Task: Vectorization of scanned document



- Improved reading experience and photocopying quality; reduced file size
- **Difficulty**: Low scanning image resolution and severe loss of character details
- **Observation**: The same character will appear multiple times in the document



copying quality; reduced file size on and severe loss of character details opear multiple times in the document

Algorithm Overview





1. Character extraction based on connected domain

IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 14, NO. 2, FEBRUARY 1992

global shape matching or registration have addressed limited polynomials. This demonstrated some interesting theoretical classes of shapes, namely, 1) polyhedral models, 2) piecewise- results but remains to be demonstrated for practical use on (super)quadric models [2], [27], and 3) point sets with known complex surfaces. correspondence. The reader may consult [6] and [14] for pre- Szeliski [54] also describes a method for estimating motion 1985 work in these areas. For a sampling of other more recent from sparse range data without correspondence between the related work not discussed below, see [8], [10], [12], [13], points and without feature extraction. His primary goal was [19], [20], [24], [26], [34], [35], [37], [39], [44], [46], [48], to create a method for estimating the motion of the observer [53], [58], [59].

done earliest by Faugeras and his group at INRIA [18], where assumption to create a smoothing spline approximation of the they demonstrated effective matching with a Renault auto part points. Then, a conventional steepest descent algorithm is used the use of quaternions for least squares registration of cor- the sum of the covariance-weighted z differences between the responding 3-D point sets in the computer vision community. points and the surface. His approach is based on a regular The alternative use of the singular value decomposition (SVD) xy-grid structure, and true 3-D point-to-surface distances are algorithm [23], [1], [49] was not as widely known in this time not computed. The steepest-descent approach is a slower on the probable existence of reasonably large planar regions ICP algorithm described below. Szeliski uses optimal Bayesian within a free-form shape.

form space curve matching problem without feature extraction range imaging sensors, the uncertainty in data points vary in late 1985. They used a nonquaternion approach to comput- significantly from the foreground to the background. For highing the least squares rotation matrix. The method works well accuracy sensors with shallow depths of field, the uncertainty with reasonable quality curve data but has difficulty with very variation between points is orders of magnitude less and is noisy curves because the method uses arclength sampling of of much less concern. Szeliski provides experimental results the curves to obtain corresponding point sets.

timation problem using robust methods combined with the problem. His block test did involve six degrees of freedom, least squares SVD registration approach, which provided a but the block is a very simple shape. Overall, this work robust statistical alternative to the least squares quaternion or presents some interesting ideas, but the experimental results SVD point set matching. This algorithm is able to handle are unconvincing for applications. statistical outliers and could theoretically be substituted for Horn and Harris [33] also addressed the problem of esour quaternion-based algorithm as long as the determinant of timating the exact rigid-body motion of the observer given the orthonormal matrix is strictly a positive one. A recent sequentially digitized range image frames of the same terrain. conference proceedings [47] contains new contributions on They describe a range rate constraint equation and an elevation this subject.

method [18] of least squares quaternion matching that uses estimate as long as the motion between frames of data is the maximum eigenvalue of a 4×4 matrix instead of the relatively small. This method is much quicker than the one minimum eigenvalue. Horn [30] and Brou [11] also developed proposed by Szeliski, but it is not clear that this method the extended Gaussian image (EGI) methods allowing the generalizes to arbitrary rotations and translations of a shape. matching of convex and restricted sets of nonconvex shapes Kamgar-Parsi et al. [36] also describe a method for the based on surface normal histograms.

implicit algebraic nonplanar 3-D curve and surface estimation using the level sets of 2.5-D range data but is essentially with applications to position estimation without feature ex- restricted to the three degrees of freedom in the plane since the traction. He describes a method of approximating data points work was addressed toward piecing together terrain map data. with implicit algebraic forms up to the tenth degree using Li [38] addressed free-form surface matching with arbitrary an approximate distance metric. Global shapes (not occluded rotations and translations. His method forms an attributed shapes) can be identified based on generalized eigenvalues, and relational graph of fundamental surface regions for data and the registration transformation can be recovered. The method model shapes and then performs graph matching using an is shown to be useful for complete planar curve and space inexact approach that allows for variability in attributes as curve shapes, but it is unclear that the effectiveness generalizes well as in graph adjacency relationships. This seems to be well to more complicated surfaces, such as terrain data or a a reasonable approach but relies on extraction of derivativehuman face. Taubin has stated that the numerical methods of based quantities. Experimental results are shown for a coffee the approximate distance fit tend to break down above the tenth cup and the Renault auto part; see also Wong et al. [60] for

free-form shapes. Most of the existing literature addressing by investigating shape matching based on generalized shape

between two range image frames of the same terrain. Given Historically, free-form shape matching using 3-D data was the set of points from one frame, he applies a smoothness (steering knuckle) in the early 1980's. This work popularized to rotate and translate the second data set so that it minimizes frame. The primary limitation of this work was that it relied alternative to reaching the local minima than our proposed mathematics to allow him to downweight noisier values at Schwartz and Sharir [50] developed a solution to the free- longer ranges from a simulated range finder. For navigation for synthetic terrain data and a block. The terrain data motion Haralick et al. [28] addressed the 3-D point-set pose es- test was a simple translation along one axis: a 1-D correlation

rate constraint equation. The result is a noniterative least Horn [31] derived an alternative formulation of Faugeras's squares method that provides a six-degree-of-freedom motion

registration of multiple overlapping range images without Taubin [55] has done some interesting work in the area of distinctive feature extraction. This method works very well degree. He later [56] extended his work in shape description other related work using attributed graphs for 3-D matching.

Scan document images



2. Clustering based on character similarity





3. Merge-based super-resolution



4. Sketching based on Potrace



Merged Character

Vector Character



5. Result & Comparison





Adobe Illustrator



This method



Remained Issue: Loss of Serif Details



Merged Image

Binarization Vectorization

Over-Smoothen Serifs

Improved Based on Machine Learning









Target Vector



Construct Training-Set: Merged Image X_i



SVG Vector Character

Low-res Noisy Images

Construct Training-Set: Target Vector Y_i

Y_i is represented in a format resembling the Marching-Square Algorithm



SVG Vector Character Divide into H × W Grid



Each grid cell corresponds to a Case above



Construct Training-Set: Target Vector Y;





Optimize vertex coordinates

Encode to fixed-length vector representation (length = 12)

H × W cells in total, so the entire grid is represented as a H × W × 12 tensor



Construct Training-Set: Target Vector *Y*_{*i*}







SVG Vector Character





Learn the mapping of $X_i \rightarrow Y_i$ with CNN









Results & Comparison



Input





Reference

[1] PyMuPDF. 2023. An enhanced Python binding for MuPDF, a lightweight PDF, XPS, and E-book viewer, renderer, and toolkit.
[2] Peter Selinger. 2019. A tool for tracing a bitmap, which means, transforming a bitmap into a smooth, scalable image.
[3] World Wide Web Consortium. 2011. Scalable Vector Graphics (SVG) 1.1 (Second Edition). https://www.w3.org/TR/SVG11/
[4] William E Lorensen, Harvey E Cline. 1987. Marching cubes: A high resolution 3D surface construction algorithm. SIGGRAPH Comput. Graph. 21, 4 (July 1987), 163–169.

[5] Rumelhart D E, Hinton G E, Williams R J. 1986. Learning representations by back-propagating errors. Nature, 323, 533–536.
[6] LeCun Y, Boser B, Denker J S, Henderson D, Howard R E, Hubbard W, Jackel L D. 1989. Backpropagation applied to handwritten zip code recognition. Neural computation, 1(4), 541-551.

[7] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan et al. 2019. Pytorch: An imperative style, high-performance deep learning library. Advances in neural information processing systems 32 (2019).
[8] He Kaiming, Zhang Xiangyu, Ren Shaoqing, Sun Jian. 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
[9] Chen Zhiqin, Zhang Hao. 2021. Neural marching cubes. ACM Trans. Graph. 40, 6, Article 251 (December 2021).
[10] Kurilin Ilya V et al. 2013. Generation of PDF with vector symbols from scanned document.
[11] Ncraun. 2013. SmoothScan is a tool to convert scanned text into a vectorized output form.
[12] Azriel Rosenfeld, John L Pfaltz. 1966. Sequential Operations in Digital Picture Processing. J. ACM 13, 4 (Oct. 1966), 471–494.
[13] FontForge. 2023. Free (libre) font editor for Windows, Mac OS X and GNU+Linux.
[14] Maas A L. 2013. Rectifier Nonlinearities Improve Neural Network Acoustic Models.

en