Exploiting Depth Discontinuities for Vision-based Fingerspelling Recognition

Rogerio Feris, Matthew Turk, Ramesh Raskar, Kar-Han Tan, Gosuke Ohashi

Rogerio Feris, University of California, Santa Barbara
Matthew Turk, University of California, Santa Barbara
Ramesh Raskar, Mitsubishi Electric Research Labs
Kar-Han Tan, University of Illinois, Urbana-Champaign
Gosuke Ohashi, Shizuoka University
Outline

Motivation
Capturing Depth Edges with Multi-Flash Shape Descriptor and Classification
Experimental Results
Discussion
Conclusions
Motivation

Automatic Sign Language Translation

Two Major Components:

- Word Level Sign Vocabulary
- Fingerspelling
Motivation

Automatic Sign Language Translation

Two Major Components:

- Word Level Sign Vocabulary
- Fingerspelling
Problem

Vision-based methods: good for small subset of hand configurations

Challenge: Handle high amounts of finger occlusions

Glove-based methods: often expensive and intrusive
Our Approach:

Capture Depth Discontinuities (Depth Edges) with Multi-Flash Imaging
Depth Edges with Multi-Flash


Flashes may be infra-red
Implementation

Static Scenes

Dynamic Scenes
Shape Descriptor and Classification

(a) Depth edge image

(b) Number of pixels in each region

(c) Normalization

(d) Thresholding

(e) Histogram

Nearest-neighbor classifier
Experimental Results

Comparison with Canny Edges

Letter ‘R’

Input  Canny Edges  Depth Edges
Experimental Results

Discriminative Signatures

Worst cases reported by a glove-based system

Original  Canny  Depth Edges
Experimental Results

Quantitative Analysis
ASL Alphabet (except ‘I’ and ‘J’) – 24 letters
72 images (3 examples per letter)

Main goal: show the importance of internal edges
Plain background
Slight variations in pose and scale

96% of correct matches compared with 88% with Canny edges
(using a leave-one-out scheme)
Experimental Results

Efficiency
Depth edges: 77 ms
Shape Descriptor: 16ms
640x480 images, P4 3GHz

Clutter Removal

[Images of clutter removal before and after processing]
Discussion

Difficult cases

Lack of Background

Stereo Methods
Ongoing Work

Variable Wavelength

Flashes are triggered simultaneously

Video-based, continuous sign language analysis
Conclusions

Depth Edges: effective signature for fingerspelling recognition.

Promising direction for reliable, glove-free sign language analysis.
THANK YOU!

Four Eyes Lab, UCSB
http://ilab.cs.ucsb.edu/