Combinatorial Models for Sign Language Recognition: Graph Poses, Sequence Alignment, and Discrete Labeling

Luke Wachholz*, Srikanth Vemula, Human-Centric eXplainable Intelligence Lab (hcXiL), Computer Science Department (CSCI), College of Saint Benedict & Saint John's University

Automatic sign language recognition can be cast in a discrete framework that invites combinatorial analysis. We represent each video as a pose graph whose nodes are hand/body keypoints and whose edges encode intra-frame constraints; temporal edges link successive frames to form a trajectory graph with partial observations via a visibility mask. We study alignment of keypoint trajectories as a shortest-path problem in a DAG equivalent to dynamic time warping, with gap penalties weighted by missing-keypoint uncertainty; this yields efficient $O(nm)$ alignment and stable decoding. For word-level recognition, we reduce gloss prediction to discrete path labeling on the product of the trajectory graph and a finite-state gloss automaton, and analyze when k-path covers suffice for exact recovery under bounded noise. We formalize keyframe selection as a submodular set function over frames using pairwise motion diversity and visibility, giving a $(1 - 1/e)$-approximation via greedy selection and improved robustness to occlusion in practice. Under degree and length constraints on the trajectory graph, we derive bounds on the number of distinguishable sign patterns and on the combinatorial ambiguity introduced by coarticulation and viewpoint changes. On a subset of WLASL with MediaPipe-extracted keypoints, our alignment-first pipeline improves word-level recognition over a baseline encoder by emphasizing normalization, fixed-length resampling, and visibility-aware costs. These results highlight tractable connections between SLT components and classical topics—labeled paths on dynamic graphs, sequence alignment, submodular optimization, and complexity bounds—suggesting principled evaluation protocols independent of particular neural architectures.

Keywords: Pose Graphs, Trajectory Graphs, Dynamic Time Warping (DTW) Sequence Alignment, Discrete Labeling, Graph-Based Sign Recognition Combinatorial Path Structures, Keypoint Trajectory Analysis Gloss Prediction Models