

Received August 23, 2017, accepted October 10, 2017, date of publication October 13, 2017, date of current version February 14, 2018. Digital Object Identifier 10.1109/ACCESS.2017.2762901

Joint-Saliency Structure Adaptive Kernel **Regression with Adaptive-Scale Kernels for Deformable Registration of Challenging Images**

BINJIE QIN \blacksquare , (Member, IEEE), ZHUANGMING SHEN ¹, ZESHAN FU¹,

ZIEN ZHOU², YISONG LV³, AND JINSONG BAO⁴

¹School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai 200240, China

²Department of Radiology, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200240, China

³School of Mathematical Sciences, Shanghai Jiao Tong University, Shanghai 200240, China

⁴College of Mechanical Engineering, Donghua University, Shanghai 200051, China

Corresponding authors: Binjie Qin (bjqin@sjtu.edu.cn) and Jinsong Bao (bao@dhu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61271320 and Grant 51475301 and in part by the Medical Engineering Cross Fund of Shanghai Jiao Tong University under Grant YG2014MS29.

ABSTRACT This paper proposes a locally adaptive kernel regression with adaptive-scale kernels for deformable image registration with outliers (i.e., missing correspondences and large local deformations). The adaptive kernel regression locally constructs dense deformation elds from the weighted contributions of each pixel's surrounding discrete displacement elds in a moving anisotropic kernel by exploiting the contextual deformations of the corresponding saliency structures in the two images. Speci cally, we rst propose an effective superpixel-based structure scale estimator to estimate the boundary-aware structure scale of each reference structure. We further propose an edge-aware mismatch scale measuring the mismatch degree of the edge structures to be matched in the images. By combining the boundary-aware structure scale with the edge-aware mismatch scale of the underlying saliency structures to be matched, we de ne edgeaware adaptive-scale kernels for the locally adaptive kernel regression to efficiently construct deformations for deformable registration with outliers. The experiments show that the proposed method achieves not only state-of-the-art matching accuracy for normal corresponding structures but also the best matching ef ciency for outlier structures in deformable image registration.

INDEX TERMS Deformable registration, structure scale, mismatch scale, joint saliency map, outliers, spatially adaptive nonparametric regression, joint-saliency structure.

I. INTRODUCTION

registration. The missing correspondences and local large Deformable image registration [1], or optical ow [2] com- deformations of local structures are called outliers in this putation (referred to as monomodal deformable image regispaper, robustly determining the deformation elds or distration), is the task of spatially aligning the points of every placement elds representing the correct alignment of the corresponding local structure by minimizing the feature-local structures is still a challenging unsolved problem in based and/or intensity-based differences between two imagemany elds, such as machine learning [3], signal process-Accurately matching corresponding local structures in twoing, medical imaging and image guided surgery [4]. Fig. 1 images by deformable image registration has numerous applilustrates both missing correspondence and large local cations in computer vision, image processing and patterndeformation problems in the two images (Figs. 1(a)-(b)) to recognition [1] [4]. However, because of the image content be registered [28]. Compared with traditional registration changes over a period of time and the different imagingapproach (Fig. 1(c),(e)) that cannot align the normal and mechanisms of multimodal sensors, some local structuresutlier structures via realistic and reasonable deformations, presented in one image appear partially or even disappearur method (Fig. 1(d),(f)) accurately aligns the normal correcompletely in another image. These local structures withsponding regions, and maps the outlier regions to the right missing correspondences are closely intermixed with thelocation but relax the deformations in the outlier-affected structures' large local deformations in the deformable imageregions.



FIGURE 1. Comparison between using traditional and our image registration methods [28]. (a) and (b) The reference and moving images. (c) and (e) Traditional method introduces eye distortions (red arrow) in registered moving image, producing the conflicts of the deformation fields. (d) and (f) Our method accurately aligns the normal and outlier structures, relaxing the deformations in the outlier regions.

than its own scale. Therefore, deformation models with spatially adaptive regularization [7] [9] have been proposed to address the varying deformation properties of local structures. Certain image segmentation-based works [10] [12] use an informative deformation prior for a speci c region or tissue type to locally adapt the deformation eld at various structures. The image segmentation is also used to address the missing correspondence problem by creating local arti cial correspondences [13], [14], discarding the missing correspondences via cost-function masking [15], [16], or developing geometric metamorphosis [17] to separate the normal deformations from the outlier changes. While effective, these methods require explicit structure segmentations or initial outlier localizations. Recently, a low-rank and sparse decomposition technique [18], [19] has been able to separate the outlier structures from the "healthy" parts in a collection of images to be registered. Despite the success, this

Deformable image registration can be formulated as themethod may be limited in image sequence applications when problem of globally searching for the optimal transforma- separating out the sparse components that are not consistent tion T that minimizes the cost function $(I_R, I_M \circ T) + S(T)$ with the low-rank structures [20]. between the reference image and the moving image_M.

The global cost function consists of two terms: the dataularization to distinguish motion differences for different term D quanti es the difference and level of alignment regions. This data-driven strategy exploits a spatially adaptive between the two images, and the regularization term transformation prior [21], local changes in intensities and regularizes the transformation toward favoring realistic and deformation elds [22], or measures of local image reliabilreasonable deformation solutions and seeks to address the illey [23] to affect the local regularization strength. Recently, an posed problem of deformable image registration. The datapptical ow estimation was able to integrate the sparse match term is referred to as a matching criterion and includespropagation or aggregation [24] [26] into a global optimizaintensity- and/or feature-based approaches. Feature-based framework to estimate the large displacements of small methods [5], [6] usually establish dense deformation elds structures, while a large deformation diffeomorphic metric by interpolating the sparse correspondences between locanapping [27] was successfully used to address large deformafeatures. Locating reliable local invariant features from thetion problems but was highly in uenced by the image intenoutlier structures remains an open problem in feature-basedity pro le with missing correspondences. Nonetheless, these methods. Intensity-based approaches use the informatiomethods do not fully consider both the outlier structures (i.e., of all image pixels to directly estimate the most exible missing correspondences and large local deformations) and dense deformation (or displacement) eld for each pixel (or motion boundary removal nor do they consider the globalvoxel), which can better quantify and represent the matchingo-local contextual information of the corresponding saliency accuracy of every point in local structure pairs. structures in the two images during the registration procedure.

Because global regularization introduces excessive exi- Actually, the corresponding saliency structures [28] convey bility, the intensity-based approach may favor unrealistic and the most useful global-to-local contextual information during unreasonable local deformations when it diffuses transformaimage registration. Finding the correspondences between the tions from the structural to non-structural regions. In partic- corresponding saliency structures is not only the starting point ular, the local structural regions make stronger contributionsbut also the ultimate goal of image registration. To efin the cost function minimization than do the non-structural ciently match the structures from outliers, the joint saliency regions, and thus, the transformation computation is easilystructures' contextual consistency is exploited in [28] for the affected by over-smoothing and is limited to the deformationsspatially adaptive deformation construction. The idea of the in these structural regions. Furthermore, the multi-resolutionspatially varying treatment [31] of joint-saliency and outlier strategy for the large local deformation problem has the fol-voxels has also been successfully adopted in the feature-based lowing inherent disadvantages: the basic sub-sampling proDRAMMS approach [32] for challenging registration probcedure in the multi-resolution strategy causes some displacements involving pathology-induced outliers. However, there ment details of the edge structures to be removed; inaccuratexists limitation of utilizing a xed kernel scale (or kernel initialization and outlier effects propagated from the coarsewidth). By determining the sample size of the displacement level lead to incorrect displacement estimations at ner levels vectors participating in the deformation construction, the in the re nement procedure; and the multi-resolution strategysptially varying kernel scale for nonparametric regression cannot correctly predict the relatively large motion of a is very important in controlling the balance between the small-scale structure that exhibits a larger scale deformationstructure matching accuracy and the smoothness of the local

deformation elds. This assumption is con rmed by the fact II. BACKGROUND AND MOTIVATION that in density estimation studies in the literature, almost all *A. JOINT-SALIENCY STRUCTURE ADAPTIVE* the adaptive-scale kernels [33], [40] have been shown to be *NONPARAMETRIC REGRESSION* superior to xed-scale kernels. Inspired by the success of nonparametric-regression-

To design an appropriate kernel scale for JAKR, we assume ased [37] machine learning for signal reconstruction, we that the kernel scale is adaptively selected according to the onsider the deformable image registration as a nonparacontextual information about the underlying structures and metric regression [28] to construct dense deformation elds their displacement vectors. Generally, a large structure hat mode deformation elds. This kernel-regressionmore contextual sample pixels for propagating their defor-based strategy is also implemented in deformable image mations to construct the current sample pixel's deformation, registration [38] and has been recently adopted in optical whereas a small structure is con ned to a small neighbor- ow estimation [39]. Suppose that we have some sparse and hood to prevent the neighboring structures' distortions from irregularly distributed deformation vectors y_i , x_i $_{i=1}^{p}$ given spreading into the current estimate and smearing the motion the form

boundaries. Moreover, large mismatches of local saliency structures require large kernel scales to include more contex-

tual displacement vector samples for the deformation con_{wherey} is a sparse displacement vector (asponse variable) struction, whereas small mismatches of saliency structures to position (explanatory variable) and $z(\cdot)$ and z

need small kernel scales. Therefore, the kernel scale adaptively depends on not only the local size of the underlying structures to be matched but also the degree of mismatch between the local structures. Assuming that the mismatches of the local structures can be appropriately measured by the edge alignment degree of the structures, we propose an edge-aware adaptive-scale kernel for edge-aware deformation construction in JAKR to handle outlier structures and motion boundaries two common and dif cult issues facing deformable image registration.

With the above-mentioned thoughts in mind, the proposed method represents three contributions: 1) After presenting a concise review on structure scale estimation for image processing, we propose a simple but effective boundaryaware local structure scale estimator: the estimator rst segments the reference image into superpixel-based [34] multi-resolution structural regions; then, it calculates the boundary-aware structure scales of these regions in terms of the local variance of Gaussian smoothing through the Bayesian estimation and minimal description length criterion (MDL) [35], [36]. 2) We present an edge-aware mismatch scale of the overlapping structure pairs of two images, whereby we can judge and control the registration inaccuracy for the underlying structure pairs during the deformable registration procedure. 3) We propose an adaptive edge-aware kernel scale by combining the mismatch scale with the structure scale into the JAKR for deformable image registration. Therefore, the JAKR with the adaptive-scale kernels (JAKRAK) can iteratively guide the local structure deformations to not only achieve the accurate matching of small edge structures but also maintain smooth deformation elds for deformable registration with outliers. The experimental results demonstrate that the proposed JAKRAK method not only achieves state-of-the-art intensity-based registration performance but also achieves the best alignment of all challenging outlier structures. The background and the proposed method are elaborated in Section 2 and Section 3, respectively, followed by the experimental results in Section 4. The whole paper is discussed and concluded in Section 5.

 $\mathbf{y}_i = \mathbf{z}(\mathbf{x}_i) + \mathbf{e}_i, \quad \mathbf{x}_i \in \Omega, \ i = 1, \cdots, P$



FIGURE 2. Multi-resolution flowchart of the proposed algorithm.

variance in the local estimation of the nonparametric regression. The are two types of approaches for kernel scale estimation in nonparametric regressionalug-in methods [37], [40] calculate the ideal scale by estimating the bias and the variance in the estimation of the mean squared error (MSE) between the real signal and its approximation. The *quality-of-t* statistics [37], [40], such as cross-validation and generalized cross-validation, are widely applied for the direct optimization of the estimation accuracy. The second estimation is de ned by the accuracy criteria and is always related to data-driven methods disregarding the bias estimates or formulas for the ideal kernel scale selection, with the main goal to achieve an optimal accuracy that balances the bias and the variance of estimation. This work uses this accuracybased estimation by taking the local structural matching contexts to boost the accuracy of deformable image registration with outliers.

We note that there are very different scale-estimation problems for 2D and/or 3D image analysis in pattern recognition and computer vision, where one goal is to describe the coarseness (or the optimal size for most spatial structures) of an image by any monotonically changing parameter. For example, the gradually changing time parametersed in the diffusion process [41] [44] of an image in scale space is

Fig. 2, which shows the three-step multi-resolution structure commonly treated as a scale parameter to globally control matching framework, with the different levels having their the smoothness of the whole image (or gradually remove own resolutions but following the same procedure. First, thethe object detail within the image). The single global scale moving image I_M is deformed with an initial displacement is widely used in many applications of multi-scale analyeld obtained via spatial interpolation of the output defor- sis: A single optimum scale [45], [46] based on Laplacian mation eld obtained on the previous level. The deformed of Gaussian (LoG) analysis of an image is identi ed as moving image and the reference image on the current level arthe smoothing parameter for a normalized LoG Iter to registered using block matching, with the point-wise mutual delineate blobs with similar sizes in medical images. The information serving as the local similarity measure. In the optimal scale based on a pre-estimation of the spatial and second step, with the JSM highlighting the overlapping JSS spectral statistics achieves satisfactory segmentation results for the deformation construction, this work estimates the with high homogeneity within the segments and high heteroscale of every reference structure and the scale (or degregeneity between the segments in multi-scale image segmenof mismatch between every pair of the underlying JSSs. Withtation [47]. An optimal scale determines the stopping of the anisotropic kernel representing the shape/orientation of V- ow [48]-based diffusion to reduce image noise while the reference structure, we estimate the edge-aware adaptivereserving the maximally stable extremal region features scale kernels for JAKRAK by combining the structure scale for computer-aided detection. Recently, by decomposing the with the edge-aware mismatch scale; then, we use JAKRAKimages into compact and region-boundary-aware superpixto construct the current deformation elds from the discrete els, the structure-guided statistical textural distinctiveness displacement elds. Finally, the resulting global deformation approach [49] illustrates that considering texture at a single for the iteration at the next level is composed of the initial scale is sufficient for reliable salient region detection in deformation and current deformation from sampling the ini- natural images. However, the global scale estimation has an tial deformation elds. intrinsic limitation: the single coarseness for image structures

in the whole image. The results of this class of algorithms might not be suf cient when the underlying ne and coarse

B. SCALE ESTIMATION IN NONPARAMETRIC REGRESSION

The kernel scale of the nonparametric regression [37], [40] is mage structures should be discriminatively analyzed in spacrucial for signal reconstruction when addressing noisy dataially adaptive schemes.

and outliers. A small scale corresponds to a smaller moving Rather than assuming a single global scale or multiple kernel for the nonparametric regression and therefore to noisscales for a whole image using prior knowledge of the scales ier estimates, with higher variance and typically decreased of the various objects of interest in the image, researchers estimation bias. A large kernel scale corresponds to smoother lways de ne a local scale to measure the size of local estimates, greater bias, and lower variance. Therefore, the tructures for each location of the image [50]. In estimation kernel scale controls the trade-off between the bias and pace-variant local scales, linear [41] [43] and nonlinear

(such as morphological operations) [51] [53] scale spaces deformation contexts). For example, the matching of large Laplacian and Gaussian pyramids [54] are widely used tdocal structures can use large kernel scales to reduce the achieve multi-scale image representation. Among the vari-deformation variance or increase the deformation smoothous local scale estimation methods, the method proposed byess compared with matching local small structures using a Lindeberg [50] is widely used for image structures such assmall kernel scale to reduce registration (or deformation) bias blobs, edges and ridges. The detected local scales may netrors. Moreover, the nonparametric regression of deformahave realistic meaning, as they simply detect local extremation elds may also blur the boundaries and motion details of over scales of normalized differential operators for the localstructures if using kernels crossing the boundaries between image representation of certain sparse locations. The detectedifferent structures. By considering a boundary-aware kerpoints, referred to as scale-space extrema, are sparsely disel scale in the kernel regression of deformation elds, we tributed in the image to represent interest points, blobs, corcan preserve the intra-structure deformation smoothness and ners, edges, ridges and valleys, and they do not consider the void inter-structure deformation smearing. actual structural information of the whole image. Alterna-On the other hand, the gradually re ned deformation tively, certain methods have utilized probabilistic approachesof small-scale structures with relatively large deformations to estimate the local scales in an image for edge detection [55] will be mistakenly predicted by the deformation of large and other low-level tasks [36] such as texture segmentationsaliency structures at coarse resolution levels in multi-An original strategy of local meaningful scale [56] detection resolution registration. To achieve a correct prediction, testrelies on the asymptotic properties of perfect shape digi-ing the local structures' matching early in the registration tizations to detect what the relevant scales at which eachprocedure and as often as possible is the best way to guide point of the digital contours should be considered. The localthe kernel-regression-based deformation construction toward adaptive scales for local pattern representation and texturaccurate structure matching. Assuming that the local strucsegmentation are also explored in several works by maxiture's matching can be validated by the alignment degree of mizing the changes between the average gradients for difthe corresponding edges of the overlapping structures, this ferent sizes of image blocks [57] or using total variation work proposes an edge-aware mismatch scale estimation to

ows [53], Gabor Itering [58], and energy minimization design an adaptive kernel scale in the registration ini320(an)-ac models [59]. Recently, some segment-based scale selection strategies [60] [63] have been proposed to determine the varying sizes of local segments (or regions) in an image such that all image pixels within a local segment satisfy a homogeneity or uniformity criterion. However, these localsegment-based scale estimation methods cannot automatically detect locally varying structures for spatially adaptive image processing.

As for image registration, most current image registration and optical ow approaches implicitly assume that the structures in both images are from the same scene and appear at the same scale. Nevertheless, image deformations often occur at different scales. Recently, Petial. [64] proposed multi-scale ow-based deformations by exploiting multiple kernels at different scales. A deformable spatial pyramid matching [65] was proposed to match pixels across scale differences coming from a discrete, pre-determined set of scales. The Deep-Matching [25] approach matched patches at several scales to overcome the lack of distinctiveness that affects small patches for optical ow computations. Tau and Hassner [66] established dense correspondences across structures with different scales to estimate the motion of small structures with large displacements and occlusions. However, the spatially varying scales of different geometric structures in the images are not considered in the above-mentioned works.

The locally adaptive scale in nonparametric regression is crucial in searching for an appropriate support of the local estimator for controlling the deformation smoothness and matching accuracy for the underlying saliency structures. On the one hand, a kernel scale is assumed to be spatially adapted to the underlying local structures (and their in Section 3.2 to quantify the mismatch of underlying local The residual component σ_k can be modeled as a zerostructures.

III. METHODS

A. STRUCTURE SCALE ESTIMATION

The structure scale is considered as the size of every image structure corresponding to each image segment [60]. With the image structure being de ned as a group of connected pixels with homogeneous features, structure scale estimation is formulated as a scale labeling assignment for each structure

in optimal multi-scale segmentation, which contains the most where $P(S_i | \sigma_k)$ is the likelihood of the observed structural homogeneous structures and the least edge-smearing mixegion S_i at scale σ_k , and $p(\mathbf{x} | \sigma_k) = P(I(\mathbf{x}) | \sigma_k)$ is the heterogeneous structures. To achieve scale invariance, the elihood of the observed image at each pixelt scales. structure scale is computed in a neighboring region adaptive To estimate the likelihood of the observed image at each to the local structures in a multi-resolution image pyramid. pixel, we use the well-known MDL criterion [35], [36] to Thus, the optimal structure scale refers to the optimal spatial elate the probability of an item with the length of the ideal extent or the optimal size of every local structure at every code used to describe it, namely, pyramid level.

To estimate the structural scales at every pyramid level,

we rst segment the image of every pyramid level into a where $L(I|\sigma_k)$ denotes the description length bbased on set of superpixel structural units that adhere to the structures decomposition at scale. This description length can be boundaries. We denote the whole image regior and the expressed $as(I|\sigma_k) = L(I_{\sigma_k}) + L(\varepsilon_{\sigma_k})$. On the one hand, the local structures as $(i = 1, \dots, n)$, with $\Phi = \bigcup_{i=1}^{n} S_i$. sampling theorem states that the number of samples needed The various structure units are then optimally smoothed to rdescribing a Gaussian smoothed image is proportional to be internally homogeneous by the spatially varying Gaus-the Gaussian Iter bandwidth in frequency space. Due to the sian Iters, with some variances in a discrete scale space uncertainty principle, this bandwidth is inversely proportional The variance σ^2 of the Gaussian lter controls the amount to σ_{k}^2 . Therefore, the number of samples needed for describof Gaussian smoothing and thus the homogeneity of eaching the Gaussian smoothed image is controlled by the spastructural region. With the minimal and maximal amounts of tial variance of the Gaussian kernel. The description length smoothness being controlled by the and σ_m , respectively, of the smoothed component (I_{σ_k}) is thus assumed [36] to be inversely proportional to c_k^2 and can be written as in the discrete scale set $(k \in \{1, \dots, m\})$, the optimal structure scale for each smoothed structure unit is obtained $(I_{\sigma_k}) \sim \frac{1}{\sigma_k^2}$. On the other hand, because the probability by maximizing its posterior probability from Bayes' theorem. Considering the scale coherence between neighboring σ_{σ_k} is modeled as a zerostructure units, we also use a Markov Random Field (MRF) mean Gaussian distribution, the description length of the model constraint to create a single large-scale labeling for z = 2 . The second straint $(\varepsilon_{\sigma_k}) = -\log_2 P(\varepsilon_{\sigma_k})$ is proportional the neighboring structure units with similar appearances. The $\varepsilon_{\sigma_k}^2$. Therefore, the local description length *b*based on its decomposition at the scale and is estimated as follows: nal structure scale estimation is an optimal labeling image,

with its segments achieving the most homogeneity within structural regions and the least edge-smearing in mixed structural regions.

Speci cally, a scale space of the imag/ex) is rst constructed by a the convolution operation $h_{\mu}(\mathbf{x})$ $(I_0 * G_\sigma)(\mathbf{x})$, where $G_\sigma(\mathbf{x}) = \frac{1}{(2\pi\sigma^2)^{N/2}} e^{-|\mathbf{x}|^2/2\sigma^2}$ denotes the Gaussian kernel and the variance is a certain scale

parameter from the scale set. In this work, we assume that the largest scale in the scale set is 15 pixels and that the small-

est scale is 1 pixel. Because an image can be decomposed inthere A is a normalizing constant, and and β are empiria smoothed component and a residual component through a ally set to 1 in this work. anisotropic diffusion lter, the intensity of a local superpixel S_i can be represented by the smoothed compole (\mathfrak{R}_i) and the residual component

$$I(\mathbf{x}) = I_{\sigma_k}(\mathbf{x}) + \varepsilon_{\sigma_k}(\mathbf{x}), \quad \mathbf{x} \in S_i$$
(4)

mean Gaussian random variable by the central limit theorem. Thus, the local structure scale estimation assigns a scale $\sigma_k, k \in \{1, \dots, m\}$ from the scale space generated for each local structures, such that the following posterior probability achieves the maximum value

$$P(\sigma_k | S_i) = \frac{P(\sigma_k) P(S_i | \sigma_k)}{P(S_i)}$$

$$\propto P(S_i | \sigma_k) = \prod p(\mathbf{x} | \sigma_k), \mathbf{x} \in S_i$$
 (5)

$$P(I|\sigma_k) = 2^{-L(I|\sigma_k)}$$
(6)

$$L(I|\sigma_k) = \alpha(\frac{\beta}{\sigma_k^2} + \varepsilon_{\sigma_k}^2(\mathbf{x}))$$
(7)

where α and β are positive parameters [36] that depend on the coding precision in bits used to represent the smoothed image and on the assumed noise variance. Using equation (6), we further estimat $\varphi(\mathbf{x} | \sigma_k)$ by the following equation:

$$\hat{p}(\mathbf{x}|\sigma_k) = A \mathbf{e}^{\left[-\alpha \left(\frac{\beta}{\sigma_k^2} + \varepsilon_{\sigma_k}^2(\mathbf{x})\right)\right]}, \quad \mathbf{x} \in S_i$$
(8)

The scale eld for the neighboring similar pixels is assumed to be inherently smooth due to the intra-structure homogeneity being usually visible in the natural world. Considering the scale coherence between similar neighboring pixels, we implemented the MRF model in the structure scale

1

estimation. As a result, the nal structure scale is estimated coarse resolution (Fig. 3(f)). With the increasing image resoas

$$\sigma_{S} = \arg \max_{\sigma_{k}} P(\sigma_{k}|S_{i}) + \lambda \sum_{\langle i,j \rangle} \delta(\sigma_{k},\sigma_{l}) \exp(-(\mu(S_{i}) - \mu(S_{j}))^{2})$$
with $\delta(\sigma_{k},\sigma_{l}) = \begin{cases} 1, & \text{if } \sigma_{k} = \sigma_{l} \\ 0, & \text{otherwise} \end{cases}$
(9)

where *i*, *j* are the indices of local neighboring structure S_i and $\mu(S_i)$ are the mean intensities of the local structures and S_i , respectively; and σ_i and σ_i are the scales of j_i and S_i from the scale set, respectively. In equation (9), the rst term always least informative and assumed to continuously have is the posterior probability $o f_i$, and the second term is a smoothness function of the local struct frend its neighboring pyramid levels; thus, they maintain smooth deformation ing local structures;. The second term prefers the same scale elds and result in the smallest driving in uences in the labeling for neighboring pairs of similar superpixel regions multi-resolution deformation construction. Conversely, edge and avoids creating the same scale labeling for neighboringstructures, with their narrow areas of high contrast and ne pairs of very dissimilar superpixel regions. The impact of detail, are most informative in driving the deformations and MRF is controlled by the parametier which is usually set to a are easily changed or bleared during multi-resolution regissmall value (0.05) to avoid the over-smoothness that increases ation such that their deformation con icts (and the topology the structure scales of small local structures.

changes in the structures) can be widely found in the discrete In Figs. 3(a)-(b), the reference and moving ower images displacement elds. Under the above-mentioned considerare 384× 288 pixels, having a stamen lament with both ations, the edge structures' mismatches must be evaluated missing correspondences and large local deformations induring the registration procedure to guide the kernel scale the top-right corner of the images. Figs. 3(f)-(h) show the estimation for the nonparametric regression of the deforma-

superpixel-based structure scales for the multi-resolutiontion elds. saliency structures. The process roughly segments the fore- With the JSM representing the matching degree of the ground structural regions and background regions at theunderlying saliency edge structure pairs [28], the mismatch

scales are inversely related to the JSM values for the adaptive nonparametric regression during the multi-resolution registration procedure. Figs. 3(c)-(e) show the multi-resolution edge-aware JSM with the color scale representing different joint saliency values. The high joint saliency values (in red) mean that the underlying pixel pairs come from the matched edge structures (or JSSs), whereas the low JSM values (blue and yellow-green) are from either unmatched structural regions (including outlier regions) or homogeneous regions. At every pyramid level, a zero or very small mismatch scale value is thus assigned to the corresponding structural regions with a high JSM value, whereas a large mismatch scale value is given to the unmatched structural regions with a low JSM value. Because they contribute the least to driving

FIGURE 3. Flower images and their multi-resolution JSM, structure scales, mismatch scales and kernel scales. (a)-(b) The reference and moving images at the 384 x 288 pixels resolution. (c)-(e) multi-resolution JSMs, (f)-(h) multi-resolution structure scales, (i)-(k) multi-resolution mismatch scales, (I)-(n) multi-resolution kernel scales

lution reducing the size of the superpixels and enhancing the image details, a small number of small structure scales are appropriately assigned to the small structures (e.g., the small petals, the petal boundaries and the stamen lament in the upper-right corner of the reference image in Figs. 3(g)-(h)), while a large number of moderate structure scales and the maximal structure scales are displayed for the foreground structural regions and the homogeneous regions, respectively.

B. MISMATCH SCALE CALCULATION USING JSM

Generally, mismatches often coincide with and are driven by intensity changes. Speci cally, homogeneous regions are large areas of smooth intensity variations at the neighborwhere JS represents the JSM at each overlapping pixel the mismatched saliency structures and outlier structures pair in the two images. The normalized mismatch scalesgradually reduce their support regions (see the increasingly (Figs. 3(i)-(k)) are thus computed to generally display converging regions in the upper-right corners in Figs. 3(I)-(n)) three types of regions during the image registration: in the image space to achieve the transformation from smooth zero-mismatch-scale regions (in blue); low-mismatch-scaledeformation to accurate structure matching during the regisregions (in green and yellow-green) being from highly tration procedure.

matched edge structure regions; and high-mismatch-scale According to the aforementioned analysis, the moving regions (in red), indicating that the underlying regions are image's local saliency structures are gradually matched to the from mismatched structure regions or outlier regions. corresponding reference structures by iteratively selecting a

C. LOCAL ADAPTIVE KERNEL SCALE

locally adaptive scale for the local nonparametric regression. Fig. 4 illustrates why we prefer the edge-aware adaptive ker-

As mentioned above, the mismatch scale of the underly nel scales to the xed kernel scale in the proposed JAKRAK ing edge structure pair is estimated to indicate the extentramework. The black stripes in the local `E' pattern at of possible deformation improvement by the kernel regres-the top-center region of the hat are small-scale structures sion. Speci cally, the moving saliency edge structures with with large local deformations (see the reference and moving high normalized mismatch scales require large deformation mages in Figs. 4(a)-(b)). Figs. 4(e)-(f) show the two zoomedimprovements so that the mismatch scales can be used as versions of the black stripes for the E' patterns registered weights to linearly enlarge the underlying kernel scales gath by the JAKR (Fig. 4(c)) and JAKRAK methods (Fig. 4(d)). ering more sparse displacement vector samples for the desiGompared with the JAKR method introducing local irregular able deformation construction. On the other hand, the movdistortion in the stripes, the JAKRAK method can obtain ing saliency edge structures with low normalized mismatchaccurate and smooth deformations of these local stripes. scales need small deformation adjustments to achieve theigs. 4(g)-(h) present a performance comparison overview of desired deformation accuracy. Because the structure scale mesh deformation process (10-pixel vertex spacing) for already indicates the size of the contextual structure, thethe JAKR and JAKRAK methods with xed-scale and edgekernel scale is not only proportional to the structure scaleaware adaptive-scale kernels. JAKRAK can ensure a smooth but also weighted by the mismatch scale of the underlyingadaptivity of the local mesh deformation to local structures structure pair. Given the structure scale and the mismatch of varying sizes. Speci cally, the edge-aware adaptive-scale kernels for JAKRAK obtain smooth mesh deformations that scale σ_m , we are ready to design the local kernel scales

$$\sigma_d = \max\{\sigma_s \times \sigma_m, 1\} \tag{11}$$

are seamlessly consistent with the boundaries of local structures with varying sizes, while the xed-scale kernels can

where 1 avoids the local kernel scale being less than 1 pixelproduce more or less irregular mesh deformations that are not Figs. 3(I)-(n) illustrate the local kernel scales for the refer- smoothly adaptive to the local structures (see Fig. 4(g)).

ence and moving images (Figs. 3(a)-(b)) for multi-resolution registration, with the color scale representing different normalized scale values. The large corresponding saliency structures with their surrounding homogeneous regions cover a relatively large range of kernel sizes (see the central regions of Fig. 3(I)) that correspond to the large areas of real image contents at the coarse resolution. These areas initialize the smooth deformation construction, while the small saliency structures at the ne resolution re ne these deformations to increase the matching accuracy. With the small structures being gradually joined and assigned relatively large_{FIGURE 4. Performance comparison of using fixed-scale and} kernel scales in the iterative nonparametric regression, the daptive-scale kernels. (a)-(b) The reference and moving images. (c) the background and homogeneous foreground regions gradually reduce their kernel scales to the smallest values for their expanding overlapping areas (see Fig. 3(n)) so that the deformation construction can be gradually adjusted to achieve the transition from deformation smoothness to deformation (or matching) accuracy.

Meanwhile, the multi-resolution kernel scales of outlier structures and small saliency structures are mostly dependent on their mismatch scales (Figs. 3(i)-(k)). Speci cally, the outlier structures and the mismatched saliency structures always have relatively large kernel scales in the multiresolution scheme. These relatively large kernel scales for

EMPIRE10 [69² data sets have been set up speci cally Due to the missing correspondences preventing the correfor thoracic image registration. However, these data sets depending landmark de nition, the landmark selection cannot not include outlier structures with both missing correspon-include outlier features with missing correspondences when dences and large local deformations for challenging imageocusing on the easily identi able corresponding locations registration. The objective and rigorous evaluation of theat the JSS pairs. Lower average error distances and lower performance of challenging image registration is demon-standard deviations imply a more accurate alignment of norstrated in the work [70] using an in-house database containingnal local structures. In most cases, the visual valuation can eight patients with recurrent brain tumors. These pathology be perfectly consistent with the landmark-based registration bearing images introduce outliers with both missing corre-evaluation. However, due to the inability to de ne the corspondences and large local deformations and include tworesponding landmark pairs within and around the outlier independent expert decisions on the corresponding landmanlegions with missing correspondences, the landmark-based de nitions and ROIs. Those landmarks and ROIs served as egistration evaluation cannot fully evaluate the real perforreferences for measuring the registration accuracy. Howevemances of these methods in matching outlier structures. This due to the HIPPA regulation (Health Insurance Portability and limitation is compensated by visually evaluating the zoomed-Accountability Act), this database was publicly unavailable in display of the outlier structures in the following section. during the preparation of this manuscript. The rst experiment involves aligning two grayscale

Our algorithm has been implemented to support 2D/3DMickey images (Figs. 5(a)-(b)) with an outlier doctoral cap deformable image registration. In this section, we usein the moving image. There are large deformations charactera set of typical challenging 2D image pairs to validate izing Mickey's left thumb, left hand, right thumb (see the red the performance of the proposed JAKRAK methoday boxes in Fig. 5), and right shoe as well as the right button comparing it with the JAKR method, SparseFlow method on Mickey's belly. Therefore, the registration performance (SF)⁴ [26], DeepFlow (DF[§]) [25], Advanced Normalized evaluations are largely dependent on the deformation results Tools (ANTs)⁶ [71] with greedy symmetric normalization of these structures. Figs. 5(c)-(h) show that the JAKRAK diffeomorphic transformation and mutual information as sim- (Fig. 5(c)) and DF methods (Fig. 5(f)) outperform other ilarity measure (AGS), and exible variational non-linear methods by perfectly deforming the local structures to the intensity-based (FVNI) method72]. The JAKR, AGS and desired positions. However, the DF method inadequately dif-FVNI methods have demonstrated [28] state-of-the-art perfuses the deformation into the wrist of the left hand. The formances for deformable registration on challenging imageSF method achieves good structure matching performance with outliers. The parameters of the JAKRAK method are in the corresponding regions but de cient deformations in the same as those of the JAKR method [28] so that all thethe left thumb, which resulted in a large variance in the algorithms are set with the default parameters for achievingfollowing landmark-based registration evaluation. Obviously, the FVNI method has introduced a rounding image artifact their best performances.

We use both landmark-based [70] registration erroraround the right button on Mickey's belly. In contrast, the measurements and visual valuation to fully evaluate the perstructure of Mickey's left hand is abnormally distorted by the formances of the six competing methods in the seven chalAGS (Fig. 5(g)) method.

lenging image registrations. The landmark-based registration error measurement task measures the matching accuracy for the normal corresponding structures in the two images, while the visual valuation is simply for the outlier structures with both missing correspondences and large local deformations. Speci cally, we not only zoom in on some small local structures in the registered moving images, displaying their deviation from the desired locations with several red crosses, but also manually select a large number of densely distributed landmark pairs from two experts in the two images for measuring the registration errors. Considering

the uncertainty of manual landmark selection, we use the FIGURE 5. Mickey image registration with missing correspondences and mean registration error (MRE) and standard deviation (SD)arge local deformations in the upper-right region. The boxed regions between the landmark pairs as the standard for registration deformations. (a)-(b) The reference and moving images, (c) JAKRAK, evaluation. (d) JAKR, (e) SF, (f) DF, (g) AGS, (h) FVNI.

²http://empire10.isi.uu.nl

³http://www.escience.cn/people/bjqin/research.html

⁴http://www.vision.ee.ethz.ch/~timofter/software/SparseFlow.zip

⁵http://lear.inrialpes.fr/src/deep ow/

⁶http://www.picsl.upenn.edu/ANTs

⁷http://hdl.handle.net/10380/3460

The second experiment, displayed in Fig. 6 for ower image registration, includes both missing correspondences and large local deformations of small structures, where the outlier stamen lament in the right part of the reference image (Fig. 6(a)) has large local deformations driven by



FIGURE 6. Flower image registration with the upper-right outlier regions of the stamen filament. (a)-(b) The reference and moving images, (c) JAKRAK, (d) JAKR, (e) SF, (f) DF, (g) AGS, (h) FVNI, (l)-(p) the corresponding zoomed versions of the red box regions for the stamen filaments (defined at (a)) in images (a)-(h), with the stamen filaments having the desired positions indicated by red crosses.



FIGURE 7. Hat image registration with the large local deformations of thin strips. (a)-(b) The reference and moving images, (c) JAKRAK, (d) JAKR, (e) SF, (f) DF, (g) AGS, (h) FVNI, (i)-(p) the zoomed versions of the boxed

the movement of the center owers. In addition, some buds regions (defined at (a)) in images (a)-(h), with the thin stripes having the behind the stamen lament in the moving image (Fig. 6(b)) desired positions indicated by red crosses.

disappear in the reference image but appear in the moving

image. Except for the AGS method (Fig. 6(g)) introducing

excessive deformations in the bottom petal, the JAKRAK of its neighboring local tiny structures and thus lead to poor method and the other methods in Figs. 6(c)-(h) achieve goos tructure alignment in a certain region. Figs. 7(c)-(h) show registrations, which properly deform the small-scale stamenthe registered moving images obtained by the six methods. Iament and the large-scale petals simultaneously. However,The zoomed versions (Figs. 7(i)-(p)) of the stripes of `E' the zoomed versions of the stamen laments in Figs. 6(i)-(p) demonstrate that only the JAKRAK, DF and SF methods can be used to distinguish the best performances of the (Figs. 7(k), (m), (n))) accurately match every small-scale JAKRAK method (Fig. 6(k)) in matching small-scale struc- structure (e.g., in the red crosses) of the stripes, whereas the tures from outliers compared with the other methods because AKR, AGS and FVNI methods introduce excessive distor-JAKRAK achieves precise structure matching in the tip of tions in the stripes of `E'.

the stamen lament (in the red cross at the top-right corner). Four experiments involving matching pre- and post-Although the JAKR, SF and DF methods (Figs. 6(I)-(n)) can operative brain tumor resection images were performed. obtain smooth registrations around the stamen lament, theyIn these experiments, surrounding normal brain tisare unable to achieve the desired large deformations in the tipues suppressed by tumor in the preoperative image of the stamen lament. Figs. 6(o)-(p) show that the AGS and (Figs. 8(1-a)-(1-b), 8(2-a)-(2-b)) expand after tumor resec-FVNI methods introduce more or fewer artifacts and unac-tion, which introduces not only the missing correspondences ceptable deformations around the stamen lament. Due toof the tumor in the post-operative images but also the large the subsequent landmark-based evaluation having dif cultylocal deformations caused by the brain shift. A desirable in de ning suf cient landmarks in the small-scale structures, registration method should smoothly deform the tumor region the zoomed-in display of the visual evaluation performs betterand surrounding preoperative brain tissues (see the red boxes than the landmark-based evaluation in evaluating the chalin Fig. 8) according to the post-operative image struclenging registration of small-scale structures with missingtures regardless of tumor resection. Figs. 8(1-c)-(1-h) and correspondences and large deformations. (2-c)-(2-h) are the registration results of the JAKRAK, JAKR,

A more challenging experiment is shown in Fig. 7, where SF, DF, AGS and FVNI methods. In general, visual inspection the hat distortion deformed all the letters, with the black shows that the JAKRAK and JAKR methods apparently stripes in `E' in particular having large local deformations. perform better than the other four methods. Due to the Moreover, the missing `I' in the reference image (Fig. 7(a)) intensity-based driving force having unexpected effects on appears in the moving image (Fig. 7(b)). The main chal-the brain tumor resection regions, the AGS and FVNI (Figs. lenge in this experiment lies in the reasonable alignment8(1-g)-(1-h), 8(2-g)-(2-h)) methods produce more or less of local small-scale structures such as the stripes in `E' excessive deformation diffusions within the tumor region Because many tiny structures are close to each other, onænd/or some non-smooth distortions across the tumor region structure's mismatching will directly affect the deformations boundaries (e.g., in the red boxes or in the red arrows), while

the SF and DF methods introduce some artifacts in certain 0.93 ± 0.63 , 0.97 ± 0.58 , 0.89 ± 0.55), while the registration edge structures (red arrows in Figs. 8(1-e)-(1-f)) as well aærrors of the DF, SF, JAKR and AGS methods are approxproduce an inappropriate and/or insuf cient contraction of imately $(1.11\pm0.51, 0.92\pm0.83, 0.92\pm0.66, 0.98\pm0.71,$ the tumor region and surrounding brain tissues (red boxes in $0.93\pm0.52, 0.97\pm0.62, 0.96\pm0.62$), $(1.25\pm1.23, 0.95\pm0.68,$ Figs. 8(2-e)-(2-f)). Some diffusion artifacts are clearly dis- $0.93\pm0.82, 1.13\pm0.87, 0.92\pm0.68, 1.05\pm0.57, 0.99\pm0.59$), played in the results of the FVNI method (Figs. 8(1-h) and $(1.43\pm0.87, 0.97\pm0.81, 1.16\pm0.85, 0.96\pm0.63, 0.96\pm0.61, 1.02\pm0.74, 0.92\pm0.58)$) and $(1.86\pm1.37, 1.08\pm0.92, 0.92\pm0.58)$

1.02 \pm 0.74, 0.92 \pm 0.58) and (1.86 \pm 1.37, 1.08 \pm 0.92, 1.05 \pm 0.79, 0.91 \pm 0.61, 1.01 \pm 0.53, 1.15 \pm 0.73, 0.94 \pm 0.63), respectively. The FVNI method cannot perform well in three cases of these seven challenging image registrations. The landmark-based registration errors printed in italic indicate that these methods produced unrealistic distortions and/or artifacts in the deformable registration results based on visual inspection (Figs. 5-8). Nevertheless, the registration errors printed in bold indicate that these methods achieved excellent performances in terms of both visual inspection and

FIGURE 8. Two cases of brain tumor image registration. The red arrows indicate unrealistic distortions and/or some artifacts in the results. (1-a)-(1-b) and (2-a)-(2-b) The reference and moving images, (1-c)-(2-c) JAKRAK, (1-d)-(2-d) JAKR, (1-e)-(2-e) SF, (1-f)-(3-f) DF, (1-g)-(2-g) AGS, (1-h)-(2-h) FVNI.

Table 1 demonstrates the landmark-based evaluation of all these methods in the above experiments. Although the landmark-based evaluation is unable to re ect the matching performances of the methods for the outlier structures, Table 1 compares the matching accuracy for the normal corresponding structures in terms of the average registration errors, with standard deviations of approximately 40-50 landmarks.

TABLE 1. Landmark registration errors (Mean+SD) of the six methods for the corresponding structures in the images. The registration errors printed in italic indicate that these methods produced unrealistic artifacts in the deformable registration results according to visual inspection, whereas the registration errors printed in bold indicate the methods that achieved excellent performances in terms of both visual inspection and landmark-based evaluation.

The JAKRAK method achieved satisfying registration performances for all seven experiments, with registration errors of $(1.23\pm0.81, 0.91\pm0.59, 0.93\pm0.81, 0.91\pm0.64,$



to achieve a steady transition from smooth deformation to[12] S. Heldmannet al., "An image registration framework for sliding motion accurate structure matching during the registration procedure.

In general, the JAKRAK method is an effective deformation construction method for accurately matching small structures and outlier structures with smooth defor-[13] mations compared with state-of-the-art methods. Many other deformable image registration methods for establishing accur14 rate structure correspondences exist and may conceivably be used instead of the block matching method with JAKRAK [15] for challenging image registration with missing correspondences and large local deformations. The challenging 2D/3D deformable image registration problem with missing cor-[16] respondences and large local deformations is well known171 to be far from solved in many research elds. At present, there is no doubt that methods and algorithms from intel-[18] ligent computing and machine learning for addressing this challenging outlier problem in deformable image registration are in high demand. Furthermore, we believe that fur-[19] ther experimental studies are required to build ground truth[20] M. Jin, R. Li, J. Jiang, and B. Qin, "Extracting contrast- lled vessels in 2D/3D image datasets with outlier structures.

ACKNOWLEDGMENT

The authors thank all the cited authors for providing the [22] B. W. Papie», M. P. Heinrich, J. Fehrenbach, L. Risser, and J. A. Schnabel, source codes used in this work. They are thankful to the anonymous reviewers for their valuable comments, which greatly helped to improve this paper.

REFERENCES

- image registration: A survey,"IEEE Trans. Med. Imag., vol. 32, no. 7, pp. 1153 1190, Jul. 2013.
- [2] D. F. P. Bouthemy and C. Kervrann, "Optical ow modeling and computation: A survey," Comput. Vis. Image Understand., vol. 134, pp. 1 21, May 2015.
- [3] X. Yang, R. Kwitt, M. Styner, and M. Niethammer, "Quicksilver: Fast predictive image registration A deep learning approach, Weuro Image,
- modalities and compensation for brain shift in tumor resection surgery, Int. J. Biomed. Imag., vol. 2017, Jun. 2017, Art. no. 6028645.
- [5] G. Wu, Q. Wang, H. Jia, and D. Shen, "Feature-based groupwise regis-tration by biasenthics." tration by hierarchical anatomical correspondence detection Brain Mapping, vol. 33, no. 2, pp. 253 271, 2012.
- [6] O. Lobachev, C. Ulrich, B. S. Steiniger, V. Wilhelmi, V. Stachniss, and M. Guthe, "Feature-based multi-resolution registration of immunostained serial sections," Med. Imag. Anal., vol. 35, pp. 288 302, Jan. 2017.
- [7] E. Suárez, C.-F. Westin, E. Rovaris, and J. Ruiz-Alzola, "Nonrigid registration using regularized matching weighted by local structure, Pinc. MICCAI, vol. 2489. 2002, pp. 581 589.
- [8] R. Stefanescu, X. Pennec, and N. Ayache, "Grid powered nonlinear image registration with locally adaptive regularization (Med. Imag. Anal., vol. 8, no. 3, pp. 325 342, 2004.
- [9] I. J. Simpsonet al., "Probabilistic non-linear registration with spatially adaptive regularisation, Med. Imag. Anal., vol. 26, no. 1, pp. 203 216, 2015.
- [10] M. Staring, S. Klein, and J. P. W. Pluim, "Nonrigid registration with tissuedependent Itering of the deformation eld," Phys. Med. Biol., vol. 52, no. 23, pp. 6879 6892, 2007.
- [11] D. F. Pace, S. R. Aylward, and M. Niethammer, "A locally adaptive regularization based on anisotropic diffusion for deformable image registration of sliding organs," IEEE Trans. Med. Imag., vol. 32, no. 11, pp. 2114 2126, Nov. 2013.

with piecewise smooth deformation," iRroceedings in Scale Space and Variational Methods in Computer Vision (Lecture Notes in Computer Science), vol. 9087, J.-F. Aujol, Eds. Berlin, Germany: Springer, 2015, pp. 335 347.

M. Foskey et al., "Large deformation 3D image registration in imageguided radiation therapy, Phys. Med. Biol., vol. 50, no. 24, pp. 5869 5892, 2005

M. Sdika and D. Pelletier, "Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping/h. Brain Mapping, vol. 30, no. 4, pp. 1060 1067, 2009.

- M. Brett, A. P. Leff, C. Rorden, and J. Ashburner, "Spatial normalization of brain images with focal lesions using cost function maskinguro Image, vol. 14, no. 2, pp. 486 500, 2005.
- P. Ripolléset al., "Analysis of automated methods for spatial normalization of lesioned brains,"NeuroImage, vol. 60, no. 2, pp. 1296 1306, 2012.
- M. Niethammeret al., "Geometric metamorphosis," inMedical Image Computing and Computer-Assisted Intervention MICCAI. Berlin, Germany: Springer, 2011, pp. 639 646.
- X. Liu, M. Niethammer, R. Kwitt, N. Singh, M. McCormick, and S. Aylward, "Low-rank atlas image analyses in the presence of pathologies," IEEE Trans. Med. Imag., vol. 34, no. 12, pp. 2583 2591, Dec. 2015.
- Z. Tang, Y. Wu, and Y. Fan, "Groupwise registration of MR brain images with tumors," Phys. Med. Biol., vol. 62, no. 17, pp. 6853 6868, 2017.
- X-ray angiography by graduated RPCA with motion coherency constraint," Pattern Recognit., vol. 63, pp. 653 666, Mar. 2017.
- [21] P. Risholm, F. Janoos, I. Norton, A. J. Golby, and W. Wells, III, "Bayesian characterization of uncertainty in intra-subject non-rigid registration," Med. Image Anal., vol. 17, no. 5, pp. 538 555, 2013.
 - An implicit sliding-motion preserving regularisation via bilateral Itering for deformable image registration, Med. Image Anal., vol. 18, no. 8, pp. 1299 1311, 2014.
- [23] L. Tang, G. Hamarneh, and R. Abugharbieh, "Reliability-driven, spatially adaptive regularization for deformable registration," Arco. Int. Workshop Biomed. Image Registration (WBIR), vol. 6204. 2010, pp. 173 185.
- [1] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical [24] T. Brox and J. Malik, "Large displacement optical ow: Descriptor matching in variational motion estimation,"IEEE Trans. Pattern Anal. Mach. Intell., vol. 33, no. 3, pp. 500 513, Mar. 2011.
 - J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, "DeepMatching: [25] Hierarchical deformable dense matching/t. J. Comput. Vis., vol. 120, no. 3, pp. 300 323, 2016.
 - [26] R. Timofte and L. Van Gool, "Sparse Flow: Sparse matching for small to large displacement optical ow," in Proc. IEEE Winter Conf. Appl. Comput. Vis., Jan. 2015, pp. 1100 1106.
- [4] S. Bayer, A. Maier, M. Ostermeier, and R. Fahrig, "Intraoperative imaging [27] M. F. Beg, M. I. Miller, A. Trouvé, and L. Younes, "Computing large J. Comput. Vis., vol. 61, no. 2, pp. 139 157, 2005.
 - B. Qin, Z. Shen, Z. Zhou, J. Zhou, and Y. Lv, "Structure matching driven by joint-saliency-structure adaptive kernel regression/pl. Soft Comput., vol. 46. pp. 851 867. Sep. 2016.
 - [29] B. Qin, Z. Gu, X. Sun, and Y. Lv, "Registration of images with outliers using joint saliency map,"IEEE Signal Process. Lett., vol. 17, no. 1, p. 91 94, Jan. 2010.
 - Z. Gu and B. Qin, "Nonrigid registration of brain tumor resection MR [30] images based on joint saliency map and keypoint clusteridg/sors, vol. 9, no. 12, pp. 10270 10290, 2009.
 - [31] X. Zhuang, S. Arridge, D. J. Hawkes, and S. Ourselin, "A nonrigid registration framework using spatially encoded mutual information and free-form deformations," IEEE Trans. Med. Imag., vol. 30, no. 10, pp. 1819 1828, Oct. 2011.
 - [32] Y. Ou, A. Sotiras, N. Paragios, and C. Davatzikos, "DRAMMS: Deformable registration via attribute matching and mutual-saliency weighting," Med. Image Anal., vol. 15, no. 4, pp. 622 639, 2011.
 - [33] A. Serag et al., "Construction of a consistent high-de nition spatiotemporal atlas of the developing brain using adaptive kernel regression," NeuroImage, vol. 59, no. 3, pp. 2255 2265, 2012.
 - [34] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods F. Trans. Pattern Anal. Mach. Intell., vol. 34, no. 11, pp. 2274 2281, Nov. 2012.
 - J. Rissanen, "Modeling by shortest data descriptionatica, vol. 14, [35] no. 5, pp. 465 471, 1978.

- of local scale," in Proc. 15th Int. Conf. Pattern Recognit. (ICPR), vol. 3. Sep. 2000, pp. 790 793.
- [37] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising tht. J. Comput. Vis., vol. 86, no. 1, pp. 1 32, 2010.
- [38] R. Gallea, E. Ardizzone, R. Pirrone, and O. Gambino, "Three-dimensional fuzzy kernel regression framework for registration of medical volume data," Pattern Recognit., vol. 46, no. 11, pp. 3000 3016, 2013.
- [39] J. Revaud, P. Weinzaepfel, Z. Harchaoui, and C. Schmid, ``EpicFlow: Edge-preserving interpolation of correspondences for optical ow," in Proc. IEEE CVPR, May 2015, pp. 1164 1172.
- [40] M. Köhler, A. Schindler, and S. Sperlich, "A review and comparison of [65] bandwidth selection methods for kernel regressidnt," Stat. Rev., vol. 82, no. 2, pp. 243 274, 2014.
- [41] A. Witkin, "Scale-space Itering: A new approach to multi-scale description," in Proc. Int. Joint Conf. Artif. Intell., vol. 9. 1983, pp. 1019 1023.
- [42] J. Koenderink, "The structure of images fiol. Cybern., vol. 50, no. 5, pp. 363 370, 1984.
- [43] J. Canny, "A computational approach to edge detection EE Trans. Pattern Anal. Mach. Intell., vol. PAMI-8, no. 6, pp. 679 698, Nov. 1986.
- [44] T. Lindeberg, "Scale-space theory: A basic tool for analyzing structures at different scales," J. Appl. Statist., vol. 21, nos. 1 2, pp. 225 270, 1994.
- [45] M. Zhang, T. Wu, and K. M. Bennett, "Small blob identi cation in medical images using regional features from optimum scale E Trans. Biomed. Eng., vol. 62, no. 4, pp. 1051 1062, Apr. 2015.
- [46] A. Basset, J. Boulanger, J. Salamero, P. Bouthemy, and C. Kervrann, "Adaptive spot detection with optimal scale selection in uorescence [70] microscopy images," IEEE Trans. Image Process., vol. 24, no. 11, pp. 4512 4527. Nov. 2015.
- [47] D. Ming, J. Li, J. Wang, and M. Zhang, "Scale parameter selection by spatial statistics for GeOBIA: Using mean-shift based multi-scale segmentation as an example, "SPRS J. Photogramm. Remote Sens., vol. 106, pp. 28 41, Aug. 2015.
- [48] J. Liu, S. Wang, E. B. Turkbey, M. G. Linguraru, J. Yao, and R. M. Summers, "Computer-aided detection of renal calculi from noncon-[72] trast CT images using TV- ow and MSER features Uted. Phys., vol. 42, no. 1, pp. 144 153, 2015.
- [49] C. Scharfenberger, A. Wong, and D. A. Clausi, "Structure-guided statistical textural distinctiveness for salient region detection in natural images," IEEE Trans. Image Process., vol. 24, no. 1, pp. 457 470, Jan. 2015.
- [50] K. Ikeuchi, Ed., Computer Vision: A Reference Guide. New York, NY, USA: Springer, 2014, pp. 701 713.
- [51] P. Perona and J. Malik, "Scale-space and edge detection using anisotrop diffusion," IEEE Trans. Pattern Anal. Mach. Intell., vol. 12, no. 7, pp. 629 639, Jul. 1990.
- [52] P. T. Jackway and M. Deriche, "Scale-space properties of the multiscale morphological dilation-erosion." IEEE Trans. Pattern Anal. Mach. Intell. vol. 18, no. 1, pp. 38 51, Jan. 1996.
- [53] T. Brox and J. Weickert, "A TV ow based local scale estimate and its application to texture discrimination,J. Vis. Commun. Image Represent., vol. 17, no. 5, pp. 1053 1073, 2006.
- [54] P. Burt and E. Adelson, "The Laplacian pyramid as a compact image From 2012 to 2013, he was a Visiting Professor with the Department of code," IEEE Trans. Commun., vol. COM-31, no. 4, pp. 532 540,
- blur estimation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 7, pp. 699 716, Jul. 1998.
- contours for unsupervised local noise estimation EEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 12, pp. 2379 2392, Dec. 2012.
- [57] D. Liu et al., "Medical image classi cation using spatial adjacent histogram based on adaptive local binary patternsomput. Biol. Med., vol. 72, pp. 185 200, May 2016.
- [58] J. Fdez-Valdivia, J. A. Garcia, J. Martinez-Baena, and X. R. Fdez-Vidal, The selection of natural scales in 2D images using adaptive Gabor Itering," IEEE Trans. Pattern Anal. Mach. Intell., vol. 20, no. 5, pp. 458 469, May 1998.
- [59] B.-W. Hong, S. Soatto, K. Ni, and T. Chan, "The scale of a texture and its application to segmentation," inProc. IEEE Conf. CVPR, Jun. 2008, pp. 18.
- [60] X. Zhang and S. Du, "Learning selfhood scales for urban land cover mapping with very-high-resolution satellite images Pemote Sens. Environ., vol. 178, pp. 172 190, Jun. 2016.

- [36] G. Gomez, J. L. Marroquin, and L. E. Sucar, "Probabilistic estimation [61] A. Souza, J. K. Udupa, and A. Madabhushi, "Image Itering via generalized scale,"Med. Image Anal., vol. 12, no. 2, pp. 87 98, 2008. J. Sun and Z. Xu, "Scale selection for anisotropic diffusion Iter
 - [62] by Markov random eld model," Pattern Recognit., vol. 43, no. 8, pp. 2630 2645, 2010.
 - [63] Z. Ji, Y. Xia, Q. Sun, Q. Chen, and D. Feng, "Adaptive scale fuzzy local Gaussian mixture model for brain MR image segmentation, Neurocomputing, vol. 134, pp. 60 69, Jan. 2014.
 - A. Pai, S. Sommer, L. Sørensen, S. Darkner, J. Sporring, and M. Nielsen, [64] Kernel bundle diffeomorphic image registration using stationary velocity elds and Wendland basis functions/EEE Trans. Med. Imag., vol. 35, no. 6, pp. 1369 1380, Jun. 2016.
 - J. Kim, C. Liu, F. Sha, and K. Grauman, "Deformable spatial pyramid matching for fast dense correspondences,"Pinoc. IEEE Conf. CVPR, Jun. 2013, pp. 2307 2314.
 - [66] M. Tau and T. Hassner, "Dense correspondences across scenes and scales," IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 5, pp. 875 888, May 2016.
 - [67] M. P. Heinrich, I. J. A. Simpson, B. W. Papie», S. M. Brady, and J. A. Schnabel, ``Deformable image registration by combining uncertainty estimates from supervoxel belief propagation del Image Anal., vol. 27, pp. 57 71, Jan. 2016.
 - [68] R. Castillo et al., "A framework for evaluation of deformable image registration spatial accuracy using large landmark point sets/s. Med. Biol., vol. 54, no. 7, pp. 1849 1870, 2009.
 - [69] K. Murphy et al., "Evaluation of registration methods on thoracic CT: The EMPIRE10 challenge,"IEEE Trans. Med. Imag., vol. 30, no. 11, pp. 1901 1920, Nov. 2011.
 - Y. Ou, H. Akbari, M. Bilello, X. Da, and C. Davatzikos, "Comparative evaluation of registration algorithms in different brain databases with varying dif culty: Results and insights,"IEEE Trans. Med. Imag., vol. 33, no. 10. pp. 2039 2065. Oct. 2014.
 - [71] B. B. Avants, N. J. Tustison, G. Song, P. A. Cook, A. Klein, J. C. Gee, "A reproducible evaluation of ANTs similarity metric performance in brain image registration,"NeuroImage, vol. 54, no. 3, pp. 2033 2044, 2011.
 - R. Werner, A. Schmidt-Richberg, H. Handels, and J. Ehrhardt, "Estimation of lung motion elds in 4D CT data by variational non-linear intensitybased registration: A comparison and evaluation stuBty s. Med. Biol., vol. 59, no. 15, pp. 4247 4260, 2014.



BINJIE QIN received the M.Sc. degree in measuring and testing technologies and instruments from the Nanjing University of Science and Technology, Nanjing, and the Ph.D. degree in Biomedical Engineering from the Shanghai Jiao Tong University, Shanghai, China, in 1999 and 2002, respectively. He was a Lecturer and also an Associate Professor with the Department of Biomedical Engineering, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China.

Computer Science, University College London, U.K. He is currently an

[55] J. H. Elder and S. W. Zucker, "Local scale control for edge detection and the Tree Hutter in the School of Biomedical Engineering, Shanghai Jiao Tong University, Shanghai, China. His group has developed image guided surgery system and small animal optical imaging system that are [56] B. Kerautret and J.-O. Lachaud, "Meaningful scales detection along digital successfully used in clinical applications. His current research interests include biomedical imaging, image processing, machine learning, computer vision, and biomedical instrumentation.

> ZHUANGMING SHEN received the B.S. degree in biomedical engineering from Tianjin Medical University, Tianjin, and the M.S. degree in biomedical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2010 and 2013, respectively. His current research interests include medical image processing and visualization, MRI data analysis, and 3-D printing.

IEEEAccess



ZESHAN FU received the B.S. degree in biomedical engineering from Nanchang University, Nanchang, China, in 2014, and he is currently pursuing the master's degree with the School of Biomedical Engineering, Shanghai Jiao Tong University. His current research interests include medical image registration and computer vision.



YISONG LV received the M.Sc. degree in automatic instruments from the Kunming University of Science and Technology, Kunming, and the Ph.D. degree in biomedical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1999 and 2003, respectively. His current research interests include image analysis, machine learning, and computer vision.



ZIEN ZHOU received the B.Eng. degree in biomedical engineering, and the M.D. degree in imaging and nuclear medicine from Shanghai Jiao Tong University, Shanghai, China, in 2007 and 2011, respectively. His current research interests include medical imaging, clinical trials, and systematic reviews of randomize control trials. JINSONG BAO received the M.Sc. degree in mechanical engineering from Northeastern University, Shenyang, and the Ph.D. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1999 and 2002, respectively. His current research interests include intelligent manufacturing, machine learning, and computer vision.

...