

# Geometric Reasoning via Internet CrowdSourcing

A.P. Jagadeesan, A. Lynn,  
J.R. Corney, X.T. Yan,  
C. Torres-Sanchez  
DMEM, University of Strathclyde  
James Weir Building  
75 Montrose St  
GLASGOW, G1 1XJ

jonathan.corney@strath.ac.uk

J. Wenzel, A. Sherlock  
The King's Buildings  
University of Edinburgh  
Mayfield Road  
Edinburgh, EH9 3JL  
wenzel.jan@gmx.net

W. Regli  
Drexel University  
3201 Arch Street  
Philadelphia, PA 19104  
regli@drexel.edu

## ABSTRACT

The ability to interpret and reason about shapes is a peculiarly human capability that has proven difficult to reproduce algorithmically. So despite the fact that geometric modeling technology has made significant advances in the representation, display and modification of shapes, there have only been incremental advances in geometric reasoning. For example, although today's CAD systems can confidently identify isolated cylindrical holes, they struggle with more ambiguous tasks such as the identification of partial symmetries or similarities in arbitrary geometries. Even well defined problems such as 2D shape nesting or 3D packing generally resist elegant solution and rely instead on brute force explorations of a subset of the many possible solutions.

Identifying economic ways to solving such problems would result in significant productivity gains across a wide range of industrial applications. The authors hypothesize that Internet Crowdsourcing might provide a pragmatic way of removing many geometric reasoning bottlenecks.

This paper reports the results of experiments conducted with Amazon's mTurk site and designed to determine the feasibility of using Internet Crowdsourcing to carry out geometric reasoning tasks as well as establish some benchmark data for the quality, speed and costs of using this approach.

After describing the general architecture and terminology of the mTurk Crowdsourcing system, the paper details the implementation and results of the following three investigations; 1) the identification of "Canonical" viewpoints for individual shapes, 2) the quantification of "similarity" relationships with-in collections of 3D models and 3) the efficient packing of 2D Strips into rectangular areas. The paper concludes with a discussion of the possibilities and limitations of the approach.

## ACM Classification Keywords

H5.2 Information interfaces and presentation (e.g., HCI) --- User Interfaces: Evaluation/methodology, Theory and Methods; H.5.3

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

2009 SIAM/ACM Joint Conference on Geometric and Physical Modeling (SPM '09), October 4-9, 2009, San Francisco, CA.

Copyright 2009 ACM 978-1-60558-711-0/09/10...\$10.00.

[Information Interfaces]: Group and Organization Interfaces – Web-based interaction.

## General Terms

Algorithms, Performance, Human Factors

## Keywords

Canonical view, 3D Similarity, 2D Strip packing, Geometric reasoning, Micro-outsourcing, Crowdsourcing, mTurk

## 1. INTRODUCTION

A recurring challenging in geometric modeling is to reproduce algorithmically human levels of discrimination in the generation, modification and retrieval of 3D shapes. The development of such computational tools is often difficult because the problems are frequently either ill-defined (e.g. location of depressions on a model) or producing multiple solutions (e.g. local operations such as surface tweaking) that require user input to select the correct answer. Examples of these sorts of problems are:

- Feature recognition: where interacting features often require a degree of subjective judgment to define their boundaries.
- Surface fitting and partitioning: where researchers aim to algorithmically generate "faces" that define a "good" or "natural" subdivision of a complex surface associated with scan data.
- Similarity measurement: where attempts are made to reproduce human perceptions of what constitutes 3D similarity in content based retrieval systems.

In these and many other cases, researchers require robust benchmark data against which to assess the performance of their automated systems in comparison to the aggregate judgments of humans.

This paper reports the results of three experiments designed to investigate the feasibility of using Internet Crowdsourcing for geometric reasoning and establish the speed, costs and quality of results associated with these investigations. Although the work started with the belief that the approach might be useful for the establishment of benchmark data-sets, Crowdsourcing has proved so effective that in many cases the authors have questioned if automated solutions are really required.

The rest of the paper is structured as follows: the next section details the experimental methodology, implementation and results of the following investigations: 1) the identification of

“Canonical” viewpoints for individual shapes, 2) the quantification of “similarity” relationships within collections of 3D models and 3) the efficient packing of 2D Strips into rectangular areas. The paper concludes with a general discussion of the accuracy and limitations of the approach, draws some conclusions and describes some of the authors’ future objectives. Full details of mTurk’s Crowdsourcing terminology can be found at [1] but the essential terms are: *Requesters*, the human beings that post tasks known as HITs (Human Intelligence Tasks) on the website; *Workers* (called *Providers*) who browse the list of HITs and after “accepting” one, complete it for a monetary payment set by the Requester.

## 2. GEOMETRIC REASONING HITs

We report 3 HITs with the following objectives

- 1) **Crowdsourcing of Canonical Viewpoints:** To investigate the Crowds ability to understand 2D images of 3D CAD models. This was a relatively constrained task involving selection from a limited number of alternatives.
- 2) **Crowdsourcing of Shape Similarity:** To investigate the Crowds ability to make subjective judgments about the relative properties of shapes. This was an open task involving several hundred models, with effectively a large number of possible answers.
- 3) **Crowdsourcing of Strip Packing:** To investigate the Crowds ability to optimize interactively a problem with many complex interactions and numerical measure of success.

The following sections describe each of these

### 2.1 Canonical Views

One of the important results established by psychophysical studies of image understanding in the 1980’s was the realisations that people recognised objects more easily from some orientations than others. Interestingly Palmer, Rosch and Chase, [2] established that such canonical views existed regardless of the degree of an observer’s familiarity with a 3D shape. In other words, studies found that certain views of familiar objects were recognized, or “understood”, consistently easier and faster than randomly chosen views of the same objects.

Today, although the phrase canonical (a.k.a *characteristic* or *representative*) view was originally coined by biological vision researchers, the term has been widely adopted by the computer vision community [2][3] in two principle areas of work:

- 1) Algorithms for automatically determining canonical view-points for a given model [5]
- 2) Algorithms for object identification based on matching against a discrete set of canonical view-points [6].

Despite the myriad of algorithms already reported, they are limited in their effectiveness and their used definition. For example researchers studying “aspect graphs” have developed definitions of canonical views based on concepts of “view stability” [4] while others have suggested a solid’s principle components can be used to define and compute such properties [7]. However, like many geometric problems, the informal and intuitive definition captures the concept much more precisely than complex mathematical representations. Because of this the

authors thought it is credible that Crowdsourcing could be used to identify canonical viewpoints with-out recourse to complex calculations.

#### 2.1.1 Experimental Methodology

Using the mTurk API, a HIT to determine the canonical view of a 3D object was designed and implemented (Figure 1). The task showed an animation of the object rotating alternately about two axes on the top left of the pages and to the right of this image lay a grid of 12 images. The HIT provider was asked to select the three “most representative views” in order. Each HIT consisted of 5 tasks (i.e. five different shapes) arranged vertically on a web page. A set of 20 HITs consisted of the same five shapes being presented (in randomly varied order) 20 times to different workers. Providers were offered \$0.15 per HIT.

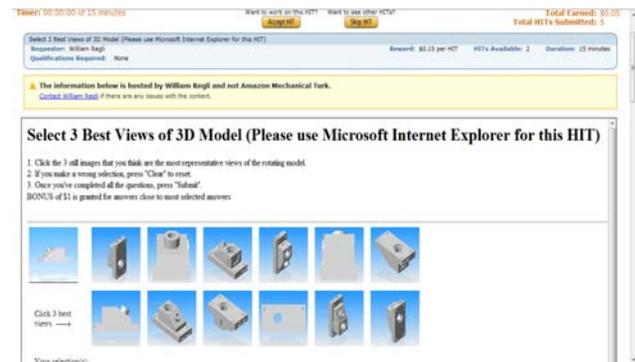


Figure 1. The Canonical Viewpoint HIT

The actual process of going online with four different sets of 20 HITs took place in two steps. First, we uploaded only one set of HITs in “mturk-Sandbox” (HIT simulation environment) to check HIT presentation, to check if results come back in expected order and to check any other error in code. When this simulation is successful, HIT is finally loaded in “mturk” for workers to answer. The first set worked without any problems and so further sets of HITs were uploaded.

#### 2.1.2 Results

**Speed:** All the HITs were accepted and results returned extremely quickly. Answers to the first 20 tasks (each classifying five shapes) were given within 1h 21min 53sec. A subsequent set of three further HITs was processed equally fast (Set 2: 23min45sec; Set 3: 41min22sec; Set 4: 42min41sec). This could be because returning workers “accept” the HITs faster without the need to read the instruction again.

**Acceptance rate:** Out of the 80 submitted answers only eight had to be rejected. Six of the rejections were based on technical problems relating to the type of browser being used (the HIT had to be done in Internet Explorer). The other two rejections were necessary because the workers (apparently) misunderstood the task and exclusively selected three orthogonal face views for all the questions. Overall 90% of HITs were approved and none of the work submitted appeared to be the product of a random choice.

**Quality of Individual Workers:** With each result submitted by a worker, the requester receives an answer, including other information about how the task was processed. One element of

this data is a unique “workerID” allowing the requester to distinguish between individual workers. Using this “workerID” it is possible to analyze how many different HITs each worker completed. At first glance, evaluating the quality of the results provided by an individual worker appears difficult because there are no absolute answers; the difference between “good” and “bad” views being subjective. However in the context of this HIT a good overall result would be a strong consensus between the workers.

Because of this we decided to motivate the workers by offering a bonus \$1.00 for the “best” results, where “best” is defined as being “closest to the average result of all the workers”. Our aim was not only to increase the general quality of results but also to make workers think about what others would consider as the best canonical views of the 3D models. In order to identify the best result for every set of HITs we first had to determine the average results of all the workers.

To do this we calculated how often each view was selected as 1st, 2nd or 3rd best view. According to these positions we assigned “points” and summed these “points” for every candidate view. Based on this calculation we could create a ranking of the best views of the model. Afterwards we compared the three best views of this ranking to the selection of each individual worker and granted the bonus to the worker with the minimum deviation.

Figure 2 shows results for one of the components (for more results refer section 4.3.2 in [21]) whose viewpoints were assessed by a HIT. The numbers in brackets are the number of mTurk workers who selected each view as most representative of the components shape. In this case, a bonus was provided to workers who selected images 9, 12 and 8 or 4.

Option 1 (6)	Option 2 (1)	Option 3 (2)	Option 4 (6)	Option 5 (3)	Option 6 (2)
Option 7 (5)	Option 8 (6)	Option 9 (9)	Option 10 (2)	Option 11 (5)	Option 12 (7)

Figure 2. Typical Results For Canonical Views

## 2.2 Shape Similarity

Various lines of development have been pursued to create optimal search-by-shape software which supports engineers looking for specific CAD-models in databases. Obviously machines have an enormous advantage in the speed with which they can analyze data. The challenge is to generate software that is able to mimic the human perception of shape [8]. Loncaric [9] describes the human visual system as one of the most sophisticated and versatile in nature. He states that its ability to understand the organization of surrounding nature is unsurpassed by artificially created reasoning systems. To do this the software developer would have to define the highly subjective term of “shape similarity”. Until that is done it is not possible to create a program that interprets a user's 3D search query correctly and retrieve appropriate “matching” models.

However neither of the words “shape” or “similar” have clear definitions in casual English. So a potential problem for a Crowdsourcing approach is that people’s interpretation of both “shape” and “similarity” could be dependent on their context. For example, a designer could think that two components are similar, while a manufacturing engineer might perceive no similarity because they are formed by totally different manufacturing processes [14]. Moreover the search-by-shape for particular parts is complicated by the fact that often similar looking parts fulfill completely different functions. Consequently, it is difficult for the search-by-shape software to distinguish between the different types of similarity desired. Given this, the authors were keen to determine if the judgment of anonymous mTurk workers could cope with the level of ambiguity inherent in the problem.

### 2.2.1 Experimental Methodology

Using the mTurk API, crowdsourced workers were given a HIT with a pool (i.e. grid) of still images showing isometric views of different CAD-models. It was requested that the workers “put similar looking models together into groups” by clicking on their images. The clustered images appeared below the initial pool of images (for detailed descriptions refer section 6.3 in [21]). The workers were asked to continue this process until there were no images left in the pool.

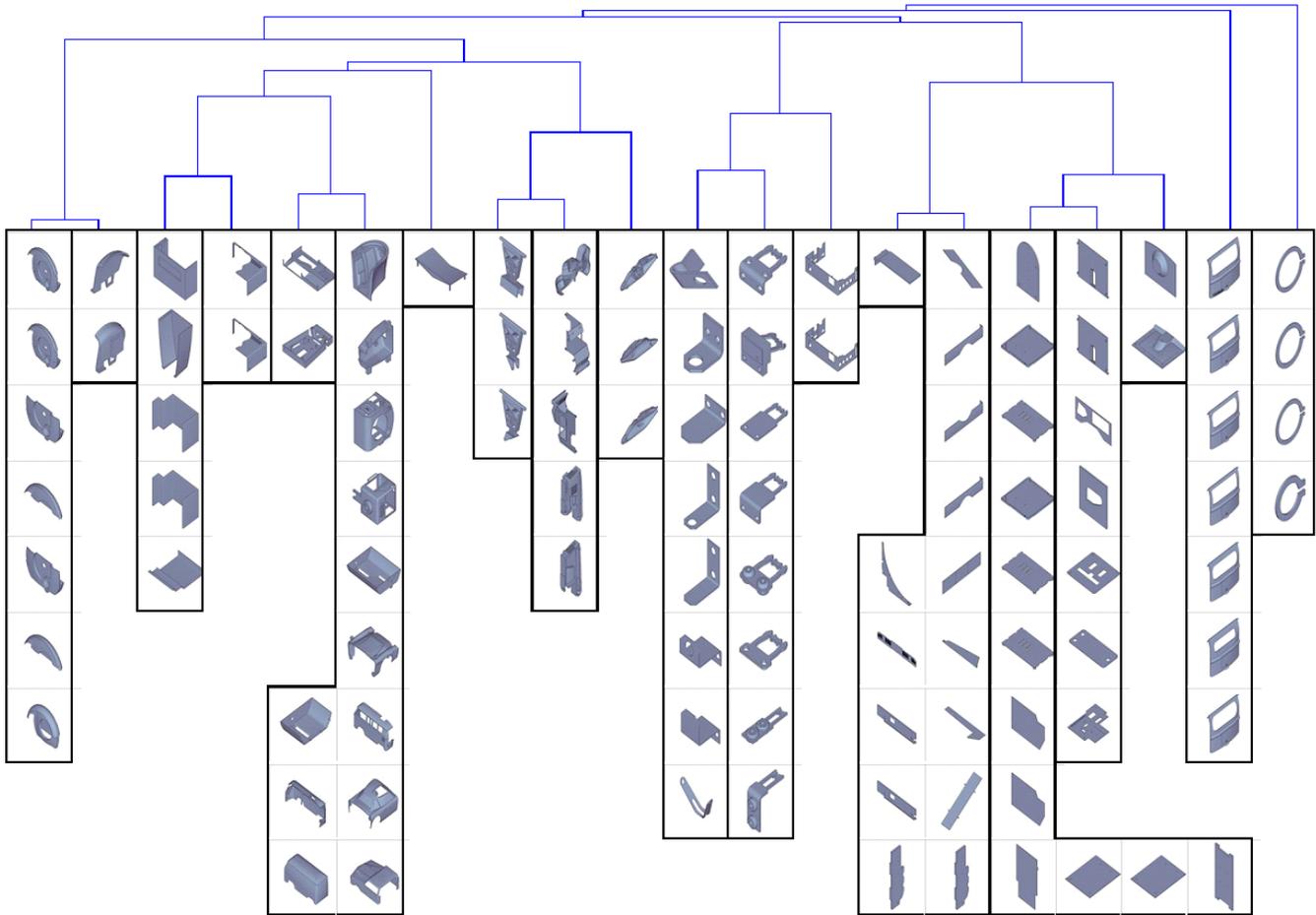
15 HITs (containing pools of 107 models) were posted and accepted within 8min 48sec, with results returned in an average of 1hr and 21mins. However the rejection rate was 29.4% (only 1% of these being for browser compatibility problems). Workers were paid \$4.00 for each HIT accepted.

The clusters identified by the workers were compiled in a similarity matrix, where the number of times each pair of parts was clustered together was counted. Given n different CAD-models, the matrix contained  $n*(n-1)/2$  values[12]. The cells of the matrix held a similarity measure (on a range from 0 to 10) for each pair of CAD-models used. Consequently, two parts could reach a maximum similarity of 10 (in cases where all workers grouped them together).

The “Statistics Toolbox” of MATLAB was used to generate clusters from the similarity matrix using the “average” linkage method [20] and dendrogram plotted. Figure 3 shows the clusters identified for the thin walled class of parts (107 parts). Similarly for large “rectangular prism” (281 parts) refer Figure A.2 in [20] and for much large “solids of revolution” (479 parts) refer Figure A.3 in [20].

To determine the effectiveness of these HITs, the clustering results of both crowdsourced workers and the published families of similar shapes (i.e. the “Purdue Engineering Shape” Benchmark’s collection of 3D models) were compared. The entire class of “flat-thin wall components” at the Engineering Shape Benchmark (ESB) is formed by 107 parts [14]. Within the ESB, these 107 flat thin wall components are divided into nine sub-clusters. These sub-clusters are the ones to be compared with this HIT in MTurk.

To discover if the approach can be scaled up, a second class of the ESB was tested. The “rectangular-cubic prism” contains 281 parts and these were presented to workers as a single HIT. 15 HITs were posted and all the HITs were accepted in 14min26sec and returned in 1hr1min8sec. As there are many (i.e. on one web



**Figure 3. 107 Parts in ESB “Flat-Thin Wall components” classified in an average of 21min 57sec**

page) parts to cluster, workers had to scroll the HIT page up and down many times. In order to encourage the worker and to produce a good result, every worker with approved answer was rewarded \$8.00. People who have done exceptional work were rewarded a bonus of \$4.00<sup>1</sup>. A dendrogram of the result is shown in figure 3.

Similarly, the third class of the ESB, “solid of revolution” was tested. It contained 479 parts. 15 HITs were posted and all the HITs were accepted in 15min02sec and returned in 1hr49min40sec. For every approved answer, a reward of \$12 was paid and for very detailed work a bonus of \$6.00 was granted. A dendrogram of the results is shown in figure A.3 in [20].

### 2.2.2 Results

A Comparison between ESB data and cluster-HIT data was made using a similarity matrix based method described in [13][21]. The results produced the following % similarity between the clusters in ESB data and the result clusters (Table 1).

The results show that the judgment of the Internet “Crowd” is not dramatically different from the Purdue University researchers and students, who originally classified the content of the ESB. Indeed

<sup>1</sup> These are generous pay rates in comparison to many HITs on mTurk, but we used the UK’s minimum wage rate of around \$9 an hour to set a “fair” rate

close examination of the differences in the identified cluster (ie family) membership between ESB and Crowdsource revealed that these were ambiguous shapes that could quite reasonably be assigned to either group (see figure 60, 61 and 62 in [21]). It was also significant that the HIT scaled from identification of similarities between 100 shapes to over 400 without any issues. Given an increased reward, the workers were happy to devote the time.

**Table 1 Percentage of difference in clusters.**

**“ESB data” vs “Cluster-HIT”**

ESB Benchmark Data	Approximate similarity between clusters-ESB vs HIT
Flat-Thin Wall Components	71.5%
Rectangular-Cubic Prism	85.8%
Solid Of Revolution	80.6%

### 2.3 Strip packing

Numerous manufacturing applications need to arrange variable numbers of arbitrary shapes into limited areas or volumes. For example, in the stamping of sheet metal, material cost can represent 75% or more of the total cost, consequently even small inefficiencies in material utilization can lead to large amounts of wasted material [15] Computation of a theoretically optimum

solution is known to be NP-complete and consequently numerous “good” and “near optimum” solution are used in practice. However the ability to improve on even a good solution can often have significant economic benefits (eg. less waste, higher productivity).

The packing problem occurs in many different applications from container transportation to the 2D stamping of complex profiles on CNC machines. For the purposes of this study we have focused on a particular sub-set of the general problem known as “strip packing” or “online packing”, where the goal is to pack a number of rectangular strips (of differing dimensions) into a rectangular area in the most compact way (ie minimize the height of the 2D packed shapes)[17].

### 2.3.1 Experimental Methodology

This HIT page was developed using “Flash” software (see [22]). The workers were provided with the task of packing rectangular boxes within a defined area. The criteria is, when boxes are stacked in the given space, the height of the space used should be as small as possible.

The dimensions of the boxes used in this HIT are from a strip packing benchmark data set [16]. The dark-green boxes can be dragged into light-green space provided. Workers can rotate the box and nudge the boxes together when closely packing. The height of the space used will be displayed dynamically while the boxes are packed, so that the worker can manipulate their moves to get an efficient packing. Once they are satisfied with the packing they can send back the results using the “Submit” button on the HIT page. Because of technical issues with the Flash implementation, workers had to input a code number to ensure they could be linked to their results.

### 2.3.2 Results

Initially 10 HITs were posted, paying \$1.00 per correctly answered HIT. We had some good results where workers managed to pack all boxes using an average of 32% height of the rectangular packing space, compared to a theoretical minimum of 26% (% of sum of areas of all boxes plus pixel boundary, divided by the area of packing space). On the second run (which included improved instruction), 15 HITs were posted, with further illustrated instructions, paying \$1.00 and bonus of \$0.50 for each 0.5% of reduction of *the best result established in earlier runs* (i.e. this HIT’s rectangular box dimensions and their positions in the packing space were taken from the best HIT returned to date).

After receiving a best result of 29.4%, another set of 10 HITs were posted. This time the best result came back as 28.125%.

Average time the workers took to complete this task was 18min57sec. Some sample results are shown in figure 4.

## 3. DISCUSSION

Discussions usually focus on inherent experimental errors and limitations of the data presented. However in the context of this work the computational boundaries are the distributed capabilities of the workers engaged with the mTurk site.

Although at first glance the results in sections 2.1 and 2.2 reveal surprising levels of consensus in subjective areas, anyone familiar with Surowiecki’s seminal work on Crowdsourcing [18] will have

read about numerous examples of a large group’s ability to arrive at accurate judgments even when faced with vague and ambiguous data.

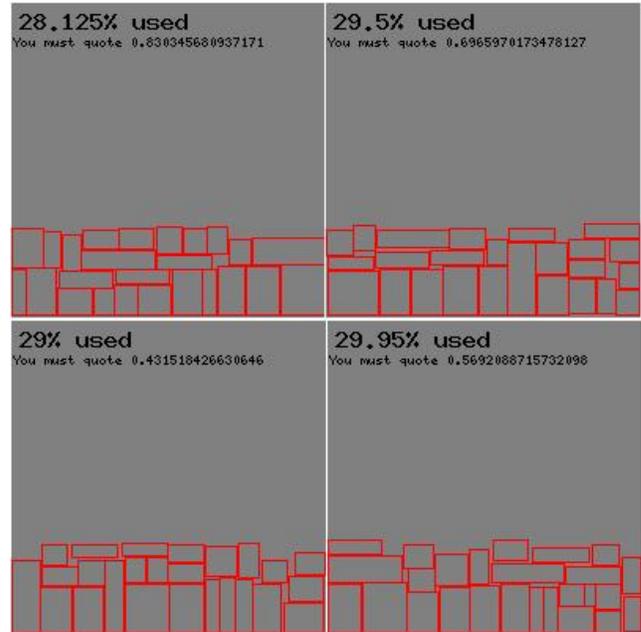


Figure 4 Sample Strip packing results

The strip packing HIT is different from the other in that it is not looking for an average, or consensus, solution, but rather it is seeking the best amongst many attempts. The HIT went through several re-designs; early versions which simply asked the user to pack the rectangles into the smallest area possible did not produce any solutions better than 31%. However when the best score and three best arrangements (so far submitted) were displayed to users and the payment level linked to the level of improvement (e.g 0.5% better \$1, 1% better \$2, 2% better \$4) the results pushed much closer towards the theoretical best, reaching 28.1% against 26% possible. Ongoing work has extended this HIT to allow the packing of complex 2D profiles, which will allow benchmarking of the results against published optimums.

Interestingly, the workers carrying out the HITs are keen to specialize in tasks they find easy or rewarding. For every one of our tests we received emails from workers requesting notice when more similar jobs were posted. Our suspicion is that given a regular stream of jobs, a pool of workers would form who were particularly talented at the task being offered.

## 4. CONCLUSION

The results demonstrate that there is a large, responsive work force available 24/7 capable of carrying out complex geometric reasoning tasks. Furthermore, our experience has suggested that this resource could cope with a large volume of work without reduction in the quality or response time.

Beyond the specifics of the results presented (ie times and costs) here, this work is believed to have established that Crowdsourcing provides a credible way in which a “human algorithm” (rather than a purely computational one) could be implemented in practical applications. Indeed the nature of the examples

presented implicitly demonstrates the limits of "algorithm alone" solutions. Because of this, the authors believe that a development of the methodology described here could create transformational processes for applications such as:

- Data mining: recognition of partial patterns in 2D and 3D models and assemblies.
- Reverse Engineering: Determination of an intuitive partitioning of an object's surface.
- Data Translation: Sanity checks and visual verification of results.
- Detail Removal: Feature recognition and suppression of visually or functionally "unimportant" geometry.

In the same way that research into the design of GPU algorithms (for exploitation of a powerful "black box") is regarded as a valid academic endeavor, the authors sustain that the production of methodologies for exploiting the "HPU" (human processing unit) will become an increasingly important area of research. As a first contribution to this new research topic we have described three different designs of HIT. Each of these went through several iterations to identify a workable way of posing the questions in a manner that could be crowdsourced. Our effort in the future will focus on designing HITs for geometric tasks that require true 3D interaction from the mTurk workers.

## 5. ACKNOWLEDGEMENTS

This work was funded by EPSRC Grant EP/F067291/1, "Micro-Outsourcing for Mechanical CAD/CAM".

## 6. REFERENCES

- [1] Amazon MTurk: DOI=[http://en.wikipedia.org/wiki/Amazon\\_Mechanical\\_Turk](http://en.wikipedia.org/wiki/Amazon_Mechanical_Turk) (Accessed 30th July 2008)
- [2] Palmer, S. E., Rosch, E., and Chase, P. 1981. Canonical perspective and the perception of objects. In *Attention and Performance IX*, J. Long and A. Baddeley Ed. Erlbaum, Hillsdale, NJ, 135--151.
- [3] Denton, T., Demirci, M.F., Abrahamson, J., Shokoufandeh, A., and Dickson, S. 2004. Selecting canonical views for view-based 3-d object recognition. In *Proceedings International conference on Pattern Recognition*. 23-26.
- [4] Weinshall, D, and Werman, M. 1997. On view likelihood and stability. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(2):97–108.
- [5] Yang, C.C., Marefat, M.M., and Johnson, E.J. 1998. Entity-based aspect graphs: Making viewer centered representations more efficient. *Pattern Recognition Letters*. 19,3-4(Mar. 1998), 265-277.
- [6] Mahmoudia, S., and Daoudib, M. 2007. A probabilistic approach for 3D shape retrieval by characteristic views. *Pattern Recognition Letters*. 28, 13(Oct. 2007), 1705-1718.
- [7] Perrett, D.I., and Harries, M.H. 1988. Characteristic views and the visual inspection of simple faceted and smooth objects: "tetrahedra and potatoes". *Perception*. 17(1988), 703–720.
- [8] Carlin, M. 1999. Measuring the Performance of Shape Similarity Retrieval Methods. *Computer Vision and Image Understanding*. 84 (2001), 44-61.
- [9] Loncaric, S. 1998. A survey of shape analysis techniques. *Pattern Recognition*. 31, 8 (1998) , 983-1001.
- [10] Jain, A.K., Murty M.N., Flynn P.N. 1999. Data clustering: A review. *ACM Computing Surveys*, 31, 3, 264-323.
- [11] Kuiper, F.K., and Fisher, L. 1975. A Monte Carlo comparison of six clustering procedures. *Biometrics*. 31, 3, 777-783
- [12] Johnson, S.C. 1967. Hierarchical clustering schemes. *Psychometrika*, 32, 3, 241 – 254.
- [13] Wenzel J., (2009), Feasibility study for the "crowdsourcing" of geometric reasoning in mechanical CAD. Institut fuer Fertigungstechnik und Werkzeugmaschinen, Leibniz Universitaet Hannover (Institution).
- [14] Jayanti, S. , Kalyanaraman, Y. , Iyer, N. & Ramani, K. Developing An Engineering Shape Benchmark For CAD Models. *CAD*. 38, 9(Sep. 2006), 939-953.
- [15] Venkata R.R. 2007. Evaluation of Metal Stamping Layouts Decision Making. ISBN 978-1-84628-818-0.
- [16] Mumford-Valenzuela benchmark data (2001): DOI=<http://dip.sun.ac.za/~vuuren/repositories/levelpaper/WebpageData/nice1.25.xls> (Accessed 3<sup>rd</sup> Feb 2009)
- [17] Hopper, E., and Turton, B.C.H. 2001. An empirical investigation of meta-heuristic and heuristic algorithms for a 2D packing problem. *European Journal of Operational Research*. 128 (2001), 34–57
- [18] Surowiecki, J. 2005. *The Wisdom of Crowds*. Anchor.
- [19] Validation of Purdue Engineering Shape Benchmark Clusters by Crowdsourcing. Jagadeesan, A., et al. PLM paper: DOI = <http://www.strath.ac.uk/media/departments/dmem/crowdsourcing/plm-paper.pdf> (Accessed 23rd June 2009)
- [20] Average, Complete and Single linkage - Dendrogram Appendix-A: DOR=<http://www.strath.ac.uk/media/departments/dmem/crowdsourcing/plm0-appendix-A.pdf> (Accessed 23rd June 2009)
- [21] Feasibility Study for the "Crowdsourcing" of geometric reasoning in mechanical CAD. Wenzel J. Tech Report: DOI =<http://www.strath.ac.uk/media/departments/dmem/crowdsourcing/TechRep08-01.pdf> (Accessed 23rd June 2009)
- [22] Screenshots of HITs described in this paper: DOI = <http://www.strath.ac.uk/dmem/research/crowdsourcing/> (Accessed 23rd June 2009)