Basic description of algorithm

1. *Search Space:* RST. All of our examples feature images at the same scale, so effectively, our search space is the space of rotations and translations (RT).
2. *Features*: Wavelet features in one case and shearlet features coupled with wavelet features in another.
3. *Similarity Metric*: Unconstrained least squares. That is, if F_R and F_I are the reference and input features, N the number of relevant pixels, (x_i, y_i) the integer coordinate of each pixel, and T_p the transformation associated to parameters p, we seek to minimize the similarity metric given by

$$\chi^2(p) = \frac{1}{N}\sum_{i=1}^{N}(F_R(T_p(x_i, y_i)) - F_I(x_i, y_i))^2$$

4. *Search Strategy:* Modified Marquadt-Levenberg method of solving non-linear least squares problem.

Algorithm details (1/2)

1. Input a reference image, I_r, and an input image I_i. These will be the images to be registered.
2. Apply shearlet feature algorithm and wavelet-like feature algorithm to I_r and I_i. This produces a set of shearlet features for both, denoted $S_1^r, ..., S_n^r$ and $S_1^i, ..., S_n^i$, respectively, as well as a set of wavelet features for both, denoted $W_1^r, ..., W_n^r$ and $W_1^i, ..., W_n^i$. Here n refers to the level of decomposition chosen.
3. Match S_1^r with S_1^i with a least-squares optimization algorithm to get a transformation T_1^S. Using T_1^S as an initial guess, match S_2^r with S_2^i, to acquire a transformation T_2^S. Iterate this process by matching S_j^r with S_j^i using T_{j-1}^S as an initial guess, for $j = 2...., n$. At the end of this iterative matching, we acquire our *final shearlet-based registration*, call it T^S.

Algorithm details (2/2)

4. Using T^S as our initial guess, match W_1^r with W_1^i with a least-squares optimization algorithm to acquire a transformation T_1^W. Using T_1^W as an initial guess, match W_2^r with W_2^i, to acquire a transformation T_2^W. Iterate this process by matching W_j^r with W_j^i using T_{j-1}^W as an initial guess, for $j = 2, ..., n$. At the end of this iterative matching, we acquire our *final registration*, call it T.
5. Output $T = (T_x, T_y, T_\theta)$.

Computing harmonic-analytic features

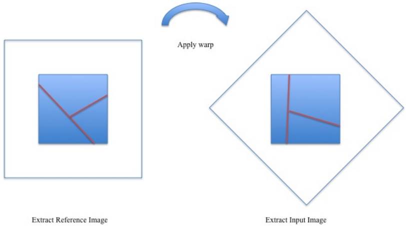
- The wavelet features to be used come in three classes, all implemented in C:
  1. Spline wavelets.
  2. Simoncelli band-pass wavelet-like features.
  3. Simoncelli low-pass wavelet-like features.
- The shearlet features are based on the Kaiserslautern MATLAB package, coded by S. Häuser.
- All coefficients are thresholded before computing the registration transformation via the Marquadt-Levenberg optimization scheme.

Overview of experiments

- We shall discuss three classes of experiments: one synthetic, two multimodal.
- We shall perform, for each image pair, many iterations of our algorithm. Each iteration, we shall *move the images farther apart*.
- The distance shall be parametrized by rotation and translation in the x and y directions. For convenience, these are coupled together as RT. So, $RT = 1.8$ means a rotation of 1.8 degrees and a translation of 1.8 pixels in both the x and y direction. Fraction translations and rotations are interpolated by splines.

Introduction to synthetic experiments

- The synthetic experiments are performed by fixing a reference image, applying a warp (rotation and translation), then extracting a corresponding input image.



- This is convenient, since we know the exact registration between the input and reference images. However, the images are of the same mode (the same image, in fact), so this is not a true test for multimodal image registration.

Landsat image for synthetic experiments

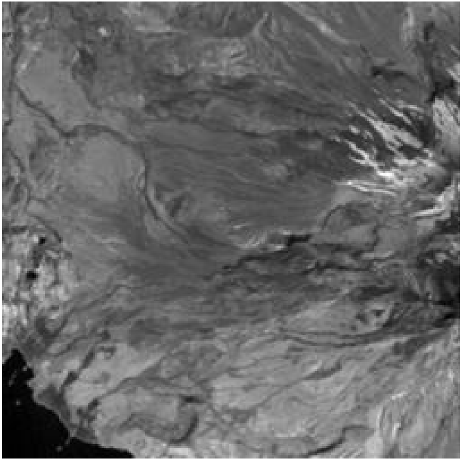


Figure: Landsat-TM scene used for synthetic experiments.

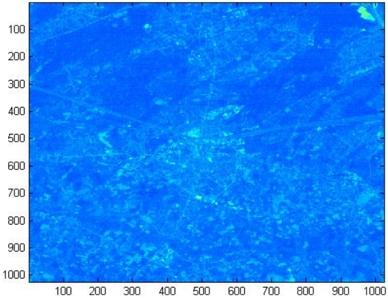
Reg. Technique	# Conv. (out of 200)	RMS	% Conv.	Rel. Impro.
Spl. Wavelets	108	.0019	54%	-
Sim. Band-Pass	21	.0045	10.5%	-
Sim. Low-Pass	113	.0040	55.5%	-
Shear.+ Spl. Wave.	154	.0058	77%	42.45%
Shear.+ Sim. L-P	154	.0080	77 %	633.33%
Shear. + Sim. B-P	154	.0081	77 %	30.83%

Table: Comparison of registration algorithms for Landsat-TM synthetic experiment.

Introduction to multimodal experiments

- We now consider experiments with registering two real images, with different modalities and radiometries.
- This represents a more realistic test of the functionality of our algorithms, since in reality, image registration will be between two different images, not an image and a synthetic translation and rotation of itself.
- To test our algorithm, we warp our input image according to the RT parameter, and attempt to recover the registration that will bring the images back into alignment.
- The images are not in alignment to begin with, so the true registration was computed by hand.

Belgium multimodal images



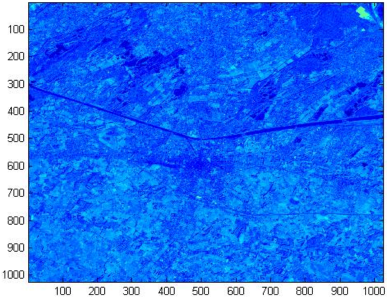


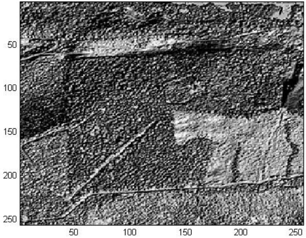
Figure: Multispectral and panchromatic images of Hasselt, Belguim.

Belgium multimodal results

Reg. Technique	# Conv.	RMS	% Conv.	Rel. Impro.
Spl. Wavelets	8	.6376	7.92%	-
Sim. Band-Pass	19	.7534	18.81%	-
Sim. Low-Pass	14	.6034	13.86%	-
Shear. + Spl. Wave.	20	.5185	19.80%	150%
Shear. + Sim. L-P	27	.6494	26.73%	42%
Shear. + Sim. B-P	20	.5513	19.80%	43%

Table: Comparison of Registration Algorithms for panchromatic to multispectral experiment.

WA multimodal images



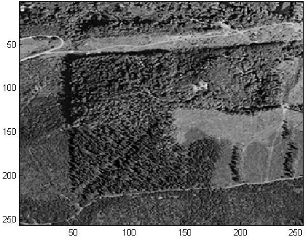


Figure: LIDAR and optical images of a scene in WA state.

WA multimodal results

Reg. Technique	# Conv. (out of 101)	RMS	% Conv.	Rel. Impro.
Spl. Wavelets	55	1.3439	54.46%	-
Sim. Band-Pass	61	1.5862	60.40 %	-
Sim. Low-Pass	86	1.4848	85.15%	-
Shear.+ Spline Wave.	60	1.3144	59.41 %	9%
Shear. + Sim. L-P	65	1.5836	64.36 %	7%
Shear. + Sim. B-P	88	1.4861	87.13 %	2%

Table: Comparison of Registration Algorithms for WA Lidar to optical experiment.

Experiment conclusions

- Overall, using shearlets and wavelets together outperforms using only wavelets.
- The improvement appears more pronounced when there are substantial edges.
- When the image is texturally dominant, there is minimal improvement.

Future directions

1. Use other anisotropic methods besides shearlets: composite wavelets, curvelets.
2. Instead of using LIDAR data in the format of an image, use the “raw data,” the *digital elevation model (DEM)*.
3. Registration of hyperspectral data cubes, i.e. registering more than two images at a time.

Thank you for your time!