NON-RIGID POINT SET REGISTRATION FOR CHINESE CHARACTERS USING STRUCTURE-GUIDED COHERENT POINT DRIFT

Hao Sun[⋆†] *Zhouhui Lian*[*⋆†] *Yingmin Tang*[⋆†] *Jianguo Xiao*[⋆†]

[⋆] Institute of Computer Science and Technology, Peking University, Beijing, PR China
[†] Center for Chinese Font Design and Research, Peking University, Beijing, PR China

ABSTRACT

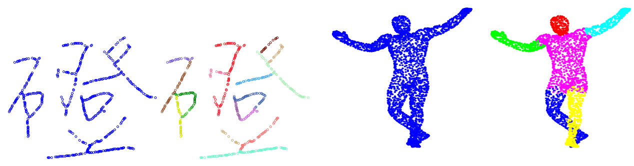
This paper proposes a non-rigid point set registration method called Structure-Guided Coherent Point Drift (SGCPD). The key idea of our method is to utilize structural information and combine the global and local point registrations together to improve the original Coherent Point Drift (CPD) algorithm. Specifically, given two point sets, we first align them using the CPD method with Localized Operator (CPDLO). Then we divide the target point set into several subsets and apply CPDLO to each subset. Finally, we implement the above two procedures until convergence. In this manner, more detailed information can be well exploited and thus higher registration accuracy can be achieved. Experimental results demonstrate that our method outperforms the original CPD approach on both point registration accuracy and skeleton decomposition accuracy for Chinese characters.

***Index Terms*—** Point set registration, Coherent Point Drift, Chinese character skeleton decomposition

1. INTRODUCTION

Point set registration is the process of finding a spatial transformation that aligns two point sets. Typically, the transformed point set is called model point set while the other one is named target point set. Point set registration plays an important role in pattern recognition and image processing. It is also a crucial step for many applications such as fingerprint recognition, motion tracking and computer-aided surgery [1, 2]. One tough task among them is the skeleton point registration of Chinese characters, which can be widely used in stroke extraction, shape morphing, etc [3, 4]. Current point set registration algorithms are not effective enough for the skeleton point registration of Chinese characters. That is mainly because existing methods all treat the task as a normal point set registration problem without taking the special properties of Chinese characters into account. One most important property of Chinese characters is their structural information. Here, the structural information denotes the subset division of model point set. It is a widespread feature for both 2D and 3D point sets, which is demonstrated in Fig. 1.

*Zhouhui Lian is the corresponding author.



(a) (b)

Fig. 1. Demonstration of structural information of model point set in 2D (a) and 3D (b), respectively.

There exists a large body of research work on point set registration. A well-known classic method is the Iterated Closest Point (ICP) algorithm [5, 6], which iteratively updates the correspondence until reaching the local minimum. The Thin-Plate Spline Robust Point Matching (TPS-RPM) [7, 8] was proposed in order to solve ICP's occlusion problem. One state-of-the-art approach is the Coherent Point Drift (CPD) [9] algorithm, which employs Gaussian Radial Basis Functions (GRBF) instead of TPS to get better registration results for high-dimensional point sets. Recently, several variants of CPD have been proposed to meet specific requirements. Hu *et al.* [10] added landmark information in the registration, while [11] and [12] focused on automatic parameter selection and outlier modeling.

Essentially, CPD and its existing variants are all global registration algorithms and thus have the drawback of neglecting local detailed features. In this paper, we propose a novel method called Structure-Guided Coherent Point Drift (SGCPD), which utilizes structural information and combines the global and local point registrations together to improve the original CPD algorithm. Experimental results demonstrate the effectiveness and superiority of our method in applications of non-rigid point set registration for Chinese characters.

2. METHOD DESCRIPTION

2.1. Mathematical Background

The CPD framework [9] is a probabilistic method which considers the point registration problem as a probability estimation problem. The model points represent the Gaussian Mixture Model (GMM) centroids, while the target points represent data points. The Expectation Maximization (EM) [13] algorithm is employed to maximize the likelihood function and optimally align the model point set to the target point set.

Let $\mathbf{X}_{N\times D} = (\mathbf{x}_1, ..., \mathbf{x}_N)^T$ be the target point set with N points and D dimensions, and $\mathbf{Y}_{M\times D} = (\mathbf{y}_1, ..., \mathbf{y}_M)^T$ be the GMM centroids set with M points. Thus, the GMM probability density function is defined as

$$p(\mathbf{x}) = \omega \frac{1}{N} + (1-\omega)\sum_{m=1}^{M} \frac{1}{M} p(\mathbf{x}|m), \tag{1}$$

in which the uniform distribution $p(x|M+1) = \frac{1}{N}$ with the weight of ω accounts for noise and outliers. Each Gaussian distribution has equal membership probability $\frac{1}{M}$ and takes the form $p(\mathbf{x}|m) = \frac{1}{(2\pi\sigma^2)^{D/2}} \exp^{-\frac{\|\mathbf{x}-\mathbf{y}_m\|^2}{2\sigma^2}}$, where σ^2 is the equal isotropic covariances of Gaussian distribution.

Then the model points' locations can be re-parameterized by a set of parameters θ, which can be estimated by minimizing the following negative log-likelihood function

$$E(\theta, \sigma^2) = -\sum_{n=1}^{N} \log \sum_{m=1}^{M+1} \frac{1}{M} p(\mathbf{x}|m). \tag{2}$$

The EM algorithm can be used to determine θ and σ in (2) (see more details in Algorithm 1). The CPD algorithm iteratively implements E and M steps until convergence.

Yet another problem is how to define non-rigid transformation for model point set. According to the Tikhonov regularization framework [14], the transformation is defined as $\mathcal{T}(\mathbf{Y}, v) = \mathbf{Y} + v(\mathbf{Y})$, where v is a displacement function. Based on [9], the optimal displacement function $v(\mathbf{z})$ follows the form

$$v(\mathbf{z}) = \sum_{m=1}^{M} \mathbf{w}_m G(\mathbf{z}, \mathbf{y}_m) + \psi(\mathbf{z}), \tag{3}$$

where the coefficients $\mathbf{w}_m$ can be evaluated by the following formula at $\mathbf{y}_m$ points

$$(\mathbf{G} + \lambda\sigma^2 d(\mathbf{P1}^{-1})\mathbf{W} = d(\mathbf{P1}^{-1})\mathbf{PX} - \mathbf{Y}, \tag{4}$$

where $\mathbf{P}$ is the correspondence matrix calculated in E step. According to the Motion Coherence Theory (MCT) [15], matrix $\mathbf{G}_{M\times M}$ can be calculated by

$$g_{ij} = G(\mathbf{y}_i, \mathbf{y}_j) = e^{-\frac{1}{2}\|\frac{\mathbf{y}_i - \mathbf{y}_j}{\beta}\|^2}, \tag{5}$$

which represents the affinity among points.

2.2. Structure-Guided CPD

As shown in Fig. 2, the proposed SGCPD algorithm is an iterative procedure which consists of the following three substeps: CPDLO, target point set decomposition and localized

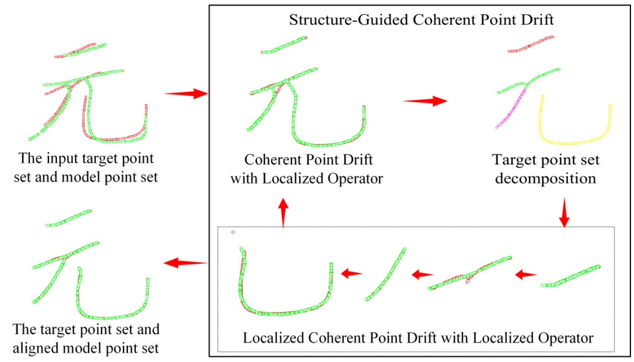


Fig. 2. Overview of our method.

CPDLO. In this section, we first give a formal description for the structural information and describe each sub-step elaborately. Then we summarize the SGCPD algorithm and provide a discussion.

In SGCPD, the target point set $\mathbf{X}$ and model point set $\mathbf{Y}$ are the same as those in the original CPD. The additional structural information is that the model point set can be divided into K subsets, namely $\mathbf{Y}_{M\times D} = \bigcup_{i=1}^{K} \mathbf{Y}^i_{M_i\times D}$, where $M = \sum_{i=1}^{K} M_i$. We introduce the structure vector to describe structural information. Here, the structure vector is defined as $\mathbf{s}^{\mathbf{Y}} = \{s_1^{\mathbf{Y}}, \ldots, s_M^{\mathbf{Y}}\}$ where $s_i^{\mathbf{Y}} = k$ when $\mathbf{y}_i \in \mathbf{Y}^k$.

In MCT [15], the velocity field is defined everywhere in the image, which causes a problem that the registration accuracy of some points may be unsatisfactory since the movement of a point is influenced by all other points. In fact, a point's movement should be mainly influenced by the points that are in the same subset. To solve this problem, we define a velocity field for each model subset and introduce a parameter ξ to represent the impact of other velocity fields. We implement the above ideas by multiplying a localized operator to affinity matrix $\mathbf{G}$. The localized operator $\mathbf{L}_{M\times M}$ is defined as

$$l_{ij} = \begin{cases} 1, & if\ s_i^{\mathbf{Y}} = s_j^{\mathbf{Y}} \\ \xi, & otherwise \end{cases}, \tag{6}$$

and it can be easily calculated using the structure vector $\mathbf{s}^{\mathbf{Y}}$. Then the localized affinity matrix $\mathbf{G}$ can be calculated by

$$g_{ij} = G(\mathbf{y}_i, \mathbf{y}_j) = l_{ij} e^{-\frac{1}{2}\|\frac{\mathbf{y}_i - \mathbf{y}_j}{\beta}\|^2}. \tag{7}$$

Algorithm 1 shows the pseudo code of our CPDLO algorithm. Note that the only difference between CPDLO and the original CPD algorithm is the calculation of $\mathbf{G}$.

After implementing CPDLO, we get the correspondence matrix $\mathbf{P}_{N\times M}$, in which p_{ij} denotes the correspondence probability of the ith target point and the jth model point. Given a target point $\mathbf{x}_i$, we denote its correspondence model point as $\mathbf{y}_{c_{\mathbf{x}_i}}$, in which

$$c_{\mathbf{x}_i} = \arg\max_{m\in[1,M]} \{p_{im}\}. \tag{8}$$

Algorithm 1 CPDLO

$\mathbf{W} \leftarrow 0$; $\sigma^2 \leftarrow \frac{1}{DNM}\sum_{m,n=1}^{M,N} \parallel \mathbf{x}_n - \mathbf{y}_m \parallel^2$;
$\mathbf{G}(\mathbf{y}_i, \mathbf{y}_j) \leftarrow l_{ij} e^{-\frac{1}{2}\parallel \frac{\mathbf{y}_i - \mathbf{y}_j}{\beta} \parallel^2}$;
while $l > l_{thr}$ and $\sigma > \sigma_{thr}$ **do**
    $p_{mn} \leftarrow \frac{\exp^{-\frac{1}{2}\parallel \frac{\mathbf{x}_n - (\mathbf{y}_m + \mathbf{G}_m \cdot \mathbf{W})}{\sigma} \parallel^2}}{\sum_{k=1}^{M} \exp^{-\frac{1}{2}\parallel \frac{\mathbf{x}_n - (\mathbf{y}_k + \mathbf{G}_k \cdot \mathbf{W})}{\sigma} \parallel^2} + (2\pi\sigma^2)^{D/2} \frac{\omega}{1-\omega} \frac{M}{N}}$;
    Solve $(\mathbf{G} + \lambda\sigma^2 d(\mathbf{P1})^{-1})\mathbf{W} = d(\mathbf{P1})^{-1}\mathbf{PX} - \mathbf{Y}$;
    $\mathbf{T} \leftarrow \mathbf{Y} + \mathbf{G} \cdot \mathbf{W}$;
    $N_{\mathbf{P}} \leftarrow \mathbf{1}^T\mathbf{P1}$;
    $\sigma^2 \leftarrow \frac{1}{N_{\mathbf{P}}D}(tr(\mathbf{X}^T d(\mathbf{P}^T\mathbf{1})\mathbf{X}) - 2tr((\mathbf{PX})^T\mathbf{T}) + tr(\mathbf{T}^T d(\mathbf{P1})\mathbf{T}))$;
end while
return $\mathbf{T}, \mathbf{P}$;

Now for the target point set $\mathbf{X}$, the structure vector $\mathbf{s}^{\mathbf{X}}$ takes the form $s_i^{\mathbf{X}} = s_{c_{\mathbf{x}_i}}^{\mathbf{Y}}$. Thus, the target point set can be divided into K subsets.

Now we get a one-to-one mapping from target point subsets to model point subsets. Current state is a globally optimal state since the global negative log-likelihood function is minimized by the model points. However for each target-model subset pair, current model points' spatial distribution might not be optimal. Hence, our key observation for the next step is to reorganize model points in each subset to make them locally optimal. To achieve this, we apply CPDLO to each pair of subsets, which is called localized CPDLO.

After localized CPDLO, all subset pairs are aligned. Note that the localized CPDLO breaks the globally optimal state of model points. Thus, it is essential to conduct a global CPDLO to the entire model points, which can again make model point set globally optimized. By combining CPDLO, target point set decomposition and localized CPDLO together, we can get an integrated iterative procedure, which is called SGCPD algorithm. When the SGCPD algorithm is converged, the model points are in a both globally and locally optimal state. Compared with the globally optimal state of the original CPD method, the registration accuracy can be markedly improved. Algorithm 2 shows the pseudo code for SGCPD.

Algorithm 2 SGCPD

Input:
Target point set $\mathbf{X}$, model point set $\mathbf{Y}$, model structure vector $\mathbf{s}^{\mathbf{Y}}$, noise rate ω, regularization parameters λ and β, localized weight ξ, iteration threshold i_{thr}, likelihood threshold l_{thr}, σthreshold σ_{thr}.
Output:
Aligned point set $\mathbf{T}$, correspondence matrix $\mathbf{P}$.

$iter \leftarrow 0$;
Compute localized operator $\mathbf{L}$ using $\mathbf{s}^{\mathbf{Y}}$ and ξ.
while $iter < i_{thr}$ **do**
    $\{\mathbf{T}, \mathbf{P}\} \leftarrow$ CPDLO $(\mathbf{X}, \mathbf{Y}, \mathbf{L}, \omega, \lambda, \beta, l_{thr}, \sigma_{thr})$;
    for $i \in \{1, \ldots, N\}$ **do**
        $pmax \leftarrow -\infty$;
        for $j \in \{1, \ldots, M\}$ **do**
            if $p_{ij} > pmax$ **then**
                $pmax \leftarrow p_{ij}$; $c_{\mathbf{x}_i} \leftarrow j$; $s_i^{\mathbf{X}} \leftarrow s_{c_{\mathbf{x}_i}}^{\mathbf{Y}}$;
            end if
        end for
        $\mathbf{X}^{s_i^{\mathbf{X}}} \leftarrow \mathbf{X}^{s_i^{\mathbf{X}}} \bigcup \mathbf{x}_i$;
    end for
    for $k \in \{1, \ldots, K\}$ **do**
        Compute $\mathbf{L}^k$ for the kth subset.
        CPDLO $(\mathbf{X}^k, \mathbf{Y}^k, \mathbf{L}^k, \omega, \lambda, \beta, l_{thr}, \sigma_{thr})$;
    end for
    $iter \leftarrow iter + 1$;
end while
return $\mathbf{T}, \mathbf{P}$;

3. EXPERIMENTAL RESULTS

In this section, we present experimental results on point set registration and character skeleton decomposition application. Due to the page limit only a small part of results are shown here, the full experimental results can be found at `www.github.com/sunhao2014/SGCPD/`.

3.1. Results on Point Set Registration

We carried out experiments on five Chinese font libraries (i.e., Kaiti (KT), Hanyi Kaiti (HK), Fongsong (FS), Hard-tripped Kaiti (HTK) and Founder Jinglei (FJ)). We chose KT as model font and the other four as target fonts. 100 characters were randomly selected from GB2312 Chinese character set, and the selected 100 characters' skeleton point sets were extracted using [16] for model font and each target font. Note that in order to test our method's robustness, we chose four target fonts quite different in style. HK and FS are printed fonts whose structures are relatively close to KT while HTK and FJ are handwriting fonts whose structures are quite different from KT.

To evaluate the accuracy of point set registration, we define Average Assignment Probability (AAP) as $AAP(\mathbf{X}, \mathbf{T}) = \frac{\sum_{i=1}^{N} \max_{1 \leqslant j \leqslant M}\{p_{ij}\}}{N}$. Fig. 3 provides the comparison of Chinese character skeleton point set registration between the original CPD and our method. We observe that our method outperforms the original CPD method in terms of registration accuracy for both printed and handwriting fonts.

3.2. Results on Character Skeleton Decomposition

After point set registration, the target points can be divided into several subsets, which make up the skeleton decomposition results. Since character skeleton decomposition is often used for stroke decomposition [3, 4], the false match of some

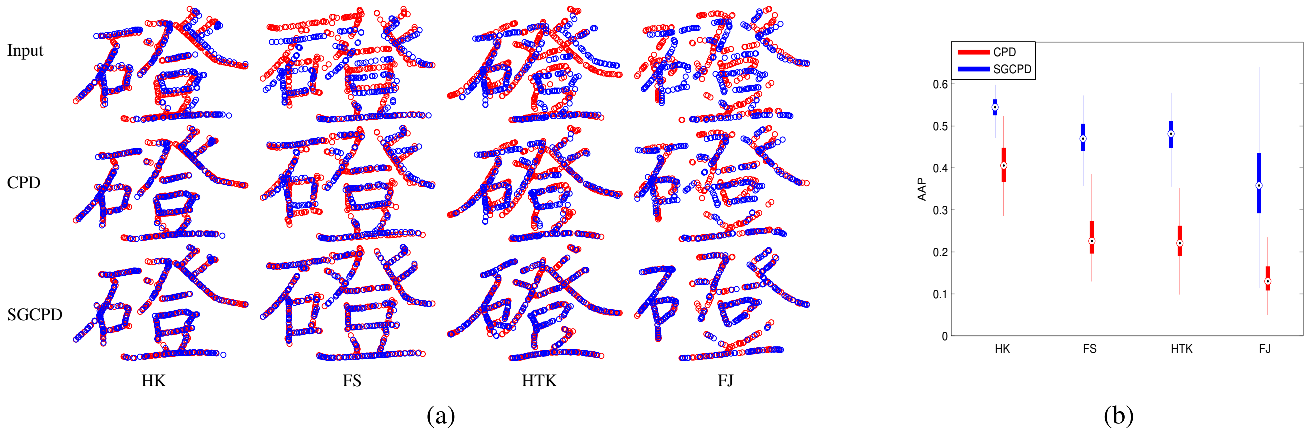


Fig. 3. Point set registration results. (a) shows registration results for the Chinese character "Deng" in four different styles. (b) is the statistical box plot for 4 data sets. Each set of box bars indicate the maximum, 75th percentile, median, 25th percentile and minimum AAPs for each data set with CPD in red and SGCPD in blue.

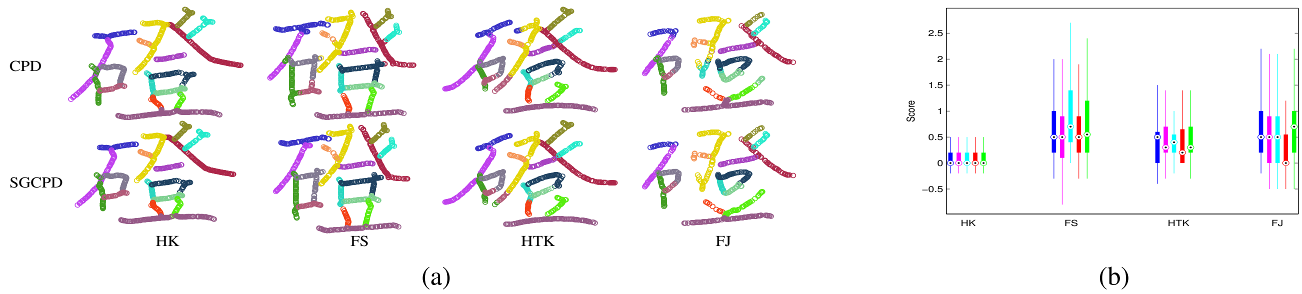


Fig. 4. Chinese character skeleton decomposition results. (a) shows the decomposition results for the Chinese character "Deng" in four different styles. (b) is the statistical box plot for four data sets, and each person has a specific color.

crucial points would lead to serious damages on the whole stroke structure. Thus it is improper to simply use the error rate to evaluate decomposition accuracy. We designed a scoring system and got assessment results from five participants. There are four levels in the scoring system, which are significant optimization, slight optimization, slight deterioration and significant deterioration, with the score of 0.5, 0.2, -0.2, -0.5 respectively. For each character in data sets, each participant estimated the differences between decomposition results of the original CPD and SGCPD, and assigned a level to each difference. If the score summation is positive, we could conclude that our method performs better than CPD. Fig. 4 shows final results for character skeleton decomposition, which indicates that our method can effectively improve the skeleton decomposition accuracy for Chinese characters.

4. CONCLUSION

In this paper, we introduce a structure-guided non-rigid point set registration algorithm. We take advantage of the structural information of model point set by adding the localized operator and combining global and local point set registration together. Besides the Chinese character skeleton decomposition, our method can have many other applications. Taking the hot research topic of 3D model decomposition [17] as an example, the proposed method which utilizes structural information can deal with the problem efficiently. Experimental results demonstrate that the proposed approach performs better than the original CPD method in terms of registration accuracy and character skeleton decomposition accuracy.

5. ACKNOWLEDGEMENT

This work was supported by National Natural Science Foundation of China (Grant No.: 61202230), National Hi-Tech Research and Development Program (863 Program) of China (Grant No.: 2014AA015102) and China Postdoctoral Science Foundation (Grant No.: 2013T60038).

6. REFERENCES

[1] Lisa Gottesfeld Brown, “A survey of image registration techniques,” *ACM computing surveys (CSUR)*, vol. 24, no. 4, pp. 325–376, 1992.

[2] Barbara Zitova and Jan Flusser, “Image registration methods: a survey,” *Image and vision computing*, vol. 21, no. 11, pp. 977–1000, 2003.

[3] Chenxi Wang, Zhouhui Lian, Yingmin Tang, and Jianguo Xiao, “Automatic correspondence finding for chinese characters using graph matching,” in *Image and Graphics (ICIG), 2013 Seventh International Conference on*. IEEE, 2013, pp. 545–550.

[4] Zhouhui Lian and Jianguo Xiao, “Automatic shape morphing for chinese characters,” in *SIGGRAPH Asia 2012 Technical Briefs*. ACM, 2012, p. 2.

[5] Paul J Besl and Neil D McKay, “Method for registration of 3-d shapes,” in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[6] Zhengyou Zhang, “Iterative point matching for registration of free-form curves and surfaces,” *International journal of computer vision*, vol. 13, no. 2, pp. 119–152, 1994.

[7] Haili Chui and Anand Rangarajan, “A new algorithm for non-rigid point matching,” in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. IEEE, 2000, vol. 2, pp. 44–51.

[8] Haili Chui and Anand Rangarajan, “A new point matching algorithm for non-rigid registration,” *Computer Vision and Image Understanding*, vol. 89, no. 2, pp. 114–141, 2003.

[9] Andriy Myronenko and Xubo Song, “Point set registration: Coherent point drift,” *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 32, no. 12, pp. 2262–2275, 2010.

[10] Yipeng Hu, Erik-Jan Rijkhorst, Richard Manber, David Hawkes, and Dean Barratt, “Deformable vessel-based registration using landmark-guided coherent point drift,” in *Medical Imaging and Augmented Reality*, pp. 60–69. Springer, 2010.

[11] Peng Wang, Ping Wang, ZhiGuo Qu, YingHui Gao, and ZhenKang Shen, “A refined coherent point drift (cpd) algorithm for point set registration,” *Science China Information Sciences*, vol. 54, no. 12, pp. 2639–2646, 2011.

[12] Yuan Gao, Jiayi Ma, Ji Zhao, Jinwen Tian, and Dazhi Zhang, “A robust and outlier-adaptive method for non-rigid point registration,” *Pattern Analysis and Applications*, pp. 1–10, 2013.

[13] Arthur P Dempster, Nan M Laird, and Donald B Rubin, “Maximum likelihood from incomplete data via the em algorithm,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 1–38, 1977.

[14] Andreĭ Tikhonov, *Numerical methods for the solution of ill-posed problems*.

[15] Alan L Yuille and Norberto M Grzywacz, “A mathematical analysis of the motion coherence theory,” *International Journal of Computer Vision*, vol. 3, no. 2, pp. 155–175, 1989.

[16] TY Zhang and Ching Y. Suen, “A fast parallel algorithm for thinning digital patterns,” *Communications of the ACM*, vol. 27, no. 3, pp. 236–239, 1984.

[17] Arjun Jain, Thorsten Thormählen, Tobias Ritschel, and Hans-Peter Seidel, “Exploring shape variations by 3d-model decomposition and part-based recombination,” in *Computer Graphics Forum*. Wiley Online Library, 2012, vol. 31, pp. 631–640.