Visualization Viewpoints

Editor:
Theresa-Marie Rhyne

Google's Geospatial Organizing Principle

Michael T. Jones
Google

This article reviews the motivations, approaches, and accomplishments in applying genus loci—the sense of place—to advance Google's mission to organize the world's information and make it universally accessible and useful. Delivering information in a meaningful spatial context has proven an ambitious undertaking. Ideas from philosophy, psychology, and sociology shaped the design, and technical and logistical challenges abounded in the implementation. For this project, we (the Google geospatial team) wove long-standing wisdom, fresh technical thinking, and our vision of the future to create a global information landscape where people experience a uniquely real understanding of the answers to their questions. That the work is used millions of times per hour shows the result to be useful. We believe this usefulness, popularity, and power derive from our attention to the nature of human cognition.

Editor's Note

This Visualization Viewpoints presents the thinking behind Google Maps and Google Earth. Here, I provide some background on the author, who is now the chief technologist of Google Maps, Earth, and local search, to put his comments in proper context.

Michael became interested in computers at an early age and taught himself assembly programming at age 10. He began his contract programming career in junior high school, and was the mechanical engineering departmental programmer at a university while attending high school. In 1980, he joined Ikonas Graphics Systems to write microcode for the RDS-3000, the first programmable frame buffer and origin of the alpha channel. By the 1990s, he was the director of advanced graphics software at Silicon Graphics, building APIs including OpenGL, Performer, the Image Vision Library, Cosmo3D, and OpenGL Optimizer. In 1998, Michael and other colleagues left SGI to build a new kind of graphics platform (Intrinsic Graphics Alchemy) and a demo to exploit its power. That demo became a company named Keyhole, and when Google acquired Keyhole in 2004, the concept of Google Earth came into being.

Some view Google Earth as the most widely used 3D visualization tool ever created. It has introduced hundreds of millions of users to what have, in the years since Ivan Sutherland's Sketchpad, been concepts at the core of visualization research—projections, 3D user interaction, user feedback, motion models, level of detail, frame-rate management, view-dependent rendering, streamed textures, multimodal data composition, data-driven extensibility, and direct manipulation. Here, Michael shares the thinking behind Google's effort to fit decades of academic and commercial research into simple tools usable by anyone with a computer.

—Theresa-Marie Rhyne

The five w's

Guiding my childhood writing efforts, my father—a journalist by training—taught me to respect the five w's: who, what, where, when, and why. "These questions," he insisted, "must be answered before other details are introduced." Why? "Because that is how you earn the reader's attention!" His admonition was reinforced by Rudyard Kipling's *The Elephant's Child*. Concluding his tale of curiosity's dual-edged sword, Kipling shares a poem beginning this way:

> I Keep six honest serving-men:
> (They taught me all I knew);
> Their names are What and Where and When
> And How and Why and Who.

Like the poem, life shows the five w's to be more than grammar school or journalistic tenets on lead-paragraph information density. They mirror human nature itself—for, as both my father and Kipling realized, these questions constitute the rod by which we divine the important from the unimportant in every aspect of our lives.

Where is where?

All too often, the notion of where is the least usefully answered of the five w's even when it's the most important; not because the location is necessarily withheld, but because the answer—when provided—is given without sufficient context or tools for interpretation.

Consider the question, where is the IEEE corporate office? A Google Web search for "ieee corporate office" finds 1,050,000 matching documents in 0.12 second. Due to Google's PageRank, the first set of search results has the link to the IEEE's homepage. Exploring further yields the address as "3 Park Avenue, 17th Floor, New York, N.Y., 10016-5997 USA." Thanks to Web document search, we have our answer in a fraction of a second—but how well do we understand it?

Several addresses appear on the first page of the search results, with locations ranging from New Jersey to California and even Singapore. Seeing them refines our question and inspires further curiosity. Are any locations within reasonable driving distance of each other? How might we best navigate between them? What nearby businesses might be of interest?

 Published by the IEEE Computer Society

Moreover, how will we recognize the IEEE headquarters when we get there?

The knowledge

Such understanding is known as *the knowledge* among London taxicab drivers, whose reverence for the term reflects its essential value in that occupation. The knowledge is what separates locals from visitors and ignorance of it makes the simplest trip an ordeal, so why should only locals and taxi drivers understand these things?

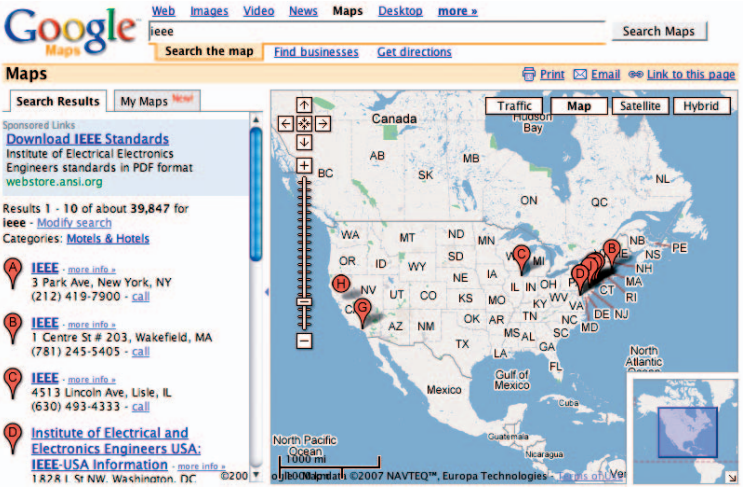
Researching this issue as application developers, our team found that a rich sense of where requires providing an interactive, exploration-oriented user modality; presenting information visually in its natural spatial context; and supporting search and search refinement in this same context—in essence, we learned that context is everything.

We built a place search engine that would search and rank a Web of places having the name "ieee" rather than using a document search engine to search and rank all documents with the text "ieee corporate office." Since the quest for where is intrinsically graphical, we search this Web of places directly from the map itself. Figure 1 shows what is displayed when we enter "ieee" in the "search the map" box of Google Maps. We find the addresses with an immediate contextual understanding and can analyze the results using the overview map inset, distance scale, map, imagery, and street-view display modes, and a responsive grab-drag-zoom map exploration user interface.

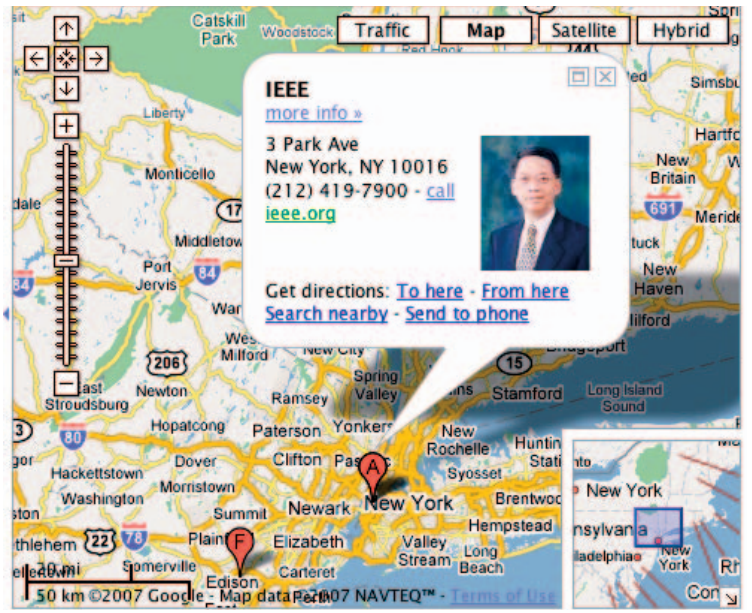
We see at a glance that locations G and H (IEEE Computer Society publications and an *IEEE Spectrum* editorial office, respectively) are on the west coast of North America; that a row of IEEE facilities is located along the eastern seaboard; and that C, the IEEE office in Illinois, is too remote to conveniently reach by car from either cluster. Due to PageRank, the corporate headquarters, A, appears first.

We can learn more by focusing our interest on one of these places. Let us click on A to explore the headquarters. Since where you look is where you search, this click is the place-searching analog of Web search refinement. It opens a Web page info window with further details, photos, relevant actions, reviews, and links to related data, as Figure 2 shows. At this closer zoom level, we can see major highways; connect to the headquarters via our telephone; send the details by SMS to a mobile phone for future reference; create a route plan to or from the location; access resources about this location such as its associated Web page ("more info"); and refine our search to the area near the HQ ("search nearby") for further information, such as where to buy a cup of coffee.

The three closest coffee vendors of the 24,075 nearby are Starbucks, Lavazza, and Guy & Gallard, as Figure 3 on the next page shows. A click on any of these results will connect the user via telephone, and a click on "more info" accesses the detailed information and further action options. This closer view reveals contextual details surrounding the IEEE HQ, such as detailed building shapes, street names, vehicle travel directions, real-time traffic status, and nearby subway station locations.



1 The facts in context.



2 Information window for IEEE corporate headquarters.

For a real-world visual context, the hybrid view shows photographic imagery overlaid with roads and markers. In this mode, we see the neighborhood. Now we know that the diagonally oriented office building of Figure 4 has a brown tone and stairway access from the Park Avenue sidewalk to the 33rd Street subway station.

We can learn even more. By clicking on "street view" we can navigate the streets of New York City at eye level. As shown in Figure 5, we are now "standing" at the sidewalk, where we can see what people see when exiting the subway station at the IEEE headquarters. From this view, we see that the brown tone of the building is a brick façade. The interactive nature of the panoramic images means that we can "walk" up and down streets of the neighborhood, turning in all directions to see what the locals see. Notice how real our sense of place has become, and correspondingly, how confidently we could make informed decisions within this realistic context.

This quick exploration immediately gives us precise understanding and general insights about the data.

3 Mapping the search for "coffee" near the headquarters.

4 Roads and coffee shops layered above aerial imagery.

Interactive place searching means that we can have the knowledge about any "where" we choose, be it New York or New Delhi; coffee or vaththalkuzhambu; the IEEE HQ, Shibuya Hachiko exit, or the Eiffel Tower.

Precision and generalization

The five w's help advance "the art of the utilisation of knowledge" when we embrace Alfred North Whitehead's educational theory as propounded in *The Organisation of Thought* and the essay collection, *The Aims of Education*.[1,2] We aim to help users rapidly select important information from less relevant information, and we offer information analysis tools so that users can discover implications contextually. Both tasks lie within Whitehead's second and third stages of learning—termed precision and generalization—and as in the IEEE HQ example, are facilitated by the five w's through the following thought processes.

Consider interpreting answers to the five w's as points or intervals along the respective *w*-axis. This view facilitates information rejection and selection by masking details using axially bounded regions easily compared to a 5D w-space bounding volume of personal interest. In this way, we first consider gross properties of the conceptual volume before specific *w*-axis values or the volume's content. (Readers with experience in efficient scene graph traversal and view-frustum culling might find this view of human cognition pleasingly familiar.[3])

For example, when lost keys are sought, the *when*-axis might be limited to times since the keys were last in hand, avoiding irrelevant recollection of places visited earlier. Framing our interest along the *w*-axes creates a framework to efficiently perceive, w-encode, and select the important from the unimportant even when presented with an otherwise overwhelming mass of data. Journalists who address the five w's in their headlines and introductory sentences let us know at a glance which newspaper story to read first.

Insight development—Whitehead's generalization—uses the *w*-axis framework as structure to exploit context. We need not step from an abstract bounding volume to specific details. We can perform contextual studies by partitioning the relevant axes into two subsets and consider one subgroup in the context of the other, using bounding volumes before details. On finding lost keys, we might recall other times when small items were misplaced and imagine points along the *when*-axis at the time of day when each loss occurred. Corresponding patterns in the *when*- and *what*-axes might tell us never to lay down small objects until after a morning cup of coffee from Lavazza; a what and where pattern might warn us not to place anything on the counter while paying at Guy & Gallard.

Partitioning the five w's into two groups (5:0 in one way, 4:1 in five ways, 3:2 in 10 ways) yields 16 opportunities to review subsets of the information in the context of the remaining data, the tasks implied by the phrase, I'll think about it.

The power of romance

Using the five w's this way means asking for information, getting relevant answers, selecting pertinent data from the results, and studying those contextually—using partitioning to compare and contrast subsets. For many people, this happens without conscious effort; awareness of importance and derived insights pop into our heads spontaneously. We believe this is precisely the aha! experience reported by our users and that it begins with the key initial precept of Whitehead's educational model.

Whitehead saw education as a cyclical process of stages—romance, precision, and generalization—with a reverent emphasis on romance, the emotional involvement on the part of the learner that initiates rewarding educational experiences. Just as children play with building blocks in simple wonder for years before noting patterns in their gravity- and friction-driven behavior and decades before learning an engineer's appreciation for stress and strain, Whitehead's romance is the play and wonder in our initial steps toward precision's deeper understanding and generalization's broader insights.

Leveraging the role in learning played by interest and curiosity, our user experience is meant to enhance the romance of unguided personal exploration, thereby enticing users toward a journey of understanding.

The planet of where

Our ultimate expression of the roles of romance and context in answering where? is Google Earth, a network-based browser for searching and displaying information in situ. It lets users navigate a geospatial Web of places in a planetary application metaphor with where and when as the organizing principles.

The vision for Google Earth inverts the roles of Web browser as application and map as content, resulting in an experience where the planet itself is the browser. Three-dimensional navigation lets users fly and swoop down to the Earth, where they search by looking, clicking, and typing. Search results show the world's information in context at its proper place as layers of georeferenced data; temporal 2D/3D content on the Earth; and pop-up Web page info windows defined using KML, the OGC best-practice way for users to author, find, publish, and share information.

Experience shows that users not only tolerate this inversion of the ordinary Web browsing experience, they joyfully embrace it, with more than 200 million unique activations to date. People can explore places from their past; family, ethnic, religious, and national histories; and sights of and above the Earth (see Figures 6a and 6b). Governments are now creating and publishing KML geospatial content for their citizens, visitors, and administrative purposes (for a 44,000-building, textured model of Berlin in KML, see http://www.3d-stadtmodell-berlin.de).

Cause and affect

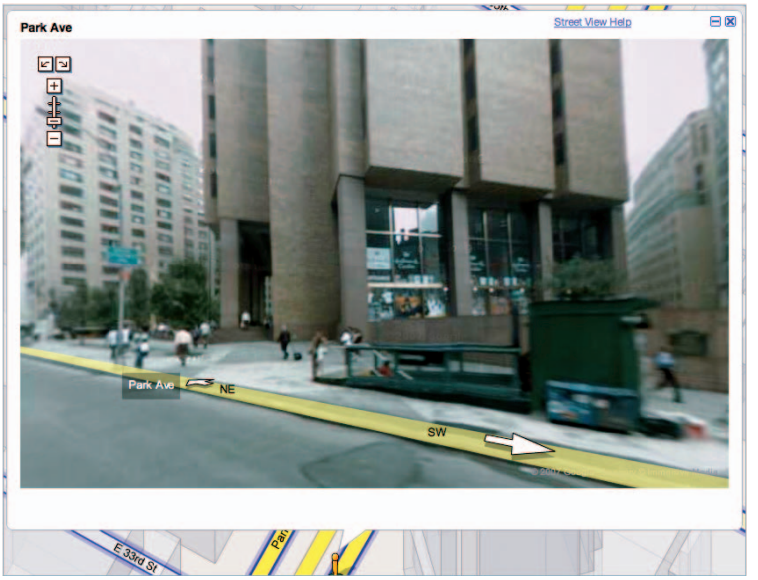
Frequently, users move from an initial quest for information toward deeper understanding and then onward to action. Examples include the following:

- natural disaster management and relief administration at the post-tsunami Indonesian shoreline, the hurricane and flooding sites in New Orleans, and the Kashmir earthquake damaged areas;
- environmental risk assessment and advocacy, such as a logging review in California's Santa Cruz Mountains, mountaintop removal in Kentucky, and primate habitat encroachment in Tanzania; and
- monitoring the spread of the avian flu's H5N1 virus and other world health risks on a global basis by *Nature* magazine and others.

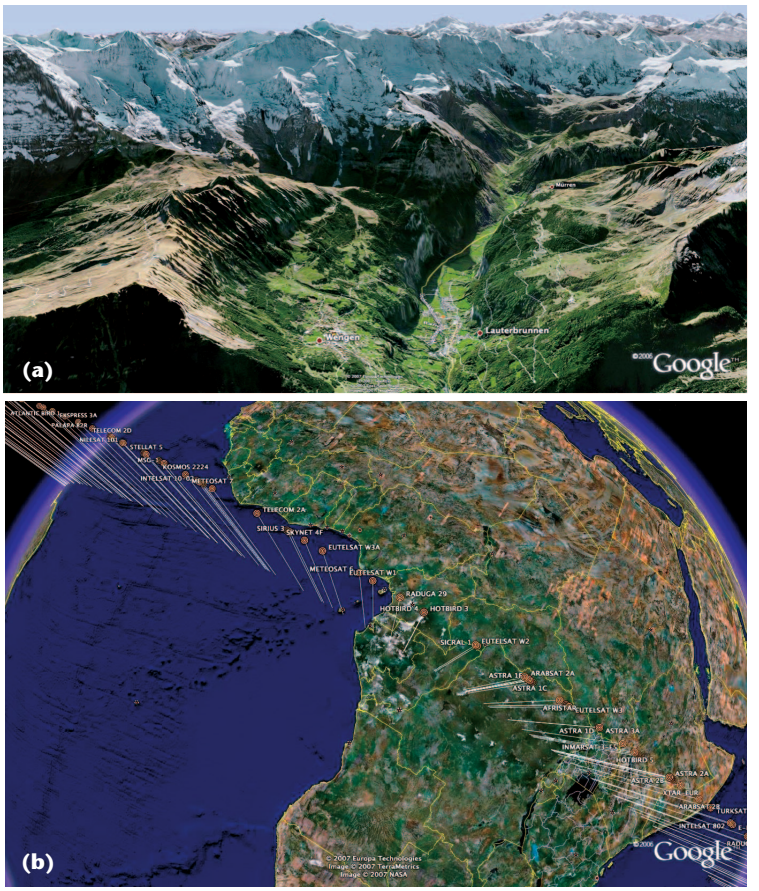
The constant in these examples is self-guided exploration of world events at both global and human scales—in the overview to sense the scope and in the minutia to feel the impact. In this way, Google Earth provides an experience where users understand remote, large-scale events as collections of individual, small-scale incidents with which they can relate and empathize, as well as possibly be inspired to action.

Nowhere is this affective experience more valuable than when exploring claims of genocide, global warming, and other topics that overwhelm our senses by their vastness, remoteness, or protracted timescale. To support such understanding we have worked with the National Geographic Society, United Nations Environment Programme, Jane Goodall Institute, Discovery Networks, the US Holocaust Memorial Museum, Smithsonian Institution, US Geological Survey, NASA, and others to bring substantive information about major issues into this understandable geospatial context.

As an example of the psychological processes behind these efforts, consider the situation in the Darfur area of

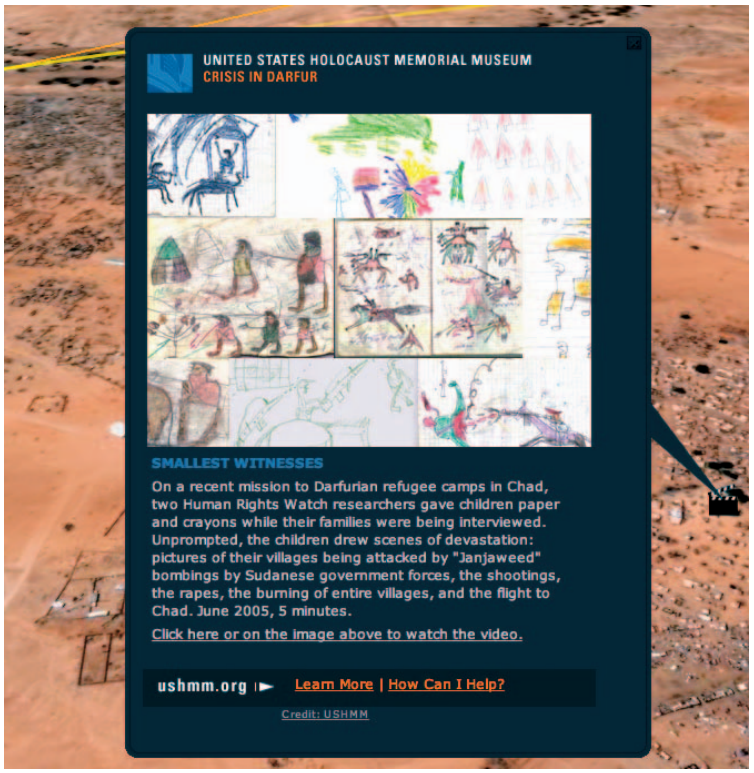


5 Street view of IEEE HQ at 3 Park Avenue.



6 (a) Switzerland's Wengen and Lauterbrunnen in Google Earth. (b) Geosynchronous satellites above the Earth's equator.

7 Destroyed and damaged villages in Darfur.



8 Visual storytelling by refugee children at this camp.

> Sudan. The summary of the situation is this: 2,500,000 civilians displaced, 300,000 people killed, and 1,600 villages destroyed. By collecting and publishing geolocated descriptions of places, events, photographs, and videos, Google Earth lets 200 million users explore this area personally. Flying above west Darfur yields the macro-level view of Figure 7, with flame icons receding to the horizon. Each flame represents a burned village, and clicking the icon reports the number of burned homes—as many as 1,200 per marker. Zooming closer reveals wells, burnt homes, empty cattle pens, and former social areas in these abandoned villages.

Moving to just above a refugee camp, we see where relief workers interviewed a group of survivor families (see Figure 8). As the adults spoke, their children drew on paper. These drawings are in the pop-up info window and linked video. Both show the children's unprompted depictions of shootings, burnings, bombings, rape, and murder.

Zooming to the human level of Figure 9, we move from an abstract awareness of 300,000 deaths to the remembrance of just one, M, as told by her brother, a Zaghawa man from Miski. Does news of her death feel less, more, or exactly 1/300,000 as powerful as the total?

This journey from western Sudan to M's home reflects established psychological research relating the importance of affect in associating meaning with information and motivating responsive action.[4] Affect is at the core of many dual process theories of thought, where the processes are broadly experiential and analytical systems.[5,6] In these models, the experiential system is more immediate and likely to lead an individual to act, but only when that person feels an affective, personal experience, such as when contributing to the welfare of an identified individual rather than a larger but impersonal group.[7] Affect theory explains why one death, as in Figure 9, can have greater power to motivate action than the much larger total figure of 300,000 deaths. Commenting on these data layers in Google Earth, US President George Bush remarked, "No one who sees these pictures can doubt that genocide is the only word for what is happening in Darfur—and that we have a moral obligation to stop it."

Judging the effort

Our geospatial products create an information landscape spanning global and human scales using visual context, direct manipulation, and Alfred North Whitehead's insights to advance Google's mission. We trust the sense of place to entice romance, facilitate precision, and encourage generalization as users search and explore our maps and globes. Google's networks and servers support millions of users. Our tools support databases that can exceed 10.92 petabytes.[8] We implement software for geospatial exploration within and without a Web browser, gather the best available data, and invent indexing and searching tools to place the world's information into a geospatial context.

Whitehead sought to remedy the woes of education to encourage a wiser society by inculcating an active knowledge throughout the world. He wrote, "The justification for a university is that it preserves the connection between knowledge and the zest of life, by uniting the young and the old in the imaginative consideration of learning."[2] We see Whitehead's objective being advanced as Google Maps and Google Earth answer where? millions of times hourly. By offering every Internet user a world of information—using genus loci as the ultimate university—we find our users both informed and affected. From this, we judge our

effort justified and the initial phase of our mission accomplished. ■

Acknowledgments

I thank Theresa-Marie Rhyne, the diligent anonymous reviewers, and my father, Barrett Glenn Jones.

References

1. A.N. Whitehead, *The Organisation of Thought*, Williams and Norgate, 1917.
2. A.N. Whitehead, *The Aims of Education and Other Essays*, MacMillan, 1929.
3. J. Rohlf and J. Helman, "IRIS Performer: A High Performance Multiprocessing Toolkit for Real-Time 3D Graphics," *Proc. ACM Siggraph*, ACM Press, 1994, pp. 381-394.
4. J.P. Forgas, ed., *Feeling and Thinking: The Role of Affect in Social Cognition*, Cambridge Univ. Press, 2000.
5. S. Epstein, "Integration of the Cognitive and the Psychodynamic Unconscious," *American Psychologist*, vol. 49, no. 8, 1994, pp. 709-724.
6. K.E. Stanovich and R.F. West, "Individual Differences in Reasoning: Implications for the Rationality Debate?" *Behavioral and Brain Sciences*, vol. 23, no. 5, 2000, pp. 645-665.
7. T. Kogut and I. Ritov, "The 'Identified Victim' Effect: An Identified Group, or Just a Single Individual?" *J. Behavioral Decision Making*, vol. 18, no. 3, John Wiley & Sons, 2005, pp. 157-167.
8. C.C. Tanner, C.J. Migdal, and M.T. Jones, "The Clipmap: A Virtual Mipmap," *Proc. ACM Siggraph*, ACM Press, 1998, pp. 151-158.

9 **Remembering a woman from the indicated village.**

Contact author Michael T. Jones at mtj@google.com.

Contact editor Theresa-Marie Rhyne at tmrhyne@ncsu.edu.