

Agile multi-scale decompositions for automatic image registration

James M. Murphy^a, Omar Navarro Leija^b, Jacqueline Le Moigne^c,

^aDuke University, Department of Mathematics, Durham, NC 27708

^bUniversity of Nevada, Las Vegas, Department of Computer Science, Las Vegas, NV 89154

^cNASA Goddard Space Flight Center, Software Engineering Division, Greenbelt, MD 20771

ABSTRACT

In recent works, the first and third authors developed an automatic image registration algorithm based on a multiscale hybrid image decomposition with anisotropic shearlets and isotropic wavelets. This prototype showed strong performance, improving robustness over registration with wavelets alone. However, this method imposed a strict hierarchy on the order in which shearlet and wavelet features were used in the registration process, and also involved an unintegrated mixture of MATLAB and C code.

In this paper, we introduce a more agile model for generating features, in which a flexible and user-guided mix of shearlet and wavelet features are computed. Compared to the previous prototype, this method introduces a flexibility to the order in which shearlet and wavelet features are used in the registration process. Moreover, the present algorithm is now fully coded in C, making it more efficient and portable than the MATLAB and C prototype. We demonstrate the versatility and computational efficiency of this approach by performing registration experiments with the fully-integrated C algorithm. In particular, meaningful timing studies can now be performed, to give a concrete analysis of the computational costs of the flexible feature extraction. Examples of synthetically warped and real multi-modal images are analyzed.

Keywords: Image registration, shearlets, multiscale representations, agile algorithms, multimodal images

1. INTRODUCTION

Image registration is the process of aligning two or more images of approximately the same scene, possibly captured with different sensors or at different times.¹ The registration of multimodal images is of particular significance in medical imaging² and remote sensing,³ where different sensors often produce very different types of images. A variety of approaches to the multimodal registration problem have been proposed, including those based on SIFT and related features,⁴ as well as those that attempt to efficiently represent the images to be registered in a common feature space. While local features like SIFT do well in many cases, for images with very different information content, there is often very little local similarity between the two images. This renders local feature descriptors ineffective for image registration, though robust outlier detection can compensate to some extent.⁵

Wavelets⁶ have been successful in extracting global features in common between two images that appear quite different.⁷ By extracting some of the most significant features in the images, thresholded wavelet features remove much of the noise and local differences that pose a challenge in multimodal registration. However, wavelets are known to be theoretically suboptimal for sparse feature extraction for a large class of image signals.^{8–10} This known suboptimality motivated the field of anisotropic harmonic analysis, which developed a number of

Further author information: (Send correspondence to J.M.M.)
J.M.M.: E-mail: jmmurphy11@gmail.com

constructions that incorporate directionality, including contourlets,¹¹ curvelets,⁹ directional Gabor systems,^{12,13} and shearlets.¹⁴

It is known that shearlets and curvelets are theoretically near-optimal in generating sparse representations of *cartoon-like* images, which can be understood as images that are smooth except for smooth boundaries. Many remotely sensed images fall to some extent into this regime, which suggested the use of such anisotropic dictionaries in the registration of remotely sensed images. Recent work^{15,16} proposed both a stand-alone shearlet feature registration algorithm, and a novel two-stage registration algorithm using shearlets for the first stage of registration, and wavelets to refine. The hybrid algorithm acquired a robust, though occasionally somewhat inaccurate, first stage shearlet registration, which was refined with the less robust, but sometimes more precise, wavelet registration. This algorithm provided substantial improvement over wavelet-only registration, on a variety of synthetically warped and real multimodal images.

However, this prototype algorithm was unsatisfactory in several respects. While the wavelet decomposition and feature-matching portions of the algorithm were coded in C, and were thus fast and portable, the shearlet features component was coded in MATLAB.¹⁷ This reduced speed considerably, and limited the potential use of the prototype algorithm by remote sensing scientists. Moreover, the prototype algorithm tested only a specific combination of shearlet and wavelet features: shearlet features, from coarse to fine, followed by wavelet features, from coarse to fine. While this had heuristic motivation and provided strong numerical results, other ordering and combinations of the shearlet and wavelet features are possible.

The present article addresses these concerns. We demonstrate the fully-integrated C algorithm on a synthetic and real multimodal dataset. Moreover, we consider other combinations of the shearlet and wavelet features, to evaluate the potential benefit in using different orders of these features.

2. PROPOSED METHODS

The proposed algorithms were prototyped in recent works of the first and third authors.^{15,16} These algorithms were based on the observation that while wavelet-based registration was usually accurate and efficient, even for large classes of multimodal images, it was often not robust to the choice of initial registration guess. By examining the features produced in the wavelet registration algorithm currently used at NASA,⁷ it was hypothesized that this was due to the speckled and diffuse features often produced by wavelet algorithms. Indeed, it appeared that the features produced were not particularly sparse. This observation is in accordance with the theory of wavelets. In particular, wavelets are known not to be optimally sparse for an important class of images, namely cartoon-like images, which are composed of smooth images with smooth edges; see¹⁶ for a fuller exposition on the mathematical limitations of wavelets for image registration, and¹⁰ for a more general discussion.

The proposed algorithms aimed to improve on this deficiency of wavelets by incorporating a generalization of wavelets known as shearlets.^{8,14} Shearlets are one of many redundant anisotropic frames that incorporate both multiscale and directional decomposition; in comparison, wavelets are multiscale but not directional. While ridgelets and curvelets are earlier methods that incorporate anisotropy into the multiscale regime, shearlets were considered over these other methods for two major reasons. First, shearlets incorporated directionality not by rotations, which are often difficult to interpolate in the discrete setting, but by the action of shearing. Shearing preserves the digital grid, meaning that one need not perform complicated interpolation when incorporating directionality.¹⁰ Second, many implementations of shearlets exist, which gave us flexibility in how to implement of shearlet features algorithm. While the ShearLab software package¹⁸ and the implementation of King¹⁹ quite popular, we chose to use the fast finite shearlet transform (FFST),¹⁷ due its simple and user-friendly implementation.

Originally, a shearlet-based registration approach was proposed in which only shearlet features were used to register the images.¹⁵ This had theoretical justification, and numerical experiments confirmed that it offered improved robustness to the initial registration guess when compared to wavelets, but suffered from a small loss

in accuracy in some cases. The loss in accuracy is due to the lack of translation invariance for shearlets, and also a double-walling artifact that appears around thin edges in the shearlet features. Thus, a more sophisticated registration hybrid algorithm was proposed, in which images would first be registered with shearlets, then with wavelets.¹⁶ While this algorithm enjoyed improved robustness, it suffered from at least one major computational weakness. While the wavelet features and registration optimization portions of the algorithm of the algorithm were coded in C, the shearlet features portion of the algorithm was coded in MATLAB, because it utilized the FFST MATLAB library¹⁷ to generate features. This limited its use for remote sensing scientists, since it required MATLAB to perform. Speed also became a significant problem, because MATLAB does not enjoy the benefits of a compiled language. The goal was to provide an automatic, robust, and portable registration algorithm to remote sensing scientists. Thus, the MATLAB construction of the shearlets needed to be translated to C.

To do so, new shearlet code was written, based on the existing shearlet features code from the prototype and the libraries of the FFST.¹⁷ During this process, new efficiencies with regard to memory storage were discovered, improving the existing prototype algorithm. Moreover, the optimization procedure, which was designed for the decimated wavelet transform, was modified for the redundant shearlet transform. This further improved the numerical performance of both the shearlet-only and hybrid shearlet-wavelet algorithm.

In addition to presenting this fully-integrated C algorithm, the present article also considers ways in which the order of the wavelet and shearlet decompositions may be permuted, to perhaps acquire better registration results. The prototype hybrid algorithm¹⁶ registered first with shearlets, then with wavelets. Intuitively, this exploits the high degree of robustness of shearlets by acquiring first an approximate registration with a large radius of convergence, followed by a precision adjustment from the wavelet registration. However, it was of interest to consider more general orders of the wavelets and shearlets, so as to produce a more agile registration algorithm. Some initial results in this direction are also presented in this article.

In particular, we consider a registration algorithm in which the coarsest-scale shearlet feature is decomposed by a decimating wavelet transform. Intuitively, this makes the first step of the registration algorithm a matching with very low-pass features, since the already low-pass shearlet feature is being decomposed further in an isotropic low-pass filter. This technique is referred to as *hybrid shearlet-wavelet with decomposition*. It is a novelty over the methods explored previously,^{15,16} in which shearlet features did not mix with wavelet features explicitly. We summarize this algorithm:

1. Input a reference image, I^r , and an input image I^i . These will be the images to be registered.
2. Input an initial registration guess $(\theta_0, T_{x_0}, T_{y_0})$. This is sometimes set at $(\theta_0, T_{x_0}, T_{y_0}) = (0, 0, 0)$. This is rather arbitrary, as this algorithm is fully automatic and assumes no a priori knowledge of the images to be registered. If a priori knowledge is available, or if manual registration has been computed, this information can be input for the initial guess at this stage. In many experiments, we will vary the initial registration guess relative to the true registration in order to evaluate the robustness of the algorithm.
3. Apply shearlet features algorithm and wavelet features algorithms to I^r and I^i . This produces a set of shearlet features for both, denoted S_1^r, \dots, S_n^r and S_1^i, \dots, S_n^i , respectively, as well as a set of wavelet features for both, denoted W_1^r, \dots, W_n^r and W_1^i, \dots, W_n^i . Here n refers to the level of decomposition chosen. In general, n is bounded by the resolution of the images as

$$n \leq \lfloor \frac{1}{2} \log_2(\max\{M, N\}) \rfloor, \quad (1)$$

where I_r, I_i are $M \times N$ pixels. For example, for images of size 256×256 , $n \leq 4$. The bound (1) is determined by the elongated, anisotropic support of the shearlet functions at higher scales. In order for the support of the shearlet function used to compute the shearlet coefficients at the n^{th} level of decomposition to fit inside of an $M \times N$ image, it is necessary that (1) hold; see²⁰ for details. The order of the coefficients for

both wavelets and shearlets is from coarsest to finest, i.e. from the coefficients containing mostly global features to those containing mostly local features. This is because the coarse features should produce the most robust but least precise matching in general, and this guess will be iteratively refined by matching with increasingly fine scale coefficients.

4. Apply the wavelet features algorithm to S_1^r, S_1^i to acquire decompositions of these coarse shearlet features. These are denoted $S_{1,1}^r, \dots, S_{1,k}^r$ and $S_{1,1}^i, \dots, S_{1,k}^i$, respectively. Here, k denotes the number of scales used in this wavelet decomposition; for the present experiments, $k = 2$.
5. Match $S_{1,1}^r$ with $S_{1,1}^i$ with a least squares optimization algorithm and initial guess $(\theta_0, T_{x_0}, T_{y_0})$ to get a transformation $T_{1,1}^S$. More precisely, we solve

$$T_{1,1}^S = \arg \min_{T_p} \frac{1}{K} \sum_{j=1}^K (S_{1,1}^r(x_j, y_j) - S_{1,1}^i(T_p(x'_j, y'_j)))^2 \quad (2)$$

with a Marquadt-Levenberg optimization scheme. Here, the sum is over all K pixels in the features and T_p is the registration transformation, determined by parameters p . The parameter p could refer to the rotation, scale, and translations in an RST transformation, or to the coefficients in an affine transformation. Using $T_{1,1}^S$ as an initial guess, match $S_{1,2}^r$ with $S_{1,2}^i$ as in (2) to acquire a transformation $T_{1,2}^S$. Iterate this process by matching $S_{1,j}^r$ with $S_{1,j}^i$ using $T_{1,j-1}^S$ as an initial guess, for $j = 2, \dots, k$. At the end of this iterative matching, we acquire a *decomposed shearlet-based registration*, call it $T_1^S = (\theta^{S,1}, T_x^{S,1}, T_y^{S,1})$.

6. Using T_1^S as an initial guess, match S_2^r with S_2^i as in (2) to acquire a transformation T_2^S . Using T_2^S as an initial guess, match S_3^r with S_3^i as in (2) to acquire a transformation T_3^S . Iterate this process by matching S_j^r with S_j^i using T_{j-1}^S as an initial guess, for $j = 2, \dots, n$. At the end of this iterative matching, we acquire a *full shearlet-based registration*, call it $T^S = (\theta^S, T_x^S, T_y^S)$.
7. Using T^S as an initial guess, match W_1^r with W_1^i as in (2) to acquire a transformation T_1^W . Using T_1^W as an initial guess, match W_2^r with W_2^i as in (2) 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 $j = 2, \dots, n$. At the end of this iterative matching, we acquire the *final hybrid registration*, call it $T^H = (\theta^H, T_x^H, T_y^H)$.
8. Output T^H .

As in the prototype algorithm, three wavelet implementations are considered: Simoncelli band-pass filters, Simoncelli low-pass filters, and spline wavelet filters. These filters do not implement a frame, but rather filter the images in an isotropic and decimating way. The shearlets algorithm we implement has the benefit of being a redundant frame, which is useful in extracting meaningful features from the images.

3. EXPERIMENTS AND RESULTS

We consider two sets of images that were previously analyzed with the prototype hybrid-shearlet registration algorithm.¹⁶ We consider synthetic experiments consisting of radiometrically warped images of Mossy Rock, in the Mount St. Helens region of WA, USA, and real lidar and optical images of a rural scene in WA, USA. Experiments with nine algorithms were considered: wavelets-only, shearlets+wavelets, and shearlets+wavelets with decomposition. To test robustness of the algorithm, many different choices of starting registration value $(\theta_0, T_{x_0}, T_{y_0})$ were considered; to ease computation, these starting values were parametrized as RT , where $RT = \alpha$ means the initial guess was $(\theta_0, T_{x_0}, T_{y_0}) = (\alpha, \alpha, \alpha)$. The algorithm also registers images at different scales, but the present images are at the same scale.

3.1 Mossy Rock Synthetic

The original image is a 512×512 shaded relief lidar image captured in 2002, generated from an airborne laser swath mapping conducted by Terrapoint, LLC under contract with the USGS. This image was synthetically altered via convolution with a point spread function (PSF). The PSF is implemented by the 512×512 matrix M , given by:

$$M(i, j) = \begin{cases} 1, & 254 \leq i, j \leq 258. \\ 0, & \text{else.} \end{cases} \quad (3)$$

This matrix is then convolved with the reference image to generate an input image that simulates a radiometrically varied image of the same scene. This can be considered as a simulation of the challenges of multimodal registration: many of the same features appear in the images, but not all, and the common features are often rendered differently. The original image of Mossy Rock, together with the radiometrically warped version, appear in Figure 1.

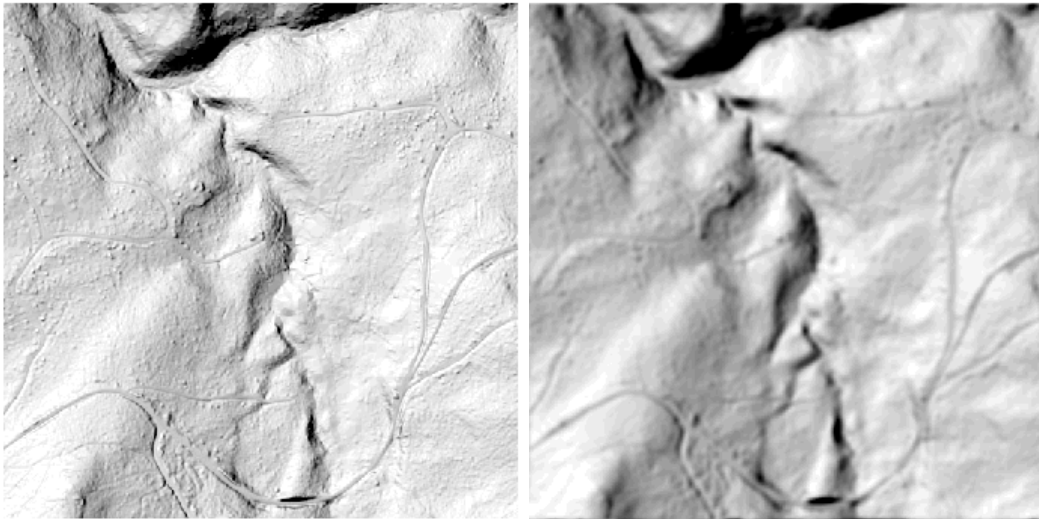


Figure 1: 512×512 lidar shaded relief images of Mossy Rock without (left) and with (right) synthetic radiometric distortion. The images have been converted to grayscale.

These experiments illustrate the value of the new optimization scheme for registering shearlet features. Compared to the results for these images already analyzed,¹⁶ the radius of convergence for the shearlets+wavelets algorithm is larger. These images also display a relatively straightforward pattern for how each algorithm performs relatively, as can be seen in Figure 2 and Table 1. In particular, nothing is gained over applying the decomposition to the low-pass shearlet features, though the shearlet+wavelets with decomposition still outperforms wavelets-only.

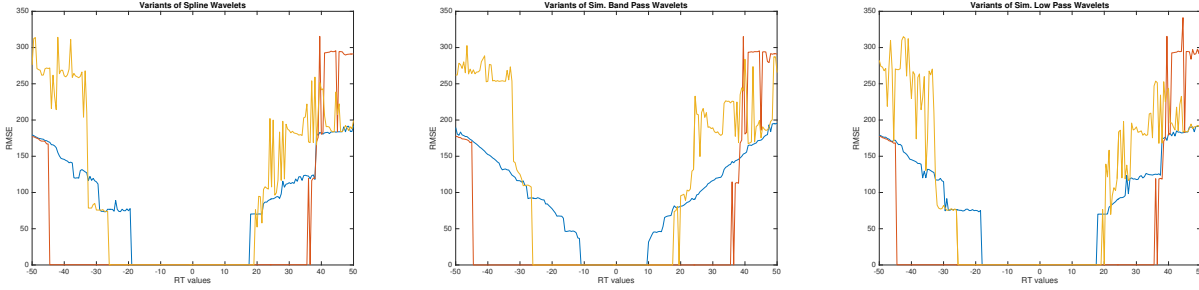


Figure 2: Comparison of algorithms for Mossy Rock synthetically warped experiments (from left to right: splines, Simoncelli band-pass, Simoncelli low-pass); blue is wavelets, yellow is hybrid shearlets+wavelets with decomposition, and red is shearlets+wavelets without decomposition.

Registration Technique	Number of Converged Experiments (out of 200)	Percentage of Converged Experiments	Mean RMSE	Relative Improvement
Spline Wavelets	74	36.82%	.3591	-
Simoncelli Band-Pass	42	10.50%	.0074	-
Simoncelli Low-Pass	72	55.50%	.2412	-
Shearlets	162	80.60%	.0748	-
Shearlet+Spline Wavelets	162	80.60%	.3287	118.92%
Shearlet+ Simoncelli Band-Pass	162	80.60%	.0074	285.71%
Shearlet + Simoncelli Low-Pass	162	80.60%	.2432	125.00%
Shearlet+Spline Wavelets Decomp	91	45.27%	.2404	22.97%
Shearlet+ Simoncelli Band-Pass Decomp	89	44.28%	.0069	111.90%
Shearlet + Simoncelli Low-Pass Decomp	91	45.27%	.2202	26.39%

Table 1: Comparison of registration algorithms for Mossy Rock synthetically warped experiments.

3.2 Lidar-to-optical multimodal experiments

We also considered experiments with registering a lidar image to an optical image, a problem known to be quite challenging. These images are shown in Figure 3. The lidar data was acquired in 2003 by Terrapoint, Inc., under contract to NASA, using a multi-return airborne laser swath mapping (ALSM) instrument. The optical data is a natural color aerial photograph, presented as a grey-scale image, obtained by the Google Earth database from the United States Geological Survey in 2006. These data sources have related, but quite different information content. A lidar image, commonly referred to as a digital elevation model (DEM), is a measure of the elevation of the components making up the surface. The data we used is a highest surface DEM. This represents vegetation canopy tops where vegetated, and ground, roads, and building tops where not vegetated. On the other hand, an

optical image records solar radiance reflected from the surface. The latter is a function of the reflectance of the surface components and their three-dimensional organization. Together, these define the patterns and brightness of illuminated and shadowed patches seen in optical images.

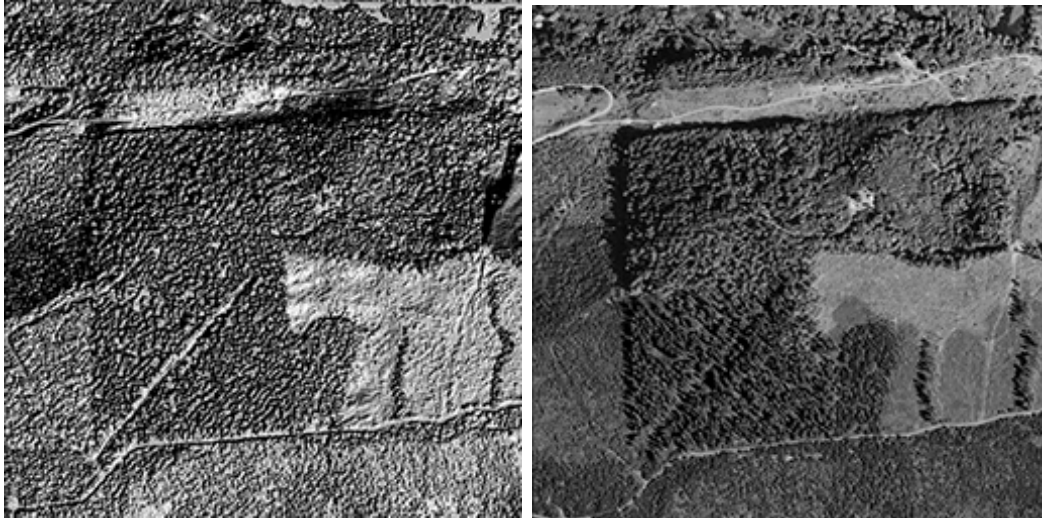


Figure 3: Lidar ALSM elevation image (top left), the derived shaded relief image (top right), and aerial photograph (bottom) for a scene in WA state. The shaded relief image, illuminated in the same direction as in the optical image, depicts similar patterns of textures and edges. All images are 256×256 . The images have been converted to grayscale.

The results of our algorithm in this case are substantially more mixed than in the synthetic example. While in general the shearlet-wavelet hybrid algorithm outperforms wavelets-only, there are some cases in which it does not. In particular, the shearlet-wavelets plots in Figure 4 are not monotonic. This indicates the complicated image changing the initial value has for the shearlets features. Moreover, the decomposition of the shearlet low-pass features appears to hurt the registration algorithm greatly, as performance for shearlet+wavelet with decomposition is clearly the worst; see Table 2. Indeed, it shows the erraticism of the shearlets-wavelets algorithm, without its larger radius of convergence. This suggests that the shearlet features are perhaps unhelpful for this example, and that the wavelet features do most of the registration work in the shearlets+wavelets algorithm. This is corroborated by earlier findings.¹⁶

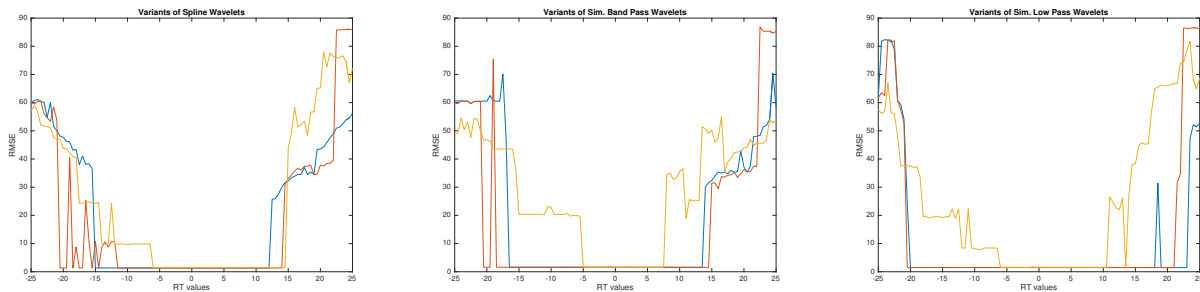


Figure 4: Comparison of algorithms for WA lidar-to-optical experiments (from left to right: splines, Simoncelli band-pass, Simoncelli low-pass); blue is wavelets, yellow is hybrid shearlets+wavelets with decomposition, and red is shearlets+wavelets without decomposition.

Registration Technique	Number of Converged Experiments (out of 200)	Percentage of Converged Experiments	Mean RMSE	Relative Improvement
Spline Wavelets	55	54.46%	1.3446	-
Simoncelli Band-Pass	61	60.40%	1.5862	-
Simoncelli Low-Pass	86	85.15%	1.4868	-
Shearlets	68	67.33%	6.1300	-
Shearlet+Spline Wavelets	60	59.41%	1.3143	9.09%
Shearlet+Simoncelli Band-Pass	70	69.31%	1.5831	14.75%
Shearlet + Simoncelli Low-Pass	84	83.17%	1.4934	-2.33%
Shearlet+Spline Wavelets Decomp	42	41.58%	1.4411	-23.64%
Shearlet+Simoncelli Band-Pass Decomp	26	25.74%	1.5915	-57.38%
Shearlet + Simoncelli Low-Pass Decomp	35	34.65%	1.5860	-48.53%

Table 2: Comparison of registration algorithms for lidar-to-optical.

4. CONCLUSIONS AND SUMMARY

The experiments confirm the effectiveness of coupling shearlets with wavelets for image registration. On both synthetic and real images, the hybrid algorithm outperforms using wavelets alone. The effect of applying a wavelet transform to the low-pass shearlet component appears, however, uncertain. In some cases, it appears to perform better than wavelets alone, but it never performs as well as the original shearlet-wavelet hybrid algorithm. Indeed, in some cases, it does quite poorly, even when compared to wavelets alone. One possible reason is that we are essentially applying a low-pass filter to a feature that is already low-pass. This would degrade the low-pass information even more, losing an unacceptably large amount of information. The impact of applying the wavelet features algorithm on high pass shearlet features remains of interest.

5. ACKNOWLEDGEMENTS

The first and second authors would like to thank the NASA Goddard Space Flight Center for generously funding the summer internships during which the present research transpired. The first named author would also like to thank Mauro Maggioni of Duke University for his support during the writing of this article. The authors would also like to thank David J. Harding of the NASA Goddard Space Flight Center for supplying the data of Washington state.

REFERENCES

1. Brown, L. G., "A survey of image registration techniques," *ACM computing surveys* **24**(4), 325–376 (1992).

2. Maintz, J. B. and Viergever, M. A., “A survey of medical image registration,” *Medical image analysis* **2**(1), 1–36 (1998).
3. Le Moigne, J., Netanyahu, N. S., and Eastman, R. D., eds., [*Image registration for remote sensing*], Cambridge University Press (2011).
4. Lowe, D. G., “Object recognition from local scale-invariant features,” in [*The proceedings of the seventh IEEE international conference on computer vision*], **2** (1999).
5. Gonçalves, H., Corte-Real, L., and Gonçalves, J. A., “Automatic image registration through image segmentation and SIFT,” *IEEE Transactions on Geoscience and Remote Sensing* **49**(7), 2589–2600 (2011).
6. Daubechies, I., [*Ten lectures on wavelets*], Society for industrial and applied mathematics (1992).
7. Zavorin, I. and Le Moigne, J., “Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery,” *IEEE Transactions on Image Processing* **14**(6), 770–782 (2005).
8. Guo, K. and Labate, D., “Optimally sparse multidimensional representation using shearlets,” *SIAM journal on mathematical analysis* **39**(1), 298–318 (2007).
9. Candès, E. and Donoho, D. L., “New tight frames of curvelets and optimal representations of objects with piecewise C^2 singularities,” *Communications on pure and applied mathematics* **57**(2), 219–266 (2004).
10. Dahlke, S., Mari, F. D., Grohs, P., and Labate, D., [*From Group Representations to Signal Analysis In Harmonic and Applied Analysis*], Springer International Publishing (2015).
11. Do, M. and Vetterli, M., “Contourlets: a directional multiresolution image representation,” in [*Proceedings of 2002 IEEE International Conference on Image Processing*], (2002).
12. Grafakos, L. and Sansing, C., “Gabor frames and directional time-frequency analysis,” *Applied and Computational Harmonic Analysis* **25**(1), 47–67 (2008).
13. Czaja, W., Manning, B., Murphy, J., and Stubbs, K., “Discrete directional Gabor frames,” *arXiv preprint arXiv:1602.04336* (2016).
14. Labate, D., Lim, W. Q., Kutinyok, G., and Weiss, G., “Sparse multidimensional representation using shearlets,” in [*Proceedings of International Society for Optics and Phototonics: Optics and Phototonics*], (2005).
15. Murphy, J. M. and Moigne, J. L., “Shearlet features for registration of remotely sensed multimodal images,” in [*Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*], (2015).
16. Murphy, J. M., Moigne, J. L., and Harding, D. J., “Automatic image registration of remotely sensed data with global shearlet features,” *IEEE Transactions on Geoscience and Remote Sensing* **54**(3) (2016).
17. Häuser, S. and Steidl, G., “Fast finite shearlet transform,” *arXiv preprint arXiv:1202.1773* (2012).
18. Kutyniok, G., Shahrnam, M., and Zhuang, X., “Shearlab: A rational design of a digital parabolic scaling algorithm,” *SIAM Journal on Imaging Sciences* **5**(4), 1291–1332. (2012).
19. King, E., Kutyniok, G., and Zhuang, X., “Analysis of inpainting via clustered sparsity and microlocal analysis,” *Journal of mathematical imaging and vision* **48**(2), 205–234 (2014).
20. Häuser, S., “Fast finite shearlet transform,” *arXiv preprint arXiv:1202.1773* (2012).