DRAFT COPY PLEASE DO NOT CITE

Perturbation Search for Nonlinear Shape Registration

Robo G. Grey
roboggrey@glys.com

Ham H. Let
hamhlet@glys.com

HAM-LB-TR-13-01

December 2013

DRAFT

Ham Laboratories
H.L. Ham Pte Ltd
Hamburg, Hamerica 03977

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

Abstract

Shape registration is a longstanding subfield in image processing that involves aligning an *observation* image to a known *template*. Kato et al. [1] have recently proposed a novel correspondence-free nonlinear shape registration (NLSR) framework expressing the registration problem as that of finding an approximate solution to a system of low-order polynomial ω equations, which is both generic and quick. However, while generally robust, NLSR does not always handle registrations reliably, particularly where the choice of ω functions happens not to suit the particular template-observation pair. We propose a simple and effective tunable procedure to mitigate this issue, and demonstrate its utility on various datasets.

DRAFT

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

1 Introduction

In [1], an unrestricted featureless registration approach is suggested, whereby a sufficiently-large nonlinear system of equations is integrated over the image domain, and solved by Levenberg-Marquardt (LM) [2]. In general, the target of registration is to obtain the parameters of some arbitrary $\varphi : \mathbb{R}^2 \to \mathbb{R}^2$ diffeomorphism. In the absence of predefined landmarks, area-based methods that consider all (normalised and centered) shape pixels equally as such are popular [3, 4, 5, 6]. The approach taken in [1] is to introduce a set of functions $\{\omega : \mathbb{R}^2 \to \mathbb{R}\}$, such that

$$\omega(\mathbf{y}) = \omega(\varphi(\mathbf{x})) \Leftrightarrow \omega(\mathbf{x}) = \omega(\varphi^{-1}(\mathbf{y})) \tag{1}$$

where $\mathbf{x}$ is the template, $\mathbf{y}$ is the observation, and φ^{-1} is guaranteed to exist – intuitively, the template and observation can always be swapped. From the chosen set of functions,

$$\int_{F_o} \omega_i(\mathbf{y}) d\mathbf{y} = \int_{F_i} \omega_i(\varphi(\mathbf{x})) |J_\varphi(\mathbf{x})| d\mathbf{x}, i = 1, \ldots, l \tag{2}$$

where F denotes foreground pixels and $l > k$, with k being the number of parameters in φ (e.g. $k = 9$ under affine transformation assumptions, as is the default in the provided demonstration code[1]). Various ω_i sets were tried by the authors, who concluded that registration quality is almost entirely unaffected by the choice of function, as long as they are rich enough. Therefore, low-order polynomials were selected for their computational tractability. They further show that the Levenberg-Marquardt solver can be adapted to run in time independent of the number of pixels from the polynomial property of the basis functions.

2 Observations

It can be noted that the NLSR framework, while largely robust, remains approximative. There is no theoretical assurance that ω will work for any particular $\{\mathbf{x}, \mathbf{y}\}$ pair, or that the possibly overdetermined set of integrals has an exact solution. In fact, it is probable that registration accuracy can be improved with diminishing returns by increasing l and/or the number of LM iterations. However, there remain occassional glaring failures, that we suspect can be avoided by perturbing either ω_i or the original template domain. In practice, we employ the latter approach due to potential complications in parameter perturbation [7].

Essentially, registration failure can be easily detected post-LM from the absolute similarity $\delta = \frac{|F_r \cup F_o|}{|F_r| + |F_o|}$ between the template and registered observation. if δ lies below a certain threshold, a perturbation search (PS) over the transformed template can be attempted, i.e. we introduce a set of γ_j functions to Equation 2:

$$\int_{F_o} \omega_i(\mathbf{y}) d\mathbf{y} = \int_{F_i} \omega_i(\gamma_j(\varphi(\mathbf{x}))) |J_\varphi(\mathbf{x})| d\mathbf{x}, i = 1, \ldots, l, j = 1, \ldots, m \tag{3}$$

[1] `http://www.inf.u-szeged.hu/%7Ekato/software/planarhombinregdemo.html`

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

Since we know the parameters of γ, the parameters of the original φ can be deduced if required, but in most applications their composition $\gamma(\varphi(.))$ should suffice. Again, there is no particular restriction on the properties of γ, but a good starting point would seem to be simple rotation, as parameterized by θ. A basic wrapper would then be:

Algorithm 1 NLSR wrapper algorithm for perturbation search

Input: Binary images of the *template* and *observation*
Output: k parameters of the estimated transformation $\hat{\varphi}$

Execute original NLSR algorithm
if $\delta > \delta_t$ **then**
  Return obtained parameters
else
  $j \leftarrow 0$
  $\delta_{best} \leftarrow 0$
  while $\delta_{best} <= \delta_t$ AND $j < m$ **do**
    $j \leftarrow j + 1$
    Generate γ_j-transformed template
    Execute NLSR on γ_j-transformed template
    if $\delta_{\gamma_j} > \delta_{best}$ **then**
      $\delta_{best} \leftarrow \delta_{\gamma_j}$
      Save current parameters as best known
    end if
  end while
  Return best known parameters
end if

DRAFT

Note that this formulation is not the same as incorporating rotation into the ω-functions, as suggested in [1]; for example, rotated power functions were proposed:

$$\omega_i(\mathbf{x}) = (x_1 cos\alpha_i - x_2 sin\alpha_i)^{n_i}(x_1 sin\alpha_i + x_2 cos\alpha_i)^{m_i} \tag{4}$$

with $\alpha_i \in \{0, \frac{\pi}{6}, \frac{\pi}{3}\}$ and $(n, m) \in \{(1, 2), (2, 1), (1, 3), (3, 1)\}$, for a total of 12 equations. Simply put, by introducing the "search" within ω, the LM minimization is on multiple distortions of the template simultaneously, but none in particular, and therefore the final optimization is expected to be less precise, being in a sense a compromise. This is borne out by experimental results showing that rotated power functions produce registrations with higher error than the same number of non-rotated ones.

Therefore, if the objective is to eliminate gross misalignments, it would be reasonable to separate the higher-level transformation from the minimization procedure, which can be viewed as attempting specialised optimizations on multiple (intrinsically equivalent) templates and picking a good-enough registration, instead of a general optimization on a single template. Additionally, perturbation search is also distinct from adapting φ, as was demonstrated with $\varphi = (\mathbf{P} \circ \gamma \circ \mathbf{S})(\mathbf{x})$ in the case of industrial inspection, as no prior knowledge is required to be assumed about the actual transformation. It can therefore be employed under any circumstances.

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

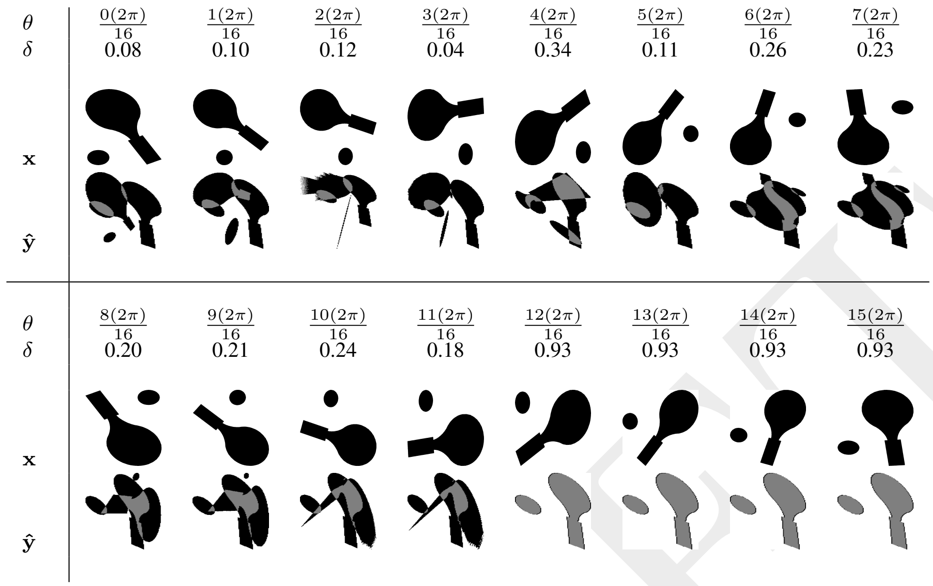

Figure 1: Example of perturbation search on Image 36, Observation 29
NLSR by default returns the registration $\hat{\mathbf{y}}$ with $\theta = \frac{0(2\pi)}{16}, \delta = 0.08$, which is poor.
In this case ($\delta_t = 0.9$), search would stop after $\theta = \frac{12(2\pi)}{16}$, as $\delta = 0.93 > \delta_t$

3 Experimental Results

For our first experiment, we use the same benchmark dataset of 1517 images as used in [1], comprising 37 different images each with one template and 40 independent observations. The images are binary, with background pixels in black and foreground pixels in white. The observations were generated synthetically from the template by randomly chosen projective transformations.

We first execute the provided software with default settings on all 1480 observations, to confirm that NLSR is indeed robust, with results shown in Figure 2. 91.1% of observations have $\delta > 0.9$, and 98.0% have $\delta > 0.7$, which indicates generally successful registration. However, there is also a minority of obviously incorrectly registered images. Note that NLSR has previously been demonstrated to be superior to the well-known Shape Context method [8] on this dataset in [1].

We next execute the provided software incorporating perturbation search with $\delta_t = 0.9$ and $\gamma = \{\left(\begin{smallmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{smallmatrix}\right), \theta = \frac{2\pi}{16}, \ldots, \frac{15(2\pi)}{16}\}$, using the wrapper described in Algorithm 1 (example output in Figure 1). The improved registration results are shown in Figure 3. 94.7% of observations have $\delta > 0.9$, and 99.9% of observations have $\delta > 0.7$. The single exception, observation 7 of image 20, has δ discounted due to its thinness, with almost all its error attributable to translation. The effect of perturbation search on the thirty observations for which $\delta <= 0.7$, roughly corresponding to registration failure, under the original NLSR is cataloged in Figure 4. The average time required per image does increase from 2.11s to 5.50s.

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

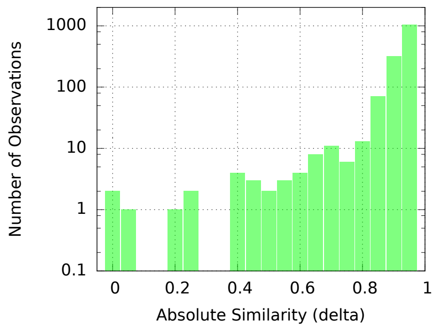

Figure 2: Histogram of δ_{NLSR}

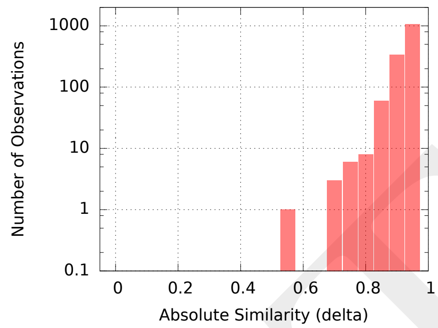

Figure 3: Histogram of $\delta_{NLSR+PS}$

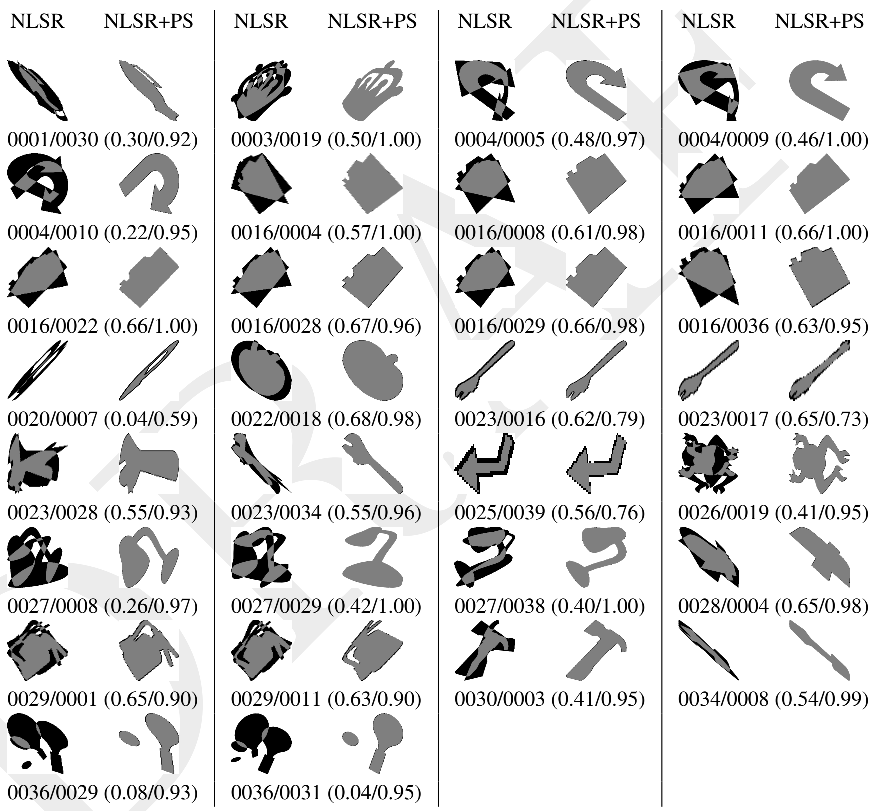

Figure 4: Comparism of registrations under NLSR and NLSR+PS
$\delta_{NLSR} <= 0.7$; overlapping pixels in grey, non-overlapping pixels in black
Legend format: [Image #]/[Observation #] (δ_{NLSR}/$\delta_{NLSR+PS}$)

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

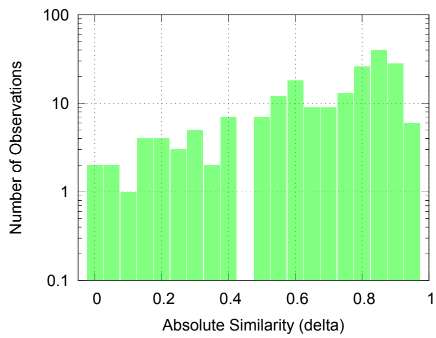

Figure 5: Histogram of δ_{NLSR}

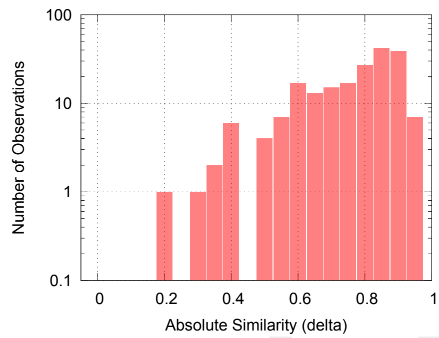

Figure 6: Histogram of $\delta_{NLSR+PS}$

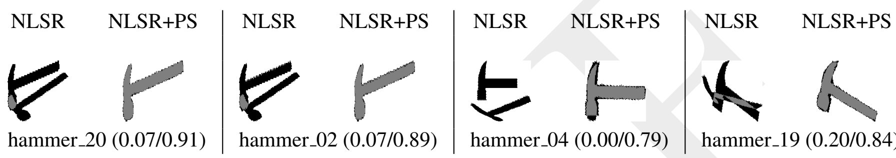

Figure 7: Images with largest improvement in δ for SIID

To test the performance of NLSR+PS against NLSR on observations that are not pure projective transformations of the template, we use the 216-shape version of the Shape Indexing of Image Databases (SIID)[2] [9]. It consists of 18 categories of objects, each with 12 binary image observations. While objects of the same category are similar in appearance, they cannot be mapped perfectly to each other in general. For our purposes, we select the first observation in each category as the template, and report the registration results on the remaining 11 observations. Average δ improved from 0.711 for NLSR to 0.785 for NLSR+PS, and the worst misregistrations were again averted (see Figures 5 and 6)

Finally, we compare the performance of NLSR+PS against NLSR on degraded images, specifically occluded and disoccluded images, as these non-uniform localised errors were the types that NLSR were found to be vulnerable against [1]. For each observation, we created an occluded and disoccluded version, with a square-shaped region of size equal to 10% of the shape removed or added at a random position respectively.

Although the average δ error was slightly reduced for NLSR+PS from NLSR, perturbation search does not qualitatively improve the registrations much in general, which can be explained by the underlying NLSR remaining an area-based method sensitive to changes in relative area. Further, the improvement in δ is accompanied by a worsening of the true error $\epsilon = \frac{1}{F_t} \sum_{x \in F_t} |\varphi(x) - \hat{\varphi}(x)|$ in some image pairs; see Figure 8 for an example where the area of occlusion dominates the area of the pen clip. This inability to take such features into account is an inherited limitation of the proposed extension.

[2] `http://www.lems.brown.edu/vision/researchAreas/SIID/`

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

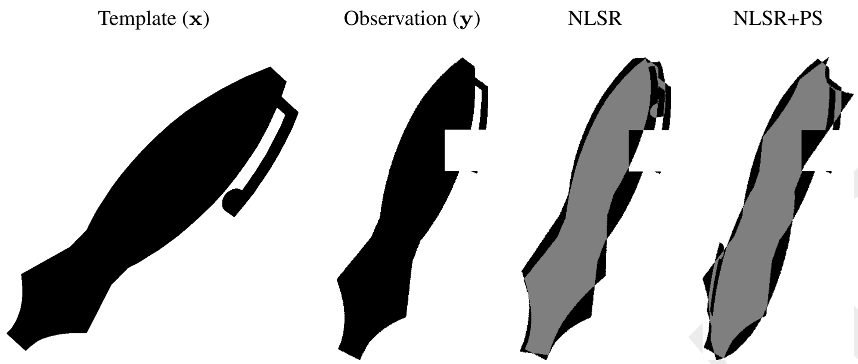

Figure 8: Example of occlusion registration on Image 1, Observation 24 $\delta_{NLSR+PS} = 0.75 > 0.73 = \delta_{NLSR}$, but $\epsilon_{NLSR+PS} > \epsilon_{NLSR}$ as registration by NLSR+PS is off by approximately 180 degrees.

4 Conclusion

We have demonstrated that searching over methodical fixed perturbations on the template domain upon initial nonlinear registration failure can nearly always reconcile the discrepancy for two-dimensional image pairs. Furthermore, the amount of time allocated for such a search can be fully controlled by the end-user, although in practice a small number of additional registrations appears to be sufficient. The procedure might be further sped up by tighter integration within the LM optimization step, in particular early stopping when convergence is slow [10]. We expect similar behaviour to be exhibited for registration both with other φ models, and also in higher dimensions. However, in such cases more dedicated methods [11] of selecting the γ_j functions would likely be very helpful, since the space to be explored expands in analogy to the "curse of dimensionality". The theoretical question of how the characteristics of γ affects successful registration remains open.

Supplementary Material

The code for all experiments in this technical report may be downloaded from `http://hamlab.glys.com/perturbsearch/`.

DRAFT COPY PLEASE DO NOT CITE

DRAFT COPY PLEASE DO NOT CITE

References

[1] C. Domokos, J. Nemeth, and Z. Kato, “Nonlinear shape registration without correspondences,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 34, no. 5, pp. 943–958, 2012.

[2] D. W. Marquardt, “An algorithm for least-squares estimation of nonlinear parameters,” *Journal of the Society for Industrial & Applied Mathematics*, vol. 11, no. 2, pp. 431–441, 1963.

[3] S. Mann and R. W. Picard, “Video orbits of the projective group a simple approach to featureless estimation of parameters,” *Image Processing, IEEE Transactions on*, vol. 6, no. 9, pp. 1281–1295, 1997.

[4] X. Huang, N. Paragios, and D. N. Metaxas, “Shape registration in implicit spaces using information theory and free form deformations,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 28, no. 8, pp. 1303–1318, 2006.

[5] M. S. Hansen, M. F. Hansen, and R. Larsen, “Diffeomorphic statistical deformation models,” in *Computer Vision, 2007. ICCV 2007. IEEE 11th International Conference on*. IEEE, 2007, pp. 1–8.

[6] K. Fujiwara, K. Nishino, J. Takamatsu, B. Zheng, and K. Ikeuchi, “Locally rigid globally non-rigid surface registration,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1527–1534.

[7] Y. Jiang, E. Edmiston, F. Wang, H. P. Blumberg, L. H. Staib, and X. Papademetris, “Shape comparison using perturbing shape registration,” in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 683–690.

[8] S. Belongie, J. Malik, and J. Puzicha, “Shape matching and object recognition using shape contexts,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 24, no. 4, pp. 509–522, 2002.

[9] T. B. Sebastian, P. N. Klein, and B. B. Kimia, “Recognition of shapes by editing shock graphs.” in *ICCV*, vol. 1, 2001, pp. 755–762.

[10] F. M. Dias, A. Antunes, J. Vieira, and A. M. Mota, “Implementing the levenberg-marquardt algorithm on-line: A sliding window approach with early stopping,” in *2nd IFAC Workshop on Advanced Fuzzy/Neural Control*, 2004.

[11] J. Elseberg, S. Magnenat, R. Siegwart, and A. Nüchter, “Comparison of nearest-neighbor-search strategies and implementations for efficient shape registration,” *Journal of Software Engineering for Robotics*, vol. 3, no. 1, pp. 2–12, 2012.

DRAFT COPY PLEASE DO NOT CITE