TOWARD SMART CITIES THROUGH AUTONOMOUS DETECTION AND RECONSTRUCTION: THE BRAZILIAN PANORAMA ON THE 3D MAPPING

Rodolfo Georjute Lotte, Luiz Eduardo Oliveira e Cruz de Aragão e Yosio Edemir Shimabukuro

Instituto Nacional de Pesquisas Espaciais (INPE)
Avenida dos Astronautas, 1758, Jardim da Granja, São José dos Campos, São Paulo
Divisão de Sensoriamento Remoto (DSR)

ABSTRACT

Nowadays, by the use of intelligent techniques, every single letter, word, face, car, urban object, and many others, can be easily interpreted by a computer once it is filmed or pictured. Observing people movements, car traffics, are examples of information that can be automatically mapped by a computer. Once interpreted, thousands of insights are generated, and the way the things are currently implemented could be positively redirected. In case of strategic planning, the mapping of urban elements can reveal much about the urban context, thereby helping to determine the application of environmental and urbanization policies. In this paper we discuss the benefits on the availability of devices, platforms, the wide-spread technologies in terms of cloud services, intelligent algorithms on detection and reconstruction. What has been done to associate all those alternatives to better project urban planning? Or how to associate each of these extracted information and use them to analyse or even predict daily life situations? The discussion presented in this document, takes Brazil as object of study to show how the recent technologies will help the government to better serve the population.

Key words *– 3D mapping, Machine Learning, 3D reconstruction, Smart cities.*

1. INTRODUCTION

In the last decade, there has been a maturation in research for large-scale 3D reconstruction of cities, such that the crucial point was to acquire rough buildings representations and their characteristics of roof (coverage), occupation, area, height and volume [1]. As there is a technological growth of sensors, of course, there are also positive changes in the quality of the imaged scenes. Consequently, new lines of research are created and exploited in order to establish new domains and understanding of urban space. Recently, and thus mature, a new demand in the areas of Photogrammetry and Remote Sensing is leading the research to further analysis of these urban objects, in which a mix of technologies, and more sophisticated remote sensing devices to analyze details surrounding this environments, then, to acquire more consistent maps.

Not long ago, the representation of cities through virtual scenarios were mostly directed to entertainment application, rather than Cartography or with support to any detailed analysis. However, since mid-2010 to the present day, the combination of two research fronts has motivated the mapping of cities, with resources ranging from the use of more sophisticated hardwares to large-scale detection and 3D reconstruction softwares. The appearance of LiDAR (Light Detection and Ranging) [2], Structure-from-Motion (SfM) and Multi-View Stereo (MVS) workflows [3], for instance, brought to real the possibility to acquire the volume of things, since small until bigger urban elements. With this remarkable stage, today, new branches of research try not faithfully represent the scene, but mitigate new ways to add knowledge to it, increasingly toward to semantic cities.

The multifaculties that 3D urban mapping has been passing on, has made a merge to other fields of research, which also use optical data. For example, Computer Vision and Pattern Recognition community had always a permanent problem on acquiring knowledge from remotely sensed data, which basically has the mission of interpreting huge amounts of data automatically.

Until mid-2012, extracting any kind of information from images would require methodologies that would certainly not fully solve the problem. With the resurgence of Machine Learning (ML) technique in 2012 [4], built on top of the original concept from 1989 [5] about convolutional operations, has changed the way to interpret images due its high accuracy and robustness. Today, the robustness of interpreting images can be rebuilt with the volumetry of the same objects detected before, with a concept a bit more ambitious: organized these huge amounts of interpreted data in semantic and well defined maps.

2. 3D MAPPING: WHAT IS NEEDED AND WHAT IS AVAILABLE

We splited the next sections in three sub-sections: focusing in equipment (Section 2.1), in application (Section 2.2), and, finally, in computational facilities (Section 2.3).

2.1. Structural data

Structural data can be interpreted as every dataset that carries information about shape and volume of objects, e.g. point clouds from laser scanners (LiDAR) [6], from radar (RAdio Detection And Ranging) by the use of Persistent Scatter Interferometry (PSI) technique [7], and others. These datasets, usually represent geometric values very close to the reality, varying very little between one sensor to another.

In Table 1, is shown a schema regarding conventional types of structural data acquisition and the maximum Level of Details (LoD) [8] that each configuration could reach in terms of urban objects. The list makes a balance between their respective costs, knowledge required to manage the equipment, accuracy, spatial and spectral resolutions, among others. It is important to note the respective table is a non-exhaustive list, which highlights only common forms of acquirement. Other forms of acquisition, such as civil construction cranes, onboard balloons, helicopters, and any other resources are not listed.

The imaging of urban environments, are carried out conventionally by orbital, aerial (long, medium and close-range) and terrestrial platforms. Orbital platforms usually have a common goal, the imaging of areas whose characteristics does not vary often, e.g. forests and oceans. Although in orbit, some sensors have centimeter scales, which places it as one of the alternatives in the urban monitoring. Observing Table 1, three different orbital configurations have characteristics whose facade observation is possible, but only fine resolutions would allow them to observe the details. When available, both images and laser measurements are rare and inaccessible to the urban context.

Aerial platforms, are categorized here into three segments: long, medium and close-range. Surveys whose sensor is onboard airplanes, are usually considered as long-range surveys. They are, in fact, the most used category over the years, allowing sophisticate imaging, sometimes, with hybrid sensors and configurations that adequately serve the innumerable urban activities.

The estimated cost shown in Table 1 is relative and may vary according to area coverage. For example, the cost of orbital laser imaging or high-resolution optical sensors requires a high cost, but may benefit from covering an area whose terrestrial imaging would cover only in parts. The estimated cost in table, therefore, is the absolute cost.

On medium-range, remote data naturally has a gain in resolution, in addition to enabling large-scale imaging, e.g. farms, forests and others. For urban purposes, only sensors with wide FOV or multi-views gives guarantees of imaging these areas. In Brazil, however, this category still goes through standardization and other bureaucratic procedures for practice, although it is regularly offered by numerous companies.

Close-range surveys consists of small and medium-sized platforms. This specific configuration has gained interest in areas such as Agriculture and Cartography. UAV (commonly called drones), are usually used for its flexibility, low cost and stability in flying over narrow paths. These qualities, makes this a right alternative to image all the faces of a particular building (e.g. roof, facades, inner gardens). These small devices are easily purchased on the market for recreational use, but depending on their physical characteristics, they can also be adapted for scientific studies. The handling of laser scanners, however, requires a certain expertise, not just onboard UAVs, but on any other platform.

The terrestrial platforms complete the range of alternatives in close-range imaging. Although the ground survey does not require the use of an aerial platform, this type of imaging has almost the same properties as the close-range aerial survey. In this category, the images are taken with lateral geometry, at the lower sight level, interesting configuration for facade analysis, but not feasible when the purpose requires the complete coverage of the building.

The green color, shows the features benefited by the respective configuration, while yellow determines the dependence of some technical factor. In blue, the maximum LoD reachable by the respective configuration, where the lighters denote non-detailed 3D models. In indoor surveys, measures are taken as a complement to lower levels, for instance, LoD2 and LoD3. Of course, the LoD4 model shown would be obtained in addition to an existing model. Measures indoor by itself, do not support the generation of LoD1, LoD2 or LoD3.

2.2. Building geometry in Brazil

Technological development models in some countries, such as Germany, allow accelerated integration with research institutes and companies, which consequently leads the population faster access to innovative instruments. In a cascading effect, making new devices or technologies available to a community enables you to evaluate more detailed situations and improve the solution. In other words, the solution becomes feasible and accessible on a larger scale

In Brazil, some sectors related to Cartography still face challenges regarding the use of close-range instruments on urbanized areas. In addition, there are efforts by Geographic Service Directorate (in Portuguese, *Diretoria de Serviço Cartográfico* (DSG))[1] to standardize and elaborate specifications for urban 3D mapping. Even though, Brazilian urban mapping is moving towards more sophisticated levels [9].

2.2.1. What has been done

Driven by the need to understand the current state of Brazilian 3D urban mapping, we perform a sort of questionary among all Brazilian capitals. The goal in this study was to understand more the local mapping infrastructure, the availability of resources for their urban planning, and in how an accurate 3D urban model would help in the management of the city.

Then, the poll was elaborated on the use of 3D maps for the strategic planning of municipalities and sent to the secretaries of each Brazilian capital. The capitals were used as reference for the research, therefore, when reporting as being a "non-user" of any 3D information, it was considered that all cities referring to the respective

[1] Available at http://www.dsg.eb.mil.br. Accessed October 18, 2018.

Table 1: Maximum LoD and quality of 3D urban models according to platform and sensor.

		Plataform	Spectral region	Spatial resol.	Sensor view	Building parts: Roof	Facade	Indoor	Max. LoD	Costs	Operated	PC	Software	Accuracy
Orbital		Satellite	Optical	•••∘	Multi-view	✓	×	×	LoD2	$$$	🎓	⋆⋆⋆	-	•••∘
		Satellite	Laser	•••∘	Multi-view	✓	×	×	LoD2	$$$	🎓	✓	📥	•••∘
		Satellite	Hybrid	•••∘	Multi-view	✓	×	×	LoD2	$$$	🎓	✓	📥	•••∘
Aerial (multiple ranges)	long	Airplane	Optical	••∘∘	Nadir	✓	✓ ⋆	×	LoD2⋆	$$$	🎓	⋆⋆⋆	-	••∘∘
		Airplane	Laser	••∘∘	Nadir	✓	✓ ⋆	×	LoD2⋆	$$$	🎓	✓	📥	••∘∘
		Airplane	Hybrid	••∘∘	Multi-view	✓	✓	×	LoD3	$$$	🎓	✓	-	••∘∘
		Airplane	Optical	••∘∘	Multi-view	✓	✓	×	LoD3	$$	🎓	⋆⋆⋆	-	•••∘
	medium	UAV	Optical	••∘∘	Multi-view	✓	✓	×	LoD3	$	🎓	⋆⋆⋆	-	••∘∘
		UAV	Laser	••∘∘	Nadir	✓	×	×	LoD2	$	🎓	✓	📥	••∘∘
		UAV	Hybrid	••∘∘	Multi-view	✓	✓	×	LoD3	$	🎓	✓	📥	••∘∘
	close	UAV	Optical	••••	Multi-view	✓	✓	×	LoD3	-	👤	⋆⋆⋆	-	••••
		UAV	Laser	••••	Multi-view	✓	✓	×	LoD3	$	🎓	✓	📥	••••
Terrestrial		User	Optical	••••	Multi-view	×	✓ ⋆⋆	×	LoD3	-	👤	⋆⋆⋆	-	••••
		User	Laser	••••	Multi-view	×	✓ ⋆⋆	×	LoD3	$	🎓	✓	📥	••••
		User	Optical	••••	Multi-view	×	×	✓	LoD4	-	👤	⋆⋆⋆	-	••••
		User	Laser	••••	Multi-view	×	×	✓	LoD4	$	🎓	✓	📥	••••

Low, **Medium**, **High**, and **Very high**, respectively (• • • •);
Estimated cost ($) - 0 to **3**; **Specialist** (🎓); **Common user** (👤); **Embedded** (📥).

⋆ Parameter is a function of the sensor's Field of View (FOV). In cases of comprehensive FOVs, it is possible to observe not only the characteristics of roofs, but also of facades, allowing an acquisition of values above LoD2. In cases of FOVs with narrow angles, only the building footprint are observable, consequently, only LoD1 is reachable.
⋆⋆ The quality of facades images by terrestrial platforms, depends directly on the height of the building. The imaging of buildings with a height greater than 50 meters, for instance, is affected by the acquisition geometry in its upper part, as well as inner structures.
⋆ ⋆ ⋆ Control points are required.

state were also considered as non-users. The questions of the poll were:

1. Does the infrastructure/planning department have urban 3D city maps?
2. If so, how did this benefit the management and urban planning service?
3. If yes, what applications are used today?
4. If not, how is urban planning currently done?
5. How could 3D urban maps contribute?

Among the 26 Brazilian capitals and 1 federal district, only 8 of them answered the poll (Table 2). Columns with no information, were omitted by the interviewed.

Even though only a few secretaries have replied, the answers from the largest capitals have been obtained. Despite the different responses, it is believed that the vast majority still adopt the same technologies: long-range aerial surveys through LiDAR, to roughly observe the urban structuring. Still, the datasets acquired by some of these cities are out of operation for urban planning or strategic purposes.

2.3. The development trend

Much has been discussed about cloud technologies and sophisticated algorithms for detecting and classifying any kind of information. What would be the real gain of these technologies looking for smart cities? Is the integrity of this information compromised? To address these issues, three topics are raised: (i) the advantages of interpreting large amounts of information; (ii) the storage capacity; and (iii) the impacts of autonomous systems on human life.

As mentioned before, recognize tiny objects seems a difficult task for human eyes, but extremely easy to computers. The resurgence of the Convolutional Neural Network has changed the direction in the development of applications and services to execute them. For example, Deep-Learning (DL) methodologies [10], despite demanding high graphics processing capacity, have shown robustness levels never reached before, weather applied to forests, cities or indoor. Not only, the approach has also been a gold-standard on face recognition [11], speech recognition [12], natural language [13, 14], and now, constantly applying on cities for multiple purposes [15].

Since mid-2010, the amount of digital information has grown drastically due to popularization of mobile devices such as smartphones, tablets, watches, among others. These devices are easily accessed, but bring a very high individual volume every day, for example, a simple smartphone generates images, textual information on social networks, mobility information, or many kind of statistics. Access to these devices and their full integration with the Internet, has brought advantages in terms of volume of information, which leads to better estimatives.

Many of the objects that surround us will be on network. Sensor network technologies will rise to meet this new challenge, in which information and communication systems are invisibly embedded in the environment around us [16]. All kinds of things are being used in Internet of Things (IoT)[2] applications, including the most unusual products such as refrigerators, TVs or security cameras. In commercial devices, it is already reality, e.g. the use of sensors on traffic signals [17]. Concluding, any device able to be plugged on could be

[2]IoT aim to connect physical world to the Internet, thus, data from any eletronic device can be use to increase productivity and efficiency over a specific environment.

Table 2: Poll answers from each Brazilian Capital regarding the use of 3D maps on urban planning.

City	Question 1.	Released or Planned?	Product	Year	Sensor	Platform	Scale	Coverage
Fortaleza, CE	×	Planned	–	–	Laser	Airplane	Large	–
Vitória, ES	×	Planned	–	–	Laser	Airplane	Large	–
Porto Alegre, RS	✓	Planned	DSM, DTM	–	Laser	Airplane	Large	Full
São Paulo, SP	✓	Released	DSM, DTM	2017	Laser	Airplane	Large	Full
Belo Horizonte, MG	✓	Released	–	–	Laser	Airplane	Large	Medium
Rio de Janeiro, RJ	✓	Released	–	–	Laser	Airplane	Large	Medium
Curitiba, PR	✓	Released	–	–	Laser	Airplane	Large	Medium
Recife, PE	✓	Released	–	–	Laser	Airplane	Large	Full

part of an IoT system.

But what all this amount of information or even its intelligent interpretation have with the 3D urban models? Remote sensing has always been the key tool for systematic, automatic and large-scale urban mapping. Although it could count on complementary resources like crownsourcing, the use of smartphones [18], traffic data from geolocation reports, among others, this is only reality now. The connection between different sources potentiates and gives life to the quality map, establishing a new barrier to urban planning.

3. CONCLUSIONS

This article aimed to present the conditions of the 3D mapping of cities, with a brief discussion about Brazil. The technologies to observe cities, such as sophisticated sensors, reconstruction and classification techniques, evolve as the numerous architectural styles change according to local culture and way of life. Moreover, it is essential to think that the multiplicity of architectural styles is not the only problem. Studies, such as those carried out at Eidgenössische Technische Hochschule (ETH) Zürich [19], show that materials used in construction might become dynamic and therefore do not present a single static structure of a building.

In Brazil, there are several issues that makes 3D mapping even more challenging. Most common in metropolis, for instance, São Paulo (SP) and Rio de Janeiro (RJ), some characteristics such as the complexity of shapes, narrow paths, the uncountable building layouts: industrial, businesses, religious, commercial, among the particular residential areas: commonly known as *favela*, makes this a tough task, not only for applications, but also for data acquisition.

Alternatives on image analysis have already changed the way that people interact, consequently, the way the cities are built can equally be transformed to adapt to this changing. Many of questions are still not answer, anyway. How these amount of data can be use to solve critical problems, for instance, the construction of new highways by analyzing the annual traffic over that region, minimize the waste of water, food and others simply through the behavioral study of a particular industry or community, control epidemies, etc.

ACKNOWLEDGEMENT

This work was supported by CAPES (Coordination for the Improvement of Higher Education Personnel) and by the grant PDSE, Process No. 88881.132115/2016-01 (author 1), and CNPq (National Council for Scientific and Technological Development), Process No. 305054/2016-3 (author 2), which we kindly acknowledge.

4. REFERENCES

[1] SALEHI, A.; MOHAMMADZADEH, A. Building roof reconstruction based on residue anomaly analysis and shape descriptors from lidar and optical data. *Photogrammetric Engineering & Remote Sensing*, v. 83, n. 4, p. 281–291, 2017.

[2] VOSSELMAN, G.; DIJKMAN, S. et al. 3d building model reconstruction from point clouds and ground plans. *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, v. 34, n. 3/W4, p. 37–44, 2001.

[3] SNAVELY, N.; SEITZ, S.; SZELISKI, R. Photo tourism: exploring image collections in 3d. *ACM Transactions on Graphics*, 2006.

[4] KRIZHEVSKY, A.; SUTSKEVER, I.; HINTON, G. E. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*. [S.l.: s.n.], 2012. p. 1097–1105.

[5] LECUN, Y. et al. Backpropagation applied to handwritten zip code recognition. *Neural computation*, v. 1, n. 4, p. 541–551, 1989.

[6] VOSSELMAN, G. Fusion of laser scanning data, maps, and aerial photographs for building reconstruction. In: IEEE. *Geoscience and Remote Sensing Symposium, 2002. IGARSS'02. 2002 IEEE International*. [S.l.], 2002. v. 1, p. 85–88.

[7] FERRETTI, A.; PRATI, C.; ROCCA, F. Permanent scatterers in SAR interferometry. *IEEE Transactions on Geoscience and Remote Sensing*, v. 39, n. 1, p. 8 –20, jan 2001. ISSN 0196-2892.

[8] KOLBE, T. H.; GRÖGER, G.; PLÜMER, L. CityGML: interoperable access to 3D city models. In: *Geo-information for disaster management*. [S.l.]: Springer, 2005. p. 883–899.

[9] LOTTE, R. G. et al. 3D façade labeling over complex scenarios: a case study using Convolutional Neural Network and Structure-From-Motion. *Remote Sensing*, v. 10, n. 9, 2018. ISSN 2072-4292. Disponível em: <http://www.mdpi.com/2072-4292/10/9/1435>.

[10] LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. *Nature*, v. 521, n. 7553, p. 436–444, 2015.

[11] LAWRENCE, S. et al. Face recognition: A convolutional neural-network approach. *IEEE Transactions on Neural Networks*, v. 8, n. 1, p. 98–113, 1997.

[12] ABDEL-HAMID, O.; DENG, L.; YU, D. Exploring convolutional neural network structures and optimization techniques for speech recognition. In: *Interspeech*. [S.l.: s.n.], 2013. p. 3366–3370.

[13] KALCHBRENNER, N.; GREFENSTETTE, E.; BLUNSOM, P. A convolutional neural network for modelling sentences. *arXiv preprint arXiv:1404.2188*, 2014.

[14] CIRESAN, D. C. et al. Convolutional neural network committees for handwritten character classification. In: IEEE. *Document Analysis and Recognition (ICDAR), 2011 International Conference on*. [S.l.], 2011. p. 1135–1139.

[15] BILJECKI, F. et al. Applications of 3D city models: State of the art review. *ISPRS International Journal of Geo-Information*, v. 4, n. 4, p. 2842–2889, 2015.

[16] GUBBI, J. et al. Internet of things (iot): A vision, architectural elements, and future directions. *Future generation computer systems*, Elsevier, v. 29, n. 7, p. 1645–1660, 2013.

[17] MESE, J. C. et al. *Smart traffic signal system*. [S.l.]: Google Patents, 2006. US Patent 6,989,766.

[18] OVEREEM, A. et al. Crowdsourcing urban air temperatures from smartphone battery temperatures. *Geophysical Research Letters*, Wiley Online Library, v. 40, n. 15, p. 4081–4085, 2013.

[19] ADRIAENSSENS, S. et al. *Advances in Architectural Geometry 2016*. [S.l.]: vdf Hochschulverlag AG, 2016.