Clive Fraser
Department of Infrastructure Engineering & CRCSI
University of Melbourne, Australia

"SLAM, SFM AND PHOTOGRAMMETRY: WHAT'S IN A NAME?"

"Clive, why so many names?"

In order to answer this question satisfactorily, it is first useful to consider the nature of the problem/concept/processes/procedures implied by SLAM, SfM and Photogrammetry within the respective scientific communities. They use these names to cover distinct approaches to solving essentially the one problem: the simultaneous localisation (position and orientation) of a sensor with respect to its surroundings, while at the same time building a 3D map of those same surroundings. Simultaneous Localisation and Mapping, or SLAM, is a concept that is largely sensor agnostic, but in this commentary the sensor will be assumed to be a camera, so we will be addressing so-called monocular visual SLAM. As is well recognised, both Structure from Motion, or SfM, and photogrammetry are also image-based localisation and mapping processes, since they both involve the use of multiple overlapping images in geometric arrangements varying from linear sequences to 3D largely unordered assemblages. They establish the relative orientation (RO) between the images, while at the same time building 3D point cloud data (generally sparse as opposed to dense at the RO stage), all within a single coordinate reference frame.

"Clive, what makes SLAM SLAM, SfM SfM and photogrammetry photogrammetry?"

The answer to this question is nuanced, although the case of SLAM is perhaps the easiest to address. The unique feature of monocular visual SLAM is not so much that the camera pose and map structure are recovered when neither is initially known, but it is that this concurrent process is carried out recursively, in real-time. This implies dynamic sensing, with the map being built in such a way that it supports in an on-line process the pose determination of newly acquired images, along with the triangulation of new points to support further new image orientation, and so on. Does this sound a bit like on-line aerial triangulation from the 1980s or real-time single-sensor photogrammetry which is feasible today? Well, yes, it does to some extent, though the use of 'structured' image measurement is at odds with the SLAM notion of completely unknown initial surroundings, and neither of these two techniques has 'taken off' anyway because of a lack of compelling need within the photogrammetric community. It is not surprising, therefore, to see that SLAM concepts have emerged from and are being further developed within areas such as robot self-exploration, autonomous vehicle navigation (e.g. planetary rovers) and, in a non-photographic geomatics context, hand-held laser scanning.

In removing the real-time aspect of monocular visual SLAM, which can also be referred to as visual odometry, we are left with SfM and photogrammetry, but more-so SfM. The term Structure from Motion infers the building of 3D structure using camera movement. The localisation/relative orientation/extrinsic parameter determination of the resulting multiple image assemblages is then established through feature detection, subsequent feature correspondence determination through descriptor-based matching, and orientation algorithms, desirably linear. SfM is seen as one of the crowning achievements of computer vision, and rightfully so, but it must be kept in mind that the development of this feature-based matching and network localisation technique had little to do with image-based 3D measurement, i.e. photogrammetry, and more to do with spatial archiving of unordered assemblages of images, often from unknown cameras – recall the Virtual Tourism initiative. SfM has stressed camera localisation rather than point cloud generation (more the domain of multi-view stereo within the computer vision community), and there has been recognition of the need for a certain measure of camera calibration. However, the engineering principles forming the foundation of photogrammetry, such as geometric network design, metric camera considerations, accuracy optimisation and variance propagation, systematic error compensation and gross error detection, scene-independent camera calibration, the adoption of observational redundancy to enhance

reliability, and quality control procedures at all stages of the photogrammetric data processing pipeline, all underpinned by rigorous, invariably non-linear mathematical models, have been conspicuous by their absence from SfM developments. It could be well argued that a data processing pipeline that is initiated with SfM-based camera localisation, but then follows the principles of photogrammetry, is no-longer SfM per se, but indeed standard photogrammetry. To call a photogrammetry measurement an SfM solution not only implies that the solution is less 'metric' (i.e. accurate and reliable) than may be the case, but also attributes more to the SfM process than what is actually there.

"Clive, what has photogrammetry gained from SfM?"

This might provide an insight as to why we see the two names being used interchangeably, and often erroneously, in some areas of image-based 3D reconstruction. The simple answer is that SfM has provided a powerful new approach to solving the image point correspondence problem, starting from the point of 'unknown surroundings'. The subsequent localisation aspects have not had much impact in photogrammetry, since the necessary rigorous sensor orientation algorithms have long been available. Exposure in photogrammetry to the utility of RANSAC processes for initial value determination and data filtering can also justifiably be associated with SfM. The raison d'etre for photogrammetry has always been 3D measurement of a photographed object or scene, and SfM does not and never really has inferred a process of dense 3D point cloud generation to specified metric tolerances. It is the aspect of metric measurement within object space that remains the inherent distinction between photogrammetry and SfM, as is indicated implicitly even in computer vision by the separation of Multi-View Stereo from SfM.

"Clive, what is the take-away message?"

The names of SLAM, SfM and photogrammetry do mean different things, notwithstanding the presence of conceptual and indeed algorithmic overlaps. And, given that we are currently at a photogrammetry conference, we should acknowledge unequivocally that neither SLAM nor SfM encompass the full process of metric-quality imaged-based 3D measurement, object reconstruction and mapping that is today's highly automated photogrammetry. Moreover, the 'lazy' use of the term SfM to describe black-box, automated exterior orientation and 3D point cloud generation process that can be opaque to its user, for example in the generation of a 3D textured model of a scene imaged from a UAV, should be strongly discouraged within a conference such as this. **'SfM' may sound sexy, but it ain't photogrammetry**!