2014 IEEE International Conference on Robotics & Automation (ICRA)
Hong Kong Convention and Exhibition Center
May 31 - June 7, 2014. Hong Kong, China

Early Failure Characterization of Cantilever Snap Assemblies using the PA-RCBHT.

Juan Rojas
School of Software
Sun Yat Sen University
Guangzhou, Guangdong, 510006, China

Kensuke Harada, Hiromu Onda, Natsuki Yamanobe, Eiichi Yoshida, and Kazuyuki Nagata.
Intelligent Sys. Research Institute, AIST
Tsukuba, Ibaraki, 305-8568, Japan

***Abstract*— Failure detection and correction is essential in robust systems. In robotics, failure detection has focused on traditional parts assembly, tool breakage, and threaded fastener assembly. However, not much work has focused on sub-mode failure classification. This is an important step in order to provide accurate failure recovery. Our work implemented a novel failure characterization scheme for cantilever snap assemblies. The approach identified exemplars that characterized salient features for specific deviations from a nominal trajectory. Then, a rule based approach with statistical measures was used to identify failure and classify failure sub-modes. Failure sub-mode classification was evaluated by using a reliability measure. Our work classified failure deviations with 88% accuracy. Varying success was experienced in correlating failure deviation modes. Cases with only 1-deviation had 86% accuracy, cases with 2-deviations had 67% accuracy, and cases with 3 deviations had 55% accuracy. Our work is an important step in failure characterization of complex geometrical parts and serves as a stepping stone to enact failure recovery.**

I. INTRODUCTION

Failure characterization (FC) and correction is essential in robust systems. FC initial developments focused on detecting abrupt changes [1].Since then, numerous filters and rule-based methods emerged. Overtime, statistical and machine learning methods were implemented to detect failure [2]. FC can be divided into model or model-free approaches. The former use theoretical system designs to identify failure, while the latter use experimental data. Model-free approaches can be both computationally expensive and damaging to robot equipment.

In robotics, FC has traditionally focused in parts assembly [3], [4], [2], tool breakage [5], [6], and threaded fastener assembly [7]. Recent work has used support vector machines to detect task outcome [2], [5]. Less work has focused on the harder problem of further classifying failure into failure modes. The latter identifies not only failure but the failure type, which is a necessary to implement failure recovery.

Our work, focuses on performing FC on Snap Assemblies. These are of great interest in industry and personal service robots but are challenging due to their elastic nature and geometrical complexity. There are three classes of snap fasteners: cantilever, annular, and torsional [8]. Cantilever snaps, are the most common and range from 1 to 2 or more fasteners in their mechanism. For this reason, they pose a higher degree of complexity than traditional assemblies.

Formerly, we implemented a framework that allowed us to: (i) autonomously assemble parts through the Pivot Approach (PA) [9], (ii) execute state reasoning through the Relative-Change-Based Hierarchical Taxonomy (RCBHT) [10], along with bayesian filtering, calibration, and a late stage FC analysis [11], [12], [13]. A snapshot of the humanoid robot HIRO performing a cantilever snap assembly with 4 fasteners is seen in Fig. 1(a).

To this stage, however, no work has been done to perform early failure characterization in snap assemblies. These assemblies yield complex force signatures that stem both from intricate hardware configurations and selected motion strategies to accomplish a task. The complexity is such that assembly planning methods like contact-state graphs [14] have yet to be used in this context.

This work studies if FC can be performed for limited noise deviations that occur at early stages of a snap assembly. FC encompasses both failure detection and failure mode classification. We hypothesize that if by associating salient features with deviations in 1-direction, then when multiple deviation directions are present, those same salient features can linearly superpose and be used to identify the failure deviation directions and magnitude.

We propose a novel rule-based approach that works in concert with the PA and the RCBHT. The PA partitions the task into four automata states that aid in the task's contextualization. The RCBHT yields five abstraction levels of increasing intuitiveness. Each level possesses labeling for actions or behaviors and quantitative data used to identify salient features. Our FC scheme first identifies key exemplars that capture salient features in the first two automata states of the PA. Exemplars are used to determine if (i) the task failed?, and (ii) if so, what failed and by how much?

Exemplars were discovered by contrasting, user-verified, successful snap assemblies with failed assemblies. Specifically, good trajectories were modified by introducing noise in only one of 3 directions (from here on named deviations). By contrasting successful assemblies with 1-direction deviations salient features were extracted as exemplars.

Later, training was used to compute statistical measures that help identify: (i) failure detection for a specific deviation direction, and (ii) the reliability of failure mode detection results. The testing phase evaluated the scheme's accuracy

978-1-4799-3684-7/14/$31.00 ©2014 IEEE

in both failure and failure mode detection. The testing was performed off-line.

Training consisted of 20 trials in which deviations were applied in three separate directions and combinations of these three in world coordinates. The testing phase used 34 new trials. Testing results showed that our scheme detected failure with 88% accuracy and failure modes with an average of 76% accuracy. Our approach hypothesis discovered that the contextual identification of exemplars based on constrained trajectory motion can aid in the identification of failure and failure modes, although with some limitations. Our work provides a simple and flexible approach to FC of cantilever snap assemblies.

The rest of the paper is organized as follows: the experimental set-up is presented in Sec. II, the Pivot Approach and RCBHT are presented in Sec. III, the FC scheme is presented in Sec. IV, Experiments and Results are presented in Sec. V, the discussion is presented in Sec. VI, and the conclusion is presented in Sec. VII.

II. Simulation Setup

HIRO, a simulated 6 DoF dual-arm anthropomorph robot, was used in the OpenHRP 3.0 environment [15]. CAD derived male and female camera parts were used. The male part was rigidly mounted on the robot's wrist, while the 4-snap female part, was fixed to the ground as in Fig. 1(a). For more details see [11].

III. The Pivot Approach and the Relative Change-Based Hierarchical Taxonomy

The PA exploits snap parts' hardware design to constraint the task's motion and generate similar sensory-signal patterns across trials and systematically discretize the assembly into intuitive states [9]. In a successful assembly, four states are completed: the Approach, Rotation, Insertion, and Mating states. Fig. 1(b) shows the assembly using the physical HIRO robot. For the purposes of this work, it is suffice to know that the Approach trajectory, a quasi-linear motion trajectory, will be modified to produce failed assemblies.

The RCBHT yields state representations by hierarchically abstracting snap assembly force-torque (FT) data in increasingly intuitive ways [10]. The hierarchical taxonomy is composed of four increasingly abstract layers that encode relative-change in the task's force signatures with a fifth layer used for state reasoning. The taxonomy is built on the premise that relative-change patterns can be classified through a small set of categoric labels and aided by contextual information. The RCBHT analyzes FT signatures from all force axes independently and contextualizes the state according to automata state participation. Previously, the RCBHT had not explicitly analyzed the approach state in its 5th layer, but now for early FC it will. The system's layered analysis is shown in Fig. 1(c).

In this section, we will briefly describe the layer's key aspects but will place more emphasis on the 2nd and 5th layers. While the description is brief, the FC process will particularly identify salient features from the 2nd layer and make use of the quantitative information there in to produce statistical measurements.

A. Primitive Layer

The first layer, the Primitives Layer, partitions FT data into linear data segments and classifies them according to gradient magnitude. Linear regression was used with a correlation measure to segment data when a minimum correlation threshold is flagged. The gradient classification starts separating data by looking for contact or mating events. Contact phenomena is characterized by abrupt changes in force signals almost approximating an impulse. To this end, positive and negative impulse labels were assigned to large gradients (see [12] for threshold parameters). For mating events, there is relatively small or no change in FT data. For this reason, constant labels are assigned to small gradients. In between these two extremes, another six sets of gradient classifications and quantitative data were issued to inform about the relative gradient change.

B. Composites

The next layer analyzes ordered-pair Primitives sequences to create "motion compositions" (MC's). By studying patterns in the ordered-pairs, seven sets of higher-level abstraction are extracted. These actions can represent force-torque signal's: adjustments, increases and decreases, contacts, and constant and unstable forms. Besides the assigned label, each MC posses quantitative data: average magnitude values, maximum signal values, average amplitude values, and starting and ending times for each of the primitives[1].

Filtering is executed after the motion compositions creation. Filtering merges motion compositions based on three different criteria: (i) a time-duration context, (ii) repeated MCs, and an (iii) amplitude value context. Detailed filtering rules are outlined in [13], here we limit to mention that their goal is to keep meaningful actions by filtering noisy ones.

The MC layer is key to the FC exemplar identification. The MC's quantitative data is general enough that we do not tamper with noisy signal dynamics yet not abstract enough that salient feature details are lost. All of our exemplars consist of either key motion composition average magnitudes or average amplitude values.

C. Low-Level Behaviors

The taxonomy's third layer considers MC ordered pairs and uses the same contextual information (signal duration and amplitude values) to yield another set of higher-abstraction classifications called low-level behaviors (LLB's). Eight LLB classifications were derived and labeled as: push, 'PS', pull, 'PL', contact, 'CT', fixed, 'FX', alignment, 'ALIGN', shift, 'SH', and noise, 'N'. The LLB formulation criteria is similar to those at the MC level. The process of continuing to abstract in the same way, reveals

[1]For this work the RMS value field in the RCBHT was changed for a maximum signal value

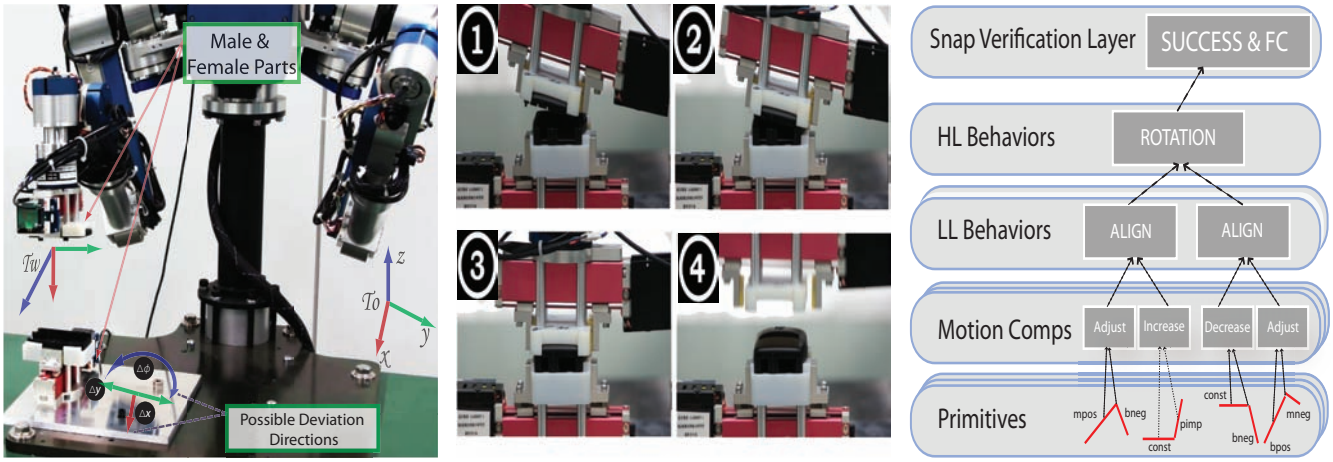

Fig. 1. HIRO assembles the snap assemblies through the PA's four states: Approach, Rotation, Insertion, and Mating. The RCBHT helps identify primitives, MC's, LLB's, HLB's, and the task's outcome.

high-level details that are not apparent at more granular stages. This level also has a filtering cycle and is detailed in [10].

D. High-Level Behaviors

The fifth layer, the high-level behaviors (HLB's) layer previously focused in determining whether an assembly task was successful or not. It did so by analyzing whether empirically selected LLB's were present in the 2nd-4th automata states. If so, it would flag the task as successful. Later in [11], this level also generated belief states about the task's success. In this work, the HLB layer was used to serve as an entry point for FC. FC analysis takes place as soon as the HLB layer analyzes the Approach State. In this case, instead of looking for key LLB's, it follows the FC scheme presented in Sec. IV. If the scheme detects a failure scenario, the rest of the analysis is aborted and relevant FC data is updated.

IV. Failure Characterization Scheme

Our FC method is a novel model-free rule based approach that uses the PA approach and the RCBHT. The PA constraints the assembly motion and partitions the task into four intuitive automata states that are in turn used by the RCBHT to richly contextualize the assembly task. The PA/RCBHT approach facilitate the identification of key actions or behaviors in a given automata state in a given force axis. The RCBHTs actions or labels have average magnitudes, amplitudes, and durations that can be used to characterize both successful and failed tasks. Before presenting the FC scheme, a set of fundamental principles for FC is presented.

A. Failure Characterization Requirements

FC systems consist of at least three components [16]:
(a) General Failure Identification: the system identifies if failure is present. The system should be robust against false-positive situations.
(b) Failure Modes Identification: the system should not only be able to identify if the task failed, but what specific task aspect's failed.
(c) Diagnostics: the system should provide quantitative data about the failure. When and by how much did something fail?
This information can then be used by a failure recovery mode and attempt to return the task's state to normal operation.

B. Failure Sources

Current state-of-the-art robotic systems are still liable to introduce error into a task. Main failure sources in manipulators are attributed to: upper bounds on payloads and speeds; absolute position accuracy which is dependent on temperature, load, speed, and workspace position; and repeatability error. Similarly, mechanical parts contribute to failure due to differences in design within the permitted tolerance range. And another important failure source is the uncertainty in pose of both robot effectors and parts, especially in less structured environments [3].

C. Constraints

This sections presents key constraints assumed in our work. Central to our approach is the assumption that the failures we study are ones that slightly miss the mark. We hypothesize that failure correction is most meaningful when one "slightly fails." In cases, where the deviation is large, one still has to correct the motion, but this correction is characterized by "starting from the beginning" and not with slight corrective motions. This work does not yet focus on failure recovery but the assumptions are important groundwork for future work.

In our attempt to define "slight failures" we do so relative to the dimensions of the snap parts (0.175m x 0.0575m x 0.03m). We also select world coordinate frames that are relevant and upper and lower bounds for those deviations. The coordinate-frame constraints selection is tied to the way in which the original approach trajectory, from the homing position to the contact of parts (known as the Approach State ine the PA strategy) works. [9]. In the linear domain, $+x$ and $\pm y$ were identified as free variables. In the rotational domain, Yall rotations about the z-axis ($\pm\phi$)

were identified and shown in Fig. 1(a). The following bounds were used for each of the free variables: $0.0075m \leq x \leq 0.0105m$; $0.0075m \leq y \leq 0.0105m$ and $-0.0105m \leq y \leq -0.0075m$; $+0.1745$rads $\leq \theta \leq +0.5066$rads and -0.5066rads $\leq \theta \leq -0.1745$rads.

In the Approach State of the PA, the male part always hits the posterior wall of the female part, hence we only consider motions that go past that wall. In the rotational domain we only considered pivot motions about the z-axis. We do not explicitly consider rotations about the *x-* or *y-axis* since the Approach State the male part already contact's its counterpart at an angle about the y-axis. Changing this angle, θ, would only change the duration of the task. We do assume bounds of $\langle 0.087\ rads^\circ \leq \theta \leq 3.145\ rads^\circ \rangle$. The lower limit prevents premature contact between the male snaps and the female walls.

D. Failure Scheme

The Failure Scheme presented in this work is divided into three stages: (i) exemplar identification; (ii) training for statistical measures computation, and (iii) testing. The exemplar identification stage finds salient features that characterize motions that deviate from nominal trajectory paths in one, two, and up to three simultaneous directions. The exemplars are used to derive three statistical measurements for successful and failure assemblies: average exemplar values across trials and associated upper and lower bounds. Finally, we test the accuracy of the scheme to identify (i) failure modes and (ii) the reliability of failure mode detection results.

1) Exemplar Identification: A key to identify salient features is to constraint the problem's dimensionality. We constrained failure cases by modifying the original PA Approach State trajectory motion (see Fig. 1(b)), in one of the three directions introduced in Sec. IV-C. Namely, $\langle +x, \pm y$ and $\pm \phi \rangle$. Noise was inserted within established bounds and deviation directions, *one-at-a-time*.

Exemplar Identification takes place by considering possible factors: (i) MC presence in a given state (RCBHT's 2nd layer); (ii) their average magnitude value throughout an automata state or a selected portion of it; (iii) similarly, their average amplitude value, or (iv) the duration of a given MC. This analysis can similarly take place for LLB's (the RCBHT's 3rd layer). These four factors can characterize a task's key dynamics and identify failure and corresponding failure modes.

Five exemplars were discovered by contrasting, user-verified, successful snap assemblies with failed assemblies. Each exemplar was correlated with: (i) presence in a given force axis; (ii) presence in a given PA automata state; (iii) a selected parameter from the list presented in the previous paragraph; (iv) salient feature duration in a state (starting from the beginning of a state), and (v) an abbreviated notation for the exemplar. Discovered exemplars are shown in Fig. 2. The exemplars, use local wrist-coordinates and their selection is explained below.

X-axis deviations are mostly characterized by the vertical

Dir	Num	Force Axis	PA State	Parameter (Avg Val)	State Duration	Notation
+x	1	My	Rotation	Magnitude	50%	My.Rot.AvgMag
	2	Mz	Rotation	Magnitude	50%	Mz.Rot.AvgMag
±y	1	Mz	Rotation	Amplitude	100%	Mz.Rot.AvgAmp
±ϕ	1	Fx	Approach	Magnitude	100%	Fx.App.AvgMag
	2	Fz	Approach	Magnitude	100%	Fz.App.AvgMag

Fig. 2. Exemplar Identification: Five exemplars were identified for trajectory deviations in three directions. Two exemplars were correlated with deviations in the x-direction, one with deviations in the y-direction, and two with deviations in the Yall or ϕ direction.

contact that the male snap makes with the female snap (z-axis in local coordinates). It's also interesting to note that My signatures strongly correlate Fz in the first half of the Rotation state. This is so, due to the fact that upon impact, the angle at which the male snap makes contact, allows the My axis to point in a similar direction to that of Fz. As the wrist aligns with the horizontal (the second part of the Rotation state), Mz tends to correlate more with Fx (as that axis points parallel to the base frame). By selecting the magnitude values of MCs in the first-half of the Rotation state (and thus the exemplar describes the average magnitude value) signatures between successful cases and those with deviations in the x-axis are distinguished well. Note that Fx is not used because even in the presence of a large number of deviations in x, the angle at which contact is made, allows the male part to remain within the friction cone of impact and maintains part quite static. This makes the Fx signature very similar to that of successful assemblies (though we noted this not to be true in situations where the x-deviation was closer to 0.0100m).

Y-axis deviations first consider that the PA is attempting the assembly while applying a downward push during the Rotation state. After the approach, the male parts tries to rotate about y_W. That is, in the absence of horizontal alignment, the male parts pivots significantly about the x_W giving rise to a clear signature. Average magnitude values are not used for this exemplar because force signals tend to oscillate about the origin cancelling each other out.

Yall deviations, consider that as the male part rotates about the z_W, the contact magnitudes both in the x- and z-coordinates are greatly diminished. Furthermore, the transition condition to move from the Approach to the Rotation state is designed to reach a certain threshold upon the proper contact of the male and female parts at the docking pivot. By visualizing force signatures at this transition, a clean change from a fixed signal to a quick and abrupt change is seen. It is asymptotic to the transition boundary between the Approach and Rotation states. However, when Yall deviations occur, noisy signals are seen in the Approach state before the transition condition is fulfilled. The behavior is identified by looking at the average magnitude values in the Approach state of both Fx and Fz.

2) Training: During training, three statistical parameters were computed for exemplars x_k of successful and unsuccessful tasks: (i) The Mean Averaged Value $\bar{x}$: it is the

mean value of all training example averaged magnitudes or maximum amplitudes; (ii) The Upper and Lower Boundaries: that is the maximum exemplar average value ($\bar{x}_{max}$) and the minimum exemplar average value ($\bar{x}_{min}$). Note that the exemplar for an individual trial already represents an averaged quantity, either of magnitudes or amplitudes, but the training average is of exemplars over a number of trials. The average value for exemplars in the training phase is show in Eqtn. 1.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \tag{1}$$

Additionally, if Π is defined as the set of all exemplars in the training phase, then the maximum value x_{max} for all exemplar averages $\bar{x}$ and the minimum value x_{min} is defined according to Eqtn. 2,

$$\bar{x}_{max} = \max_{i\in\Pi} \bar{x}_i \quad \text{and} \quad \bar{x}_{min} = \min_{i\in\Pi} \bar{x}_i. \tag{2}$$

3) Testing: Once the statistical measures were generated, the testing of new trials with deviations (in one, two, and three directions) were executed. Testing was designed to evaluate the system's performance in (i) failure mode detection, and (ii) correlation of failure detection results as a reliability measure. For the latter, we look at wether one or more exemplars were flagged, and if so, do they fall within the computed statistical measurements derived for them in the training phase. This last step correlates an exemplar value with averages computed during training. It's a reliability measure that informs whether the failure mode detection was a false-positives or not.

V. Results and Analysis

Using the experimental setup described in II, a series of successful and failed cantilever assembly tasks were assessed using a supervised approach. A training phase was used to compute statistical parameters while a testing phase was used to assess the accuracy of the failure and failure mode detection.

A. Training

In the first part of training, six trials for successful assemblies were run and their average exemplar values computed along with their upper and lower bounds. The results are shown in Fig. 3. The statistical values are used to identify failure during testing. That is, if an exemplar average value is within the bounds of Fig. 6, it is within range of successful cases and thus flagged as a non-failure. Each of the five exemplars is analyzed independently. In the second part of training, 20 trials were run. Deviations were introduced to the nominal Approach State motion ([9]) in one of three possible directions as stated in Sec. IV-D.2. An effort was made to have good data coverage so as to aid the generalizability of the algorithm in the testing phase. Out of the twenty trials, twelve of these included deviations 1 direction (4 in $+x$, 2 in $+y$, 2 in $-y$, 2 in $+\phi$, and 2 in $-\phi$, six included deviations in 2 directions (2 for x,y-deviations, 2 for x-ϕ deviations, and 2 for y-ϕ, and two trials included

Statistics		Average	Max	Min
x	MyRot	0.32	0.41	0.22
	FzRot	10.50	13.65	7.35
y	MzRot	0.16	0.21	0.11
ϕ	FxApp	1.86	2.42	1.30
	FzApp	-1.26	-1.63	-0.88

Fig. 3. Successful case exemplars: average values for selected exemplars along with maximum and minimum values.

Δx	Δ'y	Δϕ
0.0074	0.0097	0.2345
0.0080	0.0104	0.3333
0.0089	-0.0086	-0.4793
0.0101	-0.0092	-0.5066
Δx,Δy	**Δx,Δϕ**	**Δy,Δϕ**
0.0087, -0.0078	0.0077, 0.2087	0.0075,0.2673
0.0100, 0.0090	0.0081,-0.3311	-0.0079,0.2812
Δx,Δy,Δϕ		
0.0083,0.0075,0.17		
0.0085, 0.0085, 0.1848		

Fig. 4. Training Divergence Directions and Magnitudes in 20 trials: 12 trials test deviations in 1-direction (x, y, ϕ,), 6 trials test deviations in 2-directions, and 2 trials test deviations in 3-directions.

deviations in three directions. For success cases, six trials were used. Training trial's deviations are shown in Fig. 4.The summary of trials with trajectory motion deviation's are shown in Fig. 4. Exemplar averages, lower and upper bounds were computed for these trials. At this stage we noted that the Fx.App.AvgMag exemplar values were composed of both *positive* and *negative* quantities. These quantities cancelled each other during averaging and affected classification. Hence, the exemplar was subdivided into a positive version (Fx.App.AvgMag.Pos & Fz.App.AvgMag.Pos) and a negative version (Fx.App.AvgMag.Min & Fz.App.AvgMag.Min). Both still inform about the Yall axis deviation, but the valency separation now gives information about the deviation direction. Results are shown in Fig. 5. Note that average values between failure case exemplars and successful ones are markedly different and ratify their use for classification. The ratio of *average-success* value to *average-failure* value for the seven exemplars is as follows: $\langle 0.71, 1.24, 0.45, 2.73, -1.05, -1.38, -0.29\rangle$. With regards

Statistics	x		y	ϕ			
				Positive		Negative	
	MyRot	FzRot	MzRotPos	FxApp	FzApp	FxApp	FzApp
Average	0.45	8.48	0.36	0.68	1.19	-1.35	4.38
Max	0.68	11.24	1.37	1.28	2.87	-0.56	8.82
Min	0.33	6.39	0.10	0.03	-0.92	-3.69	2.09
maxRatio	1.51	1.33	3.80	1.87	2.41	0.41	2.01
minRatio	0.73	0.75	0.29	0.04	-0.77	2.74	0.48

Fig. 5. Training Case Statistics: the average values over all training trials for the five (seven with a sub-division of exemplars for ϕ) are shown, along with maximum and minim recorded values that are used to compute maximum and minimum bounds for testing cases.

+x	+y	-y	+ϕ
0.0075	0.0075	-0.0075	0.1745
0.0085	0.0085	-0.0085	0.3490
0.0095	0.0095	-0.0095	0.5235
0.0105	0.0105	-0.0105	
-ϕ	**x,y**	**x,ϕ**	**y,ϕ**
-0.1745	0.0073,0.0081	0.0081, 0.2812	0.0083, 0.1826
-0.3490	0.0092,0.0102	0.0097, -0.1800	0.0088, 0.1809
-0.5235	0.0085,-0.0081	0.0077, 0.4073	-0.0094, 0.1901
	0.0077, -0.0100	0.0100,-0.3199	-0.0083, -0.3283
x,y,ϕ			
0.0080, 0.0081, 0.2169		0.0082, 0.0077, 0.3894	
0.0080, 0.0081, 0.3087		0.0083, 0.0079, -0.3894	

Fig. 6. Testing Divergence Directions and Magnitudes in 32 trials: 18 trials test deviations in 1-direction (x, y, ϕ,), 12 trials test deviations in 2-directions, and 4 trials test deviations in 3-directions.

to upper bound variation, the latter was large across exemplars. For Mz.Rot.AvgAmp, upper bound values averaged amplitudes upto 3.8 times the mean value. This result sheds light on the large force-contact value variation that exists in the $\pm$y-directions as a result of intricate geometry within snap parts (and which is originally designed to facilitate the assembly).

B. Testing

In the testing phase, 32 trials were used to evaluate the accuracy and generalizability of the FC scheme. From the 32 trials, 18 considered deviations in 1-direction, 12 considered deviations in 2-directions, and 4 considered deviations in 3-directions. The testing trials and their corresponding deviations are shown in Fig. 6 (all deviations were within the bounds introduced in Sec. IV-C). We now present the system's effectiveness in detecting deviation-specific failures, correlation modes, and false positives. In Fig. 7, two tables are presented. Both correspond to testing trials organized by deviation directions and by exemplar classification (Exemplars 1-5 refer to those introduced in Fig. 2). On the top side, the table records: (i) Correct diagnostic's percentage for individual exemplars and (ii) two evaluation measures for the overall combined effect of multiple exemplars. On the bottom side, the table records the percentage of failed exemplars that were able to be correlated. To clarify, consider an assembly where there were deviations in the x- and y-directions, and in which My.Rot.AvgMag, Fz.Rot.AvgMag, and Mz.Rot.AvgAmp were flagged. We then consider if these exemplar's average values fell within the bounds derived during training? If so, we state that these exemplars are correlated and have a more reliable measure that the exemplars are not false-positive's.

C. Analysis

The analysis section studies failure detection and failure modes on three levels: (i) Individual Deviation Analysis: which exemplars did better when a given deviation was enacted?; (ii) Individual Exemplar Analysis: how did a given exemplar do across different deviations subgroups; and (iii) Output Computation Analysis: a study of the overall combined effect of exemplars for a given deviation.

1) Failure Detection: The first analysis, Individual Deviation Analysis, consists in taking the: x, y, ϕ, xy, $x\phi$, $x\phi$, and x, y, ϕ deviations as separate subgroups and detect which exemplars have the most salient features for a given deviation. It can be considered a row-by-row analysis and whose main points are highlighted below:

- x My.Rot.AvgMag was more consistent than Fz.Rot.AvgMag and was 100% accurate.
- y Mz.Rot.AvgAmp was very effective at 88%.
- ϕ Exemplars 4 and 5 performed similarly.
- x,y Fz.Rot.AvgMag was accurate 100% of the time.
- x,ϕ Fz.Rot.AvgMag was not effective compared to xy deviations. Fz.App.AvgMag was more effective than Fx.App.AvgMag and tended to be so in general.
- y,ϕ Mz.Rot.AvgAmp and Fz.App.AvgMag were 100% accurate and Fz.App.AvgMag did well 75% of the time.
- x,y,ϕ Mz.Rot.AvgAmp performed poorly because with deviations in x and ϕ forces originally experienced in y due to internal friction cannot be reached.

The second analysis, Individual Exemplar Analysis, consists in looking at an individual exemplar across all deviation subgroups. It can be considered a column-by-column analysis whose main points are highlighted below:

- My.Rot.AvgMag was the most effective exemplar having an efficiency of 88% across all subgroups.
- Exemplar 1,4,5 were correct over 75% of the time.
- Mz.Rot.AvgAmp is low but deceptive due to involvement in x, y, ϕ. Otherwise the average would be 88%.

The third analysis, Output Computation Analysis, looks at two approaches for modeling the exemplars combined

	Explr1	Explr2	Explr3	Explr4	Explr5	Comb	Max
X	100%	50%				75%	100%
Y			88%			88%	88%
ϕ				83%	83%	83%	83%
X,Y	75%	100%	75%			83%	88%
X,ϕ	75%	50%		50%	75%	63%	75%
Y,ϕ			100%	100%	75%	92%	100%
X,Y,ϕ	100%	75%	50%	75%	100%	80%	83%
Avg	88%	69%	78%	77%	83%	80%	88%
	Crl 1	Crl2	Crl 3	Crl 4	Crl 5	Comb	Max
X	100%	50%				75%	100%
Y			75%			75%	75%
ϕ				83%	83%	83%	83%
X,Y	75%	100%	25%			67%	63%
X,ϕ	75%	50%		50%	75%	63%	75%
Y,ϕ			50%	50%	75%	58%	63%
X,Y,ϕ	75%	75%	25%	50%	50%	55%	50%
Avg	81%	69%	44%	58%	71%	68%	81%

Fig. 7. Testing Phase Results: On the top table accuracy percentages for failure detection by deviation subgroups are shown. On the bottom, accuracy percentages for correlation detection, that is, the number of times exemplars identified with failure detection were corroborated by statistical measures akin to that type of fault.

effect: (i) the Combination Output Measurement and (ii) the Maximum Value Output Measurement. The former approach, considers one deviation with multiple exemplars and takes the exemplar's average value. *I.e.* , for deviation x, y, ϕ the approach averages the five exemplar values. The latter approach, looks at a deviation subgroup and first finds exemplar pairs that describe deviation in one direction, it then computes their maximum value (*i.e.* My.Rot.AvgMag and Fz.Rot.AvgMag, or Fx.App.AvgMag and Fz.App.AvgMag); and finally, it averages the paired exemplar's values with the remaining exemplars. *i.e.* , for deviation x, y, ϕ we compute the maximum value of two pairs of exemplars: 1 & 2 and $4 \& 5$. Then exemplar 3 is averaged with the maximum value of both pairs. The Maximum Value Output Measurement has overall higher rates than the Combination Output Measurement. It assumes that only the best exemplar that characterizes a deviation is needed, not both. In our results, we see that deviations in the $x-$ and y,Yall directions, were correctly identified 100% of the time. The Combination Measurement has similar patterns but with lower percentages.

No false-positives were recorded. All failure case scenarios were detected correctly.

2) Failure Correlation: This section analyzes failure correlation modes. It determines failure correlation mode performance by evaluating how well detected failed exemplars could be correlated with their corresponding statistical measures. Our analysis shows that failure mode correlation is more difficult than failure detection. Failure detection only detects exemplar values outside corresponding successful statistical ranges while correlation requires that failure values fall within bounds of corresponding statistical ranges. Our analysis also shows that as deviations are introduced in more axes, there is more variability in the failed exemplar's value range. As in Sec. V-C.1, we perform the same three types of analysis.

For Individual Deviation Analysis, 1-deviation subgroups scored higher than 2- or 3-deviations groups. Similarly, 2-deviations groups scored higher than those with 3-deviations. For 1-deviation subgroups, x and Yall were better correlated than y. For 2-deviations groups, y,Yall was the worst. Again, the data points show that when the male part is moved in the $x-$ or Yall directions, the force signals in y that originate from internal horizontal friction disappear. This is further corroborated in the x, y,Yall subgroup in which Mz.Rot.AvgAmp was correlated only 25% of the time.

For Individual Exemplar Analysis, My.Rot.AvgMag (x-dir.) was the best, Mz.Rot.AvgAmp (y-dir.) the worst, and with Fx.App.AvgMag and Fz.App.AvgMag accuracy diminishes as more deviations are introduced.

For Output Computation Analysis, the Yall direction had the highest results. y,Yall and x, y,Yall had the lowest results due to the negative contribution from the y-deviation. Now, if the y-correlation were removed, accuracy levels would go up to 67%. And as stated earlier, the tendency is to have better correlation for 1-deviation groups than for 2- or 3-deviation groups. One interesting measure for data spread is shown by taking the ratio between the standard deviation's mean value and the average exemplar value's mean. In this case, Mz.Rot.AvgAmp (y-dir) had a 1.15 ratio, and Fz.App.AvgMag had a 1.25 ratio, while My.Rot.AvgMag and Fz.Rot.AvgMag had ratios of 0.15 and 0.14 respectively. This ratio provides insight into the wide spread of data across Mz.Rot.AvgAmp and Fz.App.AvgMag.

The problem of decreasing correlation accuracy however, could be addressed by computing means and upper and lower bounds not just for an exemplar across all deviation subgroups but for more localized means and bounds associated with each of the subgroups.

3) Averages and Std. Deviations: 3) The final analysis looks at the average values and standard deviations of exemplars within deviation subgroups. The data is seen in Fig. 8 The following trends were noted: for Mz.Rot.AvgAmp values for all subgroups (except deviation in the y-direction) had similar magnitudes values as success cases. This explains the system's inability to correlate with these values. Also, for tasks that exhibit purely y-direction deviation's, the mean values are much larger. And, Fx.App.AvgMag.Pos has comparatively low values for Yall compared to other deviation subgroups. Exemplars across $x, Yall$ and x, y,Yall and also for y,Yall, and x, y,Yall all share closely correlated values. Fz.App.AvgMag.Min has similar value for Yall and y,Yall but not x, y,Yall. Fz.App.AvgMag.Pos has similar values for x, y,Yall and y,Yall but not for $x, Yall$. Finally, Fx.App.AvgMag.Min has similar values for y,Yall and x, y,Yall but not for Yall. Hence, contextualizing means and bounds by training them for the specific sub-groups may increase failure detection and failure mode detection significantly.

VI. DISCUSSION

Our work demonstrates that early FC was implemented with modest success for cantilever snap assemblies based on the PA approach and the RCBHT system as 88% of cases were correctly detected. Failure correlation modes were detected with varying degrees of success; ranging from more accurate detection for cases in which only one deviation existed (86% accuracy) to lower accuracy as the deviation number increased for trajectory motions (67% for 2-deviations and 55% for 3-deviations).

We were also interested in understanding if by discovering exemplars corresponding to 1-deviation directions we could then superpose deviation directions and expect the same exemplar statistical measures to capture failure. To this end, it seems that deviations in the x and Yall directions did better than the y direction. This was specially true when y-deviations were combined with x-deviations. The reason was clear. Upon, diverging in the forward x-direction, the snap parts would no longer be in position to exert the same friction patterns due to hardware geometry configuration that are possible when both parts are spliced.

This work is unique as it allows the early stage FC of snap parts with complex geometry. By combining the PA strategy

Axis	Stat	MyRot	FzRot	Axis	MzRot	Axis	FxAppP	FzAppP	FxAppM	FzAppM
x	Avg	0.55	9.54	y	0.43	ϕ	0.61	0.32	-2.41	4.22
	Dev	0.04	0.97		0.43		0.64	1.45	0.02	0.54
x,y	Avg	0.47	8.05	x,y	0.17	x,ϕ	1.40	1.56		
	Dev	0.07	0.41		0.09		0.39	0.79		
x,ϕ	Avg	0.47	8.83	y,ϕ	0.16	y,ϕ	1.09	0.88	-0.42	6.58
	Dev	0.06	1.29		0.08		0.15	0.94	0.00	0.00
x,y,ϕ	Avg	0.52	8.21	x,y,ϕ	0.16	x,y,ϕ	1.01	0.95	-0.52	0.47
	Dev	0.08	1.32		0.08		0.32	0.81	0.00	0.00

Fig. 8. Exemplar average values and standard deviations for deviation subgroups during testing.

with the snap verification RCBHT system and the current FC scheme, promising results were obtained in correctly characterizing failure and corresponding failure modes. Our work like those of [6] and [7] work not only on failure detection but also on characterizing failure modes. In our work, due to the geometrical complexity of the parts, the number of exemplars and possible recovery strategies are more numerous.

The overall strength of the approach is the identification of key parameters to characterize failure. The approach is flexible enough that it can be refined: both by computing more contextualized exemplars based on deviation subgroups and by employing more advanced classification methods. The approach also does not suffer from the curse of dimensionality as its based on a linear computation of mean values and upper and lower bounds. And while the approach performed relatively well with a limited set of coverage points, a more robust set of data should be explored and tested. It is not yet clear if an exhaustive coverage process would be desirable to design failure recovery methods in snap assemblies and similar tasks due to time costs and potential equipment damage in the process.

As par of our future work we will explore the contextualization of means and bounds by further subdividing exemplars across deviation groups. We will also explore better classification algorithms and probabilistic learning and use this work for actual failure recovery at a later time.

VII. CONCLUSION

In conclusion, a novel early failure detection scheme was implemented for cantilever snap assemblies. Snap assemblies are challenging due to elastic components and complex geometric configurations. The method is able to provide early identification of situations in which the trajectory motion of an assembly has been deviated from normal cases. We identified not only what types of failure occurred but we also correlated them with historic statistical data to obtain reliability measures for the detected failures. This information is significant in order to implement failure recovery schemes at a later time.

REFERENCES

[1] A. Willsky, "A survey of design methods for failure detection in dynamic systems," *Automatica*, vol. 12, no. 6, p. 601:611, 1976.

[2] A. Rodriguez, D. Bourne, M. Mason, G. F. Rossano, and J. Wang, "Failure detection in assembly: Force signature analysis," in *IEEE Conference on Automation Science and Engineering*, 2010.

[3] R. A. Brooks, "Symbolic error analysis and robot planning," *The International Journal of Robotics Research*, vol. 1, no. 4, pp. 29–78, 1982.

[4] D. M. Fullmer, "Parts assembly using signature analysis," US Patent 4,855,923, Aug. 1989.

[5] S. Cho, S. Asfour, A. Onar, and N. Kaundinya, "Tool breakage detection using support vector machine learning in a milling process," *International Journal of Machine Tools and Manufacture*, vol. 45, no. 3, pp. 241–249, 2005.

[6] Y.-W. Hsueh and C.-Y. Yang, "Prediction of tool breakage in face milling using support vector machine," *The International Journal of Advanced Manufacturing Technology*, vol. 37, no. 9-10, pp. 872–880, 2008.

[7] K. Althoefer, B. Lara, Y. H. Zweiri, and L. D. Seneviratne, "Automated failure classification for assembly with self-tapping threaded fastenings using artificial neural networks," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 222, no. 6, pp. 1081–1095, 2008. [Online]. Available: http://pic.sagepub.com/content/222/6/1081.abstract

[8] R. Soddhi and M. Sonnenberg, "Use of snap-fit fasteners in the multi-life-cycle design of products," in *IEEE Intl. Symp. on Electr. & Env.*, 1999.

[9] J. Rojas, K. Harada, H. Onda, N. Yamanobe, E. Yoshida, K. Nagata, and Y. Kawai, "Cantilever snap assembly automation using a constraint-based pivot approach," in *IEEE Intl. Conf. on Mechatr. & Automation*, 2012.

[10] ——, "A relative-change-based hierarchical taxonomy for cantilever-snap assembly verification," in *IEEE Intl Conf. on Robots and Systems*, 2012.

[11] ——, "Probabilistic state verification for snap assemblies using the relative-change-based hierarchical taxonomy (in-print)," in *IEEE-RAS Intl. Conf. on Humanoid Robots*, 2012.

[12] ——, "Gradient calibration for the rcbht cantilever snap verification system." in *IEEE Int'l Conference on Robotics and Biomimetics*, 2012.

[13] ——, "Towards snap sensing," *International Journal of Mechatronics and Automation*, vol. 3, no. 2, pp. 69–93, 2013.

[14] W. Meeussen, J. Rutgeerts, K. Gadeyne, H. Bruyninckx, and J. D. Schutter, "Contact-state segmentation using particle filters for programming by human demonstration in compliant-motion tasks," *IEEE Trans. on Robotics*, vol. 23, no. 2, pp. 218–231, 2007.

[15] F. Kanehiro, H. Hirukawana, and S. Kajita, "Openhrp: Open architecture humanoid robotics platform," *Intl. J. of Robotics Res.*, vol. 23, no. 2, pp. 155–165, 2004.

[16] E. Lopez-Mellado and R. Alami, "A failure recovery scheme for assembly workcells," in *Robotics and Automation, 1990. Proceedings., 1990 IEEE International Conference on*. IEEE, 1990, pp. 702–707.