# A Billion Points of Heat

by Jane Dietrich

**Left: The Orion Nebula, or Messier 42, one of 2MASS's greatest hits (which required only 10 minutes of observing time). The Trapezium Cluster in the center contains more than 3,000 bright, hot stars, the densest concentration of young stars in our solar neighborhood. Even younger, protostellar, objects appear in the small red region at the top. The bright, wispy clouds come from molecular hydrogen in dust-scattered starlight, detectable at two microns.**

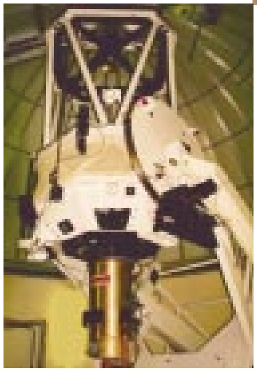
Infrared radiation—that segment of the electromagnetic spectrum with wavelengths longer than visible light, between visible light and microwaves—can pierce through the dusty universe to disclose quantities of unseen stars and galaxies; it reveals cool stars that radiate heat but no visible light, failed stars, and stars just being born in gaseous clouds. Although water vapor and carbon dioxide in the atmosphere absorb much of the infrared, satellites above the atmosphere, such as the Hubble Space Telescope and the Infrared Astronomical Satellite (IRAS), have extended our sight spectacularly. But there's also one window, at about 2 microns in what is called the near infrared (the infrared extends from wavelengths of 1 to 300 microns), where waves can pass relatively easily through the atmosphere and be detected from the ground. And at 2 microns, you're seeing light given off directly by stars, whereas at longer infrared wavelengths you see starlight that has been absorbed and reradiated by dust glowing in the space between stars.

Astronomers first surveyed the sky through the 2-micron window in the 1960s. But dramatic advances in detector technology in the last 20 years have made it possible to detect objects more than 80,000 times fainter than those discovered by the first 2-micron sky survey, and in the 1990s the National Aeronautics and Space Administration (NASA) and the National Science Foundation (NSF) funded the Two Micron All Sky Survey (2MASS), a collaboration between the University of Massachusetts and Caltech, that just finished gathering its data in February 2001.

The first survey was also a Caltech undertaking. Back in the early 1960s, Bob Leighton and Gerry Neugebauer aimed a telescope equipped with infrared detectors to scan as much of the sky as could be seen from Mount Wilson—about 70 percent of it. Leighton designed the 62-inch telescope's epoxy reflecting dish based on a principle he had first observed in his mother's mop bucket as a child, and built it in the back of his office (see *E&S,* No. 4, 1998). The infrared detectors were samples donated to Neugebauer by a friend in the defense industry; they had been developed after World War II for the heat-seeking guidance system of the Sidewinder missile. Their Two-Micron Sky Survey (TMSS), published in 1969, noted 20,000 infrared sources and cataloged about 5,700 previously unseen celestial objects.

"We thought it was a fun thing to do," says Neugebauer, now the Robert Andrews Millikan Professor of Physics, Emeritus. (Leighton, BS '40, PhD '47, the Valentine Professor of Physics, Emeritus, died in 1997.) In an era when infrared

**In April 1964, Bob Leighton's 62-inch two-micron infrared telescope (right) prepares to move into its "dome" on Mount Wilson, which is shown under construction in September 1963 at far right. The carpenter on the left is Jerry Nelson, '65, who later went on to design the 10-meter Keck Telescope's segmented mirror.**



**Top: The 2MASS 1.3-meter telescope in the southern hemisphere sits high in the Andes at the Cerro Tololo Interamerican Observatory. Inside the dome (above) the camera, with its massive arrays of detectors and filters for three wavelengths, is attached to the back of the telescope.**

astronomy was starting to grow rapidly, they also "happened to be there first," according to Leighton's oral history. Leighton's telescope is now a piece of history enshrined in the Smithsonian Institution.

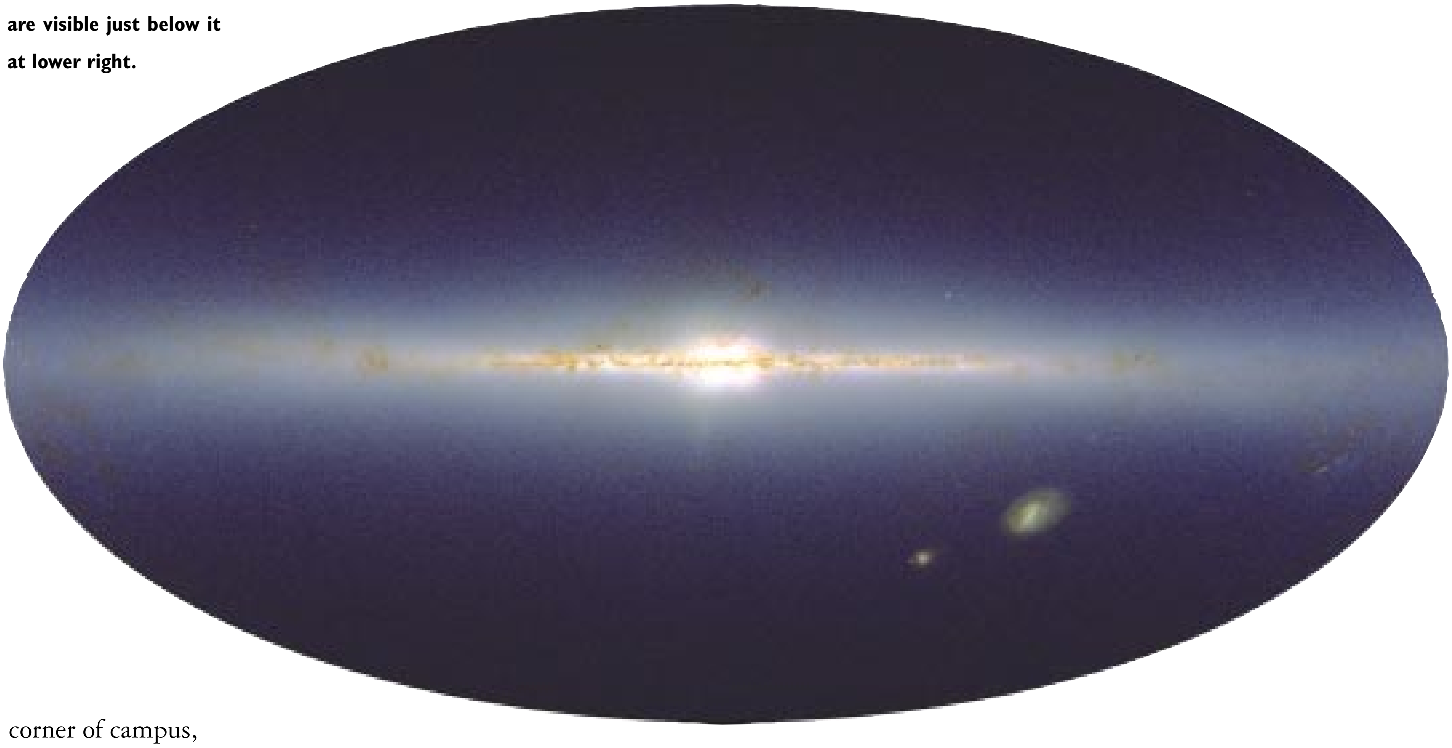
Infrared astronomy has long since left behind seat-of-the-pants technology and missile leftovers. 2MASS, says Neugebauer, "is everyone's dream of how you should do something, really do it right." The completely automated twin telescopes—one for each hemisphere—were designed and constructed by the University of Massachusetts team under astronomer Michael Skrutskie, principal investigator of the sky survey. At 51 inches (1.3 meters) they're a bit smaller than Leighton's telescope; smaller telescopes with large fields of view are better suited for covering the larger swaths of the heavens needed for a survey of the whole sky; huge mirrors like the 10-meter Keck Telescope, which gather faint light from the edges of the universe, peer through a pinhole in comparison. The northern-hemisphere instrument, on 8,550-foot Mount Hopkins (even at the 2-micron window, less atmosphere is better) in Arizona, started sweeping the sky in June 1997 and finished its observations in November 2000; its twin at the Cerro Tololo Interamerican Observatory, at about 7,000 feet in the Chilean Andes, started up in March 1998 and sent its last data to Caltech this past February. Each telescope's camera scanned strips of sky running 8.5 arcminutes wide and 6 degrees long before moving on to the next strip, or "tile." The whole sky is tiled by 59,650 such strips.

The heart and soul of the telescopes are their sensitive detectors, which represent the greatest leap in technology. Where the Mount Wilson model contained eight detector elements, a 2MASS telescope camera sports three arrays of 64,000 such elements, each with a filter for a particular wavelength. Like the earlier telescope, the new arrays are also a legacy, if not an actual hand-me-down, of defense research (for sensing heat from Earth rather than the sky). In their transfer from the military to astronomy, such infrared detectors made possible the Hubble Space Telescope's spectacular high-resolution images of infant stars forming in glowing, billowing clouds of space dust.

Each telescope has been producing a 20-gigabyte tape per night, and once a week for the past four years a stack of these digital linear tapes has been arriving (via Fed Ex) at the Morrisroe Astroscience Laboratory, home of Caltech's Infrared Processing and Analysis Center, or IPAC. Altogether there are hundreds of tapes, amounting to 25 terabytes ($10^{12}$ bytes) of data. For comparison, the photographic plates of the 1998 Palomar optical sky survey (this was also the second one, the first having been done in the 1950s), when digitized, comprised 3 terabytes of data, already several hundred times more than what was collected by IRAS, which was at that time one of the largest collections of data ever. How do you take such an incomprehensibly huge data set and turn it into something comprehensible: high-resolution images and an accurate, reliable catalog of half a billion objects? You turn it over to IPAC, which was created to analyze the IRAS data.

Caltech had stayed involved in infrared astronomy after Leighton and Neugebauer's pioneering work. Neugebauer, in fact, was U.S. cochairman of the joint science working group (which also included scientists from the Netherlands and the United Kingdom) for IRAS, which was launched in 1983 and collected data for 10 months. IRAS was managed for NASA by the Jet Propulsion Laboratory, but when it came to analyzing the mountains of data—a task that was to take several years—it became more convenient to move the project to campus as IPAC because of JPL's access restrictions, especially for noncitizens, and because of the great interest on campus in the data and their scientific interpretation. In the southwest

**Below: The whole sky, as seen in a 2MASS three-color, composite image. The Milky Way, edge-on, stretches almost all the way across, its center the bright bulge. Filaments of dust cut through the galactic plane, while the Magellanic Clouds, our nearest galaxy neighbors, are visible just below it at lower right.**



corner of campus, behind Braun gym, and lying low to appease its residential neighbors along Arden Road, the astroscience laboratory has been in operation since 1986, largely a mystery to the rest of campus.

"IPAC is unique as a piece of campus with special connections to JPL," says George Helou, IPAC's director. Although the funding comes from NASA, and the lab carries a bit of JPL "flavor," all of the lab's approximately 80 staff members are employees of the Division of Physics, Mathematics and Astronomy. "IPAC straddles the two cultures," Helou says, "the science and research culture of campus and the project and engineering culture of JPL. And 2MASS combines the science and engineering in unique ways."

2MASS is the first digital, electronic survey of the sky at relatively high resolution. While the recent Palomar Sky Survey turned to JPL's Artificial Intelligence group to develop a program to turn its photographic images into computer-digestible data for cataloging, 2MASS was born digital. What this project needed was robust software to turn its terabytes into images, as well as into useful and reliable catalog information.

The task fell to Roc Cutri, who, as task leader as well as project scientist, actually does live in both cultures. Cutri oversaw the design and implementation of the automated software pipeline, called 2MAPPS, which stands for 2MASS Production Processing System. To convert the raw digital data into calibrated images in three colors and to extract the quantitative information of the position and brightness of each object, Cutri and his team wrote 300,000 lines of computer code—not the usual pastime for an astronomer, but intellectually challenging all the same, says Cutri.

The three colors represent three different wavelengths in the near infrared: the J band at a wavelength of 1.25 microns, the H at 1.65 microns, and the $K_s$ at 2.17. (The infrared's photometric "bands"—the natural windows astronomers make use of—follow a rather loose alphabetical scheme; the longer wavelengths after K continue in a more regular fashion: LMNOP.) Classes of astronomical objects can vary in brightness at the different wavelengths, so simultaneously observing them at all three wavelengths makes identifying and interpreting them easier, says Cutri.

**Below: The Mount Hopkins telescope begins its night's work surveying the northern sky.**

**Right: In the Flame Nebula, NGC 2024, part of the Orion Molecular Cloud Complex, the near infrared reveals a dense stellar cluster, comprising stars likely surrounded by the accretion disks that could be sites of where planets form. Below: The Cat's Paw, or Bear Claw, Nebula (NGC 6334).**

But before the identifying and interpreting begin, the data must be fed through the software pipeline. While the telescopes finished their observations this past February, it will take six months to a year to finish crunching through the raw data to produce the visual and quantitative results. The full set of images—the whole sky—and catalog will be published next spring. Analyzing the data will take decades; Cutri reckons there's enough fodder for a 50-year legacy. "2MASS offers the statistical context to study a very large number of objects, and also the detail to follow up the interesting ones," he says.

In the meantime, half the sky has been accessible on line since March 2000, the single largest collection of astronomical data from a NASA mission. The final catalog will contain half a billion objects. The team actually tallied one and a half billion sources, but have winnowed down the number to those whose brightness and position can be guaranteed with great precision. "Astronomers need to be able to depend on the accuracy of the data," says Cutri, who as project scientist is responsible for ensuring the scientific quality of the data. He compares it to a phone book, which would be useless if one out of ten digits were wrong. With 2MASS, extraterrestrials could phone home—confidently—to any of nearly half a billion stars.

The best way to insure the scientific quality of the data is "to do the science yourself," according to Cutri. He was a collaborator in one of the early discoveries to come out of 2MASS: brown dwarfs, cooler than any ever before seen. Brown dwarfs are stars that never ignited, and therefore give off no visible light. They do give off some heat, however, which can be detected in the infrared. Already in 1998, in only one percent of what would be the total data, a team of 2MASS researchers led by Davy Kirkpatrick at IPAC found a number of these objects, which they named L dwarfs, revising a century-old alphabetic system

**Emissions from the planetary nebula NGC 3132 (above) are mostly in the $K_s$ band, which gives the ring around the wind-blown center its reddish color. In the center is a low-mass star like our sun, losing its outer envelope as it dies.**

**Top left: The remnant of a supernova (IC443) exhibits a bright blue arc of excited iron in the J band and a red ribbon of hydrogen in the $K_s$ band.**

**Top right: 2MASS reveals the nearby spiral galaxy Maffei 2, which lies in the "zone of avoidance" and was previously almost lost in the dust.**

(even more mysterious than the designation of wavelengths) in which stars are ranked from hottest to coolest as OBAFGKM—with L now at the end). Then came Caltech grad student Adam Burgasser's discovery of still cooler, fainter, methane-rich brown objects called T dwarfs. (Make that OBAFGKMLT.) Together, L and T dwarfs, scarcely bigger than Jupiter and collapsing under their own weight, are probably the most populous stellar objects in our galaxy, outnumbering real stars two to one, and some of them may also be our solar system's nearest neighbors.

Because the Milky Way itself contains so much dust, it has been heretofore impossible to see neighboring galaxies in the wonderfully named "zone of avoidance," which extends to 30 degrees on either side of the galactic plane. But 2MASS has cut through that murky zone and revealed many more galaxies in our local universe. It has also produced a high-resolution map of our galaxy, including the galactic center, which at other wavelengths is obscured by dust—and an exquisite census of the Milky Way.

One of the challenges of such an immense data set, says Helou, "is how to search through a billion sources to find the really interesting objects—the unusual stars and quasars." So perhaps the greatest importance of 2MASS lies in overlapping it with surveys at other wavelengths (for example the very short X rays on one end of the spectrum and the much longer radio waves on the other), the "synergy," as Cutri describes it, "of putting data sets together. The value of combined surveys goes far beyond that of a single one."

Postdoc Robert Brunner agrees. "The sum is greater than the parts," says Brunner. "You gain more from data by joining them together than using them independently." As project scientist for the Digital Sky Project, he's trying to mesh a bunch of surveys at different wavelengths—a computationally challenging task matching billions of sources at different resolutions all over the sky. He and his group are creating software to improve the accuracy of the matches. Brunner is also currently working with Microsoft Research on a Sky Server to merge data from 2MASS, the Palomar Sky Survey, and other surveys, approaching the problems less from a computer-science angle and more from the viewpoint of the scientific user. "We're starting to move on to something of service to the broader community." Ultimately, the information-rich dataset of these combined digital surveys will be accessible on line in a National Virtual Observatory.

Sometimes the overlap hasn't been foreseen. Richard Ellis, professor of astronomy and director of Palomar Observatory, is part of an international team conducting the 2-degree-Field Galaxy Redshift Survey (2dF) of a swath of the southern sky, using a novel instrument built by Keith Taylor (now a member of the professional staff at Caltech) at the Anglo-Australian Telescope near Coonabarabran, Australia. The team is seeking to measure the stellar population density and mass distribution of galaxies on large scales. They mapped the positions of 170,000 galaxies, but optical wavelengths cannot accurately discern the stellar content of galaxies. Infrared radiation provides a much better handle on how many stars are in a galaxy, because the infrared output is directly proportional to the number of stars. But until 2MASS, no infrared telescope had seen deep enough into space to get a fair sample. Combining the digital catalogs of both surveys (a process that was accomplished literally overnight) matched redshifts/distances from 2dF with 2MASS luminosity/population data for 17,000 galaxies. With that large a sample Ellis and his group could determine how many stars there are in the universe per unit volume of space, and concluded that the amount of mass in stars is only about 0.2 percent of the total mass needed to stop cosmic expansion. "The overlap led to the most accurate census of stars in the local universe," says Ellis,

**The infrared shows up a bright core in the radio source Cepheus A. Its massive young stars and molecular gas are completely invisible to optical wavelengths.**

**Opposite page: The Carina, or Keyhole, Nebula (NGC 3372) contains an unusually high concentration of young massive stars and stars that are still forming.**

"an achievement that neither survey ever imagined when it was originally conceived."

"It's gratifying to develop a product that other scientists can use," says Cutri. And not just scientists. 2MASS can turn anyone's computer into a "desktop observatory" through IPAC's Infrared Science Archive (IRSA), the Web-based system for accessing the databases. Besides the catalog of data, about 5 million images in the digital Image Atlas will also be available on line. Just enter the name of your favorite galaxy or pick a point on the sky. You can start off at http://www.ipac.caltech.edu/2mass/ and be sure to catch "2MASS Galactic Center: The Movie" on the home page. This is not its only role in show business: the *Star Trek Voyager* TV series has also snapped up some 2MASS images for space backdrops.

A topical session ("The Big Picture: Latest Science Results from 2MASS") at the American Astronomical Society's meeting in early June, held conveniently in Pasadena, highlighted the already substantial science that is emerging from the data on half the sky. The session spanned a range of topics including—besides brown dwarfs—asteroids, stellar populations in the Milky Way and the Large Magellanic Cloud (its nearest neighboring galaxy), other nearby galaxies, distant active galactic nuclei, and the cosmic near-infrared background radiation. As of the beginning of June, 174 papers had been published using 2MASS data.

Funding for 2MASS will continue for another year and a half, when the project will deliver its final products and then fold, says IPAC's Helou. "Then we'll find something else interesting and worth doing, something with a specifically Caltech point of interest." Already well under way is SIRTF, the Space InfraRed Telescope Facility, a satellite to be launched in July 2002 to observe the sky at wavelengths from 3 to 180 microns. 2MASS data are critical in laying the groundwork for the mission, which is administered by JPL for NASA. There's considerable Caltech scientific interest in this new venture, and the SIRTF Science Center will be located on campus, sharing the Keith Spaulding building with the remaining segments of Business Services that have not moved elsewhere. And IPAC will be analyzing the data. ☐

**The 2MASS team in the galactic center. Front row, from left: Schuyler Van Dyk, Diane Engler, Ron Beck, Eugene Kopan, Roc Cutri, Tracey Evans, William Wheaton, Robert Hurt, Sherry Wheelock, and Jeonghee Rho. Back row: Ken Marsh, Cong Xu, Howard McCallon, Tom Jarrett, Laurent Cambresy, J. Davy Kirkpatrick, John Gizis, and Raymond Tam. Not pictured: Brant Nelson, Helene Huynh, Tom Chester, and John Fowler.**