Clickworkers results:
Crater Marking activity

July 3, 2001

The clickworkers project was funded as a pilot study to test whether distributed human volunteers are *willing* and *able* to collectively perform massive image analysis tasks for planetary science. The two main questions it is designed to answer are (1) whether the *quantity* of participation is worth the effort of setting up a site, and (2) whether the results are of sufficient *quality* to be usable. This report deals with crater marking in the Mars Digital Image Map (MDIM) made by USGS from Viking Orbiter images.

1 Quantity

In six months (December 2000 to June 2001), over 85,000 people visited the site, and many of them participated in crater marking and/or crater classification. Over 1.9 million crater-marking entries were submitted (1.5 million in the MDIM from Viking Orbiter, the rest in new MOC images) — enough for every sizable crater in the MDIM to be marked by dozens of different clickworkers. (The site contains roughly half of the MDIM, 30°S to 30°N, and regions were assigned randomly with an initial preference for one area.)

2 Quality

2.1 Comparison of clickworker consensus vs. expert

Figure 1 shows that the automatically-computed consensus of a large number of clickworkers is virtually indistinguishable from the inputs of a geologist with years of experience in identifying Mars craters (Nadine Barlow).

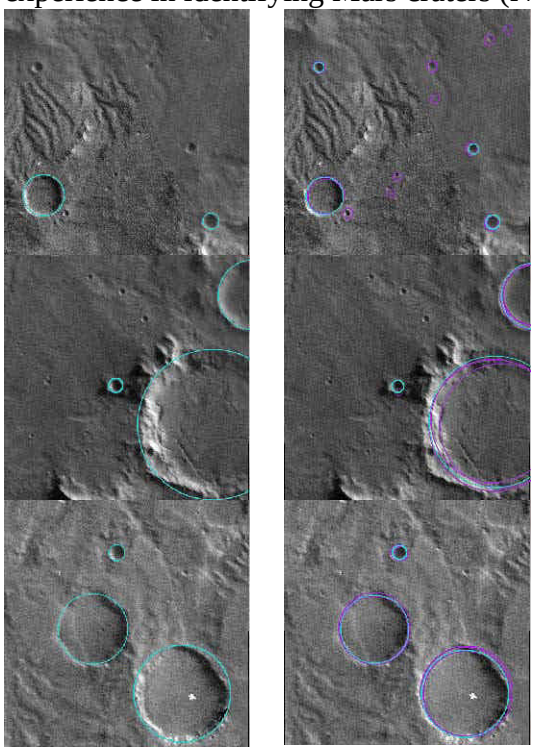
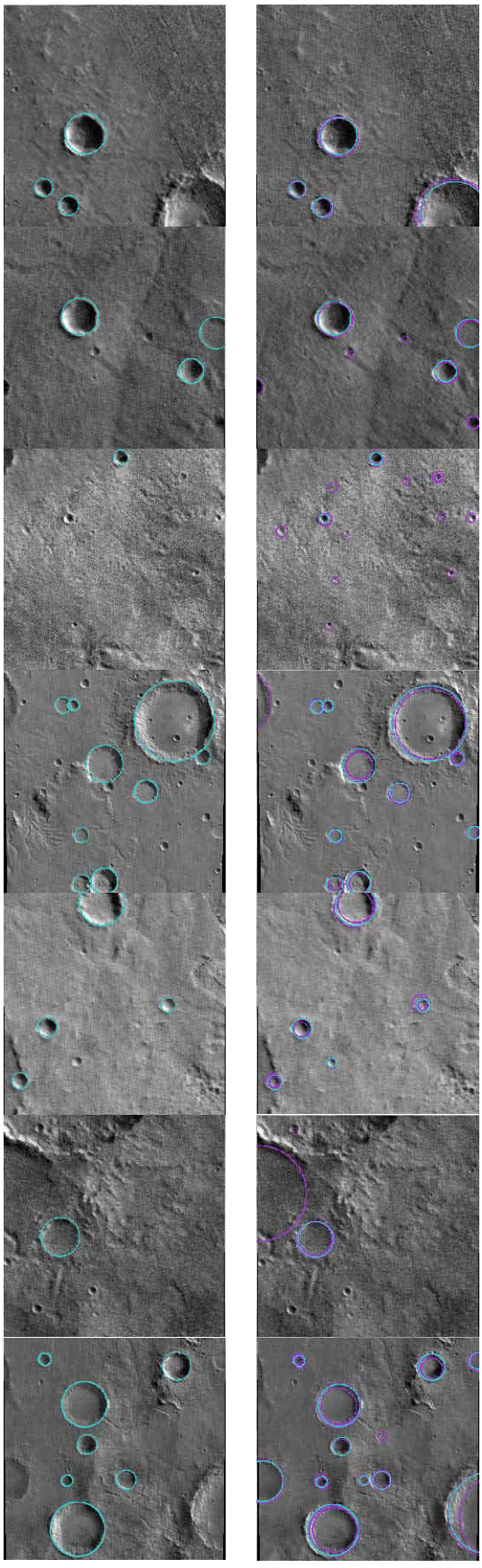


Figure 1.

Ten small regions of the Mars Digital Image Mosaic, all randomly chosen from within Margaritifer Terra, were marked by a crater expert, Nadine Barlow (left column) and by a large number of volunteer clickworkers (right column). Note that the cyan (light blue) circles match each other very closely. The cyan circles on the left are Barlow's input. The cyan circles on the right are the consensus of 200 clickworkers.

Since this was probably overkill, the consensus of a smaller number is also shown in purple as a sort of error bar indicating the amount of disagreement among the clickworkers. There are up to three purple circles on each crater, usually coextensive or nearly so. Each represents the consensus of a different group of 5



the consensus of a different group of 5 clickworkers.

The instructions suggested ignoring craters smaller than 16 pixels in diameter, but craters as small as 8 pixels were accepted. Collectively, the clickworkers erred on the side of thoroughness. Craters with purple circles and no cyan circles were marked by too small a fraction of the 200 to be considered legitimate with the weighting parameters chosen.

2.2 Handling of frivolous inputs

One failure scenario that was often imagined for this project was that some people would submit frivolous inputs. This has, of course, happened from time to time, but not often enough to affect the combined results. For example, several individuals (as many as 32) decided to fill the image with a set of concentric circles; 58 crater-marking sessions (out of over half a million) show this pattern. The site is not programmed to recognize this type of input or treat it any differently, but it's straightforward to find it offline. A colleague called our attention to it as a result of a web search. One of the 32 individuals who used this particular type of frivolous input did so as an assignment for an art class in college, and reported the results on the college's public web site. However questionable the artistic value of a set of concentric circles no human was likely to see, these cases serve as a good example showing that frivolous inputs are easily weeded out. Given enough redundant coverage, the system has no trouble distinguishing art from reality. A typical example is shown in Figure 2.

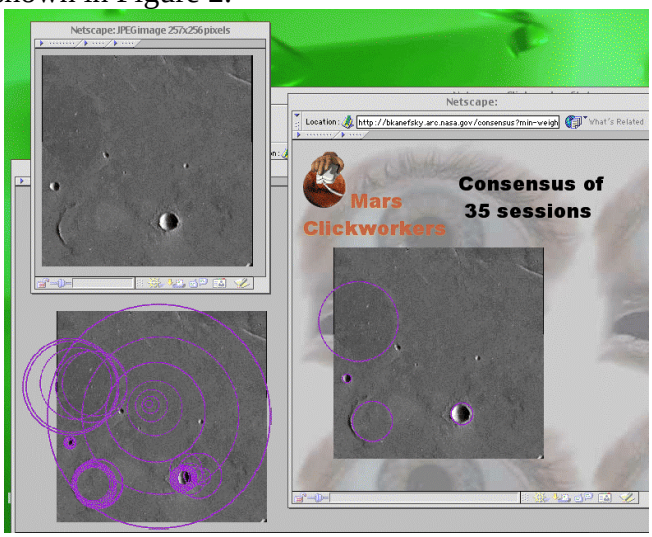


Figure 2. Frivolous input outvoted by the 34 other clickworkers marking the same image.
Lower left: all 35 clickworkers' inputs superimposed (concentric circles, plus circles aiming for four legitimate craters).
Right: Consensus.