Robotic Assembly of Naturally Varying Food Items via Teaching by Example

T. G. Williams, J. J. Rowland* and M. H. Lee
Department of Computer Science,
University of Wales,
Aberystwyth SY23 3DB,
Wales, U.K.

Abstract

This paper reports the use of range images for autonomous replication of food products whose desired configuration is taught by presenting to the robot hand-assembled examples. Range images are acquired as items on a conveyor pass through a laser stripe, and are used to achieve visually acceptable robot positioning and orientation of discrete components in relation to the pre-taught range-image template. We present results of an extensive series of experiments that illustrate the effectiveness of the approach when applied to different types of food product whose size and shape are subject to natural variation. The technique is capable of accommodating food items that are asymmetric and that require placement at specific orientation as well as position.

1 Context and Aims

This paper builds on earlier work on teaching by example [19] and on autonomous grasping of variable products [12]. Our overall aim is to permit flexible systems to learn to assemble new products by being shown examples of the results required, and thus to accommodate a useful range of product types without re-programming or skilled reconfiguration. The system is intended to permit a person without specific robot expertise to specify an individual assembly stage of a previously untaught product and to enable its subsequent autonomous execution by robot. Our application case studies currently relate to the food industry and we are concerned with products such as sandwiches, pizzas and pre-packed salads. Most stages in assembling these products can be achieved with 'pick and place' operations, and we report extensive experiments on teaching from examples of products from which the robot determines position and orientation for component placement. The literature reports many examples of robotic research for food-related applications, and we are aware of unreported work in the UK and New Zealand. However, we are not aware of any other work similar to that described here. A recent review of sandwich production technology [8] emphasises the dominance of manual operations and, in relation to the potential of robotics, discusses hygiene and quality benefits, as well as cost justification and the need for configurability.

Much research into food robotics has centred on hard coded solutions, sometimes with sensor assistance. In meat processing, for example, work has been done on packing fresh chicken pieces [9]. The use of advanced sensing in robots for food manufacture is relatively rare even in the research environment, although there are notable examples, particularly relating to fish. Esprit project 6901 used vision and robotics in fish processing and inspection, and fish head removal is described in [5]. Other uses of advanced sensing in food-related applications, for robot control or for inspection purposes, include weight estimation of flatfish [15], fruit inspection and harvesting (e.g. [16, 7, 20]) and packing of horticultural produce (e.g. [17]). In all of these the flexibility of the

*email: jjr@aber.ac.uk

automation is limited to accommodating natural variation in size and shape between individual items being processed, rather than between different categories of product or ingredient; a carefully formulated representative model of the items being manipulated is an integral part of the system, with the natural variations being accommodated via sensing or mechanical compliance. Such reliance on a hard model cannot readily accommodate different, but related, tasks or a diversity of food items or ingredients; the resultant need for skilled reprogramming renders this approach inappropriate where frequent product modifications are required as is the case in the snack food industry, to which our work relates.

A particular challenge concerns the need to accommodate the variation between different product types, and the natural variation between (food) components of the same type. We have addressed this using a search strategy that achieves a visually acceptable match between nominally equivalent examples of naturally varying items. Handling the components to be placed is another challenge; for the present we use simple suction grippers and we accept their limitations, thus allowing us to concentrate on the teaching process and appropriate sensing techniques. Other essential features of the work are sensing methods that can implicitly detect relevant product and component features, and algorithms for assembly that do not require explicit product knowledge other than that derived during the teaching phase.

In teaching via hand-prepared examples of the desired end-product, our work contrasts with the substantial body of work on 'teaching by showing' or 'teaching by demonstration' that concerns teaching robot motion or transfer of human skills (e.g. [3, 4, 6, 10, 13]).

2 Experimental Environment

The likely operating context of an eventual system based on this technique would be a robot that could be moved into place over a conveyor belt and which, after being taught by example, would undertake a specific food assembly operation; several such robots could be used in sequence along the production line to form a final composite product. Currently work is based on one of our existing industrial robots, an Adept 1. This is not a food-grade robot but serves for experimental purposes; an industrially viable system would be based around low-cost actuators and our methods are consistent with this. Control of our robot is via an IBM-compatible PC interfaced via serial line to the Adept controller. Products are manipulated on a conveyor belt whose control permits easy control of speed and operation in both directions to permit cyclic tests.

The sensory data upon which the technique operates is obtained via 2.5D profile sensing, which readily allows determination of the location, shape, thickness, orientation and surface profile of a variety of food products of relevance to the sandwich and snack-food industry. A 'smart' camera views the objects at an oblique angle while they are passed through a laser stripe. There is a large literature on this and similar techniques, relating to the food industry and otherwise (e.g. [1, 15]), and commercial systems are readily available. A specific advantage of the method, in addition to its 2.5 dimensional capability is its relative insensitivity to variations in ambient lighting conditions and to tonal variations in the objects being examined.

The range sensor is the IVP 'Ranger' Smart Camera[1] with a laser that projects a fan of red light through which the objects to be scanned are moved. Laser and sensor are mounted above the work area as shown in Figure 1. The camera is a 'smart' device which incorporates the essential image processing functions so that the interface to the PC carries only the x, y, z profile data in response to commands issued by the PC. This sensor is effective for the purpose, with a resolution of about 0.5mm in x (the direction of conveyor travel), 1.4mm in y (across the conveyor) and 0.8mm in z (vertical height above the conveyor), over a sensed width of approximately 250mm. There is inevitable shadowing as a result of the angle between illumination and camera and some signal processing is required to compensate for this.

The eventual system based on this technique would, after being taught by example, undertake a specific food assembly operation; several such systems could be used in sequence along the production line, perhaps interspersed with human operators to provide maximum versatility. Each

[1] Integrated Vision Products, Sweden.

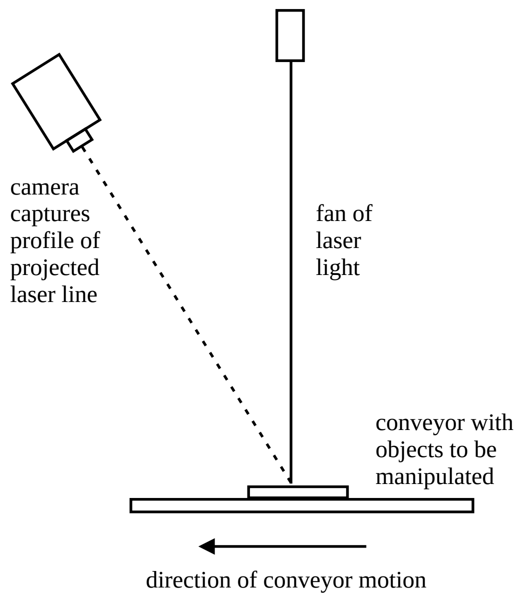

Figure 1: *The conveyor and sensor arrangement.*

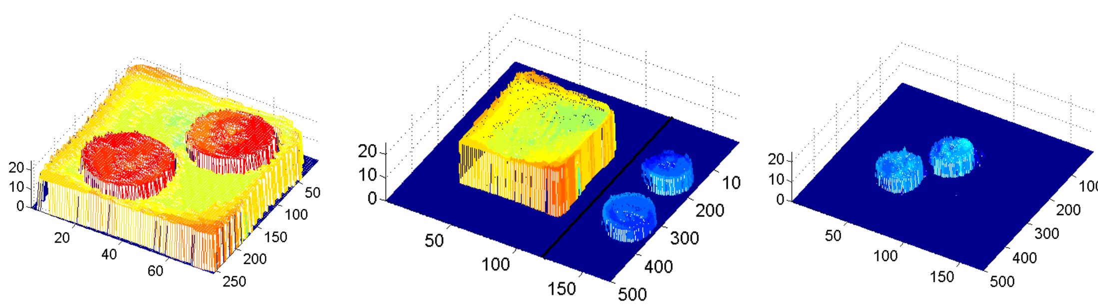

Figure 2: *From left to right: range scans of the required product, the substrate on which components are to be placed, with the components adjacent, and the difference surface that is used to determine the required component placement locations.*

of these assembly stages is therefore designed as a stand-alone system comprising of an actuator (robot), range sensor and controlling computer and each will at any one time manipulate one type of component, e.g. cucumber slices, chicken legs or lettuce, and exchangeable end-effector tooling will be provided to deal with the particular component type.

3 Teaching by Example

To teach the product, the substrate (the component upon which other components are to be assembled) is placed on the conveyor and scanned (see Figure 2). The scanning operation is then repeated after each component is added to the substrate by hand. Subsequently, placement positions and orientations are derived by navigating the difference surface between these two scans, which represent the actual and desired product states. Dfference surfaces, one for each added component, are derived by applying appropriate signal processing techniques to remove noise and artefacts from the scan data and then subtracting the relevant scans. The data representing the substrate and the difference surfaces are then saved using a pre-defined label for future retrieval by the assembly algorithm.

In addition to learning the properties of the example product, the system must learn about the component type it is to assemble at a given stage. It is envisaged that in a production system the components will be dispensed from purpose-built dispensers at an approximately constant orientation. To simulate this, we perform an additional learning stage (additional to the example product learning stage) in which the component is placed on the conveyor by the robot in the (assumed) orientation provided by the dispenser; it is then scanned to provide the part acquisition information for the robot.

There is the possibility of forming composite templates by scanning a range of examples so as to permit generalisation over the range of expected variation in component and product characteristics. However, as discussed below, our current work makes successful use of alternative techniques for accommodating variation.

3.1 Matching Component and Template

The problem of matching components to the taught template is one of approximate alignment, given that no two nominally similar components are precisely identical. The literature reports several approaches to this type of problem based, for example, on principles of symmetric difference [2], on fuzzy techniques [11] or genetic search [14, 18]. In our application we use a matching algorithm based on a hill climb via directed search that involves component rotation and centroid registration. By this process, the profile of the component to be placed is matched to the 'hole' in the difference surface and the position and orientation for placement by the robot of the actual components is determined.

This matching operation accommodates variation in component shape and size and achieves visually acceptable position and orientation of the new component in relation to the template. This is entirely appropriate for component placement within the natural variation expected of the components but it is not appropriate for accommodating variation in the substrate. In this case the component locations need to be automatically adjusted so that the visual appearance is retained despite substrate size variation; for this a form of morphing is used. The substrate template is therefore scaled to the size of the current substrate, along with the previously determined component placement locations. Consequently, substrate size variation leads to a correctly proportioned scaled product rather than to a product with apparently inappropriately placed components.

The algorithm uses a two-stage search strategy, with a global search to produce the starting point. This initial stage conducts a coarse search of the complete search area and provides what is intended to be a good starting point within the region of the global minimum for the subsequent algorithms in order that they can refine the match locally. This stage rotates the template through $-90°$ to $+90°$ in increments of $45°$ and, for each rotation, translates the template so that its centroid is coincident with that of the product. The match which produces the least error is passed forward as the starting point for a subsequent search.

The algorithm then employs a combination of targeted and steepest descent searches in order to improve the match presented to it by the initial search. It relies solely on the area in pixels of the mis-matched regions. Following the first stage (above) the algorithm enters an iterative loop which applies local translation, rotation and resizing iteratively, in that order, until the error is minimised.

Orientation is found via a binary search within $40°$ of the starting orientation. The template is initially rotated $20°$ either side selecting the alternative that has the higher match quality. The rotation step is then halved to $10°$ and so on, until a search resolution of $1°$ is achieved. The translation search utilises displacements by one pixel in each of four orthogonal directions; the translation which produces the match of highest quality is then adopted and the procedure repeated until the error is minimised. To assist in achieving best rotational and translational matching, the template is re-sized one pixel at a time in the direction that reduces the error. This algorithm has been found to be effective and fast in relation to other approaches considered.

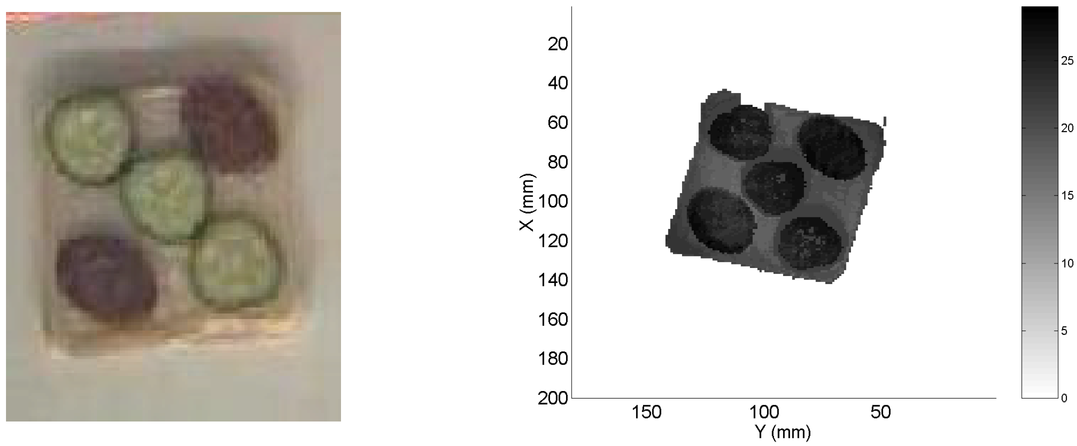

Figure 3: *Product B: conventional image (left) and range image (right). Note that two of the components have specific orientation.*

3.2 System Calibration

The system operates in three distinct coordinate frames: that of the robot, that of the scanner and a third, system defined, coordinate frame that is a corrected variant of the scanner coordinate system that allows components to be rotated and resized without distortion. Calibration procedures are adopted that allow transformation between these coordinate frames. The accuracy of the system is reliant on the pick-up location of a component from a supply station whose location is taught by the user as part of the calibration procedure.

Calibration is verified by a 'target practice' procedure that uses the complete set of transformations from the robot via the scanner to the data processing coordinate frame and back to the robot. It also uses the data from the component learning stage and employs the sensory data processing and matching algorithms. Thus the majority of the procedures employed during actual assembly are exercised, and this is used to verify system operation.

4 Pick and Place Experiments

These took the form of a series of tests, each of which investigated the performance of the system under increasingly challenging circumstances. Each test was performed on each of three product types, with the exception of product 'C' where test (i) (see below) is not applicable:

Product A: Three approximately circular cucumber slices are placed onto an approximately square slice of bread.

Product B: Two elliptical slices of sausage are placed onto a substrate consisting of a slice of sandwich bread and three previously placed cucumber slices. See Figure 3.

Product C: Two pieces of chicken breast meat are placed onto a bed of lettuce in a transparent plastic salad container.

In product types [B] and [C], component orientation as well as position is important.

There were four tests, one of which was used as a control to verify the system calibration:

Control Test. This used a single artificial substrate on which the exact desired positions of specific artificial components were marked. In the case of the chicken salad product, a horizontal plane was used in place of the lettuce, onto which was placed an artificial chicken breast slice.

Score	Component placement
0	Completely off substrate
1	Less than 50% of component on substrate
2	Greater than 50% of component on substrate
3	100% of component area on substrate, none on target
4	Less than 50% of component area on target
5	More than 50% of component on target but visibly off target
6	Small overlap with another component or over the edge of the substrate
7	Very good match, error only detected with respect to substrate edges or other components
8	Visually Perfect; exact placement with respect to substrate edges and other components

Table 1: *Scoring scheme for the quality assessment of assembled components. Note that errors resulting from components crossing the substrate boundary are always considered worse than other errors. Thus, where two scores are potentially applicable, the lower of them is awarded.*

Test (i): Constant Real Substrate, Artificial Components. This used the same slice of bread for both the learning and assembly stage. Artificial components were employed for both the learning and assembly stages. (This was perhaps the least useful test in relation to the eventual application but provided accurate measurements of the placement accuracy of the system.)

Test (ii): Varying Real Substrates, Artificial Components. Examples of different naturally varying substrates of different dimensions were employed. In the case of a salad container, the same container was used but the lettuce substrate was redistributed between each run of the experiment; thus the substrate also varied between learning and assembly stages. The same set of artificial components was used as for test (i).

Test (iii): Varying Substrate, Varying Components. Substrates and components were real food items, all of which exhibited natural variation between the items used for the learning and the assembly stages.

4.1 Assembly Quality and Accuracy Measurements

To assess the effectiveness of the system, criteria are required by which the quality of the resulting assemblies can be measured. In the case of the control test and test (i), measurements on the positional accuracy of the component placements can be made. When dealing with a range of naturally varying components, it is not practical or meaningful to measure placement accuracy in terms of positional and rotational errors in mm and degrees respectively. A scoring scheme to record the quality of component placement was therefore defined and, to allow subsequent verification of the score awarded, a visual record was kept for all taught products and assemblies by means of an independent conventional video camera and image capture system.

Marks for position and orientation of each component range from 0, worst case, to 8, perfect placement; the criteria are summarised in Table 1 and illustrated in Figure 4. In addition to the scoring scheme, textual notes are made explaining the reasons for a particular score being awarded.

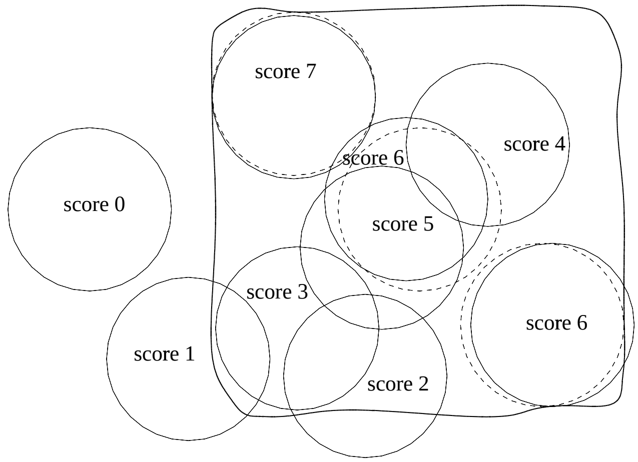

Figure 4: *Example scores for component placement quality. The dotted circles are the target locations for the three circular components used in Product A. Displacement of a component from its target position such that it crosses the edge of the substrate is considered to be worse than an equal displacement that retains the component within the substrate boundary (see top left and bottom right).*

5 Results

Results of all experiments are summarised in Table 2 and demonstrate that the system's accuracy is more than sufficient for food assembly tasks. Food products need to be assembled so that they 'look right' to a human observer. This typically translates to a placement accuracy of within about 5mm, although the tolerance is much less in critical areas, such as overlap regions, and can be greater in other circumstances. For the scoring system, scores of 6, 7 and 8 are well within the application's requirements. Scores of 5 can be classified as 'acceptable', but performance would be questioned if a significant percentage of components scored 5.

5.1 Comparison of qualitative and quantitative measures

It is informative to examine the relationship between the quantitative and qualitative scoring for the placement accuracy; this can be achieved for test (i). The mean quantitative measure for each qualitative score is presented in Table 3 and demonstrates that top scores of 8 indicate, on average, placement accuracies of less than 1mm while lower scores of 6, on average, translate to placement errors of approximately 3mm, well within the specified requirements of 5mm. For product B, the quantitative average for a score of 6 is less than that for a score of 7. This reflects correctly the subtlety in the qualitative measures whereby overlapping components are marked down to a 6 and considered of poorer quality than those which may exhibit a larger error but do not overlap. The scoring system relates to a judgement of appearance based on the criteria tabulated above, so it does not necessarily correlate with absolute placement error.

For product C in the control test and in test (ii) artificial components were used that corresponded to slices of chicken breast cut diagonally along the short length of the breast. These shapes have a unique orientation and therefore provided a stern test for the placement accuracy of the system. Because of the considerable variation of the characteristics of different chicken breasts, test (iii), was carried out with strips of meat cut along the length of the chicken breast. These approximated to elongated rectangles and components placed in either of the two possible orientations were accepted; the artificial product items served to indicate the capability of the system in dealing with components that have a unique orientation.

Product	Test	Assembled	Pos. err (mm)		Rot. err (deg)		Score								
			Mean	Max	Mean	Max	0	1	2	3	4	5	6	7	8
A	ctrl	40 × 3	1.44	7.00											
A	i	25 × 3	2.29	5.83			0	0	0	0	0	0	3	26	46
A	ii	50 × 3					0	0	0	0	0	1	24	56	69
A	iii	50 × 3					5	0	0	0	0	1	32	65	47
B	ctrl	45 × 2	2.13	5.83	3.26	15.00									
B	i	25 × 2	2.21	5.83	2.14	10.00	0	0	0	0	0	0	3	26	21
B	ii	50 × 2					0	0	0	0	0	0	8	60	32
B	iii	50 × 2					0	0	0	0	0	0	9	65	26
C	ctrl	30 × 2	1.96	7.07	2.50	10.00									
C	ii	50 × 2					0	0	0	0	0	8	0	49	43
C	iii	50 × 2					8	0	0	0	0	2	0	40	50

Table 2: *Results of the tests. The 'assembled' column indicates the number of assemblies undertaken for each test and the number of components placed for each of these assemblies. The mean and maximum placement error figures are in mm and the mean and maximum of the absolute rotation error is also noted for asymmetric components. The remaining columns indicate the number of components to which the respective scores were allotted. Note that the scores of 0 given to five components in test A (3.3% of those placed) and eight compnents in test C (8% of those placed) were the result of the components being dropped by our less-than-optimal gripper.*

Product	Average Component Score	Mean err. mm
A	6	3.18
A	7	1.96
A	8	0.80
B	6	2.77
B	7	3.01
B	8	0.60

Table 3: *Correlation of scored results with quantified error measurements for test (i). Placement accuracy was measured and the scoring system used for each component placed. Note that no components were given a score of less than 6 during the experiments and no test (i) experiments were conducted for product C.*

6 Discussion and Conclusions

We have conducted an extensive series of experiments on both artificial and real food items that are representative of the snack-food industry, with the aim of demonstrating the effectiveness of our approach to teaching an assembly system simply by presenting to it hand-prepared examples of the required output of an assembly stage.

Under control conditions (with artificial components) the system demonstrated an average placement error of 1.44mm, 2.13mm and 1.96mm for products A, B and C respectively, with the maximum errors within acceptable limits. With real food items, 96%, 100% and 90% of assemblies were given 'good' scores of 6 or higher. 3.3%, 0% and 8% of assemblies scored 0, resulting from failure of the gripper to acquire the component correctly. This is a consequence of our current policy of concentrating on the sensing and algorithmic aspects of the work, rather than developing specific tooling, which will be pursued separately via specialist collaborators.

Scores of 5 were given to only 0.7%, 0% and 2% of the assemblies, all relating to real food products and there were no occurrences at all of 'poor' scores 1 to 4. The few cases that led to the marginally acceptable scores of 5 were found to result from noisy difference surfaces leading the search into local minima.

Overall, the results indicate that, provided a component is grasped successfully, the system's accuracy in replicating product examples is well within the requirements for the snack food assembly applications at which the work is aimed. It therefore provides the basis of a flexible food assembly system in which a range of product types may be accommodated by showing hand-prepared examples to the system, without the need for reprogramming.

Acknowledgements

We thank Dr. Mark Neal for valuable and stimulating discussions, and Ian Izett, Colin Mudie and Rhys Hughes for assistance with workcell configuration. This work is undertaken in collaboration with Solway Foods Ltd. and RHP, with financial support from the Agri-Food Directorate of the UK BBSRC (grant 2/D09343).

References

[1] G. J. Agin and P. T. Highnam. Movable light-stripe sensor for obtaining 3-dimensional coordinate measurements. *Proc. Soc. Photo-Optical Instrumentation Engineers*, 360:326–333, 1982.

[2] H. Alt, U. Fuchs, G. Rote, and G. Weber. Matching convex shapes with respect to the symmetric difference. *Algorithmica*, 21(1):89–103, 1998.

[3] C. G. Atkeson and S. Schaal. Learning tasks from a single demonstration. In *1997 IEEE International Conference On Robotics and Automation - Proceedings*, pages 1706–1712, Albuquerque, NM, 1997.

[4] P. Chen. On intelligent robot motion planning via learning. *Journal of Intelligent & Robotic Systems*, 19(3):299–320, 1997.

[5] C. W. Desilva, R. G. Gosine, Q. M. Wu, N. Wickramarachchi, and A. Beatty. Flexible automation of fish processing. *Engineering Applications of Artificial Intelligence*, 6(2):165–178, 1993.

[6] H. Friedrich, S. Munch, R. Dillmann, S. Bocionek, and M. Sassin. Robot programming by demonstration (RPD): Supporting the induction by human interaction. *Machine Learning*, 23(2-3):163–189, 1996.

[7] S. Gunasekaran. Computer vision technology for food quality assurance. *Trends in Food Science & Technology*, 7(8):245–256, 1996.

[8] B. Heaton. The technology behind sandwich manufacture. *European Sandwich & Snack News - Review '96*, pages 39–41, 1996.

[9] K. Khodabandehloo. Robotic handling and packaging of poultry products. *Robotica*, 8:285–297, 1990.

[10] V. Klingspor, J. Demiris, and M. Kaiser. Human-robot communication and machine learning. *Applied Artificial Intelligence*, 11(7-8):719–746, 1997.

[11] J. S. Marques. A fuzzy algorithm for curve and surface alignment. *Pattern Recognition Letters*, 19(9):797–803, 1998.

[12] M. J. Neal, J. J. Rowland, and M. H. Lee. A behaviour-based approach to robotic grasp formulation: experimental evaluation in a food product handling application. In *IEEE Intl. Conf. on Robotics and Automation, Albuquerque*, pages 304 – 309, 1997.

[13] H. Ogata and T. Takahashi. Robotic assembly operation teaching in a virtual environment. *IEEE Transactions On Robotics and Automation*, 10(3):391–399, 1994.

[14] E. Ozcan and C. K. Mohan. Partial shape matching using genetic algorithms. *Pattern Recognition Letters*, 18(10):987–992, 1997.

[15] F. Storbeck and B. Daan. Weight estimation of flatfish by means of structured light and image- analysis. *Fisheries Research*, 11(2):99–108, 1991.

[16] J. D. Tedford. Developments in robot grippers for soft fruit packing in New-Zealand. *Robotica*, 8:279–283, 1990.

[17] N. D. Tillett, W. He, and R. D. Tillett. Development of a vision-guided robot manipulator for packing horticultural produce. *J. Agricultural Engineering Research*, 61(3):145–154, 1995.

[18] P. W. M. Tsang. A genetic algorithm for aligning object shapes. *Image and Vision Computing*, 15(11):819–831, 1997.

[19] T. G. Williams, J. J. Rowland, M. H. Lee, and M. J. Neal. Teaching by example in food assembly by robot. In *IEEE Intl. Conf. on Robotics and Automation, San Francisco*, pages 3247 – 3252. IEEE, April 2000.

[20] Q. S. Yang. Apple stem and calyx identification with machine vision. *Journal of Agricultural Engineering Research*, 63(3):229–236, 1996.