Geometric Analysis of SL(2,C) and Biologically-Mediated Computational Vision

Jacek Turski
University of Houston-Downtown
Department of Computer and Mathematical Sciences

The group SL(2,C) of 2×2 complex matrices of determinant one occupies a truly remarkable place in mathematics and sciences. For example, it is inherently relevant to non-Euclidean geometries, modern complex analysis, and Einstein's special theory of relativity. In our work, SL(2,C) provides unified geometrical and numerical framework for computational vision, including visual neuroscience and machine vision systems.

The conformal camera, which models eyes imaging functions, produces image projective transformations generated by the linear-fractional mappings of the group SL(2,C). Thus, the camera's underlying geometry can be dually described as (1) one-dimensional complex projective geometry and (2) the conformal geometry imposed by the holomorphic complex structure of the Riemann sphere, also called Möbius, or inversive geometry. Although this geometry does not possess a distance, it provides a full set of descriptors for the Gestalt rules used in grouping fragmented contours into global shapes that primate visual system must solve when viewing natural images—one of the most difficult problems to model numerically. The unity of geometrical and numerical methods is established by the fact that the conformal camera has its own projective Fourier analysis, geometric Fourier analysis constructed on the group SL(2,C) in the framework of representation theory of semisimple Lie groups—a great achievement of the 20th century mathematics. Projective Fourier transform (PFT) provides image representation well adapted to both perspective transformations of retinal images and the retinotopy of the brain's visual and oculomotor pathways. Thus, PFT integrates the head, eyes, and visual cortex into one computational system. We use this binocular system to process visual information during fast scanning eye movements called saccades, employed to build up understanding of scenes despite the acuity limitations of foveate vision. We make about three saccades per second at the eyeball's maximum speed of 700 deg/sec. Visual sensitivity is markedly reduced, as we do not see moving retinal images. Despite these incisive eye movements, the fragmented pieces of visual information are integrated in the brain into a stable percept of the world. This visual constancy is maintained by neuronal receptive field shifts prior to saccade onset in various retinotopically organized cortical areas. These shifts give the brain access to visual information at the impending saccade target prior to the eyes' arrival. It integrates visual information across saccades and eliminates the need for starting visual information acquisition anew three times per second at each fixation. However this remapping is not perfect; around the time of saccades, the flashed probes are not perceived in veridical locations by humans in laboratory experiments, a phenomenon called perisaccadic mislocalization.

In our modeling of perisaccadic perception, we utilize basic properties of PFT. First, the PFT can be efficiently computed by a fast Fourier transform in logarithmic coordinates that approximate the retinotopy. Second, a simple translation in retinotopic (logarithmic) coordinates, modeled by the standard shift property of Fourier transform, remaps the presaccadic scene into a postsaccadic reference frame. This shift also accounts for the perisaccadic mislocalization.

This research program is guided by a strategy important in the contemporary neurocomputing research: linking known anatomical and physiological details with efficient computational modeling should be vital not only to emerging field of neural engineering but also to interpreting relevant neurophysiological data.