Image Formation and Representation

CSC 249/449
Sensory Motor Systems

Randal C. Nelson

What is an Image?

Intuitively, a picture - i.e., a 2D representation of the 3D world.

Formally defined by an *image function*, prototypically f(x,y).

Produced from the world by a process of *projection*.

Represented in computer by process of *digitization* and *Quantization*

Generalizations:

- f may be vector valued (e.g., color images)
- f may use other coordinates systems (e.g., spherical)
- f may use modalities other than light (e.g. radar, x-rays)
- f may represent information other than radiation intensity (e.g., depth or density or motion)
- f may depend on time (i.e., an image sequence)
- f is almost always digitized.

What is image function?

How can the image function be characterized in general mathematical terms?

- Not continuous (image contain boundaries)
- Not piecewise continuous (boundaries show up at all scales)
- Fractal??

In real images, there are few regions without apparent discontinuity at any scale - they appear at one scale, vanish at others, and new ones show up.

- Fractal-like but fractal models too specific and hard to use
- Usual practical model is piecewise continuity (that changes with scale)

Projection

Projection geometry takes *world coordinate system* (3D, XYZ)) into *image coordinate system* (2D, xy)

Simplest model is planar images with linear projection, often referred to simply as *planar projection*. Several common models:

- perspective (pinhole model)
  - image plane
  - point of projection
  - focal length
- orthographic (parallel, engineering)
  - image plane
  - direction of projection

- scaled orthographic
  - image plane
  - direction of projection
  - scale factor
- para-perspective (multiple scaled orthographic, e.g., one per object)

Generalizations to basic notion of projection

- Non-planar images (spherical, cylindrical, hyperbolic etc.)
- projection via integration along curves (tomography)
- Use of non-linear projection curve families. Not often seen, comes up occasionally, e.g. trying to untangle gravitational distortion or refraction through atmosphere.

Representation of Images

Question: How do we represent an image function in a computer (or other symbolic medium)

Basic problem is that we need to represent a function, which mathematically can represent an infinite amount of information, with a finite number of symbols.

(Comment: Relevant infinity is aleph 2, where aleph 0 is the cardinality of integers, and aleph 1 is the cardinality of reals, i.e., very big)

Two basic processes:

- *Quantization*: Discretization in value
- *Digitization* (Sampling): Discretization in space (and time)

Quantization

Idea is that function can take on only finitely many values. Three general domains:

- Large: 32-64 bits per channel, a billion or more intensity values, "real number" representations. Not usually needed, but sometimes convenient for computational purposes)
- Moderate: 8-12 or sometimes 16 bits per channel, 256-4096 intensity values, "integer" representations. Good enough for almost all display purposes and many others)
- Small: 1 or 2 bits - binary, few-valued images, special purposes.

Digital numerical value is sometimes linear wrt. measurables (e.g. brightness); but more often not, especially with 8-bit reps.

Digitization (spatial sampling)

Digitization involves both a *sampling pattern*, and a *sampling process*

Sampling Patterns: image space is tesslated into discrete, local, compact, regions - neighborhoods, and value is associated with each one

- Regular, rectangular planar grid. Most common, 99% of images.
  - Good indexing characteristics
  - Problems with connectivity - two types of neighbors
- Other patterns (hexagons, triangles, offset rectangles)
  - Hexagons have better connectivity
  - Indexing is more difficult

- Non-regular grids, e.g., log-polar - inspired by retina.
- Non-planar sampling patterns (e.g. for sphere), often desired, no clear best solution.
  - Azimuthal (latitude/longitude)
  - Grid-mapped hemispheres
  - Image cube
  - Tessalated triangles (icosohedron)
  - Overlapping flat images

Sampling Process: How is value associated with grid location defined in terms of the underlying function? Not clear what best method is

- Point in neighborhood (e.g. center)
  - Convenient mathematically, theoretically
  - Serious aliasing problems
  - Hard to obtain with real sensor
- Average over neighborhood
  - Some aliasing
  - Easier to approximate with real sensor
  - Poor mathematical/theoretical properties
- Weighted average in overlapping neighborhoods (e.g. Gaussian)
  - Best solution to aliasing
  - Can have good mathematical properties
  - Not produced by basic physical sensors

Generalizing the Notion of Value

Digitizing the “same” scene twice generally does not produce exactly the same digital image. Always some uncontrollable variation (uncertainty) in values.

- Many sources of variation
  - Quantization error
  - Thermal noise in sensor
  - Poisson distribution of photon counts
  - Illumination flicker (esp. with fluourescent lights)
  - Platform jitter
  - Sensor temperature, external voltage fluctuations, dust in air, atmospheric distortion, etc. etc.

- It is often useful to quantify this uncertainty statistically
- Simplest method is specifying mean and variance.
  - Model may use constant variance or differ across the image.
  - May be fixed a-priori, or adapted on-line.
  - Pixel variance may be usefully correlated with intensity or neighborhood measurements (e.g. gradient), which allows variance values to be assigned to (predicted for) single images

- If more information is desired, image can be represented with vector mean and covariance for local neighborhood; Again, covariance may be usefully related to measured neighborhood values.
- Full mapping, even for 3x3 is to big to learn, so usual practice would be to functionally relate covariance matrix to smaller space (e.g. neighborhood mean and gradient)
- This approach allows certain potentially useful information to be captured. For instance if variation is due to jitter, then neighboring points covary in a way that could make gradient estimates more accurate than would be predicted on the basis of individual pixel variances that are assumed independent.

Other Image Representations

There are a number of other ways of representing an image function, not all of them spatially localized or sampled. None are produced directly from sensors. AMost have applications in pattern recognition and indexing, or texture analysis.

- Multi-scale/scale space
- Fourier transform
- Wavelets (think of local Fourier description in space and scale)
- Covariance representations
- Random field models (Markoff/Gibbs)
- Fractal representations

Digital-image acquisition devices

Hard-copy photograph + scanner

- Ordinary 35mm film can produce images on the order of 1000-2000 pixels square.
- Scanners
  - Drum, flatbed: old mechanical technology, slow, usually monochrome
  - Flying spot, laser: faster, monochrome or separate bands
  - CCD linear array: new, cheap, can handle color pics

Electronic cameras

- Old electron-bean scanning technology (TV cameras)
  - Image orthocon, vidicon
  - Big, expensive, tempermental
- Newer solid state technology - TV and consumer video and digital cameras
  - CCD (Charge Coupled Device) chips, CID (Charge Injection Device) chips
  - small, cheap, robust

More about CCD chips

- Array of cells, rows connected bucket-brigade fashion
- Charge builds up in cells, then is clocked out, row by row
- Standard is to produce analog signal, but digital signals can easily be produced.
- Color can be put on single ship
- 512x512 cameras the size of a pencil eraser are inexpensively available

Electronic Camera Parameters

- Chip physics/electronics
  - Spectral response - how does output value depend on wavelength
  - Photometric response - how is output value related to incident light energy (usually non-linearly)
  - Shutter speed - many modern devices can vary this
  - Gain - variable in many devices

- Comments on electronics
  - Modern cameras often utilize automatic gain control where shutter speed and chip gain are varied in an attempt to keep the output within some range.
  - Photometric response is also sometimes adjusted in an effort to maintain contrast.
  - Digitizers, when separate, sometimes perform additional processing; adjusting gain and photometric curve in order to produce a digital image that is useful for display purposes

- Timing (For TV/video cameras)
  - Frame rate (30 Hz is TV standard)
  - Interlaced versus progressive scan
  - Blanking periods (vertical and horizontal)
- Chip geometry
  - Number of rows
  - Number of columns/horizontal resolution (may be defined by digitizer)
  - Aspect ratio (may be defined by digitizer)
  - Skew

- Camera geometry
  - Focal length (field of view). May vary for zoom lens.
  - Center of projection
  - Rotation of chip coordinate system wrt. camera body
- Lens effects
  - Focus
  - Aperture (F-stop)
  - Depth of field (combined effect of focus and aperture)
  - Distortion ("fisheye" lens is extreme case)
  - Dimming towards boundary