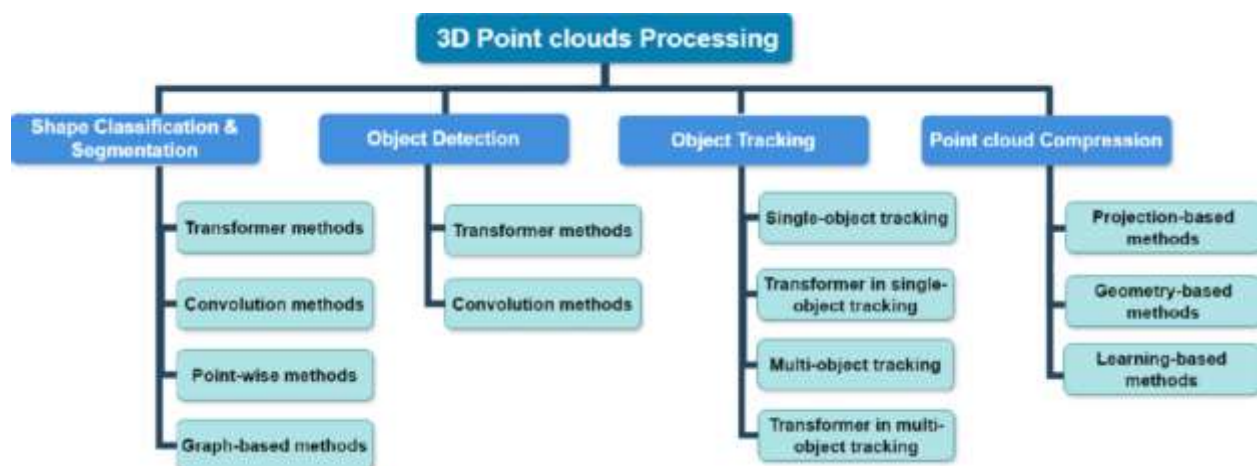


پردازش ابر نقاط سه بعدی: تأثیر بالقوه ترانسفورماتورها بر پردازش ابر نقاط

تقاضای رو به رشد برای مدل‌سازی سه‌بعدی، منجر به اهمیت فرمت‌های نمایش بصری سه‌بعدی، به ویژه ابر نقاط، در زمینه‌های مختلفی مانند پزشکی، رانندگی خودکار و برنامه‌ریزی شهری شده است. در سال‌های اخیر، روش‌های کلاسیک پردازش ابر نقاط جای خود را به تکنیک‌های یادگیری عمیق داده و پیشرفت‌های قابل توجهی در دقت ارائه کرده‌اند. پردازش خودکار داده‌های ابر نقاط سه‌بعدی برای تشخیص، ردیابی و تقسیم‌بندی اشیاء، یک حوزه تحقیقاتی مهم در کاربردهایی مانند رانندگی خودکار، رباتیک و واقعیت مجازی است. علاوه بر این، توسعه الگوریتم‌های فشرده‌سازی کارآمد داده‌های ابر نقاط به دلیل حجم بالای این داده‌ها نیز از اهمیت زیادی برخوردار است. در این پروژه ابتدا، پیشینه پژوهشی کاملی از پردازش ابر نقاط، از جمله تشخیص اشیاء، طبقه‌بندی، تقسیم‌بندی، ردیابی و فشرده‌سازی مورد بررسی قرار می‌گیرد. در این بررسی، تکنیک‌های جدید، از جمله ترانسفورماتورها به صورت خاص مورد توجه قرار می‌گیرد. در زمینه تقسیم‌بندی و طبقه‌بندی اشیاء روش‌های مبتنی بر ترانسفورمرها شامل رویکرد توکن سازی [2] [1]، رویکرد مدل‌سازی توالی [4] [3] و رویکرد ترکیبی توکن سازی-توالی [6] [5] به صورت جامع بررسی خواهند شد. سپس به بررسی روش‌های کانولوشن شامل کانولوشن گسسته و پیوسته و ترکیبی [9] [8] [7] خواهیم پرداخت. در ادامه روش‌های نقطه‌ای شامل رویکرد فقط رمزگذار و رویکر رمزگذار - رمز گشا [11] [10] بررسی می‌شوند. در نهایت نیز روش‌های مبتنی بر گراف شامل رویکرد دامنه فضایی و دامنه طیفی بررسی می‌شوند. در بخش تشخیص اشیاء، روش‌های مبتنی بر کانولوشن و ترانسفورم‌های مبتنی بر توجه [14] [13] [12] بررسی می‌شوند. در بخش ردیابی اشیاء، ترانسفورم‌ها در ردیابی تک شی و چندشی [17] [16] [15] بررسی می‌شوند. در نهایت در قسمت فشرده‌سازی، روش‌های مبتنی بر projection، روش‌های مبتنی بر هندسه و روش‌های یادگیری محور شامل کدگذار خودکار [19] [18] بررسی خواهد شد.

هدف از این بخش، برجسته کردن تأثیر بالقوه ترانسفورم‌ها در پردازش و فشرده‌سازی ابر نقاط است. به طور قابل توجهی، ترانسفورم‌ها با بهره‌گیری از مکانیسم‌های توجه، نتایج پیشرفته‌ای را در این حوزه به دست آورده‌اند. بررسی جامع انجام شده در این پروژه، نقاط قوت و محدودیت‌های آن‌ها را برجسته می‌کند و همچنین چالش‌های موجود در این حوزه را مورد بحث قرار می‌دهد. طبقه‌بندی رویکردهای فعلی در پردازش ابر نقاط در شکل ۱ نشان داده شده است هدف این است که نتایج این بررسی جامع در یکی مقاله مروری به چاپ برسد.



شکل ۱ - طبقه‌بندی رویکردهای فعلی در پردازش ابر نقاط

در ادامه به بررسی راه کارها جهت بهبود کارایی این شبکه ها برای پردازش ابر نقاط از جهت دقت و زمان پردازش خواهیم پرداخت. در طرح سال قبل، با بهره گیری از ساختار آرایه های هندسی فشرده (CGA) برای نمایش ابر نقاط، کارایی الگوریتم های نمونه گیری دورترین نقطه (FPS) و K-نزدیک ترین همسایه (KNN) را بهبود بخشیدیم. روش پیشنهادی از لحاظ زمان اجرا و حافظه مصرفی بسیار کارآمدتر از روش های قبلی است. در نهایت، یک روش وصله گذاری کارآمد برای فشرده سازی ابر نقاط توسعه داده شد که به طور قابل توجهی تعداد محاسبات و مصرف انرژی مورد نیاز برای رمزگذاری را بسیار کارآمد می کند. در طرح فعلی هدف این است که روش ارائه شده برای سایر پردازش های ابرنقاط شامل تشخیص و طبقه بندی اشیا و ردگیری اشیا نیز تعمیم داده شود و پردازش ابرنقاط با کارایی بهتر و پیچیدگی محاسباتی کمتر انجام شود.

مراجع

- [1] C. Zhang, H. Wan, X. Shen, and Z. Wu, "Patchformer: An efficient point transformer with patch attention," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 11799–11808, 2022.
- [2] Z. Liu, X. Yang, H. Tang, S. Yang, and S. Han, "Flatformer: Flattened window attention for efficient point cloud transformer," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 1200–1211, 2023.
- [3] H. Zhao, L. Jiang, J. Jia, P. H. Torr, and V. Koltun, "Point transformer," in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 16259–16268, 2021.
- [4] X. Wu, Y. Lao, L. Jiang, X. Liu, and H. Zhao, "Point transformer v2: Grouped vector attention and partition-based pooling," Advances in Neural Information Processing Systems, vol. 35, pp. 33330–33342, 2022.
- [5] D. Lu, Q. Xie, K. Gao, L. Xu, and J. Li, "3dctn: 3d convolution transformer network for point cloud classification," IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 12, pp. 24854–24865, 2022.
- [6] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," Advances in neural information processing systems, vol. 30, 2017.
- [7] X.-L. Li, M.-H. Guo, T.-J. Mu, R. R. Martin, and S.-M. Hu, "Long range pooling for 3d large-scale scene understanding," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10300–10311, 2023.
- [8] W. Wu, L. Fuxin, and Q. Shan, "Supplementary of pointconvformer: Revenge of the point-based convolution," 2023.
- [9] J. Park, C. Kim, S. Kim, and K. Jo, "Pcscnet: Fast 3d semantic segmentation of lidar point cloud for autonomous car using point convolution and sparse convolution network," Expert Systems with Applications, vol. 212, p. 118815, 2023.
- [10] S. Cheng, X. Chen, X. He, Z. Liu, and X. Bai, "Pra-net: Point relation aware network for 3d point cloud analysis," IEEE Transactions on Image Processing, vol. 30, pp. 4436–4448, 2021.
- [11] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 652–660, 2017.
- [12] Y. Zhou and O. Tuzel, "Voxelnet: End-to-end learning for point cloud based 3d object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4490–4499, 2018.
- [13] Y. Yan, Y. Mao, and B. Li, "Second: Sparsely embedded convolutional detection," Sensors, vol. 18, no. 10, p. 3337, 2018.

- [14] J. Mao, Y. Xue, M. Niu, H. Bai, J. Feng, X. Liang, H. Xu, and C. Xu, "Voxel transformer for 3d object detection," in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3164–3173, 2021.
- [15] B. Li, W. Wu, Q. Wang, F. Zhang, J. Xing, and J. Yan, "Siamrpn++: Evolution of siamese visual tracking with very deep networks," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 4282–4291, 2019.
- [16] Z. Wang, Q. Xie, Y.-K. Lai, J. Wu, K. Long, and J. Wang, "Mlvsnet: Multi-level voting siamese network for 3d visual tracking," in Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 3101–3110, 2021.
- [17] D. Wang, C. Huang, Y. Wang, Y. Deng, and H. Li, "A 3d multiobjecttracking algorithm of point cloud based on deep learning," *Mathematical Problems in Engineering*, vol. 2020, 2020.
- [18] D. Graziosi, O. Nakagami, S. Kuma, A. Zaghetto, T. Suzuki, and A. Tabatabai, "An overview of ongoing point cloud compression standardization activities: Video-based (v-pcc) and geometry-based (gpcc)," *APSIPA Transactions on Signal and Information Processing*, vol. 9, 2020.
- [19] M. Quach, G. Valenzise, and F. Dufaux, "Learning convolutional transforms for lossy point cloud geometry compression," in 2019 IEEE international conference on image processing (ICIP), pp. 4320–4324, IEEE, 2019.