

# Accelerating Feature Based Registration Using the Johnson-Lindenstrauss Lemma<sup>\*</sup>

Ayelet Akselrod-Ballin<sup>1</sup>, Davi Bock<sup>2</sup>, R. Clay Reid<sup>2</sup>, and Simon K. Warfield<sup>1</sup>

<sup>1</sup> Computational Radiology Laboratory, Children's Hospital, Harvard Medical School, Boston, USA

<sup>2</sup> Department of Neurobiology, Harvard Medical School, USA

**Abstract.** We introduce an efficient search strategy to substantially accelerate feature based registration. Previous feature based registration algorithms often use truncated search strategies in order to achieve small computation times. Our new accelerated search strategy is based on the realization that the search for corresponding features can be dramatically accelerated by utilizing Johnson-Lindenstrauss dimension reduction. Order of magnitude calculations for the search strategy we propose here indicate that the algorithm proposed is more than a million times faster than previously utilized naive search strategies, and this advantage in speed is directly translated into an advantage in accuracy as the fast speed enables more comparisons to be made in the same amount of time. We describe the accelerated scheme together with a full complexity analysis. The registration algorithm was applied to large transmission electron microscopy (TEM) images of neural ultrastructure. Our experiments demonstrate that our algorithm enables alignment of TEM images with increased accuracy and efficiency compared to previous algorithms.

## 1 Introduction

Image registration is a fundamental process in medical imaging applications aimed at establishing spatial correspondences between images [1]. Registration algorithms must satisfy demanding requirements of speed, robustness and accuracy depending on the specific application involved. Alignment of Electron Microscopy (EM) images of neural tissue involves consideration of the high resolution and overwhelming size of the data, the large amount of detail, the acquisition artifacts and the deformation induced by the intrinsic deformation of the slices [2]. Thus, the reconstruction of neural circuitry from EM remains a substantial and challenging problem [3].

Existing algorithms for registration can be broadly classified into approaches that directly operate on image intensities or to feature based alignments seeking to identify features that should be aligned and an optimal transformation

---

<sup>\*</sup> This investigation was supported in part by a research grant from CIMIT, grants RG 3478A2/2 and RG 4032A1/1 from the NMSS, and by NIH grants R03 CA126466, R01 RR021885, R01 GM074068, R01 EB008015 and P30 HD018655.

that brings them into alignment [1]. Similar to previous registration methods for alignment of microscopy images our method relies on matching local image patches across successive slices. A recent study [2] addressed matching as part of a complete algorithm for assembling 3D volumes from EM data. This approach first identifies feature descriptors based on a gradient vector pyramid and then exploits these features to match adjacent slices. An approach for 3D reconstruction based on a block matching strategy was presented by [4] where the local displacements were utilized to robustly estimate a global transformation. [5] developed a solution for 3D reconstruction of a series of EM images based on the finite support properties of the cubic B-splines, where the initial estimate for the affine registration was based on the technique described in [4].

In this work we presents a novel algorithm designed for alignment of large cross-sectional slices of EM by introducing a search strategy which to our knowledge has not been used for image registration before. The key advance we propose is based on the realization that the search for corresponding features can be dramatically accelerated by utilizing Johnson-Lindenstrauss (JL) [6] dimension reduction. This enables us to consider more regions in the images and to consider all of the potential correspondences, ensuring that the search does not overlook the correct set of correspondences, and thus dramatically improves the robustness and accuracy of the registration. Combining this together with recent advances in approximate nearest neighbor (ANN) techniques [7] for the search strategy reduces the computational complexity significantly.

The paper is organized as follows. Section 2 describes the algorithm, section 3 presents experimental results and conclusions are presented in section 4.

## 2 Methods

The volume reconstruction is obtained by composing pairwise 2D alignments of consecutive slices by taking as reference the middle of the stack. Thus we focus on the 2D registration of successive sections, although the algorithm can be used for 3D registration as well. The input includes the fixed scene image  $I_S$  and a moving model image  $I_M$ . Our aim is to find the transformation  $T$  aligning the scene with the model. The algorithm consists of three main steps. **In the first step**, we extract image patches, also called blocks, fragments or templates [4], as the features for matching. The patches are rectangular sub-images of  $d = 10^4$  voxels (i.e.  $100 \times 100$  sized image regions) extracted from the scene and model images. Then using JL-embedding [6] with the random projections approach by Achlioptas [8] we generate a low dimensional representation of these patches. **In the second step** the patches of reduced dimension are compared across successive slices. So that, given a projected scene and model patch we search for correspondences based on the Euclidean distance between the patches. **Finally**, we construct two sets of points based on the center coordinates of the projected scene and model patches. We compute the transformation with an extension to the Expectation Maximization Iterative Closest Point (EM-ICP) [9] algorithm which includes the similarity measure of the patches. Table 1 presents an outline of the algorithm proposed.

**Table 1.** Outline of the algorithm

---

Given a pair of successive slices, a fixed (scene) and moving (model) image:

**Step 1: JL dimension reduction of features:** State of the art feature matching approaches to registration commonly perform feature detection by correlation, followed by a truncated search and a transform estimation. In contrast with previous methods we first extract small patches at different locations for both images and then the features are projected on to a lower dimension based on the JL-Lemma.

**Step 2: Search for correspondences:** The projected features extracted from both images are compared based on the correlation measure or Euclidean distance.

-2a. Brute Force search strategy.

-2b. Approximate nearest neighbor (ANN) search where each near neighbor is reported with a certain probability. For this step there are several solutions such as tree based search or locality sensitive hashing [10].

**Step 3: Transformation estimation:** Apply extended EM-ICP [9] approach based on the corresponding features between the images.

---

## 2.1 Step 1: Johnson-Lindenstrauss (JL) Dimension Reduction

The essential property enabling accelerated search has been the realization that efficient high dimensional search can be achieved by creating randomized projections into low-dimensional spaces, and then using efficient low-dimensional search. The JL Lemma [6] asserts that any set of  $n$  points in  $d$ -dimensional Euclidean space can be projected down to  $k$ -dimensional Euclidean space, where  $k = O(\epsilon^{-2} \log n)$ , while maintaining pairwise distances with a low distortion. Recent research [8] has shown that random projection matrices can be used for JL projection. Following this proof, given the initial set of  $n$  points in  $R^d$ , represented as an  $n \times d$  matrix, where each feature-patch is represented by a row, let  $R$  be a  $d \times k$  random matrix with  $R(i, j) = r_{ij}$ ; where the independent random variables  $r_{ij}$  are:  $\{1$  with probability 0.5, and  $-1$  with probability 0.5  $\}$ . Naively, the random projection can be performed by constructing a  $k \times d$  random matrix; so that mapping each point takes  $O(dk)$ , however recent theoretic work suggests that a projection from dimension  $d$  to dimension  $k$  can be computed with  $O(d)$  operations [11].

## 2.2 Step 2: Search for Correspondences

The transformation computation is based on maximizing the similarity of corresponding patches. Currently we utilize the brute-force approach which involves computing the distances between the patch and all the patches in the neighboring image. Our experiments compare the results of using the full-size patches and their projections. We use normalized correlation (NC) as a similarity measure since NC is invariant to linear intensity transformation and it is assumed that for small corresponding image patches across the two successive slices, the intensities are locally related by some linear intensity transformation [4]. Another

attractive property of NC is that it is equivalent to a squared Euclidean distance [12] which meets the requirements of the JL Lemma.

Recently it has been discovered that allowing approximate rather than exact search enables dramatically accelerated search [7]. In the approximate nearest neighbor (ANN) formulation, given a query point  $q$  and a specified constant  $c > 0$ , the algorithm returns a point whose distance from  $q$  is at most a  $c = (1 + \epsilon)$ -factor larger from the distance of the nearest point  $p$  in  $P$  from  $q$ . This approach results in efficient algorithms which are based on data structures such as tree based search (spill trees) and locality sensitive hashing (LSH)[10]. Thus instead of performing the naive search described above, additional significant acceleration can be obtained by casting the registration problem as an ANN search problem.

### 2.3 Step 3: Transform Estimation

Given the set of patches represented by their center coordinates this step determines the transformation that matches a set of model and scene points. The approach exploits the Expectation Maximization (EM) scheme to optimize simultaneously for correspondences and the registration transformation. It extends the EM-ICP approach to go from point matching to patch matching. Let  $s_i$  be the points of the scene set  $S \in R^2$  and  $m_i$  the points of the model patch set  $M \in R^2$ , with  $n_s$  and  $n_m$  determining the number of points respectively.  $T$  represents the rigid transformation from the scene to the model. The probability of a point  $s_i$  to correspond to the model points  $m_i$  is modeled by a Gaussian probability distribution. In the case of homogeneous isotropic Gaussian noise the probability is modeled by

$$p(s_i|m_j, T) = \exp(-\|T * s_i - m_j\|^2/2\sigma^2) \quad (1)$$

where  $\sigma$  represents the noise in the measurement.

The idea is to maximize  $\log p(S, A|M, T)$  the log-likelihood of the data distribution where the unknown correspondences  $A \in R^{n_s \times n_m}$  are considered as a hidden random variables. The algorithm starts by initialization of the transformation ( $T$ ), and repeats until convergence of the two EM steps. In the E-step,  $T$  is fixed and the probability of matches  $\overline{(A_T)_{ij}}$  are computed as follows (for more details see [9])

$$\overline{(A_T)_{ij}} = \frac{\overline{\pi_{ij}} \exp(-\|T * s_i - m_j\|^2/2\sigma^2)}{\sum_k \overline{\pi_{ik}} \exp(-\|T * s_i - m_k\|^2/2\sigma^2)} \quad (2)$$

In the classic EM-ICP, the prior probability of the matches is based on the uniform law:  $\pi_{ij} = \frac{1}{n_m}$ . In contrast, the EM-ICP-NC used here extends the prior to account for the NC similarity of the patches taking advantage of the similarity measure in addition to the spatial information. Hence, the prior is based on the normalized NC:

$$\overline{\pi_{ij}} = \frac{NC(p_i, p_j)}{\sum_k NC(p_i, p_k)} \quad (3)$$

The NC-based prior  $\overline{\pi_{ij}}$  is also used in the initialization step to determine the initial transformation  $T$ . In the M-step,  $A$  is fixed and the likelihood is optimized w.r.t to  $T$ .

$$T_{n+1} = \arg \max_T (E_A[\log P(S, A|M, T)]) \quad (4)$$

Thus, by ignoring constant terms the criterion optimized by the EM-ICP-NC yields

$$C_{AT}(T) = \sum_i^{n_s} \sum_j^{n_m} \overline{(AT)_{ij}} \log p(s_i|m_j, T) \quad (5)$$

## 2.4 Computational Complexity

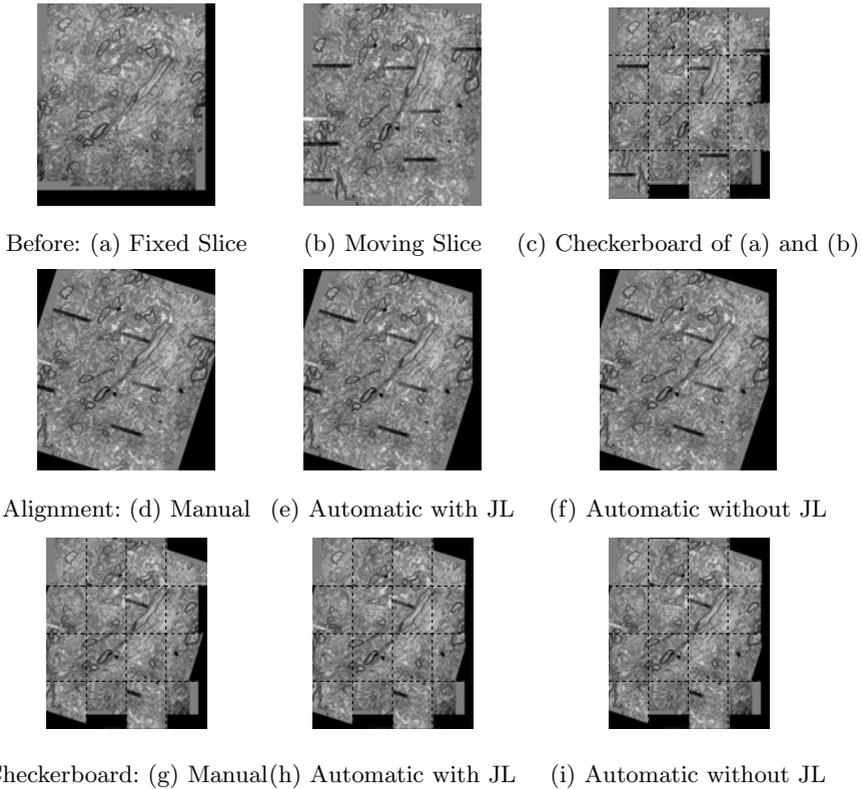
The computational complexity of the entire process is determined by the number of images ( $T \approx 100$ ) in the data base, the image size  $n$  (we assume that the number of patterns per slice is proportional to  $n$ ) and the feature dimension  $d$ . Our preliminary results were performed on slice sections of  $n = 10^8 = 10^4 \times 10^4$ , which were downsampled using Gaussian smoothing plus bi-linear interpolation to a size of  $n = 10^6 = 10^3 \times 10^3$ . Improved TEM techniques utilizing multiple camera arrays have lead to datasets of  $n = 10^{10} = 10^5 \times 10^5$ . The size of a typical local image patch is  $d = 10^4 = (100 \times 100)$ , by employing JL lemma the initial  $n \times d$  matrix is projected to an  $n \times k$  feature matrix (where  $k = 100$  in our experiments). We focus on the search for correspondences since it is the most expensive operation of the algorithm. The naive searching method involves comparing each projected patch to all the projected patches in the neighboring image. The complexity of the naive search time per query is  $O(dn)$ . Thus, the overall complexity for using the full  $d$ -dimensional patches is  $O(Tn^2d)$ . Hence, projection to lower dimension alone reduces the complexity to  $O(Tn^2 \log n)$  saving  $O(d - k)$  operations ( $10^4$  operations per slice in our experiments).

Additional significant acceleration can be obtained by employing recent advances in approximation algorithms for performing ANN. The authors in [7] have shown that the time required for retrieving similar features is reduced to  $O(dn^{1/c^2})$  where for  $c = 2$  this becomes  $O(dn^{1/4})$ . Consequently, the proportion between the naive and accelerated approaches is  $O(n^{1-1/c^2})$  and for  $c = 2$  this becomes  $O(n^{3/4})$  which in the typical conditions of the data we are dealing with becomes  $O(10^{7.5})$ . Thus we claim that the proposed approach to registration is more than a million times faster than using naive approaches. We also note that the query time quoted assumes the need to project from high dimension to low dimension before doing the query. However, since our patches participate as both search points and query points, we can benefit from pre-computing the projection of all the points ahead of the query, which reduces the query cost by  $O(d)$ . Recall  $d$  is  $O(10^4)$  in our case, so this is a substantial improvement.

## 3 Experiments and Results

The algorithm was tested on a series of EM images of the lateral geniculate nucleus of a ferret. Each image is about  $10000 \times 10000$  pixels large with a pixel

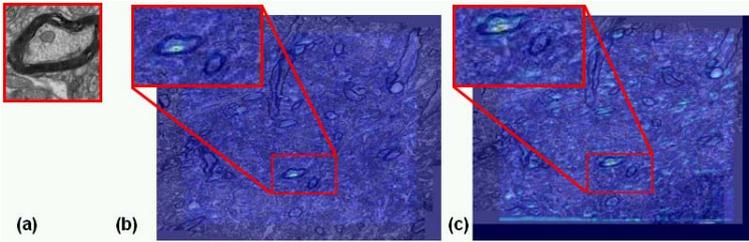
resolution of  $3nm$  and a slice thickness of  $60nm$ . We evaluated our experiments on 114 slices which were manually preprocessed to correct or exclude images with significant artifacts or severe rotations. The approach was validated by comparison to a manual ground truth registration. The manual registration denoted as ( $T^*$ ), was performed by manually selecting corresponding points in a pair of consecutive images and computing the pairwise transformation based on Horn's method [13]. Given a set of corresponding points in two systems, Horn's method finds the closed form solution to the "least square" problem relating these points. We performed two sets of experiments. The first set of experiments was obtained with the JL projection of features while the second set was obtained without the projection. Both experiments compare the results obtained by three automatic approaches including our EM-ICP-NC approach (see Sec. 2.3), the classic EM-ICP [9] and Horn's [13] approach. The  $L_2$  difference between the transformations



**Fig. 1.** Successive pair of slices before and after alignment. (a) and (b) present the fixed and moving image before alignment and (c) shows their checkerboard composite. (d),(e),(f) show the alignments results of the manual, and automatic algorithms with JL projection and without projection respectively. (g),(h),(i) demonstrate the checkerboard of the fixed image in (a) and the aligned moving image in (d), (e) and (f) respectively.

**Table 2.** Comparing our results to the manual transformation and to two other automatic approaches. Our algorithm has higher accuracy than previously described methods. The advantage is obtained both with and without the JL projection.

Dimension	EM-ICP-NC	EM-ICP	Horn
$k$ (with JL projection)	$5.41 \pm 3.32$	$45.024 \pm 28.56$	$93.63 \pm 56.9$
$d$ (without JL projection)	$4.93 \pm 2.94$	$20.55 \pm 26.74$	$66.86 \pm 53.6$



**Fig. 2.** Effectiveness of template matching for identifying correspondences. The fusion of the correlation map computed for the patch (a) with the original slice, from which the patch was extracted (b) and with the successive slice (c). The maxima regions in both slices (b,c) respectively, are highlighted in red rectangles and enlarged. The red maxima values in the fusion image correspond to higher correlation values, showing that the features are a sharp local maxima of the correlation function.

of all three automatic approaches in both sets of experiments are compared to the ground truth transformation computed ( $T^*$ ). Figure 1 shows the results obtained with and without the JL projection compared to the manual registration. Table 2 presents the results of both experiments. Our EM-ICP-NC automatic algorithm has a higher accuracy than previously described methods. The advantage is achieved in both set of experiments with and without the JL projection, with similar results. Thus, we conclude that the projection to low dimensional space did not reduce the ability to recognize corresponding patches.

To evaluate the effectiveness of template matching for identifying correspondences in various cases of neurobiological objects, we tested patches with different features, such as myelinated white matter, dendrites, synapses, and microtubules. The NC function was computed between the features visible in the image, with patches within the same image and with patches in a successive image. Figure 2 illustrates the results, demonstrating that the local maxima in the NC function within the same image is detected in the exact same spot, while in the next slice the detection is in the area of similar anatomy.

## 4 Summary

We present a novel algorithm for alignment of large EM images. To our knowledge this is the first attempt to provide the registration algorithm a low

dimensional representation of the data by utilizing the JL embedding and to demonstrate the dramatic speed up of the identification of correspondence which at the same time guarantees the robustness and accuracy of the alignment. In this work we have demonstrated these contributions on real EM data and also constructed a full scheme with significant computational savings. We assessed the impact of dimensionality reduction on the transformation accuracy and shown for the first time the feasibility and effectiveness of this approach, showing that the projected features are as effective for registration as the full-dimensionality features. Future work will evaluate alternative search strategies for the acceleration of accurate correspondence estimation. This will include evaluation of data structures that support ANN search strategies, such as LSH, and spill-trees. We will explore faster variants of algorithms to project into  $k$ -dimensional space.

## References

1. Maintz, J.B.A., Viergever, M.A.: A survey of medical image registration. *Medical image analysis* 2(1), 1–16 (1998)
2. Koshevov, P., Tasdizen, T., Whitaker, R., Jones, B., Marc, R.: Assembly of large three-dimensional volumes from serial-section TEM. In: MICCAI (2006)
3. Smith, S.J.: Circuit reconstruction tools today. *Curr. Opin. Neuro.* 17(5) (2007)
4. Ourselin, S., Roche, A., Subsol, G., Pennec, X., Ayache, N.: Reconstructing a 3D structure from serial histological sections. *IVC* 19, 25–31 (2000)
5. Dauguet, J., Bock, D., Reid, C.R., Warfield, S.K.: Alignment of large image series using cubic b-splines tessellation: Application to TEM data. In: MICCAI (2006)
6. Johnson, W., Lindenstrauss, J.: Extensions of lipschitz mapping into a hilbert space. *Contemp. Math.* 26, 189–206 (1984)
7. Indyk, P., Motwani, R.: Approximate nearest neighbors: towards removing the curse of dimensionality. In: STOC 1998, pp. 604–613 (1998)
8. Achlioptas, D.: Database-friendly random projections: johnson-lindenstrauss with binary coins. *Journal of Comp. and Sys. Sci.* 66, 671–687 (2003)
9. Granger, S., Pennec, X.: Multi-scale EM-ICP: A fast and robust approach for surface registration. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2353, pp. 418–432. Springer, Heidelberg (2002)
10. Andoni, A., Indyk, P.:  $e^2$ lsh: Exact euclidean locality-sensitive hashing. Implementation (2004), <http://web.mit.edu/andoni/www/LSH/index.html>
11. Liberty, E., Ailon, N., Singer, A.: Dense fast random projections and lean walsh transforms. LNCS, pp. 512–522 (2008)
12. Arya, K., Gupta, P., Kalra, P., Mitra, P.: Image registration using robust m estimators. *Pattern Recognition Letters* 28, 1957–1968 (2007)
13. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America* 4, 629–642 (1987)