**NBD(NIST Big Data) Requirements WG Use Case Template Aug 11 2013**

|  |  |  |  |
| --- | --- | --- | --- |
| **Use Case Title** | | Organizing large-scale, unstructured collections of consumer photos | |
| **Vertical (area)** | | (Scientific Research: Image Management) | |
| **Author/Company/Email** | | David Crandall, Indiana University, djcran@indiana.edu | |
| **Actors/Stakeholders and their roles and responsibilities** | | Computer vision researchers (to push forward state of art), media and social network companies (to help organize large-scale photo collections), consumers (browsing both personal and public photo collections), researchers and others interested in producing cheap 3d models (archaeologists, architects, urban planners, interior designers…) | |
| **Goals** | | Produce 3d reconstructions of scenes using collections of millions to billions of consumer images, where neither the scene structure nor the camera positions are known a priori. Use resulting 3d models to allow efficient and effective browsing of large-scale photo collections by geographic position. Geolocate new images by matching to 3d models. Perform object recognition on each image. | |
| **Use Case Description** | | 3d reconstruction is typically posed as a robust non-linear least squares optimization problem in which observed (noisy) correspondences between images are constraints and unknowns are 6-d camera pose of each image and 3-d position of each point in the scene. Sparsity and large degree of noise in constraints typically makes naïve techniques fall into local minima that are not close to actual scene structure. Typical specific steps are: (1) extracting features from images, (2) matching images to find pairs with common scene structures, (3) estimating an initial solution that is close to scene structure and/or camera parameters, (4) optimizing non-linear objective function directly. Of these, (1) is embarrassingly parallel. (2) is an all-pairs matching problem, usually with heuristics to reject unlikely matches early on. We solve (3) using discrete optimization using probabilistic inference on a graph (Markov Random Field) followed by robust Levenberg-Marquardt in continuous space. Others solve (3) by solving (4) for a small number of images and then incrementally adding new images, using output of last round as initialization for next round. (4) is typically solved with Bundle Adjustment, which is a non-linear least squares solver that is optimized for the particular constraint structure that occurs in 3d reconstruction problems. Image recognition problems are typically embarrassingly parallel, although learning object models involves learning a classifier (e.g. a Support Vector Machine), a process that is often hard to parallelize. | |
| **Current**  **Solutions** | **Compute(System)** | | Hadoop cluster (about 60 nodes, 480 core) |
| **Storage** | | Hadoop DFS and flat files |
| **Networking** | | Simple Unix |
| **Software** | | Hadoop Map-reduce, simple hand-written multithreaded tools (ssh and sockets for communication) |
| **Big Data  Characteristics** | **Data Source (distributed/centralized)** | | Publicly-available photo collections, e.g. on Flickr, Panoramio, etc. |
| **Volume (size)** | | 500+ billion photos on Facebook, 5+ billion photos on Flickr. |
| **Velocity**  **(e.g. real time)** | | 100+ million new photos added to Facebook per day. |
| **Variety**  **(multiple datasets, mashup)** | | Images and metadata including EXIF tags (focal distance, camera type, etc), |
| **Variability (rate of change)** | | Rate of photos varies significantly, e.g. roughly 10x photos to Facebook on New Years versus other days. Geographic distribution of photos follows long-tailed distribution, with 1000 landmarks (totaling only about 100 square km) accounting for over 20% of photos on Flickr. |
| **Big Data Science (collection, curation,**  **analysis,**  **action)** | **Veracity (Robustness Issues)** | | Important to make as accurate as possible, subject to limitations of computer vision technology. |
| **Visualization** | | Visualize large-scale 3-d reconstructions, and navigate large-scale collections of images that have been aligned to maps. |
| **Data Quality** | | Features observed in images are quite noisy due both to imperfect feature extraction and to non-ideal properties of specific images (lens distortions, sensor noise, image effects added by user, etc.) |
| **Data Types** | | Images, metadata |
| **Data Analytics** | |  |
| **Big Data Specific Challenges (Gaps)** | | Analytics needs continued monitoring and improvement. | |
| **Big Data Specific Challenges in Mobility** | | Many/most images are captured by mobile devices; eventual goal is to push reconstruction and organization to phone to allow real-time interaction with the user. | |
| **Security & Privacy**  **Requirements** | | Need to preserve privacy for users and digital rights for media. | |
| **Highlight issues for generalizing this use case (e.g. for ref. architecture)** | | Components of this use case including feature extraction, feature matching, and large-scale probabilistic inference appear in many or most computer vision and image processing problems, including recognition, stereo resolution, image denoising, etc. | |
| **More Information (URLs)** | | http://vision.soic.indiana.edu/disco | |
| **Note:** <additional comments> | | | |