Defining Gestures from Optical Flow: Week 9

Jon Harter

University of Central Florida

July 17, 2009

Current Progress and Goals

Recent Progress

- Approximated rotation and translation values for each gesture in test data
- Designed a feature vector representation of rotation and translation data
- Converted Matlab implementation to C# for better performance and compatibility with four camera setup

Current Goal: Create a working gesture recognition system

- Ensure that current implementation of gesture classifier is working properly
- Collect gesture data on a large scale

Gesture Representation

Gesture 1: Zorro

R_x	R_y	R_z	t_x	t_y	t_z
2	0	0	-1	0	-1
2	1	-1	1	-2	-2
1	0	1	0	-1	-2
1	0	2	0	-1	-2
1	0	1	0	-1	-2
1	0	2	0	-1	-2
-1	-1	0	-1	0	2
-1	-1	0	-1	0	2
-2	-1	0	-1	0	2
-1	-1	0	-1	0	2
-1	-1	0	-1	0	2
-1	-1	0	-1	0	2
-1	-1	0	-2	0	2
-2	0	-1	0	-1	1
1	1	-2	2	-2	-1

Gesture 2: S-Shape

R_x	R_y	R_z	t_x	t_y	t_z
-1	-1	0	-2	0	2
-1	-1	0	-2	0	2
-1	-1	0	-1	0	2
1	0	0	0	-1	-2
1	1	0	1	0	-2
1	0	1	0	-2	-2
2	1	0	1	0	-2
1	1	0	2	0	-2
2	1	-1	1	-2	-2
2	-1	0	2	0	1
1	1	0	1	0	-2
1	1	0	1	0	-2
2	1	0	1	0	-2
2	1	0	1	0	-2
2	1	0	1	0	-2

Extracting Features

Representation

- Each row of rotation and translation representation is one frame in gesture
- We cannot assume all gestures utilize the same number of frames

Important Properties

- Order of frames is crucial (ie. Zorro and S-Shape share similar rows but in different order)
- Runs of zeros, positives, and negatives are prevelant
- Magnitudes of rotation and translation should play key role

Feature Selection

- Feature 1: **Length**
  - Assumption: Similar gestures should be similar lengths
  - Particularly important in distinguishing between gestures such as stab and zorro that varry drastically in length.

- Features 2-7: **Longest negative run** for each column
  - Assumption: A run represents a prevelant type of motion within the gesture. Similar gestures should have similar prevelant motions.

Feature Selection

- Features 8-13: **Number of negative values in first half** of gesture for each column
  - Assumption: Adds order of motion to the gesture, distinguishing gestures with similar length, runs, and magnitudes from each other
- Features 14-19: **Number of negative values in second half** of gesture for each column
  - Assumption: Same as above.

Feature Selection

- Features 20-37: Extract features 2-19 with **zeros instead of negatives**
  - Assumption: Positive values need not be extracted for these types of features as the positive data should be the converse of the negative and zero data
- Features 38-43: **Sum of absolute magnitudes** for each column
  - Additional Assumption: Certain gestures may have consistant variation in the direction of a certain parameter, this ignores direction while still noting the motion (eg. a wrist may rotation in either direction unintentionally while performing the same gesture repeatedly)

Feature Analysis

- Total Features: 43
- Classifier is undergoing debugging so features are subject to change based on interaction with said classifier.
- Given a set of three gestures (Zorro: z_i, S-Shape: s_i, Stab: st_i), compared the difference between each vector as a measure of distinguishment

Comparing Feature Vectors

	z1	z2	z3	z4	z5	s1	s2	s3	s4	st1	st2	st3
z1	0	14	13	13	15	26	26	27	26	35	37	33
z2	14	0	19	14	21	31	31	31	30	41	42	39
z3	13	19	0	17	14	23	22	23	23	32	35	30
z4	13	14	17	0	14	27	28	27	26	30	32	29
z5	15	21	14	14	0	24	24	23	23	23	25	21
s1	26	31	23	27	24	0	12	10	10	35	36	34
s2	26	31	22	28	24	12	0	14	14	36	38	35
s3	27	31	23	27	23	10	14	0	10	32	34	31
s4	26	30	23	26	23	10	14	10	0	33	34	32
st1	35	41	32	30	23	35	36	32	33	0	9	7
st2	37	42	35	32	25	36	38	34	34	9	0	10
st3	33	39	30	29	21	34	35	31	32	7	10	0

Figure: Euchlidian Distance of Each Feature Vector to the Others

Comparing Feature Vectors

	z1	z2	z3	z4	z5	s1	s2	s3	s4	st1	st2	st3
z1	100	62	66	65	58	30	31	27	30	5	0	10
z2	66	100	54	67	49	27	27	26	29	3	0	7
z3	64	45	100	50	60	34	38	35	33	7	0	14
z4	60	57	45	100	57	15	13	16	19	3	0	9
z5	39	16	45	46	100	5	6	8	8	7	0	17
s1	29	16	37	26	34	100	67	72	72	5	0	7
s2	33	19	43	28	38	69	100	62	63	6	0	9
s3	22	10	34	22	32	70	58	100	71	8	0	11
s4	24	13	32	25	32	70	58	71	100	4	0	7
st1	13	0	21	25	42	15	12	22	19	100	78	83
st2	12	0	17	24	40	13	9	18	18	79	100	77
st3	14	0	23	26	46	13	11	21	18	83	75	100

Figure: Larger Values Denote Higher Similarities

Feature Analysis (Cont.)

- In every comparison, the top n coorelations (where n is the anticipated number of matching gestures) belonged to the target category (i.e. All Zorros matched the closest to each other Zorro ect...)
- For at least this small set of gestures, selected features are very descriptive
- Cannot weigh true quality of feature vectors until utilized within the classifier

Future Goals

- Have working gesture recognition within today or tommorrow
- Hone feature vectors to ensure the best gesture description
- Expand gestures to more than four distinct cases